The Wear and Tear on Health: What Is the Role of Occupation?

Bastian Ravesteijn[ab] Hans van Kippersluis[bc]
Eddy van Doorslaer [bd]

Abstract

Health is well known to show a clear gradient by occupation. While it may appear evident that occupation affects health, there are multiple sources of selection that preclude the strong association to be interpreted as exclusively deriving from a causal effect of occupation on health. Despite abundant literature documenting the association, quantification of the relative importance of selection into occupation and the effect of occupation on health is scarce. We link job characteristics to German panel data spanning 29 years to characterize occupations by their physical and psychosocial burden. Employing a dynamic model to control for factors that simultaneously affect health and selection into occupation, we find that selection into occupation accounts for at least 60 percent of the association between health and both physical strain and job control, while selection accounts for nearly 100 percent of the association between psychosocial workload and health. The residual effect of occupational characteristics such as physical strain and low job control is negative and increases with age. The effects of late-career exposure of one year to high physical strain and low job control are equivalent to the health deterioration from ageing 16 and 6 months, respectively.

[a]Department of Health Care Policy, Harvard Medical School, 180A Longwood Avenue, Boston, MA 02115, USA (corresponding author, e-mail: ravesteijn@hcp.med.harvard.edu, telephone: +1 (617) 432 3138).

[b]Department of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands.

[c]Department of Economics, Chinese University of Hong Kong, Hong Kong, China.

[d]Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, the Netherlands.

Keywords: occupational stressors, health economics, labour economics, dynamic models.
JEL: I0, J0, C23.

1 INTRODUCTION

It is well known that average health and life expectancy display a clear gradient by occupation (e.g. Marmot et al. 1991). Manual workers in the US, for example, are 50 percent more likely to die within a given year than workers in managerial, professional and executive occupations (Cutler et al. 2008). The mortality rate for manual workers in Europe is higher than for non-manual workers throughout the age distribution, and this gap has widened over time (Mackenbach et al. 2003). For the Netherlands, Ravesteijn et al. (2013) find a strong gradient in self-assessed health by occupational class—particularly at an older age—and note that 20 percent of elementary workers[1] at the age of 60 have exited the workforce into disability, as opposed to eight percent of workers in occupations that require academic training.

While the occupation/health gradient is widely documented, less is known about the underlying mechanisms that generate it. Occupation may exert a causal effect on health, but the strong correlation may also stem from reverse causality, with health constraining occupational choice. Moreover, individuals in different occupational groups can differ in other—observed and unobserved—"third factors" that influence health. For example, manual workers are less educated or

[1] Elementary occupations consist of simple and routine tasks which mainly require the use of hand-held tools and often some physical effort.

possess different genetic predispositions than non-manual workers. Both reverse causality and third factors may lead to selection effects: people with good health prospects are selected into certain types of occupations. As a result, the magnitude of the association between occupation and health is likely to be (much) higher than the magnitude of the causal effect of occupation on health.

Identification of health effects of occupations is policy relevant for at least two reasons. From a fairness perspective, health disparities that result from a heavily constrained occupational choice set may be more objectionable than differences in health due to free choice behaviour, such as smoking or drinking. From a productivity perspective, it is vital for both regulators and employers to know which specific occupational characteristics are most harmful to health. For example, occupations with harmful ergonomic workplace conditions may simultaneously be characterized by low control possibilities at work, which may exert an independent effect on health. Consequently, separating out these effects is crucial for better-targeted efforts at reducing sickness absenteeism and disability by adjusting specific labor conditions.

Many studies have documented strong associations between occupational characteristics and health (see e.g. Kunst et al. 1999, Goodman 1999), but few have attempted to obtain estimates of a causal effect, and those that do often focus on specific occupations or specific types of exposure to unhealthy circumstances.[2] The relationship between occupation and health has received surprisingly little attention in the economics literature, but interest in the topic has grown in recent years. Case & Deaton (2005) show that the self-reported health of manual workers is lower and declines more rapidly with age than that of non-manual workers.

[2]For example, Bongers et al. (1990) study back pain among helicopter pilots.

Choo & Denny (2006) report similar patterns for Canadian workers while controlling for a more extensive set of lifestyle factors and suggest that manual work has an independent effect on health over and above any differences in lifestyle across occupations. Using the longitudinal Panel Study of Income Dynamics (PSID), Morefield et al. (2012) estimate that five years of blue-collar employment predicts a four to five percent increase in the probability of moving from good health to poor health.[3]

The most comprehensive attempt to estimate the health impact of occupation is Fletcher et al. (2011) who combine information on the physical requirements of work and environmental conditions taken from the Dictionary of Occupational Titles (DOT) with occupational information in the PSID. Their aim is to estimate the health impact of five-year exposure to physical and environmental conditions as the DOT lacks information on psychosocial stressors. Controlling for first-observed health and five-period lagged health in their empirical model, they estimate negative health effects of physical requirements and environmental conditions. They acknowledge that the potential endogeneity of occupation and occupational change does not allow for a causal interpretation of their random effects estimates. Their data also do not permit to disentangle the contributions of physical and psychosocial occupational stressors.

[3]Apart from current occupation, a worker's entire occupational history is likely to affect current health. Thus, Fletcher & Sindelar (2009) use father's occupation during childhood and the proportion of blue-collar workers in the state as instrumental variables for first occupation and find that a blue-collar first occupation negatively affects self-assessed health. Kelly et al. (2012) question the statistical relevance of the two instrumental variables used in Fletcher & Sindelar (2009) and instead propose methods developed by Lewbel (2012) and Altonji et al. (2005) to investigate the causal effect of first occupation on health. They find that entering the labor market as a blue-collar worker raises the probabilities of obesity and smoking by four and three percent, respectively, which indicates that the effect of occupation on health may—at least in part—be transmitted through lifestyles.

In sum, the literature so far has failed to establish to what extent the strong association between occupation and health reflects a causal effect of occupation on health, and to what extent it reflects the selection of unhealthy individuals into occupations with harmful job characteristics. In part this is due to the difficulty of finding credible sources of exogenous variation in occupation. Previous studies have also had only limited success in disentangling the health effects of different types of occupational stressors.

Our contribution to the literature is threefold. First, we derive an empirical specification that is grounded in a theoretical model of occupation and health over the life cycle. The explicit link between the theory and the empirical specification (i) identifies the sources of health-related selection into occupation, (ii) shows how our econometric estimators relate to the structural parameters, and (iii) details the conditions under which our dynamic panel data estimates allow for a causal interpretation. These insights provide a theoretical foundation for the dynamic panel data estimation.

Second, we estimate an empirical model, on German longitudinal data, that can account for various sources of selection: (i) unobserved time-invariant variables due to the inclusion of individual fixed effects, (ii) time-varying observed variables such as age and wave dummies, and (iii) time-varying unobserved shocks that exponentially die out through the inclusion of the lagged dependent variable. We argue that with panel data spanning 29 years and in the absence of credible sources of exogenous variation in occupation, our model provides the most promising estimates of the relative importance of selection into occupation versus effects of occupation on health.

Third, we show that blue-collar occupations are both more physically demanding and more often characterized by low job control. Previous studies have often characterized occupation with a binary indicator of manual versus non-manual occupation or have focused only on the physical aspects of occupation. This approach has left the contributions of the various ergonomic and psychosocial stressors unseparated and made clear policy conclusions difficult to draw. By linking German data on occupational stressors to individual-level longitudinal data, we are able to unravel the health effects of job characteristics in greater detail.

