DE GRUYTER BE J. Econ. Anal. Policy 2016; 16(2): 577–617

Nabanita Datta Gupta*, Daniel Lau and Dario Pozzoli

The Impact of Education and Occupation on Temporary and Permanent Work Incapacity

DOI 10.1515/bejeap-2015-0055
Published online March 15, 2016

Abstract: This paper investigates whether education and working in a physically demanding job causally impact temporary work incapacity (TWI), i. e. sickness absence and permanent work incapacity (PWI), i. e. the inflow to disability via sickness absence. Our contribution is to allow for endogeneity of both education and occupation by estimating a quasi-maximum-likelihood discrete factor model. Data on sickness absence and disability spells for the population of older workers come from the Danish administrative registers for 1998–2002. We generally find causal effects of both education and occupation on TWI. Once we condition on temporary incapacity, we find again a causal effect of education on PWI, but no effect of occupation. Our results confirm that workers in physically demanding jobs are broken down by their work over time (women more than men) but only in terms of TWI.

Keywords: work incapacity, education, occupation, factor analysis, discrete factor model
JEL Classification: J18, I12, I20, C33, C35

1 Introduction

Despite the impressive medical breakthroughs that have improved aggregate health over the last decades, large numbers of people of working age persistently leave the workforce and rely on health-related income support such as sickness absence and disability benefits (Autor and Duggan 2006; Pettersson-Lidbom and Thoursie 2013). This trend is witnessed in virtually all Organisation for Economic Co-operation and Development (OECD) countries today, including Denmark, and the lost work effort of these individuals represent a great cost to

***Corresponding author: Nabanita Datta Gupta,** Department of Economics and Business Economics, Aarhus University, Fuglesangs Alle 4, DK-8210 Aarhus, Denmark, E-mail: ndg@econ.au.dk
http://orcid.org/0000-0001-7904-445X
Daniel Lau, Department of Economics, Cornell University, Ithaca, NY, USA
Dario Pozzoli, Department of Economics, Copenhagen Business School and IZA, Copenhagen, Denmark

society.[1] Health-related exit from the labor market thus continues to present an obstacle to raising labor force participation rates and keeping public expenditures under control.

Previous, mainly descriptive, research shows a significant *educational* gradient in the onset of long-term sickness absence and disability (see e. g. Fried and Guralnik 1997; Freedman and Martin 1999; Cutler and Lleras-Muney 2007) and the few studies that estimate causal effects of education on self-reported disability status also find significant negative effects (Kemptner et al. 2011; Oreopoulous 2006). An *occupational* gradient in health is also well documented (Case and Deaton 2005; Fletcher, Sindelar, and Yamaguchi 2011; Morefield, Ribar, and Ruhm 2012; for Denmark, see Christensen et al. 2008). Yet no study has tried to simultaneously uncover the relationships between these variables and health deterioration of older workers.

The aim of this paper is to empirically model and estimate the relationships between education, working in a physically demanding job and individual health-related exit from the labor market. Our study adds to the literature in this field in several respects. Our main contribution is to distinguish between whether lower labor market attrition among elderly educated workers is due to less wear-and-tear and fewer accidents on their jobs, or whether it is due to the protective effect of education leading to greater health knowledge, more caution and greater efficiency in health investments, the so-called *causal* effect of education. This is a valuable addition to current econometric studies that estimate either the causal effect of education on disability or estimate the effect of occupational status or cumulative effect of job characteristics on self-assessed health or disability in a panel framework (Oreopoulous 2006; Kemptner et al. 2011; Case and Deaton 2005; Fletcher, Sindelar, and Yamaguchi 2011, Morefield, Ribar, and Ruhm 2012). We are not aware of any previous study that attempts a simultaneous estimation of the causal effects of both education and occupational demands on disability exit, exploiting data across individuals and time. For the purpose of identifying the appropriate policy response to curb the high disability exit in many countries, we need to know what types of investments are needed, i. e. is it sufficient to raise the educational level of the population or do we need to improve working conditions as well, and if both effects are important, what are their relative magnitudes?

---

1 On average, expenditures on disability and sickness benefits account for about 2 % of GDP in OECD countries, about 2.5 times what is spent on unemployment benefits. This ranges from 0.4 % for Canada to 5 % in Norway. In the United States, it is 1.4 % of GDP, while in Denmark 3.1 % of GDP is spent annually (OECD 2007).

By applying clear identification strategies based on a major educational reform and on pre-sample characteristics of local labor supply, we aim to uncover causal impacts of both education and working in a physically demanding job on sickness absence and disability. Specifically following Arendt (2005, 2008), we exploit the change in urban–rural differences in education due to the 1958 reform to identify the education effect on work incapacity. Whereas as an instrument for having a physically demanding job, we use a historic measure of the average physical demands of the occupation (separately for male and female) of all workers in an individual's current local labor market. These causal effects are estimated by applying a quasi-maximum-likelihood (quasi-ML) estimator with discrete factor approximations (Mroz 1999; Picone et al. 2003; Arendt 2008). This econometric approach compares favorably with instrumental variable (IV)-type estimators in terms of precision and bias, especially in the presence of weak instruments (Mroz 1999).

Since the institutional setup is one in which sickness absence is a precondition for being considered for disability pension, we model the pathway of disability exit going through sickness absence and we also explore similarities and differences in the factors affecting behavior for these two processes.

Furthermore, to construct a highly reliable measure of the physical demands of jobs, we use very detailed skill information in the Dictionary of Occupational Titles (DOT) data in conjunction with the information provided in the US census occupation codes and the Danish register's occupation codes.[2]

A further advantage of using detailed administrative data is that, whereas much of the previous research on disability is based on error-prone self-reported measures, we access more objective register-based indicators.[3] On the other hand, of course subjective measures may capture latent health factors (for instance, how ill the individual actually feels, how well they are able to function with their ailment) which the cruder objective measures cannot do.

Our analyses are carried out in the homogeneous welfare-state setting of Denmark. Denmark has disability-related exit, universal health-care access and coverage and generous health benefits provision. Thus, any socioeconomic gradient in health in Denmark is less likely to be due to differential access,

2 In a sensitivity analysis, we try an alternative measure of physical demands using the DWECS data which has less detailed skill measures compared to the DOT.

3 The reliability of self-reported measures has in fact been questioned because of reporting heterogeneity (Bago d'Uva et al. 2008; Kaptyen, Smith, and van Soest 2007; Datta Gupta, Kristensen, and Pozzoli 2010) as individuals may have an incentive to overplay their own health problems for the purpose of gaining disability benefits.

allowing for cleaner identification of the effects of education and occupation on health-related exit from the labor market.

The comprehensiveness of the Danish register data enable us to estimate a recursive model of work incapacity, education and occupation with correlated cross-equation errors and unobserved heterogeneity. Our findings show that both education and occupation have causal effects on work incapacity. Our results for temporary work incapacity (TWI) (sickness absence) show that education generally reduces sickness absence, although more strongly for women than men. At the same time, we find an independent and stronger role for occupation, but which also differs by gender. More specifically, a stronger causal effect of the physical demands of the occupation increasing TWI is seen for women than men, despite the average probability of sickness absence being fairly similar by gender. The effects of education and occupation on permanent work incapacity (PWI) or disability, on the other hand, are imprecisely estimated, conditional on TWI. These results may be relevant for other welfare state countries and can inform policy in these cases of the ways in which to stem the inflow into sickness absence and disability, respectively, temporary and permanent incapacity.

The paper is organized as follows. Section 2 reviews the related literature, derives the main theoretical hypotheses and describes the institutional setting. Section 3 introduces the data set and presents main descriptive statistics. Section 4 provides details on the empirical strategy. Section 5 explains the results of our empirical analysis, and Section 6 concludes.

2 Background

In the first part of this section we explore the theoretical channels linking education and the physical demands of the job to health. In the second part, we present information on the institutional setting.

2.1 Theoretical Foundation

A possible theoretical basis for a socioeconomic gradient in health arising from both education and the nature of jobs is provided in the modified version of Grossman's (1972) intertemporal model of health (Muurinen 1982; Muurinen and Le Grand 1985; Case and Deaton 2005). According to the Grossman's original theoretical setup, health deteriorates over time but it is maintained by education through various channels. First, education raises the efficiency of inputs needed to restore health by following directions on medicine packages (Goldman and Lakdawalla 2001), i. e. education increases productive efficiency. Second, education gives people greater health knowledge (Kenkel 1991) and enables them to choose the right input mix, i. e. enhancing allocative efficiency. A third channel proposed by Fuchs (1982) and Becker and Mulligan (1997) is that education can change people's rate of time or risk preferences and thereby lead them to invest in better health. A final pathway between education and health is that educated individuals can avoid physically demanding jobs and this reduces the rate at which their health deteriorates (Muurinen 1982; Muurinen and Le Grand 1985; Case and Deaton 2005). That is, health is also affected by the extent to which "health capital" is used in consumption and in work. Some consumption activities are harder on the body than others and manual or physically demanding work is harder on the body than nonmanual work. Allowing for the health deterioration rate to depend on physical effort provides an explanation for why health may deteriorate faster with age among the low-educated or manual workers. In our setup, we are able to isolate the independent causal effects of both education and occupation on health (disability).

In a health-care system with universal access and insurance such as the Danish one, we would expect financial constraints to be less binding for the purpose of meeting health needs. Maurer (2007) analyzes comparable cross-national data from the SHARE on 10 European countries. Estimating health-care utilization using flexible semi- and nonparametric methods, there is no evidence of a socioeconomic gradient in health-care usage in Denmark. A characteristic of the Danish labor market that works against finding work-related health deterioration is a relatively high degree of employer accommodation paid for by the municipalities. Danish workplaces are reportedly among the most accommodating in Europe, especially after the twin pillars of corporate social responsibility and activation of marginalized groups were incorporated starting from the mid-1990s (Bengtsson 2007). Larsen (2006) reports that almost half of the private and public workplaces in Denmark with at least 10 employees and with a least one employee above the age of 50 make an effort to retain older workers. These institutional features imply that any work-related health deterioration in manual jobs found in Denmark is probably only a lower bound of what can be found in other settings.

At the same time, prior research has convincingly shown that sickness absence and disability can be forms of moral hazard behavior in settings where agents are fully insured, e. g. in welfare states with generous benefit provision for health-related problems (Johansson and Palme 2005; Larsson 2006; Hofmann 2013). Compensation-seeking behavior can be of two types: ex ante moral hazard, implying individuals do not undertake sufficient measures or investments to prevent diseases/disability or ex post moral hazard, according to which individuals claim compensation for hard to verify diseases, overplay their symptoms or even withhold their labor supply during the awarding phase such as continuing to be on sickness absence while waiting for a decision on a disability application (Bolduc et al. 2002). In Denmark, municipalities pay for necessary workplace adaptations, provide employers with wage subsidies and give sick-listed workers a form of social support that facilitates continued work without risk of benefit loss, barring for the first few weeks which are covered by the employer. Educated people and individuals working in nonmanual jobs have lower expected payoffs from disability insurance or sickness benefits, i. e. they would be less prone to compensation-seeking behavior. We refrain from including any measure of the (expected) benefit replacement rates in our models out of concerns for endogeneity, but assume instead that education, age, gender, region and occupation are good proxies for income, and hence replacement rates, in a setting with low-income dispersion. Unmeasured individual propensities to engage in moral hazard constitute an important source of unobserved heterogeneity that needs to be accounted for. Apart from time-invariant moral hazard tendencies, time-varying preferences for compensation-seeking behavior could induce a (unobserved) correlation between investing in education/working in a physically demanding job and exiting via disability, which will also be captured in our framework.

