Journal of Health Economics 58 (2018) 151–175

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Employment, job skills and occupational mobility of cancer survivors

Eskil Heinesen[a,*], Susumu Imai[b,c], Shiko Maruyama[c]

[a] Rockwool Foundation Research Unit, Sølvgade 10, 1307 Copenhagen K, Denmark
[b] Hokkaido University, Kita 9 Nishi 7, Kita-ku, Sapporo, Hokkaido 060-0809, Japan
[c] University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia

ARTICLE INFO

Article history:
Received 20 March 2017
Received in revised form 25 January 2018
Accepted 26 January 2018
Available online 10 February 2018

JEL classification:
I1
I14
J21
J24
J62

Keywords:
Health shock
Return to work
Human capital
Earnings
Disability pension

ABSTRACT

Previous studies find significant negative effects of cancer on employment, with stronger effects for less-educated workers. We investigate whether the effect of cancer varies by skill requirement in the pre-cancer occupation, whether such heterogeneity can explain educational gradients, and whether cancer is associated with changes in job characteristics for cancer survivors who remain employed four years after the diagnosis. We combine Danish administrative registers with detailed skill requirement data and use individuals without cancer as a control group. Our main findings are the following: the negative effect of cancer on employment is stronger if the pre-cancer occupation requires high levels of manual skills or low levels of cognitive skills; the educational gradient diminishes substantially if we allow the effects of cancer to also depend on pre-cancer skill requirements; and cancer is not associated with occupational mobility, indicating potential for policies that reduce labour market frictions for cancer survivors.

1. Introduction

Each year many people of working age are diagnosed with cancer and survival has increased considerably due to screening programmes and improved treatments (Cutler, 2008). Labour market outcomes for cancer survivors are therefore important for society. Cancer is a serious health shock, and it is liable to exert important effects on various economic outcomes (e.g., Lee and Kim, 2008; García-Gómez et al., 2013; Lundborg et al., 2015). Previous studies have found that cancer has significant negative effects on labour market participation, although the majority of cancer survivors return to work (Bradley et al., 2002a,b, 2005, 2007; Steiner et al., 2004; Moran et al., 2011; Short et al., 2008; Datta Gupta et al., 2011; Heinesen and Kolodziejczyk, 2013; Candon, 2015).

Information about which groups of cancer patients are at greater risk of leaving the labour force is important in designing effective labour market policies for cancer survivors. Some dimensions of heterogeneity in the effects of cancer have already been studied in the literature. Bradley et al. (2007) found that the adverse effect of cancer on the labour supply of married females in the US is larger for women who have health insurance through their spouse's employer than for women who have it through their own employer. Comparison of the results in Short et al. (2008) and Moran et al. (2011) indicates that the effect of cancer on the probability of working is similar for workers above and below 55 years of age. The effect of cancer is greater where the cancer is found to have metastasized at diagnosis (Thielen et al., 2015) and with recurrence/further cancers (Heinesen and Kolodziejczyk, 2013). Heinesen and Kolodziejczyk (2013) found significantly larger effects of breast and colorectal cancer on labour market participation for workers with less education than for better-educated workers. They also found significantly larger effects for blue-collar workers than for white-collar workers, although the sample size involved was too small for investigation of the combination of educational and blue-collar/white-collar gradients and their underlying mechanisms. The binary blue-collar/white-collar distinction may also be too coarse to provide useful policy implications.

* Corresponding author at: Rockwool Foundation Research Unit, Sølvgade 10, DK-1307 Copenhagen K, Denmark.
E-mail address: esh@rff.dk (E. Heinesen).

https://doi.org/10.1016/j.jhealeco.2018.01.006
0167-6296/

In this paper, we utilise detailed data on job characteristics to investigate heterogeneity in the effect of cancer on labour market outcomes. Our variables describing job characteristics measure skill and ability requirements in each specific occupation. We construct these variables using the American Occupational Information Network (O*NET) database. The O*NET database and its earlier version, the Dictionary of Occupational Titles, have been used in empirical research into various topics: returns to skills (Ingram and Neumann, 2006; Bacolod and Blum, 2010; Yamaguchi, 2012); the association between wage losses and the extent of skill switching for displaced workers (Poletaev and Robinson, 2000); skills characteristics of immigrants' jobs compared to natives' jobs (Ottaviano et al., 2013; Imai et al., 2014); and the effect of immigration on job skill characteristics of native workers (Foged and Peri, 2016). We are not aware that these data have been used in analyses of the effect of cancer or other health shocks.

We combine detailed O*NET job characteristics data with a large longitudinal dataset of cancer survivors and matched control groups drawn from Danish administrative registers. This allows control for a large number of important baseline characteristics, including health indicators and previous labour market outcomes. We study labour market outcomes four years after the diagnosis of cancer, focusing on workers who were employed at baseline (two years before the diagnosis) and who survived for five years after the diagnosis. We consider all types of cancers (except skin cancer); many earlier studies looked only at the effects of specific cancer (e.g., breast or colorectal cancer).

In particular, we study three new questions. First, we investigate whether the effect of cancer varies with baseline job characteristics. We investigate several aspects of skill requirements of the pre-cancer job, focussing in particular on the hypothesis that the effect of cancer on employment is greater where the pre-cancer job involved high physical and manual demands, because cancer and its treatment may reduce physical strength, and previous studies indicate that cancer survivors are less likely to return to work if they had physically demanding jobs at baseline (see the survey in Spelten et al., 2002). Second, we investigate whether the educational gradient in the effect of cancer on employment is due to differences in occupation or some other reason (e.g., lifestyle and health knowledge) by allowing the effect of cancer to depend on both pre-cancer job characteristics and education levels. Third, we study whether cancer affects the overall probability of switching occupation, workplace or industry, and whether cancer affects job characteristics for those who remain employed after cancer.[1] One may expect, on the one hand, that reduced ability to work due to cancer not only reduces labour market participation but also increases mobility towards less demanding jobs (including jobs requiring less physical strength). On the other hand, even if there is a need for cancer survivors to move to less demanding jobs, this may be difficult in practice due to labour market frictions. Moving to another employer and finding a new job may be particularly difficult after a long period of absence due to a serious illness such as cancer. Information about whether workers with reduced ability to work can readjust their work situation has important policy implications, in view of the large social cost of workers leaving the labour force, possibly to receive disability pensions.

Our results can be summarised as follows. For both genders, cancer reduces the probability of being employed in the fourth year after diagnosis by about 7 percentage points, relative to the non-cancer group. Cancer also leads to an increase in the probability of receiving disability pension (which implies that the person has left the labour force permanently) by 5–6 percentage points. Earnings fall by about 10% because of cancer (not conditioned on being employed after cancer), and before-tax gross income falls by about 3%. We also find significant educational gradients in the effect of cancer on employment status. These results are consistent with earlier studies.

We also find significant gradients in the effect of cancer according to baseline job characteristics. An increase in cognitive (analytical and interpersonal) job skill requirements by 1 standard deviation reduces the negative effect of cancer on employment by about 2 percentage points for males and 1 percentage point for females. An increase in manual skill requirements (physical strength and fine motor skills) by 1 standard deviation increases the negative effect of cancer on employment by about 1.4 percentage points for females, but for males no statistically significant gradient is found. Gradients in cognitive and manual skill requirements in the effect of cancer are statistically significant for both males and females when the outcome is the probability of receiving a disability pension, however. For earnings and income, we find no statistically significant gradients in relation to skill requirements or education. When the effects of cancer are taken to depend on both baseline skill requirements and education, estimates of educational gradients are reduced considerably, especially for females. Consequently, pre-cancer job characteristics are important in explaining why the effect of cancer on labour market outcomes is larger for the low-educated than for the high-educated. This finding probably reflects the fact that low-educated persons must accept jobs characterised by high requirements of physical strength (and low requirements of analytical and other cognitive skills), and that cancer and its treatment often reduce physical strength in particular.

We find no effect of cancer on the probability of moving to a different occupation, plant or industry after cancer (conditional on remaining employed). We also find no effect of cancer on job characteristics (for those who remain employed), which implies that cancer survivors do not switch to less demanding occupations any more than the control group. This is an unexpected result because for many people cancer reduces skills and ability to work. A possible explanation is that some cancer survivors could have tried unsuccessfully to switch to a less demanding occupation. The results therefore indicate very limited opportunities in the labour market for occupational adjustment after a health shock, and thus a potential for policies that enhance labour market mobility of cancer survivors, so as to alleviate the negative effect of cancer on labour-market participation.

Finally, we show that our main regression estimates of the effects of cancer are very similar to treatment effect estimates obtained by using inverse probability weighting, and we address the plausibility of the unconfoundedness assumption in three supplementary analyses. These are: using later cancer patients as an alternative control group; using difference-in-differences; and estimating the effect of cancer on lagged outcomes (falsification test).

2. Empirical methods

We use a dataset of cancer survivors, and a control group of workers without cancer, to estimate the effects of being a cancer survivor on labour market outcomes four years after the year of diagnosis. We pool observations on cancer patients from several base years (years of diagnosis) and select a matched control group for each base year. Let t denote the base year (which is constant over time for a given individual) and consider a model for the outcome four years after the base year, as follows:

$$Y_{i,t+4} = \beta_0 + \beta_1 C_i + \beta_2 C_i X_{i,t-2} + \beta_3 A_{it} + \beta_4 Z_{i,t-2} + \varepsilon_{i,t+4}, \quad (1)$$

where $Y_{i,t+4}$ is the outcome of individual i in calendar year t+4 (four years after the base year), C_i is a dummy variable which is unity

[1] This part of the analysis is related to van de Mheen et al. (1999) who find no significant association between ill health conditions and subsequent upward or downward occupational mobility based on eight ordered occupational classes.

if individual i belongs to the cancer group (if he was diagnosed with cancer for the first time in one of the base years) and zero for the control group, $Z_{i,t\text{-}2}$ is a vector of covariates, $X_{i,t\text{-}2}$ is a subset of $Z_{i,t\text{-}2}$ that interacts with the cancer variable, A_{it} is a vector of base year dummies and dummies for age in year t, and $\varepsilon_{i,t+4}$ represents the error term.[2] The covariates in $Z_{i,t\text{-}2}$, such as lagged labour market outcomes and health indicators, are measured in year t-2 and earlier, but not in year t-1, since some of the individuals who were diagnosed with cancer in year t could have been affected by symptoms in year t-1.

We focus on the effects of cancer four years after diagnosis, but results are very similar if instead we measure outcomes at t+3 or t+5. It is interesting to focus on long-term outcomes, since for most cancer survivors the negative side effects of cancer treatment are much smaller 3–4 years after diagnosis than earlier.[3] We estimate models with different specifications of $X_{i,t-2}$ (skill requirements and/or education variables), and also models without interaction terms.

In our main analysis, we estimate Eq. (1) using OLS. We do not use matching or weighting techniques because we are interested in effects of interactions between cancer and job characteristics (which are continuous variables), and interactions between cancer and education. We investigate the extent of similarity between the cancer and control groups with respect to baseline covariates. We also show that, based on models without interaction terms, OLS estimates of the effect of cancer (β_1) do not differ significantly from estimates of the average treatment effect (ATE) or the average treatment effect on the treated (ATT), using Inverse Probability Weighting (IPW).[4]

The OLS estimator, as well as matching and IPW estimators, assumes selection on observables, i.e. unconfoundedness given the control for the observed baseline characteristics. The administrative register data enables us to take account of an extensive set of control variables, including lagged outcomes and baseline health indicators. We do not observe lifestyle variables (e.g., smoking, alcohol consumption, physical activity, and overweight/obesity), which are presumably important confounders affecting the risk of cancer, survival, and potential labour market outcomes. We expect, however, that the potential bias in estimated effects of cancer due to unobserved lifestyle variables is relatively small given the large set of control variables, including lagged labour market outcomes. If the assumption of unconfoundedness does not hold, then the sign of the bias is unclear because individuals with 'weak' unobserved characteristics presumably have a higher risk of having cancer, but a lower chance of surviving it.

We conduct three robustness checks related to the unconfoundedness assumption. First, we use an alternative control group of later cancer patients. For the treatment group diagnosed with cancer in the base year 2000, for example, we use as control group individuals diagnosed with cancer in 2006 who survived to at least 2011. For the base year 2000, outcomes are measured in 2004, where the control group is not affected by cancer. This control group of later cancer patients is presumably closer to the treatment group regarding unobserved characteristics than a control group mainly consisting of individuals who did not have cancer, or had it many years later.

Second, for some of our outcomes, we use difference-in-differences (DID) methods, in which identification is based on the common trends assumption instead of unconfoundedness. Consider the following model for years t+4 and t-2 without the time-invariant Z variables but with individual-specific fixed effects (η_i):

$$Y_{i,t+4} = \gamma_{0,4} + \gamma_{1,4}C_i + \gamma_{2,4}C_iX_{i,t-2} + \gamma_{3,4}A_{it} + \eta_i + u_{i,t+4}, \quad (2)$$

$$Y_{i,t-2} = \gamma_{0,-2} + \gamma_{3,-2}A_{it} + \eta_i + u_{i,t-2}, \quad (3)$$

where we assume that having cancer in year t may affect outcomes in year t and later but not in t-2. By subtracting (3) from (2), we obtain the DID model

$$Y_{i,t+4} - Y_{i,t-2} = \delta_0 + \delta_1 C_i + \delta_2 C_i X_{i,t-2} + \delta_3 A_{it} + \left(u_{i,t+4} - u_{i,t-2}\right), \quad (4)$$

where $\delta_0 = \gamma_{0,4} - \gamma_{0,-2}$, $\delta_1 = \gamma_{1,4}$, $\delta_2 = \gamma_{2,4}$ and $\delta_3 = \gamma_{3,4} - \gamma_{3,-2}$. The term $\delta_3 A_{it}$ allows for differential time trends. The coefficient δ_0 captures the trend for the reference group (age 30, base year 2005, no cancer). The coefficient on a given age dummy in A_{it}, e.g. for age 40, is the trend from t-2 to t+4 for those aged 40 in year t relative to the trend of the reference group. Similarly, the coefficient on a given base-year dummy, e.g. for base year 2002, is the trend from t-2 to t+4 for those with base year 2002 relative to the trend for the reference group.