Our findings suggest that at least 60 percent of the association between physical demands at work and self-reported health stems from the selection of individuals with worse health (prospects) into occupations with high physical demands, and that the same holds for the degree of job control. This leaves at most 40 percent that could potentially stem from the causal effect of both stressors on health. In contrast, our estimate of the effect of psychosocial workload is close to zero. Under the, admittedly stringent, assumptions laid out in the theoretical framework, we estimate that the average effect of one year exposure to a one standard deviation increase in the degree of physical strain (e.g., working as a toolmaker instead of as a teacher) is comparable to the effect of ageing nine months, and the effect increases with age. A lower degree of control over daily activities at work (e.g., a secretary versus a librarian) is harmful to health at older ages but not at younger ages. We estimate that exposure to a one standard deviation increase in handling heavy burdens between the ages of 50 and 54 leads to a health deterioration that is comparable to ageing 16 months. In other words, exposure to physical strain increases the biological health deterioration rate by 130 percent. The estimated effect of exposure to low job control between the ages of 50 and 54 is comparable to ageing 6 months, or an increase of the biological aging rate by

50 percent.

2 OCCUPATION AND HEALTH OVER THE LIFE CYCLE

In the economics literature, health is treated as a durable capital stock that depreciates with age and can be increased with investment (Grossman 1972). The age-related health depreciation rate is exogenous, but an individual can invest in his health by purchasing preventive and curative medical care. The effect of behaviour on health can be positive or negative. Occupational choice can be understood as a form of health disinvestment/erosion: an individual chooses an occupation that is characterized by a set of potentially harmful occupational stressors (Case & Deaton 2005, Galama & van Kippersluis 2010). Occupations with more harmful characteristics may yield higher earnings than other less harmful occupations in the choice set of the individual, which is known as the compensating wage differential (Smith 1974, Viscusi 1978). The additional earnings may be used to partially offset the detrimental effect of work on health by investing in health or to increase consumption. This economic paradigm is useful for distinguishing between the sources of health-related selection into occupation.

Our empirical investigation is based on a theoretical model of an individual maximising the expected present value of lifetime utility, which is derived from consumption c and health h, by choosing levels of consumption c, occupational stressors in vector $\mathbf{o}$, and health investment m. Each occupation is characterized by physical and psychosocial occupational stressors that tend to be clustered, i.e., occupations with low psychosocial workload are often characterized by high physical demands. Future utility is discounted at discount rate β. The information set $\mathcal{I}$ includes endowments e and permanent health h_p, all state and choice variables up to time t, and all future values of the ageing rate, but not future unanticipated health shocks η.

$$max_{\{c_{t+j},\mathbf{o}_{t+j},m_{t+j}\}_{j=0}^{T-t}} E\left[\sum_{j=0}^{T-t} \beta^j u(c_{t+j}, h_{t+j})|\mathcal{I}_t\right] \tag{1}$$

The health production function depends on (i) characteristics and circumstances that remain constant over time that are embodied by permanent health $h_p = f(e)$, which is a function of endowments and reflects all circumstances and personal characteristics that remain constant over the life cycle; (ii) anticipated health deterioration due to ageing a; (iii) a vector of (physical and psychosocial) occupational characteristics $\mathbf{o}$;[4] (iv) medical investment m; and (v) exogenous health shocks η. The effect of occupational characteristics on health, $\boldsymbol{\gamma}_o$, is nonpositive, and $0 \leq \theta \leq 1$ reflects diminishing marginal benefits to health investment. The effects of occupational stressors, health investments and shocks are assumed to decay at the same rate ϕ, which lies between 0 and 1. Total lifetime T is exogenous and known to the individual.

$$h_{t+j} = h_p + \sum_{k=2}^{t+j} \left(a_k + \phi^{t+j-k}(\boldsymbol{\gamma}_{\mathbf{o}}'\mathbf{o}_{k-1} + \gamma_m m_{k-1}^{\theta} + \eta_k)\right) \tag{2}$$

Expenditures on consumption and health investment, at prices p_c and p_m, respectively, should not exceed the net value of wage earnings. The individual can lend and borrow at real interest rate r, but he must repay any remaining debt at the

[4] In section 3 we will link this to the seminal work by Karasek (1979) on occupational stressors.

end of his life. Wage w is a function of (i) current occupational choice $\mathbf{o}$, (ii) current health h, and (iii) endowments e.

$$s.t. \sum_{k=1}^{T} (p_c c_k + p_m m_k) \leq \sum_{k=1}^{T} (1+r)^{k-1} w(\mathbf{o}_k, h_k; e) \tag{3}$$

Consumption, health investment and occupational choice are chosen by equating marginal benefit with marginal cost. The marginal utility of consumption is equal to the shadow price of income λ multiplied by the price of consumption.

$$\frac{\partial u_t}{\partial c_t} = \lambda p_c \tag{4}$$

For each occupational attribute o_l in vector $\mathbf{o}$, the marginal benefit of occupational stress is represented by the product of λ and the instantaneous wage premium. The marginal cost includes the marginal deterioration of health in all future periods multiplied by (i) the discounted marginal utility of future health and (ii) the product of λ and the present value of the marginal wage returns to future health.

$$\lambda \frac{\partial w_t}{\partial o_{t,l}} = -\sum_{j=1}^{T-t-1} \frac{\partial h_{t+j}}{\partial o_{tl}} \left[\beta^j \frac{\partial u_{t+j}}{\partial h_{t+j}} + \lambda \left(\frac{1}{1+r} \right)^j \frac{\partial w_{t+j}}{\partial h_{t+j}} \right] \quad \forall l \tag{5}$$

Health investment is the 'mirror image' of occupational choice. The marginal benefit (the product of the marginal effect of health investment on health and both the discounted marginal utility of health and the marginal wage returns to health in all future periods) is equated with marginal cost (the product of the shadow price of income and the price of medical care).

$$\sum_{j=1}^{T-t-1} \frac{\partial h_{t+j}}{\partial m_t} \left[\beta^j \frac{\partial u_{t+j}}{\partial h_{t+j}} + \lambda \left(\frac{1}{1+r} \right)^j \frac{\partial w_{t+j}}{\partial h_{t+j}} \right] = \lambda p_m \tag{6}$$

The theoretical framework shows how an individual takes the future consequences of his decisions into account while deciding on the optimal levels of harmful occupational stressors. Three insights from the theory are particularly noteworthy. First, both time-invariant initial endowments e—in the form of, for example, physical ability, intelligence or taste for adventure—and time-varying factors such as health shocks η—e.g. a car accident or the onset of a disease— may influence *both* occupational choice *and* health status through (i) the marginal utility of health, (ii) the marginal wage returns to health, and (iii) the shadow price of income λ. This finding indicates that workers may select themselves into certain types of occupations depending on exogenous factors that directly influence health. Observed health differences across occupational classes should therefore not be interpreted as evidence of a causal effect of occupation on health.

Second, health-related selection into occupation is not only exogenously determined by endowments and shocks: individuals choose their levels of health investment. Health investment may be correlated with occupational choice because (i) exogenous factors influence both health and occupational choice and (ii) workers may choose to offset occupation-related health damage by investing in health (e.g., a bricklayer may seek physiotherapeutic treatment for his back pain, or a manager may take yoga classes to help handle psychological stress).

Third, the relationship between work and health may change over the life cycle for three reasons. First, as equation 6 illustrates, the expected wage returns on health investment decrease as the individual approaches retirement age, which implies that individuals have fewer incentives to offset occupational damage to health by medical investment.[5] Second, $\boldsymbol{\gamma_o}$ may change over the lifetime, for example if health at older ages is more susceptible to wear and tear at the workplace. Third, the marginal effect of health repair may decrease with age to such an extent that full health repair is no longer feasible at older ages.[6]

In sum, the theory imposes the following conditions on our empirical identification strategy: it should (i) account for factors that can influence selection into type of occupation and may also be related to health, (ii) allow for behavioural adjustments that affect health may coincide with occupational choice, and (iii) accommodate the changing relationship between occupation and health over the life cycle. This requires individual-level data that includes information on health, observes occupational stressors for working individuals, and repeated measures of these variables over an extensive period of time.