2.2 Institutional Setting and Data

2.2.1 The 1958 School Reform

Our identification strategy in the case of education is based on a Danish school reform in 1958 that abolished a division of schooling that occurred after the fifth grade and affected attendance particularly beyond the seventh grade, i. e. beyond primary school. The reform also eliminated a distinction between rural and urban elementary schools wherein only urban schools could offer classes from 8th to 10th grades. Prior to the reform, schooling in the sixth and seventh grades was determined on the basis of a test conducted at the end of fifth grade, and the curriculum level and the number of lessons were lower in rural schools than urban schools. These differences were abolished in 1958, which we use to identify causal effects of increasing years of education (and to some extent also quality) for individuals born in rural areas from 1946 onwards, compared to their urban counterparts. The impact specific to schools in rural areas came about because a distinction between rural and city elementary schools was eliminated. The city schools were placed in 87 cities of reasonable size and historical or geographical significance. Prior to the reform the curriculum level and the number of lessons were lower in the rural schools, and only city schools could offer classes for 8th to 10th form (preparing for upper secondary schooling).[4] It is a common view that the 1958 reform helped to alleviate barriers to further education that existed prior to 1958, which were pronounced for children from less educated backgrounds or living outside areas with city schools (Bryld et al. 1990). Thus, the impact of the reform on individuals is based on their year and area of birth (on the parish level). Whether the area had an urban or rural school is specified in the same way as in Arendt (2008). And as in Arendt (2008), we use a binary variable denoting education beyond primary schooling as the endogenous covariate. Measuring the effect on beyond primary education exploits changes in the rural–urban difference from before to after the reform.

2.2.2 TWI and PWI in Denmark

Some institutional background is necessary in the case of the outcome measures as well. TWI is defined as sickness absence for more than 8 consecutive weeks. Workers can receive the benefit for up to 52 weeks, but the benefit period may be extended under certain circumstances, e. g. if the worker has an ongoing workers' compensation or disability benefit claim. The sick-listed worker reports the onset of the injury or sickness to the employer who in turn registers the spell with the municipality. Both the employer and municipality can require verification of the condition from the medical authorities.[5] The threshold of 8 weeks is chosen as municipalities in Denmark are obliged to follow up on all cases of sickness benefit within 8 weeks after the first day of work incapacity. This threshold also aligns with other studies using Danish register data

4 Note that following the earlier 1937 middle school reform, rural schools could join forces and provide 6th–9th grades as in the urban areas (the so-called centrale skoler), but because of wartime shortages school consolidation was delayed in the rural areas, so leading up to the 1958 reform, rural educational opportunities still lagged considerably behind those of the urban areas. Still, it is likely that there was selection into educational quantity, so the exogenous variation induced by the 1958 reform could potentially be arising from the lower end of ability/income distribution. Source: http://danmarkshistorien.dk/leksikon-og-kilder/vis/materiale/lov-om-folkeskolen-18-maj-1937/.

5 These rules are dictated by the national sickness absence law and applies uniformly across all occupations. Employers finance their workers' sickness benefits for the first 3 weeks, and public authorities finance the remaining period.

(Christensen et al. 2008, Lund, Labriola, and Christensen 2006) and combined survey register data (Høgelund and Holm 2006 and Høgelund, Holm, and McIntosh 2010). After the first 8 weeks of sickness, the municipality must perform a follow-up assessment every 4th week in complicated cases and every 8th week in uncomplicated cases. The primary goal of the assessments is to restore the sick-listed worker's labor market attachment. The assessments must take place in cooperation with the sick-listed worker and other relevant agents, such as the employer and medical experts. If return to ordinary work is impossible because of permanently reduced working capacity, the municipality may refer the sick-listed worker to a Flexjob, the latter being wage-subsidized jobs with tasks accommodated to the worker's working capacity and with reduced working hours. The criterion for being assigned to such a Flexjob is that there has been a permanent reduction in working capacity as assessed by the county medical examiners, and all other channels of obtaining regular unemployment have been exhausted. Depending on the reduction of the worker's working capacity, the wage subsidy equals either one-half or two-thirds of the minimum wage as stipulated in the relevant collective agreement. If a person with permanently reduced working capacity is incapable of working in a wage-subsidized job, the municipality may award a permanent disability benefit, which is financed entirely by public authorities. Workers assigned wage-subsidized jobs retain these permanently in our sample period, transiting only to full-term disability or old age pension. A recent reform, however, will grant wage-subsidized jobs for a 5-year period only, though renewable. In 2012, 60,000 disabled persons annually were on wage-subsidized jobs but demand exceeded supply, so currently there is wait unemployment of around 30 %.

3 Data Description and Variables

The data sample begins with all individuals born between 1943 and 1950, which are cohorts a few years prior to and after the school reform, and followed over the period from 1998 to 2002. The year 1998 is selected as the first year because reliable information on sickness spells is only available starting from that point on. Prior to this, sickness spells are indistinguishable from maternity leave spells for women. Furthermore, because we are concerned about encountering major changes in the disability policy scheme in 2003, 2002 is the last year in the sample period. Moreover, this period is chosen to allow the chosen cohorts prior and after the school reform to reach a target age range, which is relevant in order to study how education and occupation affects individuals' labor market exit due to sickness or disability. Individuals selected for the sample are required to be employed at least in 1997.[6] Data are taken from administrative registers from Statistics Denmark. These are sourced from the Integrated Database for Labor Market Research (IDA) and include characteristics of individuals such as highest educational attainment, marital status and occupation for each year. Work incapacity data come from another administrative dataset, DREAM (Danish Register-Based Evaluation of Marginalisation), which contains information on weekly receipts of all social transfer payments for all citizens in Denmark since 1991, and separately records sickness absence since 1998.[7]

In this paper, we estimate two different sets of models. The first model focuses on the probability of TWI, and it uses the entire population of Danish workers who meet the selection criteria stated earlier. The second model uses the full sample of individuals who experience TWI. It considers PWI as the receipt of disability benefit or the transit to wage-subsidized jobs and uses data for welfare and employment status following TWI, which is also taken from the DREAM register. We look at the longest spell within the first 24 weeks after experiencing TWI to determine the destination status. The latter is grouped into four categories: PWI, self-sufficiency, early retirement and other welfare receipts (see Table 1).[8] PWI includes disability benefits and wage-subsidized jobs while self-sufficiency is largely full-time employment. About 68 % of the sample goes back to ordinary work after experiencing TWI and a considerable share (22 %) transits to a status of PWI.

As a measure of education, we use a binary variable denoting education beyond primary schooling. Whether or not an individual has a physically demanding job is proxied by an occupational brawn score. The latter is constructed on the basis the individual's contemporaneous occupation codes linked to detailed job characteristics from the US DOT. The fourth edition (1977) of the DOT provides measures of job content, reporting 38 job characteristics for over 12,000 occupations. Rendall (2010) performs factor analysis on the 38 job characteristics in DOT, taken across occupations and employment from the US 1970 census.[9] Following Ingram (2006), Rendall arrives at three factors, which

---

6 Around 5 % of the sample is lost as a result of this sample selection criteria.

7 No observation is lost by merging IDA to DREAM.

8 In principle, mortality may be another pathway that competes with disability risk. In practice, not so many in our age group (max 59 years in the sample period) would use it, as on overage only 4,600 individuals within this age group die every year.

9 Given that we cannot rule out that technological advances may have changed the main job characteristics for some occupations since 1970, we also propose an alternative measure of brawny job, based on responses to a more recent survey conducted in Denmark, as a robustness check.

Table 1: Destinations state after temporary work incapacity (1998–2002).

	Permanent work incapacity (1)	Self-sufficiency (2)	Early retirement (3)	Other destinations (4)
All	0.226	0.679	0.003	0.095
Women	0.221	0.707	0.003	0.078
With education beyond primary	0.198	0.708	0.002	0.091
With a physically demanding job	0.228	0.698	0.003	0.079
N	14,482	4,4112	183	6,150

Note: (1) Disability benefit (førtidspension) and subsidized jobs (***flexjob*** and ***skånejob***); (2) mostly employment; (3) early retirement scheme (efterløn); (4) unemployment benefits, active labor market policies and sickness benefits.

are labeled brain, brawn and motor coordination.[10] We use the results of factor analysis by Rendall (2010) on these characteristics and obtain a single measure of brawn for each occupational code provided in the Danish register set.

We then use crosswalks to match census occupation codes to the Danish register's occupation codes, called DISCO (Danish ISCO – International Standard Classification of Occupations). Specifically we use a crosswalk of 506 occupations at Ganzeboom (2003) to map 1980 census codes to ISCO-88. The four-digit ISCO code is organized hierarchically in four levels and is almost identical to DISCO (390 vs. 372 codes).[11] The first digit aligns broadly with the education level required for the job, and the job classification becomes more detailed with the second and third digits. If a direct match from a DISCO code to 1980 census code is not found, we use the nearest category, above or immediately adjacent (based on preceding digits). One to many matches receive an average of the brawn score. Over 92 % of workers have an occupation code to which a brawn score can be attributed.

We proceed with two alternative measures of physical demands of the job. One is related only to the brawn level of the current job being above the 75th

10 The orthogonal factor analysis results in three factors explaining 93 % of the total covariance. A partition of job characteristics into the three factors is chosen, according to high coefficients from the analysis. This partition is then used to construct a second factor analysis that allows for correlation across factors. We sum the scores on the six variables contributing to the brawn factor, weighted by the coefficients from factor loading, so that for each 1970 census occupation, we obtain a single value denoting brawn level. Additional details about the factor loadings are available on request from the authors.

11 http://www.dst.dk/Vejviser/Portal/loen/DISCO/DISCO-88/Introduktion.aspx.

percentile of the gender-specific brawn distribution while the other is based on having a history of brawny work in the last 3 years.[12]

3.1 Descriptive Statistics

The full sample consists of 2.3 million observations of individuals born between 1943 and 1950, employed in 1997, and subsequently observed in 1998–2002. Table 2 presents descriptive statistics for both the full sample and conditional on TWI split by gender, and within-gender group, separately for all and for the higher of the two educational groups (above primary level). Means are taken over all person years 1998–2002 when workers are between 48 and 59 years of age.

Around 3 % of women and men aged 48–59 experience TWI. For those with education beyond primary level, the corresponding shares are on average slightly lower. Years of education are, by definition, lower for all workers compared to educated workers, almost 10 years versus 12 years. Women and men in the full sample are similar in terms of age, the average being 53 years. Looking at the shares of workers employed in physically demanding occupations, a striking gender difference is seen. While 12 % of women work in such occupations, 25 % of all men work in brawny jobs, reflecting the high degree of occupational segregation by gender in the Danish labor market. Moreover, the average brawn level in women's commuting areas (see footnote 15) is also considerably lower, 67, compared with men, 98. Marriage rates among older working women are about 2 percentage points lower than among their male counterparts. Educated workers – both women and men – are more likely to be born in a city. Table 2 also shows that the proportion which is not affected by the education reform instrument and which is used as a control group is approximately 26 % for both the sample of women and men. Due to mortality risk increasing with age, there is slightly more representation of the younger cohorts in the sample; however, differences in cohort sizes are not large.