Our basic sample is restricted to individuals who were employed in year t-2 because our interest is in whether the effects of cancer show dependence on job characteristics at baseline. Our DID analysis therefore focuses on the effect of cancer on job characteristics, earnings and income rather than on employment status.

Finally, we conduct an analysis suggested in Imbens (2015) to assess the plausibility of the unconfoundedness assumption. Using a modified sample that includes individuals who were not employed at baseline, in addition to the main sample of employed individuals, we regress employment status and earnings in year t-2 on a dummy for cancer (in year t), controlling for covariates (including lagged labour market variables) in year t-3 and earlier; we then test whether the effect of cancer is zero. Thus, we estimate models of the form

$$Y_{i,t-2} = \theta_0 + \theta_1 C_i + \theta_2 A_{it} + \theta_3 Z_{i,t-3} + \upsilon_{i,t-2}, \quad (5)$$

and test if $\theta_1 = 0$.

3. Data

3.1. Cancer and control groups

The selection of the cancer group is based on the Danish cancer and hospitalisation registers (Gjerstorff, 2011; Lynge et al., 2011). We select individuals who were diagnosed with cancer for the first time in 2000–2005 according to the cancer register. We end up with a sample of 25,094 individuals after having excluded the following categories: persons who were diagnosed with skin cancer; those who according to the hospitalization register had any diagnosis of cancer or benign tumour in an earlier year; those who were not in the basic population registers in the year of diagnosis and 1–5 years prior to diagnosis; those who were not 30–60 years of age at the beginning of the year of diagnosis; and those who were not employed two years prior to diagnosis. We exclude skin cancer because previous research showed that skin cancer does not significantly affect labour market outcomes (Heinesen and Kolodziejczyk, 2013). Table A1 in the Appendix A contains more details on the selection of the cancer group.

We select the control group in two steps. First, for each base year t in the period 2000–2005, we select the basic control group that consists of all 30–60-year-olds who were not diagnosed with cancer or benign tumours in 2000–2005 or earlier, who were employed in year t-2, and who were in the registers in all years from t-5 to t.

[2] We have 5 base year dummies and 30 dummies for age in the base year.

[3] Heinesen and Kolodziejczyk (2013) consider the short-term effects of breast and colorectal cancer.

[4] For IPW methods see, e.g., Hirano and Imbens (2001), Hirano et al. (2003) and Wooldridge (2010).

Table 1
Cancer group by gender and by type of cancer: The number of new cases 2000–2005 among 30–60-year-olds, numbers surviving 5 years, and survival rates.

ICD-10	Type of cancer	New cases				5-year survivors				Survival rates (%)	
C codes		Males	%	Females	%	Males	%	Females	%	Males	Females
16	Stomach	388	3.3	114	0.9	78	1.4	28	0.3	20.1	24.6
18–21	Colorectal	1827	15.5	1110	8.3	997	17.3	670	7.2	54.6	60.4
25	Pancreas	404	3.4	214	1.6	27	0.5	15	0.2	6.7	7.0
17, 23–24, 26	Other digestive system	132	1.1	60	0.5	50	0.9	20	0.2	37.9	33.3
34	Lung and bronchus	1590	13.5	1048	7.9	183	3.2	164	1.8	11.5	15.6
30–33, 35–39	Other respiratory	354	3.0	71	0.5	185	3.2	51	0.6	52.3	71.8
50	Breast	32	0.3	6140	46.1	27	0.5	5249	56.6	84.4	85.5
53–55	Uterine	0	0.0	1596	12.0	0	0.0	1358	14.6		85.1
51–52, 56–58	Other genital females	0	0.0	808	6.1	0	0.0	425	4.6		52.6
61	Prostate	1251	10.6	0	0.0	877	15.3	0	0.0	70.1	
60, 62–63	Other genital males	938	8.0	0	0.0	879	15.3	0	0.0	93.7	
64–68	Urinary system	924	7.9	264	2.0	517	9.0	150	1.6	56.0	56.8
70–72	Brain, other nervous system	500	4.3	250	1.9	94	1.6	75	0.8	18.8	30.0
81–90, 96	Lymphomas, multiple myeloma	531	4.5	275	2.1	356	6.2	208	2.2	67.0	75.6
91–95	Leukaemia	477	4.1	217	1.6	323	5.6	150	1.6	67.7	69.1
(1)	Other (except skin cancer)	2416	20.5	1163	8.7	1157	20.1	719	7.8	47.9	61.8
	Total	11,764	100	13,330	100	5750	100	9282	100	48.9	69.6

The numbers in the first column refer to the International Classification of Diseases (ICD-10, C codes).
(1) 0–15, 22, 40–41, 45–49, 69, 73–80, 97.

Next, we select the final control group: for each base year t, gender g, and age a (a = 30, 31,. . ., 60), we draw randomly $10 \times N_{gat}$ persons from the basic control group, where N_{gat} is the number of people in the cancer group with age a and gender g who were diagnosed with cancer in year t. This random draw is repeated sequentially without replacement so that control persons in the final control group of a given base year do not appear in the final control group of any other base years.

Table 1 shows the size of the cancer group by gender and type of cancer. In the main analysis, we condition on survival until the end of the fifth year after the year of diagnosis. Therefore, we also show these figures for the group of five-year survivors. Among female cancer survivors, 57% have breast cancer, whereas for male survivors colorectal cancer is the most common type (with 17%). About 49% of males and 70% of females survive five years. The main reason for this large gender difference in overall survival rates is the high survival rate for breast cancer; a further reason is the higher female survival rates for cancers which are common for both genders (such as colorectal, lung, and brain cancer, and lymphomas, leukaemia, and 'other' cancers). These numbers suggest that male cancer survivors are a more selective group than female cancer survivors. Any gender difference in the estimated effects of cancer may be due to gender differences in the type and stage of cancer or gender differences in labour market behaviour in response to a health shock. It is beyond the scope of this paper to distinguish between these different mechanisms.

3.2. Outcomes: labour market status and income

We focus on labour market outcomes in the fourth year after the base year (t+4) and condition on survival up to the end of year t+5. This ensures that outcomes for the cancer group are not negatively affected by those who are near to death in the year we measure outcomes. Mortality rates for the cancer groups are only 3–4% from the end of year t+4 to the end of year t+5, however, and our results do not change significantly if we condition on survival to the end of t+4 instead of t+5; Table A2 in the Appendix A shows cumulative survival rates by gender, year, and treatment status.

Table 2 shows the means of main outcomes at t+4 by gender and treatment status and two sample t-test statistics for the equality of means between the cancer and control groups. The first three outcomes are mutually exclusive dummy variables for the dominant labour market status during the year: employed, unemployed, or out of the labour force. A person is categorised as unemployed in a given year if he was unemployed for half of the year or more, and as employed if he was not categorised as unemployed and earnings were the dominant source of income (where earnings are annual wage income and income from business activity for the self-employed).[5] The remaining category, "out of the labour force", includes people receiving a disability pension, early retirement benefits, and long-term sickness benefits. The t+4 employment rate of male cancer survivors is about 72%, whereas for the control group it is about 78%. The difference by 6 percentage points is matched by a similar difference, but with opposite sign, for the out-of-labour-force state; there is no significant difference in unemployment rates. The fourth outcome is a dummy for being employed full time. For both the cancer and control groups, about 96% of employed males and 88% of employed females are full-timers. For both genders, the difference between the cancer and control groups in the proportion of full-timers is approximately the same as the difference in the proportion employed.

The fifth outcome is a dummy for receiving disability pension (most of the year), which typically indicates that the individual has left the labour force permanently. For the male cancer group, 26.5% are out of the labour force at t+4, and 8.8% receive a disability pension. For the control group, 20.2% are out of the labour force, and only 2.4% receive a disability pension. The descriptive statistics indicate that male cancer survivors are about 6 percentage points more likely to be out of the labour force at t+4 than the control group, and also 6 percentage points more likely to receive a disability pension. For females, these differences are about 5 percentage points.

The next three outcomes in Table 2 indicate immobility from t-2 to t+4 with respect to occupation, plant, and industry. They are conditional on being employed at t+4. We have information on the industry for all employed persons, and "same industry t+4 as t-2" is based on Statistics Denmark's 111-grouping of industries.[6] Occupation codes and plant identification numbers are missing in t-2 or

[5] Earnings have to be larger than public transfers including disability pension, early retirement benefits, and old-age pension. For those who are not self-employed, earnings also have to be above a minimum level, which is about DKK 45,000 (≈ USD 7500 in 2000 prices). This categorization is based on Statistics Denmark's socioeconomic classification of the population.

[6] Details of data for occupational and industry mobility are given in the Appendix, Section A3.

Table 2
Labour market status and income in the cancer and control groups at t+4 conditional on survival to t+5.

	Cancer	Control	Difference	t test	N cancer	N control
	Males					
Employed t+4	0.718	0.779	−0.060***	−9.98	5750	114,322
Unemployed t+4	0.016	0.019	−0.003+	−1.70	5750	114,322
Out of labour force t+4	0.265	0.202	0.063***	10.67	5750	114,322
Full-timer t+4	0.689	0.746	−0.057***	−9.18	5750	114,322
Disability pension t+4	0.088	0.024	0.064***	17.09	5750	114,322
Same occupation t+4 as t-2 given employed t+4	0.523	0.527	−0.004	−0.45	3427	73,783
Same plant t+4 as t-2 given employed t+4	0.524	0.530	−0.006	−0.65	3050	65,433
Same industry t+4 as t-2 given employed t+4	0.638	0.653	−0.015+	−1.91	4076	87,355
Earnings in t+4 (DKK 1000)	256.533	273.534	−17.001***	−4.25	5750	114,322
Wages in t+4 (DKK 1000)	215.427	228.943	−13.516***	−4.36	5750	114,322
Income in t+4 (DKK 1000)	347.007	346.201	0.807	0.11	5750	114,322
Disposable income t+4 (DKK 1000)	217.565	216.151	1.413	0.35	5750	114,322
Hourly wage rate t+4 (DKK)	202.046	197.820	4.226*	2.00	3358	72,313
	Females					
Employed t+4	0.725	0.772	−0.047***	−9.93	9282	131,521
Unemployed t+4	0.022	0.021	0.001	0.37	9282	131,521
Out of labour force t+4	0.253	0.207	0.047***	10.08	9282	131,521
Full-timer t+4	0.639	0.679	−0.040***	−7.74	9282	131,521
Disability pension t+4	0.080	0.028	0.051***	18.01	9282	131,521
Same occupation t+4 as t-2 given employed t+4	0.604	0.603	0.001	0.17	5859	89,164
Same plant t+4 as t-2 given employed t+4	0.527	0.532	−0.005	−0.72	5228	77,517
Same industry t+4 as t-2 given employed t+4	0.680	0.683	−0.004	−0.64	6605	99,389
Earnings in t+4 (DKK 1000)	191.922	202.994	−11.072***	−6.54	9282	131,521
Wages in t+4 (DKK 1000)	181.722	190.116	−8.394***	−5.36	9282	131,521
Income in t+4 (DKK 1000)	254.823	258.088	−3.265	−1.35	9282	131,521
Disposable income t+4 (DKK 1000)	171.833	173.844	−2.011	−1.40	9282	131,521
Hourly wage rate t+4 (DKK)	165.588	163.322	2.265**	2.75	5804	87,013

+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

t+4 for about 15% and 25% of those employed in t+4, respectively, reducing the number of observations for "same occupation" and "same plant".[7]

The probability of having the same occupation in year t+4 as in t-2 is biased upwards because some firms under-report changes in occupation codes. In contrast, the probability of working at the same plant in t+4 as in t-2 is biased downwards, because the workplace identification number in the register may change from one year to the next when the organisational structure is changed, e.g. when workplaces are merged. However, these biases are presumably similar for the cancer and control groups. According to Table 2 there are no significant differences between the cancer and control groups in mobility with respect to occupation, plant, or industry. Among those who are employed in year t+4, 52–53% of males and 60% of females have the same occupation as in t-2, 52–53% are at the same plant, and about two-thirds are in the same industry.

Earnings, wages, income (before tax), and disposable income (after tax) are measured in 1000 Danish kroner (DKK) per year, and the hourly wage rate is defined conditional on employment at t+4 and measured in DKK (USD 1 ≈ DKK 6); these variables are adjusted to 2000 prices using the consumer price index.[8] In the year t+4, earnings and wage income are about 6% lower for male cancer survivors than for the control group. For females, the figure is about 5%.

3.3. Job characteristics

We construct variables describing job skill requirements and other job characteristics using the US O*NET data, which contain information on more than 200 aspects of skills and other job characteristics. These are based on professional assessments by job evaluation analysts as well as self-reported assessments by workers, for about 1000 different occupations; for details, see Tippins and Hilton (2010). We link the O*NET data to Danish occupation codes and construct six variables that measure basic skill requirements and job characteristics for each occupation: analytical skills, interpersonal skills, physical strength, fine motor skills, visual skills, and customer contact. The first five of these match those used by Imai et al. (2014). Following Imai et al. (2014), we reduce the dimension of the O*NET data by conducting factor analysis and extracting the first principal component for each of the six variables. Factor analysis is conducted on individual-level data for the Danish workforce, in effect using the occupational distribution in Denmark as weights in the analysis. The six variables are normalised to have mean zero and standard deviation unity in the distribution for the Danish workforce. Details of the factor analysis and crosswalks between occupation classifications are provided in Appendix A2.