3 THE GERMAN SOCIOECONOMIC PANEL

The German Socioeconomic Panel (SOEP) is a representative longitudinal household survey that started in 1984. We use data from the 29 subsequent annual waves. Respondents are followed over multiple waves, but the panel is unbalanced as many respondents enter the sample after 1984 or leave the sample before 2012. The sample is restricted to 222,726 person-wave observations for which we

[5] However, a model that endogenizes length of life as a function of health can explain an increase in medical investment at older ages.

[6] Our model does not incorporate real-world labor market rigidities, but such rigidities may also prevent individuals from switching occupations at older ages to optimize their exposure to occupational stressors.

observe occupation in the previous year, educational attainment, and health in the previous and in the current year. It includes individuals between 16 and 65 years old and employed in the previous period. Sample sizes per wave range between 4,702 in 1989 and 11,798 in 2003. Figure 1 shows that we observe 31,216 individuals for at least one period and that 10,577 individuals are observed for at least nine periods, a restricted sample that we will use in a robustness check.

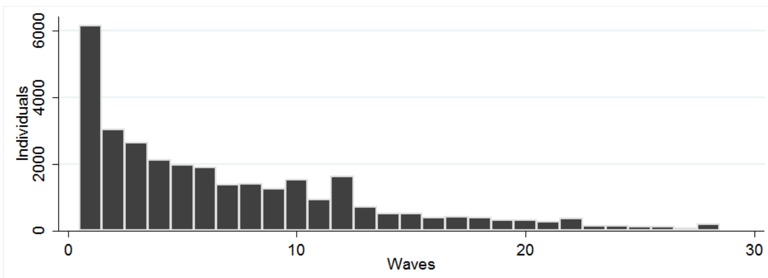


Figure 1: Number of individuals by number of observed waves.

Each bar shows the number of individuals by the total number of (not necessarily consecutive) observed waves in which the individual was employed in the previous period and between 16 and 65 years old. Source: SOEP.

3.1 *Health*

The SOEP has two general health measures: We use health satisfaction (HSAT), measured on an integer scale from 0 to 10, as our main outcome variable. Self-assessed health (SAH), measured on a five-point scale ranging from bad to very good, is used for a robustness check because it was only included since 1992 (and missing in 1993).

3.2 *Occupation*

Occupational titles were coded into the International Standard Classification of Occupations of the OECD (ISCO-88). This gives us 306 occupational titles that were grouped into nine major occupational groups ranked by the OECD classifications, white-collar workers include legislators, senior officials, managers, professionals, technicians, associate professionals, and clerks. We define blue-collar workers as service workers and shop and market sales workers, skilled agricultural and fishery workers, craft and related trades workers, plant and machine operators, assemblers, and workers in elementary occupations. These definitions are consistent with the distinction between manual and non-manual work of Case & Deaton (2005), but the blue-/white-collar terminology better reflects the fact that these occupations differ both in terms of physical strain and psychosocial demands. This classification gives a total of 119,456 person-wave observations for white-collar occupations and 103,270 observations for blue-collar occupations.

Figure 2 shows that on average, blue-collar workers report better health at younger ages, whereas the opposite is true after the age of 28. HSAT decreases for both blue-collar and white-collar workers over most of the age range but increases after the age of 57. One should keep in mind that these patterns only reflect the HSAT ratings of those who are employed. At older ages, unhealthy workers exit out of employment, whereas healthy workers remain employed which explains the upward slope after the age of 60. Consistent with Case & Deaton (2005), we find that predicted health decline associated with age in the pooled sample is much stronger among blue-collar than white-collar workers. This begs the question why we observe this.

Panel A of table I shows that on average blue-collar workers report worse health

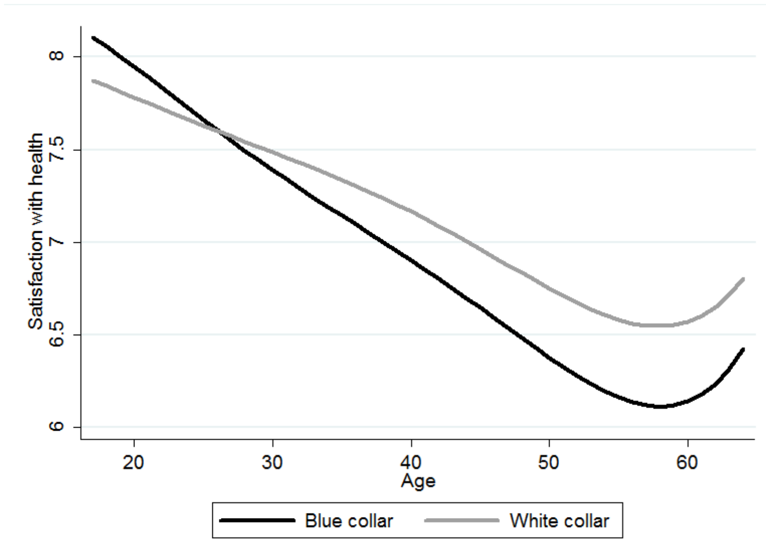


Figure 2: Health for blue- and white-collar workers.

Predicted satisfaction with health for blue- and white-collar workers over the life cycle. Source: SOEP.

(6.86) than white-collar workers (7.01).[7] Blue-collar workers are slightly younger and less likely to be female, and have an average of three years of schooling less than white-collar workers.

Panel B of table I shows descriptive statistics for a restricted sample that we will use in a robustness check. The average age in the restricted sample of individuals who were observed in at least nine waves is approximately one year higher than in the full sample. Average health and the proportion of women are slightly lower. The full and restricted samples are similar in terms of education and (blue-collar) employment.

[7]Health worsens from the top to the bottom of the OECD occupational ladder: 23 percent of legislators, senior officials and managers rate their health with a five or less, as opposed to 31 percent of elementary workers, and 49 percent of legislators, senior officials and managers rate their health with at least an eight, as opposed to 42 percent of elementary workers. This pattern is monotonic across the nine ranked major OECD occupational groups.

Table I: Summary statistics for the German Socioeconomic Panel.

	HSAT	Age	Female	Schooling	Observations
A. Baseline sample					
All workers	6.94	41.39	.45	12.10	222,726
	(2.05)	(11.47)		(2.71)	
White collar	7.01	42.51	.51	13.34	119,456
	(1.99)	(11.07)		(2.83)	
Blue collar	6.86	40.10	.35	10.66	103,270
	(2.11)	(11.78)		(1.65)	
B. Individuals who were employed in at least nine annual waves					
All workers	6.92	42.16	.43	12.14	151,752
	(2.00)	(10.45)		(2.70)	
White collar	6.98	43.07	.51	13.34	68,585
	(1.96)	(10.12)		(2.82)	
Blue collar	6.85	41.05	.33	10.68	83,167
	(2.05)	(10.74)		(1.61)	

HSAT, age, female proportion, years of schooling and monthly labor earnings in the German Socioeconomic Panel. Each wave is viewed as a separate observation. Standard deviations are in parentheses. Source: SOEP.

