Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Bastian Ravesteijn

The Wear and Tear of Health

Does Manual Work Matter for Health at an Older Age?

RM Thesis 2011-055

The wear and tear of health: does manual work matter for health at an older age?

Bastian Ravesteijn

August 19, 2011

Abstract

Using German panel data spanning 26 years, we find that manual work leads to faster deterioration of health than nonmanual work for the group of workers between 50 and 65 years old. We control for time-varying and time-invariant omitted variable bias and reporting heterogeneity and we estimate an upper and a lower bound of the effect of manual work on health, as a consistent estimator is unavailable for our data. We do not find evidence of an effect of manual work on health for the group of workers between 16 and 50 years old.

1 Introduction

Lower socioeconomic status is associated with worse health. In the U.S., for example, manual workers are fifty percent more likely to die within a given year than workers in managerial, professional and executive occupations (Cutler et al. 2008). The effect of socioeconomic status on health may operate through more knowledge about healthy behavior and the health care system, healthier peers, healthier consumption, higher income and the possibility of avoiding jobs that are harmful to health. The effect could also run the other way, as an adverse health shock may limit the possibilities of obtaining a good education, a good job and high income. Our results indicate that manual work is harmful to health for older workers.

Economists and epidemiologists have investigated whether type of occupation by itself can have an effect on health. We distinguish three possible causal pathways. The first causal pathway is through physical workplace conditions as suggested by Case and Deaton (2005). Manual workers could face higher accident risk or have worse ergonomic workplace conditions, such as repetitive work movements or inconvenient work postures.

The second possible causal pathway is through psychosocial workplace conditions. It is ex ante unclear how psychosocial factors affect the gradient between socioeconomic status and health. For example, both managers who face a high workload and assembly workers who have low control possibilities at work may be more likely to suffer from mental disorders such as depression (Marmot 1997). Karasek (1979) emphasizes the role of job decision latitude, the degree to which a worker has control over the events at the workplace. Low job control possibilities may have a negative effect on health.

The third possible pathway is through lifestyle choices. Peer effects could lead to unhealthy behavior such as smoking and drinking or a lack of physical exercise (Marmot 1997). In this paper, we look at the difference in health status associated with doing manual versus nonmanual work. This means that we can not yet isolate the role of each of the three transmission mechanisms individually. Linking occupational titles to measures of physical and psychosocial workplace conditions could allow for a more detailed analysis of these causal pathways.

Several theoretical and empirical results suggest that manual work is most harmful to health at older ages. According to the Grossman (1972) health capital model, workers can use a health repair technology to offset a deterioration of their health capital. Unhealthy working conditions could lead to a faster deterioration of health at older ages because health at older ages is more vulnerable to wear and tear at the workplace. Another explanation is that the health depreciation rate because of hard work is equal at all ages, but the cost of health repair increases with age, such that full repair of health is not feasible at an older age. Case and Deaton (2005) describe regularities in the U.S. National Health Interview Survey data that suggest that manual work causes health to decline more rapidly with age than nonmanual work. Fletcher et al. (2011) find a negative effect of physical workplace demands on health status for old men but not for young men. We present evidence that supports the hypothesis that manual work leads to worse health only at an older age.

We want to identify whether there is a causal relationship between type of occupation and health. This is not a straightforward exercise. For example, the fact that mortality rates for manual workers in the U.S. are higher than for people who do nonmanual work does not automatically mean that manual work is harmful to health. Four identification problems may prevent us from properly identifying the causal relationship between type of occupation and health. First, manual workers may have lower education, unhealthier lifestyles or a worse genetic predisposition. These are examples of selection effects: people with bad health prospects are selected into certain types of occupation. If unhealthy people do manual work, then naïve estimation strategies will overestimate the negative effect of manual work on health. But the selection effect could also run in the opposite direction in the presence of a healthy manual worker effect. If strong and healthy workers, with an innate ability for good health, are selected into manual occupations, a simple comparison of the health outcomes between manual and nonmanual workers would show a positive relationship between manual work and health, even if manual work in fact is harmful to health.

Second, apart from factors that are constant throughout the lifetime, time-varying factors may influence both occupational choices and health outcomes. For example, the death of a loved one or a divorce may affect performance at work and we can not rule out the possibility that this leads to job switching. The event may also have an effect on mental wellbeing and health in general. This means that a naïve analysis would erroneously attribute the deterioration of health to the new job, while in fact it was caused by a third factor such as the death of a loved one or a divorce. This could lead to either an underestimation or an overestimation of the effect of manual work on health, depending on the exact impact of the event on type of occupation and on health.

Third, a change in health may in fact cause a person to switch jobs, such that the causal relationship runs from health to work instead of from work to health. For example, a physical disability that was caused by a car crash could lead to a switch from manual work to a desk job. Or a person who suffers from a burnout could decide to switch from a nonmanual job with managerial responsibilities to a manual job with fewer responsibilities. So we can think of many ways in which a naïve analysis would yield biased results because of reverse causality, and we do not know the direction of the bias.

Fourth, the use of a subjective health measure may lead to reporting bias. Researchers may use a subjective health measure as a proxy for true health status, as a variable indicating subjective health is often available in survey data. Several studies show that self-assessed health is a reliable predictor of mortality and morbidity (Mossey and Shapiro 1982, Idler and Benyamini 1997, Mackenbach et al. 2002). However, different subgroups may assess the same objective health status differently (Lindeboom and Van Doorslaer 2004). If, for example, men report worse health than women who have the same objective health status, and if men are more likely than women to do manual work, a naïve analysis would yield results that suffer from a downward bias. Reporting heterogeneity could also lead to an upward bias if people who report health more optimistically than others are more likely to do manual work.