When conditioning on TWI, 21–23 % of women and 19–21 % of men are on PWI, so the differences are not large. The TWI sample is also less educated than the full sample but is about the same age. Those on TWI are more likely to work in brawny jobs than those in the full sample. The brawniness of their (the TWI)

12 Ideally, we would like to go further back than the past 3 years when defining the history measure. However, the cross-walk available in the O*NET that allows us to merge Danish registers with the American DOT is only available for DISCO codes after 1994. To measure the brawniness we need to match for each DISCO code the task content retrieved from O*NET for each occupation.

Table 2: Descriptive statistics.

Variables	Women		Men	
	All	Education beyond primary	All	Education beyond primary
Full sample				
Temporary work incapacity	0.0340	0.0281	0.0283	0.0202
Years of education	10.7272	12.0451	10.8406	12.5959
Education beyond primary	0.5918		0.4855	
Age	53.3294	53.1016	53.4342	53.2460
Physically demanding job	0.1198	0.0818	0.2477	0.1274
Average brawn at the commuting area	67.4503	66.57455	97.8857	96.0558
Married	0.7493	0.7402	0.7735	0.7857
Urban	0.4069	0.4681	0.4043	0.4898
Reform × urban	0.2684	0.3180	0.2602	0.3245
Cohort (1943)	0.0996	0.0851	0.1075	0.0947
Cohort (1944)	0.1138	0.1002	0.1207	0.1098
Cohort (1945)	0.1258	0.1152	0.1313	0.1218
Cohort (1946)	0.1352	0.1308	0.1384	0.1353
Cohort (1947)	0.1381	0.1362	0.1344	0.1338
Cohort (1948)	0.1318	0.1372	0.1272	0.1337
Cohort (1949)	0.1265	0.1410	0.1203	0.1303
Cohort (1950)	0.1293	0.1542	0.1203	0.1406
N	110,5809	647,598	123,0614	597,433
Sample of individuals with temporary work incapacity				
Permanent work incapacity	0.2349	0.2059	0.2219	0.1904
Years of education	10.1411	11.7620	9.8879	11.7150
Education beyond primary	0.2256		0.1636	
Age	53.7681	53.8967	53.8150	54.0297
Physically demanding job	0.1711	0.1298	0.3457	0.2148
Average brawn at the commuting area	67.9019	66.8856	99.1773	97.0483
Married	0.7142	0.6510	0.7262	0.7442
Urban	0.3985	0.4737	0.3810	0.4672
Reform × urban	0.2591	0.3143	0.2352	0.2918
Cohort (1943)	0.1008	0.0948	0.1216	0.1161
Cohort (1944)	0.1199	0.1075	0.1313	0.1249
Cohort (1945)	0.1308	0.1167	0.1362	0.1326
Cohort (1946)	0.1362	0.1377	0.1430	0.1458
Cohort (1947)	0.1406	0.1403	0.1327	0.1499
Cohort (1948)	0.1266	0.1279	0.1179	0.1157
Cohort (1949)	0.1232	0.1305	0.1104	0.1149
Cohort (1950)	0.1220	0.1446	0.1070	0.1001
N	33,297	16,027	31,760	10,593

commuting area is, however, similar to that of all workers. They are also less likely to be married compared to the full worker sample, perhaps indicative of the health benefits that can be obtained from marriage (Wilson and Owsald 2005). Means of the remaining factors in this table (born in urban, reform × born in urban, cohort dummies) appear to be largely the same for the TWI subsample as for the full sample.

Not surprisingly, Figure 1 shows a negative, even though imperfect, correlation between having a physically demanding job and our education variable for both genders. The share of individuals with a brawny job is in fact higher among those with below primary education for both genders. The same share does not decrease with age for those with the lowest educational level and it is on a negative trend for both women and men with above primary education.

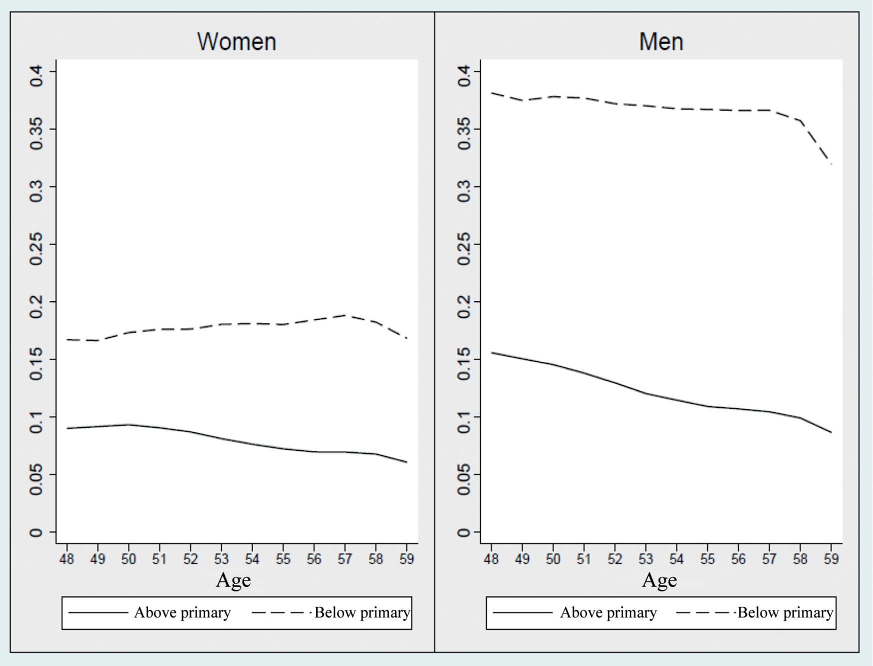

Figure 1: Share of individuals with physically demanding job by education and age.

4 Empirical Strategy

Our point of departure for describing the determinants of either TWI or PWI is a set of modified versions of Grossman's intertemporal model of health (Muurinen 1982; Muurinen and Le Grand 1985; Case and Deaton 2005). According to this theoretical setup, utility maximization determines the optimal level of health inputs and yields a reduced form model of the demand of health, such as

$$H_i = X_i\beta + \gamma_1 E_i + \gamma_2 PJ_i + \xi_i, \qquad [1]$$

where H_i is the health stock of individual i, X_i include a number of exogenous regressors in the health equation, E_i is education, PJ_i is having a physically demanding job and ξ_i is an error term. Education may be correlated with the unobserved component as those with better "endowment" to obtain more education and are at the same time more healthy as adults (Card 1999; Rosenzweig and Schultz 1983). Alternatively a correlation between education and ξ_i may stem from the assumption that individuals with higher preferences for the future are more likely to engage in activities with current costs and future benefits such as education and health investments (Fuchs 1982; Grossman and Kæstner 1997). Similarly, underlying preferences for compensation-seeking behavior could affect both education and health and later as we shall see, also working in a physically demanding job and health. To take account of this endogeneity issue, we assume a model, where education, measured by a dummy variable taking value of 1 for those having any education greater than 8 years of registered schooling, and the health outcome, WI_i (work incapacity), in our case the probability of either TWI or PWI, is related through the following two equations:

$$\begin{cases} WI_i = f\left(X_{1i}\beta_1 + \gamma_1 E_i + \gamma_2 PJ_i + \xi_{1i}\right) \\ E_i = g\left(X_{2i}\beta_2 + \phi_1 URB_i + \phi_2 (URB_i \times REF_i) + \phi_3 \times REF_i + \xi_{2i}\right). \end{cases} \qquad [2]$$

The vector X_{1i} includes age, age squared, gender[13], whether married and born in an urban area, the average work incapacity (TWI or PWI) annual growth rate for each level of brawn, cohort and region of living dummies. Education is a function of cohort dummies (X_{2i}) and whether born in an urban area (URB_i) and whether affected by the 1958 school reform (REF_i), where the latter is equal to 1 for cohorts born in 1946 and later. Following Arendt (2005, 2008), the interaction of the urban and reform dummies constitute our exclusion restrictions. That is, the change in urban–rural differences in education due to the 1958 reform is exploited to identify the education effect on work incapacity[14]. Figure 2 on the number of

13 To allow the effects of education and occupation to be gender specific, we will provide estimation results separately for women and men.

14 The coefficient on the interaction of the urban and reform dummies thus should be estimated with negative sign if urban–rural differences narrow after the reform. Note that this model is exactly equivalent to interacting the reform with living in a rural area and including a control for rural area in addition. The estimated coefficient on the interaction is simply of the opposite sign in that case.

years of education registered for each cohort in the rural areas reveals that education shows a small jump in correspondence with the 1946 cohort and accelerates its growth with time for women and it increases at decreasing rates prior to the reform year and at increasing rates thereafter for men.[15]

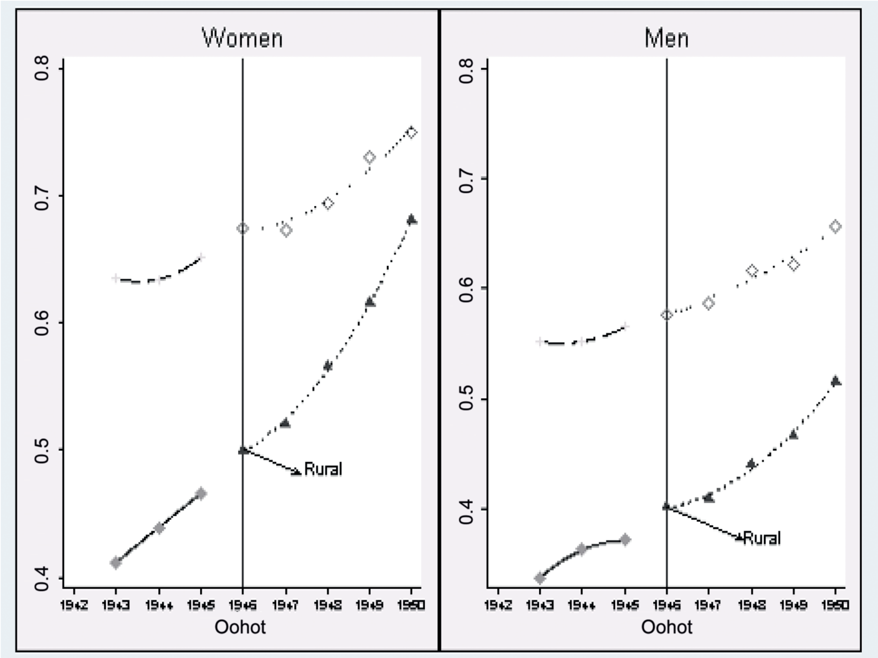

Figure 2: The school reform and the share of individuals with at least primary education.
Note: Trends before and after the school reform are calculated with a quadratic fit.

The previous model is then extended to also endogenize the physically demanding job variable (PJ):

$$\begin{cases} WI_i = f\left(X_{1i}\beta_1 + \gamma_1 E_i + \gamma_2 PJ_i + \xi_{1i}\right), \\ E_i = g(X_{2i}\beta_2 + \phi_1 URB_i + \phi_2(URB_i \times REF_i) + \xi_{2i}), \\ PJ_i = h\left(X_{3i}\beta_3 + \delta_1 E_i + \delta_2 BRAWN\_AVE_i + \xi_{3i}\right). \end{cases} \quad [3]$$

15 Note that Figure 1 is fitted with a quadratic trend, as suggested by Holmlund (2008). Since the quadratic terms did not improve the fit considerably over the linear terms (the figure also shows mainly a linear trend) the regression model includes 14 region dummies and cohort dummies to capture differences in trends.

