Industry Wage Differentials: How Many, Big and Significant Are They?*

Kevin T. Reilly and Luisa Zanchi
Economics Division
Leeds University Business School
University of Leeds
Leeds LS2 9JT
United Kingdom

December 2000

Major Revision: May 2002
Minor Revision: June 2002

Abstract: In this paper we examine three implementation and interpretation issues associated with Krueger and Summers's (1988) method for calculating interindustry wage differentials. The literature tends to report a less than complete set of industry wage differentials; use the wrong standard errors; and misinterpret the meaning of the industry wage differentials. The solution to the first two issues follows from making explicit the restriction that the employment-weighted average of all industry wage effects is zero, the same restriction that Krueger and Summers are implicitly imposing on industry wage effects. All industries have thus a wage effect relative to an average worker net of any industry effect and correct standard errors are available via the Delta Method. Finally, we propose a method for analysing interindustry wage differentials as actual differences between wage levels expressed in percentage points and not as log points, which is the current misleading standard. Our procedure calculates actual average percentage wage differences by industry and avoids the distortion in differences across industries that log point comparisons engender. An application is provided, using the United States Outgoing Rotation Files of the Current Population Survey for 1989 and 1996 and so updates the work by Krueger and Summers (1988). (JEL: C12 and J31)

* The authors thank participants at seminars at the University of Manchester and Università di Padova and at the conference "Econometrics of Wages" organised by the Applied Econometrics Association, Université Libre de Bruxelles, 28-29 May 2002, for comments received. All errors that remain are the authors' responsibility.

(I) Introduction

Since Krueger and Summers's (1988) seminal work on industry wage differentials there has been a proliferation of studies for specific countries and international comparisons that examine the existence and size of these differentials.[1] In this literature the industry wage effect is measured as the difference between the least squares coefficient and the employment-weighted average of all estimated coefficients on the industry dummy variables in the context of a log-linear wage equation specification.[2] This measure has the double advantage of being more easily interpretable from an economic point of view (Suits, 1984) and independent of the arbitrarily chosen omitted industry dummy. This paper is concerned with three issues that will improve the implementation and interpretation of this important innovation in the industry wage differential literature.

The first issue pursued is the interindustry wage effect associated with the industry that is excluded from the least squares regression used in the Krueger and Summers's transformation. Following Krueger and Summers, researchers have failed to report an interindustry wage effect for this industry (e.g. Goux and Martin, 1999). Yet, as Kennedy (1986) shows in his independent derivation of this dummy variable effect, the industry excluded does have an interindustry wage differential. In particular the restriction that the researcher is imposing is that the employment-weighted average industry wage effect over *all* industries, including the industry that will be excluded from the least squares regression, is zero. This implies that *all* industries, including the industry excluded from the least squares estimation have an industry wage effect. As we will show below the industry wage effect of the industry

[1] A partial list includes: Borland and Suen (1990); Edin and Zetterberg (1992); Winter-Ebmer (1994); Zanchi (1992, 1995); Lausten (1995); Vainiomäki and Laaksonen (1995); Goux and Martin (1999).

[2] This transformation is one possibility that can be derived from the idea of Suits (1984). Kennedy (1986) provided independently a direct precursor of Krueger and Summers's technique.

the researcher chooses to exclude from the least squares regression is equal to the negative of the employment-weighted average of all the industry coefficients that are estimated.

A second problem with many of the studies in the industry wage differential literature is their failure to provide correct standard errors for the Krueger and Summers's transformed industry wage effects (Goux and Martin, 1999). Researchers have tended to use the least squares standard errors associated with the untransformed coefficients. As discussed above, when calculating these industry wage differentials, researchers are making an explicit restriction which implies that one can use the Delta Method (Zanchi, 1998) to provide the proper standard errors needed for inference.

The final issue is that industry differentials in the context of a semilogarithmic wage equation are measured as log points and not as percentage differences in wage levels as they are usually (mis)interpreted. We present the correct transformation for conversion of log points to percentage points for Krueger and Summers's industry wage differentials building on the work of Halvorsen and Palmquist (1980) and Kennedy (1981).

The next section will outline in detail the three points we are making about implementation and interpretation of industry wage effects. Section III then provides an application using the 1989 and 1996 Outgoing Rotation Samples from the Current Population Survey (CPS-ORG) of the United States. The specification and sample restrictions used are identical to those in the Krueger and Summers's (1988) article to allow for a comparison with their results. We have two basic conclusions in this section. First, the traditional use of least squares standard errors tends to underestimate the significance of industry wage effects. Second, for the majority of industries failure to transform the log points into a percentage wage effect does not significantly distort the results. However, in each year there are a reasonable number of large industry wage effects such that the failure to do the transformation results in reporting a wrong percentage wage effect. Section IV provides a brief conclusion.

(II) Theory

A. The Dummy Variable Trap Problem

In the context of an economy with $K+1$ industries, the researcher who wants to estimate industry wage differentials can write a standard log-linear wage regression with the industry effects measured by the parameters of $K+1$ dummy variables representing industry affiliation:

$$\ln w_i = \alpha + x_i{}'\beta + \sum_{j=1}^{K+1} d_{i\,j}\delta_j + u_i \quad i = 1,\ \ldots,\ N, \tag{1}$$

where $\ln\, w_i$ is the log of the wage rate, x_i is a vector of human capital and working condition controls, $d_{i\,j}$, $j = 1, \ldots, K+1$, are a set of exhaustive and mutually exclusive industry dummies, and u_i is a disturbance term $\text{IID}\left(0,\ \sigma_u^2\right)$. In this context the researcher has in mind estimating a “pure” industry effect on log wages, in the sense of separate effects for all $K+1$ industries compared with some “generic” log wage unaffected by industry affiliation.

Equation (1) cannot be estimated by least squares due to the standard dummy trap problem. However, the assumption that the industry dummy variables are mutually exclusive and exhaustive implies that we can write the $(K+1)^{\text{th}}$ industry’s dummy variable as:

$$d_{i\,K+1} = 1 - \sum_{j=1}^{K} d_{i\,j} \tag{2}$$

Substituting equation (2) into equation (1) yields:

$$\begin{aligned} \ln w_i &= \alpha + \delta_{K+1} + x_i{}'\beta + \sum_{j=1}^{K} d_{i\,j}\left(\delta_j - \delta_{K+1}\right) + u_i \\ &= \alpha^* + x_i{}'\beta + \sum_{j=1}^{K} d_{i\,j}\delta_j^* + u_i. \end{aligned} \tag{3}$$

This implies that the omitted industry's effect on wages is captured by the intercept in the regression and the parameters δ^* measure log points wage differentials for all the other industries compared with the omitted industry, i.e. $\delta_j^* = \delta_j - \delta_{K+1}$. In the case of more than one set of mutually exclusive dummies the intercept captures the aggregate effect of all the excluded dummy variables – as well as the continuous explanatory variables – so one cannot estimate the separate effects of the various excluded dummy variables. Further, this implies that the effect on wages of a particular industry can only be estimated as a relative effect with respect to the excluded $(K+1)^{\text{th}}$ industry. As it is well known, this makes the results obtained from equation (3) sensitive to the choice of the excluded industry.