We aim to contribute to the literature on the relationship between occupation and health by presenting an estimation strategy that explicitly controls for reporting bias and for factors that caused selection into type of occupation and that are related to health outcomes. We look at the effect of manual work versus nonmanual work on satisfaction with health using German Socioeconomic Panel data. Our model includes an individual-specific effect to control for the bias that stems from individual-specific characteristics that do not vary over time and for reporting bias. The model also includes satisfaction with health in the previous year as a control variable. We identify the effect of at most one year of manual work by including a variable that takes has value 1 if the person was doing manual work and value 0 if the person was doing nonmanual work. We control for any events that could have simultaneously affected type of occupation and health if we assume that they are captured by last years health.

Our econometric model does not rely on sources of exogenous variation in occupational choice, such as instrumental variables or a (pseudo) experimental setting. An analysis based on exogenous variation in occupation, such as a plant closing, may produce results that only hold very locally for a limited class of occupations and within a very specific institutional setting. We look at the health effects of doing manual versus nonmanual work. We view the degree to which work is of a manual nature as one characteristic of work. We look at all job categories except for the armed forces. The advantage of this broad brush approach is that we can include all nonmilitary workers in our analysis instead of focusing on the health effects of just one class of occupations. This broad brush approach allows us to make statements that generalize throughout the labor market.

Several recent studies provide empirical evidence on the relationship between working place conditions and health. Fletcher et al. (2011) report results from an ordered probit random effects regression of self-assessed health on physical demands and environmental conditions of type of occupation and on a set of control variables, including lagged health. They combined information from the U.S. National Longitudinal Study of Youth with occupational characteristics in the Dictionary of Occupational Titles. They conclude that physical demands and environmental conditions can harm the health of women and older workers. Morefield et al. (2011) look at the effect of a history of employment status on a binary outcome variable indicating whether people reported bad health, which they define as reporting poor, fair or good health. They define good health as reporting very good or excellent health. They report that five years of blue collar employment predicts a four to five percent increase in the probability of moving from excellent or very good health to good or fair or poor health.

In particular, first occupation may have an effect on health later in life. First occupation is predetermined during the remainder of the lifetime, so it may solve issues of simultaneity between health and occupational choice. Sindelar et al. (2007) report that people whose first occupation was as a professional or a manager are less likely to report fair or poor health and to suffer from a heart attack. Fletcher and Sindelar (2009) acknowledge the endogeneity of occupational choice and use fathers occupation during childhood and the proportion of blue collar workers in the state of residence as instrumental variables for first occupation. They report that first occupation in a blue collar job has a negative effect on self-assessed health if one assumes that the instrumental variables are valid. Kelly et al. (2011) use the method of internal instruments as developed by Lewbel (2007) and a method that was developed by Altonji et al. (2005) to deal with selection into occupation on the basis of unobservable factors that may also be related to health status by imposing a correlation between unobservable factors and assuming that the degree of sorting on the basis of observables is equal to the degree of sorting on the basis of unobservables. They find that entering the labor market as a blue collar worker raises the probabilities of obesity by 4 percent and of smoking by 3 percent. This indicates that the effect of occupation on health may be transmitted through lifestyle factors.

We aim to contribute to the literature on how certain characteristics of work are detrimental to health, and which groups of workers are most vulnerable. Policymakers may be interested to learn more about how work affects health for two main reasons. First, people at the bottom of the income distribution suffer most from bad health and well-targeted labor market policies may play a role in reducing inequality and increasing social welfare. Second, detailed knowledge on the relationship between work and health may allow policymakers to efficiently allocate scarce resources. They can focus on implementing policies that achieve the highest health gains at the lowest cost. Policies that reduce the adverse effects of hard work can be targeted at vulnerable groups of workers that have the highest chance of dropping out of the labor force because of health problems, thus keeping them at work. Knowledge on the relationship between work, aging and health can inform the debate on the retirement age in many Western countries with an aging population. Policymakers can choose to exempt certain groups of workers who are most vulnerable to hard work at an older age from an increased retirement age.

The paper is organized as follows: section 2 introduces the German Socioeconomic Panel. Section 3 shows how we estimate the effect of manual work on health. Section 4 presents the results. Section 5 discusses how our results contribute to the literature and concludes.

2 German Socioeconomic Panel data

The German Socioeconomic Panel (GSOEP) started in 1984 and data from 26 yearly waves is currently available. For the purpose of the present study, we limit our sample to observations of persons who are between 16 and 65 years old. Our sample consists of 358,281 person-wave observations. Sample sizes per wave range from 8,681 in the year 1989 to 20,912 in 2000. Respondents are followed over multiple waves but the panel is unbalanced, which means in our case that many respondents enter the sample after the year 1984 or leave the sample before 2009. The GSOEP includes detailed yearly measures of level of education, individual and household earnings and other sources of income. Table 1 shows some summary statistics for the pooled person-year observations. Standard conversion techniques were used by the German Institute for Economic Research to map educational degrees into years of schooling. Respondents were asked to rate their satisfaction with their own health on an integer scale from 0 to 10. Table 2 shows the distribution of satisfaction with health.