The vector X_{3i} includes age, age squared, whether married or born in an urban area and cohort dummies. As before, the identification of the occupation effect is obtained by the nonlinearity of the functions, but it is strengthened using an exclusion restriction. As an instrument for having a physically demanding job, we use the 1994 average brawn (separately for male and female) of all employed individuals in the current "commuting area."[16] The lagged average brawn at the commuting area level presents a suitable supply-driven instrument for the individual decision to select a brawny job herself which is not correlated with the risk of work incapacity. This argument is further reinforced by the role of networks in the employment process (Montgomery 1991; Munshi 2003). Thus, individuals placed in areas with high lagged average brawn are also more likely to have a physically demanding job. Thus, we are assuming here that a lagged average brawn level in an area impacts current choice of a physically demanding job but has no impact on working incapacity conditional on a physically demanding job for the individual. This would rule out any impact from other workers through peer effects (if, say, the occupational structure of the industries located in an area impact WI, independently of the individual, e. g. that more manufacturing industries about 4–8 years ago would impact current WI in these areas). Recall, however, that the model includes in addition the average work incapacity (TWI or PWI) annual growth rate for each level of brawn, thereby controlling for the tendency of workers in jobs of the same brawn level to go on disability. Thus, any contagion effects should arise net of the individual's own choice of a physically demanding job and also net of job-type level disability rates. To have a visual counterpart of our identification strategy, Figure 3 maps the spatial distribution of the lagged average brawn level across commuting areas. Eyeball evidence seems to suggest that there is considerable spatial variation in our IV. This is the variation in the data used by the third equation in model [3] to predict our endogenous variable.

It is important to emphasize that although the commuting areas are not closed economies in the sense that workers are free to move out, there is a clear evidence of low residential mobility for Denmark (Deding, Filges, and Van Ommeren 2009), which seems to support the appropriateness of our IV strategy. In theory, it may be preferable to use as instrument the brawn level of the

16 The so-called functional economic regions or commuting areas are identified by using a specific algorithm based on the following two criteria. First, a group of municipalities constitute a commuting area if the commuting within the group of municipalities is high compared to the commuting with other areas. Second, at least one municipality in the area must be a center, i. e. a certain share of the employees living in the municipality must work in the municipality, too (Andersen 2000). In total 51 commuting areas are identified.

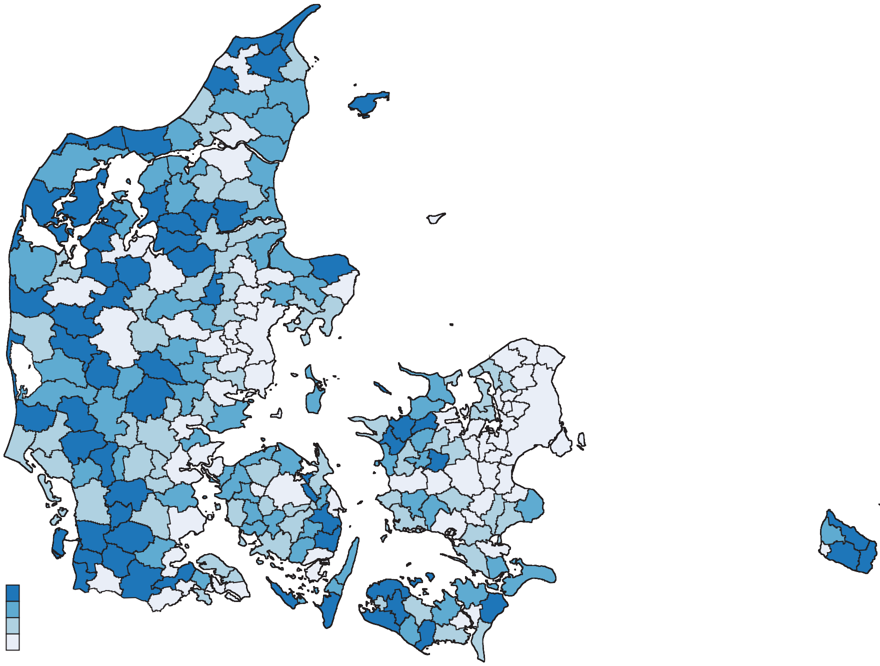

Figure 3: Average brawn level across commuting areas in 1995.

commuting area when the individuals made their first job decisions, say at ages 20–25. Unfortunately, the DISCO occupational code used to identify the commuting area average brawn is only available from 1994 onward, meaning that we cannot go back earlier than that for the purpose of constructing the instrument. At the median, the average brawn of males' commuting areas is 45 % greater than the average brawn of females' commuting areas.[17] Finally, another important feature of our identification strategy is that "brawny" commuting areas in our sample are fairly comparable to less "brawny" commuting areas in terms of average labor income, gender, age, marital status and education. Table 3 reveals that the average socioeconomic characteristics are fairly

---

17 The correlation between having had a brawny job in the past 3 years and the average brawn in the commuting area is 14 % for males and 9.5 % for females. The correlation between the current brawn of individual occupation and the average brawn in the commuting area is 15 % for males and 13 % for females. In principle, of course, a person could live on the border between two commuting areas. To address this potential measurement error problem, we would need more detailed residential address information to define each individual's commuting area. As we show later, however, this instrument turns out to be highly relevant.

Table 3: Average socioeconomic characteristics across commuting areas with different levels of occupational brawn.

	Distribution of occupational brawn			
Variables	**25th percentile**	**50th percentile**	**75th percentile**	**100th percentile**
Years of education	11.0494	10.8141	10.8274	10.4481
Age	53.3695	53.3843	53.40492	53.3933
Men	0.4387	0.4587	0.4906	0.8726
Married	0.7683	0.7683	0.7735	0.7893
Log of annual labor income	12.3303	12.3303	12.3697	12.3804
N	59,9205	93,4421	23,4468	56,8330

Note: The column 25th percentile reports the average of the socioeconomic variables for those commuting areas with occupational brawn up to the 25th percentile of the brawn index distribution. The column 50th percentile reports the average of the socioeconomic variables for those commuting areas with occupational brawn between the 25th and the 50th percentiles of the brawn index distribution. The column 75th percentile reports the average of the socio-economic variables for those commuting areas with occupational brawn between the 50th and the 75th percentiles of the brawn index distribution. The column 100th percentile reports the average of the socioeconomic variables for those commuting areas with occupational brawn above the 75th percentile of the brawn index distribution.

comparable across commuting areas with different brawn levels, although somewhat lower average education and a much larger share of men are recorded in those areas with the highest brawn levels.[18]

To estimate model [3], we apply a quasi-ML estimator with discrete factor approximations (Mroz 1999; Picone et al. 2003; Arendt 2008). The advantages of this approach are twofold. First, it is particularly suited to correct for endogeneity and selection bias caused by unmeasured explanatory variables in dichotomous-dependent variable models (Cameron and Taber 1998). Second, it avoids the need to evaluate multivariate normal integrals and it is less expensive

18 The latter should not be a problem as the analysis is conditional on gender, but the former may indicate some spatial segregation between typically white- and blue-collar geographic areas. To further reassure ourselves that this is not a problem for the interpretation of the results, we have performed the analysis by excluding metropolitan areas (Copenhagen, Frederiksberg, Aarhus and Odense) and altering the definition of job physical demands (by taking the 80th percentile threshold rather than the 75th percentile threshold). The results obtained from these additional analysis are fairly similar to those reported in the paper and are available on request from the authors.

computationally. According to Mroz (1999), the discrete factor method compares favorably with efficient estimators in terms of precision and bias. Therefore, the identification of the effects of education and occupation is obtained by the nonlinearity of the $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$ functions, but it is strengthened using exclusion restrictions.

Specifically, we implement this estimation method by assuming that the error terms ξ_1, ξ_2 and ξ_3 are given by:

$$\begin{cases} \xi_1 = \alpha_1 \times v + u_1 \\ \xi_2 = \alpha_2 \times v + u_2 \\ \xi_3 = \alpha_3 \times v + u_3 \end{cases} \quad [4]$$

and where u_1, u_2 and u_3 are independently distributed and v is a univariate random variable with equation-specific factor-loading parameters α_1, α_2 and α_3. The joint density function l of ξ_1, ξ_2 and ξ_3 conditional on v can be written as

$$\begin{cases} l(\xi_1, \xi_2, \xi_3) = \\ \frac{1}{\sigma_1}\phi\left(\frac{\xi_1 - \alpha_1 \times v}{\sigma_1}\right) + \frac{1}{\sigma_2}\phi\left(\frac{\xi_2 - \alpha_2 \times v}{\sigma_2}\right) + \frac{1}{\sigma_3}\phi\left(\frac{\xi_3 - \alpha_3 \times v}{\sigma_3}\right) \end{cases} \quad [5]$$

where σ_1, σ_2 and σ_3 are the standard deviations of u_1, u_2 and u_3 and ϕ is the standard normal density function.[19] The cumulative distribution of v is approximated by a step function, where $Pr(v = \eta_k) = P_k \geq 0$ for $k = 1,...,K$ and $\sum_{k=1}^{K} P_k = 1$. The number of steps, K, the parameters η_k and the probabilities P_k are estimated jointly with the other parameters of our model, i. e. eq. [3]. The discrete factor, quasi-likelihood function associated with our model is

$$\begin{aligned} L = \prod_{i=1}^{N} \sum_{k=1}^{K} & P_k \Phi\left(X_{1i}\beta_1 + \gamma_1 E_i + \gamma_2 PJ_i + \alpha_1 \eta_k\right)^{WI_i} \\ & \times \left[1 - \Phi\left(X_{1i}\beta_1 + \gamma_1 E_i + \gamma_2 PJ_i + \alpha_1 \eta_k\right)\right]^{1-WI_i} \\ & \times \Phi(X_{2i}\beta_2 + \phi_1 URB_i + \phi_2(URB_i \times REF_i) + \alpha_2 \eta_k)^{E_i} \\ & \times [1 - \Phi(X_{2i}\beta_2 + \phi_1 URB_i + \phi_2(URB_i \times REF_i) + \alpha_1 \eta_k)]^{1-E_i} \\ & \times \Phi\left(X_{3i}\beta_3 + \delta_1 E_i + \delta_2 BRAWN\_AVE_i + \alpha_3 \eta_k\right)^{PJ_i} \\ & \times \left[1 - \Phi\left(X_{3i}\beta_3 + \delta_i E_i + \delta_2 BRAWN\_AVE_i + \alpha_3 \eta_k\right)\right]^{1-PJ_i} \end{aligned} \quad [6]$$

where N is the sample size, K the number of points of support, WI_i is the binary health outcome, whereas E_i and PJ_i are the binary treatment variables, i. e. education and physically demanding job. Because the number of points K is unknown

19 The σ_s and the α_s parameters are not uniquely identified as we have three binary outcomes.

and there can be many local maxima, it is important to look for good starting values. We use single equation estimates and add one point of support for the mixing distribution at a time. Unfortunately, there is no standard theory indicating how to select K in a finite sample. Following Cameron and Taber (2004) and Heckman and Singer (1984), we add additional points of support until the likelihood (and AIC) fails to improve. We added up to three points of support.[20]

5 Results

5.1 The Effects of Education and Occupation on TWI

We begin by exploring the effects of education and working in a physically demanding job on TWI. In Tables 4 and 5, the dependent variable is TWI, i. e. the probability of having a sickness spell of at least 8 weeks, for, respectively, women and men. Coefficient estimates from the quasi-ML model with discrete factor approximation are shown together with standard errors. For education and physically demanding job, we also report the calculated marginal effects, in brackets below. Five models (columns) are discussed. Model 1 treats both education and occupation as exogenous variables and does not account for time-invariant unobserved heterogeneity. Model 2 is the same as model 1 but accounts for unobserved heterogeneity. Model 3 simultaneously models the endogeneity of both education and occupation variables and accounts for unobserved heterogeneity. Finally, models 4 and 5 are the same as model 3 but where alternative definitions of occupation are applied. As clarified in the previous section, to minimize other sources of potential endogeneity, we include only few covariates in the specification of the main equation: age, age squared, marital status, born in a city, the average work incapacity annual growth rate for each level of brawn, seven cohort dummies and 14 regional dummies.