Suits (1984) argues that this non-unique result can be solved if we think about imposing restrictions on the industry coefficients in equation (1)[3]. Kennedy (1986), picking up on this suggestion in the regional dummies context, suggested that a reasonable approach to recovering the pure dummy variable effect is to select a reasonable comparison "group" and to base the restriction on this group. Examining equation (1) carefully, the reasonable restriction to impose is one that implies that comparisons are made between log wages in all the various industries and a log wage that is "net" of all industry effects:

$$\ln w_i = \alpha + x_i ' \beta + u_i \,. \qquad (4)$$

Note that α and β in equation (4) are required to be the same as in equation (1). This reflects the particular meaning given in this context to the expressions "average" employee and log wage "net" of all industry effects. In particular, equation (4) is not the equation that would describe a model that simply ignores industry effects (i.e. where $\delta_j = 0, \forall \; j$), since such a model would have different intercept and slope parameters (i.e. different α and β). The restriction on the industry coefficient implied by choosing a global average net of industry effect on wages is that the employment-weighted average of the industry dummies parameters in equation (1) is zero (Kennedy, 1986: 175):

$$\sum_{j=1}^{K+1} \delta_j \left(\frac{n_j}{N} \right) = \sum_{j=1}^{K+1} \delta_j s_j = 0, \quad (5)$$

where n_j is employment in industry *j* and $N = \sum_{j=1}^{K+1} n_j$, so that s_j is the employment share of industry *j*. Equation (5) is satisfied when industry differentials δ_j represent differences between the log wage in industry *j* and the overall mean log wage net of industry effects, so that the sum of (employment-weighted) wage deviations from the mean wage is zero, which is the implication of equation (5). Implicitly this idea was implemented by Krueger and Summers (1988) in the industry dummy context.

Under the restriction expressed in equation (5), we can then recover the parameters of equation (1) using the parameters of equation (3) as:

$$\begin{aligned} \delta_j &= \delta_j^* - \sum_{j=1}^{K} \delta_j^* s_j \quad j = 1, \ \ldots, \ K \\ \delta_{K+1} &= -\sum_{j=1}^{K} \delta_j^* s_j \quad \& \\ \alpha &= \alpha^* + \sum_{j=1}^{K} \delta_j^* s_j . \end{aligned} \quad (6)$$

The implication of equation (6) is therefore that each δ_j measures the industry wage effect for industry *j* relative to the "average" employee in the economy, as suggested by Krueger and Summers (1988, p. 263) and Kennedy (1986, p. 174).[4] A simple application of OLS to equation (3) and the imposition of the relationships in equation

[3]Suits suggests and implements the restriction that the sum of the dummy variable coefficients is zero.
[4]All three sub-equations in equation (6) are consistent with both equations (1) and (3). If you substitute all the equations in equation (6) into equation (1) and simplify, you obtain equation (3).

(6) gives the researcher – in log points form – estimates of the industry wage effect for a given industry relative to the average employee in the economy, for *all* industries. In fact, the parameters generated by this transformation include a parameter $-\delta_{K+1}$ for the industry omitted from equation (3).[5]

B. Standard Errors for δ_j 's

The next step is to establish appropriate inference procedures so that we can test hypotheses about the transformed industry wage differentials δ_j, $j=1,...,K+1$ of equation (6). Zanchi (1995, 1998) proposes to transform the standard errors of the estimates obtained from equation (3) using the Delta Method.[6] To see how this will yield the appropriate standard errors, we write the first two equations in equation (6) in matrix form as:

$$\delta = Z\delta^* - es'\delta^* = (Z - es')\delta^*, \tag{7}$$

where δ is the $((K+1)\times 1)$ vector of all the $K+1$ parameters δ_j included δ_{K+1}, δ^* is the $(K\times 1)$ vector of the *K* parameters δ_j^*, *Z* is a $((K+1)\times K)$ matrix constructed as the stack of a $(K\times K)$ identity matrix and a $(1\times K)$ row of zeros, *e* is a $((K+1)\times 1)$ vector of ones, and *s* is a $(K\times 1)$ vector with elements s_j representing the employment shares of each of the first *K* industries. The matrix *Z* has the effect of transforming the $(K\times 1)$ vector δ^* into a $((K+1)\times 1)$ vector the $(\text{K+1})^{\text{th}}$ element of which is zero. $es'\delta^*$ is a $((K+1)\times 1)$ vector of elements all equal to $\sum_{j=1}^{K}\delta_j^* s_j$.

[5]Some might find this derivation a little too detailed; however, in the industry wage differential literature there exists confusion as to the correct interpretation of the excluded dummy. In particular this literature has adopted the assumption that excluding the dummy implies that this is equivalent to setting its coefficient to zero (e.g. Krueger and Summers, 1988; Goux and Martin, 1999) and, as we have now shown, this is an incorrect interpretation.

[6]See also Haisken-DeNew and Schmidt (1997), who developed the same estimator of the standard errors in the context of an explicit Restricted Least Squares estimator, rather than the two-step Delta Method developed by Zanchi (1998).

Viewing the problem in the light of equation (7) provides an immediate solution to obtaining correct standard errors for the pure industry wage differential effects, the Delta Method. The researcher can view the matrix $(Z - es')$ as the restrictions being imposed on the least squares industry parameters and an estimate of the standard errors of the δ_j's can be obtained by suitably transforming the estimated standard errors associated with the δ_j^*'s.

When equation (3) is estimated by OLS, we can obtain the estimate of the variance-covariance matrix of its coefficients $\hat{\delta}^*$, $\mathrm{var}(\hat{\delta}^*)$. And using equation (7), the variance-covariance matrix of the estimates of the transformed industry wage differentials $\hat{\delta}$ can be derived as:

$$\mathrm{var}(\hat{\delta}) = (Z - es')\mathrm{var}(\hat{\delta}^*)(Z - es')'. \qquad (8)$$

The square roots of the diagonal elements of equation (8) are the correct estimates of the standard errors of industry wage effects as defined by equation (6). Using these standard errors we can test, in a standard significance framework, whether the wages of employees in each of the total K+1 industries are significantly different from the wage of the "average" employee in the economy.

A final issue that arises from equation (8) is the magnitude of the "mistake" that the researcher makes by using the least squares standard errors associated with the $\hat{\delta}_j^*$'s, as opposed to the correct ones for the $\hat{\delta}_j$'s, as is traditional in the interindustry wage differential literature. As equation (8) makes clear, the discrepancy between $\mathrm{var}(\hat{\delta})$ and $\mathrm{var}(\hat{\delta}^*)$ will be a function of all the elements of the variance-covariance matrix $\mathrm{var}(\hat{\delta}^*)$ (included the covariance terms across the estimated least squares parameters) and the employment shares s_j's observed in all industries. In a situation without explicit restrictions on both the elements of the variance-covariance matrix and the employment shares, there is no unambiguous sign for the difference between $\text{var}(\hat{\delta}_j)$ and $\text{var}(\hat{\delta}^*{}_j)$ (i.e. whether the $\text{var}(\hat{\delta}_j)$'s are smaller or larger than the $\text{var}(\hat{\delta}^*{}_j)$'s) and we leave this as an empirical issue that we will pursue in the next section.

C. Log Point versus Percentage Industry Wage Effects

The problem with the results of equation (6) is that interindustry wage differentials are still expressed in log point form and not as percentage differences in wage levels. This is because the industry variables in equation (1) are dichotomous dummy variables rather than continuous variables and therefore the derivative of the dependent variable with respect to an industry dummy, $\partial \ln w / \partial d_j$, does not exist (Halvorsen and Palmquist, 1980). Thus the usual interpretation of this type of interindustry wage differentials as the percentage effect of industry *j* affiliation on wages is incorrect. According to Halvorsen and Palmquist (1980), industries with positive δ_j's will have their actual industry wage differentials compressed, thus appearing to be closer in wage levels than is actually true. Further, this compression effect increases with the size of the log point industry differential. The reverse is true for industries with negative δ_j's: log point wages will make them look less similar due to the expansion property of the log transformation in the negative range of δ_j.