Occupational titles were coded into the International Standard Classification of Occupations of the OECD (ISCO-88). These are 311 occupational classes that were grouped into ten major occupational groups by the OECD. Table 3 shows the frequencies for each of these occupational categories with the exception the armed forces. We drop observations from the armed forces from the sample because of the specific nature of this occupational class. On the basis of the OECD classifications, we define nonmanual workers as legislators, senior officials, managers, professionals, technicians, associate professionals, clerks, service workers and shop and market sales workers. We define manual workers as skilled agricultural and fishery workers, craft and related trades workers, plant and machine operators, assemblers and workers in elementary occupations. According to these definitions, we have a total of 147,053 observations for nonmanual work and 86,149 observations for manual work. In our sample there are 3,211 switches from manual work to nonmanual work between one year and the next and 2,580 switches from nonmanual work to manual work. Table 4 shows how workers switch between the nine major occupational groups from one year to the next. Labor earnings were censored at a monthly labor income of 100,000 Euro. If we disregard censoring, average net monthly labor income for nonmanual workers was 1,463 Euro and for manual workers this was 1,098 Euro.

3 Modeling and estimating the effect of manual work on health

Based on the Grossman (1972) health capital model, we model health status as a function of time-invariant individual characteristics, type of occupation in the previous period, time-varying individual characteristics, health status in the previous period and an error term. We would like to obtain a consistent estimate of the effect of type of occupation in the previous period on health.

$$HEALTH_{it}^* = \alpha_i + \beta OCC_{it-1} + \gamma HEALTH_{it-1}^* + \delta CHAR_{it} + \varepsilon_{it} \tag{1}$$

Where $HEALTH_{it}^*$ is a latent variable indicating true health status, α_i is a time-invariant individual-specific effect, OCC_{it-1} is a scalar or vector indicating occupational characteristics, $CHAR_{it}$ is a vector of time-varying individual characteristics and ε_{it} captures the non-deterministic variation in latent health status. In this model, β is a scalar or vector that captures the impact on health of being exposed at most one year to one or more characteristics of occupation.

For now, we assume that equation (1) describes the true model and that true health status can be directly observed. Naïve estimations of the effect of type of occupation on health can suffer from two types of omitted variable bias. The first type is time-invariant omitted variable bias. The estimate of

β is biased if

$$Cov[\alpha_i, OCC_{it-1}] \neq 0,$$

but α_i is omitted from the regression equation and instead a common intercept α is included. Then the error term in this regression equation becomes

$$\tilde{\varepsilon_i} = \varepsilon_i + \alpha_i - \alpha$$

and thus

$$Cov[\tilde{\varepsilon_i}, OCC_{it-1}] \neq 0.$$

This means that the OLS estimator is biased and inconsistent. For example, time-invariant omitted variable bias would occur if a person suffers from a disability since birth and we do not observe this disability and we do not account for time-invariant individual-specific effects. The disability affects occupational choices and health, and failing to account for this leads to biased estimates of β.

The second type is time-varying omitted variable bias. The estimate of β is biased if time-varying variables that are correlated with OCC_{it-1} are omitted from the regression equation. Here we view $HEALTH^*_{it}$ as a stock variable that is generated by an AR(1) process, such that

$$HEALTH^*_{it} = \theta HEALTH^*_{it-1} + \eta_{it}.$$

If we omit $HEALTH^*_{it-1}$ from the model,

$$\bar{\varepsilon_i} = \gamma HEALTH^*_{it-1} + \varepsilon_{it}$$

and

$$Cov[OCC_{it-1}, \bar{\varepsilon_i}] \neq 0$$

because

$$Cov[OCC_{it-1}, HEALTH^*_{it-1}] \neq 0.$$

For example, time-varying omitted variable bias would occur if an unobserved event, such as breaking a leg, led to a change in health and a change in type of occupation in period t-1. If we would omit the lagged value of the dependent variable $HEALTH^*_{it-1}$ from our regression equation, the estimate of β would be biased.

We have shown that we should account for α_i and $HEALTH^*_{it-1}$ in our estimation strategy. We assume that all time-varying factors that affect current health are either not correlated with job switching in the previous period, or are transmitted by last periods health, which is a control variable. This implies that excluding time-varying individual characteristics $CHAR_{it}$ reduces the explanatory power of the model and decrease the precision of our estimates, but it does not lead to omitted variable bias in the estimate of β.

Now we relax the assumption that true health status $HEALTH^*_{it}$ can be directly observed. Instead, we observe satisfaction with health, $SATHEALTH_{it}$ on an integer scale from 0 to 10. Health status is both a dependent variable and an independent variable in our model. This may lead to three possible complications. First, satisfaction with health could be viewed as the sum of the true health status and white noise measurement error. We use satisfaction with health as an independent variable, so this type of measurement error will lead to an attenuation bias toward zero of our estimate of γ. Following Bertrand and Mullainathan (2001) we assume that measurement errors in the independent variable are not dominant, and while we should be cautious to interpret our biased estimate of γ, we can still include $SATHEALTH_{it-1}$ as a control variable. For the dependent variable, this type of measurement error does not lead to bias but only to bigger standard errors.

Second, individuals with an identical level of objective health may report differently on the basis of unobservable characteristics. We argue that our estimate of β is not biased even if this reporting heterogeneity is correlated with type of occupation because we have included α_i in the model. If for example people with low education would rate an identical level of objective health higher than highly educated people, but they do this consistently over their lifetime, the shift that is due to this reporting heterogeneity will be absorbed by α_i.