The results for women, in Table 4, are discussed first. Model 1 shows that education exceeding primary school level reduces their likelihood of TWI significantly by 1.4 pp. Working in a physically demanding job, as expected, increases the likelihood of females going on TWI by 1.3 pp, and this effect is

20 Similarly to Mroz (1999) and Picone et al. (2003), we set $\eta_0 = 0$ and $\eta_K = 1$. The other K–2 points of support are given by $\eta_k = \frac{\exp(\theta_{1k})}{1+\exp(\theta_{1k})}$ $k = 1,\ldots,K–1$. The probabilities P_k are also restricted to be between 0 and 1: $P_k = \frac{\exp(\theta_{2k})}{1+\exp(\theta_{2k})}$, $k = 1,\ldots,K-1$, where θ_{2k} is the log odds for location k. When K points are specified, $K-1$ log odds are estimated to give the K probabilities.

Table 4: Education and occupation effects on temporary work incapacity, women.

Temporary work incapacity	Model 1	Model 2	Model 3	Model 4	Model 5
	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]
Education beyond primary	−0.188***	−0.196***	−0.095***	−0.096***	−0.077***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
	[−0.014]	[−0.014]	[−0.007]	[−0.007]	[−0.006]
Physically demanding job	0.151***	0.169***	0.138***		0.143***
	(0.008)	(0.008)	(0.007)		(0.006)
	[0.013]	[0.013]	[0.012]		[0.012]
Physically demanding job in the last 3 years				0.121***	
				(0.007)	
				[0.011]	
Education beyond primary					
Born in city			0.497***	0.643***	0.507***
			(0.005)	(0.043)	(0.005)
Reform			0.168***		0.176***
			(0.005)		(0.005)
Reform × born in city			−0.148***	−0.139***	−0.154***
			(0.007)	(0.047)	(0.006)
Physically demanding job					
Education			−1.157***	−1.654***	−1.222***
			(0.004)	(0.065)	(0.003)
Average brawn at the commuting area			0.017***	0.027***	0.038***
			(0.000)	(0.001)	(0.003)

(*continued*)

Table 4: (*continued*)

Temporary work incapacity	Model 1	Model 2	Model 3	Model 4	Model 5
	(S.E.) [M.E.]	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**
Unobserved heterogeneity					
First support		0	0, 0.395, 1	0, 0.345, 1	0, 0.395, 1
Prob.		0.2504	0.333	0.31	0.352
Second support		0.395	0, 0.617, 1	0, 0.534, 1	0, 0.617, 1
Prob.		0.3252	0.082	0.10	0.059
Third support		1	0, 0.937, 1	0, 0.837, 1	0, 0.937, 1
Prob.		0.4244	0.585	0.586	0.587
Covariance (temp., work inc., education)			−0.216	−0.198	−0.223
Covariance (temp., work inc., phys. dem. job)			0.212	0.231	0.216
Covariance (phys. dem. job, education)			−0.206	−0.201	−0.220
N	910,763	910,763	910,763	910,763	910,763
Pseudo-R^2; LL	0.019	−129,247.151	−886,911.394	−815,600.124	−1,102,045.951
AIC		258,526.30	1,773,854.79	1,631,232.25	2,204,123.90

Notes: The dependent variable is the probability of having a sickness spell of at least 8 weeks. In columns 1, 2, 3 and 4, the variable “physically demanding job” is cc instructed on a DOT-based brawn index, while in column 5 is calculated on a DWECS-based brawn index. The “reform” dummy is equal to 1 for individuals born in 1946 and above, whereas the “urban” dummy is equal to 1 for individuals born in an urban area. In all specifications: (i) the “temporary work incapacity” equation also includes age, age squared, a “married” and an “urban” dummies, the average work incapacity annual growth rate for each level of brawn, cohort and regional dummies; (ii) the “education” equation also includes cohort dummies; (iii) the “physically demanding job” equation also includes age, age squared a “married” dummy, an “urban” dummy and cohort dummies. Marginal effects are shown in brackets. Significance levels: ***1 %, **5 %, *10 %, standard errors in parentheses.

Table 5: Education and occupation effects on temporary work incapacity, men.

Temporary work incapacity	Model 1	Model 2	Model 3	Model 4	Model 5
	(S.E.) [M.E.]	] (S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]
Education beyond primary	−0.224***	−0.282***	−0.059***	−0.061***	−0.062***
	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)
	[−0.014]	[−0.018]	[−0.002]	[−0.002]	[−0.002]
Physically demanding job	0.174***	0.178***	0.141***		0.123***
	(0.007)	(0.008)	(0.018)		(0.016)
	[0.009]	[0.011]	[0.007]		[0.006]
Physically demanding job in the last 3 years				0.117***	
				(0.017)	
				[0.006]	
Education beyond primary					
Born in city			0.519***	0.519***	0.519***
			(0.005)	(0.005)	(0.005)
Reform			0.558***	0.499***	0.378***
			(0.005)	(0.005)	(0.005)
Reform × born in city			−0.081***	−0.079***	−0.077***
			(0.006)	(0.006)	(0.006)
Physically demanding job					
Education			−1.461***	−1.223**	−1.006***
			(0.003)	(0.526)	(0.012)

(*continued*)

Table 5: (*continued*)

Temporary work incapacity	Model 1	Model 2	Model 3	Model 4	Model 5
	(S.E.) [M.E.]	] (S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]
Average brawn at the commuting area			0.008***	0.008***	0.007***
			(0.000)	(0.001)	(0.005)
Unobserved heterogeneity					
First support		0	0, 0.395, 1	0, 0.403, 1	0, 0.399, 1
Prob.		0.8573	0.379	0.36	0.39
Second support		0.617	0, 0.617, 1	0, 0.444, 1	0, 0.369, 1
Prob.		0.0092	0.133	0.133	0.133
Third support		1	0, 0.937, 1	0, 0.654, 1	0, 0.945, 1
Prob.		0.1335	0.488	0.506	0.477
Covariance (temp., work inc., education)			−0.212	−0.202	−0.213
Covariance (temp., work inc., phys. dem. job)			0.206	0.213	0.205
Covariance (phys. dem. job, education)			−0.215	−0.211	−0.283
N	959,899	959,899	959,899	959,899	959,899
Pseudo-R^2; LL	0.017	−117,723.735	−659,706.621	−647,534.40	−774,356.01
AIC		235,479.47	1,319,445.24	1,295,100.80	1,548,744.02

Note: The dependent variable is the probability of having a sickness spell of at least 8 weeks. In columns 1, 2, 3 and 4, the variable “physically demanding job” is constructed on a DOT-based brawn index, while in column 5 it is calculated on a DWECS-based brawn index. The “reform” dummy is equal to 1 for individuals born in 1946 and above, whereas the “urban” dummy is equal to 1 for individuals born in an urban area. In all specifications: (i) the “temporary work incapacity” equation also includes age, age squared, a “married” and “urban” dummies, the average work incapacity annual growth rate for each level of brawn, cohort and regional dummies; (ii) the “education” equation also includes cohort dummies; (iii) the “physically demanding job” equation also includes age, age squared a “married” dummy, “urban” dummy and cohort dummies. Marginal effects are shown in brackets. Significance levels: ***1 %, **5 %, *10 %, standard errors in parentheses.

also statistically significant.[21] Accounting for time-invariant unobserved heterogeneity in model 2 hardly changes the results and produces, if anything, even larger effects of education and occupation on TWI.

In model 3, when instrumenting education by the urban–rural difference in education following the 1958 school reform, we see that the instrument is significantly related to the education measure at the 1 % level and shows, as expected, that the urban–rural difference in the share attaining education beyond primary level decreased after the reform.[22] The effect of education on reducing TWI likelihood now is reduced by half to 0.7 pp. An almost identical effect of education is estimated from an augmented specification, where we include a full set of interactions between the urban and the cohort dummies in the education equation. This allows us to dismiss the issue that our instrument is confounded by a potential relative trend between treatment and control groups. This is also confirmed by the fact that, when we plot the residuals from a regression of cohort dummies on the education variable for each of the four groups (men control, men treated, women control, women treated), there is still a visible discontinuity within and across groups as in Figure 2.[23]. In model 3, current occupation is at the same time instrumented by the (lagged) average brawn in a worker's commuting area. The average brawn is positively related to the brawniness of a female worker's job and is also highly significant, i. e. at the 1 % level. Instrumenting occupation, as can be expected, reduces its effect on TWI but only slightly to 1.2 pp. The marginal effects of education and occupation are equivalent to, respectively, a decrease and a rise in the probability of TWI by 21 % and 35 %, respectively.[24] We cannot, however, conclude anything about the statistical significance of those figures, as in nonlinear models significance of

---

21 It is important to note that our measure of physically demanding job relies on a brawn index, which is estimated and therefore has a variance of its own. Therefore, the standard errors on this variable are likely to be underestimated. We think, however, this may be less of a issue in our case, as we specify the physically demanding job as a dummy variable.

22 To assess the validity of our instrument, we have also run a placebo test, by estimating model 3 on the pre-reform period, i. e. the one including the cohorts 1943, 1944, 1945 and 1946. For this sub-sample, we have specified a "placebo" reform as a dummy variable equal to 1 for the cohorts 1945 and beyond and we have interacted the latter with the urban dummy. The results from this test, available on request, clearly corroborate the appropriateness of our main results as this modified version of the instrument loses its power in terms of predicting individuals' education.

23 All these additional results are available on request from the authors.

24 The figures are obtained using the average sample probability of TWI. From Table 2 the average probability of TWI is approximately 3%. Therefore, the changes in the probability of TWI, in percentage terms, are $(-0.007/0.034)\times 100 = -20.58$ and $(0.012/0.034)\times 100 = 35.3$, respectively.