To eliminate this distortion in interindustry wage differentials we are required to examine equation (1) in the context of its anti-logs:

$$w_i = \left\{ \prod_{j=1}^{K+1} \left[\exp\left(\delta_j\right) \right]^{d_{ij}} \right\} \exp\left(\alpha + x_i{}'\beta + u_i\right). \tag{9}$$

For the employees in a particular industry *j*, for whom $d_{ij} = 1$ and $d_{ik} = 0$ for any $k \neq j$, equation (9) reduces to:

$$w_i^j = \exp(\delta_j)\exp(\alpha + x_i'\beta + u_i), \tag{10}$$

while for the "average" employee in the economy, whose log wage is given by equation (4), equation (9) reduces to:

$$\tilde{w}_i = \exp(\alpha + x_i'\beta + u_i). \tag{11}$$

The relative industry wage effect in percentage points when we compare the employees in industry *j* with the "average" employee in the economy will be therefore:

$$\begin{aligned} g_j = \frac{w_i^j - \tilde{w}_i}{\tilde{w}_i} &= \frac{\exp(\delta_j)\exp(\alpha + x_i'\beta + u_i) - \exp(\alpha + x_i'\beta + u_i)}{\exp(\alpha + x_i'\beta + u_i)} \\ &= \exp(\delta_j) - 1 \quad \forall \quad j = 1,\ \ldots,\ K+1 \ , \end{aligned} \tag{12}$$

and using the relationships in equation (6), we can recover these percentage differences in wage levels from the parameters of equation (3) as:

$$\begin{aligned} g_j &= \exp\left(\delta_j^* - \sum_{j=1}^{K} \delta_j^* s_j\right) - 1 \qquad j = 1, \ldots, K \\ g_{K+1} &= \exp\left(-\sum_{j=1}^{K} \delta_j^* s_j\right) - 1. \end{aligned} \tag{13}$$

Equation (13) thus tells us that, compared to the average employee in the economy, employees in industry *j* obtain a relative wage gain (or loss) of 100* g_j percent for being employed in industry *j*.

However, Kennedy (1981) points out that the relationship expressed by equation (12) is only true for population parameters. When the parameter δ_j is replaced by its estimate $\hat{\delta}_j$, equation (12) will produce a biased estimate of g_j even when $\hat{\delta}_j$ is an unbiased estimate of δ_j. Using work by Goldberger (1968), Kennedy

(1981) suggests an alternative estimator of g_j which is less biased than the one produced by equation (12). Accordingly, the least biased estimator of the percentage effect of industry affiliation on wages is:[7]

$$\hat{g}_j = \exp\left[\hat{\delta}_j - \frac{1}{2}\hat{\text{var}}\left(\hat{\delta}_j\right)\right] - 1 \qquad \forall \quad j = 1, ..., K+1, \tag{14}$$

where $\hat{\text{var}}\left(\hat{\delta}_j\right)$ is an estimate of the variance of $\hat{\delta}_j$. The importance of this modified transformation will depend on the size of the variance term; however, its inclusion raises two issues associated with the relationship between log point and percentage point estimates relative to what would be obtained using equation (12). First, since the variance term is positive, the estimates of g_j obtained by using equation (14) – as suggested by Kennedy (1981) – are (algebraically) smaller than those that would be obtained by using equation (12) – as implicitly suggested by Halvorsen and Palmquist (1980). Second, the inclusion of the variance term removes the monotonic, rank preserving relationship between percentage point and log point differentials established by equation (12). According to equation (12), percentage point differentials are always (algebraically) larger than log point differentials ($g_j > \delta_j$). Equation (14), instead, can generate any relationship between percentage point estimates and log point estimates ($\hat{g}_j \gtreqless \hat{\delta}_j$), depending on the relative size of the variance term and $\hat{\delta}_j$. In theory, a positive log point differential can even become a negative percentage point differential. Our estimates of percentage point differentials are all calculated using equation (14) rather than equation (12). While the results reported in the next section do not contain any such extreme examples, we do report

[7] This equation follows from: $E\left[\exp\left(\hat{\delta}_j\right)\right] = \exp\left[\delta_j + \frac{1}{2}\text{var}\left(\hat{\delta}_j\right)\right]$ (Goldberger, 1968).

cases in which positive log point estimates are indeed larger than the corresponding percentage point estimates.

(III) Interindustry Wage Effect

In this section we purse the issue of industry wage differentials as an example of the procedures proposed in the previous section. In particular, the exercise will have the added benefit of updating and extending the results reported in the seminal article by Krueger and Summers (1988). We will first briefly discuss the data and sample selection procedures and then turn to the estimates of the industry wage differential effects.

Krueger and Summers (1988) used the Out-Going-Rotation sample from the Current Population Survey (ORG-CPS) for the month of May for the years 1974, 1979 and 1984. We have chosen to update their results by drawing samples from ORG-CPS for the month of May in 1989 and 1996.[8] Following Krueger and Summers (1988: 263), the main restrictions that we have imposed on the sample consist in excluding all individuals who work in the agriculture sector, those aged less than 16 years and those who report a usual hourly wage that is less than $1.00 or greater than $250.00. These restrictions, along with a number of other standard restrictions, result in sample sizes of 11,575 males and females for May 1989 and 9,391 individuals of both genders for May 1996.

The specification of the wage equation also follows Krueger and Summers (1988). We use as a dependent variable the log of the (nominal) usual hourly wage, which is usual weekly earnings divided by usual weekly hours of work. The regressors include: years of education and its square; six age dummies; three regional dummies; eight occupation dummies; dummy variables for race, union membership, marital status, gender, veteran status and central city location; and interactions

[8]We use the National Bureau of Economic Research version of these data files. Details are contained in the Appendix and include a precise accounting for sample restrictions on sample size in Table A1.

between gender and marital status, education and age variables.[9] Our treatment of the set of mutually exclusive industry dummy variables is for two levels of aggregation, the one-digit and two-digit levels, and follows the Krueger and Summers's (1988) classification scheme.[10] Complete sets of least squares estimates (using equation (3)) for both levels of aggregation of the industry dummies are reported in the Appendix (in Table A5 for one-digit industry dummies and in Table A6 for two-digit industry dummies). In both cases the excluded industry dummy is for Forestry and Fishery, as in Krueger and Summers's (1988) study.[11]

Table 1 reports our estimates of industry wage differentials for 1989 and 1996, as well as Krueger and Summers's (1988: 264, Table I) results for 1974, 1979 and 1984, in both log point and percentage terms for the one-digit level of aggregation. In terms of sign and relative size the estimates of industry wage differentials are quite similar across all five years. For example, our results confirm the trend observed in Krueger and Summers's results that the Mining and Construction industries have switch positions at the top of the industry differentials ranking from 1979 onwards, although the gap between the two industries seems to have shrunk substantially in most recent years. Our results for Manufacturing in both 1989 and 1996 are very close in both log points and percentage terms to what Krueger and Summers obtained for 1984. Wholesale and Retail Trade is confirmed as the lowest paying industry in the economy. These results are not surprising, since a well known stylised fact about industry wage differentials in the United States is their extreme stability (Zanchi, 1995).

[9] A complete set of descriptive statistics is available in the Appendix, Table A2.

[10] In the Appendix, Table A3 reports our aggregation criteria and sample proportions for the one-digit level. Table A4 reports aggregation rules and sample proportions in each industry at the two-digit level.

For the 1989 and 1996 results the correct standard errors obtained from equation (8) are reported in brackets below our log point parameter estimates. With the exception of the Forestry and Fishery industry in both years, all the parameter estimates are significant at the 95 percent level of confidence and all but one (Mining in 1996) also at the 99 percent level of confidence or better. These results allow us to conclude that significant and large industry wage differentials are paid in the United States for this period relative to the average wage that would be paid net of industry effects.[12] Note that if the least squares standard errors were used instead, as Krueger and Summers (1988) do, none of the industry differentials but that for Mining in 1989 would be statistically significant.

The results of the transformation from log point estimates to percentage differences in wage levels indicate relatively small changes. This follows directly from the nature of the transformation when log point estimates lie in the range $(-0.3, 0.3)$, as our results do. When a log point estimate is in this range the transformation from log points to percentage points will result in a difference in the second decimal place at most. However, with such transformation we now have easy to interpret results. For example, relative to the average employee, in 1989 industry wage differentials range from 30 percent above the average in Mining to 10 percent below the average in Wholesale and Retail Trade.

Table 2 reports Krueger and Summers's (1988: 265-266, Table II) results for 1974, 1979 and 1984 along with our estimates for 1989 and 1996 for the two-digit level of aggregation. Like the one-digit industry results reported in Table 1, our

---

[11] The only difference between our procedure and Krueger and Summers's procedure is that our least squares and Delta Method standard errors are robust (heteroscedasticity consistent), as is now standard practice in the labour econometrics literature.