Third, it is not ex ante clear why we could assume that satisfaction with health is a cardinal variable. If satisfaction with health is an ordinal variable, an ordered probit or ordered logit approach should be preferred over a linear regression model. Ferrer-i-Carbonell and Frijters (2004) show that for variable that measure satisfaction with life on an eleven point scale, GSOEP, assuming ordinality or cardinality makes little difference. They also find that allowing for fixed effects does change results substantially. Based on their findings we conclude that using satisfaction with health rather than an objective health measure does not threaten the validity of our results.

The preceding section shows that instead of equation (1) we can estimate

$$SATHEALTH_{it} = \alpha_i + \beta OCC_{it-1} + \gamma SATHEALTH_{it-1} + \varepsilon_{it} \tag{2}$$

As we will show, a consistent estimation strategy of equation (2) is unavailable. Our data permit us only to estimate an upper and a lower bound of the effect of type of occupation on education. We use two estimation specifications that are biased in opposite directions and we argue that the true effect is bounded between the results of these two regressions.

Our model in specification (1) includes a time-invariant individual-specific effect and a lag of the dependent variable to control for events that simultaneously affected health status and selection into certain types of work. Fixed and random effects estimation is biased and inconsistent for finite T if the assumption of strict exogeneity is violated. Strict exogeneity implies that $E[\varepsilon_{it}|X_{i1}, ..., X_{iT}, \alpha_i] \neq 0$, that the error term is uncorrelated with the regressors X_{it} in all time periods and with the individual-specific effect. Inclusion of a lagged value of the dependent variable as a regressor leads to a violation of this assumption and therefore to inconsistency of the random effects estimator and the class of fixed effects estimators.

A possible solution to the problem of inconsistent estimators is to use an instrumental variable approach to the first-difference estimator in a GMM framework (Arellano and Bover, 1995; Blundell and Bond, 1998). Firstdifferencing equation (2), gives

$$SATHEALTH_{it} - SATHEALTH_{it-1} = \beta(OCC_{it-1} - OCC_{it-2}) \\ +\gamma(SATHEALTH_{it-1} - SATHEALTH_{it-2}) + \varepsilon_{it} - \varepsilon_{it-1} \qquad (3)$$

The individual specific effect α_i has been differenced out. We can not use OLS to estimate equation (3), as

$$E[\varepsilon_{it} - \varepsilon_{it-1} | SATHEALTH_{it-1} - SATHEALTH_{it-2}] \neq 0$$

because $SATHEALTH_{it-1}$ is correlated with ε_{it-1} in equation (3). The first-difference estimator is thus biased and inconsistent.

The so-called Arellano-Bond estimator, which is based on the first-difference estimator, does provide consistent estimates under the assumption that ε_{it} is serially uncorrelated. $SATHEALTH_{it-2}$ and further lags are uncorrelated with $\varepsilon_{it} - \varepsilon_{it-1}$ and can be used as instrumental variables for $SATHEALTH_{it-1} - SATHEALTH_{it-2}$. Tests show that the assumption of no autocorrelation does not hold for our sample. Generally, one way to try to overcome this problem is to include more lags of the regressors in the model and use further lags of the instruments to try to get rid of the autocorrelation.

Unfortunately, four data limitations reduce our sample size to the point that the Arellano-Bond estimator is not feasible for our analysis. First, the first-difference estimator of the coefficient of OCC_{it-1} is based only on observations for which people have switched between manual and nonmanual work between years t-1 and t-2. As we mentioned before, we have 6,791 of these occupational switches between one year and the next in our data. Second, we need observations for manual and nonmanual work. This means that we lose observations if people left the labor force for at least a year. Third, as we have explained, the GSOEP is an unbalanced panel data set and this leaves us with insufficient observations to successfully implement the Arellano-Bond estimator because we need lagged values of $SATHEALTH_{it}$ to get rid of the autocorrelation in the error terms. For example, there are only 1,967 observations for which we have nine or more lagged values of the dependent variable and for which there was a switch between manual and nonmanual work between years t-1 and t-2. Fourth, health is measured on an eleven point integer scale. This limits the variation in the outcome variable.

We conclude that although we have a very large data set, we can only use a small number of observations because there is autocorrelation in the error term, we use a dichotomous classification of type occupation and there is little variation in the outcome variable because of the way health was measured. This leaves us with an insufficient number of switches between manual and nonmanual work to implement the Arellano-Bond estimator.

Instead of equation (2) we estimate equations

$$SATHEALTH_{it} = \alpha + \beta OCC_{it-1} + \gamma SATHEALTH_{it-1} + \varepsilon_i \quad (4)$$

and

$$SATHEALTH_{it} = \alpha_i + \beta OCC_{it-1} + \varepsilon_i \quad (5)$$

Following Angrist and Pischke (2009), we argue that we can estimate the bounds of the effect of manual work on health using two estimators that are biased, but in opposite directions such that the true coefficient is bounded by the results of the two biased estimates. To show the direction of the bias of the first estimate, assume that equation (5) describes the true model, but we estimate equation (4). Then the resulting estimator $\hat{\beta}$ has probability limit

$$\frac{Cov(SATHEALTH_{it}, OCC_{it} - \hat{\rho} SATHEALTH_{it})}{Var(OCC_{it} - \hat{\rho} SATHEALTH_{it})} = \beta + \frac{\hat{\rho}\sigma_\varepsilon^2}{Var(OCC_{it} - \hat{\rho})}$$

where $\hat{\rho}$ is the regression coefficient of a regression of OCC_{it} on $SATHEALTH_{it}$ and σ_ε^2 is the variance of ε_{t-1}. Since manual workers have worse health than nonmanual workers, $\hat{\rho} < 0$. Our estimate of $\hat{\rho}$ is -.0074 (.0005). If, as we assume, selection into type of occupation is correlated to α_i, estimating equation (4) leads to an estimate of β that is too small.