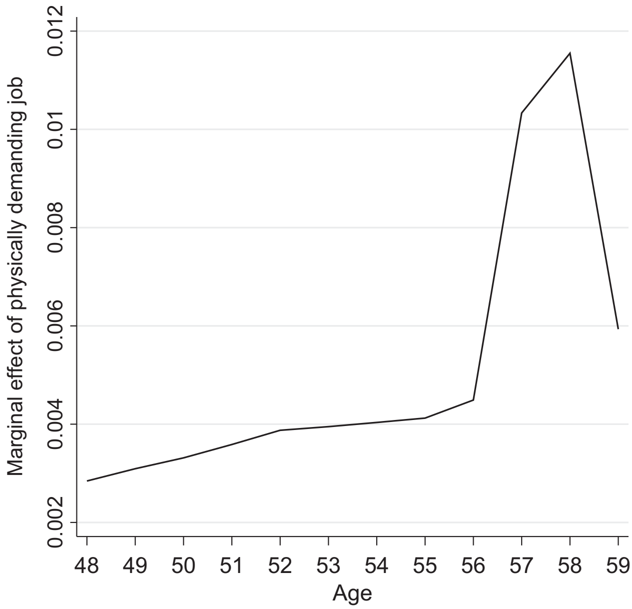

Figure 4: Marginal effects of physically demanding job on temporary work incapacity by age, women.
Note: Marginal effects estimated from model 3 of Table 4.

the underlying parameters does not necessarily imply the significance of the marginal effects. Interestingly, when computed separately by age, we also find that the marginal effect of having a physically demanding job increases with age, except for a flattening off between ages 52 and 55. These results are reported in Figure 4 and seem to confirm the "broken down by work" hypothesis formulated in Case and Deaton (2005). However, given that it is not possible to control for the age, cohort and time factors at the same time, we cannot rule out that these age effects are artifacts of general trends in TWI.

What if it is not the brawniness of the current occupation but rather having a past experience of brawny jobs that matters for TWI? Workers approaching the end of their careers, particularly those experiencing health deterioration, may switch to softer "bridge" jobs or partial retirement before retiring that may even be outside the occupation of the career job (Ruhm 1990). This would imply that the physical demands of the job are erroneously proxied by current occupation brawn.[25] Furthermore, it may be that it is the cumulative exposure to physical

25 In our sample, around 30% of workers changes their occupation on average. Among those switchers, around 80% changes their occupation only once over the sample period. This demands and/or environmental conditions on the job that matters for health breakdown (Fletcher, Sindelar, and Yamaguchi 2011). However, when applying the lagged measure of the physical demands of the job (having a brawny job in the last 3 years) in model 4, it can be seen that occupation still has a causal effect on females' probability of TWI and that the marginal effect is fairly similar to the one estimated from model 3. Thus, for women, the attributes of the current job matter for sickness absence as much as having had physically demanding work in the past. Presumably, this result shows that bridge jobs are not as important a mechanism in high-wage, high labor cost regimes such as in Denmark, where it may be more difficult for older workers to switch jobs at the end of their careers and/or the incentives to do so may be lower. Furthermore, Danish workplaces are said to be among the most accommodating in the world, and the possibility to decrease work hours or negotiate more flexible conditions would be tried before switching employment. Education remains significant in this specification.

Moving to the results of the effect of education and occupation on TWI for men, in Table 5 we see in model 1 that the effect of working in a physically demanding occupation is slightly smaller than it is for women, increasing the chances of TWI by 0.9 pp. Similar to women, education beyond primary decreases men's likelihood of TWI by 1.4 pp (model 2). As in the case for women, accounting for time-invariant unobserved heterogeneity does not change the results substantially.

Proceeding to the most general specification in which both education and occupation are modeled as endogenous in model 3, we find that both education and occupation are statistically significant. Education exceeding primary school level reduces the likelihood of TWI by 0.2 pp, which is more than half as large as the effect for women. Working in a physically demanding occupation has also a lower marginal effect compared to the one estimated for women, 0.7 pp compared to 1.2 pp. Relative to the mean, these figures are equivalent to a decrease and a rise in the probability of TWI by 7 % and 25 %, respectively.[26] As in the case for women, the impact of occupation magnifies with age, as clearly indicated in Figure 5; however, the incline is not as steep. Trying out the alternative definition of occupation based on history of physically demanding work in

suggests that both the current and the lag measure may be a good proxy for the physical demands experienced at work by the individual in our sample.

26 From Table 2, the average probability of TWI is 2.8%. Therefore, the changes in the probability of TWI, in percentage terms, are $(-0.002/0.028) \times 100 = -7.14$ and $(0.007/0.028) \times 100 = 25$, respectively.

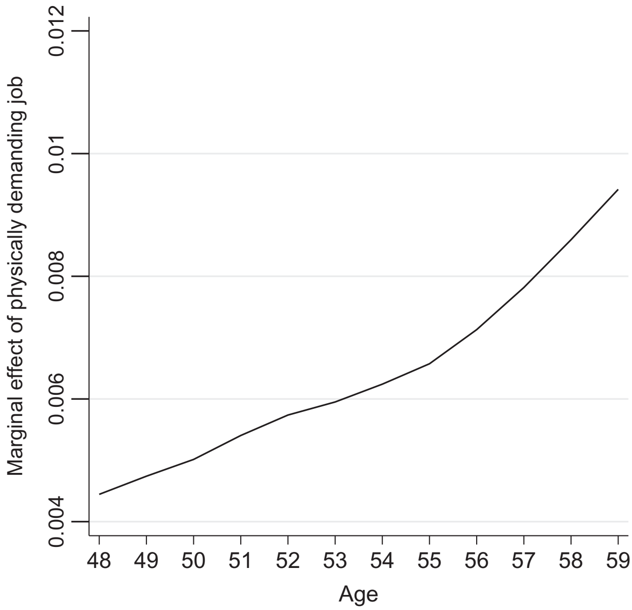

Figure 5: Marginal effects of physically demanding job on temporary work incapacity by age, men.
Note: Marginal effects estimated from model 3 of Table 4.

model 4 does not change the results for men: the coefficient on education remains negative and significant and occupation exerts a statistically significant effect on increasing TWI with a marginal effect of 0.6 pp.

Finally, for both men and women, in model 3, the covariance between education and TWI is estimated to be negative (i. e. the unobserved factors affecting a greater tendency to go on TWI vary negatively with the unobserved factors affecting more education), the covariance between working in a physically demanding job and TWI is positive, and the covariance between education and physically demanding job is negative, as may be expected. The negative (positive) correlation between TWI and education (physically demanding job) could reflect the impact of unobservables such as time preferences or compensation-seeking behavior that was mentioned earlier.

5.2 The Effects of Education and Occupation on PWI

To probe further the effects of occupation and education on work incapacity, in Tables 6 and 7 we estimate their impact on the probability of PWI (either the

Table 6: Education and occupation effects on permanent work incapacity, women.

Permanent work incapacity	**Model l**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model ö**	**Model 7**
	(S.E.) [M.E.]	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**	**(S.E.) [M.E.]**
Education beyond primary	−0.146***	−0.156***	−0.069***	−0.060***	−0.063***	−0.045***	−0.069***
	−0.017	−0.019	−0.021	−0.025	−0.022	−0.02	−0.019
	[−0.044]	[−0.0441	[−0.017]	[−0.016]	[−0.016]	[−0.013]	[−0.017]
Physically demanding job	0.031	0.034	0.031		0.03	0.033	0.026
	−0.023	−0.024	−0.025		−0.026	−0.036	−0.02
	[0.010]	[0.011]	[0.010]		[0.010]	[0.010]	[0.009]
Physically demanding job in the last 3 years				0.028			
				−0.028			
				[0.012]			
Education beyond primary							
Born in city			0.552***	0.573***	0.556***	0.558***	0.565***
			(0.030)	(0.032)	(0.030)	(0.030)	(0.028)
Reform			0.244***	0.253***	0.240***	0.282***	0.260***
			(0.033)	(0.036)	(0.033)	(0.033)	(0.030)
Reform × born in city			−0.244***	−0.235***	−0.225***	−0.227***	−0.212***
			(0.033)	(0.039)	(0.037)	(0.037)	(0.035)

(*continued*)

Table 6: (*continued*)

Permanent work incapacity	Model l	Model 2	Model 3	Model 4	Model 5	Model ö	Model 7
	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]
Physically demanding job							
Education			−0.730***	−0.742***	−0.743***	−0.729***	−0.924***
			(0.023)	(0.025)	(0.023)	(0.023)	(0.019)
Average brawn at the commuting area			0.019***	0.021***	0.019***	0.019***	0.176***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.014)
Unobserved heterogeneity							
First support		0	0, 0.563, 1	0, 0.756, 1	0, 0.339, 1	0, 0.801, 1	0, 0.899, 1
Prob.		0.285	0.487	0.4538	0.486	0.467	0.4671
Second support		0.585	0, 0.739, 1	0, 0.937, 1	0, 0.389, 1	0, 0.644, 1	0, 0.514, 1
Prob.		0.288	0.043	0.0945	0.0121	0.007	0.0122
Third support		1	0, 0.651, 1	0, 0.519, 1	0, 0.781, 1	0, 0.398, 1	0, 0.631, 1
Prob.		0.427	0.51	0.4517	0.5019	0.5260	0.5207
Covariance (work disability, education)			−0.023	−0.023	−0.019	−0.025	−0.022
Covariance (work disability, phys. dem. job)			0.239	0.227	0.246	0.246	0.246

Covariance (phys. dem. job, education)			0.221	0.227	0.246	0.247	0.246
N	26,794	26,794	26,794	26,794	26,794	26,794	26,794
Pseudo-R^2; LL	0.006	−14377.90	−44163.05	−37022.888	−41971.268	−35718.437	−53094.534
AIC		28787.80	88358.10	74077.78	83974.54	71468.87	106221.07

Note: In columns 1, 2, 3, 4 and 7 the dependent variable is the probability of being granted a permanent disability benefit or of transit to a subsized job, conditional on temporary work incapacity. In column 5, the dependent variable is the probability of being granted a permanent disability benefit, conditional on temporary work incapacity. In column 6, the dependent variable is the probability of transit to a subsized job, conditional on temporary work incapacity. In columns 1, 2, 3, 4, 5 and 6 the variable "physically demanding job" is constructed on a DOT-based brawn index, while in column 7 it is calculated on a DWECS-based brawn index. The "reform" dummy is equal to 1 for individuals born in 1946 and above, whereas the "urban" dummy is equal to 1 for individuals born in an urban area. In all specifications: (i) the "permanent work incapacity" equation also includes age, age squared, a "married" and a "born in an urban area" dummies, the average work incapacity annual growth rate for each level of brawn, cohort and regional dummies; (ii) the "education" equation also includes cohort dummies; (iii) the "physically demanding job" equation also includes age, age squared a "married" dummy, a "born in a urban area" dummy and cohort dummies. Marginal effects are shown in in brackets. Significance levels: ***1 %, **5 %, *10 %, standard errors in parentheses.

Table 7: Education and occupation effects on permanent work incapacity, men.

