# 294
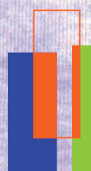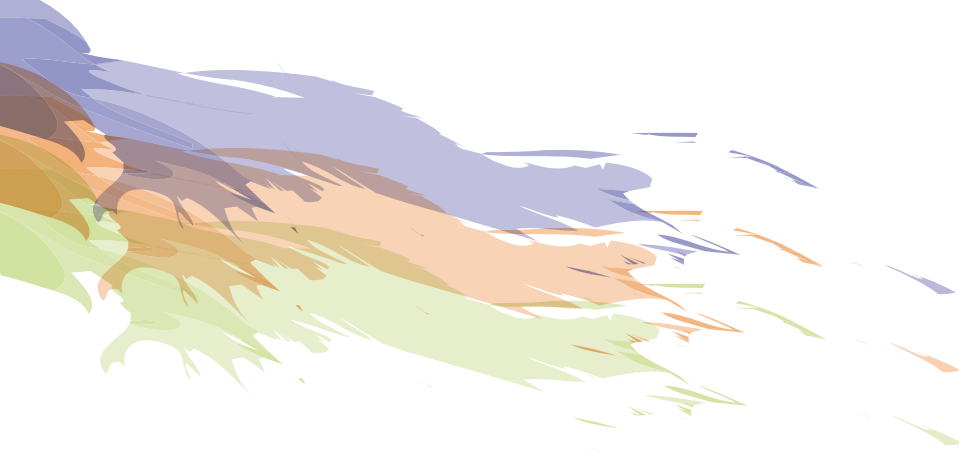
# Using twins to resolve the twin problem of having a bad job and a low wage*

Petri Böckerman
Pekka Ilmakunnas
Jari Vainiomäki

**Corresponding author. Turku School of Economics, Labour Institute for Economic Research, and IZA. Address: Pitkänsillanranta 3A, FI-00530 Helsinki, Finland. E-mail: petri.bockerman@labour.fi

*** Department of Economics, Aalto University School of Business, P.O. Box 21240, FI-00076 Aalto, Finland. E-mail: pekka.ilmakunnas@aalto.fi

**** Department of Economics, School of Management, FI-33014 University of Tampere, Finland. E-mail: jari.vainiomaki@uta.fi

Helsinki 2014

**Tiivistelmä**

Tutkimuksen tavoitteena on tarkastella ns. kompensoivien palkkaerojen esiintymistä työmark-kinoilla. Tutkimuksessa käytetään hyväksi suomalaista kaksosaineistoa. Kaksosten vertaamisen avulla voimme kontroloida työolojen valintaan liittyviä havaitsemattomia tekijöitä. Tulokset antavat tukea sille, että epämukavat työolot korvataan ainakin jossakin määrin työntekijöille korkeamman palkan muodossa.

**Abstract**

We use data on twins matched to register-based information on earnings to examine the long-standing puzzle of non-existent compensating wage differentials. The use of twin data allows us to remove otherwise unobserved productivity differences that were the prominent reason for estimation bias in the earlier studies. Using twin differences we find evidence for positive compensation of adverse working conditions in the labor market.

**1. Introduction**

The idea that adverse working conditions should be compensated in the form of higher pay is an old one. It dates back at least to Adam Smith's Wealth of Nations. Although the notion is intuitively appealing, verifying it has turned out to be difficult. Over the years, empirical

research has investigated this issue in various ways, but with conflicting results (see Rosen, 1986, for a survey).

The main approach has been to estimate hedonic wage functions, where earnings are explained by personal characteristics and various indicators for working conditions. In a competitive labor market with differing tastes and technologies, workers with a high distaste for a negative aspect of work are matched with firms that have a low cost of avoiding it, and obtain a low wage. Correspondingly, workers with a lower distaste for a disamenity are matched with firms with a high cost of producing lower levels of it, and obtain a higher pay.

This forms the theoretical basis for expecting that estimation of a wage equation augmented with job characteristics would yield estimates of compensating wage differentials.

A variety of different aspects of work have been investigated, including risk of death, physical and mental harms, different hazards, shift work, part-time work, night work, promotion prospects, conflicts etc. Sometimes the analysis is conducted at an aggregate level – for example using industry or occupation as the observation unit – but mostly using individuals as observation units. Even in the latter case the measures of working conditions may be taken from a more aggregate level (firm or occupation). In the most appealing circumstances the analyses use information on the actual working conditions that the workers face (e.g. Duncan and Holmlund, 1983).

The clearest result is that occupations with a high risk of death are better compensated, given the characteristics of the individual (e.g. Dorman and Hagstrom, 1998). However, in many studies no compensating differential has been found or even have negative returns appeared. This finding has initiated work that has aimed to explain the unexpected findings. Hwang et al.