To show the direction of the bias of the second estimator, assume that equation (4) describes the true model but we estimate equation (5). Then the resulting estimator $\hat{\beta}$ has probability limit

$$\beta + (\gamma - 1)[\frac{Cov(SATHEALTH_{it}, OCC_{it-1})}{Var(OCC_{it})}]$$

where γ is larger than 0 and smaller than 1 because $SATHEALTH_{it}$ is stationary and $Cov(SATHEALTH_{it}, OCC_{it-1})$ is negative. The sample covariance is $-.0344$. Since manual workers have worse health, the estimate of β is too big if we use the first-difference estimator for equation (5) but selection into type of occupation depends on events that also affected health directly. We conclude that the true parameter lies between these two biased estimates. In the next section we present the estimation results for these bounds.

4 Results

Columns (1) and (2) of table 5 show the bounds of the effect of less or equal than one year of exposure for manual work on health for all age groups. Column (2) shows that our estimate of the upper bound of the effect is not significantly different from zero. Columns (3) and (4) of table 5 show that we do not find evidence of an effect for the group of workers who are younger than 50 years. Our estimation results in column (5) and (6) in table 5 show that there is sufficient evidence to conclude that manual work leads to a faster deterioration of health than nonmanual work for the group of workers between 50 and 65 years old at the 5 percent significance level. The point estimate of the upper bound of the effect of nonmanual work on satisfaction with health in column (4) is $-.098$ and the point estimate of the lower bound is $-.213$. To give an indication of the importance of this effect, we performed a fixed effects regression of satisfaction with health on age and we find that aging decreases satisfaction with health by $-.048(.002)$. A very casual interpretation of the effect size of the point estimate in column (6) is that doing manual work instead of nonmanual work at an older age leads to a health deterioration that is comparable to the total effect of aging at least two years for the same age group. As we explained before, the regression coefficient of satisfaction with health in the previous period should be interpreted with care because it represents the sum of true health status and measurement error. It was included as a control variable.

Table 6 focuses on the results for older workers. We argued earlier that omitting time-varying individual characteristics $CHAR_{it}$ in equation (1) would not lead to bias if we could estimate β consistently. However, our current estimates are already biased in opposite directions. Possibly, the inclusion of additional control variables could remove part of the bias, thereby tightening the bounds. Columns (1) and (2) in table 6 show the bounds of the effect for manual workers who are between 50 and 65 years old, but now including monthly labor income in thousands of Euro in the previous year as control variables. The specifications that are presented in column (3) and (4) in table 6 include monthly labor income, an age polynomial of the third degree and wave dummies. Additionally, the specification in column 3) controls for level of education, measured as years of schooling, and gender. The coefficients for level of education and gender are not defined in the fixed effects specification of column (4) as these are time-invariant and absorbed by the individual-specific time-invariant effect.

The regression coefficients for manual work at t-1 for people between 50 and 65 years old in the fixed effects specifications in columns (2) and (4) of table 6 do not differ much from their counterpart in column (6) of table 4. The same holds for the regression coefficient for manual work at t-1 for people between 50 and 65 years old in column (1) of table 6 and the coefficient in column (5) of table 4. However, the inclusion of covariates in column (3) of table 5 leads to an upward shift of the estimate of the lower bound of the effect of manual work in comparison to column (3) in table 4. Our explanation for this upward shift is that the additional control variables removed part of the bias because our control variables pick up part of the fixed effect that was omitted from the model that we used to estimate the lower bound. Therefore, our best estimate of the lower bound of the effect of manual work on satisfaction with health for people between 50 and 65 years old is -.126 (.018) instead of -.213 (.017).

We define permanent labor earnings as expected lifetime labor earnings and transitory labor earnings as temporary deviations from permanent labor earnings. The coefficients of monthly labor income in columns (2) and (4) of Table 6 are purged of the effect of permanent labor earnings because of the removal of the fixed effect and estimate the effect of transitory labor earnings. The negative regression coefficients in columns (2) and (4) indicate that transitory labor earnings has a negative effect on health, so health deteriorates faster if current labor earnings are higher than permanent labor earnings. The coefficients of monthly labor income in columns (1) and (3) of Table 6 are positive. They represent the estimated total effect of permanent and transitory labor earnings, but the estimates are biased because of omitted variable bias and reporting heterogeneity as was mentioned earlier in the context of type of occupation. Therefore these coefficients should be interpreted with care.

No other suitable control variables are available in the GSOEP. For example, the variable that indicates the presence of a steady partner has 236,180 missing values which leads to an unacceptable reduction of the sample size. Results for workers who are between 16 and 49 years old are similar to those that we reported in table 5 for older workers but not shown here. Output is available upon request.