The upper panel of Table 3 shows the means of the job characteristics variables at t-2 for cancer survivors and for the control group. For males, there are no significant differences in pre-cancer job characteristics, but female cancer survivors held jobs with more analytic and interpersonal skill requirements and less manual skill requirements than the control group.[9] The lower panel of

[7] One reason for the large number of observations with missing plant identification is that this information is based on the job held at the end of November, and some of those employed for most of the year may not be employed at that moment. A further reason is that our dataset does not contain plant information for self-employed persons.

[8] Earnings include wages (for employees) and income from business activity for the self-employed. Income refers to gross income (before tax) including earnings and transfer payments.

[9] If we do not condition on survival to t+5, males with cancer held jobs at t-2 with higher levels of manual skill requirements and lower levels of analytical and interpersonal skill requirements than the control group, whereas for females no significant differences are found between the cancer and control groups. Thus, for

Table 3
Means of job characteristics for cancer and control groups in years t-2 and t+4 by gender.

	Males				Females			
	Cancer	Control	Diff.	DID	Cancer	Control	Diff.	DID
Not conditional on being employed in t+4								
Analytical skills t-2	0.053	0.035	0.018		−0.113	−0.168	0.055***	
Interpersonal skills t-2	−0.101	−0.118	0.017		0.007	−0.056	0.063***	
Physical strength t-2	0.140	0.156	−0.016		−0.109	−0.061	−0.049***	
Fine motor skills t-2	0.401	0.409	−0.008		−0.220	−0.197	−0.022*	
Visual skills t-2	0.566	0.573	−0.008		−0.359	−0.337	−0.022**	
Customer contact t-2	−0.252	−0.255	0.003		0.066	0.024	0.041***	
N	5229	104,362			8650	122,895		
Conditional on being employed in t+4								
Analytical skills t-2	0.168	0.113	0.055**		0.003	−0.074	0.078***	
Analytical skills t+4	0.118	0.053	0.065***	0.010	0.007	−0.065	0.072***	−0.006
Interpersonal skills t-2	0.013	−0.038	0.051**		0.125	0.039	0.086***	
Interpersonal skills t+4	−0.020	−0.083	0.063***	0.012	0.136	0.059	0.077***	−0.009
Physical strength t-2	0.052	0.097	−0.044*		−0.173	−0.108	−0.065***	
Physical strength t+4	0.048	0.096	−0.048**	−0.004	−0.189	−0.122	−0.067***	−0.002
Fine motor skills t-2	0.307	0.350	−0.043*		−0.270	−0.237	−0.032**	
Fine motor skills t+4	0.274	0.318	−0.044*	−0.001	−0.281	−0.256	−0.025*	0.007
Visual skills t-2	0.486	0.526	−0.041*		−0.396	−0.372	−0.024**	
Visual skills t+4	0.427	0.459	−0.032	0.009	−0.425	−0.399	−0.026***	−0.002
Customer contact t-2	−0.208	−0.222	0.014		0.125	0.068	0.056***	
Customer contact t+4	−0.235	−0.239	0.004	−0.010	0.115	0.067	0.048***	−0.008
N	3391	73,055			5854	89,090		

Two-sample t-test:
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table 3 shows means of the job characteristics variables at both t-2 and t+4 for individuals who are employed at t+4 and for which information on job characteristics is available in both years.[10] Conditional on employment at t+4, occupations of cancer survivors are characterised by greater analytical and interpersonal skill requirements but lower physical strength, fine motor skills, and visual skill requirements than occupations in the control group. This applies to both years t-2 and t+4. Gender differences are noticeable: on average, jobs held by males are characterised by higher levels of analytical and manual skill requirements (physical strength, fine motor, and visual), whereas jobs held by females have higher levels of interpersonal skill requirements and customer contact intensity.[11] The simple DID statistics are small and insignificant indicating that cancer incidence is not associated with a change in job characteristics.

Table A8 in the Appendix A shows the means of job characteristics at t-2 by the level of education and cancer status. Correlations between educational levels and job characteristics are high, and patterns are similar for the cancer and control groups. Means of analytical and interpersonal skill requirements, for instance, are about 1.2 standard deviations higher for workers with a further or higher education qualification than for those with no education beyond compulsory school. On the other hand, means for physical strength and fine motor skill requirements are larger for workers with low education than for those with further or higher education. The difference is about 1 standard deviation for males and about 0.6 standard deviations for females.

Some correlations between the job characteristics variables are large, in particular the correlation between analytical and interpersonal skills (0.93–0.94) and – especially for men – the correlation between fine motor skills and physical strength requirements (0.86), and between physical strength requirements and interpersonal skills (−0.86); see Table A9 in the Appendix A for details. To reduce collinearity problems, in parts of our analysis we combine the two 'cognitive' skill variables into one by defining: Cognitive skills = (Analytical skills + Interpersonal skills)/2. Similarly, we define: Manual skills = (Physical strength + Fine motor skills)/2. These two variables are used in interaction terms between job skills and the cancer dummy.

3.4. Control variables, balancing properties, and common support

We use an extensive set of control variables: Job skill requirements in year t-2 (6 variables), missing job skills (1 dummy), educational level (3 dummies), age (30 dummies), base year (5 dummies), county of residence (14 dummies), family type (2 dummies), hospitalisation in at least one of the years t-5 to t-2 by type of diagnosis (16 dummies), consumption of selected categories of prescription drugs during t-5 to t-2 by type of drug (20 dummies), number of contacts with the primary health care sector in year t-2 (3 variables: GPs, specialists and dentists), industry of employment in t-2 (9 dummies), not employed in t-3 and t-5 (2 dummies), log earnings in t-2 and t-5, some unemployment in t-2 and t-5 (2 dummies), degree of unemployment in t-2 and t-5, and full-timer in t-2 and t-5. For females, we also control for a dummy for having no child at age 30; this is a predictor of breast cancer and may also be associated with labour market outcomes.

Variables for educational level are used both as control variables and in interactions with the cancer dummy. In all regressions, we include dummies for three education levels as controls: Vocational education, further education (upper-secondary education and up

both genders there is a higher risk of not surviving cancer among workers with high manual job requirements.

[10] Information on job characteristics is not available for all employed persons, because the occupation code is sometimes missing. Also, the O*NET system does not provide information on jobs within the military. In the regressions presented in this paper where job characteristics in t-2 are included as explanatory variables, we set missing values of these variables to zero and include a dummy for missing job characteristics.

[11] Fig. A1 in Appendix A shows distributions of the job characteristics variables at t-2 for the cancer group by gender. Distributions for the control group (not shown) are very similar.

Table 4
Heterogeneous effects of cancer on the probability of being employed in year *t*+4.

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer	−0.075***	−0.074***	−0.099***	−0.091***	−0.067***	−0.068***	−0.079***	−0.072***
	(0.005)	(0.006)	(0.012)	(0.012)	(0.004)	(0.004)	(0.009)	(0.009)
Cancer*cognitive skills		0.022*		0.017+		0.013**		0.008
		(0.008)		(0.009)		(0.005)		(0.006)
Cancer*manual skills		−0.002		−0.000		−0.014*		−0.016**
		(0.008)		(0.008)		(0.006)		(0.006)
Cancer*vocational edu.			0.020	0.016			0.004	−0.005
			(0.014)	(0.014)			(0.011)	(0.011)
Cancer*further edu.			0.050***	0.030+			0.031**	0.013
			(0.014)	(0.017)			(0.010)	(0.013)
Analytical skills *t*-2	0.009*	0.009*	0.009*	0.009*	−0.003	−0.004	−0.003	−0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Interpersonal skills *t*-2	0.007	0.006	0.007	0.006	0.019***	0.018***	0.019***	0.018***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Physical strength *t*-2	−0.014***	−0.014***	−0.014***	−0.014***	−0.023***	−0.022***	−0.023***	−0.022***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Fine motor skills *t*-2	0.004	0.004	0.004	0.004	0.015***	0.016***	0.015***	0.016***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Visual skills *t*-2	0.007***	0.007***	0.007***	0.007***	−0.001	−0.001	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Customer contact *t*-2	0.009***	0.009***	0.009***	0.009***	0.003*	0.003*	0.003+	0.003*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Vocational education	0.015***	0.015***	0.014***	0.014***	0.034***	0.034***	0.033***	0.034***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Further education	0.036***	0.036***	0.033***	0.034***	0.036***	0.036***	0.034***	0.035***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Higher education	0.079***	0.079***	0.077***	0.078***	0.067***	0.067***	0.065***	0.066***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
N	120,072	120,072	120,072	120,072	140,803	140,803	140,803	140,803

Note: All specifications include the following additional controls: Age (30 dummies), base year (5 dummies), county of residence (14 dummies), family type (2 dummies), hospitalisation during *t*-5 to *t*-2 by type of diagnosis (16 dummies), consumption of selected categories of prescription drugs during *t*-5 to *t*-2 by type of drug (20 dummies), number of contacts with primary health care sector in *t*-2 (3 variables: GPs, specialists and dentists), industry of employment at *t*-2 (9 dummies), not employed in *t*-3 and *t*-5 (2 dummies), log earnings in *t*-2 and *t*-5, some unemployment in *t*-2 and *t*-5 (2 dummies), degree of unemployment in *t*-2 and *t*-5, full-timer in *t*-2 and *t*-5 (2 dummies), no child at age 30 (1 dummy, for females only), a dummy for missing job skills information, and a constant term. In the interaction term with cancer, 'further education' includes all post-secondary education degrees, i.e. also higher education degrees. Heteroskedasticity-robust standard errors in parentheses.

+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

to 4 years of post-secondary education), and higher education (5 or more years of post-secondary education); the reference group is no education beyond compulsory school (9 years). In interactions with the cancer dummy, we combine further and higher education into one category which we refer to as further education.

Table A10 in the Appendix A shows means of the control variables by gender and treatment status. The *t*-test rejects equality of means for many variables because of the large sample size, but the cancer and control groups have rather similar characteristics, as shown by the normalised differences, most of which are less than 0.03. The assumption of common support is not violated as shown in Fig. A2 in the Appendix, which displays the distributions of the propensity scores for the cancer and control groups estimated by a probit model for the probability of cancer, using the full set of control variables listed above.

4. Results

4.1. Heterogeneous effects of cancer on employment and disability pension

Table 4 presents OLS estimates of the effects of cancer, pre-cancer job characteristics (in year *t*-2) and education on the probability of being employed in year *t*+4. All specifications in Table 4 include all the baseline covariates listed in Section 3.4. The results in (1) and (5) indicate that cancer in year *t* reduces the probability of being employed in *t*+4 by 7.5 and 6.7 percentage points for males and females, respectively. These average effect estimates are not sensitive to the choice of covariates. Even if we control only for age and base year, the estimates are very close to those in Columns (1) and (5): −0.078 and −0.064, respectively. The fact that the estimate of the effect of cancer on employment does not change significantly upon conditioning on a large set of baseline covariates indicates the randomness of cancer. Job characteristics at baseline clearly affect employment probability in year *t*+4. If the job held in *t*-2 required high levels of physical strength, the probability of being employed in *t*+4 is smaller; we find that analytical and visual skill requirements for males, and interpersonal and fine motor skill requirements for females, and customer contact intensity for both genders, have opposite effects. The probability of being employed at *t*+4 also increases with the education level.

The next columns show that the effect of cancer depends on the job characteristics at baseline. In Columns (2) and (6) the estimates of the interaction term between cancer and cognitive skill requirements indicate that an increase in cognitive skill requirements in the job at *t*-2 by 1 standard deviation reduces the negative effect of cancer on employment by about 2 percentage points for males and by about 1 percentage point for females. The interaction term between cancer and manual skill requirements is also significant for females; an increase in manual skill requirements of 1 standard deviation enhances the negative effect of cancer on employment by 1.4 percentage points.

Columns (3) and (7) include interaction terms between cancer and education, instead of interactions between cancer and job

Table 5
Effects of cancer on the probability of receiving a disability pension in t+4.

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer	0.063***	0.057***	0.101***	0.083***	0.052***	0.051***	0.078***	0.068***
	(0.004)	(0.004)	(0.009)	(0.009)	(0.003)	(0.003)	(0.007)	(0.007)
Cancer*cognitive		−0.019***		−0.015*		−0.014***		−0.007+
		(0.006)		(0.006)		(0.004)		(0.004)
Cancer*manual		0.012*		0.011*		0.017***		0.016***
		(0.005)		(0.005)		(0.004)		(0.004)
Cancer*vocational			−0.041***	−0.035**			−0.027***	−0.018*
			(0.011)	(0.011)			(0.008)	(0.008)
Cancer*further			−0.064***	−0.033**			−0.045***	−0.027**
			(0.011)	(0.012)			(0.008)	(0.009)
N	120,072	120,072	120,072	120,072	140,803	140,803	140,803	140,803

Note: Regressions include the full set of control variables (see Section 3.4 and the note to Table 4). Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

skills. For males, the magnitude of the negative effect of cancer on employment is about twice as large for those with no education beyond compulsory schooling (the reference group) as for the further education group; the effects are about −10 and −5 percentage points, respectively. The educational gradient is also significant for females, although it is smaller. The average effects of cancer in Columns (1) and (5) and the educational gradients in (3) and (7) are consistent with the estimates for breast and colorectal cancer in Heinesen and Kolodziejczyk (2013). The educational gradient is substantially smaller for both genders when we also allow for interaction effects between cancer and job characteristics (Columns (4) and (8)).

The employment outcome in Table 4 includes both full- and part-time employment. Results for full-time employment (shown in Table B1 in Appendix B) are similar, indicating that there are only small and insignificant effects of cancer on the probability of part-time employment. This may reflect two opposing effects. Thus, the probability of leaving the workforce may be higher for those who would have been part-time workers if they did not have cancer, but the probability of working part-time may be higher for cancer survivors who would have been full-timers if they did not have cancer. The lower panel of Table B1 in Appendix B shows the effect of cancer on the risk of being out of the labour force at t+4 (i.e., neither employed nor unemployed). The results are very similar to the corresponding estimates for the employment outcome in Table 4, but with opposite signs, reflecting the fact that effects on unemployment are small and insignificant.