3.3 *Occupational stressors*

While the distinction between blue-collar and white-collar occupations helps us to characterize health differences across broad occupational groups, it does not allow us to identify which occupational stressors associated with blue-collar occupations matter the most. In addition to *physical strain*, we follow the seminal work by Karasek (1979) and distinguish between two psychosocial stressors that could affect health. First, *job control* is defined by Karasek as decision authority and intellectual discretion: the individual's potential control over his tasks and his conduct during the working day. Second, *psychosocial workload* refers to the degree to which someone is required to work very fast, hard, or to accomplish large amounts of work and whether he or she is short of time.[8]

We have linked information on occupational stressors from the German Qualification and Career Survey (GQCS) to 306 different ISCO 88 occupational titles in the SOEP.[9] In the GQCS individuals are asked about exposure to occupational stressors and can answer with 'frequently', 'sometimes', 'rarely', or 'never' (Hall et al. 2010). Each occupational title is assigned three measures of occupational stress, and each of these measures is constructed on the basis of one variable to allow for an intuitive interpretation of its coefficient. Robustness checks with composite measure of occupational stressors yield similar results.

First, *physical strain* is measured by the degree of heavy lifting that is required for a job. This variable is constructed as an exposure score: the percentage of individuals in each of the 306 occupations who reported sometimes or frequently carrying heavy stocks of more than 20 kg (men) or 10 kg (women) in the GQCS. 35 percent of all individuals surveyed in the GQCS report that they sometimes or frequently carry heavy stocks. Subsequently, each working individual in the SOEP was assigned the exposure score in the GQCS that pertains to his current occupation. Second, we measure *job control* as the percentage of workers in a given occupation who reported in GQCS that their work was frequently or sometimes stipulated in the minutest details. 46 percent of respondents reported that this was frequently or sometimes the case. Third, we measure *psychosocial workload* as the percentage of respondents in a particular occupation that frequently work

[8] Karasek additively combines four measures of decision authority and four measures of intellectual discretion (including a measure of the required skill level) into an aggregate decision latitude scale. He constructs a job demands scale by combining seven measures on not having enough time and working hard and fast.

[9] 606 person-year observations were dropped from the sample because 34 of the total of 306 occupational titles in the SOEP could not be linked to the GQCS.

Table II: Occupational stressors across the major ISCO occupational groups.

	High physical strain	Low job control	High psychosocial workload	Observations
Legislators, senior officials and managers	30	28	69	15,263
Professionals	13	29	62	40,079
Technicians and associate professionals	31	46	57	55,650
Clerks	19	56	51	30,819
Service workers and shop/market sales workers	47	50	44	28,269
Skilled agricultural and fishery workers	79	44	44	3,523
Craft and related workers	65	59	60	46,960
Plant and machine operators and assemblers	54	65	52	22,892
Elementary occupations	55	49	36	14,439

The numbers reflect average percentages for exposure to high physical strain, low job control and high psychosocial workload aggregated by major ISCO 88 occupational group, based on measures of occupational stress for each of the 306 observed ISCO 88 occupational codes. The number of observations refers to the number of person-wave observations in our sample and standard errors are reported in parentheses. White-collar occupations are above the dashed line, and blue-collar occupations are below the dashed line. Source: SOEP, GQCS.

under great deadline pressure, which is what 56 percent of all respondents in the GQCS reported. The cutoff value for each of the occupational stressors was determined by taking the value that is closest to the median. Varying the cutoff values did not alter our results.

Although in the analysis we use the mapping of occupational stressors into *all* 306 individual job titles, for illustrative purposes table II shows the exposure of the nine major OECD occupational groups in the SOEP to the three occupational stressors in the GQCS. Two important observations can be made. First, blue-collar occupations are characterized not only by higher physical strain compared with white-collar occupations but also by lower job control. This illustrates the importance of disentangling the health effects of separate occupational stressors. Psychosocial workload is somewhat higher for white-collar occupations.

Second, there is ample variation in occupational characteristics even within the major occupational groups. Even though blue-collar workers are generally more likely to work under more demanding ergonomic conditions and have lower job control compared with their white-collar counterparts, this may not necessarily be the case for many specific occupations. A simple division into blue-collar or white-collar occupations therefore hides the considerable heterogeneity within these groups and the clustering of occupational stressors. We will first look at the blue-collar/white-collar distinction before investigating the effects of the three occupational stressors on health.

4 ESTIMATION OF THE EFFECT OF OCCUPATIONAL STRESSORS ON HEALTH

4.1 *Model specification*

We aim to estimate the structural parameter $\boldsymbol{\gamma}_{\mathbf{o}}$ in equation 2, which refers to the health effects of exposure to occupational stressors $\mathbf{o}$ in the previous year. Note that the one-period lag of the health production function (equation 2), which includes permanent health h_p, the health effects of ageing a, health investment m and shocks η, is:

$$h_{t+j-1} = h_p + \sum\nolimits_{k=2}^{t+j-1} \left(a_k + \phi^{t+j-1-k} (\boldsymbol{\gamma}_{\mathbf{o}}' \mathbf{o}_{k-1} + \gamma_m m_{k-1}^{\theta} + \eta_k) \right) \tag{7}$$

Substituting equation 7 into equation 2, we obtain:

$$h_{t+j} = (1-\phi)\left(h_p + \sum\nolimits_{k=1}^{t+j-1}(a_k)\right) + a_{t+j} + \boldsymbol{\gamma}_{\mathbf{o}}'\mathbf{o}_{t+j-1} + \gamma_m m_{t+j-1}^{\theta} + \phi h_{t+j-1} + \eta_{t+j} \tag{8}$$

Switching to individual notation and demeaning the covariates to eliminate the time-invariant factors, we obtain a fixed effects within estimator:

$$h_{i,t+j} - \bar{h}_i = \phi(h_{i,t+j-1} - \bar{h}_i) + \boldsymbol{\gamma}_{\mathbf{o}}'(\mathbf{o}_{i,t+j-1} - \bar{\mathbf{o}}_i) + \boldsymbol{\delta}'(\mathbf{x}_{i,t+j} - \bar{\mathbf{x}}_i) + \varepsilon_{i,t+j} \tag{9}$$

This specification controls for various sources of selection. First, any unobserved heterogeneity that is constant over time and may be correlated with occupation (such as permanent health h_p in equation 2) is eliminated: $(1-\phi)h_p - (1-\phi)\bar{h}_p = 0$. Moreover, we go beyond the traditional fixed effects estimator by controlling for lagged health. As detailed in the theoretical framework, this purges the occupational effect from all time-varying unobserved shocks, to the extent that the impact of these unobserved shocks decays exponentially at the same rate over time (see 2). This decay rate is the interpretation of the coefficient of the lagged dependent variable in an individual fixed effects model.[10] The coefficient ϕ of the demeaned one-period lag of health can be interpreted as the decay parameter through which occupational choice $\mathbf{o}$, health investment m, and unanticipated shocks η in period t-2 and earlier periods affect current health.

$\mathbf{x}$ is a vector of control variables consisting of a fifth order age polynomial and wave dummies to control for a common time trend. A less flexible linear approxi-

[10] Only time-varying shocks that are correlated with occupational choice and that do not decay at the average rate may lead to a bias in the estimated effects of occupation. We cannot rule out this source of selection.

mation of the age effect would bias our estimates of $\boldsymbol{\gamma_o}$ if health deteriorates more rapidly at older ages, or if workers at older ages would be more or less likely to be exposed to certain occupational stressors.

The error term is $\varepsilon_{i,t+j} = \gamma_m(m^{\theta}_{i,t+j-1} - \bar{m}^{\theta}_i) + \eta_{i,t+j} - \bar{\eta}_i$, which implies two things. First, the ordinary least squares estimator of the coefficient of the lagged dependent variable ϕ is biased because $h_{i,t+j-1}$ is correlated with $\bar{\eta}_i$, and $\bar{h}_i$ is correlated with $\eta_{i,t+j}$. Importantly, however, the estimator is consistent for large T (Nickell 1981, Bond 2002). Plausibly, the 29 waves of the SOEP panel satisfy this criterion and we check the robustness of our results to restricting the sample to individuals who were observed in nine or more waves.