The dichotomization of type of occupation is somewhat arbitrary. Table 6 shows the monotonic relationship between five ordered categories of occupation and health. These five categories are based on the ISCO-88 classification and each category contains roughly twenty percent of the observations of type of occupation. The results in table 6 suggest that the results are not an artefact of our dichotomization.

We were not able to determine whether the effect of manual work was different for people who were born in East or in West Germany as data on the region of birth was unavailable. Analyses for each gender separately did not yield different results, but we can not conclude anything from these findings because we have not taken into account the possibility that women may be selected into a different subset of manual occupations than men.

The effect of manual work on health may be different for people who have done manual work for most of their lives. We were not able to determine whether occupational history interacted with manual work because we have very few observations of people from the age that they started working and we dont have the complete occupational history.

5 Discussion and conclusion

In this paper we present the bounds of the effect of doing manual work instead of nonmanual work. We contribute to the literature in several ways. Studies by Marmot (1991), Case and Deaton (2005), Choo et al. 2006, Fletcher and Sindelar (2007), Morefield et al. (2011) may suffer from time-invariant omitted variable bias. We account for this by including individual-specific effects.

We account for unobserved factors that affect work in period t-1 and satisfaction with health in period t by including satisfaction with health in period t-1. This is not taken into account by Marmot (1991), Case and Deaton (2005), Choo et al. (2006) and Sindelar et al. (2007). We make no assumptions about the correlation between selection on observables and on unobservables such as Kelly et al. (2011).

We account for reporting heterogeneity by controlling for individual-specific effects in contrast to Case and Deaton (2005), Sindelar et al. (2007), Fletcher and Sindelar (2009), Fletcher et al. (2011), Morefield et al. (2011).

As reported earlier, Fletcher et al. (2011) use an ordered probit random effects model that included the lagged dependent variable as a regressor. This yields inconsistent estimates for two reasons. First, random effects estimation depends on the strict exogeneity assumption. Including the dependent variable as a regressor leads to a violation of this assumption and to inconsistent estimation. Second, the random effects framework relies on invalid distributional assumptions of the error term. We have shown that the error terms in a regression of satisfaction with health on type of occupation are serially correlated. The violation of the distributional assumptions is the second source of inconsistency of the random effects estimator.

Finding sources of exogenous variation of type of occupation could shed more light on the causal relationship between type of occupation and health. Fletcher and Sindelar (2009) use father's occupation and proportion of blue collar workers at the U.S. state level as instrumental variables for type of occupation. Exogeneity of these instrumental variables seems implausible since fathers occupation and state level macroeconomic variables are likely to affect health through other channels than occupation. Fathers occupation may be correlated with household income throughout childhood, which may affect health through an unhealthy diet pattern or other channels. High state employment levels may be associated with low per capita statewide healthcare spending, which could indicate a lower level of preventive and curative care. The instrumental variables are not exogenous if this is the case. Kelly et al. (2011) cast doubt on the relevance of these two instrumental variables, while Fletcher and Sindelar (2009) refrain from reporting the results from the first stage of the two stage least squares estimator. The present study does not rely on invalid instruments.

A weakness of the present paper is that we only make a distinction between manual and nonmanual labor. We intend to extend the analysis by constructing continuous scales of physiological, ergonomic and psychosocial workplace conditions. We are in the process of obtaining detailed information on sixteen of these workplace conditions that are available in the Finnish Job Exposure Matrix, under the assumption that the ordering of workplace conditions in Germany and Finland is roughly the same. The variables in FINJEM are based on the assessment of occupational health specialists. The U.S. Occupational Information Network also has data on workplace conditions. However, this information is not gathered by experts but based on worker assessments, which leads to concerns about reporting heterogeneity. The U.S. Dictionary of Occupational Titles has much less detailed information and hasnt been updated for more than a decade. This is why using FINJEM seems the appropriate choice.

Using continuous scales of workplace conditions as measures of type of occupation solves the problem that we encountered and that was caused by the small number of job switches. Whereas an analysis that is based on the distinction between manual and nonmanual work loses a lot of information because of its simplicity, a continuum of workplace conditions enables us to exploit each job switch for analytical purposes. This may make the consistent Arellano-Bond estimation procedure feasible and could provide more insight into what it is exactly about work that makes people more healthy or unhealthy and who suffers most and why. Another advantage is that the analysis does not solely rely on the small, perhaps atypical, group of respondents who switch between manual and nonmanual work.

Summarizing, this paper provides compelling evidence that doing manual work leads to worse health than doing nonmanual work for older but not for younger workers. In contrast to earlier studies, we account for several sources of bias and present tight bounds for a consistent estimator. The point estimate of most conservative bound effect suggests that manual workers over 50 years old age three times faster than nonmanual workers.

Policymakers may take into account that manual work affects health at an older age and they may create provisions to protect people at vulnerable ages against the harmful effects of manual work. If education determines occupation throughout the lifetime, work-related deterioration of health may be an important transmission mechanism of the effect of education on health. Educational policies may have unintended health effects if they affect the career paths of students. These policy issues underline the importance of further research on this topic for determining and eventually improving the socioeconomic determinants of health.

