I Z A

IZA DP No. 4202

Wage Dispersion in a Partially Unionized Labor Force

John T. Addison
Ralph W. Bailey
W. Stanley Siebert

June 2009

Wage Dispersion in a Partially Unionized Labor Force

John T. Addison
University of South Carolina,
Queen's University Belfast and IZA

Ralph W. Bailey
University of Birmingham

W. Stanley Siebert
University of Birmingham and IZA

ABSTRACT

Wage Dispersion in a Partially Unionized Labor Force

Taking as our point of departure a model proposed by David Card (2001), we suggest new methods for analyzing wage dispersion in a partially unionized labor market. Card's method disaggregates the labor population into skill categories, which procedure entails some loss of information. Accordingly, we develop a model in which each worker individually is assigned a union-membership probability and predicted union and nonunion wages. The model yields a natural three-way decomposition of variance. The decomposition permits counterfactual analysis, using concepts and techniques from the theory of factorial experimental design. We examine causes of the increase in U.K. wage dispersion between 1983 and 1995. Of the factors initially considered, the most influential was a change in the structure of remuneration inside both the union and nonunion sectors. Next in importance was the decrease in union membership. Finally, exogenous changes in labor force characteristics had, for most groups considered, only a small negative effect. We supplement this preliminary three-factorial analysis with a five-factorial analysis that allows us to examine effects from the wage-equation parameters in greater detail.

Corresponding author:

John T. Addison
Department of Economics
Moore School of Business
University of South Carolina
1705 College Street
Columbia, SC 29208
USA
E-mail: ecceaddi@moore.sc.edu

I. Introduction

In his influential analysis of a partially unionized labor force, David Card (2001) pays close attention to the relationship between workers' pay, their levels of skill, and the likelihood of their being in a union.[1] Now pay and skill are unlikely to be independent of union membership. And if they are dependent, it is important to understand the nature of this dependence in investigating such issues as the extent to which deunionization contributes to elevated wage inequality, or the notion that unions are relatively more effective in raising wages at lower reaches of the wage distribution (see respectively Card, 2001; Lewis, 1986).

Card (2001, p. 297) seeks to accommodate two specific features of partially unionized labor markets: first, unionization rates that vary across the wage distribution, with disproportionate falls among lower-wage workers in Britain and the United States in the 1980s; and, second, differences in the relative wage effect of unions for different skill groups. He is concerned to point out how each tendency may upset the conclusions of analyses that neglect the interrelationship between unionization and skill.

Card's own approach targets the variance of the log-wage, which is an attractive way to measure dispersion because differences in the log-wage are unit-free and are therefore (like pay ratios) of direct economic meaning. Furthermore, the variance lends itself to additive decompositions via the analysis of variance, an obvious advantage it holds over other dispersion measures such as the standard deviation, Gini coefficient, and inter-quartile range. In Card's treatment, workers are split into discrete 'skill' categories on the basis of the predicted nonunion wage. The law of total variance is then used to decompose the total variance of the log-wage into the contributions from each category. Within skill categories, a further decomposition is possible; specifically, into the contributions from union and nonunion membership. In investigating a phenomenon such as deunionization, therefore, one can use Card's decompositions to accommodate the effect of deunionization rates that differ by skill level.

The motive for the present treatment is that we believe that there is scope for enhancement of Card's illuminating formulation of variance decomposition, based on an alternative handling of his important 'skill' variable.

---

[1] See also Card, Lemieux and Riddell (2004); Canal Dominguez and Rodriguez Gutierrez (2004); DiPrete (2005); Freeman (2005); Checchi, Visser, and Van de Werfhorst (2007); Addison, Bailey, and Siebert (2007).

Specifically, we shall argue that (i) Card's variance decompositions based on groups of workers in skill categories would be easier to interpret if conducted at the level of individual workers; (ii) retaining individual-level information about characteristics and 'skill' would also facilitate a more accurate analysis of wage variance, particularly if the characteristics vector (rather than the scalar 'skill') were used to explain union-membership behavior; and, (iii) rather than making 'skill' depend only on the predicted *non*union wage, it would be more natural to use a definition of skill that is even-handed between the union and nonunion sectors.

The plan of the paper is as follows. Section II outlines the salient features of Card's analysis. Section III amplifies the proposals summarized above, and explains why we see them as desirable. Section IV develops our own model of wage dispersion in partially unionized labor forces, leading to a natural three-way variance decomposition. Section V then deals with estimation matters. Section VI analyzes the components of U.K. wage variance in 1983 and 1995. Section VII shows how the three-way decomposition, allied with factor-response techniques from the theory of experimental design, can be used to investigate the causes of the variance increase identified earlier. Section VIII concludes.

II. Card's Model of a Partially Unionized Labor Force

The use of the analysis of variance for partially unionized labor forces was pioneered by Freeman (1980, p. 19, equation [6]), using the so-called *law of total variance*. The law states that if data fall into groups, then overall variance can be decomposed into the variance of the group means, plus the mean of the group variances. A similar law applies to random variables. Suppose that we are modeling the data by use of a random variable W representing the log-wage; that there are two groups, unionized and nonunionized workers, occupying proportions u and $1 - u$ respectively of the labor force; and that the first two moments of the two groups' log-wages are given by (w^u, v^u) and (w^n, v^n), respectively. Further, let $v = Var\left(W\right)$. Then the law tells us that

$$v = uv^u + (1 - u) v^n + u (1 - u) \Delta_w^2, \tag{1}$$

where $\Delta_w \equiv w^u - w^n$. The final, variance-of-group-means term is the variance of a dichotomous random variable that equals w^u with probability u, and w^n

with probability $1 - u$. One can of course alternatively obtain (1) directly from the definition of v.

Card's model of wage dispersion can be seen as a far-reaching generalization of the Freeman decomposition. But it does not confine itself to observed quantities. As noted by Card, Lemieux, and Riddell (2004, p. 521), it can also be seen as an example of potential outcomes analysis (Angrist and Krueger, 1999) in which one attempts to assign to each worker both a union wage and a nonunion wage. Only one of these is observed, while the other is estimated and used in counterfactual analysis – as, for instance, when we ask what the wage structure would have been had some nonunion workers instead belonged to the union sector.

The model presented by Card (2001, pp. 297-298) allocates workers to categories $c = 1, \ldots, K$ according to a variable called *skill*. An exogenously-determined proportion $u(c)$ of workers in skill category c are union members. The log-wages of union workers within a given category are assumed uncorrelated, with identical expectations and variances. In an analysis of variance we are interested only in the first two moments; for union worker i in category c the log-wage has moments $w_i^u(c) \sim (w^u(c), v^u(c))$, where $w^u(c)$ and $v^u(c)$ are the category mean and variance. We assume that the data are numerous enough for the two moments to be calculated with precision. Similar assumptions are made about nonunion workers, with the difference that $w_i^n(c) \sim (w^n(c), v^n(c))$.

At this stage, *skill* seems exogenous. However, Card (2001, pp. 304-305) recognizes that skill is partially determined by observed characteristics: "A necessary first step ... is to define skill groups ... I divided workers into observable groups based on their predicted wages in the nonunion sector ..." Consequently, there are two differing expected log-wages, at any rate for nonunion workers in a given skill category; namely, the category mean $w^n(c)$, and the prediction just mentioned. We might reconcile these by supposing that the second wage prediction is accessible to researchers but not to employers, who observe only the worker's skill category.

Partitioning the data into skill categories allows Card to investigate how the expectation and variance of the log-wage differ across the categories. He can attribute the variance of the log-wage to components from all of the skill categories and both sectors (union and nonunion). In Card's (2001, p. 298) analysis, the law of total variance yields the first line of his equation (3), namely

$$v = Var\left[w\left(c\right)\right] + E\left[v\left(c\right)\right], \tag{2}$$

3

where $w(c)$ and $v(c)$ are the category means and variances. Noting that each skill category contains both union and nonunion members, he obtains further decompositions. We can write $w(c)$ as the weighted mean, as follows

$$
\begin{aligned}
w(c) &= u(c)\,w^u(c) + (1 - u(c))\,w^n(c) \qquad\qquad (3) \\
&= w^n(c) + u(c)\,\Delta_w(c),
\end{aligned}
$$

where $u(c)$ is union coverage in category c, and $\Delta_w(c) \equiv w^u(c) - w^n(c)$. By a second application of the law of total variance we can decompose the within-category variance $v(c)$ into the Freeman-like form

$$
v(c) = u(c)\,v^u(c) + (1 - u(c))\,v^n(c) + u(c)\,(1 - u(c))\,\Delta_w(c)^2. \qquad (4)
$$

By combining the three expressions above, we arrive at Card's culminating variance decomposition

$$
\begin{aligned}
v &= v^n + Var\left[u(c)\,\Delta_w(c)\right] + 2Cov\left[w^n(c), u(c)\,\Delta_w(c)\right] \qquad (5) \\
&\quad + E\left[u(c)\,\Delta_v(c)\right] + E\left[u(c)\,(1 - u(c))\,\Delta_w(c)^2\right],
\end{aligned}
$$

where

$$
v^n = Var\left[w^n(c)\right] + E\left[v^n(c)\right]. \qquad (6)
$$

The decomposition (5) shows the dependence of variance on the structure of union density $u(c)$ across the skill levels. This is particularly important when we ask how v might change if the level of unionization changed. Suppose (to take an example similar to those discussed in section VII below) we ask what v would have been in a given year had overall rates of unionization been similar to those obtaining in some other year. The distinctive feature of Card's approach is that the use of skill categories in (5) allows him to produce estimates of this counterfactual v that pay attention to the differing rates of deunionization across the different skill categories[2].