[12] We cannot make with confidence the same statement for the years 1974, 1979 and 1984 since Krueger and Summers (1988) mistakenly use the least squares standard errors to evaluate the statistical significance of their industry differentials (Zanchi, 1998).

estimates of industry wage differentials are broadly similar to theirs in sign and size. For example, the Petroleum, Mining, Chemical and Transport Machinery industries are confirmed as paying large wage differentials above the average. At the opposite extreme of the ranking, Private Household, Personal Services, Welfare Services and Eating & Drinking are confirmed among the low paying industries with respect to the average. For Education Services, Krueger and Summers identified a downward trend in wages which our 1989 and 1996 results seem to corroborate.

In terms of statistical significance, relying on the transformed standard errors from equation (8), for 1989, 32 out of the 43 estimated industry wage differentials pass standard significance tests, while for the 1996 results only 18 of the 43 parameters are significantly different from zero. Note again that if the least squares standard errors were mistakenly used, only 10 industry differentials would be statistically significant in 1989 and only one differential in 1996. So relative to the average employee the evidence for both years indicates significant industry effects, although the 1996 results do suggest - unlike those at the one-digit level of aggregation - that there might be changes occurring in the structure of industry wages in the United States over time. This seems to suggest that more detailed research on the time structure of industry wage differentials in the United States may be required.

In terms of the transformation from log points to percentage points, changes are marginal in many cases since the range of log point estimates for 1989 is –0.295 (Private Household) to 0.485 (Tobacco) and for 1996 this range is –0.228 (Welfare Services) to 0.523 (Petroleum). However, for particular industries interpreting the log point estimates as percentage differences in wage levels will produce large mistakes with the two-digit level of aggregation, unlike what is observed at the one-digit level of industry aggregation. For Tobacco in 1989, for example, the log point differential underestimates the true percentage point differential by 11% and for Petroleum in 1996 the underestimate is 16%.

(IV) Conclusions

In this paper we have dealt with three implementation and interpretation issues associated with Krueger and Summers's (1988) seminal article on industry wage differentials. We show that making explicit the restriction on the coefficients – that the employment-weighted average industry wage effect over all industries is zero – produces an industry wage effect for *all* industries, including the industry that the researcher must exclude in the estimation to avoid the standard dummy variable trap. Further, by making explicit this restriction, we are able to take advantage of the Delta Method to provide proper standard errors for the Krueger and Summers's industry wage differentials. This gives us a proper framework for inference on industry wage differentials. Finally, we have reminded the literature of a point made originally by Halvorsen and Palmquist (1980) about the correct interpretation of dummy variable coefficients in semilogarithmic equations. If one wants to interpret these results as percentage differences, then a transformation of dummy variable coefficients must be performed. We have extended their result to the context of industry wage differentials, as well as taking account of Kennedy's (1981) criticism.

To explore the implications of these points we update Krueger and Summers's (1988) results for the United States in the years 1989 and 1996. We find that the least squares standard errors seriously overestimate the transformed standard errors using the Delta Method. The implication is that, especially for small industry wage effects, a researcher might mistakenly label as insignificant an effect by using the least squares standard errors. For the transformation from log point to percentage industry wage differentials we find, not surprisingly, that it is only the largest (absolute) values that are significantly affected by the transformation. However, we would argue that given the extreme simplicity of the transformation, this is what researchers in the industry wage differentials literature should get into the habit of doing.

References

Borland, Jeff and Anthony Suen (1990): "The Determinants of Individual Wages in Australia: Competitive and Non-Competitive Influences" Australian Economic Review, 4th Quarter, 33-44.

Edin, Per-Anders and Johnny Zetterberg (1992): "Interindustry Wage Differentials: Evidence from Sweden and a Comparison with the United States" American Economic Review, 82, 5, 1341-9.

Goldberger, Arthur S. (1968): "The Interpretation and Estimation of Cobb-Douglas Functions" Econometrica, 35, 3/4, 464-72.

Goux, Dominique and Eric Martin (1999): "Persistence of Interindustry Wage Differentials: A Reexamination Using Matched Worker-Firm Panel Data" Journal of Labor Economics, 17, 3, 492-533.

Haisken-DeNew, John P. and Christoph M. Schmidt (1997): Interindustry and Interregion Differentials: Mechanics and Interpretation" Review of Economics and Statistics, 79, 3, 516-521.

Halvorsen, Robert and Raymond Palmquist (1980): "The Interpretation of Dummy Variables in Semilogarithmic Equations" American Economic Review, 70, 3, 474-5.

Kennedy, Peter E. (1981): "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations" American Economic Review, 71, 4, 801.

Kennedy, Peter E. (1986): "Interpreting Dummy Variables" Review of Economics and Statistics, 68, 1, 174-5.

Krueger, Alan B. and Lawrence H. Summers (1988) "Efficiency Wages and the Inter-Industry Wage Structure" Econometrica, 56, 2, 259-93.

Lausten, Mette (1995): "Inter-Industry Wage Differentials in Denmark?" University of Aarhus, CLS Working Papers, No. 95-18.

Suits, Daniel (1984): "Dummy Variables: Mechanics V. Interpretation" Review of Economics and Statistics, 66, 1, 177-80.

Vainiomäki, Jari and Seppo Laaksonen (1995): "Inter-Industry Wage Differentials in Finland: Evidence from Longitudinal Census Data for 1975-85" Labour Economics, 2, 161-73.

Winter-Ebmer, Rudolph (1994): "Endogenous Growth, Human capital, and Industry Wages" Bulletin of Economic Research, 46, 4, 289-314.

Zanchi, Luisa (1992): "The Inter-Industry Wage Structure: Empirical Evidence for Germany and a Comparison with the U.S. and Sweden" EUI Working Papers in Economics, No. 92/76.

Zanchi, Luisa (1995): "Cross-Country Comparisons of the Inter-Industry Wage Structure: Empirical Evidence with the 'Luxembourg Income Study' Data-Bank" University of York, Discussions Papers in Economics, No.95/18.

Zanchi, Luisa (1998): "Interindustry Wage Differentials in Dummy Variable Models" Economics Letters, 60, 297-301.

Table 1
Industry Log Point and Percentage Differentials for One-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]
Forestry and Fishery							-.015 (.080)	-.018	.064 (.169)	.051
Mining	.179 {.035}**	.195	.229 {.058}**	.255	.222 {.075}**	.245	.260 {.089}** (.040)**	.296	.130 {.178} (.053)*	.138
Construction	.195 {.021}**	.215	.126 {.031}**	.134	.108 {.034}**	.113	.098 {.082} (.017)**	.103	.106 {.171} (.024)**	.111
Manufacturing	.055 {.020}**	.056	.044 {.029}	.045	.091 {.032}**	.095	.091 {.081} (.008)**	.096	.090 {.169} (.010)**	.094
Transportation, Communications And Public Utilities	.111 {.021}**	.117	.081 {.031}**	.084	.145 {.034}**	.155	.097 {.082} (.014)**	.102	.098 {.170} (.018)**	.102
Wholesale and Retail Trade	-.128 {.020}**	-.120	-.082 {.030}**	-.079	-.111 {.033}**	-.106	-.105 {.080} (.008)**	-.100	-.102 {.169} (.009)**	-.097
Finance, Insurance and Real Estate	.047 {.022}*	.048	-.010 {.035}	-.011	.055 {.034}	.056	.076 {.081} (.014)**	.079	.091 {.170} (.017)**	.095

Continued

Table 1
Industry Log Point and Percentage Differentials for One-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]
Services	-.070 {.021}**	-.068	-.055 {.030}	-.054	-.078 {.032}*	-.076	-.062 {.080} (.007)**	-.061	-.044 {.169} (.008)**	-.043
Standard Deviation of Differentials	.121	.126	.109	.116	.119	.124	.114	.124	.082	.083
Significant Differentials Using Least Squares Standard Errors	7 out of 7		4 out of 7		6 out of 7		1 out of 7		0 out of 7	
Significant Differentials Using Delta Method Standard Errors	NA		NA		NA		7 out of 8		7 out of 8	
Proportion of Standard Errors by Delta Method Smaller Than Those Obtained by Least Squares	NA		NA		NA		100%		100%	
Average Ratio of Least Squares to Delta Method Standard Errors	NA		NA		NA		7.175		12.482	
Average Absolute Value of Difference Log Points – Percent	.008		.006		.007		.009		.005	

[a] *Source*: Krueger and Summers (1988: 264, Table I).
[b] Differentials in percentage points computed according to equation (14). For 1974, 1979 and 1984, since the Delta Method variances $\hat{\text{var}}\left(\hat{\delta}_j\right)$ are not available, the least squares variances $\hat{\text{var}}\left(\hat{\delta}_j^*\right)$ have been used instead.
{} Least squares standard errors. For 1989 and 1996 they are robust standard errors.
() Standard errors transformed according to equation (8), the Delta Method.
* Significant at the 5% level and ** Significant at the 1% level.