(1992) have shown theoretically how unobservable worker productivity differences cause substantial bias in the estimates. Persons who have more human capital (and hence higher productivity) choose jobs with both a higher wage and better working conditions, although with fixed human capital there is a tradeoff between wage and job quality. Hwang et al. (1998) have demonstrated that an additional bias is caused by unobserved productivity differences in the firm side.

If there are unobservable individual characteristics that are correlated with the job characteristics, panel data could be used to eliminate time-invariant unobservables (e.g. Brown, 1980). Another alternative is to account for the endogeneity of the working conditions by using instrumental variables or selection correction (e.g. Garen, 1988). In practice, even these solutions are not always successful. It is difficult to find appropriate instruments that are significantly correlated with working conditions but are otherwise unrelated to earnings. Another issue is measurement error in the working conditions variables. This error may arise for example from the use of aggregate occupational measures, or from the use of self-reported information. Classical measurement errors tend to attenuate the estimates for compensating differentials. Using panel data with differencing has been suggested as a solution (at least partial) to this problem in case of measurement error which is correlated within individual over time (Duncan and Holmlund, 1983).

Other reasons for not finding compensating differentials (even with panel data) include wage bargaining institutions (e.g. Daniel and Sofer, 1998), or search and other frictions in the labor market (e.g. Manning, 2003). Both make the actual labor markets differ from the competitive markets that underlie the theory of compensating differentials.

We examine compensating differentials using data on twins. When wage differences within pairs of identical twins are explained by within-pair differences in working conditions, the unobservable characteristics that are similar to both twin are eliminated. This approach is an alternative to the use of panel data and purges unobservables that are same within the family, but may be time-varying e.g. by age. The effects are identified for twins who work under different conditions. Using data on identical twins it is possible to account for genetic and other family factors like parental investments that may predispose individuals to both poor working conditions and low earnings.

We use a large sample of Finnish twins born before 1958, who were surveyed in 1975, 1981, and 1990. The surveys include some work-related questions that we use to measure working conditions; assessment of how monotonous the work is, physical demands of the work, and opportunities to influence work content. The data can be linked to register data on the same individuals. We explain earnings from tax registers in 1975 and 1990 by working conditions reported in the surveys. This allows us to control for both time-invariant person-specific unobservables and family-specific, possibly time-varying effects. We find that the results support the idea that pooled OLS estimation and even fixed effect estimation may lead to an underestimation of the wage effects of adverse working conditions. In twin differencing, the estimates increase, and we find compensating differentials for having no influence on work content and for physically demanding work.

The paper unfolds as follows. Section 2 analyzes potential biases arising from measurement errors and autocorrelation of job characteristics and how they can be solved with appropriate differencing using panel data on twins. Section 3 presents our data and Section 4 the estimation results. Section 5 concludes the paper.

## 2. Econometric issues in estimating compensating wage differentials

We investigate what factors cause biases in the estimates and how they can be avoided in different types of data sets. To do this, we combine insights from Duncan and Holmlund (1983) on measurement errors of working conditions and from Isacsson (2007) on the use of twin data and panel data to remove unobservables. For ease of exposition we assume that there is only one job characteristic variable and that there are no control variables (or they have been netted out with a prior regression).

*2.1 Accounting for unobserved effects using twin data and panel data*

Following Isacsson (2007) we assume

$$y_{ijt} = \beta_0 + \beta_1 J_{ijt} + u_{ijt},$$  (1)

where $y_{ijt}$ is log earnings, $J_{ijt}$ is job characteristic, $i=1, 2$ denotes individuals within twin pairs, $j=1, \ldots , N$ denotes twin pairs and $t=1, \ldots , T$ refers to years. The characteristic $J$ is assumed to be a positive aspect, so its return $\beta_1$ is expected to be negative.

Furthermore, we assume the following structure for the error term:

$$u_{ijt} = \gamma_t A_{ij} + e_{ijt},$$  (2)

where $A_{ij}$ is an individual-specific, time-invariant unobserved characteristic ("ability"). $\gamma_t$ is the return to unobserved ability, which we allow to vary over time (e.g. because of increasing skill bias). Isacsson (2007) allows the ability term to change over time, but assumes the change to be the same for both twins. Our formulation is slightly more general and can be justified as

5

allowing the return to unobserved ability to change over time similarly to the return to education.

It is well known that the cross-section OLS estimate is biased if $A_{ij}$ correlates with $J_{ijt}$. For example, negative compensating differential for a job amenity is biased towards zero if the more productive (able) persons choose more amenable jobs. The estimate can even turn positive. This is the unobserved productivity (human capital) bias in estimating compensating differentials stressed e.g. in Brown (1980), Duncan and Holmlund (1983) and Hwang et al. (1992).