Table 5 shows estimates for the effect of cancer on the probability of receiving disability pension in t+4. The pattern of the estimates is similar to that in Table 4 with opposite signs, but point estimates indicate even more pronounced educational gradients (models (3) and (7)), and interaction terms between cancer and baseline job characteristics are more significant. Thus, the interaction term between manual skills and cancer has a positive and significant coefficient for both males and females.

4.2. Effects of cancer on earnings, wages, and income

Table 6 reports the estimated effects of cancer on earnings, wages, income (gross income before taxes including transfers), and disposable income (after tax) in year t+4. The last five columns of the table show estimated effects of cancer on these four variables and the hourly wage rate for the subsamples of males and females who are employed in t+4; causal interpretation is problematic here, because of induced sample selection. In these regressions, interaction terms between cancer and job characteristics, and between cancer and education, become insignificant; we, therefore, present only the estimation results without these interactions. Each estimate in Table 6 is the coefficient of the cancer dummy in an OLS regression of the dependent variable (e.g., disposable income in t+4) on the full set of control variables, and in addition lagged values (at t-2 and t-5) of the dependent variable. Results are shown for the full sample, and also for a sample excluding outliers (see the note to Table 6 for details).

Cancer has a significant negative effect on earnings and wages for both genders (see the first two columns of Table 6): about DKK 25,000 and 20,000 (USD 4200 and 3300) for males and females respectively, corresponding to about 10% of average earnings and wages. (See Table 2 for means of these variables.) Cancer is also associated with reduced earnings and wages for those who remain employed; see columns (5) and (6) of Table 6. This reduction is not caused by increased part-time employment for those who remain employed because the effects of cancer on full-time employment (Table B1 in the Appendix B) are smaller than the effects on employment (full- or part-time; Table 4). However, cancer might be associated with a reduction in hours of work within the two categories, which we are unable to investigate because of lack of data on hours of work, except for the part-time/full-time indicator. For females, cancer also has a significant negative effect on income (including transfer payments) before and after tax (columns (3) and (4)), and these income losses are about the same size when we condition on employment (columns (7) and (8)). For males, the effects on income before and after tax are not significant when the full sample is used; when outliers are excluded the effect on income before tax is significant. For females, there is a significant negative effect on the hourly wage rate but of only about 1% or less (DKK 1.0-1.6). For males the point estimates are of similar size, but insignificant. Thus, cancer does not seem to be associated with a large loss of productivity for those who remain employed, although the true negative effect may be larger due to the induced sample selection problem.

4.3. Effects of cancer on job mobility

Table 7 shows estimates for the coefficient of the cancer dummy in models (with the full set of covariates) for three aspects of job mobility: the probability moving to a different occupation, a different plant and a different industry from t-2 to t+4, conditional on being employed in t+4 (and having information on occupation,

Table 6
Effects of cancer on earnings, wages, and income in year t+4.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Conditional on being employed t+4				
	Earnings	Wages	Income	Disposable income	Earnings	Wages	Income	Disposable income	Hourly wage rate
	Males, full sample								
Cancer	−26.838***	−25.255***	−5.142	0.447	−7.186^{+}	−8.038**	7.108	6.941	−0.907
	(3.131)	(2.351)	(6.642)	(3.777)	(3.773)	(2.584)	(8.981)	(5.151)	(1.643)
N	120,072	120,072	120,072	120,072	93,156	93,156	93,156	93,156	75,671
	Males, excluding outliers[a]								
Cancer	−26.318***	−24.433***	−8.721***	−1.725	−6.846***	−7.376***	−0.801	3.539*	−2.004
	(2.074)	(1.944)	(1.713)	(1.224)	(2.022)	(1.928)	(1.994)	(1.495)	(1.229)
N	117,035	119,282	116,743	117,762	90,603	92,367	89,942	91,175	75,541
	Females, full sample								
Cancer	−20.999***	−18.611***	−11.823***	−6.615***	−7.998***	−6.137***	−9.293***	−5.414***	−1.577**
	(1.270)	(1.135)	(2.284)	(1.376)	(1.222)	(0.988)	(2.754)	(1.644)	(0.603)
N	140,803	140,803	140,803	140,803	108,303	108,303	108,303	108,303	92,817
	Females, excluding outliers[a]								
Cancer	−20.099***	−18.220***	−7.255***	−3.810***	−7.286***	−5.733***	−3.694***	−1.809**	−1.017^{+}
	(1.139)	(1.125)	(0.842)	(0.539)	(0.942)	(0.962)	(0.939)	(0.609)	(0.574)
N	139,948	140,720	140,082	139,983	107,636	108,223	107,624	107,744	92,793

Note: The table reports estimated coefficients on the cancer dummy. Regressions include the full set of control variables (see Section 3.4 and the note to Table 4) and lagged values of the dependent variable in t-5 and t-2.

[a] The excluded outliers are observations with values of the dependent variable below zero or above 1000; i.e., above DKK 1 m (USD 167,000) for earnings, wages, income, and disposable income, and above DKK 1000 (USD 167) for the hourly wage rate. Heteroskedasticity-robust standard errors in parentheses.

$^{+}$ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table 7
Effects of cancer on the probability of moving to a different occupation, plant or industry (in t+4 compared to t-2) conditional on being employed in t+4; OLS regressions.

	Males			Females		
	(1) Different occupation	(2) Different Plant	(3) Different industry	(4) Different occupation	(5) Different plant	(6) Different industry
Cancer	−0.006	−0.017*	0.005	−0.010	−0.005	0.003
	(0.008)	(0.009)	(0.007)	(0.006)	(0.007)	(0.006)
N	77,210	68,483	91,431	95,023	82,742	105,994

Note: The table reports coefficients on the cancer dummy in OLS regressions which include the full set of control variables; see Section 3.4 and the note to Table 4. The samples are restricted to observations having information on occupation/plant/industry in t-2 and t+4. Heteroskedasticity-robust standard errors in parentheses.

$^{+}$ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

plant, and industry in t-2 and t+4). Because of this conditioning on employment in t+4, there is an induced sample selection problem, and we cannot, therefore, interpret these estimates as causal effects of cancer. We find no association between cancer and mobility in terms of plant, occupation or industry for those who remain employed, except for the marginally significant coefficient on cancer for plant mobility for males. The estimates in Table 7 are based on OLS regressions, but average marginal effects based on logit and probit models are almost identical.

To address the concern about potential selection bias due to conditioning on employment in t+4, we consider a multinomial model with 3 mutually exclusive outcomes: employed and the same job in t+4 as in t-2; employed and different job in t+4 than in t-2; and not employed (or no job information) in t+4. Here 'job' means occupation, plant or industry. Table 8 shows estimates of the average marginal effect of cancer on these three outcomes based on multinomial logit (MNL) models. As in Table 7, the analysis is restricted to observations with job information at baseline (t-2). Cancer increases the risk of not being employed in t+4 (the third outcome in Table 8), and the size of the estimated effects are not significantly different from the corresponding OLS estimates of employment effects (with opposite sign) in columns (1) and (5) of Table 4.[12] By construction, the percentage-point decrease in the joint probability of the first two outcomes is the same as the percentage-point increase in non-employment. For instance, the estimates in column (3) indicate that, for males, cancer reduces the probability of staying in the same industry in t+4 as in t+2 by 5.1 percentage points, it reduces the probability of moving to a different industry by 2.4 percentage points, and it increases the risk of not being employed in t+4 by 7.5 percentage points. Thus, in percentage points, the negative effect of cancer on the probability staying in the same industry is about twice as large as the effect on the probability of moving to a different industry. However, this does not indicate a large effect on switching industry relative to staying in the same industry because the share staying is about twice as large as the share switching (for both the

[12] The small differences are due to different functional form assumptions (OLS versus MNL), different samples and different definitions of being employed in t+4. The samples in Table 8 are restricted to observations with information on occupation/plant/industry at baseline, whereas the samples in Table 4 include observations without this information (and the models control for dummies for missing information). Not being employed in t+4 in Table 8 includes those who are employed, but for whom we do not have information on occupation/plant/industry in t+4.

Table 8
The average marginal effect of cancer on the probability of staying in the same occupation/plant/industry, moving to a different occupation/plant/industry, and not being employed; multinomial logit models.

	Males			Females		
Outcome	(1) Occupation	(2) Plant	(3) Industry	(4) Occupation	(5) Plant	(6) Industry
Same	−0.033***	−0.033***	−0.051***	−0.037***	−0.031***	−0.045***
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)
Different	−0.036***	−0.045***	−0.024***	−0.033***	−0.034***	−0.022***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)
Not employed	0.069***	0.078***	0.075***	0.070***	0.064***	0.067***
	(0.006)	(0.006)	(0.005)	(0.004)	(0.005)	(0.004)
Different given employed	−0.006	−0.016+	0.007	−0.007	−0.007	0.003
	(0.008)	(0.009)	(0.007)	(0.006)	(0.007)	(0.006)
N	110,648	98,452	119,607	131,617	117,122	140,249

Note: The dependent variable in the multinomial models has three outcomes: Same occupation/plant/industry in *t*+4 as in *t*-2; different occupation/plant/industry in *t*+4 compared to *t*-2; and not employed (or no information on occupation/plant/industry) in *t*+4. The table shows the average marginal effect (over the estimation sample) of cancer on the probabilities of these three outcomes. In addition, we report the effect on the probability of a different occupation/plant/industry conditional on being employed in *t*+4. The models include the full set of control variables (see Section 3.4 and the note to Table 4). Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

cancer and control groups, and for both genders; see Table 2). In order to illuminate this point, we show in the last row of Table 8 the marginal effect of cancer on the probability of moving to a different occupation/plant/industry conditional on being employed in *t*+4: P(different | employed). This conditional probability is equal to P(different)/[P(different)+P(same)]. The estimates of the marginal effect of cancer on this conditional probability can be compared to the OLS estimates in Table 7. The two sets of estimates are almost identical.

An important weakness of the MNL model in this context is that it imposes the assumption of independence of irrelevant alternatives (IIA). Therefore we may consider the multinomial probit (MNP) model which does not rely on the IIA assumption. Formal identification of the MNP model is obtained by, e.g., restricting one of the elements in the 2-by-2 covariance matrix of the error terms of relative choices, for instance by setting one of the variances equal to 1 (Keane, 1992). However, in practice identification requires further restrictions. Without additional restrictions, parameter estimates become very fragile and standard errors very large. Keane (1992) shows that identification may be obtained if the set of regressors includes alternative-specific variables (i.e., variables which for a given individual vary between categories of the dependent variable). However, we do not have such variables in our data (which is a common problem when applying multinomial models in labour economics; Keane, 1992). Instead, we may impose further restrictions on the covariance matrix. If we restrict the 'structural' covariance matrix of the three outcomes to be the identity matrix (which is the restriction imposed by the Stata command mprobit), we impose IIA, but given this restriction there is little reason to use the MNP instead of the MNL model, which is computationally much more efficient (Cameron and Trivedi, 2009, p. 503-504). In fact, the marginal effects obtained for an MNP model with this restriction are almost identical to the estimates for the MNL model in Table 8. We also estimate a more general MNP model which does not impose the IIA assumption: In addition to the usual scaling restriction that one of the variances in the 2-by-2 covariance matrix of the error terms of relative choices is set equal to a constant, we also restrict the other variance term to the same constant, but allow the covariance of the error terms to be estimated freely.[13] The estimated average marginal effects of cancer are very similar to the estimates in Table 8 using the MNL model, despite the fact that the estimate of the covariance is significantly different from that imposed by the IIA assumption in four of the six models.[14]

4.4. Effects of cancer on job characteristics

Even though we do not find significant effects of cancer on the probability of moving to another occupation, it might be the case that cancer survivors are more likely to move to less demanding jobs than the control group when changing occupation. To investigate this hypothesis, we show in Table 9 the estimation results for six models where the dependent variables are the job characteristics in *t*+4. We show only the estimated coefficients of the cancer dummy even though the models include the full set of covariates. We again condition on employment in *t*+4, and thus we cannot interpret the estimates as causal. We find no significant association between cancer and the job characteristic variables in *t*+4.[15] This may indicate that a large proportion of cancer survivors, who would otherwise have had to shift to less demanding jobs, leave the labour force instead. Another possible reason why we do not observe significant associations between cancer and job characteristics in *t*+4 is that we observe a change in job characteristics only when a person shifts to a new job which is characterized by another occupation code. Within many occupation codes, there may be large variation in job characteristics, so that individuals may change job tasks without changing their occupation code.

As in Section 4.3, we also consider multinomial models which do not require conditioning on employment in *t*+4. For each of the six skill requirement variables, we define a categorical dependent variable with four mutually exclusive outcomes: (1) not employed in *t*+4; (2) employed with a less demanding job in *t*+4 than in *t*-2; (3) employed with an equally demanding job in *t*+4 as in *t*-2; and (4) employed with a more demanding job in *t*+4 than in *t*-2. The first category, 'not employed', also includes persons who are employed, but have missing information on occupation code (and therefore job skill requirements).

13 This model is estimated using the cmp command for Stata (Roodman, 2011).

14 The estimates of the marginal effect on 'different given employed' corresponding to the six models of Table 8 are: −0.005, −0.019, 0.004, −0.009, −0.005, and 0.000.

15 The lagged job characteristics coefficients (not shown) have the expected structure: the 'diagonal' elements of the matrix of coefficients are positive and highly significant.

Table 9
Effects of cancer on job skill requirements in year *t*+4, conditional on employment in *t*+4; OLS regressions.