6 Appendix A: Tables

Tables

Variable	Mean	Standard deviation	Observations
Age	40.40	13.50	358,281
Female	.51	-	358,280
Years of schooling	11.65	2.64	344,149
Monthly labor income	1,312.24	1,093.26	244,059

Table 1: Variables in the GSOEP

Satisfaction with health	Frequency
0	4,008
1	2,980
2	8,033
3	15,277
4	17,757
5	45,720
6	33,697
7	58,785
8	87,510
9	46,327
10	37,377
Total	357,471

Table 2: Satisfaction with health in the GSOEP. 0 is not satisfied, 10 is satisfied

ISCO-88 major group	Frequency
Legislators, senior officials and managers	13,351
Professionals	33,930
Technicians and associate professionals	47,772
Clerks	27,366
Service workers and shop and market sales workers	24,634
Skilled agricultural and fishery workers	3,183
Craft and related workers	43,220
Plant and machine operators and assemblers	21,114
Elementary occupations	18,632
Total	233,202

Table 3: ISCO-88 occupations in the GSOEP. We consider the four major occupational groups below the dashed line to be manual work.

	Legislators, senior officials and managers at t	Professionals at t	Technicians and associate professionals at t	Clerks at t	Service workers and shop and market sales workers at t	Skilled agricultural and fishery workers at t	Craft and related workers at t	Plant and machine operators and assemblers at t-1	Elementary occupations at t-1	Total
Legislators, senior officials and managers at t-1	8,943	504	545	230	211	18	200	94	64	10,809
	(82.74)	(4.66)	(5.04)	(2.13)	(1.95)	(0.17)	(1.85)	(0.87)	(0.59)	(100.00)
Professionals at t-1	593	25,239	1,125	240	84	18	96	37	68	27,500
	(2.16)	(91.78)	(4.09)	(0.87)	(0.31)	(0.07)	(0.35)	(0.13)	(0.25)	(100.00)
Technicians and associate professionals at t-1	635	1,196	33,802	1,300	607	25	406	177	211	38,359
	(1.66)	(3.12)	(88.12)	(3.39)	(1.58)	(0.07)	(1.06)	(0.46)	(0.55)	(100.00)
Clerks at t-1	269	282	1,479	18,944	319	14	107	131	260	21,805
	(1.23)	(1.29)	(6.78)	(86.88)	(1.46)	(0.06)	(0.49)	(0.60)	(1.19)	(100.00)
Service workers and shop and market sales workers at t-1	243	137	740	393	16,532	12	149	117	376	18,699
	(1.30)	(0.73)	(3.96)	(2.10)	(88.41)	(0.06)	(0.80)	(0.63)	(2.01)	(100.00)
Skilled agricultural and fishery workers at t-1	21	15	29	11	24	2,178	34	41	94	2,447
	(0.86)	(0.61)	(1.19)	(0.45)	(0.98)	(89.01)	(1.39)	(1.68)	(3.84)	(100.00)
Craft and related workers at t-1	290	113	574	171	171	21	31,807	1,002	605	34,754
	(0.83)	(0.33)	(1.65)	(0.49)	(0.49)	(0.06)	(91.52)	(2.88)	(1.74)	(100.00)
Plant and machine operators and assemblers at t-1	118	47	205	162	113	24	903	15,014	569	17,155
	(0.69)	(0.27)	(1.19)	(0.94)	(0.66)	(0.14)	(5.26)	(87.52)	(3.32)	(100.00)
Elementary occupations at t-1	75	127	275	317	353	93	571	586	11,352	13,749
	(0.55)	(0.92)	(2.00)	(2.31)	(2.57)	(0.68)	(4.15)	(4.26)	(82.57)	(100.00)
Total	11,187	27,660	38,774	21,768	18,414	2,403	34,273	17,199	13,599	185,277
	(6.04)	(14.93)	(20.93)	(11.75)	(9.94)	(1.30)	(18.50)	(9.28)	(7.34)	(100.00)

Table 4: Job switches between major occupational groups in two subsequent years. The table does not show respondents who were not employed in nonmilitary occupations for both years. Each observation in the table represents one respondent in two subsequent years. Percentages in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
Manual work at t-1	-.084**	-.021	-	-	-	-
	(.002)	(.028)				
Manual work at t-1 * age 16-49	-	-	-.054**	.015	-	-
			(.009)	(.029)		
Manual work at t-1 * age 50-65	-	-	-	-	-.213**	-.098*
					(.017)	(.045)
SWH at t-1	.569**	-	.558**	-	.556**	-
	(.008)		(.002)		(.002)	
Specification	Pooled OLS	Fixed effects	Pooled OLS	Fixed effects	Pooled OLS	Fixed effects

Table 5: Main results for satisfaction with health at time t. Panel-robust standard errors in parentheses. * indicates significance at 5 percent level, and ** at 1 percent level. Each of the three pairs of regression specifications show estimation results for respectively the lower and upper bound of the effect of manual work on health for the total population and each age group. Base category is nonmanual work.

	(1)	(2)	(3)	(4)
Manual work at t-1 * age 50-65	-.192**	-.099*	-.126**	-.098*
	(.017)	(.045)	(.018)	(.045)
SWH at t-1	.556**	-	.540**	-
	(.002)		(.002)	
Labor earnings at t-1	.029**	-.033**	.047**	-.022**
	(.017)	(.009)	(.005)	(.008)
Education	-	-	.020**	-
			(.002)	
Female	-	-	-.030**	-
			(.009)	
Third order age polynomial	No	No	Yes	Yes
Wave dummies	No	No	Yes	Yes
Specification	Pooled OLS	Fixed effects	Pooled OLS	Fixed effects

Table 6: Robustness checks for satisfaction with health at time t for workers between age 50 and 65. Panel-robust standard errors in parentheses. * indicates significance at 5 percent level, and ** at 1 percent level. Each of the two pairs of regression specifications show estimation results for respectively the lower and upper bound of the effect of manual work on health. Education is measured as years of schooling and labor earnings is measured in thousands of Euro of monthly labor income. Base category is nonmanual work.