Panel data fixed effects (time difference) estimator is obtained by using OLS for the equation

$$y_{ijt} - y_{ijs} = \beta_1(J_{ijt} - J_{ijs}) + (\gamma_t - \gamma_s)A_{ij} + (e_{ijt} - e_{ijs}), \qquad t > s \qquad (3)$$

If $s = t$-1, this is the first-difference (FD) model, otherwise the long difference model. This yields consistent estimates for $\beta_1$ only if ability-related wage growth components are not correlated with the changes in job characteristics. However, if more able persons' higher wage growth $((\gamma_t - \gamma_s)A_{ij} > 0)$ leads them to choose better job conditions $((J_{ijt} - J_{ijs}) > 0)$, the compensating differential estimate $\beta_1^{FD}$ is biased towards zero. The standard panel data assumption is that the return to ability is constant over time, $(\gamma_t - \gamma_s) = 0$, which eliminates the bias.

A fixed effects estimator based on twin data (twin differences) is given by

$$y_{1jt} - y_{2jt} = \beta_1(J_{1jt} - J_{2jt}) + \gamma_t(A_{1j} - A_{2j}) + (e_{1jt} - e_{2jt}) \qquad (4)$$

It is a standard assumption that the unobserved ability is equal for identical (MZ) twins, $A_{1j} = A_{2j}$, because they share exactly the same genes and they are usually brought up in

6

the same family environment and they usually also share the same peer group in their youth. This assumption yields consistent estimates for compensating differentials $\beta_1$ for MZ twins, because the ability term is eliminated due to twin-differencing. However, for DZ twins and other siblings the compensating differentials estimates are biased towards zero if twin-differences in ability correlate positively with twin-differences in job amenities.

A Difference-in-Differences (time difference – twin difference) estimator is obtained from

$$(y_{1jt} - y_{1js}) - (y_{2jt} - y_{2js}) = \beta_1[(J_{1jt} - J_{1js}) - (J_{2jt} - J_{2js})] + (\gamma_t - \gamma_s)(A_{1j} - A_{2j})$$
$$+ [(e_{1jt} - e_{1js}) - (e_{2jt} - e_{2js})] \tag{5}$$

The DiD estimator removes the ability effects only if the ability related wage growth components are fixed within twin pairs (i.e. pair specific). This prevails, for example, if $A_{1j} = A_{2j}$ for MZ twins. Otherwise, e.g. for DZ twins, a downward bias remains even in DiD estimation, if the twin with higher wage growth is experiencing larger improvement in job amenity over time.

Isacsson (2007) assumes the following structure for the ability term:

$$A_{ijt} = f_j + g_{ij} + h_j t . \tag{6}$$

In this case the wage growth component $h_j(t-s)$ is the same for both twins in the twin pair, so it is eliminated by twin-differencing. The twin-difference in individual-specific effects $\Delta g_j = g_{1j} - g_{2j}$ is time-invariant, so it is eliminated by time-differencing in the DiD estimator. In this case DiD is consistent (unbiased).

Our formulation is slightly more general and clarifies the central assumption for the consistency of DiD. The wage growth component $(\gamma_t - \gamma_s)A_{ij}$ must be pair-specific, not individual-specific, for the DiD to be consistent. Our formulation allows for wage growth to be

different for the two twins in general, but making the equal ability assumption for the MZ twins, wage growth turns out to be the same for the twins. Thus, the wage growth term is eliminated in the DiD estimation, and thereby the ability bias is avoided. But it remains, for example, for DZ twins. The DiD for MZ twins is therefore unbiased with the standard assumption of equal ability within MZ twin pairs.

*2.2 Measurement error bias with correlated errors and job characteristics*

Assume that job characteristics are measured with error (the individual and pair indexes are not shown for simplicity)

$$J_t^* = J_t + \varepsilon_t,$$
(7)

where $\varepsilon_t$ is measurement error, $J_t^*$ is the observed job characteristic and $J_t$ is the true job characteristic.

Assuming that random measurement error is not correlated with true job characteristic $J$ or with ability $A$, the cross-section OLS is (further) biased towards zero due to the well-known attenuation bias. The bias towards zero is increasing in the noise-to-signal ratio. It is often assumed that differencing increases the bias to the extent that the noise-to-signal ratio is higher compared to cross-section OLS.

Next we allow for the measurement error to be correlated over time using the same $s$ period lag in the correlation structure as in the time differencing above

$$\varepsilon_t = \rho_\varepsilon \varepsilon_s + v_t,$$
(8)

where $v_t \approx i.i.d(0, \sigma_v^2)$. It can then be shown that for $0 \leq \rho_\varepsilon \leq 1$,

$$\text{var}(\Delta \varepsilon_t) = \frac{2\sigma_v^2}{1 + \rho_\varepsilon},$$
(9)

where we have used $\text{var}(\varepsilon_t) = \text{var}(\varepsilon_s) = \text{var}(\varepsilon) = \dfrac{\sigma_v^2}{1 - \rho_\varepsilon^2}$ .