Second, our estimates should be interpreted as the net effect of occupational stressors, including health investment responses to occupational choice. This net effect is the sum of $\boldsymbol{\gamma_o}$ and an additional term resulting from a possible correlation between occupational choice and unobserved contemporaneous health investment. While this may seem restrictive, we argue that this is a relevant parameter of interest for policymakers because it captures both the direct effects of occupation and the indirect effects through health investment responses to occupation.

4.2 *Estimation details*

Our estimates are based on all within-individual deviations of occupational stressors and health from their observed averages. Within our sample, 13,150 out of 22,526 individuals have switched occupations, for a total of 31,727 times. The number of upward and downward switches in terms of physical and psychosocial demands balance out, and a large proportion of older workers are exposed to high occupational stress.[11]

Our sample selects individuals who were working in the previous year. Even for individuals not working in the current period, we can still estimate the effect of occupation in the previous period on their current health. This implies that sample selection on the outcome variable of the type that warrants a Heckman two-step approach is not an issue here. Obviously, our results should be interpreted as the treatment effect on the working population—and not on children, students, the unemployed, the disabled, and retirees—but that seems to be the policy-relevant effect.

One potential issue could be attrition due to mortality or non-response. Health-related attrition—if present—will lead to a bias toward zero of our estimators, if individuals with the highest vulnerability to occupation-related health deterioration are more likely to suffer from attrition. We find that the likelihood of attrition is at most one percent higher for blue-collar workers than for white-collar workers in our sample. Attrition bias is therefore unlikely to be of major importance to our estimates, and if anything should lead to a slight underestimation of the true effect of occupational stressors.

Self-reported health, as measured on a five-point ordinal scale from poor to

[11] 22 percent of individuals working in two subsequent years in their early twenties experienced a year-to-year job switch that resulted in changes in the level of each of the three occupational stressors. Job switches that result in changes in stressors peak at 24 percent between the age of 25 and 26, and year-to-year switches in occupational stressors still amount to 14 percent of all workers who are in their early sixties. For physical strain, there are slightly more year-to-year switches to lower levels (8.96 percent) than to higher levels (8.55 percent). 8.90 percent of switchers move to an occupation with higher job control while 8.63 percent switch to an occupation with lower control. Slightly more workers switch to occupations with a higher workload (8.94 percent) than with a lower workload (8.57 percent). These small differences are stable across all ages. A substantial proportion of workers in their early sixties are exposed to above-median levels of the occupational stressors: 40 percent for physical strain, 37 percent for low job control, 55 percent for high workload.

excellent, has been shown to be a strong predictor of mortality and morbidity (e.g. Idler & Benyamini 1997, Mackenbach et al. 2002). We use HSAT (on a 0-10 integer scale) as a proxy for health, which exhibits more variation than the five-point SAH measure which we use in a robustness check. Ferrer-i Carbonell & Frijters (2004) and Frijters et al. (2005) show that for the variable that measures satisfaction with life on a ten-point scale, assuming ordinality or cardinality makes little difference, such that a linear specification is acceptable. Reporting heterogeneity arising from different subgroups reporting the same objective health status differently (Lindeboom & van Doorslaer 2004) is eliminated by the individual fixed effect, to the extent that reporting heterogeneity is time-invariant.

5 RESULTS

5.1 *Main results*

Table III shows the main results for six different models, where we first present results for a dichotomous indicator for blue-/white-collar occupations (columns 1 to 3) and then for occupation as characterized by three occupational stressors (columns 4 to 6). To understand the order of magnitude of the coefficients, note that the average health deterioration of growing one year older (obtained from an individual fixed effects regression of satisfaction with health on age) is -.0616 (.0008) in our sample.

1. Health effects of blue-collar work. The bivariate association in column 1 between satisfaction with health and blue- or white-collar occupation in the previous year confirms that blue-collar workers are in worse health and that the size of this health gap is similar to the average effect of ageing 29 months, which is a sizable and economically meaningful difference. Column 2 shows the results for the model described by equation 9. Much of the association appears to be driven by health-related selection into blue-collar occupations because the estimate of the effect is -.0487(.0160) compared with -.1483 (.0087) in column 1. When taken at face value, the health effect of exposure to a blue-collar occupation in the previous year is comparable to the average health effect of ageing nine months.

We add an interaction between age and blue-collar work in column 3 to investigate whether the effect of blue-collar employment differs with age. The coefficient in the first row of column 3 refers to the hypothetical effect of blue-collar employment at the age of zero. The coefficients of the interaction term in the second row indicate that blue-collar employment is harmful to health and that this effect increases with age.

2. Health effects of occupational stressors. Column 4 breaks down occupation into three dimensions of occupational stressors: physical strain, job control, and psychosocial workload in the preceding year. As expected, physical strain and low job control are associated with worse health, whereas psychosocial workload is positively associated with health.

From our theoretical model, we expect health-related selection into occupation to partially drive these associations. Column 5 therefore shows estimates of the effects of these three occupational stressors according to the specification in equation 9, which controls for selection into occupation on the basis of time-invariant and time-varying factors. These results imply that approximately 61 percent of the negative association between physical strain and health can be explained by selection. Our point estimate (-.0830) suggests that a one standard deviation increase in the distribution of physical strain (e.g. being employed as a toolmaker

Table III: Results.

	Associations for blue/white collar	FE & LDV for blue/white collar	FE & LDV for blue/white collar and age interactions	Associations for stressors	FE & LDV for stressors	FE & LDV for stressors and age interactions
	(1)	(2)	(3)	(4)	(5)	(6)
Blue-collar at t-1	-.1483***	-.0487**	.1662***			
	(.0087)	(.0160)	(.0456)			
Age × blue collar at t-1			-.0056***			
			(.0011)			
Physical strain at t-1				-.2102***	-.0830***	.4163***
				(.0051)	(.0032)	(.0909)
Job control at t-1				-.0558*	-.0238	.3656**
				(.0289)	(.0434)	(.1411)
Psychosocial workload at t-1				.5978***	.0074	-.1203
				(.0333)	(.0460)	(.15349)
Age × physical strain at t-1						-.0130***
						(.0022)
Age × job control at t-1						-.0098***
						(.0035)
Age × psychosocial workload at t-1						.0014
						(.0037)
Health at t-1		.1122***	.1120***		.1120***	.1117***
		(.0030)	(.0030)		(.0030)	(.0030)
Individual FE, fifth order age polynomial and wave dummies	✗	✓	✓	✗	✓	✓
Observations	222,726	222,726	222,726	222,726	222,726	222,726
R^2	.0013	.5627	.5628	.0026	.5630	.5631

Main results for satisfaction with health. FE refers to fixed effects estimation, and LDV refers to the inclusion of the lagged dependent variable. Panel-robust standard errors are in parentheses. * indicates significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level. Fixed effects specifications are obtained by subtracting individual averages for each regressor. The reference category for columns 1 to 3 is working in a white-collar occupation. All results are similar in terms of size and significance for models with lower-order age polynomials. Intercepts not shown. Source: SOEP, GQCS.

instead of a teacher) leads to a next-year health deterioration that is comparable to ageing 16 months. Similarly, the coefficient of job control drops by 58 percent although it is no longer significant, while the estimated effect of the psychosocial workload in column 5 is close to zero, suggesting that the positive association with health may even be entirely driven by selection.