Permanent work incapacity	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]	(S.E.) [M.E.]
Education beyond primary	−0.118***	−0.126***	−0.059***	−0.057**	−0.050*	−0.059***	−0.061***
	(0.020)	(0.022)	(0.024)	(0.027)	(0.026)	(0.024)	(0.022)
	[−0.033]	[−0.036]	[−0.017]	[−0.017]	[−0.015]	[−0.017]	[−0.017]
Physically demanding job	−0.013	−0.013	−0.037		−0.028	−0.037	−0.029
	(0.020)	(0.021)	(0.222)		(0.026)	(0.032)	(0.023)
	[−0.004]	[−0.004]	[−0.012]		[−0.010]	[−0.012]	[0.010]
Physically demanding job in the last 3 years				−0.047			
				(0.029)			
				[−0.013]			
Education beyond primary							
Born in city			0.533***	0.531***	0.518***	0.453***	0.526***
			(0.032)	(0.034)	(0.031)	(0.032)	(0.030)
Reform			0.828***	0.810***	0.795***	0.758***	0.800***
			(0.037)	(0.037)	(0.035)	(0.038)	(0.033)
Reform × born in city			−0.098***	−0.084**	−0.091***	−0.075*	−0.082*
			(0.040)	(0.042)	(0.039)	(0.038)	(0.038)
Physically demanding job							
Education			−1.028***	−0.988***	−0.943***	−1.051***	−1.014***
			(0.078)	(0.026)	(0.025)	(0.022)	(0.023)
Average brawn at the commuting area			0.004***	0.009***	0.004***	0.004***	0.008***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unobserved heterogeneity							
First support		0	0, 0.768, 1	0, 0.558, 1	0, 0.35928, 1	0, 0.722, 1	0, 0.442, 1
Prob.		0.278	0.5134	0.5163	0.4363	0.5548	0.5236

Second support		0.841	0, 0.509, 1	0, 0.451, 1	0,.52463, 1	0, 0.623, 1	0, 0.507, 1
Prob.		0.357	0.0079	0.0099	0.1993	0.0124	0.0109
Third support		1	0, 0.620, 1	0, 0.8054, 1	0, 0.1839, 1	0, 0.576, 1	0, 0.447, 1
Prob.		0.365	0.4787	0.4738	0.3644	0.4328	0.4655
Covariance (work disability, education)			−0.248	−0.247	−0.197	−0.247	−0.246
Covariance (work disability, phys. dem. job)			0.248	0.247	0.187	0.234	0.246
Covariance (phys. dem. job, education)			0.247	0.247	0.206	0.257	0.246
N	23,762	23,762	23,762	23,762	23,762	23,762	23,762
Pseudo-R^2; LL	0.007	−11,952.738	−41,362.75	−32,761.57	−39,424.561	−34,235.144	−42,981.572
AIC		23,937.48	82,757.50	65,555.14	78,881.12	68,502.29	85,995.14

Note: In columns 1, 2, 3, 4 and 7 the dependent variable is the probability of being granted a permanent disability benefit or of transit to a subsized job, conditional on temporary work incapacity. In column 5, the dependent variable is the probability of being granted a permanent disability benefit, conditional on temporary work incapacity. In column 6, the dependent variable is the probability of transit to a subsized job, conditional on temporary work incapacity. In columns 1, 2, 3, 4, 5 and 6 the variable "physically demanding job" is constructed on a DOT-based brawn index, while in column 7 it is calculated on a DWECS-based brawn index. The "reform" dummy is equal to 1 for individuals born in 1946 and above, whereas the "urban" dummy is equal to 1 for individuals born in an urban area. In all specifications: (i) the "permanent work incapacity" equation also includes age, age squared, a "married" and a "born in an urban area" dummies, the average work incapacity annual growth rate for each level of brawn, cohort and regional dummies; (ii) the "education" equation also includes cohort dummies; (iii) the "physically demanding job" equation also includes age, age squared a "married", a "born in an urban area" and cohort dummies. Marginal effects are in brackets. Significance levels: ***1 %, **5 %, *10 %, standard errors are in parentheses.

receipt of disability benefits or the transit to nonregular subsidized jobs) conditional on TWI, hence, the lower sample sizes. Here too, several different models are estimated for, respectively, women and men. As was the case for TWI, the same parsimonious specification is chosen for the equation for PWI. Besides education and occupation (and their instruments), only age and its square, marital status and residence in an urban area are included as determinants. Although in principle PWI and TWI could be separate processes with each of their own determinants, we retain a sparse set of common determinants for both measures of work incapacity. This is because our focus is on uncovering the relationships between education, occupation and the relevant measure of work incapacity controlling for the most obvious sources of heterogeneity and not on accounting for all the variation in the measure of work incapacity. While considering different determinants for each type of work incapacity is certainly a worthwhile approach, it would introduce many potentially endogenous determinants which further complicate the analysis, given that we already are handling two sources of endogeneity simultaneously.

Results for women in Table 6 show that although the workers have a TWI, a higher level of education still reduces their probability of a PWI by 4.4 pp (model 1). Moreover, working in a physically demanding job increases this probability by 1 pp. When endogenizing both variables simultaneously, only the coefficient on education retains its significance. The effect of occupation is not statistically significant even in the specification based on the occupational history.[27] As in Andren (2011), we also estimate the most general specification by distinguishing between the probability of “full-time work disability,” i. e. the receipt of disability benefits, and the probability of “part-time work disability,” i. e. working as subsidized worker at reduced working hours. The results are reported, respectively, in columns 5 and 6 of Table 6 and compared to the previous ones show that the effect of education is precisely estimated for both types of transitions. Very similar effects are found for men in Table 7, except that the effect of education on PWI for men is precisely estimated for part-time disability only, although the effect for full-time disability is estimated to be roughly the same size.

For both men and women, in Tables 6 and 7, the cross-equation correlations between education and PWI are estimated to be negative (i. e. the unobserved factors affecting a greater tendency to go on PWI vary negatively with the unobserved factors affecting more education) and the covariance between

27 Additional results not reported in this paper show that the impact of occupation on PWI does not even vary with age. These additional calculations are available on request from the authors.

working in a physically demanding job and PWI is positive. As mentioned earlier, a possible explanation for these plausible correlations could be, controlling for other things, that educated people and individuals working in nonmanual jobs have lower discount rates and lower expected replacement rates from disability insurance and, therefore, would be less prone to compensation-seeking behavior.

5.3 Robustness Test

All the previous findings on occupation are based on our novel measure of a physically demanding job, so we test the sensitivity of results to this measure. To this end, we run the model using an alternative measure of brawny job based on responses to the Danish Work Environment Cohort Study (DWECS) conducted in 2000. DWECS monitors the working population for the prevalence of occupational risk factors. The study conducted phone interviews on physical, thermochemical and psychosocial exposures, health and symptoms, including a doctor's diagnosis, as well as self- and peer assessment. It is a representative sample of the Danish labor force. Details of the sampling are in Burr, Bjorner, and Kristensen (2003, 2006). The purpose of the DWECS survey is not to derive a complete description of the job or activities, so it is not appropriate to use principal component analysis on the entire questionnaire, as variation driven by any factor accounts for too little of the total variation. Instead, we select questions about more intrinsic (objective) aspects of the job (rather than of the workplace, or of the manager and colleagues), and these include all questions about physical activity in the workplace. They also include aspects of job control and autonomy, and mental or emotional demands.

The first principal component accounts for almost 30 % of the variation, whose eigenvector reflects a brawn-like dimension. Although it is driven by variation of jobs in relation to other jobs, rather than by fundamental tasks or analysis of the composition of a job, it appears that the most objective questions in the survey highlight variation in browniness as the main dimension along which to discriminate occupations. The eigenvector corresponding to the first principal component is used to score each occupation for brawn. The 61 elements in the eigenvector, each corresponding to a survey question, are the loadings given to each survey question. As most of the questions have ordered multiple choice (two or five) answers, we use the polychoric package in Stata (Kolenikov and Angeles 2009) to derive principal components from the 105 occupations for which there were at least 10 respondents. If a direct match from an occupation to a DISCO code from the full sample is not found, we use the nearest category. For each individual in the full sample, the instrument of commuting area level brawn based on the new index is generated in the same way as for the DOT-based index. Results obtained using this alternative measure of brawny job are reported in the last columns of Tables 4–7. In line with the main results, we find that women (men) working in a physically demanding job have a 1.2 (0.6) percentage points higher likelihood of TWI than otherwise comparable women (men).[28] Furthermore, for both women and men, the effects on the probability of PWI are imprecisely estimated, as we have previously found in Section 4, exploiting the DOT-based brawn index.

6 Conclusions

Large numbers of working age individuals continue to leave the labor force via disability pension enroute sickness absence in almost all OECD countries today and the lost work effort of these individuals constitutes a great cost to these economies. The objective of this paper is to empirically model the relationships between education, working in a physically demanding job and health-related exit from the labor market with a view to determining whether these factors have causal impacts on sickness absence and the inflow to disability and their relative importance for each of these processes.

Theoretically, health and disability can be impacted through both education and the nature of jobs. For this reason, we estimate a recursive model of work incapacity and education and occupation with correlated cross-equation errors and unobserved heterogeneity on comprehensive Danish longitudinal data. These data contain objective register-based sickness absence and disability spells, unlike self-reported measures used in the previous literature. The empirical model simultaneously allows for endogenous education and occupation, and identification is achieved by applying as exclusion restrictions a carefully constructed measure of the physical demands of the occupations at the commuting area level and the exogenous variation induced by a major educational reform.

Our findings show that education has a negative *causal* effect on both TWI and PWI, while working in a physically demanding occupation causally affects TWI only.[29] For both men and women, education exceeding primary school level

28 From Table 2, the average probability of TWI is 3.4% and 2.8% for, respectively, women and men. Therefore, the changes in the probability of TWI, in percentage terms, are, respectively $(0.012/0.034) \times 100 = 35.29$ and $(0.006/0.028) \times 100 = 21.4$, respectively.

29 Indeed, a sizable fraction of the causal effect of education (50–75 %) is indirect, i. e. working through the choice of occupation (compare column 2 with column 3 in Tables 4 and 5).

decreases the probability of TWI, but more so for women than men (21 % versus 7 %). The *causal* effect of education suggests that the extra education benefits women more, via more health knowledge, changed time rate of preferences or greater efficiency in health investments.

At the same time, we find a *causal* effect of occupation, which also differs by gender. More specifically, a stronger causal effect increasing TWI is seen for women than men (35 % versus 25 %), despite the average probability of TWI being fairly similar by gender. We also uncover evidence of workers in physically demanding jobs, both men and women, being “broken down by work” over time, and women more so than men, even though the average brawn level of men’s jobs was significantly higher. A number of explanations could be offered for the stronger effect of a physically demanding occupation on women’s transit to sickness absence such as biological differences in the perception of pain, a greater socialization of men to “tolerate” declining health and remain on the job or a higher degree of employer accommodation in men’s jobs. Future work could look into which of these potential explanations carry merit.

Our second major finding for both genders is that, while education significantly affects in most cases PWI, brawniness of the occupation has no effect on the inflow to disability, conditional on TWI. A mechanism could be the rise in the share of awards based on mental disorders over this period, going from 32.1 % of all grants in 2001 to 44.4 % in 2006 and a decline in the share of grants based on musculoskeletal disorders implying a decreasing role for occupational brawn. Recall that the factor analysis based on DWECS data included questions relating to job control and autonomy, and mental or emotional demands in addition to physical demands. However, occupations distinguished themselves most from each other in the DWECS according to a brawn-like aspect. A remaining explanation for the lack of an effect of brawn on PWI could be greater workplace accommodation and efforts to retain workers in Danish workplaces. A final concern could be that moral hazard and not work-induced health deterioration drives the greater transit to TWI of workers in brawny jobs who may be using sickness absence as a waiting state prior to a disability award decision (Bolduc et al. 2002). The fact that brawn did not relate to permanent receipt of disability may, on the face of it, corroborate this compensation-seeking story. However, since we control for time constant individual heterogeneity in the model, we should be able to rule out this explanation unless the individual propensity to seek moral hazard was time varying and correlated with brawny jobs. Several robustness checks on the definition of work incapacity and physically demanding job confirmed the results earlier.