Furthermore, we assume that the true job characteristics may also be correlated. With unchanged preferences for working conditions, and unchanged market price for job amenities, an individual's optimum wage-amenity combination should be stable over time, captured here by high autocorrelation (again, we use lag $s$ in the correlation structure)

$$J_t = \rho_J J_s + w_t , \tag{10}$$

where $w_t \approx i.i.d(0, \sigma_w^2)$. Then, analogously to measurement error, we get

$$\text{var}(\Delta J_t) = \frac{2\sigma_w^2}{1 + \rho_J} \tag{11}$$

Using this and the previous result for measurement error to rewrite the noise-to-signal ratio, we obtain for the probability limit of the FE (time difference) estimate of compensating differential

$$\text{plim } \beta_1^{FE} = \frac{\beta_1}{1 + \text{var}(\Delta\varepsilon_t)\big/\text{var}(\Delta J_t)} = \frac{\beta_1}{1 + \dfrac{(1 + \rho_J)\sigma_v^2}{(1 + \rho_\varepsilon)\sigma_w^2}} \tag{12}$$

This formulation implies that the bias in the estimate towards zero is larger, *ceteris paribus,* if

1) $\sigma_v^2$ increases (there is more measurement error)

2) $\sigma_w^2$ decreases (there is less idiosyncratic variation in true $J$)

3) $\rho_J$ increases (true $J$ is more highly correlated, i.e. more persistent)

4) $\rho_\varepsilon$ decreases (measurement error is less correlated, i.e. less persistent)

As noted by Duncan and Holmlund (1983), if individuals have a tendency to persistently over- or underestimate their working conditions in surveys, the job characteristics measures are strongly correlated over time. Time-differencing this time-invariant measurement error component out reduces measurement error bias. However, it does not disappear even in the case of perfect autocorrelation, as claimed by Duncan and Holmlund (1983).[1]

Note that high correlations in signal and measurement error have opposing effects on the bias. Persistence in measurement error improves, but persistence in signal worsens the consistency of the fixed effects estimate for compensating differential. The latter effect arises because the variation in the differenced explanatory variable is reduced by persistence in working conditions over time. However, using very long differences may lower the autocorrelation in $J$.

We use the job characteristics as indicator variables. The following results apply to the measurement error (classification error) in this case (Freeman, 1984)

$$\text{plim } \beta_1^{OLS} = \beta_1(1 - p_{10} - p_{01})\frac{\text{var}(J_t)}{\text{var}(J_t^*)} \tag{13}$$

$$\text{plim } \beta_1^{FE} = \beta_1(1 - p_{10} - p_{01})\frac{\text{var}(\Delta J_t)}{\text{var}(\Delta J_t^*)} \tag{14}$$

where $p_{10}$ and $p_{01}$ are the misclassification probabilities, i.e. $p_{10}$ ($p_{01}$) is the probability that 1 (0) is erroneously classified as 0 (1). Since $J$ and $J^*$ are binary, $\text{var}(J_t) = P(1-P)$ and $\text{var}(J_t^*) = P^*(1-P^*)$, where $P$ and $P^*$ are the proportions of true and observed 1's. Further, if these proportions are equal, the limit of the OLS estimate reduces to $\text{plim } \beta_1^{OLS} = \beta_1(1 - p_{10} - p_{01})$.

---

[1] Duncan and Holmlund (1983, p. 371) argue that "… in the special case with 'perfect' autocorrelation ($\rho$=1), the OLS estimate is consistent." However, in this case $\text{var}(\Delta\varepsilon_t) = \sigma_v^2$, so the signal-to-noise ratio does not disappear from (12).

For the size of bias these formulas have similar implications as in the continuous case:

1) more misclassification increases the bias towards zero

2) higher noise-to-signal ratio (lower variance ratio of true $J$ to observed $J^*$) increases the bias

3) FE bias is larger than OLS bias if noise-to-signal ratio is larger in differences

However, these results assume uncorrelated measurement error.[2]

*2.3 Comparison of twin-differences vs. time-differences FE estimators*

The results for the panel FE (time-difference) estimator above extend analogously to the twin-difference estimator, where the difference is taken between the two twins in the same pair. The correlations are now between the true job characteristics or the measurement errors of the two twins within the same pair.

Therefore, we make the following conclusions regarding the relative biases in twin-difference and panel-difference FE estimators:

1) If the within twin-pair correlation in true $J$ ($\rho_J$) is smaller than the panel correlation for individuals, then the bias from Twin FE is *smaller* than that from time difference FE, other things being equal. This should be the case, because it is likely that

---

[2] The issue of measurement errors in binary variables, combined with correlation over time has not been examined in the literature. This would involve specifying the probabilities of staying at the same value (0 or 1) over time, $q_{00}$ and $q_{11}$, and the corresponding transition probabilities, $q_{01} = 1 - q_{00}$ and $q_{10} = 1 - q_{11}$.

preferences for job characteristics vary more between twins than for an individual over time, at least when relatively short differences are used.[3]

2) If the twin correlation in measurement error ($\rho_\varepsilon$) is also smaller than the panel correlation, then the bias from Twin FE tends to be *larger* than that from time-difference FE, other things being equal. This should be the case, because it is less likely that the two twins over- or underestimate their working conditions in the same way than it is for the same individual to over- or underestimate in different time periods.