OcCat3 at t-1	-.102**
	(.011)
OcCat45 at t-1	-.141**
	(.011)
OcCat67 at t-1	-.152**
	(.012)
OcCat89 at t-1	-.233**
	(.013)
SWH at t-1	.542**
	(.002)
Third order age polynomial	Yes
Wave dummies	Yes
Specification	Pooled OLS

Table 7: The effect of five occupational categories on satisfaction with health at time t. Panel-robust standard errors in parentheses. ** indicates significance at 1 percent level. Base category includes legislators, senior officials, managers and professionals. OcCat3 includes technicians and associate professionals. OcCat45 includes clerks, service workers and shop and market sales workers. OcCat67 includes skilled agricultural and fishery workers and craft and related workers. OcCat89 includes plant and machine operators, assembles and elementary occupations

7 Appendix B: The Finnish Job Exposure Matrix

Type of occupation is indicated in the GSOEP by ISCO-88, a nominal variable that can take the values of 311 occupational titles. In future versions of this paper, we intend to map the occupational codes of the GSOEP into indices of physical and psychosocial workplace conditions which we will use to explain health status. These occupational characteristics are unavailable for Germany so we will use detailed information from Finland under the assumption that workplace characteristics in Germany and Finland are similar.

The Finnish Job Exposure Matrix (FINJEM) maps about 360 occupational titles into eight indicators of physical workplace conditions, which are defined as measures of high accident risk, inconvenient and difficult work postures, manual handling of burdens, perceived physical work load, repetitive work movements, sedentary work, standing work and work with video display units, and into nine indicators of psychosocial workplace conditions, which are defined as challenge at work, social climate at work, control possibilities at work, work load, risks at work, social demands at work, supervisor support at work, working time arrangements and night work.

These characteristics are update periodically and, with the exception of perceived physical workload, are based on the subjective assessment of occupational health specialists. Despite the subjective nature of these variables, we argue that they are superior to other available measures of workplace conditions. Measures of workplace conditions that are based on workers' own assessments may suffer from reporting heterogeneity which may be correlated to other worker characteristics. In the FINJEM, each indicator was constructed by one occupational health specialist, which rules out reporting heterogeneity. Data from 1945 until 2003 is available, work on estimates until 2009 is in progress.

References

Altonji, J. G., Elder, T., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29 – 51.

Bertrand, M. and Mullainathan, S. (2001). Do people mean what they say? implications for subjective survey data. *American Economic Review*, 91(2):67–72.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115 – 143.

Case, A. and Deaton, A. S. (2005). Broken down by work and sex: How our health declines. In *Analyses in the Economics of Aging*, NBER Chapters, pages 185–212. National Bureau of Economic Research, Inc.

Choo, E. and Denny, M. (2006). Wearing out – the decline in health. Working Papers tecipa-258, University of Toronto, Department of Economics.

Cutler, D. M., Lleras-Muney, A., and Vogl, T. (2008). Socioeconomic status and health: Dimensions and mechanisms. Working Paper 14333, National Bureau of Economic Research.

Fletcher, J. M. and Sindelar, J. L. (2009). Estimating causal effects of early occupational choice on later health: Evidence using the psid. Working Paper 15256, National Bureau of Economic Research.

Fletcher, J. M., Sindelar, J. L., and Yamaguchi, S. (2011). Cumulative effects of job characteristics on health. *Health Economics*, 20(5):553–570.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–255.

Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, 38(1):pp. 21–37.

J.P., M., J.G., S., C.W., L., and I.M., J. (2002). Self-assessed health and mortality: could psychosocial factors explain the association? *International Journal of Epidemiology*, 31(6):1162–1168.

Karasek, Robert A., J. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, 24(2):pp. 285–308.

Kelly, I. R., Dave, D. M., Sindelar, J. L., and Gallo, W. T. (2011). The impact of early occupational choice on health behaviors. Working Paper 16803, National Bureau of Economic Research.

Lewbel, A. (2010). Using heteroskedasticity to identify and estimate mismeasured and endogenous regressor models. Boston College Working Papers in Economics 587, Boston College Department of Economics.

Lindeboom, M. and van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics*, 23(6):1083 – 1099.

Marmot, M., Bosma, H., Hemingway, H., Brunner, E., and Stansfeld, S. (1997). Contribution of job control and other risk factors to social variations in coronary heart disease incidence. *The Lancet*, 350(9073):235 – 239.

Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A., Marmot, M. G., and Smith, G. D. (1991). Health inequalities among british civil servants: the whitehall ii study. *The Lancet*, 337(8754):1387 – 1393.

Morefield, G. B., Ribar, D. C., and Ruhm, C. J. (2011). Occupational status and health transitions. Working Paper 16794, National Bureau of Economic Research.

Mossey, J. M. and Shapiro, E. (1982). Self-rated health: a predictor of mortality among the elderly. *Am J Public Health*, 72(8):800–808.

Sindelar, J. L., Fletcher, J., Falba, T., Keenan, P., and Gallo, W. T. (2007). Impact of first occupation on health at older ages. Working Paper 13715, National Bureau of Economic Research.