	(1) Analytical	(2) Interpersonal	(3) Strength	(4) Fine motor	(5) Visual	(6) Customer
	Males					
Cancer	0.008	0.006	0.004	0.003	0.014	−0.011
	(0.011)	(0.010)	(0.010)	(0.011)	(0.012)	(0.012)
N	76,446	76,446	76,446	76,446	76,446	76,446
	Females					
Cancer	−0.000	0.002	−0.007	0.004	−0.003	0.005
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)
N	97,628	97,628	97,628	97,628	97,628	97,628

Note: The sample is restricted to those who are employed in *t*+4 and have information on job characteristics in *t*-2. The OLS regressions include the full set of control variables; see Section 3.4 and the note to Table 4. Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table 10
The average marginal effect of cancer on the probability of not being employed and being employed in a job with less, equally or more demanding skill requirements (in *t*+4 compared to *t*-2); multinomial logit models.

Outcome	(1) Analytical	(2) Interpersonal	(3) Strength	(4) Fine motor	(5) Visual	(6) Customer
	Males					
Not employed	0.070***	0.070***	0.070***	0.070***	0.070***	0.070***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Less demanding	−0.025***	−0.024***	−0.021***	−0.030***	−0.031***	−0.027***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)
Equally demanding	−0.021***	−0.022***	−0.022***	−0.022***	−0.022***	−0.021***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
More demanding	−0.024***	−0.024***	−0.027***	−0.018***	−0.017**	−0.022***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
More given more or less	−0.003	−0.005	−0.006	0.011	0.010	0.003
	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)
N	109,591	109,591	109,591	109,591	109,591	109,591
	Females					
Not employed	0.070***	0.070***	0.070***	0.070***	0.070***	0.070***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Less demanding	−0.019***	−0.020***	−0.022***	−0.026***	−0.023***	−0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Equally demanding	−0.024***	−0.024***	−0.024***	−0.024***	−0.024***	−0.024***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
More demanding	−0.027***	−0.026***	−0.024***	−0.021***	−0.023***	−0.028***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
More given more or less	−0.003	0.002	−0.003	0.004	−0.006	−0.003
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)
N	131,545	131,545	131,545	131,545	131,545	131,545

Note: For each of the job skill requirements, the dependent variable in the MNL models has four outcomes: (1) Not employed in *t*+4; (2) Employed in *t*+4 with less demanding job skills than in *t*-2; (3) Employed with equally demanding job skills; and (4) Employed with more demanding job skills. The table shows the average marginal effect (over the estimation sample) of cancer on the probability of these four outcomes. In addition, the table reports the effect on the probability of the fourth outcome (more demanding job skills) conditional on the second or fourth outcome (more or less demanding job skills). Heteroskedasticity−robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table 10 shows estimates of the average marginal effect of cancer on these four outcomes based on MNL models.[16] Cancer increases the risk of non-employment by about 7 percentage points for both males and females. These estimates are very similar to the OLS estimates in columns (1) and (5) of Table 4 (despite differences in functional form, sample and employment definition). The percentage-point increase in the probability of non-employment is (by construction) equal to the percentage-point decrease in the joint probability of the three other outcomes. For instance, for females' interpersonal skills, cancer reduces the probability of having an occupation in *t*+4 with less demanding, equally demanding and more demanding skills, respectively, by 2.0, 2.4 and 2.6 percentage points. However, a larger percentage-point reduction in the probability of more demanding skills compared to the percentage-point reduction in the probability of less demanding skills does not necessarily imply a reduction in the share with more demanding skills *relative to* the share with less demanding skills because the levels of these shares are different in the estimation sample (for females' interpersonal skills, these shares are 26.6 and 19.0%, respectively). To illustrate this, the last row in the two panels in Table 10 shows the effect of cancer on the probability of moving to a more demanding occupation conditional on remaining employed and moving to either a more or a less demanding occupation, i.e. the effect on P(more | more or less). This conditional probability is

[16] We obtain almost identical marginal effects using MNP models estimated by the mprobit command in Stata.

equal to P(more)/[P(more) + P(less)]. Thus, for females, conditional on remaining employed and moving to an occupation with either more or less interpersonal skill requirements, cancer *increases* the probability of more demanding skills by 0.4 percentage points according to the point estimate. However, this estimate and the other estimates of the effect of cancer on the above conditional probability are not statistically significant. Thus, again we find no effect of cancer on job skill requirements in *t*+4 conditional on being employed.

Since the four outcomes of the dependent variable have an ordered structure (not employed, employed with less demanding job skills, employed with equally demanding job skills, employed with more demanding job skills) related to a latent ability-to-work measure, one may consider ordered logit or probit models as an alternative to the multinomial models. However, the ordered models are restrictive since the coefficients of all explanatory variables (except the constant term) are assumed to be the same across the values of the dependent variable. Specifically, this parallel lines assumption implies in our model that the coefficients of all explanatory variables are the same in the three separate binary response models for the probabilities of outcomes 2–4 versus outcome 1, outcomes 3–4 versus outcomes 1–2, and outcome 4 versus outcomes 1–3. Thus, the ordered model assumes, for instance, that the coefficient of cancer is the same when considering outcomes 2–4 (being employed at t+4) versus outcome 1 (not employed) as when considering outcome 4 (more demanding job) versus outcomes 1–3 (not employed or equally or less demanding job). Therefore, because of this model restriction, the large and significant negative effect of cancer on employment will tend to result in a large negative effect of cancer on the probability of moving to an occupation with more demanding skills relative to moving to an occupation with less demanding skills. The parallel lines restriction is clearly rejected by a Brant test (based on the three separate binary response models discussed above), both for all explanatory variables taken together and for many variables separately, including the cancer and education variables.[17] This clear rejection indicates that the extensive margin response (employed versus not employed) and the intensive margin response (in this case whether individuals who remain employed move to more or less demanding jobs) are qualitatively different.

4.5. Heterogeneity of effects with respect to age, cancer site and cancer stage

The effect of cancer on the probability of being employed in year *t*+4 might differ by age, particularly because early retirement options are available from age 60, affecting the oldest workers in our sample; those aged 55–60 at the beginning of base year *t* are aged 60–65 at the end of year *t*+4. Of the cancer survivors in our sample, 41% of males and 29% of females belong to these six oldest cohorts. However, the estimated effects of cancer on employment are not significantly affected by excluding individuals aged 55–60 in the base year; see Table B2 in Appendix B which may be compared to Table 4. In specifications (1) and (5) without interaction terms between cancer and education or skills, the point estimates are almost identical in Tables B2 and B4 in Appendix B. When interaction terms are included (in columns 2–4 and 6–8), the point estimates indicate larger educational and skill gradients in the effects of cancer when the older cohorts are excluded from the sample (Table B2 in Appendix B), but these differences in point estimates are not statistically significant. Further restriction of the sample to 30–50-year-olds does not produce estimates which differ significantly from those in Table 4, either.

Heterogeneity in effects of cancer across cancer site may be important. A detailed investigation is beyond the scope of this paper since there are many types of cancer. Given our focus on gradients in cancer effects with respect to pre-cancer job characteristics and education, we need a rather large sample of cancer survivors for each cancer site. Therefore, we limit the analysis of site heterogeneity to females with breast cancer versus females with other cancers. As discussed in Section 3, breast cancer represents 46% of new cancer cases for females and, due to the relatively high survival rate, more than 56% of 5-year survivors.

Table B3 in the Appendix B shows that on average, breast cancer reduces the probability of being employed 4 years after diagnosis by 5.9 percentage points, whereas the effect is 7.8 percentage points for other cancers. The gradients in the cancer effect with respect to pre-cancer job skills and education found for females in Table 4 seem mainly driven by breast cancer, but the differences in gradients between breast cancer and other cancers are not statistically significant.

The cancer effect may also vary by cancer stage at diagnosis because it is an indication of the seriousness of the disease and the type of cancer treatment applied which affect short- and long-term ability to work for cancer survivors. We limit the analysis of stage heterogeneity to solid tumours, for which stage can be assessed using the TNM classification (i.e., we exclude tumours of the central nervous system, lymphomas, and leukaemia). These solid tumours represent 87% of the 11,764 new cancer cases for males in Table 1, and 94% of the 13,330 new cancer cases for females, but information on cancer stage is missing for about 12% of our sample (15% for males and 9% for females). We code cancer stage by the following three stages: localised (had not entered the lymphatic system), regional lymphatic spreading, and distant metastasis.[18] For new cancer cases (without conditioning on survival) the distribution is 44%, 24%, and 32% for males, and 49%, 36%, and 15% for females. Cancer stage is highly correlated with the type of cancer treatment. For instance, chemotherapy is used for about 14% of localised cancer cases and about 42% of cancers with regional or distant metastatic spread.[19] Survival rates are much higher for localised cancer than for cancers with regional or, especially, distant metastatic spread. When we condition on 5-year survival, the numbers with information on cancer stage are 4075 males and 8061 females, and the shares who had localised, regional, and distant metastatic cancers at diagnosis are 72%, 20%, and 8% for males and 61%, 36%, and 3% for females.

Tables B4 and B5 in the Appendix B show that localised cancer reduces the probability of being employed 4 years after diagnosis by 5.5 and 4.4 percentage points for males and females, respectively. As expected, the average effects of non-localised cancer (regional or metastatic spread) are considerably larger: 9.3 and 8.5 percentage points for males and females, respectively. The significantly larger employment effects of cancer for females with high levels of manual skill requirements in the pre-cancer job found in Table 4 seem

[17] The Brant test is conducted after estimation of ordered logit models. For all 12 skill models (six for each gender), the p value for the test of parallel lines is less than 0.001 both for the joint test for all control variables, and for the cancer coefficient separately. For a discussion of the Brant test, see Long and Freese (2014). According to the estimates of the ordered response models, cancer reduces the probability of moving to an occupation with higher demands (relative to moving to a job with higher or lower demands) by about 5 percentage points for all job skill variables, and these estimates are highly significant. However, using generalized ordered models (estimated by the gologit2 command for Stata; see Williams, 2006) allowing just the cancer and education variables to vary across alternatives, these effects become much smaller and insignificant or only marginally significant, and the parallel lines assumption is still clearly rejected for the remaining covariates.

[18] The TNM codes are N0, N1-N3, and M1, respectively.

[19] These numbers are based on observations with year of diagnosis 2000–2003 for which we have data on cancer treatment in the cancer register.

to be driven by those with non-localised cancer where the gradient is about twice as large as in Table 4 (see columns (6) and (8) of Table B5 in Appendix B), whereas the gradient is small and insignificant for those with localised cancer (see Table B4 in Appendix B). For males, the point estimates also indicate a larger negative gradient with respect to manual skill requirements in the effect of cancer for those with non-localised cancer than for those with localised cancer, but here standard errors are larger, and the differences are not statistically significant. For both genders, the point estimates indicate a larger positive gradient with respect to pre-cancer cognitive skill requirements for those with localised cancer compared to those with non-localised cancer, but these differences are not statistically significant. Educational gradients tend to be larger for non-localised cancer, but again not significantly so.

4.6. Robustness and specification checks

We use OLS regression in our main analysis. However, our estimates of the effects of cancer are robust to using other estimation methods that are also based on the unconfoundedness assumption, such as matching or weighting. To illustrate this, we show in Table B6 in Appendix B ATT and ATE estimates of the effect of cancer on the probability of being employed in t+4 using IPW, and compare them with our basic OLS estimates. When using IPW (or matching) it is not possible to estimate interaction effects between cancer and education or job skills, so we make comparison here with the OLS estimates of the main effects of cancer for the whole sample (columns (1) and (5) of Table 4), and effects for subsamples by educational level. The three sets of estimates (OLS, ATT and ATE) are very similar; clearly, there are no significant differences.

In Section 2, we argued that the unconfoundedness assumption is reasonable given our rich longitudinal data, which make it possible to control for lagged labour market outcomes and baseline health information, for instance. We shall nevertheless address the unconfoundedness assumption in three distinct ways. First, we use an alternative control group of persons who were diagnosed with cancer 6 years later than the treatment group. Thus, for treatment group individuals diagnosed with cancer in the base year 2000, we use as control group individuals diagnosed with cancer for the first time in 2006 (and surviving to 2011). For the base year 2000, outcomes are measured in 2004, where the control group is not affected by cancer. Similarly, for individuals in the treatment group diagnosed with cancer in 2005, we use as control group those diagnosed with cancer in 2011. The control group of later cancer patients is constructed in the same way as the treatment group, i.e., based on the cancer and hospitalisation registers.[20] As argued in Section 2, this control group of later cancer patients is presumably closer to the treatment group regarding unobserved characteristics, including unobserved health and lifestyle variables, than a control group consisting mainly of individuals who did not have cancer, or who had cancer many years later. The results using this alternative control group are very similar to our main results. As an illustration, Table B7 in the Appendix B shows estimates of the effects of cancer on the probability of employment in t+4 when this alternative control group is used. The control group here is larger than the treatment group, because the probability of being diagnosed with cancer increases with age, and the control group is 36–66 years of age in the year of diagnosis, whereas the treatment group is 30–60 years old. There are no significant differences between the estimates in Table B7 in Appendix B and our main results in Table 4. Point estimates are very similar, except that the interaction effect between cancer and manual job skills for females is about 50% larger when the control group consists of later cancer patients.

Second, we use DID methods where identification is based on the common trends assumption instead of unconfoundedness. Our basic sample is restricted to individuals who were employed at baseline (2 years before the base year) because we wish to estimate whether the effects of cancer depend on job characteristics at baseline. DID analysis of the effects of cancer on labour market status (employed, out of labour force, etc.) is not relevant using this sample, because all individuals have the same baseline value of the outcome,[21] whereas DID estimation of the effects of cancer on, e.g., job characteristics, earnings and income is relevant. Table B8 in the Appendix B shows DID estimates of the effect of cancer on earnings, wages and income (the same outcomes as in Table 6), i.e. the dependent variable is the change from t-2 to t+4 in each of these variables. Table B8 in Appendix B reports unconditional DID estimates and DID estimates controlling for age and base year. When common trends are allowed to vary by age and base year, DID estimates are very similar to the corresponding main OLS estimates of Table 6 (there are no significant differences), whereas unconditional DID estimates for earnings and wages are somewhat smaller in absolute value. Table B9 in the Appendix B shows DID estimates of the effects of cancer on job characteristics in t+4, showing both unconditional DID estimates and estimates in which we allow trends to depend on age and base year. As in our main OLS analysis, the coefficient of the cancer dummy is not significant.