Therefore, we cannot make unambiguous predictions about the size of bias in Twin FE and time difference FE due to measurement error effects.

Neither can we make definitive conclusions about the size of unobserved ability bias in these two estimators in Section 2.1. There is no unobserved ability bias in panel FE, if there are no wage growth effects from ability. Twin FE is unbiased for MZ, if their ability is equal. Both of these assumptions can in principle be questioned, so unbiasedness cannot be claimed indisputably.[4] Furthermore, not even DiD is unambiguously consistent without further restrictions on ability effects. However, for MZ twins the DiD estimator is unbiased given the standard assumption of equal ability within MZ twin pairs. For DZ twins and other siblings the bias from unobserved heterogeneity remains. The size of biases is therefore an empirical issue.

---

[3] Cesarini et al. (2009) show that genetic factors explain only about 20% of the variation in behavior in their dictator experiments. Hence social preferences are only in small part explained genetically.

[4] For example, Sandewall et al. (2014) challenge the identification power of equal ability assumption for MZ twins. They argue that controlling additionally for IQ test scores reduces within-twin pair estimates of returns to schooling by about 15%. However, their within-twin pair estimates are over 50% smaller than cross-sectional OLS estimates, so it seems that the "equal ability" assumption for MZ twins accounts for the bulk of the ability variation across individuals.

## 3. Data

The twin data used in this study are based on the Older Finnish Twin Cohort Study of the Department of Public Health at the University of Helsinki. The initial twin data gathered in 1974 contain almost all same-sex DZ (dizygotic) and MZ (monozygotic) twins in the Finnish population born before 1958 (see Kaprio et al., 1979; Hyytinen et al., 2013). The surveys were sent to the twins in 1975, 1981, and 1990. Identification of the DZs and MZs is based on the survey questions, but later this identification was confirmed for a small subsample using blood markers. The results based on the tests and surveys matched almost perfectly.

We link the twin sample to the FLEED (Finnish Longitudinal Employer-Employee Data) maintained by Statistics Finland using personal ID codes attached to every person residing in Finland.[5] This matching is exact, and there are no misreported ID codes. We therefore avoid problems associated with errors in record linkages (e.g., Ridder and Moffitt, 2007). FLEED is constructed from a number of different registers of individuals and firms maintained by Statistics Finland. FLEED includes information on individuals' labor market status and salaries and other income, taken directly from tax and other administrative registers that are collected and/or maintained by Statistics Finland. Thus, our earnings data do not suffer from underreporting or recall error, nor is it top coded. We concentrate on year 1990 in the FLEED data, as a twin survey is available from the same year. For the survey year 1975 earnings data are available from the Longitudinal Population Census of Statistics Finland. (There is no income information available for the survey year 1981.) The income measure is the logarithm of annual earnings from the tax register. Because the data on earnings contain some outliers, we have truncated the observations outside the 1st and 99th percentiles.

---

[5] Hyytinen et al. (2013, p. 63) document that the life-time labor market outcomes of the linked twin data are representative of the Finnish population.

Working conditions are self-reported in the survey waves conducted in 1975 and 1990. Monotonous work (1975 and 1990) is measured using the question "your work can be characterized as" with the alternatives: "very monotonous", "rather monotonous", "rather non-repetitive" and "very non-repetitive". To detect potential non-linear patterns in the effect of working conditions on earnings, we form three binary indicators for the degree of monotony of work, using "very non-repetitive" as the reference category. Opportunities to influence work methods (1990) are measured with the alternatives: "no influence", "some influence" and "substantial freedom in choosing work methods". We use two indicators for the level of influence, with "substantial freedom" as the reference. Physically demanding work (1975 and 1990) is measured with alternatives: "my work is physically very demanding", "my work involves lifting and carrying objects in addition to standing and walking", "my work involves standing and walking but no other physical activity" and "my work requires hardly any physical activity". Again, we use binary indicators for the types of physical demands with "hardly any" as the reference category.

To prevent education and early pension choices from affecting our earnings measures, we restrict the analyses to persons in prime working age. All empirical specifications are estimated for individuals born after 1944 but before 1958. Our analysis focuses on men, because they are more strongly attached to the labor market. Male labor supply decisions are also much less complex compared with women because men are less affected by family and fertility choices. All models include (unreported) control for observed human capital using the years of education, based on standard degree times used by Statistics Finland.

## 4. Results

Before the presentation of the estimation results, it is useful to report some unconditional correlations of our measures for working conditions in the twin data. Table 1 documents that measured self-reported working conditions are significantly less correlated between twins (Panel A) than within individuals over time (Panel B), even when the time difference is as long as from 1975 to 1990. The pattern gives some support for our conjecture above that the same would hold for the correlations of true working conditions. This would imply that the twin FE estimates are less biased than the panel FD estimates, other things being equal.