However, column 6 shows that the effects on health of low job control and handling heavy burdens vary with age. The predicted health deterioration due to a one standard deviation (.26) increase in handling heavy burdens is equal to zero at age 32. At the age of 50 the point estimate of the effect of moving up the distribution of physical strain by one standard deviation is comparable to ageing 12 months: $.26(.4163 - \times .0130 \times 50) = -.0607$. Low job control has a negative effect after age 37: being in a job with a one standard deviation lower job control (e.g. from librarian to secretary) at age 50 leads to a predicted health deterioration comparable to the effect of ageing 4 months: $.17(.3656 - .0098 \times 50) = -.0211$. We conclude that the effects of physical strain and job control are age-dependent. The coefficient of the interaction between psychosocial workload and age is not significantly different from zero, possibly because workload is only important for certain personality types or for a subset of occupations.

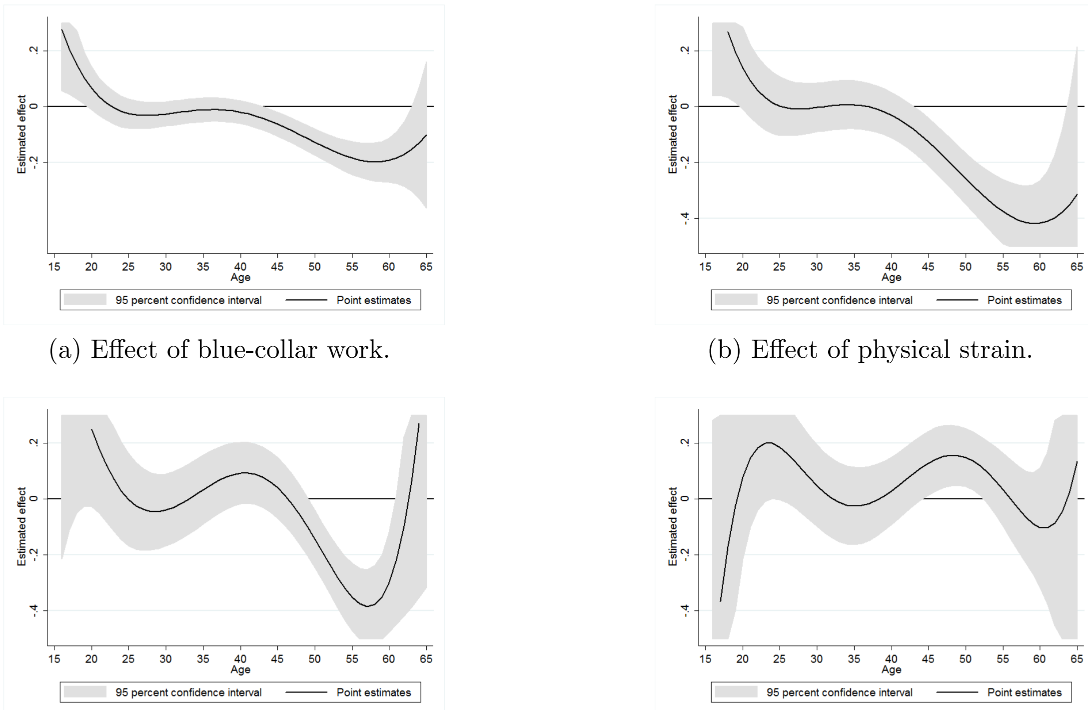
The effect of occupational stressors potentially differs by gender because the type of stressors experienced and the vulnerability to certain stressors is likely to be different for men and women. The direction and statistical significance of the estimates for the subsample of men are similar to separate estimates. However, estimated effects of physical strain and job control for women do not seem to increase with age. A regression without the age interactions gives a significant point estimate of -.1082 (s.e. .0500) for the effect of physical strain on women's health.

5.2 *Cumulative effects*

Cumulative health effects can be obtained from the estimated coefficient of the lagged dependent variable ϕ in equation 9. By assumption, ϕ is the uniform exponential decay rate at which past health investment, occupational stressors, and shocks affect current health in equation 2. The point estimates of ϕ in table III suggest that roughly ten percent of the occupation-related health deterioration in period t-2 persists in period t. Using the point estimates in column 6 of table III, the point estimate of health deterioration at the age of 55 caused by a one standard deviation increase in the physical strain between ages 50 to 54 is $\sum_{k=50}^{54} .1117^{54-k}.26(.4163 - .0130 \times k) = -.0831$, which is comparable to the average health effect of ageing nearly 16 months. Likewise, the point estimate of the effect of working in occupations with a one standard deviation lower degree of job control between the ages of 50 and 54 is -.0311, which is comparable to the effects of ageing 6 months.

5.3 *Vulnerability to occupational stressors over the life cycle*

Health effects may be nonlinear in age. Figure 3 shows the results of two regression models that are similar to the specifications of columns 3 and 6 in table III but include additional interactions between the occupational variables and age up to the fifth power. Panel A shows the estimated next-period effect of blue-collar



(a) Effect of blue-collar work.

(b) Effect of physical strain.

(c) Effect of low job control

(d) Effect of high work load

Figure 3: The effects of occupational stressors over the life cycle.

95 percent confidence intervals of the coefficients of the occupational stressors, computed using the delta method. Panel 3a refers to the coefficient of a binary variable, the other panels refer to the health effects of occupations with 100 instead of 0 percent exposure to each respective stressor. Source: SOEP, GQCS.

versus white-collar occupations at different ages.[12] The estimated next-period negative effect of blue-collar employment is statistically significant from the age of 43 onwards.[13]

Panels B, C, and D of figure 3 refer to estimates obtained from the model in column 6 of table III with added interactions between each of the three occupational stressors and age up to the fifth power. Physical strain has a significant

[12]Confidence intervals are computed using the delta method (Oehlert 1992).

[13]Caution is warranted when interpreting the estimates at the lower and upper end of the age distribution because of the lower number of observations at young and old ages and the polynomial functional form. Less than three percent of observations occur at ages below 20, and only two percent of observations occur at ages over 60.

negative effect on workers aged 42 and older.

Panel C shows that the negative effect of low job control is significant for workers aged 48 years and older. The estimated negative effects in panels A, B, and C are strongest around age 60. Surprisingly, we find a minor positive effect of high psychosocial work load around the age of 50 as shown in panel D.

5.4 *Robustness checks*

We examine the robustness of our main findings to (i) alternative measures of occupational stressors, (ii) using SAH instead of HSAT as our outcome measure, (iii) controlling for education-specific health deterioration by age, (iv) limiting the sample to individuals who are observed in nine or more waves, and (v) two additional ways of specifying our dynamic panel data model.

The results are robust to using other measures of occupational stressors and health. Column 1 in table IV includes composite measures of physical strain, job control and psychosocial workload, and shows similar results in terms of effect sizes of a standard deviation increase in the occupational stressors. We obtain composite measures of the three occupational stressors by averaging percentages of high exposure to (i) physical strain as measured by heavy lifting, working crouched down, and working standing up; (ii) job control as measured by work being stipulated in the minutest details, being allowed to plan and schedule work by oneself, and being able to influence the amount of work one has to do; and (iii) psychosocial workload as measured by working under great deadline pressure, reaching the limits of one's capacities, and working very quick. Furthermore, analyses using the Finnish Job Exposure Matrix instead of the GQCS provides results that are similar in terms of order of magnitude and significance (see Ravesteijn et al. 2014).

SAH—measured on a five-point scale—was only included in the SOEP questionnaire in 1992 and from 1994 onwards. Column 2 of table IV substitutes SAH for HSAT and the coefficient estimates reveal similar signs and sizes to our main findings in column 6 of table III, supporting our conclusion that physical strain and low job control are harmful to health at older ages.