Finally, we assess the plausibility of the unconfoundedness assumption by conducting an analysis of the type suggested by Imbens (2015). In our main analysis, the main finding is that cancer in year t has significant effects on employment and related outcomes of labour market status and earnings after cancer (in year t+4). We now conduct a 'placebo test' by checking that cancer in year t does not affect employment and related outcomes in year t-2, conditional on covariates measured in t-3 and earlier. This is not possible using our estimation sample for the main analysis, in which all individuals are employed in t-2. In this placebo test, we thus use a modified sample, which includes individuals who were not employed in t-2, in addition to the main sample of employed individuals.[22] We then regress employment status in year t-2 on a dummy for cancer in year t, controlling for covariates, including employment status and health indicators, measured in t-5 to t-3. Thus, we include in the analysis the same types of covariates as in the main analysis, except that they are measured in the period t-5 to t-3 instead of the period t-5 to t-2. The inclusion of individuals who were not employed (for most of the year) in t-2 increases the sample size by about 30%.

Upon using this sample, we find that cancer in year t has no significant effect on the probability of being employed in t-2, conditional on covariates measured in t-3 and earlier. (See the first column of Table B10 in the Appendix B). Cancer diagnosed in year t could affect labour market outcomes in t-1 if the health of some of the cancer patients were affected by cancer before the time of diagnosis. We would expect any such effects to be very small, and we, therefore, estimate the effect of cancer in year t on employ-

[20] Since we have data from the cancer register only to 2010, the 2011 control group is based only on data on cancer diagnoses from the hospitalization register. We have verified that this does not make a large difference by comparing the groups of cancer patients identified by the two alternative methods for 2010.

[21] We could extend the sample to include also persons who were not employed at t-2 and use DID to estimate main cancer effects (and educational gradients) also for these outcomes. This was done for breast and colorectal cancer in Heinesen and Kolodziejczyk (2013), where DID estimates are very similar to IPW estimates. Using data for Pennsylvania, Moran and Short (2014) find that the effects of cancer 2–6 years after diagnosis for females who were not employed at baseline are only slightly larger than effects for females employed at baseline found in other studies.

[22] In other words, we drop the last exclusion restriction of Table A1 for the cancer group, and construct a new control group which also includes individuals who were not employed in t-2.

ment status in t-1 as an additional placebo test (see column (2) of Table B10 in Appendix B); this effect is also clearly insignificant for both genders. For comparison with our main analysis, we show (in column (3) of Table B10 in Appendix B) the estimated effects of cancer on employment status in t+4. These estimates are very similar to our estimates in the main analysis (see columns (1) and (5) of Table 4), even though the samples are different (the sample in Table B10 in Appendix B includes individuals who were not employed in t-2) and the covariates are different (regressions in Table B10 in Appendix B control for covariates measured in t-5 to t-3, whereas the main analysis includes covariates measured in t-5 to t-2).

As additional placebo tests, we replace the employment dummy with other outcomes: earnings (columns (4)-(6) of Table B10 in Appendix B) and wage income (columns (7)-(9) of Table B10 in Appendix B). For males, the estimates of the effect of cancer in year t on earnings and wages in t-2 and t-1 are insignificant. For females, the estimates are more precise, and for t-2 they are positive and significant at the 5% level. The estimates are very small, however, only about DKK 1200 (USD 200) per year, corresponding to about 0.6-0.7% of average earnings for the cancer group in t-2, or to about 7.5% of the estimated effect of cancer on earnings and wages in t+4 (columns (6) and (9) of Table B10 in Appendix B). These small positive estimates might indicate a weakly positively selected female cancer group which would imply a small bias towards zero in the estimated effects of cancer for females. The estimates for females in t-1 are a little smaller than in t-2 and are not significant. For both males and females the estimates of the effect of cancer on earnings and wages in t+4 (columns (6) and (9) of Table B10 in Appendix B) are similar to the estimates of the main analysis (columns (1) and (2) of Table 6).[23] Overall, this analysis indicates that the unconfoundedness assumption is plausible.

5. Conclusion

We have estimated the effects of cancer on labour market status 4 years after diagnosis for 5-year cancer survivors who were employed 2 years before diagnosis. We find significant gradients in the effects of cancer on labour market status with respect to job characteristics 2 years before cancer. An increase in cognitive job skill requirements by 1 standard deviation reduces the negative effect of cancer on employment by about 2 percentage points for males and 1 percentage point for females. An increase in manual job skill requirements by 1 standard deviation increases the negative effect of cancer on employment by about 1.4 percentage points for females, whereas there is no significant gradient with respect to manual skill requirements for males. When the outcome is the probability of receiving a disability pension, gradients with respect to both cognitive and manual skill requirements in the effect of cancer are significant for males and for females. Gradients with respect to education and skill requirements in the effect of cancer on earnings, wages and income are not statistically significant. One might expect larger negative effects for the highly educated or for individuals with jobs characterized by large cognitive skill requirements, since they typically have higher incomes and therefore public transfer payments such as disability pension provide lower replacement rates; on the other hand, though, these groups often have access to higher supplementary compensation from labour market pension schemes and private insurance schemes, and they are more likely to remain employed.

We also investigate the effects of cancer on job mobility and change of job characteristics from year t-2 to year t+4 by estimating two types of models: OLS models based on individuals who remain employed in t+4 and multinomial models which include non-employment as an outcome. These two approaches yield very similar results: We find no association between cancer and the probability of moving to a different occupation, plant or industry, and we find no association between cancer and change of job characteristics. One reason why we do not observe cancer leading to less demanding jobs, despite the negative effect of cancer and its treatment on the ability to work, may be that it is difficult for some cancer survivors to find new jobs in other firms. Another reason may be that job characteristics for cancer survivors are adjusted to some extent even though their occupation code is not changed. This is possible due to heterogeneity in job characteristics within occupation codes. Nevertheless, it is surprising that we do not observe cancer survivors switching to less demanding jobs more than the control group because cancer is expected to have negative effects on skill endowments and ability to work. This suggests that opportunities for occupational adjustment in the labour market are limited for individuals who have experienced a negative health shock. There may therefore be a mismatch between individual skill endowments and skill demands in the occupation, and greater transition to unemployment and non-participation, including early retirement and disability pension. Policies promoting opportunities for more flexible adjustment to less demanding jobs after a reduction in ability to work (e.g., through increased rehabilitation efforts and firm incentives to provide such jobs) might therefore reduce the negative employment effects after cancer and other health shocks. In Denmark and elsewhere, lack of opportunities for occupational adjustment after health shocks are likely to have more serious consequences in the future, due to aging populations and increases in the age of retirement.

We find that the overall effect of cancer is to reduce the probability of being employed in the fourth year after diagnosis by about 7 percentage points (for both males and females) and to increase the probability of receiving disability pension by 5–6 percentage points. Cancer reduces earnings and wages by about 10% and reduces gross income by about 3%. We also find significant educational gradients in the effect of cancer on labour market status, especially for males. Thus, for low-educated males, cancer reduces the probability of being employed by 10 percentage points, whereas the effect for high-educated males is 5 percentage points. The corresponding values for females are 8 and 5 percentage points. The educational gradient is even more pronounced for the risk of disability pension; the effects for low- and high-educated males are 10 and 4 percentage points, respectively, and for females, they are 8 and 3 percentage points. These results are largely consistent with the estimates of effects of breast and colorectal cancer 3 years after diagnosis by Heinesen and Kolodziejczyk (2013). For females, this is perhaps not surprising since breast and colorectal cancer make up about 60% of the female cancer survivor sample, but for males, it is rather surprising given that colorectal cancer represents only about 17% of the total number of male cancer survivors studied in the present paper. Our estimates of employment effects are also largely consistent with the average effects of cancer in the US found by Moran et al. (2011) and Short et al. (2008).

Our estimates of educational gradients in the effect of cancer on employment are reduced considerably when interaction terms between cancer and job characteristics at t-2 are included in the model, especially for females. This indicates that pre-cancer job characteristics are important in explaining the observation that the effects of cancer on labour market status are much larger for the low-educated than for the high-educated. One reason why neg-

[23] There are no significant differences for males. For females, the estimates are about 20% smaller in Table B7 than in Table 6, and these differences are statistically significant; but, as explained above, the samples differ according to whether non-employed in t-2 are included, and covariates are lagged an extra year in the analysis of Table B7. We should not therefore expect exactly the same cancer effect estimates for t+4.

ative effects of cancer on labour market outcomes are larger for the low-educated may be that they are more obliged to accept jobs characterised by high requirements of physical strength (and low requirements of analytical and other cognitive skills) and that cancer and its treatment often reduces physical strength. Training programmes targeting workers who had specialised in physically demanding jobs before they were diagnosed with cancer might therefore be effective in providing opportunities to find other types of jobs.

We find that non-localised cancers have stronger negative effects on employment, especially for less-educated workers and workers with pre-cancer jobs requiring high levels of manual skills. Policies to improve early detection, for instance through higher participation in screening programmes, might therefore increase not only survival rates but also employment rates for survivors. The gradients we find and the fact that earlier studies find lower screening participation rates for the less educated (Jensen et al., 2012) indicate that such policies might be especially effective if targeted at these groups.

Acknowledgements

We are grateful to Finn Diderichsen, Jakob Bjørner, participants at the 2016 EEA-ESEM, 2017 EALE, 2017 IHEA, 2017 AHES, 2017 AASLE, and 2018 AEA conferences, seminar and workshop participants, anonymous reviewers and the editor Maarten Lindeboom for helpful comments and suggestions, and to Mikkel Flyger Andersen for excellent research assistance. This research was supported by the Rockwool Foundation (grant number 1137).

Appendix A. Additional data description and estimation results

Additional data description

A1. Sample selection and survival rates

A2. Job characteristics variables

This Appendix specifies how we construct job characteristics variables for Danish occupations using the US O*NET database. This involves crosswalks between various classifications of occupations, selection of types of job characteristics and specific O*NET variables, and factor analyses using these variables. We first explain the crosswalks.

A2.1. Crosswalks between classifications of occupations

In the O*NET database, job characteristics are linked to American occupation codes (SOC codes). To utilise these data on job characteristics, we link the SOC codes to Danish occupation codes, known as DISCO codes, which are the Danish version of International Standard Classification of Occupations (ISCO) codes established by the International Labour Organization (ILO). We use the most recent version of the O*NET data, which is based on the American 2010 SOC occupation codes. The most recent version of the DISCO codes is DISCO08, based on the international ISCO08 classification developed by ILO in 2008. DISCO08 codes have been used by Statistics Denmark from 2010.

We first construct a crosswalk between the most detailed 2010 SOC codes and the most detailed DISCO08 codes. A crosswalk is not one-to-one: what we need is a 'one-way' crosswalk which allows us to reassign a unique occupation code of one system to each code in the other system. A given code in DISCO08 sometimes corresponds to many codes in SOC, but this does not present a problem. When, however, DISCO08 has more detailed codes corresponding to a particular SOC code, so that there is no obvious unique DISCO08 code to assign to the SOC code, we look at the number of Danish workers in each detailed DISCO08 code and choose the most representative code. The same principle applies to the construction of other crosswalks. Second, in order to link to Danish data before 2010, we construct a crosswalk between the DISCO08 codes and the Danish codes used before 2010. These are known as DISCO88 and are based on the international ISCO88 codes developed by ILO in 1988. ILO has provided a complete link between ISCO88 and ISCO08 codes.[24] To establish the link between the DISCO08 and DISCO88 codes, we construct two additional crosswalks. One of these is between the DISCO08 and ISCO08 codes, and one is between the ISCO88 and DISCO88 codes.

To check and improve the crosswalk, we undertake panel data analysis to study how Statistics Denmark classified jobs in the two systems of classification. We identify all persons working at the same plant in 2009 and 2010 and compare their DISCO88 code in 2009 and DISCO08 code in 2010. Any pair of codes at the most detailed level with more than 300 observations is added to the crosswalk if it is not already included in it unless these are likely to reflect promotions or other forms of job changes at the same plant. This process added 134 extra combinations to the crosswalk. The resulting panel dataset is also used to calculate weights for each DISCO88 code, to be used for calculating job characteristics of DISCO88 codes given the job characteristics of DISC08 codes. These weights are equal to the proportion of observations in the corresponding DISCO08 codes in the crosswalk. For each DISCO88 code, these weights sum to unity.

The crosswalk links only 'unique' codes, i.e. codes at the most detailed level in each classification. Statistics Denmark is often unable to classify individuals' jobs very precisely, and therefore uses more aggregate ('non-unique') codes. For example, while many individuals have 'unique' job codes, 19210, 19220, and 19230, which are valid codes with the description in the documentation, some individuals have a more aggregate job code, 19200. In that case, we know only that these individuals have one of the three occupations. For both DISCO08 and DISCO88 codes, we calculate job characteristics corresponding to non-unique codes as simple averages of the job characteristics of the underlying unique codes.[25] We observe 594 unique DISCO88 codes, 564 unique DISCO08 codes and 877 unique US SOC occupation codes in our data.