===== Table 1 here =====

The central aim of this paper is to compare various estimations to evaluate different sources of bias. We first present the standard OLS estimates to obtain baseline results that are comparable to the earlier studies. Then we proceed to estimate models using twin differences to control for otherwise unobserved productivity/human capital effects. The estimations for DZs control for sibling effects i.e. family background and some of the genetics (DZ twins share about 50% of genes). The estimations using data on MZs control for family background, peer effects and genetics more completely (MZ twins are genetically identical).[6]

We also estimate panel data fixed effects models using two of our three measures for working conditions that are available in the survey waves for both 1975 and 1990. First, we estimate individual-level fixed effects models using the panel dimension of the data (i.e. long

---

[6] We focus on models in which each working condition measure is entered in a separate model, because there is substantial correlation between the measures.

differences of 15 years). These models control for all time-invariant unobservable characteristics at the individual level.

Second, we estimate fixed effects specifications for twin differences, i.e. Difference-in-Differences models. In addition to the standard twin differences for DZs and MZs, these models also control for all otherwise unobservable time-invariant characteristics at the twin pair level that may matter for workers' sorting into workplaces. For example, it is possible that the place of residence,[7] spouse's occupation or family situation in general have effects on persons' occupational choices and therefore on working conditions. To the extent that these effects vary between twins, but are invariant over time, their effects are not controlled by twin differencing, but are controlled by time differencing of twin differences.

Since the working condition variables describe negative aspects of work, the compensating differentials would imply positive coefficients (unlike in Section 2 above, where the job characteristic was a positive aspect). The standard OLS estimates show that the indicator for rather monotonous work and the indicators for somewhat and very physical work are significantly *negatively* related to earnings (Table 3, Panels A and C, Column 1). This result is contrary to the expected positive compensating wage differential for adverse working conditions. The pattern is likely to be driven by unobserved heterogeneity related to the sorting of workers into different working conditions. The OLS estimations of compensating wage differentials in other studies have also often found wrong-signed premiums for adverse working conditions.

---

[7] Finland is sparsely populated. In many regions there are only a few large firms with a fixed set of tasks and working conditions.

Next we turn to the results that use the panel dimension of our data. Table 2 reports the estimates for individual-level panel fixed effects models. These results show that controlling for time-invariant unobservable characteristics at the individual level is not able to remove the negative compensating wage differentials for our measure of physical work (Table 2, Panel B). In fact, we observe a statistically significant negative compensation. Our result is similar to Duncan and Holmlund (1983, p. 373) who also do not find a positive compensating wage differential for physically demanding work in panel FE models. However, Duncan and Holmlund (1983) find support for compensating wage differentials for dangerous work and stressful work in panel estimations.[8] Our panel results are similar to the OLS estimates. One reason may be the difference in the working conditions being studied.

===== Table 2 here =====

The use of twin differences changes the picture (Table 3, Columns 2-3). First, the statistically significant negative effects observed in the OLS models turn either insignificant or positive in the twin difference models. Second, we obtain some evidence for positive compensation for adverse working conditions. In particular, having no influence at work and being positioned in somewhat physical work both have a statistically significant *positive* effect on earnings for MZs at the 10% level (Table 3, Panels A and B, Column 3). For DZs we obtain a positive effect at the 10% level also for some (rather than substantial) influence at work. This effect is positive and larger for MZs, as expected, but it is not statistically significant, possibly due to the smaller sample size. For MZs the point estimate of no influence at work is notably larger than the estimate for some influence. These results suggest that the unobserved ability bias considerably affects the OLS point estimates. Even using twin differences we do not obtain significant positive compensation for monotonous work. We also find that the positive

---

[8] Böckerman et al. (2011) obtain a positive compensating differential for job uncertainty using Finnish panel data.

compensation for physical work does not apply to those who are in the worst working conditions.

===== Table 3 here =====

The unique feature of our data is that we are also able to perform Difference-in-Differences estimations (Table 4). These specifications control for unobservable time-invariant characteristics at the twin pair level and for the common wage growth for the twins. We find evidence for positive compensating wage differentials for both monotonous and physical work using the data on MZs (Table 4, Panels A and B, Column 3). The effects are statistically significant only at the 10% level probably because the sample size is even smaller for MZs when using information on working conditions and earnings both in 1975 and 1990.

For physical work the results in Table 4 are similar to the ones that use twin differences for MZs in Table 3. For rather monotonous work the point estimates are also similar but the effect obtains statistical significance only in Table 4. It is notable that the point estimates for significant effects are similar in Tables 3-4 and in accordance with the discussion in Section 2. Assuming that equal ability assumption holds for MZs the results support the notion that both twin difference estimates and D-in-D estimates are unbiased for MZs.

===== Table 4 here =====

## 5. Conclusions

We use twin data matched to register-based information on earnings to examine the long-standing puzzle of non-existent compensating wage differentials. The use of twin data allows us to remove otherwise unobserved productivity differences that are the prominent reason for estimation bias in the earlier studies.