---

[2]Equations (2) to (5) are the fundamental elements of Card's analysis. However, he also discusses a particular departure from them in referring to "unobserved heterogeneity," which was a key concern in his earlier investigation (Card, 1996) of union impact on the structure of wages. As Card (2001, p. 299) describes the problem: "The preceding formulas have to be modified slightly if the union and nonunion workers in a given skill category have different productivity levels and would earn different wages even in the absence of unions. Such a phenomenon will arise if workers have productivity characteristics that are known to employers but not fully captured in the observed skill categories, and if the mean level of these unobserved skills is different between union and nonunion workers in a given skill group."

4

III. Some Developments of Card's Model

We now elaborate the points listed in section I, suggesting that the model is capable of further refinement. Specifically, we shall claim the following.

Card's variance decompositions based on groups of workers in skill categories would be easier to interpret if conducted at the level of individual workers. Card's important decomposition (2), based on the law of total variance, is a striking and fundamental feature of his model, but its meaning for the dispersion of wages is hard to interpret. The problem is that the decomposition varies with respect to the arbitrary number of categories K (which Card takes equal to 10). As K increases, the categories narrow and the within-category variances decrease correspondingly. Consequently, variance is transferred in an uninformative way from the second term to the first on the right-hand side of (2). One can contrast this situation with a more economically meaningful decomposition such as that of Freeman (1), whose partitioning is informative because the union and nonunion sectors, unlike skill quantiles, are economically significant and autonomous entities.

To transmute (2) into a more informative equation, we suggest replacing skill groups with individual workers. Thus K becomes equal to the sample size N. This restores significance to (2), which now allows us to analyze total variance into economically meaningful components: within-worker variance and across-worker variance. Of course, skill 'groups' containing one worker each raise numerous problems. These are dealt with in the next section.

Retaining individual-level information about characteristics and 'skill' would also facilitate a more accurate analysis of wage variance, particularly if the characteristics vector (rather than the scalar 'skill') were used to explain union-membership behaviour. Predicted log-wages provide information at the interval-scale level of information.[3] Using categories as an alternative is cruder, although one might adopt them in a preliminary study or for reasons of convenience or necessity. The problem we face is that of assigning a nonunion wage to union workers (or vice versa), so as to be able to perform a counterfactual analysis. With the help of skill categories, we could measure the effect of a shift in union coverage by assuming that if a worker leaves a union, the new wage earned will be distributed like that of nonunion workers in the same skill category (subject to modifications in the case of unobserved

---

[3]See Stevens (1946) for a discussion of levels of measurement and, for criticisms of his typology, Velleman and Wilkinson (1993).

heterogeneity). However, replacing skill by two expected wages (union and nonunion) provides a more sensitive means of dealing with deunionization. The counterfactual wage is then predicted using all available information about the worker, not just his or her skill category and union status.

As Freeman's equation (1) shows, an important aspect of wage variance is union density. How this is determined is an important aspect of any model of wage variance in a partially-unionized labor force. Card, Lemieux, and Riddell (2004, p. 522) state: "One starting point is the assumption that union status 'is as good as randomly assigned,' conditional on observed skill characteristics." And this starting point is in fact the one adopted in Card (2001). However, it seems more probable that within each skill category, some union and nonunion members will be likelier than others to leave or join a union when conditions change. We should take such factors into account when predicting the resulting counterfactual distribution of log-wages. Accordingly, in investigating the effect of deunionization on wage dispersion, one should be as careful in modeling the causes of union membership as in deriving its effects.

In our own model, following Dinardo, Fortin, and Lemieux (1996), we explain union membership by means of a probit model, allowing us to assign a union-membership probability to each individual worker. Then, for example, in addressing the deunionization that took place in Britain between 1983 and 1995, we can investigate how much the wage variance would have increased in the absence of deunionization, by applying the 1983 probit coefficients to 1995 workers and estimating the resulting variance. In fact, we shall consider a number of counterfactual scenarios, in a systematic way, using methods from the theory of experimental design.

The point about using all available information is relevant to the problem of unobserved heterogeneity (Card, 2001, pp. 299-300). Discussion of unobserved heterogeneity is difficult in the context of Card's model precisely because of the deletion of information that takes place when worker i is assigned to a skill category, and his or her characteristics z_i are discarded. It could be that relatively productive nonunion workers tend to be bunched towards the top of a given skill category, while relatively unproductive union members are bunched towards the bottom. This bunching could happen simply because of the workers' spread of characteristics z_i, and hence be fully explicable within the model if only we made use of the available information[4].

---

[4]This is not to say that the problem of unobserved heterogeneity is easily overcome.

Rather than making 'skill' depend only on the predicted nonunion wage, it would be more natural to use a definition of skill that is even-handed between the union and nonunion sectors. It seems desirable to treat union and nonunion workers symmetrically, rather than defining the important *skill* variable as the predicted *non*union wage. We propose an even-handed approach: to define skill as a worker's expected wage, unconditional on his or her union status. To compute this expected wage we need to know his or her union-membership probability, which can be obtained as fitted values from the previously mentioned probit analysis. We also need to know the worker's expected union and nonunion wages. These can in turn be obtained as predicted values from separate regression analyses of union and nonunion wages.

IV. A Two-Wage Model

Proceeding in the manner suggested, we now present a model in which the number of skill categories K in effect equals the number of workers N. Thus, the part of $w(c)$ is played by w_i, the unconditional expectation of W_i, the log-wage of worker i. In what follows we shall abbreviate 'log-wage' as 'wage.' The part of $v(c)$ is played by $v_i \equiv E\left[(W_i - w_i)^2\right]$. Accordingly, (2) is replaced by the more transparent

$$v = Var(w_i) + E(v_i),\tag{7}$$

in which the variance and expectation are now taken over worker-subscripts i rather than over category indices c. We call w_i 'skill', which is thus synonymous with 'expected wage'. Thus we may read (7) as

{Total variance} = {Variance of across-worker skill}
+ {Average within-worker variance}

As Card notes (2001, p. 300), longitudinal data on union status changers could be used to evaluate the wage gains of union joiners and the wage losses of union leavers (see, for example, Lemieux, 1992; Card, 1996). Longitudinal studies observing the same workers in both union and nonunion contexts provide evidence that the log-wages of union workers that transfer to the nonunion sector do indeed differ systematically from predictions of their wages formed on the basis of observed characteristics alone. But our point is that the effect, although real, will be over-stated if we fail to exploit the explanatory power residing in the workers' characteristics z_i, which are necessary for a satisfactory counterfactual analysis.

Since our skill categories contain exactly one worker each, we cannot follow Card (2001) in estimating w_i and v_i from quantile data. Instead, we must analyze the dependence of the wage on union membership and skill in a more detailed way. Our model for the way in which W_i is generated is

$$
\begin{aligned}
W_i &= U_i W_i^u + (1 - U_i) W_i^n \\
&= U_i (w_i^u + \varepsilon_i^u) + (1 - U_i) (w_i^n + \varepsilon_i^n),
\end{aligned}
$$

where U_i is a dichotomous random variable indicating union membership or otherwise, and W_i^u and W_i^n are the worker's union and nonunion wages, only one of which is observed. Thus, U_i equals 1 with probability u_i and U_i equals 0 with probability $1 - u_i$. The union membership probability u_i depends on i's characteristics vector z_i. Hence, $u_i = E(U_i|z_i)$. The value w_i^u is the expected union wage for this worker, $w_i^u = E(W_i|U_i = 1, z_i)$. The deviation ε_i^u from this value reflects unobserved factors and random variation. We assume that $\varepsilon_i^u \sim (0, v_i^u)$. Similarly, $w_i^n = E(W_i|U_i = 0, z_i)$, $\varepsilon_i^n \sim (0, v_i^n)$. An effect of our observing only one wage is that the degree of association between ε_i^u and ε_i^n plays no part in the variance analysis that follows; we can and do leave it unspecified. We assume that U_i is uncorrelated with both ε_i^u and ε_i^n. Our version of (3) is therefore

$$
w_i = E(W_i|z_i) = u_i w_i^u + (1 - u_i) w_i^n = w_i^n + u_i \Delta_{w,i}, \tag{8}
$$

where $\Delta_{w,i} = w_i^u - w_i^n$ is the 'wage gap' (the difference between the individual's expected union and nonunion wages), and w_i^u and w_i^n are functions of z_i. We can now see the difference between Card's *skill*, which corresponds to our w_i^n, and our own measure w_i.