Individuals in different occupations may have different biological ageing rates. We have assumed uniform ageing effects in the preceding analyses. If the health of blue-collar workers declines more rapidly regardless of their occupation, our results overestimate the harmful effects of physical strain. In column 3 of table IV, we allow for different rates of ageing by interacting each of the levels of educational attainment with a fifth-degree age polynomial and our estimates are similar to our findings in table III.

The estimator of the coefficient of the lagged dependent variable is consistent if the number of time periods in the sample goes to infinity. Our sample spans 29 years and is unbalanced because it includes individuals who are observed for fewer waves. We repeat our analysis for a subsample of 10,373 individuals who have been employed for at least nine of the 29 years to counter the downward bias of the estimator of the lagged dependent variable that plagues short panels (Bond 2002). The number of person-wave observations drops from 222,726 in our baseline sample to 151,752 in column 4 of table IV. The coefficients of the (age-interacted) occupational stressors are similar to those in our baseline specification. However, the coefficient of lagged health is now larger, suggesting that past health investment, occupational stress, and health shocks are more persistent than they were in the full-sample analysis. We conclude that our estimates of the effects

Table IV: Robustness.

	Composite stressors	SAH	Control for education-specific ageing trends	Only individuals with $T > 8$	FE	LDV
	(1)	(2)	(3)	(4)	(5)	(6)
Physical strain at t-1	.2693***	.1775***	.4032***	.3473***	.4693***	.1324**
	(.0943)	(.0506)	(.0940)	(.1010)	(.0921)	(.0529)
Job control at t-1	.9661***	.2906***	.3584**	.4439***	.3945***	.3953***
	(.2641)	(.0760)	(.1454)	(.1605)	(.1427)	(.0918)
Psychosocial workload at t-1	.2355	.0755	.1125	.4175**	.1564	.1326
	(.2152)	(.0816)	(.1563)	(.1787)	(.1553)	(.1032)
Age × physical strain at t-1	-.0088***	-.0046***	-.0127***	-.0110***	-.0145***	-.0071***
	(.0023)	(.0012)	(.0023)	(.0024)	(.0022)	(.0013)
Age × job control at t-1	-.0028***	-.0059***	-.0095***	-.0111***	-.0107***	-.0130***
	(.0065)	(.0018)	(.0036)	(.0038)	(.0035)	(.0022)
Age × psychosocial workload at t-1	.0044	-.0027	.0014	-.0072*	-.0020	.0035
	(.0052)	(.0020)	(.0038)	(.0043)	(.0038)	(.0024)
Health at t-1	.1117***	.0723***	.1110***	.1613***		.5464***
	(.0030)	(.0034)	(.0030)	(.0034)		(.0021)
Fifth order age polynomial and wave dummies	✓	✓	✓	✓	✓	✓
Fifth order age polynomial interacted with education	✗	✗	✓	✗	✗	✗
Individual FE	✓	✓	✓	✓	✓	✗
Education and gender	✗	✗	✗	✗	✗	✓
Observations	222,726	162,595	222,726	151,752	222,726	222,726
R^2	.5631	.5906	.5635	.5220	.5576	.3389

Robustness checks for satisfaction with health. FE refers to fixed effects estimation, and LDV refers to the inclusion of the lagged dependent variable. Panel-robust standard errors are in parentheses. * indicates significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level. Fixed effects specifications are obtained by subtracting individual averages for each regressor. The fourth column refers to sample of individuals who are observed in at least nine—not necessarily consecutive—waves. Intercepts not shown. Source: SOEP, GQCS.

of occupational stressors are robust across specifications but that an analysis of the full sample leads to underestimation of the coefficient of lagged health. We may have underestimated the cumulative effects of occupational history by underestimating ϕ, and the predictions regarding the cumulative effects provide—in absolute terms—a lower bound on the health effects, which indicates that the true health effects may in fact be even larger.

Angrist & Pischke (2009) have voiced concerns about the violation of strict exogeneity in fixed effects dynamic models, particularly by utilising short panels. They propose checking robustness by separately estimating both a fixed effects and a lagged dependent variable model. Column 5 of table IV presents results from a fixed effects model without a lagged dependent variable.[14] The point estimates in column 5 suggest a somewhat stronger effect of physical strain and low job control at older ages than the baseline specification. However, these estimates may be the result of a bias caused by past events that affected health and occupational choice that are not accounted for by the lagged dependent variable, which is omitted in this specification.

In a model in which we control for a lagged dependent variable, but not for individual-specific fixed effects, the estimator of the decay parameter ϕ in equation 9 is biased toward one because h_{t-1} contains h_p (see equation 7), which has a coefficient of one and no longer drops out if we do not subtract $\bar{h}$. We can therefore no longer distinguish between the elements in h_{t-1} that are transitory and the

[14]With respect to equation 9, the error term would now include the deviations of the effects of health investment, occupational stressors, and health shocks before period t-1 from their individual averages. If a past health shock would have a negative effect on current health and lead to higher occupational stress in the previous period, we would overestimate the effect of occupational stressors because this situation leads to additional correlation between $\mathbf{o}$ and the error term.

elements that are constant over time, which explains the bias of the estimator of ϕ toward one. In this specification, we therefore overestimate the impact of past events on current health, and we only partly control for unobserved time-invariant heterogeneity.[15] To proxy for time-invariant unobserved factors otherwise picked up by the fixed effect, we control for years of schooling and gender. Our estimates are now mostly driven by variation between individuals. The coefficients of the interaction between age and occupational stressors in column 6 of table IV are similar to our earlier results. Note that the coefficients in the first three rows of the table lack a clear interpretation because of the age interaction and should therefore not be interpreted. Overall, our main conclusions do not change when estimating models that include either individual-specific fixed effects or a lagged dependent variable, which is reassuring.

Other methods have been proposed to consistently estimate $\boldsymbol{\gamma_o}$ in equation 9 in short panels, of which the so-called Arellano-Bond estimator (Arellano & Bover 1995, Blundell & Bond 1998) is the most prominent. The Arellano-Bond estimator is based on the first-difference estimator. The most important assumption is that second and further lags of health are uncorrelated with the first differences of the error term and can be used as instrumental variables for $h_{t-1} - h_{t-2}$. The Arellano-Bond test for autocorrelation rejects this assumption in our case, which is not surprising because using lagged values as instruments is difficult to justify in the case of health: chronic illnesses or the introduction of a new medical drug may progressively affect health over time, which leads to second- or higher-order serial correlation in the differenced error term and violation of the exogeneity

[15]By not subtracting averages in equation 9, the error term now includes $(1 - \phi)h_p$, which may be correlated with lagged health and occupational characteristics.

assumption. In attempting to overcome this problem, more lags of the regressors were included in the model, and further lags of regressors and instruments were used to purge the error term from autocorrelation. However, we still find higher-order autocorrelation in these models, rejecting the validity of the instruments.[16]

6 CONCLUSION

The strong association between occupation and health is widely documented. Our results confirm that German blue-collar workers report worse health than white-collar workers, and that the size of this health gap is similar to the effect of ageing 29 months. However, because of various sources of selection into occupation, the association does not necessarily reflect the causal effect of occupation on health.

In this paper, we make three contributions. First, by proposing a dynamic theoretical model as the foundation of our empirical specification, we highlight the various sources of selection into occupation, and we make explicit under which conditions the coefficient of occupational hazards can be interpreted as causal.