The standard OLS estimates show negative or insignificantly positive compensating wage differentials as are often also found in the earlier studies. In contrast to Duncan and Holmlund (1983), who found some support for compensating wage differentials in panel estimations, our panel results are similar to the OLS estimates.

The use of twin differences changes the picture. Using twin differences for MZs we find evidence for positive wage compensation for somewhat physically demanding work and for having no influence on work content. These results suggest that unobserved ability bias affects the OLS point estimates considerably.

The unique feature of our data is that we are able to estimate specifications that also control for unobservable time-invariant characteristics at the twin pair level (e.g. place of residence, spouse's occupation, or family situation) and the common wage growth for the twins. We find evidence for positive compensating wage differentials for both monotonous and physical work using the data on MZs. Assuming that equal ability assumption holds for MZs, the results imply that both twin difference estimates and Difference-in-Differences (combined twin difference – time difference) estimates are unbiased for MZs.

The pattern that positive compensating wage differential does not prevail in the worst working conditions (i.e. "very monotonous work" and "very physical work") may be due to the lack of control for workers' bargaining power, if bargaining power effects are particularly important in the most onerous low-wage jobs. Daniel and Sofer (1998) argue that positive relationship between job amenities and earnings is possible along the contract curve from joint bargaining of job amenities and earnings.

**References**

Brown, G. (1980). "Equalizing differences in the labor market", Quarterly Journal of Economics, 94, 113-134.

Böckerman, P., Ilmakunnas, P. and Johansson E. (2011). "Job security and employee well-being: Evidence from matched survey and register data", Labour Economics, 18, 547-554.

Cesarini, D., Dawes, C.T, Johannesson, M., Lichtenstein, P. and Wallace, B. (2009). "Genetic variation in preferences for giving and risk-taking." Quarterly Journal of Economics 124, 809-842.

Daniel, C. and Sofer, C. (1998). "Bargaining, compensating wage differentials, and dualism of the labor market: theory and evidence for France", Journal of Labor Economics, 16, 546-75.

Dorman, P. and Hagstrom, P. (1998). "Wage compensation for dangerous work revisited", Industrial and Labor Relations Review, 52, 116-135.

Duncan, C.J. and Holmlund, B. (1983). "Was Adam Smith right after all? Another test of the theory of compensating wage differentials", Journal of Labor Economics, 1, 366-379.

Freeman, R.B. (1984). "Longitudinal analyses of the effects of trade unions", Journal of Labor Economics, 2, 1-26.

Garen, J. (1988), "Compensating wage differentials and the endogeneity of job riskiness", Review of Economics and Statistics, 70, 9-16.

Hwang, H., Mortensen, D.T. and Reed, W.R. (1998). "Hedonic wages and labor market search", Journal of Labor Economics, 16, 815-847.

Hwang, H., Reed, W.R. and Hubbard, C. (1992). "Compensating differentials and unobserved productivity", Journal of Political Economy, 100, 835-58.

Hyytinen, A., Ilmakunnas, P. and Toivanen, O. (2013). "The return- to-entrepreneurship puzzle", Labour Economics, 20, 57-67.

Isacsson, G. (2007). "Twin data vs. longitudinal data to control for unobserved variables in earnings functions – which are the differences?", Oxford Bulletin of Economics and Statistics, 69, 339-362.

Kaprio, J., Artimo, M., Sarna, S., and Rantasalo, I. (1979). The Finnish twin registry: Baseline characteristics. Section I: Materials methods, representativeness and results for variables special to twin studies. Department of Public Health, Publications, M(47).

Manning, A. (2003). Monopsony in Motion. Imperfect Competition in Labor Markets, Princeton University Press, Princeton, NJ.

Ridder, G. and Moffitt, R. (2007). "The econometrics of data combination", in J.J. Heckman and E. E. Leamer (Eds.), Handbook of Econometrics, Vol. 6B, Elsevier, Amsterdam, pp. 5469-5547.

Rosen, S. (1986). "The theory of equalizing differences", in O. Ashenfelter and R. Layard (Eds.), Handbook of Labor Economics, Vol. 1, Elsevier, Amsterdam, pp. 642-92.

Sandewall, Ö., Cesarini, D. and Johannesson, M. (2014). "The co-twin methodology and returns to schooling – testing a critical assumption", Labour Economics, 26, 1-10.

Table 1. Correlation coefficients between the measures of working conditions.

| Panel A: within twin-pairs (1990) | DZ | MZ |
|---|---|---|
| | | |
| Monotonous work | 0.0915 (0.1056) | 0.2216*** (0.0090) |
| Opportunities to influence | 0.1144** (0.0428) | 0.1334 (0.1189) |
| Physically demanding work | 0.0696 (0.2188) | 0.1911** (0.0248) |
| | | |
| Panel B: within individuals (1975-1990) | DZ | MZ |
| | | |
| Monotonous work | 0.2258*** (0.0000) | 0.2349*** (0.0000) |
| Physically demanding work | 0.2571*** (0.0000) | 0.2251*** (0.0000) |

Notes: Spearman's rank correlation coefficients reported. Significant at *10%, ** 5%, and *** 1% level.