Table 2
Industry Log Point and Percentage Differentials for Two-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	**Log Points**	**Percent**[b]	**Log Points**	**Percent**[b]	**Log Points**	**Percent**[b]	**Log Points**	**Percent**[b]	**Log Points**	**Percent**[b]
Forestry and Fishery							-.073 (.080)	-.073	.040 (.172)	.026
Mining	.203 {.022}**	.225	.263 {.031}**	.300	.241 {.033}**	.273	.278 {.089}** (.040)**	.319	.139 {.181} (.053)**	.148
Construction	.228 {.011}**	.256	.137 {.016}**	.147	.126 {.020}**	.134	.111 {.082} (.017)**	.117	.115 {.175} (.024)**	.121
Ordnance	.202 {.040}**	.223	.091 {.067}	.093	NA	NA	.139 {.101} (.061)*	.147	-.008 {.233} (.157)	-.021
Lumber	.003 {.021}	.003	-.035 {.037}	-.035	.001 {.037}	.001	.016 {.085} (.033)	.015	-.014 {.178} (.053)	-.015
Furniture	-.059 {.025}*	-.058	-.120 {.036}**	-.114	-.006 {.048}	-.006	.045 {.087} (.034)	.045	.008 {.179} (.050)	.007
Stone, Clay and Glass	.032 {.022}	.032	.052 {.034}	.053	.085 {.044}	.089	.142 {.089} (.038)**	.152	.055 {.179} (.048)	.055

Continued

Table 2
Industry Log Point and Percentage Differentials for Two-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**
Primary Metals	.082 {.016}**	0.085	.114 {.026}**	.120	.162 {.037}**	.176	.105 {.088} (.035)**	.110	.086 {.181} (.055)	.089
Fabricated Metals	.057 {.015}**	.059	.039 {.026}	.039	.071 {.033}*	.074	.084 {.086} (.030)**	.087	.055 {.176} (.030)	.056
Machinery, Exclusive Electrical	.083 {.013}**	.086	.092 {.022}**	.096	.185 {.024}**	.203	.157 {.083} (.021)**	.170	.105 {.175} (.026)**	.110
Electrical Machinery	.055 {.013}**	.056	.045 {.021}*	.046	.107 {.025}**	.113	.152 {.083} (.022)**	.164	.099 {.175} (.029)**	.104
Transport Machinery	.120 {.014}**	.127	.156 {.021}**	.169	.191 {.025}**	.210	.217 {.083}** (.020)**	.243	.258 {.175} (.028)**	.293
Instruments	.086 {.025}**	.089	.137 {.040}**	.146	.139 {.041}**	.149	.099 {.090} (.040)*	.104	.195 {.177} (.039)**	.215
Misc. Manufacturing	-.116 {.024}**	-.110	-.110 {.042}**	-.105	.014 {.053}	.014	-.114 {.092} (.044)*	-.108	.040 {.182} (.058)	.039

Continued

Table 2
Industry Log Point and Percentage Differentials for Two-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]
Food	.010 {.015}	.010	.019 {.026}	.019	.057 {.027}*	.059	-.011 {.084} (.026)	-.011	.025 {.175} (.029)	.025
Tobacco	-.007 {.063}	-.009	-.040 {.156}	-.051	.340 {.129}**	.405	.485 {.204}* (.187)**	.596	.240 {.209} (.118)*	.262
Textiles	-.010 {.019}	-.010	-.034 {.034}	-.034	.011 {.039}	.011	.017 {.088} (.034)	.017	.025 {.183} (.059)	.024
Apparel	-.087 {.016}**	-.083	-.132 {.030}**	-.124	-.127 {.032}**	-.119	-.104 {.086} (.032)**	-.099	-.092 {.179} (.047)	-.089
Paper	.057 {.020}**	.058	.088 {.033}**	.091	.141 {.039}**	.151	.166 {.087} (.034)**	.180	.144 {.178} (.040)**	.153
Printing	.052 {.017}**	.053	.039 {.028}	.039	.092 {.028}**	.096	.013 {.085} (.030)	.012	.033 {.176} (.036)	.033
Chemical	.157 {.018}**	.170	.148 {.029}**	.159	.221 {.033}**	.247	.290 {.084}** (.026)**	.336	.182 {.175} (.034)**	.199

Continued

Table 2
Industry Log Point and Percentage Differentials for Two-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**
Petroleum	.238 {.036}**	.268	.278 {.062}**	.318	.371 {.073}**	.449	.159 {.101} (.062)*	.170	.523 {.197}** (.095)**	.679
Rubber	.007 {.021}	.007	.023 {.036}	.023	.054 {.041}	.055	.108 {.087} (.034)**	.114	.042 {.178} (.043)	.042
Leather	-.097 {.034}**	-.093	-.233 {.051}**	-.209	-.082 {.060}	-.079	-.005 {.107} (.070)	-.008	-.146 {.189} (.077)	-.139
Railroad	.200 {.023}**	.221	.120 {.037}**	.127	NA	NA	.191 {.099} (.058)**	.209	.043 {.195} (.088)	.040
Other Transport	.090 {.014}**	.094	.120 {.022}**	.127	.132 {.022}**	.141	.067 {.083} (.022)**	.069	-.002 {.174} (.025)	-.002
Communications	.159 {.016}**	.172	.064 {.027}*	.066	.171 {.029}**	.186	.120 {.085} (.029)**	.127	.188 {.176} (.035)**	.206
Public Utilities	.138 {.021}**	.148	.068 {.028}*	.070	.259 {.032}**	.296	.208 {.085}* (.027)**	.231	.289 {.177} (.038)**	.334

Continued

Table 2
Industry Log Point and Percentage Differentials for Two-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]	Log Points	Percent[b]
Wholesale Trade	.035 {.012}**	.036	-.015 {.020}	-.015	.047 {.020}*	.048	.101 {.082} (.017)**	.106	.026 {.174} (.021)	.026
Eating & Drinking	-.267 {.012}**	-.234	-.125 {.020}**	-.118	-.189 {.021}**	-.172	-.212 {.082}** (.017)**	-.191	-.127 {.173} (.019)**	-.119
Other Retail	-.141 {.030}**	-.132	-.093 {.050}	-.090	-.155 {.067}*	-.144	-.156 {.081} (.011)**	-.144	-.147 {.173} (.013)**	-.136
Banking	.081 {.014}**	.084	-.063 {.031}*	-.062	.064 {.022}**	.066	.077 {.083} (.020)**	.079	.050 {.174} (.027)	.051
Insurance (including Real Estate)	.048 {.013}**	.049	.022 {.027}	.022	.071 {.021}**	.074	.093 {.082} (.020)**	.097	.128 {.174} (.023)**	.136
Private Household	-.151 {.019}**	-.140	-.259 {.034}**	-.229	-.366 {.033}**	-.306	-.295 {.106}** (.063)**	-.257	-.123 {.185} (.077)	-.119
Business Services	-.053 {.016}**	-.052	-.067 {.028}*	-.065	.000 {.023}	.000	-.028 {.082} (.018)	-.028	-.025 {.174} (.020)	-.025