As regards the worker's wage variance conditional on z_i, we obtain the Freeman-like relation

$$
v_i = Var(W_i|z_i) = u_i v_i^u + (1 - u_i) v_i^n + u_i (1 - u_i) \Delta_{w,i}^2, \tag{9}
$$

where the conditional variances v_i^u and v_i^n are again functions of z_i. The first two terms together equal expected residual variance, while the final term is the variance deriving from uncertainty about the worker's union status. Accordingly, the equation can be read {Variance of i's wage} = {expected residual variance for i} + {wage gap variance for i}. Combining (7) and (9), and re-ordering the terms, we obtain a fundamental three-way decomposition of the variance of the wage, namely

$$
v = Var(w_i) + E\left[u_i(1 - u_i)\Delta_{w,i}^2\right] + E\left[u_i v_i^u + (1 - u_i) v_i^n\right], \tag{10}
$$

in which the first term relates to across-worker skill differences, the second to the wage gap, and the third to the residual error terms ε_i^u and ε_i^n. Thus we may think of (10) as decomposing total variance v in an economically informative way as

$$
\begin{aligned}
v \quad = \quad & \{\text{across-worker skill variance}\} \\
& + \{\text{within-worker wage-gap variance}\} \\
& + \{\text{within-worker expected residual variance}\}. \qquad (11)
\end{aligned}
$$

Card's mean and variance decompositions remain true in our model, except that his skill categories c are replaced by worker indices i. For example, Card's (2001) major formula, (5) above, is easy to modify in this way. In doing so, let us introduce notation $\pi_i \equiv u_i \Delta_{w,i} = w_i - w_i^n$ for the quantity conventionally described as *union power*, namely the extent to which the worker's expected wage (his or her skill, in our terms) exceeds his or her expected nonunion wage, both wages being conditional on characteristics z_i. We then obtain

$$
\begin{aligned}
v \quad = \quad & Var\left[w_i\right] + E\left[v_i\right] \\
= \quad & Var\left[w_i^n + \pi_i\right] + E\left[v_i^n + u_i \Delta_{v,i} + u_i \left(1 - u_i\right) \Delta_{w,i}^2\right] \\
= \quad & Var\left[w_i^n\right] + Var\left[\pi_i\right] + 2Cov\left[w_i^n, \pi_i\right] \\
& + E\left[v_i^n\right] + E\left[u_i \Delta_{v,i}\right] + E\left[u_i \left(1 - u_i\right) \Delta_{w,i}^2\right],
\end{aligned}
$$

where moments are now computed across workers i.

Card (2001, p. 305) uses his decompositions to illustrate the differing levels and effects of unionization across the skill levels. Thus, for example, his Table 4 shows how union membership $u\left(c\right)$, mean wage gap $w^u\left(c\right) - w^n\left(c\right)$, and mean variance gap $v^n\left(c\right) - v^u\left(c\right)$ are related to skill category c.[5] In general, however, it is more illuminating to plot such graphs using on the horizontal axis the interval-scale information present in our 'skill' w_i rather than the merely ordinal information present in the category number c. Doing so requires kernel-density techniques that we develop in the Appendix.

_____

[5] Similar information is presented in Addison, Bailey, and Siebert (2007), Figures 1 and 3, pp. 344-345.

V. Estimation of the Model

Given the information on U_i, W_i, z_i in the data set, our model requires us to estimate the quantities u_i, w_i^u, w_i^n, v_i^u, and v_i^n for each worker. In doing so, we make two convenient although inessential assumptions. The first of these is that the regressors in respect of union membership U_i are the same as the regressors for the log-wage W_i, namely the z_i in both cases. One might seek separate and more parsimonious specifications for each regression; but with labor force data there are typically enough observations to make this effort unnecessary.

Our second assumption relates to the variances v_i^u and v_i^n. In describing his regression procedures allowing for differences across skill groups, Card (2001, pp. 304-305) assumes homoskedasticity of individual workers, and we do likewise. Thus, our second assumption is that all v_i^u equal the same value σ_u^2, and all v_i^n equal the same value σ_n^2. (We note parenthetically that Card's $v^u(c)$ and $v^n(c)$ do vary across skill categories. However, these quantities cannot be interpreted as wage variances conditional on a given level of skill since, as already noted, when we increase the number of categories K a worker's particular $v^u(c)$ and $v^n(c)$ decrease, as the corresponding category ranges shrink. Consequently, the assumption of log-wage homoskedasticity within the union sector, conditional on characteristics z_i, does not contradict heterogeneity of the $v^u(c)$. The same applies with respect to nonunion workers and the $v^n(c)$.)

Based on these two modeling decisions, our procedure is as follows. First, for all workers in the sample run a probit regression of U_i on z_i, yielding fitted probabilities $\hat{u}_i$. Second, for union workers only ($U_i = 1$), regress W_i on z_i, yielding estimated union wage coefficients $\hat{\beta}^u$, and also an estimated union variance $\hat{\sigma}_u^2$. Using this regression equation we can predict a union wage $\hat{w}_i^u = z_i' \hat{\beta}^u$ for all workers, union and nonunion. Third, repeat the operation for nonunion workers only, allowing us to predict a nonunion wage $\hat{w}_i^n$ for all workers, with an associated variance $\hat{\sigma}_n^2$. The estimated quantities $\hat{u}_i, \hat{w}_i^u, \hat{w}_i^n, \hat{v}_i^u = \hat{\sigma}_u^2, \hat{v}_i^n = \hat{\sigma}_n^2$ can now be used to estimate the decomposition (10).

In subsequent sections we will apply the above techniques to data for the U.K. labor force. The information on earnings, unionization, and individual characteristics, and our rationale for choosing them, are described in detail in Addison, Bailey, and Siebert (2007, pp. 339-341). The 1983 data come from the General Household Survey (GHS) of that year, the only year in which

the GHS included a union membership question. In selecting the 1983 GHS we make the same choice as have other scholars, including Machin (1997), Bell and Pitt (1998), and Gosling and Lemieux (2001). For 1995 we choose the Labour Force Survey of that year. By 1995 the U.K. union movement's fortunes had declined sharply from a 'golden age' of unionism to which 1983 belonged. The contrast between the two years thus provides an intriguing natural experiment for observers concerned with the factors, including union density, that affect wage dispersion.

(Figures 1 and 2 near here)

Figure 1 ((a) through (e)) provides a smoothed version of the dichotomous union-membership indicator variable U_i, plotted against centered skill $w_i - \bar{w}$ in our two comparator years. The point of the centering is to remove the effect of changes in the general wage level across the two years. (For Card, this removal is effected by his use of categories, which are perforce unaffected by the general wage level.) In similar fashion Figure 2 plots the union-nonunion wage gap $\Delta_{w,i}$ against centered skill.

Figure 1 shows that, for most groups of full-time workers, unionization is highest at intermediate levels of skill, tending both to decrease in slope and flatten out in the transition from 1983 to 1995. For its part, Figure 2 shows, again for most groups, a strong negative correlation between skill and the wage gap, pointing to a seemingly strong incentive to belong to a union at low skill levels. Similar qualitative features are found both in Card's wage-gap graphs (Card, 2001, p. 312, Figures 3, 4) and in Addison, Bailey, and Siebert (2007, p. 344). For private-sector workers the wage-gap curve tends to shift downward in the later year.

The exception to these generalizations is provided by public-sector females, whose union density increases with skill, a relationship that strengthens across the two years. Striking also is the change in this group's wage-gap curve between 1983 and 1995, which uniquely among the wage-gap curves even shows a strong positive relationship to skill in 1995 among skilled workers. One might hypothesize from Figures 1(e) and 2(e) that some change between 1983 and 1995 increased the union premium sharply for high-skilled public-sector females, and that this had a causal effect on the willingness of these workers in to join a union. In section VII we will see that the effect of deunionization is quite different for public-sector females than for other groups in the sample.

11

VI. Decomposition of U.K. Log-Wage Variance in 1983 and 1995

In section VII we will discuss the contribution of deunionization and other factors to the increase in U. K. wage dispersion between 1983 and 1995, for various sub-groups. The counterfactual methodology adopted will depend heavily on the three-way variance decomposition (10), and estimation of the associated parameters. In the present section we illustrate the necessary techniques, using as an example the log-wage of the private-sector male sub-group.

The first estimation task is to conduct a probit analysis of union membership for this group, so as to obtain a union-membership probability for each worker in 1985 and 1993. Changes in these equations and associated worker characteristics determine the union density contribution to changes in wage dispersion. As a measure of our level of success in explaining union membership we use Efron's (1978) pseudo-R^2, defined as $R_E^2 = 1 - \Sigma \left(U_i - \hat{u}_i \right)^2 / \Sigma \left(U_i - u \right)^2$, where u is the actual union density and $\hat{u}_i$ is the probit estimate of worker i's union membership probability. In other words R_E^2 indicates the explanatory power of the probit model, compared with the model that assigns the same probability $u = \bar{U}$ to all workers.

Following Addison, Bailey, and Siebert (2007, p. 341) (see also Card 2001, p. 305) the regressors used (i.e. the components of the characteristics vector z_i) in both the probit analysis and the regression analyses are: years of education and years of experience, and their second, third, and fourth powers; interaction between experience and education; interaction between experience and experience-squared; as well as dichotomous variables indicating whether or not workers are white, married, and employed by small firms, together with five regional, eight occupational, and nine industry dummies. Details of the probit analyses for the private-sector male group years appear in Table 1.