Table 2. The effect of working conditions on earnings: individual-level panel fixed effects estimations.

| Panel A: | FE All | FE DZ | FE MZ |
|---|---|---|---|
| | | | |
| Very monotonous work | -0.1322 | -0.1726 | -0.1067 |
| | (0.2600) | (0.3344) | (0.3297) |
| Rather monotonous work | -0.0350 | -0.2266 | 0.4111 |
| | (0.1661) | (0.1959) | (0.3101) |
| Rather non-repetitive work | -0.0856 | -0.1347 | 0.0227 |
| | (0.1027) | (0.1249) | (0.1783) |
| | | | |
| N | 1232 | 855 | 377 |
| | | | |
| Panel B: | FE All | FE DZ | FE MZ |
| | | | |
| Very physical | -0.6731*** | -0.6581*** | -0.7041** |
| | (0.1770) | (0.2163) | (0.3051) |
| Rather physical | -0.3849*** | -0.3698** | -0.4158* |
| | (0.1409) | (0.1758) | (0.2274) |
| Somewhat physical | -0.1608 | -0.1629 | -0.1614 |
| | (0.1685) | (0.2151) | (0.2581) |
| N | 1303 | 906 | 397 |

Notes: Working conditions are measured in 1975 and 1990. In Panel A the reference category is very non-repetitive work. In Panel B the reference category is non-physical work. All models include (unreported) control for the years of education. Significant at *10%, ** 5%, and *** 1% level.

Table 3. The effect of working conditions on earnings.

| Panel A: | OLS | Twin differences: DZ | Twin differences: MZ |
|---|---|---|---|
| | | | |
| Very monotonous work | -0.2306 (0.3624) | -0.0999 (0.2078) | -0.0254 (0.2913) |
| Rather monotonous work | -0.3467** (0.1724) | -0.1032 (0.2484) | 0.7219 (0.4881) |
| Rather non-repetitive work | -0.0861 (0.0927) | 0.0252 (0.1227) | -0.0279 (0.2265) |
| | | | |
| N | 1988 | 994 | 324 |
| | | | |
| Panel B: | OLS | Twin differences: DZ | Twin differences: MZ |
| | | | |
| No influence | -0.4030 (0.2792) | 0.2358 (0.3047) | 0.5718* (0.2969) |
| Some influence | 0.0824 (0.0765) | 0.1966* (0.1024) | 0.2749 (0.2246) |
| | | | |
| N | 1872 | 936 | 311 |
| | | | |
| Panel C: | OLS | Twin differences: DZ | Twin differences: MZ |
| | | | |
| Very physical | -0.4593** (0.1833) | -0.2424 (0.2555) | -0.1191 (0.4529) |
| Rather physical | -0.0775 (0.1121) | -0.1105 (0.1760) | 0.1016 (0.2816) |
| Somewhat physical | -0.2832* (0.1613) | 0.0516 (0.2047) | 0.6869* (0.3564) |
| | | | |
| N | 2030 | 1015 | 334 |

Notes: All working conditions are measured in 1990. In Panel A the reference category is very non-repetitive work. In Panel B the reference category is having substantial influence. In Panel C the reference category is non-physical work. All models include (unreported) control for the years of education. Significant at *10%, ** 5%, and *** 1% level.

Table 4. The effect of working conditions on earnings: fixed effects specification for twin differences (difference in differences).

| Panel A: | FE All | FE DZ | FE MZ |
|---|---|---|---|
| | | | |
| Very monotonous work | 0.3404 | 0.3686 | 0.4236 |
| | (0.2716) | (0.3099) | (0.4153) |
| Rather monotonous work | 0.0855 | -0.1071 | 0.6141* |
| | (0.1975) | (0.2321) | (0.3568) |
| Rather non-repetitive work | 0.0277 | 0.0643 | -0.0195 |
| | (0.1042) | (0.1191) | (0.2084) |
| | | | |
| N | 806 | 554 | 252 |
| | | | |
| Panel B: | FE All | FE DZ | FE MZ |
| | | | |
| Very physical | -0.1565 | -0.2078 | -0.0694 |
| | (0.2306) | (0.2965) | (0.3487) |
| Rather physical | 0.0864 | -0.0721 | 0.4199 |
| | (0.1831) | (0.2308) | (0.2824) |
| Somewhat physical | 0.1354 | -0.0926 | 0.5946* |
| | (0.1648) | (0.1884) | (0.3121) |
| | | | |
| N | 874 | 600 | 274 |

Notes: Working conditions are measured in 1975 and 1990. In Panel A the reference category is very non-repetitive work. In Panel B the reference category is non-physical work. All models include (unreported) control for the years of education. Significant at *10%, ** 5%, and *** 1% level.