Acknowledgments: The authors would like to thank Jakob Arendt, Pilar Garcia-Gomez, Hans van Kippersluis, Petter Lundborg, Filip Pertold, David C. Ribar and Chan Shen for helpful comments and suggestions. We also appreciate comments from participants at seminars organized by Lund University, University of Maastricht, The Danish Institute of Governmental Research (AKF) and The Swedish Institute of Social Research (SOFI), and participants at the workshop on Health, Work and the Workplace, Aarhus University, May 2011, at the 32nd Nordic Health Economists Study Group Meeting, Odense, August 2011, at the ASHE conference in Minneapolis, June 2012, the invited lecture at the 10th TEPP conference in Le Mans on Research in Health and Labour Economics. The usual disclaimer applies.

Funding: Funding from the Danish Council for Independent Research – Social Sciences, grant no. x09-064413/FSE, is gratefully acknowledged.

References

Andersen, A. K. 2000. *Commuting Areas in Denmark*. Copenhagen: AKF Forlaget.

Arendt, J. N. 2005. "Does Education Cause Health? A Panel Data Analysis Using School Reforms for Identification." *The Economics of Education Review* 24 (2):149160.

Arendt, J. N. 2008. "In Sickness and in Health- Till Education Do Us Part: Education Effects on Hospitalization." *The Economics of Education Review* 27 (2):161–72.

Andren, D. 2011. "Half Empty of Half Full: The Importance of the Definition of Part-Time Sick Leave When Estimating Its Effects." Örebro University working paper, 4/2011.

Autor, D. H., and M. G. Duggan. 2006. "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." *Journal of Economic Perspectives* 20 (3):71–96.

Bago d'Uva, T., E. van Doorslaer, M. Lindeboom, and O. O'Donnell. 2008. "Does Reporting Heterogeneity Bias Health In Equality Measurement?" *Health Economics* 17 (3):351–75.

Becker, G. S., and C. B. Mulligan. 1997. "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics* 112:729–58.

Bengtsson, S., 2007. "Report on the Employment of Disabled Persons in European Countries, Country: Denmark. The Academic Network of European Disability Experts (ANED)." VT/2007/005.

Bolduc, D., B. Fortin, F. Labrecque, and P. Lanoie. 2002. "Workers' Compensation, Moral Hazard and Composition of Workplace Injuries." *Journal of Human Resources* 37 (3):623–52.

Bryld, C.-J., H. Haue, K. H. Andersen, and I. Svane. 1990. *GL 100. Skole, Stand, Forening, Gymnasieskolernes Lrerforening 1890 1990*. Kbenhavn: Gyldendal.

Burr, H., J. B. Bjorner, T. S. Kristensen, et al. 2003. "Trends in the Danish Work Environment in 1990–2000 and Their Associations withLabor-Force Changes." *Scandinavian Journal of Work, Environment & Health* 29:270–9.

Burr, H., E. Bach, H. Gram, and E. Villadsen. 2006. *Arbejdsmiljø I Danmark 2005 – Et Overblik Fra Den Nationale Arbejdsmiljøkohorte*. Kobenhavn, Arbejdsmiljø Instituttet. http://www.arbejdsmiljøforskning.dk//media/Arbejdsmiljoedata/NAK/nak2005- arbejdsmiljø.pdf.

Cameron, S., and C. Taber. 2004. “Borrowing Constraints and the Returns to Schooling.” *Journal of Political Economy* 112:132–82.

Card, D. 1999. “Causal Effects of Education on Earnings.” *Handbook of Labor Economics* 3A:1801–63.

Case, A., and A. Deaton. 2005. “Broken Down by Work and Sex: How Our Health Declines.” In *Analyses in the Economics of Aging*, edited by D. A. Wise, 195–205. Chicago: University of Chicago Press.

Christensen, K. B., M. Labriola, M. Kivimaki, and T. Lund. 2008. “Explaining the Social Gradient in Long-Term Sickness Absence: A Prospective Study of Danish Employees.” *Journal of Epidemiology and Community Health* 62:181–3.

Cutler, D. M., and A. Lleras-Muney. 2007. “The Education Gradient in Old-Age Disability.” NBER forthcoming chapter in the Research Findings in the Economics of Aging.

Datta Gupta, N., N. Kristensen, and D. Pozzoli. 2010. “External Validation of the Use of Vignettes in Cross-Country Health Studies.” *Economic Modelling* 27 (4):854–65.

Deding, M., T. Filges, and J. Van Ommeren. 2009. “Spatial Mobility and Commuting: The Case of Two-Earner Households.” *Journal of Regional Science* 49:113–47.

Fletcher, J. M., J. Sindelar, and S. Yamaguchi. 2011. “Cumulative Effects of Job Characteristics on Health.” *Health Economics* 20 (5):553–70.

Freedman, V. A., and L. G. Martin. 1999. “The Role of Education in Explaining and Forecasting Trends in Functional Limitations among Older Americans.” *Demography* 36 (4):461–73.

Fried, L. P., and J. M. Guralnik. 1997. “Disability and Older Adults: Evidence Regarding Significance, Etiology and Risk.” *Journal of American Geriatric Society* 45:92–100.

Fuchs, V. 1982. “Time Preference and Health: An Exploratory Study.” In *Economic Aspects of Health: Second NBER Conference in Standford*, edited by V. Fuchs, 93–119. Chicago: University of Chicago Press.

Ganzeboom, H. B. G., and D. J. Treiman. 2003. “International Stratification and Mobility File: Conversion Tools.” Accessed February 26. http://home.fsw.vu.nl/hbg.ganzeboom/ismf/occisko.

Goldman, D. P., and D. Lakdawalla. 2001. “Understanding Health Disparities across Education Groups.” NBER working paper, no. 8328.

Grossman, M. 1972. “On the Concept of Health Capital and the Demand for Health.” *Journal of Political Economy* 80 (2):223–55.

Grossman, M., and R. Kæstner. 1997. “The Effects of Education on Health.” The Social Benefits of Education, MI: 69–123.

Heckman, J., and B. Singer. 1984. “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data.” *Econometrica* 52 (2):271–320.

Hofmann, B. 2013. “Sick of Being Activated?” *Empirical Economics* 47 (3):1103–27.

Holmlund, H. 2008. “A Researcher’s Guide to the Swedish Compulsory School Reform.” LSE Research Online Documents in Economics 19382, London School of Economics and Political Science, LSE Library.

Høgelund, J., and A. Holm. 2006. “Case Management Interviews and the Return to Work of Disabled Employees.” *Journal of Health Economics* 25 (3):500–19.

Høgelund, J., A. Holm, and J. McIntosh. 2010. “Does Graded Return-to-Work Improve Sick-Listed Workers’ Chance of Returning to Regular Working Hours?” *Journal of Health Economics, Elsevier* 29 (1):158–69.

Ingram, B. F., and G. R. Neumann. 2006. “The Returns to Skill.” *Labour Economics* 13 (1):35–59.

Larsen, M. 2006. “Fastholdelse Og Rekruttering Af Ældre. Arbejdspladsers Indsats. (Retaining and Recruiting the Elderly: Workplace Interventions).” Copenhagen: The Danish National Centre for Social Research, WP 06: 09.

Larsson, L. 2006. “Sick of Being Unemployed? Interactions between Unemployment and Sickness Insurance.” *Scandinavian Journal of Economics* 108 (1):97113.

Johansson, P., and M. Palme. 2005. “Moral Hazard and Sickness Insurance.” *Journal of Public Economics* 89:1879–90.

Kemptner, D., H. Jurges, and S. Reinhold. 2011. “Changes in Compulsory Schooling and the Causal Effect of Education on Health: Evidence from Germany.” *Journal of Health Economics* 30:340–54.

Kenkel, D. 1991. “Health Behavior, Health Knowledge and Schooling.” *Journal of Political Economy* 99:287–305.

Kolenikov, S., and G. Angeles. 2009. “Socioeconomic Status Measurement with Discrete Proxy Variables: Is Principal Component Analysis a Reliable Answer?” *Review of Income and Wealth* 55 (1):128–65.

Lund, T., M. Labriola, K. B. Christensen, et al. 2006. “Physical Work Environment Risk Factors for Long Term Sickness: Prospective Findings among a Cohort of 5357 Employees in Denmark.” *BMJ* 332:449–52.

Maurer, J. 2007. “Socioeconomic and Health Determinants of Health Care Utilization among Elderly Europeans: A New Look at Equity, Intensity and Responsiveness in Ten European Countries.” HEDG working paper 07/26.

Montgomery, J. D. 1991. “Social Networks and Labor Market Outcomes: Toward an Economic Analysis.” *American Economic Review* 81:1408–18.

Morefield, B., D. Ribar, and C. Ruhm. 2012 “Occupational Status and Health Transitions.” The B.E. *Journal of Economic Analysis & Policy* 11:3.

Mroz, T. 1999. “Discrete Factor Approximations in Simultaneous Equations Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome.” *Journal of Econometrics* 72 (2):233–74.

Munshi, K. 2003. “Networks in the Modern Economy: Mexican Migrants in the US Labor Market.” *The Quarterly Journal of Economics* 118:549–99.

Muurinen, J. M. 1982. “Demand for Health: A Generalized Grossman Model.” *Journal of Health Economics* 1 (1):5–28.

Muurinen, J. M., and Le. Grand. 1985. “The Economic Analysis of Inequalities in Health.” *Social Science and Medicine* 20 (10):1029–35.

Kaptyen, A., J. Smith, and A. van Soest. 2007. “Vignettes and Self-Reports of Work Disability in the United States and the Netherlands.” *American Economic Review* 97 (1):461–73.

OECD, 2007. “Thematic review on reforming sickness and disability policies to improve work incentives, Country Note – Denmark.”

Oreopoulos, P. 2006. “Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter.” *American Economic Review* 96 (1): 152–75, March.

Pettersson-Lidbom, P., and T. Peter Skogman. 2013. “Temporary Disability Insurance and Labor Supply: Evidence from a Natural Experiment.” *Scandinavian Journal of Economics* 115 (2):485–507.

Picone, G. R., F. Slone, S. -Y. Chou, and D. Taylor. 2003. “Does Higher Hospital Cost Imply Higher Quality of Care?” *The Review of Economics and Statistics* 85 (1):51–62.

Rendall, M. 2010. "Brain Versus Brawn: The Realization of Women's Comparative Advantage." Working paper no.491, Institute for empirical research in Economics, University of Zurich.
Rosenzweig, M. R., and T. P. Schultz. 1983. "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs and Their Effects on Birth Weight." *The Journal of Political Economy* 91 (5):723–46.
Ruhm, C. J. 1990. "Bridge Jobs and Partial Retirement." *Journal of Labor Economics* 8 (4): 482–501.
Wilson, C. M., and A. J. Oswald. 2005. "How Does Marriage Affect Physical and Psychological Health? A Survey of the Longitudinal Evidence." The War-Wick Economics Research Paper Series (TWERPS) 728, Department of Economics, University of Warwick.