A2.2. Construction of job characteristics variables

We construct variables for skill requirements and other job characteristics of occupations using principal component analysis (PCA) of sets of O*NET variables of the levels of the particular skills required in each occupation. We use the five job characteristics variables specified by Imai et al. (2014): (1) analytical skills, (2) interpersonal skills, (3) physical strength, (4) fine motor skills and (5) visual skills. Each of the job characteristics is constructed from different sets of O*NET variables with no overlap so that each set of O*NET variables describes a unique latent aspect of a job. Following Imai et al. (2014), we extract the first principal component for each chosen set of O*NET variables. O*NET variables with loadings above 0.75 are retained in the final PCA score. We conduct the PCA on individual-level data for 2010-12, weighting occupations by their observed frequency in the Danish data using both unique and non-unique DISCO codes. Each PCA is based on the (weighted) correlation matrix of the variables, where the individual O*NET variables are effectively standardised. This standardisation is necessary since they are not all measured at the same scale in the O*NET system. The results of the five-factor analyses are summarised in Tables A3–A7 in Appendix A. We include a few extra O*NET variables which were not used by Imai et al. (2014). We also construct a further variable to describe the extent of customer contact, as the average of two O*NET variables: 4A4a8 (performing

[24] http://www.ilo.org/public/english/bureau/stat/isco/isco08/.

[25] More detailed information on DISCO codes and crosswalks is available from the authors upon request.

Table A1
Selection of the cancer group.

Selection criteria	Sample size
Persons who had their first cancer diagnosis as 29–61–year-olds in 2000–2005 according to the cancer register (any C diagnosis according to ICD10)	62,305
Exclusion of skin cancer cases (C43 and C44 diagnoses)	44,870
Exclusion of persons who, according to the hospitalisation register, had any diagnosis of cancer or benign tumour (C or D00-D49 diagnosis) in a year prior to the year of the first diagnosis in the cancer register	36,547
Exclusion of persons who are not in the basic population registers in the year of diagnosis and 1–5 years before the diagnosis	35,220
Exclusion of persons who are not 30–60 years of age at the beginning of the year of diagnosis	33,617
Exclusion of persons who were not employed (most of the year) 2 years before the year of diagnosis	25,094

Table A2
Cumulative survival rate by gender, year, and treatment status.

	Males			Females		
	Cancer	Control	Difference	Cancer	Control	Difference
End of t (base year)	0.841	0.995	−0.155	0.929	0.998	−0.069
End of t+1	0.662	0.990	−0.329	0.835	0.996	−0.161
End of t+2	0.582	0.986	−0.404	0.782	0.993	−0.212
End of t+3	0.542	0.981	−0.439	0.746	0.991	−0.246
End of t+4	0.512	0.977	−0.465	0.717	0.989	−0.271
End of t+5	0.489	0.972	−0.483	0.696	0.987	−0.290
N	11,764	117,640		13,330	133,300	

Table A3
O*NET variables included and first principal-component loadings: Analytical skills.

O*NET ID	Name	Loading	KMO
1A1b1	Fluency of Ideas	0.914	0.936
1A1b2	Originality	0.886	0.923
1A1b3	Problem Sensitivity	0.883	0.946
1A1b4	Deductive Reasoning	0.964	0.971
1A1b5	Inductive Reasoning	0.908	0.938
1A1b6	Information Ordering	0.926	0.969
1A1b7	Category Flexibility	0.894	0.968
1A1c1	Mathematical Reasoning	0.866	0.938
1A1c2	Number Facility	0.828	0.939
1A1d1	Memorization	0.856	0.942
1A1e1	Speed of Closure	0.858	0.940
1A1e2	Flexibility of Closure	0.789	0.940
1C7b	Analytical Thinking	0.883	0.982
2A1e	Mathematics	0.827	0.958
2A2a	Critical Thinking	0.940	0.971
2A2b	Active Learning	0.950	0.959
2A2c	Learning Strategies	0.836	0.927
2B2i	Complex Problem Solving	0.949	0.981
2C4a	Mathematics	0.777	0.971
4A2b1	Making Decisions and Solving Problems	0.893	0.969
4A2b2	Thinking Creatively	0.830	0.946
4A2b3	Updating and Using Relevant Knowledge	0.903	0.970
Eigenvalue		17.091	
% of variance		77.7	

Note: Variables excluded from original gross list (loading < 0.75): 1A1e3, 1A1f2, 1A1g1, 1A1g2, 1C7a, 2A1f. KMO is the Kaiser-Meyer-Olkin measure of sampling adequacy.

Table A4
O*NET variables included and first principal-component loadings: Interpersonal skills.

O*NET ID	Name	Loading	KMO
1A1a1	Oral Comprehension	0.940	0.938
1A1a2	Written Comprehension	0.959	0.947
1A1a3	Oral Expression	0.956	0.952
1A1a4	Written Expression	0.966	0.933
1A4b4	Speech Recognition	0.853	0.945
1A4b5	Speech Clarity	0.892	0.945
2A1a	Reading Comprehension	0.962	0.950
2A1b	Active Listening	0.950	0.956
2A1c	Writing	0.968	0.938
2A1d	Speaking	0.978	0.955
2B1a	Social Perceptiveness	0.852	0.919
2B1b	Coordination	0.778	0.900
4A4a1	Interpreting the Meaning of Information for Others	0.858	0.951
4A4a2	Communicating with Supervisors, Peers, or Subordinates	0.824	0.936
4A4a3	Communicating with Persons Outside Organization	0.842	0.914

Table A4 (*Continued*)

O*NET ID	Name	Loading	KMO
4A4a4	Establishing and Maintaining Interpersonal Relationships	0.864	0.928
4A4b6	Provide Consultation and Advice to Others	0.806	0.935
4C1a2h	Electronic Mail	0.871	0.951
Eigenvalue		14.505	
% of variance		80.6	

Note: Variables excluded from original gross list (loading <0.75): 1C2b, 1C3a, 1C3b, 1C3c, 4A4a5, 4A4a6, 4A4a7, 4A4a8, 4A4b1, 4A4b2, 4A4b3, 4A4b4, 4A4b5, 4C1a2c, 4C1a2f, 4C1a2j, 4C1a2l, 4C1b1e, 4C1b1f, 4C1b1g. KMO is the Kaiser-Meyer-Olkin measure of sampling adequacy.

Table A5
O*NET variables included and first principal-component loadings: Physical strength.

O*NET ID	Name	Loading	KMO
1A3a1	Static Strength	0.952	0.960
1A3a3	Dynamic Strength	0.953	0.960
1A3a4	Trunk Strength	0.954	0.962
1A3b1	Stamina	0.967	0.956
1A3c1	Extent Flexibility	0.961	0.974
1A3c3	Gross Body Coordination	0.966	0.949
1A3c4	Gross Body Equilibrium	0.924	0.961
4A3a1	Performing General Physical Activities	0.917	0.914
4A3a2	Handling and Moving Objects	0.907	0.936
4C2d1a	Spend Time Sitting	−0.887	0.894
4C2d1b	Spend Time Standing	0.860	0.901
4C2d1d	Spend Time Walking and Running	0.901	0.947
4C2d1e	Spend Time Kneeling, Crouching, Stooping, or Crawling	0.892	0.943
4C2d1f	Spend Time Keeping or Regaining Balance	0.876	0.934
4C2d1h	Spend Time Bending or Twisting the Body	0.939	0.951
Eigenvalue		12.816	
% of variance		85.4	

Note: Variables excluded from original gross list (loading < 0.75): 1A3a2, 4C2d1c, 4A3a3, 4C2d1g, 4C2d1i, 1A3c2, 4A3a4. KMO is the Kaiser-Meyer-Olkin measure of sampling adequacy.

Table A6
O*NET variables included and first principal-component loadings: Fine motor skills.

O*NET ID	Name	Loading	KMO
1A2a1	Arm-Hand Steadiness	0.877	0.860
1A2a2	Manual Dexterity	0.931	0.879
1A2a3	Finger Dexterity	0.795	0.917
1A2b1	Control Precision	0.943	0.950
1A2b2	Multi-limb Coordination	0.925	0.927
1A2b3	Response Orientation	0.922	0.906
1A2b4	Rate Control	0.914	0.951
1A2c1	Reaction Time	0.928	0.909
1A2c2	Wrist-Finger Speed	0.860	0.961
1A2c3	Speed of Limb Movement	0.864	0.916
4A3a3	Controlling Machines and Processes	0.883	0.927
4C2d1g	Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls	0.840	0.928
Eigenvalue		9.531	
% of variance		79.4	

Note: Variables excluded from original gross list (loading < 0.75): 4C2d1i. KMO is the Kaiser-Meyer-Olkin measure of sampling adequacy.

Table A7
O*NET variables included and first principal-component loadings: Visual skills.

O*NET ID	Name	Loading	KMO
1A1f1	Spatial Orientation	0.979	0.910
1A4a4	Night Vision	0.977	0.824
1A4a5	Peripheral Vision	0.973	0.844
1A4a6	Depth Perception	0.856	0.918
1A4a7	Glare Sensitivity	0.966	0.949
Eigenvalue		4.525	
% of variance		90.5	

Note: Variables excluded from original gross list (loading < 0.75): 1A1e1, 1A1e2, 1A1e3, 1A1f2, 1A4a1, 1A4a2, 1A4a3. KMO is the Kaiser-Meyer-Olkin measure of sampling adequacy.

for or working directly with the public) and 4C1b1f (dealing with external customers). Our six job characteristics variables are standardised so as to have zero mean and unit standard deviation using the individual-level data for 2010–12.

To verify the validity and robustness of this approach, we construct three variables wich correspond to those used in Ottaviano et al. (2013): cognitive_intensity, communication_intensity and manual_intensity. These variables are based largely on the same

Table A8
The means of job-skill requirements in *t*-2 by gender, education, and treatment group.

	Males			Females		
	Compulsory	Vocational	Further	Compulsory	Vocational	Further
	Cancer group					
Analytical	−0.503	−0.211	0.746	−0.757	−0.392	0.558
Interpersonal	−0.684	−0.427	0.691	−0.682	−0.214	0.654
Strength	0.591	0.422	−0.489	0.304	−0.173	−0.279
Fine motor	0.923	0.697	−0.308	0.130	−0.261	−0.404
Visual	1.120	0.747	−0.005	−0.055	−0.390	−0.523
Customer	−0.498	−0.413	0.114	−0.378	−0.050	0.467
	Control group					
Analytical	−0.452	−0.182	0.750	−0.762	−0.392	0.548
Interpersonal	−0.625	−0.391	0.696	−0.690	−0.229	0.636
Strength	0.539	0.383	−0.493	0.322	−0.150	−0.267
Fine motor	0.871	0.646	−0.319	0.135	−0.239	−0.414
Visual	1.071	0.704	−0.028	−0.053	−0.384	−0.510
Customer	−0.494	−0.379	0.124	−0.393	−0.052	0.439

Note: Educational levels: Compulsory schooling involves 9 years of primary and lower secondary school; vocational education programmes are upper-secondary programmes of 3–4 years; further education refers to academic upper-secondary school and post-secondary degrees.

Table A9
Correlation coefficients of job characteristics in *t*-2 by gender.

	Analytical	Interpersonal	Strength	Fine motor	Visual	Customer
	Males					
Analytical	1.00					
Interpersonal	0.93	1.00				
Strength	−0.70	−0.83	1.00			
Fine motor	−0.55	−0.76	0.86	1.00		
Visual	−0.33	−0.50	0.70	0.83	1.00	
Customer	0.40	0.59	−0.43	−0.53	−0.23	1.00
	Females					
Analytical	1.00					
Interpersonal	0.94	1.00				
Strength	−0.41	−0.57	1.00			
Fine motor	−0.29	−0.50	0.73	1.00		
Visual	−0.21	−0.37	0.58	0.69	1.00	
Customer	0.49	0.61	−0.36	−0.40	−0.39	1.00

Table A10
Means of covariates by gender and treatment status, given survival to the end of year *t*+5, and *t*-test statistics for equality of means across cancer and control groups, and normalised differences (ND).

	Males					Females				
	Cancer	Control	Diff.	t	ND	Cancer	Control	Diff.	t	ND
Cancer	1.000	0.000	1.000	.		1.000	0.000	1.000	.	
Manual skills *t*-2	0.246	0.258	−0.012	−0.94	−0.01	−0.153	−0.120	−0.033***	−3.92	−0.03
Cognitive skills *t*-2	−0.022	−0.038	0.016	1.30	0.01	−0.049	−0.105	0.055***	5.91	0.04
Analytical skills *t*-2	0.048	0.032	0.017	1.35	0.01	−0.106	−0.157	0.051***	5.40	0.04
Interpersonal skills *t*-2	−0.092	−0.108	0.016	1.21	0.01	0.007	−0.052	0.059***	6.25	0.05
Physical strength *t*-2	0.127	0.143	−0.015	−1.21	−0.01	−0.102	−0.057	−0.045***	−4.80	−0.04
Fine motor skills *t*-2	0.365	0.374	−0.009	−0.63	−0.01	−0.205	−0.184	−0.020*	−2.39	−0.02
Visual skills *t*-2	0.514	0.523	−0.009	−0.62	−0.01	−0.334	−0.315	−0.019**	−2.97	−0.02
Customer contact *t*-2	−0.229	−0.232	0.004	0.28	0.00	0.061	0.023	0.038***	4.28	0.03
Missing job skills *t*-2	0.082	0.078	0.004	1.07	0.01	0.068	0.065	0.002	0.92	0.01
Compulsory education	0.242	0.258	−0.016**	−2.69	−0.03	0.263	0.284	−0.021***	−4.39	−0.03
Vocational education	0.449	0.444	0.004	0.65	0.01	0.375	0.379	−0.004	−0.75	−0.01
Further education	0.219	0.217	0.003	0.45	0.00	0.305	0.288	0.017***	3.37	0.03
Higher education	0.089	0.081	0.009*	2.26	0.02	0.057	0.049	0.008**	3.23	0.03
Age	50.057	51.433	−1.376***	−12.16	−0.12	48.892	49.608	−0.717***	−8.69	−0.07
Base year 2000	0.138	0.153	−0.015**	−3.23	−0.03	0.154	0.157	−0.003	−0.77	−0.01
Base year 2001	0.157	0.161	−0.004	−0.90	−0.01	0.168	0.170	−0.002	−0.55	0.00
Base year 2002	0.161	0.167	−0.006	−1.20	−0.01	0.166	0.165	0.001	0.27	0.00
Base year 2003	0.170	0.173	−0.003	−0.68	−0.01	0.170	0.172	−0.002	−0.52	0.00
Base year 2004	0.187	0.174	0.013*	2.43	0.02	0.170	0.169	0.001	0.21	0.00
Base year 2005	0.187	0.171	0.016**	3.08	0.03	0.172	0.167	0.005	1.32	0.01
Married	0.695	0.708	−0.013*	−2.06	−0.02	0.677	0.690	−0.013*	−2.57	−0.02
Cohabiting	0.106	0.101	0.005	1.21	0.01	0.103	0.092	0.011***	3.30	0.03
Living alone	0.199	0.192	0.008	1.44	0.01	0.220	0.218	0.002	0.49	0.00
Copenhagen City	0.091	0.077	0.014***	3.56	0.04	0.094	0.082	0.012***	3.88	0.03
Copenhagen County	0.123	0.117	0.006	1.39	0.01	0.133	0.124	0.009*	2.46	0.02
Frederiksborg County	0.079	0.077	0.002	0.64	0.01	0.082	0.082	0.000	0.12	0.00
Roskilde County	0.056	0.051	0.005	1.57	0.02	0.052	0.052	0.000	0.01	0.00
West Zealand County	0.049	0.058	−0.009**	−3.01	−0.03	0.061	0.056	0.006*	2.21	0.02