(Table 1 near here)

Table 1 shows for this group a conventional pattern of union membership determinants, with the main change over the two years confined to the occupation dummies. Thus, the coefficient estimates for labor market experience are well determined throughout in both years: in general, rising experience translates into a greater probability of union membership, as usual. Another

12

well established result is the relationship between firm size and unionization: in both years, workers in small firms are less likely to be unionized than their counterparts in larger enterprises. But neither race nor marital status appear to be influential. The occupational dummies show blue-collar occupations to be more unionized in both years, but generally the coefficients are smaller. Coefficients tend also to become smaller over time for the industry dummies. Below, we will find these changes to be variance-increasing. There has been a movement of private-sector males towards the less-unionized white-collar categories, and the blue-collar categories themselves show smaller coefficients in the union probit. However, as our calculations below will show, these effects are quite small for the private-sector male group, although larger for other sub-groups such as public-sector females.

The next step is to run a least-squares regression of the log-wage on worker characteristics for the union sector, and to use the estimated coefficients β^u to predict an expected union wage for every worker – union and nonunion. Then, we need to run a similar regression for the nonunion sector, and use the resulting coefficients β^n to predict a nonunion wage for every worker. (Results of these exercises for private-sector males, and indeed all nine worker groups, are available from the authors on request.) The estimated residual variances $\hat{\sigma}_u^2$ and $\hat{\sigma}_n^2$ from these two fitted regressions are interesting in that both increase across the two years, $\hat{\sigma}_u^2$ by 22 percent (from 0.099 to 0.121) and $\hat{\sigma}_n^2$ by 30 percent (from 0.125 to 0.162). This increase suggests already that deunionization may not be the sole cause of the increase in wage dispersion, since residual variance was increasing within each sector over the sample period.

(Table 2 near here)

Table 2 illustrates the estimation of the three-way decomposition for the entire labor force in each of our two chosen years. From the last line of the table, which gives the variance of the log-wage in 1983 and 1995, observe the 34 percent increase in this value (from 0.221 to 0.297) which provides the motivation for the analysis in the preceding rows. (The question of the extent to which this material increase in dispersion is connected with the 31 percent decrease in unionization observed over the same period is addressed in section VII.)

Using the probit and regression analyses, we can estimate the components of dispersion detailed in (10) and (11). As explained in section V, this exercise

involves using the results from the probit and regression analyses to assign a union-membership probability to each worker, and a union and nonunion log-wage to each worker, as well as a variance to each of these log-wages. One then treats the log-wages actually observed as a single draw from the probability model. An important check on the model's performance is to determine how close its predicted variance (a sum of three components) is to the actual variance. From the last two lines of the table it can be seen that, for the labor force as a whole, the actual and predicted variance are close to each other in each of the two years. Looking at the components themselves for 1983, we see that the wage-gap variance $u_i \left(1 - u_i\right) \Delta_{w,i}^2$ is small in comparison with the other two components, although as the various components of Figure 2 show, $\Delta_{w,i}$ itself can be sizable.

(Table 3 near here)

Since the contribution of the union-nonunion gap variance is so small, nearly all dispersion comes from the two other components, and in roughly equal amounts. Table 3 shows that similar percentages – approximately 50:0:50 – hold in subgroups of the labor force. Comparing the model predictions of variance $\hat{v}_{tot}$ and the actual variances v_{act}, shown in the last two columns of the table, note that the two magnitudes tend to be closer in 1983 than in 1995, and closer in the private sector than the public sector. The agreement between the two is generally good: of the eighteen groups considered in Table 3, $|\hat{v}_{tot} - v_{act}|$ exceeds 0.004 only for the three smallest: public-sector males in 1995, and public-sector females in both 1983 and 1995.

VII. An Analysis of the Increase in U.K. Wage Dispersion Between 1983 and 1995

Suppose that we are supplied with a set of union-membership probit coefficients γ, a sequence of wage-regression parameters $\psi \equiv (\sigma_u^2, \beta^u, , \sigma_n^2 \beta^n)$, and a data set of worker characteristics $Z = (z_1, \ldots, z_N)$. Then, the three-way decomposition (10) yields a value

$$v = v\left(\gamma, \psi, Z\right) \tag{12}$$

for the variance of the log-wage. Notice that v is population variance (a constant), not sample variance (a random variable). It can be thought of

14

as the average wage variance from an infinite sequence of replications of the model, if we were able to set up identical labor forces and labor markets in many different worlds. Using the language of the theory of experimental design, we can think of γ, ψ, and Z as 'factors' or 'treatments' influencing the 'response' variable $\hat{v}$. We now attempt to say how important each factor was in explaining the actual increase in wage variance $v_{1995} - v_{1983}$, taking into account the non-linearity of the relation (12).

Methodologies for doing so are suggested in the experimental design literature; see, for example, Cox and Reid (2000); Box, Hunter, and Hunter (2005); Morgan and Winship (2007). A useful brief survey of terminology and methods is provided by Li, Sudarsanam, and Frey (2006). The way these ideas are applied in the present section are as follows. The model (12) provides a deterministic relationship between the model inputs and the model's response variable, log-wage variance. The theory of experimental design is helpful (even though experimental error is absent from (12)) because it is particularly concerned with the application of several differing 'treatments', each at two levels conventionally called 'low' and 'high' (in our case, 'low' corresponds to 1983, 'high' to 1995; the 'treatments' themselves will be described shortly). Moreover, as we now show, experimental design's concept of the 'main effect' proves a valuable means of addressing the index number problem of deciding at which level (1983 or 1995) to hold the different explicative factors, when we undertake counterfactual analysis.

Let us denote the 1983 and 1995 settings of γ, ψ, and Z by the subscripts 0 and 1, respectively. Write $\hat{v}_{101}$, for instance, to mean $v\left(\hat{\gamma}_{1995}, \hat{\psi}_{1983}, Z_{1995}\right)$. Thus, $\hat{v}_{101}$ can be interpreted as the predicted counterfactual log-wage variance if probit coefficients and worker characteristics are held at their 1995 levels, but wage equation parameters assume their 1983 levels. We can calculate all eight predictions $\hat{v}_{000}, \hat{v}_{001}, \ldots, \hat{v}_{111}$ using (10). Thus the three 'treatments' can be applied at any combination of 'high' and 'low' settings, and the size of the resultant 'response' can be measured in each case.

The 'effect' of a factor is its impact (however measured) on the response variable. In particular, the 'main effect' of a factor is defined in the literature as the average value of the responses when the factor is at its high setting, minus the average value when it is at its low setting. For example, the main effect of the first factor, union probit coefficients γ, is

$$ME_{\gamma} = \frac{1}{4}\left(\hat{v}_{100} + \hat{v}_{101} + \hat{v}_{110} + \hat{v}_{111}\right) - \frac{1}{4}\left(\hat{v}_{000} + \hat{v}_{001} + \hat{v}_{010} + \hat{v}_{011}\right)$$

15

with corresponding expressions for the main effect from the wage equations, ME_ψ, and the main effect from characteristics, ME_Z. The sum of the three main effects in general differs from both the actual variance increase $v_{1995} - v_{1983}$ and the fitted increase $\hat{v}_{1995} - \hat{v}_{1983}$. However, if the model is a good one and if (a separate point) the response is a roughly linear function of the three factors, we obtain

$$v_{1995} - v_{1983} \simeq ME_\gamma + ME_\psi + ME_Z \; (\simeq \hat{v}_{1995} - \hat{v}_{1983}). \qquad (13)$$

As Table 4 shows, the approximation (13) does in fact hold up well in the wage-variance context (for instance, for private-sector males the sum of the main effects is 0.077, while the actual increase in the variance of the log wage is 0.076) allowing us to use (13) to estimate the comparative effects of changes in the three factors, on wage dispersion.

(Table 4 near here)

The labor force as a whole and nearly all subgroups display the same pattern. Most of the variance increase is attributable to ψ, that is, to changes in the coefficients and residual variances of the two wage equations. Less important are changes in the union probit coefficients γ. Least influential of all are changes in exogenous worker characteristics factor Z, whose impact is both small and negative for the majority of subgroups. This last result means that if the union probit coefficients and wage parameters were held constant, changes in worker characteristics would have caused a small *decrease* in dispersion. As might have been expected from the graphs in section VI, public-sector females prove an anomalous group. For these workers, the main effects from changes in the probit and wage parameters are particularly large, relatively speaking, while the main effect from changes in characteristics is large and negative. These aspects of public-sector females have a clear influence on the public sector as a whole.

Public-sector females became more homogeneous over the period, which explains the negative effect of changes in characteristics for their wage dispersion. To understand in greater detail why this is so, we examined the wage regressions for this groups of workers. The regressions showed that a key variable affecting wage was years of education, whose variance for this group fell from 5.24 to 3.91. In contrast, the variance of years of education for private-sector females rose from 1.94 to 3.23, and for this group the main effect from changes in characteristics was positive.

16

The preliminary factorial analysis undertaken in Table 4 suggests directing closer attention to the wage-equation parameters ψ, the most active of the three factors considered. Accordingly, we now treat σ_u^2, β^u, σ_n^2, and β^n as four separate factors, and (following common practice in factorial experiments) combine the relatively unimportant factors γ and Z into a fifth factor, 'other', thus obtaining a 2^{6-1} design based on the thirty-two predicted responses $\hat{v}_{00000},\ldots,\hat{v}_{11111}$. The results from the new counterfactual experiment are shown in Table 5.