Second, we estimate the empirical equation deriving from the theory using a detailed German longitudinal dataset over many time periods (29 years). In this equation, we account for the sources of selection: (i) unobserved time-invariant

[16] Limiting the number of waves can give us the false illusion that serial correlation of the error term is not a problem simply because of the low power of the test. Blundell & Bond (1998) and Michaud & Van Soest (2008) use short panels of six waves and "use up" even more waves due to the inclusion of lagged values of the dependent variable. The autocorrelation tests in these studies do not reject the assumption of no autocorrelation in the error term, which may be the result of limited test power based on the small number of waves. If we include one- and two-period lags of the dependent variable, we find no second-order autocorrelation. However, we find autocorrelation of the third-order, which still violates the Arellano-Bond assumptions. Including third or fourth lags seems to shift the order of autocorrelation downward rather than to solve the problem. The Sargan test may not be informative because it assumes that at least one instrument is exogenous, which is an assumption we are not willing to make.

variables due to the inclusion of individual fixed effects, (ii) time-varying observed variables such as age and wave dummies, and (iii) time-varying unobserved shocks that exponentially die out through the inclusion of the lagged dependent variable. In doing so, we are able to quantify a lower bound on the selection effect, and an upper bound on the causal effect of occupation on health.[17] Moreover, our results generalize across the entire labor force, which is in contrast to local effect estimates based on a particular reform that affected only part of the employed population.

Third we estimate separate health effects of physical strain and Karasek's two dimensions of psychosocial occupational stress from the German Qualification and Career Survey. However, because we do not observe individual levels of health investment, we are unable to disentangle the effects of such occupational stressors and any health investment made in response to occupational choice. This is a policy-relevant effect, since it can be interpreted as the sum of the direct effect of occupation and the indirect effect of any behavioral response to occupational choice.

Our main finding is that selection of individuals with poor health (prospects) into occupations with heavy physical demands and low job control, accounts for 61 and 58 percent, respectively, of the observed association with health. The association between psychosocial workload and health seems to be entirely driven by selection. While we cannot rule out other sources of selection, we interpret the residual, conditional differences in health as deriving from the occupational

[17]There could be other sources of selection, like time-varying shocks that are related to occupational choice that do not exponentially decay over time or that have different decay rates. While we cannot rule out that some time-varying unobserved factors operate in the opposite direction, the sign of the bias when excluding fixed effects or the lagged dependent variable renders the case that the true selection effect is larger more plausible.

characteristics itself. Using this interpretation, we find that both high physical occupational demands and low job control have negative effects on health and that these effects increase with age. The effect of exposure to a one standard deviation increase in the degree of handling heavy burdens (e.g., working as a toolmaker instead of a teacher) between age 50 to 54 is comparable to ageing 16 months. Low job control is harmful to health but only after age 48. The effect of exposure to a one standard deviation decrease in the degree of job control (e.g., working as a secretary instead of a librarian) between age 50 to 54 is comparable to ageing 6 months.

Occupational health and safety policies, career development programs, and retirement policies should be based on the knowledge that selection into occupation explains a substantial fraction of health disparities across occupational groups. It is important to emphasize that ignoring this selection will lead to serious overestimation of the benefits of improving occupational characteristics. Our results cannot rule out, and even strongly suggest, that exposure to physical strain and low job control is harmful to health at older ages. That would mean that successfully shielding workers—especially older workers —from such conditions will prevent accelerated health deterioration and therefore also illness-related absenteeism and labor force exit due to work disability. Our findings therefore have great relevance for the many OECD countries that are aiming to extend working life careers in order to keep disability and pension systems sustainable.

References

Altonji, J. G., Elder, T. E. & 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. & Pischke, J.-S. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

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

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

Bond, S. R. (2002), 'Dynamic panel data models: a guide to micro data methods and practice', *Portuguese Economic Journal* **1**(2), 141–162.

Bongers, P. M., Hulshof, C. T. J., Dijkstra, L., Boshuizen, H. C., Groenhout, H. J. M. & Valken, E. (1990), 'Back pain and exposure to whole body vibration in helicopter pilots', *Ergonomics* **33**(8), 1007–1026.

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

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

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

Ferrer-i Carbonell, A. & Frijters, P. (2004), 'How important is methodology for the estimates of the determinants of happiness?*', *The Economic Journal* **114**(497), 641–659.

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

Frijters, P., Haisken-DeNew, J. P. & Shields, M. A. (2005), 'The causal effect of income on health: Evidence from german reunification', *Journal of Health Economics* **24**(5), 997 – 1017.

Galama, T. & van Kippersluis, H. (2010), A theory of socioeconomic disparities in health over the life cycle, Working Paper 2010-079/3, Tinbergen Institute.

Goodman, E. (1999), 'The role of socioeconomic status gradients in explaining differences in us adolescents' health.', *American Journal of Public Health* **89**(10), 1522–1528.

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

Hall, A., Braun, U., Herget, H., Krekel, E. M., Leppelmeier, I., Schade, H.-J., Troltsch, K. & Ulrich, J.-G. (2010), 'Die bibb/baua-erwerbstätigenbefragung 2006', *Arbeit und Beruf im Wandel, Erwerb und Verwertung beruflicher Qualifikationen–Abschlussbericht (Bundesinstitut für Berufsbildung) Bonn* .

Idler, E. L. & 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.

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. & Gallo, W. T. (2012), 'The impact of early occupational choice on health behaviors', *Review of Economics of the Household* **10**, 1–34.

Kunst, A. E., Groenhof, F., Andersen, O., Borgan, J. K., Costa, G., Desplanques, G., Filakti, H., Giraldes, M. d. R., Faggiano, F., Harding, S. et al. (1999), 'Occupational class and ischemic heart disease mortality in the united states and 11 european countries.', *American Journal of Public Health* **89**(1), 47–53.

Lewbel, A. (2012), 'Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models', *Journal of Business & Economic Statistics* **30**(1), 67–80.

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

Mackenbach, J. P., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., Reid, A., Hemström, Ö., Valkonen, T. & Kunst, A. E. (2003), 'Widening socioeconomic inequalities in mortality in six western european countries', *International Journal of Epidemiology* **32**(5), 830–837.

Mackenbach, J., Simon, J., Looman, C. & Joung, I. (2002), 'Self-assessed health and mortality: could psychosocial factors explain the association?', *International Journal of Epidemiology* **31**(6), 1162–1168.

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

Michaud, P.-C. & Van Soest, A. (2008), 'Health and wealth of elderly couples: Causality tests using dynamic panel data models', *Journal of Health Economics* **27**(5), 1312–1325.

Morefield, B., Ribar, D. C. & Ruhm, C. J. (2012), 'Occupational status and health transitions', *The BE Journal of Economic Analysis & Policy* **11**(3), 1–29.

Nickell, S. (1981), 'Biases in dynamic models with fixed effects', *Econometrica: Journal of the Econometric Society* **49**(6), 1417–1426.

Oehlert, G. W. (1992), 'A note on the delta method', *The American Statistician* **46**(1), pp. 27–29.

Ravesteijn, B., van Kippersluis, H. & van Doorslaer, E. (2013), The contribution of occupation to health inequality, *in* 'Research on economic inequality', Vol. 21, Emerald Group Publishing Limited, pp. 311–332.

Ravesteijn, B., Van Kippersluis, H. & Van Doorslaer, E. (2014), The wear and tear on health: What is the role of occupation?, SOEPpaper 15256.

Smith, R. S. (1974), 'The feasibility of an" injury tax" approach to occupational safety', *Law and Contemporary Problems* **38**(4), 730–744.

Viscusi, W. K. (1978), 'Labor market valuations of life and limb: Empirical evidence and policy implications', *Public Policy* **26**(3), 359–386.