Continued

Table 2
Industry Log Point and Percentage Differentials for Two-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**
Repair Services	-.126 {.021}**	-.119	-.026 {.032}	-.026	-.056 {.034}	-.054	-.065 {.087} (.034)	-.064	.001 {.178} (.041)	.000
Personal Services	-.216 {.015}**	-.194	-.107 {.025}**	-.102	-.154 {.025}**	-.143	-.166 {.082}* (.022)**	-.153	-.093 {.174} (.026)**	-.089
Entertainment	-.145 {.023}**	-.135	-.078 {.036}*	-.076	-.141 {.034}**	-.132	-.124 {.085} (.039)**	-.118	-.078 {.177} (.044)	-.076
Medical Services	-.052 {.015}**	-.051	-.039 {.022}	-.038	-.082 {.023}**	-.079	.016 {.083} (.022)	.016	.037 {.174} (.025)	.038
Hospitals (including Nursing)	.039 {.013}**	.040	.063 {.018}**	.065	.059 {.023}*	.061	.034 {.082} (.014)*	.034	.032 {.173} (.017)	.032
Welfare Services	-.333 {.022}**	-.283	-.190 {.032}**	-.173	-.246 {.027}**	-.218	-.242 {.086}** (.030)**	-.215	-.228 {.175} (.035)**	-.204
Education Services (including Libraries & Museums)	-.127 {.016}**	-.119	-.185 {.019}**	-.169	-.194 {.028}**	-.176	-.239 {.085}** (.027)**	-.213	-.216 {.174} (.027)**	-.195

Continued

Table 2
Industry Log Point and Percentage Differentials for Two-Digit Level of Aggregation

Year	1974[a]		1979[a]		1984[a]		1989		1996	
Industry Category	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**	**Log Points**	**Percent[b]**
Professional Services	.085 {.016}**	.089	.060 {.029}*	.061	.062 {.026}*	.064	-.013 {.083} (.024)	-.013	.002 {.174} (.027)	.001
Standard Deviation of Differentials	.134	.133	.122	.124	.159	.163	.157	.165	.140	.156
Significant Differentials Using Least Squares Standard Errors	36 out of 42		28 out of 42		31 out of 42		10 out of 42		1 out of 42	
Significant Differentials Using Delta Method Standard Errors	NA		NA		NA		32 out of 43		18 out of 43	
Proportion of Standard Errors by Delta Method Smaller Than Those Obtained by Least Squares	NA		NA		NA		100%		100%	
Average Ratio of Least Squares to Delta Method Standard Errors	NA		NA		NA		3.146		5.257	
Average Absolute Value of Difference Log Points - Percent	.009		.007		.013		.012		.011	

[a] *Source*: Krueger and Summers (1988: 265-266, Table II).
[b] Differentials in percentage points computed according to equation (14). For 1974, 1979 and 1984, since the Delta Method variances $\hat{\mathrm{var}}\left(\hat{\delta}_j\right)$ are not available, the least squares variances $\hat{\mathrm{var}}\left(\hat{\delta}_j^*\right)$ have been used instead.
{} Least squares standard errors. For 1989 and 1996 they are robust standard errors.
() Standard errors transformed according to equation (8), the Delta Method.
* Significant at the 5% level and ** Significant at the 1% level.

Appendix To:

Industry Wage Differentials: How Many, Big and Significant Are They?

We used the May 1989 and May 1996 Outgoing Rotation Samples of the Current Population Survey (CPS). The CPS is a monthly survey of households in the United States, covering 50 states, and is primarily used to produce an estimate of the unemployment rate. For the May 1989 CPS the initial sample size is 60,000 households and for May 1996 the initial sample size is roughly 50,000 households. The difference in basic sample sizes is the result of a change introduced in April 1994, when the CPS basic sample was reduced by 10,000 households. For a sample design reason the reduction was only completed in October 1995. Households are interviewed for four months, leave the sample for eight months and then reappear for another four monthly interviews. After the last of eight interviews the household is dropped from the sample and replaced by another household. On both occasions, when a household leaves the sample its members are asked a series of questions that makes it possible to estimate a standard wage equation as done by ourselves and many other researchers. The versions of the data files used were prepared by and distributed by the National Bureau of Economic Research.

The sample selection criteria used are identical to Krueger and Summers's (1988) and the resulting samples refer to all private non-agricultural employees 16 years of age or older. Other sample restrictions imposed included the elimination of all observations for which any of the variables used contained missing values, usual weekly earnings were reported as equal to zero, usual hours of work were reported as zero or top-coded at 99 hours per week, and the nominal hourly wage rate (usual weekly earnings divided by usual weekly hours of work) were less than $1.00 or greater than $250.00. Table A1 provides a complete account of the effect on our samples of the various restrictions.

Table A2 outlines the variable definitions and descriptive statistics of the variables used in the analysis with the exception of the industry variables.

Tables A3 and A4 outline our aggregation of the IND80 variable (1980 Standard Industrial Classification system) in the CPS-ORG files into the categories used by Krueger and Summers (1988), including the proportion of each industrial category observed in the two samples.

Tables A5 reports the least squares estimates of our wage equation for the one-digit level of aggregation of the industry dummies and Table A6 reports the results at the two-digit level of aggregation.

Table A1: Sample Size Effects of Sample Restrictions

Restriction	**Sample Size**	
	May 1989	**May 1996**
Observations Available	26,212	22,859
Eliminating all individuals not classified as employed.	16,381	14,217
Eliminating all individuals not classified as private sector employees	11,784	10,358
Eliminating all agricultural workers (SIC between 10 and 20)	11,611	10,157
Eliminating individuals with missing values all variables used in the analysis.	11,611	9,454
Eliminating individuals whose usual weekly earnings equal to zero.	11,611	9,426
Eliminating individuals with usual hours per week equal to zero or top-coded at 99.	11,594	9,412
Eliminating individuals whose hourly wage less than $1 or greater than $250	**11,575**	**9,391**

Table A2: Descriptive Statistics

Variables	**May 89**	**May 96**
Nominal Usual Hourly Wage (Usual Weekly Earnings/Usual Hours per Week)	9.91 (7.01)	12.86 (9.25)
Years of Education	12.96 (2.50)	13.01 (2.43)
Age 16 to 21	0.12 (0.32)	0.09 (0.29)
Age 22 to 29	0.23 (0.42)	0.20 (0.40)
Age 30 to 39	0.28 (0.45)	0.30 (0.46)
Age 40 to 49	0.20 (0.40)	0.23 (0.42)
Age 50 to 59	0.12 (0.32)	0.13 (0.34)
Age 60 to 64	0.03 (0.18)	0.03 (0.17)
Age > 64	0.02 (0.15)	0.02 (0.14)
Female	0.46 (0.50)	0.48 (0.50)
Single	0.29 (0.45)	0.26 (0.44)
Not a Veteran	0.86 (0.35)	0.89 (0.31)
White	0.89 (0.32)	0.87 (0.34)
Union Member	0.12 (0.33)	0.10 (0.31)
Live in City Centre	0.21 (0.41)	0.24 (0.43)

Table A2: Descriptive Statistics (continued)