(Table 5 near here)

For public-sector workers one would expect the combined effect of changes in the union parameters σ_u^2 and β^u to outweigh the combined effect of changes in the nonunion parameters σ_n^2 and β^n, and this is borne out in Table 5, with as expected the reverse result for private-sector workers. Nonetheless, changes in the non-union residual wage variance factor σ_n^2 accounts for a sizeable proportion of changes in the wage variance ($v_{1995} - v_{1983}$) for all groups, public and private, except public-sector males. In six of the nine cases, it turns out to be the most important factor in the variance increase[6]. We also see that changes in the coefficients β^u and β^n from the wage equations always contribute positively to $v_{1995} - v_{1983}$. This effect means that the wage coefficients have increased in magnitude over time, pushing dispersion upwards. The coefficients γ from the union-membership probit equations also have a positive contribution. This effect, by contrast, arises from the union coefficients becoming smaller in magnitude, which in turn prompts a fall in union density and a migration of union workers to the more dispersed nounion sector.

---

[6]The table provides checks of the decisions made above about our choice of factors. As in Table 4, the first two columns show that the sum of the main effects for each group is close to the actual $v_{1995} - v_{1983}$ for the group. Comparing $ME_U + ME_Z$ (from Table 4) with the main effect of 'other' in Table 5, we find that the absolute difference between the two exceeds one percent only in the case of private-sector males, for whom it is two percent (10.6 percent minus 1.8 percent, versus 6.8 percent). It appears that there is little interaction between γ and Z. In other words, combining them in order to reduce the number of factors has done little harm.

VIII. Conclusion

We have developed a two-wage model of wage dispersion in a partially-unionized labor force, in which dispersion is decomposed into explicable factors and residual variation. The residual component is minimized by treating union membership as endogenous and therefore partially explained by worker characteristics. Each worker has a union membership probability, a predicted union wage and residual variance, and a predicted nonunion wage and residual variance. We used the law of total variance to obtain a natural three-way decomposition of log-wage variance based on these quantities.

We then used the three-way decomposition in a counterfactually-based analysis of the increase in U. K. wage dispersion that occurred between 1983 and 1995. The analysis employed concepts and methods from the theory of experimental design. We found that the main factor in the increase in dispersion was a change in the patterns of remuneration within the union and nonunion sectors. The decline in union membership made a smaller contribution, roughly one-quarter as large. The third factor considered, exogenous changes in worker characteristics, had only a small – and usually negative – effect compared with that of the other two factors, indicating a decline in the dispersion of characteristics, and in particular a decline in the dispersion of education. In view of the large contribution from the wage equation, we split the wage effect into four factors, combining the union and characteristics effects into 'other'. Except in the exclusively public-sector groups, we found the largest factor in increasing wage variance to be an increase in the residual variance of the nonunion wage equation, but with substantial supplementary effects coming from increases in the coefficients of the two wage equations.

Along with these numerical findings, the enabling material contributions of this exercise include the two-wage model itself and the resulting three-way variance decomposition; the bivariate kernel-density-based graphical techniques elaborated in the Appendix; and the counterfactual experimental design methodology. These are versatile techniques for the analysis of wage dispersion in partially unionized labor forces.

References

Addison, John T., Ralph W. Bailey, and W. Stanley Siebert. 2007. "The Impact of Deunionisation on Earnings Dispersion Revisited." *Research in Labor Economics* 26: 337-364.

Angrist, Joshua D. and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." In Orley C. Ashenfelter and David Card, eds. Handbook of Labor Economics, Volume 3a. Amsterdam and New York: Elsevier, 1999, pp. 1277-1366.

Bell, Brian D. and Michael K. Pitt. 1998. "Trade Union Decline and Distribution of Wages in the U.K.: Evidence from Kernel Density Estimation." *Oxford Bulletin of Economics and Statistics* 60 (4): 509-528.

Box, George E., J. Stuart Hunter, and William G. Hunter. 2005. *Statistics for Experimenters: Design, Innovation, and Discovery (2nd ed.)*. New York: Wiley.

Canal Dominguez, Juan Francisco and César Rodriguez Gutiérrez. 2004. "Collective Bargaining and Within-firm Wage Dispersion in Spain." *British Journal of Industrial Relations* 42 (3): 481-506.

Card, David. 1996. "The Effect of Unions on the Structure of Wages: a Longitudinal Analysis." *Econometrica* 64 (4): 957-979.

Card, David. 2001. "The Effect of Unions on Wage Inequality in the U.S. Labor Market." *Industrial and Labor Relations Review* 54 (2): 296-315.

Card, David, Thomas Lemieux, and W. Craig Riddell. 2004. "Unions and Wage Inequality." *Journal of Labor Research* 25 (4): 519-559.

Checchi Daniele, Jelle Visser, Herman G. Van de Werfhorst. 2007. "Inequality and Union Membership: The Impact of Relative Earnings Position and Inequality Attitudes." *IZA Discussion Paper No. 2691.* Bonn: Institute for the Study of Labor/IZA, March.

Cox, David R. and N. Reid. 2000. *Theory of Design and Experiments*. Boca Raton, FL: CRC Press.

DiNardo, John, Fortin, Nicole M. and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64 (5): 1001-1044.

DiPrete, Thomas A. 2005. "Labor Markets, Inequality, and Change: A European Perspective." *Work and Occupations* 32 (2): 119-139.

Efron, Bradley 1978. "Regression and ANOVA with Zero-One Data: Measures of Residual Variation." *Journal of the American Statistical Association* 73 (361): 113-121.

Freeman, Richard B. 1980. "Unionism and the Dispersion of Wages." *Industrial and Labor Relations Review* 34 (1): 3-23.

Freeman Richard B. 2005. "What Do Unions Do? The 2004 M-Brane Stringtwister Edition." In James T. Bennett and Bruce E. Kaufman, eds., *What Do Unions Do – A Twenty year Perspective.* New Brunswick and London: Transaction Publishers, pp. 607-636.

Fukunaga, Keinosuke. 1972. *Introduction to Statistical Pattern Recognition.* New York: Academic Press.

Gosling, Amanda and Thomas Lemieux. 2001. "Labor Market Reforms and Changes in Wage Inequality in the United Kingdom and the United States." NBER Working Paper No. 8413, Cambridge, MA: National Bureau of Economic Research.

Lemieux Thomas. 1992. "Unions and Wage Inequality in Canada and the United States." In David Card and Richard B Freeman, eds., *Small Differences that Matter: Labor Markets and Income Maintenance in Canada and the United States.* Chicago, IL: University of Chicago Press, pp. 69-107.

Lewis, H. Gregg. 1986. *Union Relative Wage Effects: A Survey.* Chicago: University of Chicago Press.

Li, Xiang, Nandan Sudarsanam, and Daniel B. Frey. 2006. "Regularities in Data from Factorial Experiments." *Complexity* 11(5): 32-45.

Machin, Stephen. 1997. "The Decline of Labor Market Institutions and the Rise of Wage Inequality in Britain." *European Economic Review* 41(3-5): 647-657.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and Causal Inferences: Methods and Principles for Social Research.* Cambridge: Cambridge University Press.

Silverman, Bernard W.1986. *Density Estimation for Statistics and Data Analysis.* London, New York: Chapman and Hall.

Appendix. The Presentation of Scatter Plots as Functional Relationships.

Here we address an issue of graphical display. In Card's (2001) framework, it is easy to plot the relationship between the skill category c (on the horizontal axis) and other statistics of interest, such as the proportion of workers in a union $u(c)$ or the union-nonunion gap $\Delta_w(c)$. A natural alternative in our framework would be to replace skill category c by worker skill w_i, and to plot scatter diagrams of U_i and $\Delta_{w,i}$ against w_i, where $\Delta_{w,i}$ is the wage gap $w_i^u - w_i^n$. Such plots would however contain a large number of data points, and fail to convey a clear impression of the statistical relationship between the two variables. For example, if union membership declined between two comparator years, such shifts would be difficult to detect in intermingled scatter plots. In order to represent scatter-relationships by graphical functional relationships, we employ techniques from the theory of bivariate kernel-density smoothing.

We begin by stating our main results. Suppose that we have N pairs of points (X_i, Y_i). (Figures 1 and 2 present the examples $X_i = w_i$ (skill); and $Y_i = U_i$ (union density) or $Y_i = \Delta_{w,i}$ (union-nonunion wage gap).) These are used to construct a smooth bivariate distribution describing a pair of random variables (X, Y). Our strategy is to compute and plot $E(Y|X = x)$ against x, yielding a smooth nonlinear curve.