Variables	**May 89**	**May 96**
Managerial & Professional Occupation (OCC80>2 & OCC80<200)	0.20 (0.40)	0.25 (0.43)
Technical Occupation (OCC80>202 & OCC80<236)	0.04 (0.19)	0.03 (0.17)
Sales Occupation (OCC80>242 & OCC80<286)	0.14 (0.34)	0.14 (0.35)
Administrative Support Occupation (OCC80>302 & OCC80<390)	0.16 (0.37)	0.16 (0.36)
Services Occupation (OCC80>403 & OCC80<470)	0.13 (0.34)	0.13 (0.34)
Forestry & Fishing Occupation (OCC80>482 & OCC80<500	0.01 (0.10)	0.005 (0.07)
Process, Craft & Repair Occupation ((OCC80>502 & OCC80<550) or (OCC80>612 & OCC80<700)	0.09 (0.28)	0.08 (0.27)
Construction Occupation (OCC80>552 & OCC80<600)	0.04 (0.20)	0.04 (0.19)
Operators, Fabricators & Laborers Occupation (OCC80>702 & OCC80<889)	0.20 (0.40)	0.17 (0.37)
Northeast	0.24 (0.42)	0.23 (0.42)
Central	0.26 (0.44)	0.26 (0.44)
Southern	0.33 (0.47)	0.28 (0.45)
Western	0.18 (0.38)	0.22 (0.42)

(.) Standard Deviation

Table A3: One-Digit Level Aggregation

One-Digit Category	IND80: SIC Codes	May 1989 Proportion	May 1996 Proportion
Forestry and Fishery	22 to 32	0.006	0.002
Mining	40 to 50	0.011	0.007
Construction	60	0.062	0.056
Manufacturing	100 to 392	0.234	0.214
Transportation, Communications and Public Utilities	400 to 472	0.074	0.069
Wholesale and Retail Trade	500 to 691	0.250	0.258
Finance, Insurance and Real Estate	700 to 712	0.082	0.078
Services	721 to 893	0.280	0.316

Table A4: Two-Digit Level of Aggregation

Two-Digit Category	IND80: SIC Codes	May 1989 Proportion	May 1996 Proportion
Forestry and Fishery	22-32	0.006	0.002
Mining	40-50	0.011	0.007
Construction	60	0.062	0.056
Ordnance	292	0.001	0.001
Lumber	230-241	0.010	0.007
Furniture	242	0.007	0.007
Stone, Clay and Glass	250-262	0.006	0.007
Primary Metals	270-280	0.008	0.007
Fabricated Metals	281-291&300-301	0.012	0.014
Machinery, Exclusive Electrical	310-332	0.027	0.027
Electrical Machinery	340-350	0.021	0.019
Transport Machinery	351-370	0.027	0.022
Instruments	371-382	0.007	0.009
Misc. Manufacturing	390-392	0.006	0.005
Food	100-122	0.020	0.020
Tobacco	130	0.001	0.001
Textiles	132-150	0.008	0.005
Apparel	151-152	0.012	0.009
Paper	160-162	0.010	0.007
Printing	171-172	0.021	0.020
Chemical	180-192	0.018	0.015
Petroleum	200-201	0.002	0.002

Table A4: Two-Digit Level of Aggregation

Two-Digit Category	IND80: SIC Codes	May 1989 Proportion	May 1996 Proportion
Rubber	210-212	0.008	0.009
Leather	220-222	0.002	0.001
Railroad	400	0.004	0.002
Other Transport	401-432	0.037	0.035
Communications	440-442	0.019	0.019
Public Utilities	460-472	0.014	0.012
Wholesale Trade	500-571	0.049	0.048
Eating & Drinking	641	0.058	0.062
Other Retail	580-640 & 642-691	0.144	0.148
Banking	700-709	0.033	0.032
Insurance (including Real Estate)	710-712	0.049	0.046
Private Household	761	0.011	0.007
Business Services	721-742	0.045	0.048
Repair Services	750-752	0.015	0.014
Personal Services	762-791	0.031	0.027
Entertainment	800-802	0.011	0.018
Medical Services	812-830 & 840	0.028	0.040
Hospitals (including Nursing)	831-832	0.061	0.066
Welfare Services	862-871	0.014	0.020
Education Services (including libraries & Museums)	842-861 & 872	0.026	0.030
Professional Services	841 & 873-893	0.039	0.045

Table A5: Wage Equation Results with One-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Years of Education	0.023 (0.011) [0.037]	0.001 (0.012) [0.961]
Square of Years of Education(*1/10)	0.011 (0.005) [0.018]	0.025 (0.005) [0.000]
Age 16 to 21	-0.199 (0.020) [0.000]	-0.125 (0.026) [0.000]
Age 30 to 39	0.164 (0.015) [0.000]	0.205 (0.019) [0.000]
Age 40 to 49	0.243 (0.019) [0.000]	0.272 (0.021) [0.000]
Age 50 to 59	0.261 (0.022) [0.000]	0.306 (0.026) [0.000]
Age 60 to 64	0.188 (0.037) [0.000]	0.221 (0.058) [0.000]
Age > 64	-0.044 (0.051) [0.384]	-0.026 (0.059) [0.653]
Northeast	0.115 (0.010) [0.000]	0.115 (0.013) [0.000]
Central	0.017 (0.010) [0.105]	0.065 (0.013) [0.000]

Table A5: Wage Equation Results with One-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage

Independent Variables	**May 1989**	**May 1996**
Western	0.053 (0.012) [0.000]	0.069 (0.014) [0.000]
Technical Occupation	-0.047 (0.022) [0.035]	-0.041 (0.026) [0.112]
Sales Occupation	-0.284 (0.017) [0.000]	-0.207 (0.020) [0.000]
Administrative Support Occupation	-0.270 (0.015) [0.000]	-0.231 (0.016) [0.000]
Services Occupation	-0.493 (0.017) [0.000]	-0.395 (0.019) [0.000]
Forestry & Fishing Occupation	-0.389 (0.079) [0.000]	-0.496 (0.101) [0.000]
Process, Craft & Repair Occupation	-0.213 (0.018) [0.000]	-0.164 (0.020) [0.000]
Construction Occupation	-0.236 (0.024) [0.000]	-0.221 (0.031) [0.000]
Operators, Fabricators & Laborers Occupation	-0.414 (0.016) [0.000]	-0.366 (0.018) [0.000]
White	0.049 (0.013) [0.000]	0.085 (0.015) [0.000]
Union Member	0.216 (0.012) [0.000]	0.178 (0.015) [0.000]

Table A5: Wage Equation Results with One-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Single	-0.132 (0.016) [0.000]	-0.153 (0.019) [0.000]
Female	-0.056 (0.111) [0.617]	-0.285 (0.121) [0.019]
Not a Veteran	0.006 (0.014) [0.667]	0.019 (0.018) [0.277]
Live in City Centre	0.035 (0.010) [0.001]	0.003 (0.012) [0.799]
Single*Female	0.110 (0.022) [0.000]	0.129 (0.026) [0.000]
Years of Education*Female	-0.019 (0.017) [0.262]	0.004 (0.020) [0.843]
Square of Years of Education*Female	0.001 (0.001) [0.354]	0.000 (0.001) [0.870]
Age 16 to 21*Female	-0.026 (0.027) [0.329]	-0.020 (0.037) [0.580]
Age 30 to 39*Female	-0.057 (0.022) [0.009]	-0.056 (0.026) [0.032]
Age 40 to 49*Female	-0.119 (0.026) [0.000]	-0.099 (0.030) [0.001]
Age 50 to 59*Female	-0.191 (0.031) [0.000]	-0.116 (0.035) [0.001]

Table A5: Wage Equation Results with One-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Age 60 to 64*Female	-0.160 (0.053) [0.003]	-0.076 (0.072) [0.288]
Age > 64*Female	-0.047 (0.069) [0.490]	0.017 (0.096) [0.864]
Mining	0.275 (0.089) [0.002]	0.066 (0.178) [0.709]
Construction	0.113 (0.082) [0.168]	0.042 (0.171) [0.807]
Manufacturing	0.106 (0.081) [0.187]	0.026 (0.169) [0.879]
Transportation, Communications and Public Utilities	0.112 (0.082) [0.169]	0.034 (0.170) [0.844]
Wholesale and Retail Trade	-0.090 (0.080) [0.263]	-0.166 (0.169) [0.326]
Finance, Insurance and Real Estate	0.091 (0.081) [0.262]	0.027 (0.170) [0.873]
Services	-0.048 (0.080) [0.555]	-0.108 (0.169) [0.523]
Constant	1.588 (0.113) [0.000]	1.565 (0.183) [0.000]
R^2	0.469	0.435
Number of Observations	11,575	9,391

(.) Robust Standard Error and [.] Probability Value.