To do so, let the sample variance-covariance matrix of (X_i, Y_i) be $\Sigma = \begin{pmatrix} s_X^2 & s_{XY} \\ s_{YX} & s_Y^2 \end{pmatrix}$. Define $p_i(x) = A_i(x)/A(x)$, where

$$A_i(x) = \exp\left(-\frac{(X_i - x)^2}{2h^2 s_X^2}\right), \; A(x) \equiv \sum_{k=1}^{N} A_k(x).$$

The factor h controls the level of smoothness (high h) or detail (low h) in the resulting graphs. A high N justifies a low choice of h. Following experiment with the graphs' appearance we set $h = 20/\sqrt{N}$, implying a window width of about 0.6 standard deviations when $N = 1000$. Then (as we shall explain)

$$E(Y|X = x) = \sum_{k=1}^{N}\left[Y_k + \frac{s_{YX}}{s_X^2}(x - X_k)\right]p_k(x). \qquad \text{(A.1)}$$

To see how these results are derived, it is helpful first to consider univariate smoothing. Given a sequence $X_1, \ldots, X_N$ of scalar observations, one constructs a density function f_X whose graph provides similar information to a spike graph or a histogram of the observations, but has the advantages of being smooth, and (in common with the two-wage model itself) of being independent of any choice of bins. Thus kernel-density smoothing constructs a density function f_X that may be regarded as an approximation to the density function generating the X_i. However a useful alternative perspective is that f_X describes the *exact* distribution of a constructed random variable X that we generate by adding a random perturbation to the empirical distribution of the X_i. This random perturbation levels out the spike-graph of the empirical distribution into a smooth density function. In detail, the construction of f_X is as follows. First, define the random integer I by equiprobable choice from the population $\{1, \ldots, N\}$. Second, define $X \equiv X_I + \delta$, where δ is some suitably chosen random perturbation, which we take to be independent of I, although this is not essential. Third, compute the density f_X of the random variable X. This density may now be plotted.

To complete the specification of f_X we must choose a density function f_δ, known as the *kernel density function*, for δ. A tractable choice for the kernel density is $\delta \sim N(0, h^2 s_X^2)$, where hs_X is called the *window width* of the process, and s_X^2 is the sample variance of the X_i. The choice of smoothing parameter h is a compromise between the conflicting requirements of smoothness (large h, near 1, say) and detail (small h) in the resulting density graph.

All this can be generalized to the case where, instead of scalars X_i, we have bivariate vectors $Q_i \equiv (X_i, Y_i)$. Now the joint density f_Q takes two arguments. The perturbation δ becomes a bivariate random vector, whose distribution has to be specified. Once this is done we can define $Q = Q_I + \delta$ and compute the bivariate density $f_Q(x, y)$. So far this is a straightforward extension of scalar smoothing. However, with a bivariate distribution we can obtain the conditional distribution of Y given X: this conditional distribution exists because f_δ and f_Q are density functions in their own right, as discussed above. From f_Q we can compute the conditional distribution $f_{Y|X}(y|X = x)$ and hence the conditional expectation $E(Y|X = x)$. Finally, we can plot $E(Y|X = x)$ against x. This is our method for turning the scatter plot of, say, (w_i, U_i) into an informative functional relationship that can be plotted as a single curve.

It is important that the perturbing distribution should be tractable, al-

lowing us to compute analytically the effect of the conditioning information $X = x$. Thus, a convenient choice of density for δ is the bivariate normal. The conditional distributions are then also normal. The particular specification adopted here is $\delta \sim N_2\left(0, h^2\Sigma\right)$, where $\Sigma = \begin{pmatrix} s_X^2 & s_{XY} \\ s_{YX} & s_Y^2 \end{pmatrix}$, the sample variance-covariance matrix of the Q_i. Imposing such a distribution on δ causes the elliptical contours of f_δ to be roughly aligned with the contours suggested by a scatter plot of (X_i, Y_i), although neither these nor the contours of f_Q are in general elliptical.[7] For our actual choice of h see the main body of the text.

In detail, the computation is as follows. To specify the joint distribution of (X, Y), start from the sequence (X_i, Y_i) of N data pairs. A random integer I is chosen as described above. Conditionally on the value $I = i$, (X, Y) has a bivariate normal distribution about (X_i, Y_i)

$$\begin{pmatrix} X \\ Y \end{pmatrix} \Big| I = i \sim N\left(\begin{pmatrix} X_i \\ Y_i \end{pmatrix}, h^2 \begin{pmatrix} s_X^2 & s_{XY} \\ s_{YX} & s_Y^2 \end{pmatrix} \right). \tag{A.2}$$

The aim is to find and graph $E\left(Y|X = x\right)$ against x, so we need the distribution of Y conditional on $X = x$. According to the total probability theorem, the density of this distribution is

$$f_{Y|X}\left(y|X = x\right) = \sum_{k=1}^{N} f_{Y|X,I}\left(y|X = x, I = i\right) p_i\left(x\right), \tag{A.3}$$

where we define $p_i\left(x\right) \equiv P\left(I = i|X = x\right)$. The first factor on the right of (A.3) can be obtained from (A.2) using the main result about the conditional bivariate normal distribution, which here yields for the distribution of Y conditional upon $X = x$ and $I = i$

$$Y|X = x, I = i \sim N\left(Y_i + \frac{s_{XY}}{s_X^2}\left(x - X_i\right), s_Y^2 - \frac{s_{XY}^2}{s_X^2}\right).$$

The second factor on the right of (A.3) can be obtained from Bayes' theorem, which yields

$$p_i\left(x\right) = P\left(I = i|X = x\right) = \frac{f_{X|I}\left(x|I = i\right) P\left(I = i\right)}{\sum_{k=1}^{N} f_{X|I}\left(x|I = k\right) P\left(I = k\right)}. \tag{A.4}$$

---

[7]On this natural choice of Σ, see Fukunaga (1972); Silverman (1984, pp. 77-8).

23

Now conditionally on $I = i$, $X|I = i \sim N(X_i, h^2 s_{XX})$, thus $f_{X|I}(x|I = i) = (2\pi s_X^2)^{-1/2} \exp\left(-(x - X_i)^2 / 2h^2 s_X^2\right)$. Moreover $P(I = i) = N^{-1}$. If we remove the factor $(2\pi s_X^2)^{-1/2} N^{-1}$ common to all terms in the numerator and denominator of (A.4) we have

$$p_i(x) = \frac{A_i(x)}{A(x)},$$

where

$$A_i(x) = \exp\left(-(x - X_i)^2 / 2h^2 s_X^2\right), \ A(x) = \sum_{k=1}^{N} A_k(x).$$

Now use (A.3) and (A.4) to obtain $E(Y|X = x)$. The final result is

$$
\begin{aligned}
E(Y|X = x) &= \sum_{k=1}^{N} E_{Y|X,I}(y|X = x, I = k) \, p_k(x) \\
&= \sum_{k=1}^{N} \left(Y_k + \frac{s_{XY}}{s_X^2}(x - X_k)\right) p_k(x),
\end{aligned}
$$

as stated in section VI.

We note that the method also allows us to compute $Var(Y|X = x)$ and hence to compute and plot confidence bounds about the curve of $E(Y|X = x)$. However, in order to keep the graphs uncluttered, these bounds have been suppressed in the present context, in which the curves for two years are already plotted together.

Figure 1(a)

Union Coverage, All Full-Time Workers.

Figure 1(b)

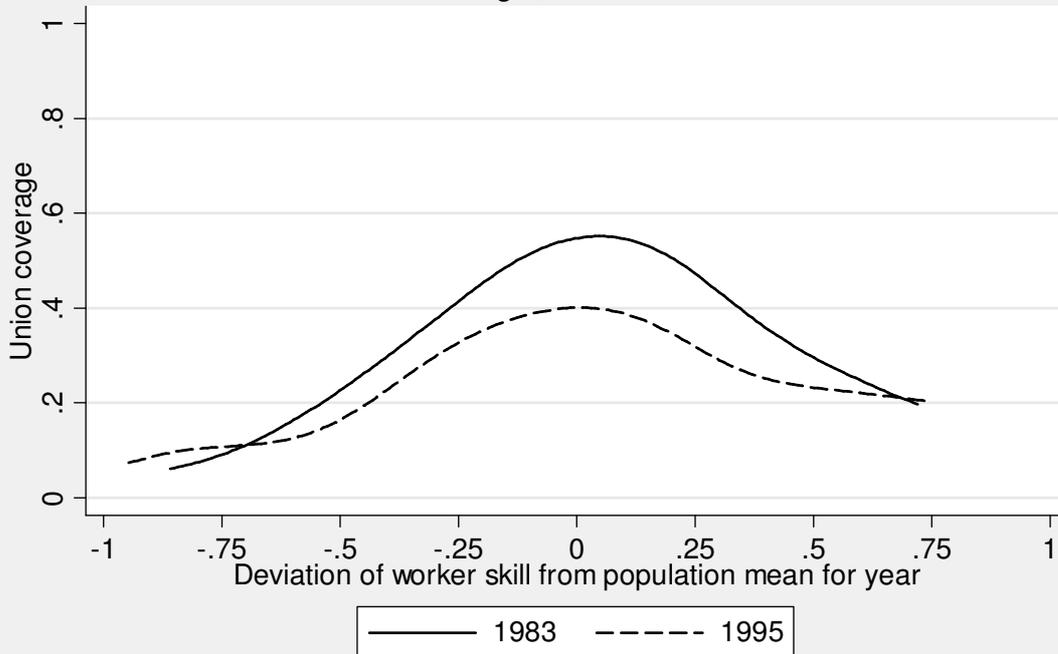Union Coverage, Private-Sector Males.