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Years of Education	0.014 (0.011) [0.218]	-0.004 (0.012) [0.712]
Square of Years of Education(*1/10)	0.014 (0.005) [0.003]	0.026 (0.005) [0.000]
Age 16 to 21	-0.181 (0.020) [0.000]	-0.121 (0.026) [0.000]
Age 30 to 39	0.157 (0.015) [0.000]	0.201 (0.018) [0.000]
Age 40 to 49	0.230 (0.019) [0.000]	0.264 (0.021) [0.000]
Age 50 to 59	0.250 (0.022) [0.000]	0.298 (0.026) [0.000]
Age 60 to 64	0.172 (0.037) [0.000]	0.222 (0.058) [0.000]
Age > 64	-0.045 (0.051) [0.378]	-0.021 (0.059) [0.715]
Northeast	0.118 (0.010) [0.000]	0.115 (0.013) [0.000]
Central	0.012 (0.010) [0.236]	0.059 (0.012) [0.000]

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Western	0.058 (0.012) [0.000]	0.069 (0.013) [0.000]
Technical Occupation	-0.094 (0.022) [0.000]	-0.088 (0.026) [0.001]
Sales Occupation	-0.278 (0.018) [0.000]	-0.199 (0.021) [0.000]
Administrative Support Occupation	-0.296 (0.015) [0.000]	-0.238 (0.016) [0.000]
Services Occupation	-0.449 (0.017) [0.000]	-0.388 (0.019) [0.000]
Forestry & Fishing Occupation	-0.315 (0.079) [0.000]	-0.474 (0.103) [0.000]
Process, Craft & Repair Occupation	-0.215 (0.018) [0.000]	-0.168 (0.021) [0.000]
Construction Occupation	-0.241 (0.024) [0.000]	-0.231 (0.031) [0.000]
Operators, Fabricators & Laborers Occupation	-0.410 (0.016) [0.000]	-0.351 (0.019) [0.000]
White	0.047 (0.012) [0.000]	0.084 (0.014) [0.000]
Union Member	0.203 (0.012) [0.000]	0.169 (0.015) [0.000]

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Single	-0.119 (0.016) [0.000]	-0.140 (0.019) [0.000]
Female	-0.015 (0.110) [0.889]	-0.225 (0.120) [0.060]
Not a Veteran	0.014 (0.014) [0.317]	0.023 (0.018) [0.201]
Live in City Centre	0.036 (0.010) [0.000]	0.004 (0.011) [0.759]
Single*Female	0.099 (0.022) [0.000]	0.119 (0.026) [0.000]
Years of Education*Female	-0.021 (0.017) [0.214]	-0.005 (0.019) [0.784]
Square of Years of Education*Female	0.001 (0.001) [0.331]	0.001 (0.001) [0.505]
Age 16 to 21*Female	-0.019 (0.026) [0.471]	-0.006 (0.037) [0.859]
Age 30 to 39*Female	-0.059 (0.022) [0.007]	-0.058 (0.026) [0.026]
Age 40 to 49*Female	-0.114 (0.025) [0.000]	-0.097 (0.029) [0.001]
Age 50 to 59*Female	-0.178 (0.030) [0.000]	-0.105 (0.035) [0.003]

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Age 60 to 64*Female	-0.138 (0.053) [0.009]	-0.061 (0.072) [0.392]
Age > 64*Female	-0.029 (0.068) [0.669]	0.036 (0.097) [0.710]
Mining	0.351 (0.089) [0.000]	0.099 (0.181) [0.584]
Construction	0.184 (0.082) [0.024]	0.075 (0.175) [0.669]
Ordnance	0.212 (0.101) [0.035]	-0.049 (0.233) [0.835]
Lumber	0.089 (0.085) [0.295]	-0.054 (0.178) [0.762]
Furniture	0.118 (0.087) [0.177]	-0.032 (0.179) [0.858]
Stone, Clay and Glass	0.215 (0.089) [0.015]	0.015 (0.179) [0.935]
Primary Metals	0.179 (0.088) [0.042]	0.046 (0.181) [0.799]
Fabricated Metals	0.157 (0.086) [0.066]	0.015 (0.176) [0.932]
Machinery, Exclusive Electrical	0.230 (0.083) [0.006]	0.065 (0.175) [0.711]

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Electrical Machinery	0.225 (0.083) [0.007]	0.059 (0.175) [0.735]
Transport Machinery	0.291 (0.083) [0.000]	0.217 (0.175) [0.214]
Instruments	0.172 (0.090) [0.054]	0.155 (0.177) [0.381]
Misc. Manufacturing	-0.041 (0.092) [0.659]	-0.000 (0.182) [0.999]
Food	0.062 (0.084) [0.459]	-0.015 (0.175) [0.930]
Tobacco	0.558 (0.204) [0.006]	0.199 (0.209) [0.339]
Textiles	0.090 (0.088) [0.302]	-0.015 (0.183) [0.935]
Apparel	-0.030 (0.086) [0.725]	-0.132 (0.179) [0.459]
Paper	0.239 (0.087) [0.006]	0.103 (0.178) [0.560]
Printing	0.086 (0.085) [0.316]	-0.007 (0.176) [0.969]
Chemical	0.364 (0.084) [0.000]	0.142 (0.175) [0.417]

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Petroleum	0.232 (0.101) [0.022]	0.483 (0.197) [0.015]
Rubber	0.181 (0.087) [0.037]	0.002 (0.178) [0.991]
Leather	0.068 (0.107) [0.523]	-0.186 (0.189) [0.325]
Railroad	0.264 (0.099) [0.008]	0.003 (0.195) [0.989]
Other Transport	0.140 (0.083) [0.092]	-0.042 (0.174) [0.810]
Communications	0.193 (0.085) [0.024]	0.148 (0.176) [0.400]
Public Utilities	0.282 (0.085) [0.001]	0.249 (0.177) [0.160]
Wholesale Trade	0.174 (0.082) [0.033]	-0.014 (0.174) [0.934]
Eating & Drinking	-0.139 (0.082) [0.091]	-0.167 (0.173) [0.336]
Other Retail	-0.083 (0.081) [0.305]	-0.187 (0.173) [0.280]
Banking	0.150 (0.083) [0.070]	0.010 (0.174) [0.956]

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Insurance (including Real Estate)	0.166 (0.082) [0.043]	0.087 (0.174) [0.615]
Private Household	-0.222 (0.106) [0.037]	-0.163 (0.185) [0.377]
Business Services	0.045 (0.082) [0.583]	-0.065 (0.174) [0.706]
Repair Services	0.008 (0.087) [0.928]	-0.040 (0.178) [0.824]
Personal Services	-0.092 (0.082) [0.262]	-0.133 (0.174) [0.445]
Entertainment	-0.051 (0.085) [0.544]	-0.118 (0.177) [0.505]
Medical Services	0.089 (0.083) [0.285]	-0.003 (0.174) [0.987]
Hospitals (including Nursing)	0.107 (0.082) [0.190]	-0.008 (0.173) [0.962]
Welfare Services	-0.169 (0.086) [0.049]	-0.268 (0.175) [0.127]
Education Services (including libraries & Museums)	-0.165 (0.085) [0.051]	-0.256 (0.174) [0.142]
Professional Services	0.061 (0.083) [0.468]	-0.038 (0.174) [0.826]

Table A6: Wage Equation Results with Two-Digit Industry Dummies

Dependent Variable: Natural Log of Usual Hourly Wage		
Independent Variables	**May 1989**	**May 1996**
Constant	1.587 (0.114) [0.000]	1.579 (0.185) [0.000]
R^2	0.494	0.453
Number of Observations	11,575	9,391

(.) Robust Standard Error and [.] Probability Value.