Figure 1(c)
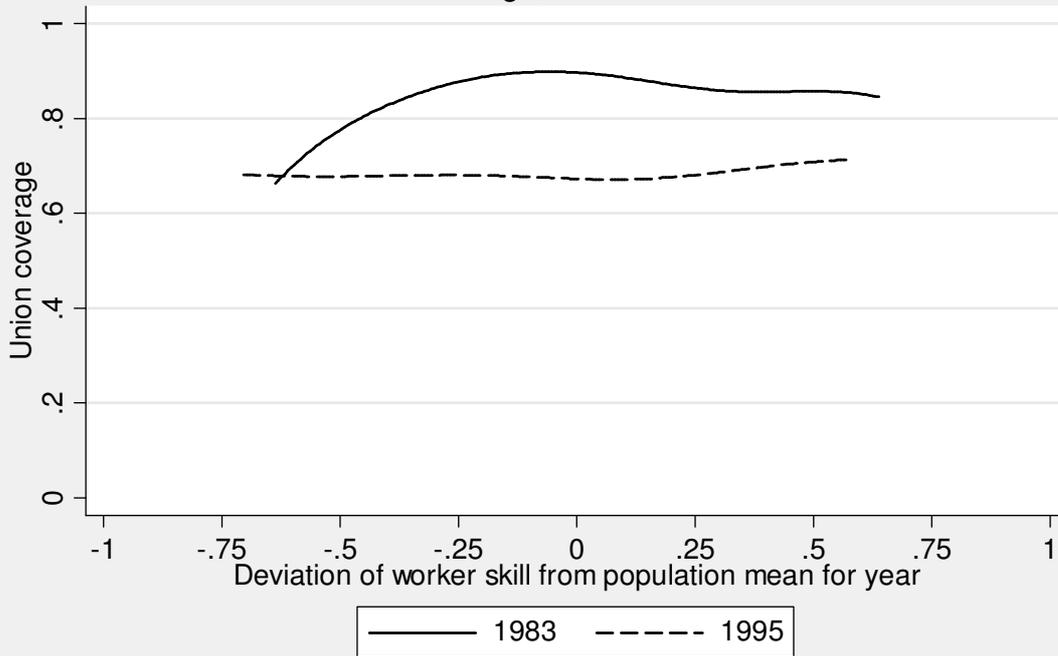### Union Coverage, Public-Sector Males.

Figure 1(d)
Union Coverage, Private-Sector Females.

Figure 1(e)
Union Coverage, Public-Sector Females.

Figure 2(a)
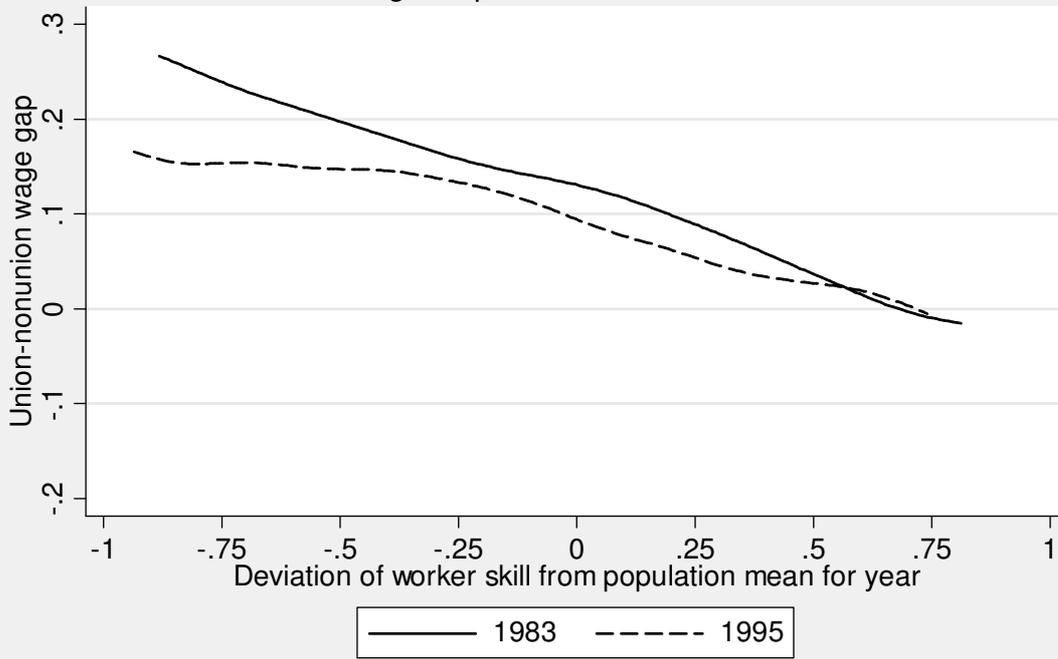### Wage Gap, All Full-Time Workers.

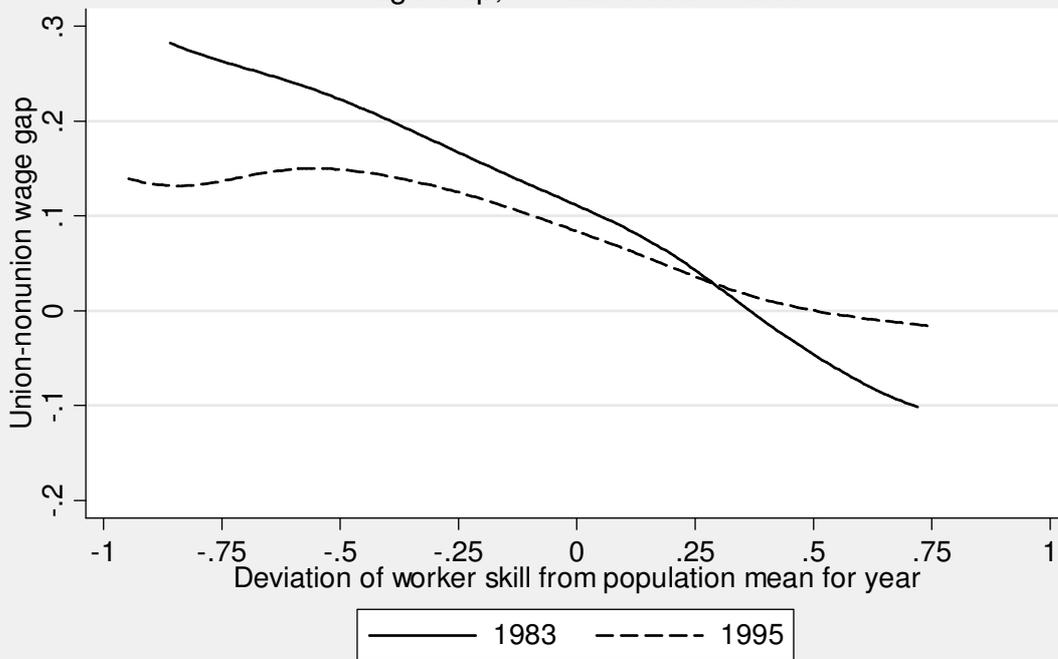## Figure 2(b)
Wage Gap, Private-Sector Males.

Figure 2(c)
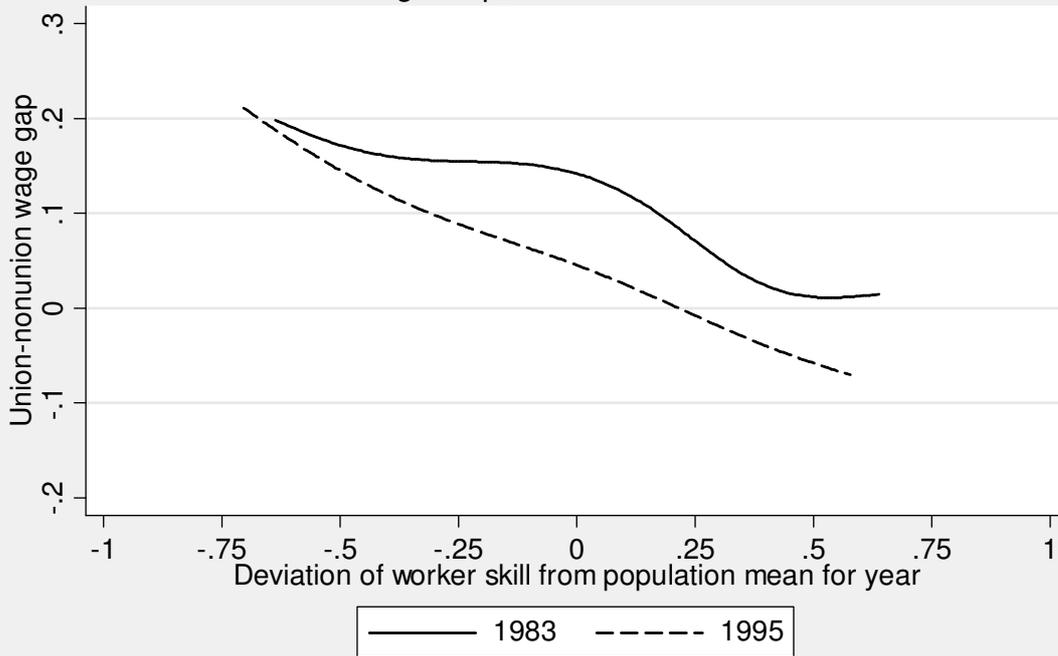Wage Gap, Public-Sector Males.

Figure 2(d)
Wage Gap, Private-Sector Females.

Figure 2(e)
Wage Gap, Public-Sector Females.

Table 1. Probit Analysis of Union Membership of Private-Sector Males, 1983 and 1995.

	1983		1995	
	Coefficient	St. Err.	Coefficient	St. Err.
Constant	31.37	67.2	-46.29	61.4
SCHOOL (= years of education)	-9.839	23.5	13.89	21.2
$SCHOOL^2$	0.9666	3.05	-1.655	2.72
$SCHOOL^3$	-0.03454	0.174	0.08849	0.154
$SCHOOL^4$	0.0002617	0.00368	-0.001781	0.00323
EXP (= years of experience)	0.3289	0.0865***	0.1765	0.0852**
EXP^2	-0.01478	0.00436***	-0.004587	0.00422
EXP^3	0.0003028	0.000114***	0.0000844	0.000119
EXP^4	-2.52×10^{-6}	1.08×10^{-6}**	-7.95×10^{-7}	1.17×10^{-6}
EXP.SCHOOL	-0.009541	0.00528*	-0.006344	0.00562
EXP^2.SCHOOL	0.0002035	0.000113*	0.0000533	0.000117
White	-0.1466	0.144	0.2588	0.163
Married	0.07417	0.0782	0.03005	0.0649
Small firm	-0.7146	0.0655***	-0.5636	0.0678***
Region				
North	0.3613	0.0696***	0.2720	0.0662***
Midlands	0.1487	0.0757**	0.1770	0.0733**
Wales	0.3405	0.141**	0.5250	0.139***
Scotland	0.3529	0.0988***	0.1838	0.0957*
Occupation				
Professional	0.1582	0.146	0.2175	0.116*
Associate professional	0.6942	0.126***	0.3970	0.104***
Clerical	0.4690	0.121***	0.4438	0.118***
Personal Service	0.3946	0.330	-0.4437	0.377
Craft	1.088	0.0936***	0.6501	0.0817***
Operative	1.130	0.111***	0.5953	0.101***
Laborer	1.047	0.163***	0.2290	0.165
Sector				
Agricultural	-0.8221	0.214***	-0.4900	0.263*
Energy and Water Utilities	-0.3005	0.260	1.253	0.207***
Mining	0.1915	0.113*	0.08231	0.226
Construction	-0.7905	0.0975***	-0.2936	0.105***
Distribution	-0.4729	0.0853***	-0.6704	0.0919***
Transport	0.03239	0.135	0.2607	0.0902***
Banking and Finance	0.1441	0.104	-0.1706	0.0792**
Government	-0.2887	0.140*	0.3808	0.177**
N	2708		3057	
Efron Pseudo-R^2	0.285		0.209	

Notes: The dependent variable is a dummy for union membership (=1, 0 otherwise).
 ***, **, * indicate two-tailed statistical significance at the .01, .05, and .10 levels, respectively.

Table 2. Computation of the Three-Way Variance Decomposition for U.K. Private-Sector
Males, 1983 and 1995.

	1983		1995	
Probit analysis of union membership				
Proportion unionized (actual)	0.427		0.296	
Proportion unionized (fitted)	0.429		0.297	
Efron Pseudo-R²	0.285		0.209	
Regression analysis				
Residual union variance σ_u^2	0.099		0.121	
Adjusted R², union sector	0.253		0.351	
Residual nonunion variance σ_n^2	0.125		0.162	
Adjusted R², nonunion sector	0.554		0.533	
Estimated variance decomposition				
Across-worker skill variance	0.104	(46.7%)	0.146	(48.8%)
Within-worker, wage gap variance	0.004	(1.8%)	0.003	(1.0%)
Within-worker, expected residual variance	0.114	(51.5%)	0.150	(50.2%)
Model estimate of total variance	0.222		0.299	
Actual total variance	0.221		0.297	
N	2708		3507	

Table 3. Three-way Decomposition for Subgroups of the U.K. Labor Force, 1983 and 1995.

Sector/Gender	Year	N	$\hat{v}_{skill}$	$\hat{v}_{gap}$	$\hat{v}_{residual}$	$\hat{v}_{total}$	v_{actual}
Private-Sector Males	1983	2708	0.104	0.004	0.114	0.222	0.221
	1995	3057	0.146	0.003	0.150	0.299	0.297
Public-Sector Males	1983	1473	0.083	0.003	0.089	0.175	0.171
	1995	933	0.102	0.006	0.099	0.207	0.198
Private-Sector Females	1983	1242	0.076	0.006	0.080	0.163	0.161
	1995	1574	0.101	0.003	0.135	0.239	0.236
Public-Sector Females	1983	708	0.081	0.007	0.068	0.156	0.148
	1995	883	0.089	0.008	0.089	0.186	0.180
All Male Workers	1983	4181	0.099	0.003	0.106	0.208	0.208
	1995	3990	0.136	0.003	0.142	0.281	0.279
All Female Workers	1983	1950	0.096	0.006	0.078	0.180	0.180
	1995	2457	0.108	0.005	0.121	0.233	0.233
All Private-Sector Workers	1983	3950	0.129	0.005	0.121	0.254	0.254
	1995	4631	0.145	0.003	0.161	0.309	0.308
All Public-Sector Workers	1983	2181	0.092	0.003	0.094	0.189	0.185
	1995	1816	0.099	0.004	0.111	0.209	0.213
All Workers	*1983*	*6131*	*0.121*	*0.004*	*0.113*	*0.239*	*0.238*
	1995	*6447*	*0.133*	*0.003*	*0.150*	*0.286*	*0.286*

Table 4. The Contribution of *U*, *WE*, and *Z* to the Increase in Wage Dispersion in the U.K., 1983 to 1995.

Sector/Gender	Actual increase in variance of the log-wage $v_{1995} - v_{1983}$	Sum of Main Effects $ME_\gamma + ME_\psi + ME_Z$	Main Effect from Changes in Union Probit Coefficients, ME_γ	Main Effect from Changes in the Union and Nonunion Wage Equations, ME_ψ	Main Effect from Changes in Worker Characteristics, ME_Z
Private-Sector Males	0.0759	0.0771	0.0082 (10.6%)	0.0703 (91.1%)	-0.0014 (-1.8%)
Public-Sector Males	0.0267	0.0314	0.0055 (17.5%)	0.0288 (91.7%)	-0.0029 (-9.2%)
Private-Sector Females	0.0760	0.0759	0.0029 (3.8%)	0.0534 (70.3%)	0.0196 (25.8%)
Public-Sector Females	0.0312	0.0304	0.0102 (33.4%)	0.0332 (109.3%)	-0.0130 (-42.7%)
All Male Workers	0.0714	0.0720	0.0119 (16.5%)	0.0535 (74.3%)	0.0066 (9.2%)
All Female Workers	0.0529	0.0539	0.0100 (18.6%)	0.0411 (76.4%)	0.0027 (5.1%)
All Private-Sector Workers	0.0536	0.0537	0.0060 (11.2%)	0.0532 (99.0%)	-0.0055 (-10.2%)
All Public-Sector Workers	0.0236	0.0244	0.0087 (35.5%)	0.0232 (95.0%)	-0.0074 (-30.5%)
All Workers	*0.0473*	*0.0472*	*0.0118 (25.0%)*	*0.0409 (86.8%)*	*-0.0056 (-11.8%)*

Table 5. The Influence of Wage Equation Parameters on the Increase in Wage Dispersion in the U.K., 1983 to 1995.

Sector/Gender	Actual $v_{1995} - v_{1983}$	Sum of Main Effects	Main Effect from σ_u^2	Main Effect from β^u	Main Effect from σ_n^2	Main Effect from β^n	Main Effect from Other Factors (γ, Z)
Private-Sector Males	0.0759	0.0759	0.0081 (10.7%)	0.0167 (22.4%)	0.0236 (31.2%)	0.0223 (29.4%)	0.0051 (6.8%)
Public-Sector Males	0.0267	0.0317	0.0014 (4.5%)	0.0163 (51.4%)	0.0043 (13.5%)	0.0070 (22.2%)	0.0027 (8.4%)
Private-Sector Females	0.0760	0.0765	0.0135 (17.7%)	0.0019 (2.5%)	0.0391 (51.2%)	-0.0006 (-0.8%)	0.0226 (29.5%)
Public-Sector Females	0.0312	0.0295	0.0114 (38.5%)	0.0178 (60.3%)	0.0074 (25.2%)	-0.0042 (-14.2%)	-0.0029 (-9.7%)
All Male Workers	0.0714	0.0714	0.0090 (12.6%)	0.0131 (18.3%)	0.0186 (26.0%)	0.0130 (18.3%)	0.0178 (24.9%)
All Female Workers	0.0529	0.0540	0.0118 (21.9%)	0.0142 (26.3%)	0.0275 (50.9%)	-0.0128 (-23.8%)	0.0133 (24.7%)
All Private-Sector Workers	0.0536	0.0537	0.0102 (18.9%)	0.0119 (22.2%)	0.0259 (48.2%)	0.0059 (11.1%)	-0.0002 (-0.3%)
All Public-Sector Workers	0.0236	0.0242	0.0042 (17.3%)	0.0100 (41.5%)	0.0087 (35.9%)	0.0000 (0.2%)	0.0013 (5.3%)
All Workers	*0.0473*	*0.0472*	*0.0096 (20.5%)*	*0.0096 (20.4%)*	*0.0205 (43.6%)*	*0.0013 (2.8%)*	*0.0060 (12.8%)*