Table A10 (*Continued*)

	Males					Females				
	Cancer	Control	Diff.	t	ND	Cancer	Control	Diff.	t	ND
Storstrøm County	0.045	0.051	−0.006*	−1.98	−0.02	0.048	0.052	−0.004	−1.53	−0.01
Bornholm County	0.008	0.008	−0.001	−0.57	−0.01	0.009	0.008	0.002+	1.73	0.01
Funen County	0.082	0.088	−0.006	−1.62	−0.02	0.090	0.088	0.002	0.72	0.01
South Jutland County	0.048	0.050	−0.002	−0.61	−0.01	0.048	0.047	0.001	0.23	0.00
Ribe County	0.050	0.042	0.008**	2.59	0.03	0.042	0.041	0.001	0.35	0.00
Vejle County	0.062	0.068	−0.006+	−1.77	−0.02	0.062	0.064	−0.002	−0.89	−0.01
Ringkøbing County	0.057	0.054	0.003	0.94	0.01	0.042	0.052	−0.010***	−4.51	−0.03
Aarhus County	0.120	0.118	0.002	0.54	0.01	0.117	0.123	−0.006+	−1.67	−0.01
Viborg County	0.040	0.046	−0.005+	−1.95	−0.02	0.040	0.043	−0.002	−1.15	−0.01
North Jutland County	0.089	0.095	−0.006	−1.62	−0.02	0.079	0.087	−0.009**	−2.95	−0.02
Hosp. *t*-2/*t*-5: infectious disease	0.014	0.010	0.004**	2.76	0.03	0.010	0.009	0.001	1.06	0.01
Hosp. *t*-2/*t*-5: endocrine disease	0.016	0.016	−0.001	−0.43	0.00	0.026	0.022	0.004*	2.14	0.02
Hosp. *t*-2/*t*-5: mental disorders	0.009	0.005	0.004**	2.86	0.03	0.005	0.004	0.001	1.54	0.01
Hosp. *t*-2/*t*-5: nervous system	0.019	0.017	0.002	1.31	0.01	0.017	0.017	−0.000	−0.03	0.00
Hosp. *t*-2/*t*-5: eye diseases	0.025	0.020	0.005*	2.40	0.02	0.018	0.017	0.000	0.10	0.00
Hosp. *t*-2/*t*-5: ear diseases	0.023	0.025	−0.002	−0.87	−0.01	0.016	0.016	−0.000	−0.03	0.00
Hosp. *t*-2/*t*-5: circulatory system	0.065	0.061	0.004	1.28	0.01	0.044	0.043	0.001	0.38	0.00
Hosp. *t*-2/*t*-5: respiratory system	0.027	0.022	0.005*	2.08	0.02	0.023	0.021	0.002	1.13	0.01
Hosp. *t*-2/*t*-5: digestive	0.079	0.067	0.012***	3.38	0.03	0.052	0.047	0.005*	2.04	0.02
Hosp. *t*-2/*t*-5: skin	0.023	0.018	0.005**	2.64	0.03	0.022	0.018	0.004*	2.46	0.02
Hosp. *t*-2/*t*-5: musculoskeletal system	0.109	0.096	0.013**	3.03	0.03	0.101	0.092	0.009**	2.88	0.02
Hosp. *t*-2/*t*-5: genitourinary system	0.041	0.028	0.013***	4.88	0.05	0.087	0.075	0.012***	4.06	0.03
Hosp. *t*-2/*t*-5: childbirth						0.078	0.068	0.010***	3.51	0.03
Hosp. *t*-2/*t*-5: ill-defined conditions	0.058	0.046	0.012***	3.67	0.04	0.050	0.047	0.004	1.56	0.01
Hosp. *t*-2/*t*-5: injury, poisoning, etc	0.302	0.281	0.020**	3.27	0.03	0.221	0.218	0.003	0.68	0.01
Hosp. *t*-2/*t*-5: other diseases	0.007	0.005	0.002+	1.65	0.02	0.006	0.006	0.000	0.46	0.00
Drugs *t*-2/*t*-5: blood pressure	0.057	0.049	0.009**	2.72	0.03	0.059	0.052	0.006*	2.54	0.02
Drugs *t*-2/*t*-5: heart disease	0.028	0.026	0.002	0.86	0.01	0.013	0.013	0.000	0.03	0.00
Drugs *t*-2/*t*-5: rheumatism	0.415	0.387	0.028***	4.24	0.04	0.424	0.410	0.014*	2.57	0.02
Drugs *t*-2/*t*-5: hormonal preparation	0.081	0.074	0.007+	1.91	0.02	0.085	0.082	0.003	1.02	0.01
Drugs *t*-2/*t*-5: diabetes	0.023	0.023	0.000	0.12	0.00	0.010	0.011	−0.001	−0.77	−0.01
Drugs *t*-2/*t*-5: antipsychotic	0.015	0.013	0.003	1.60	0.02	0.020	0.018	0.001	0.87	0.01
Drugs *t*-2/*t*-5: antidepressive	0.059	0.051	0.008**	2.65	0.03	0.086	0.081	0.005+	1.65	0.01
Drugs *t*-2/*t*-5: analgesic	0.100	0.091	0.009*	2.18	0.02	0.100	0.095	0.006+	1.72	0.01
Drugs *t*-2/*t*-5: hypertension	0.154	0.137	0.017***	3.57	0.03	0.189	0.182	0.007+	1.67	0.01
Drugs *t*-2/*t*-5: parkinson	0.002	0.002	−0.001	−1.38	−0.01	0.005	0.005	0.001	1.22	0.01
Drugs *t*-2/*t*-5: asthma	0.063	0.056	0.007*	2.08	0.02	0.083	0.082	0.001	0.26	0.00
Drugs *t*-2/*t*-5: bronchitis	0.044	0.039	0.005+	1.86	0.02	0.058	0.057	0.001	0.29	0.00
Drugs *t*-2/*t*-5: thyroid	0.004	0.004	0.000	0.31	0.00	0.020	0.021	−0.001	−0.89	−0.01
Drugs *t*-2/*t*-5: antithyroid	0.002	0.002	0.000	0.41	0.00	0.010	0.009	0.001	1.06	0.01
Drugs *t*-2/*t*-5: epilepsia	0.010	0.009	0.001	0.82	0.01	0.010	0.010	−0.000	−0.42	0.00
Drugs *t*-2/*t*-5: anxiolytic	0.075	0.067	0.008*	2.33	0.02	0.122	0.116	0.006	1.57	0.01
Drugs *t*-2/*t*-5: acid disorders	0.116	0.100	0.016***	3.71	0.04	0.104	0.097	0.008*	2.34	0.02
Drugs *t*-2/*t*-5: headaches	0.040	0.039	0.001	0.33	0.00	0.056	0.056	0.001	0.29	0.00
Drugs *t*-2/*t*-5: acid antimigraine	0.014	0.015	−0.001	−0.72	−0.01	0.065	0.066	−0.002	−0.60	0.00
Drugs *t*-2/*t*-5: alcohol dependence	0.020	0.017	0.003	1.60	0.02	0.009	0.006	0.002*	2.42	0.02
#Contacts dentist/10 *t*-2	0.125	0.123	0.002	1.52	0.01	0.140	0.137	0.003**	2.94	0.02
#Contacts specialist doctors/10 *t*-2	0.206	0.196	0.010	0.99	0.01	0.376	0.355	0.020	1.58	0.01
#Contacts GPs/10 *t*-2	0.421	0.363	0.057***	7.97	0.08	0.594	0.561	0.032***	4.77	0.04
Agriculture etc. *t*-2	0.036	0.044	−0.008**	−3.24	−0.03	0.013	0.015	−0.003*	−2.17	−0.02
Energy and water supply *t*-2	0.010	0.011	−0.001	−0.69	−0.01	0.002	0.003	−0.001	−1.62	−0.01
Construction *t*-2	0.095	0.101	−0.006	−1.58	−0.01	0.013	0.013	−0.000	−0.37	0.00
Trade, hotels, restaurants, etc. *t*-2	0.143	0.143	0.000	0.04	0.00	0.114	0.109	0.005	1.41	0.01
Infrastructure *t*-2	0.095	0.096	−0.000	−0.06	0.00	0.037	0.035	0.002	1.03	0.01
Financial services *t*-2	0.133	0.123	0.010*	2.26	0.02	0.119	0.117	0.002	0.48	0.00
Public and personal services *t*-2	0.247	0.245	0.002	0.33	0.00	0.570	0.572	−0.002	−0.42	0.00
Unknown occupation	0.005	0.004	0.001	0.75	0.01	0.004	0.004	−0.000	−0.63	0.00
No job end of November	0.023	0.024	−0.002	−0.79	−0.01	0.027	0.027	0.001	0.30	0.00
Manufacturing *t*-2	0.214	0.210	0.004	0.74	0.01	0.101	0.104	−0.002	−0.77	−0.01
Not employed *t*-3	0.027	0.026	0.001	0.28	0.00	0.037	0.042	−0.004*	−2.17	−0.02
Not employed *t*-5	0.046	0.042	0.004	1.44	0.01	0.066	0.072	−0.005*	−2.05	−0.02
Log(earnings) *t*-2	12.492	12.486	0.006	0.37	0.00	12.226	12.185	0.041***	3.67	0.03
Log(earnings) *t*-5	12.202	12.266	−0.064*	−2.23	−0.02	11.779	11.684	0.096***	3.88	0.03
Degree of unemployment *t*-2	0.018	0.019	−0.001	−0.98	−0.01	0.024	0.024	−0.000	−0.19	0.00
Degree of unemployment *t*-5	0.038	0.037	0.000	0.22	0.00	0.049	0.052	−0.003*	−1.98	−0.01
Some unemployment *t*-2	0.108	0.111	−0.003	−0.65	−0.01	0.143	0.145	−0.002	−0.57	0.00
Some unemployment *t*-5	0.134	0.132	0.002	0.46	0.00	0.177	0.184	−0.007+	−1.71	−0.01
Full-timer *t*-2	0.990	0.989	0.001	1.12	0.01	0.927	0.917	0.010***	3.65	0.03
Full-timer *t*-5	0.946	0.950	−0.004	−1.39	−0.01	0.855	0.840	0.015***	4.01	0.03
No child at age 30						0.518	0.533	−0.014**	−2.70	−0.02
N	5750	114,322				9282	131,521			

+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

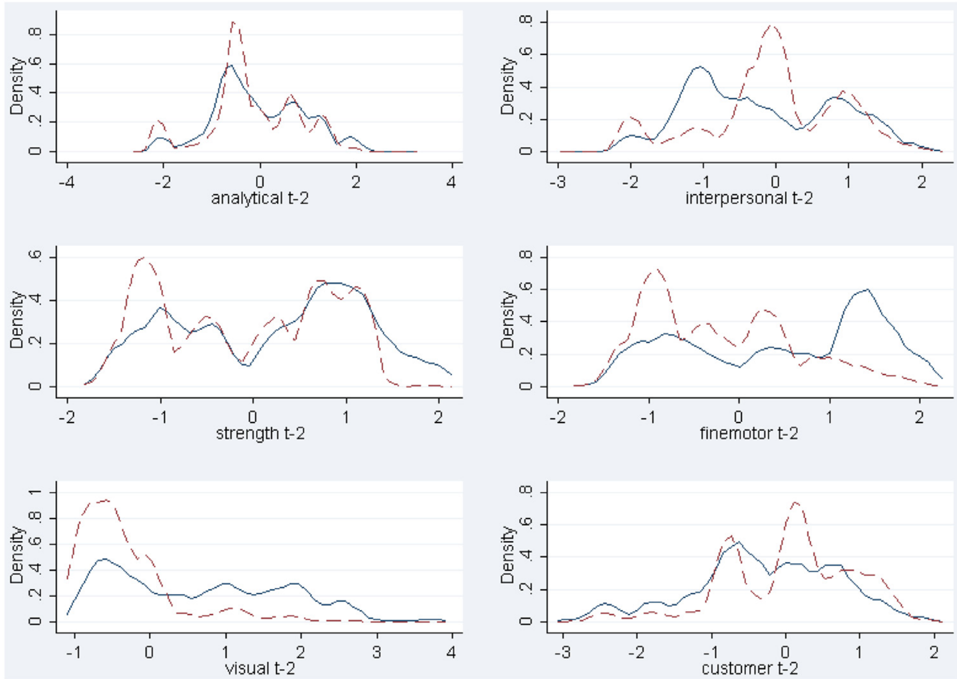

Fig. A1. Density functions for the job characteristics variables for the cancer group in *t*-2 for males (solid line) and females (dashed line).

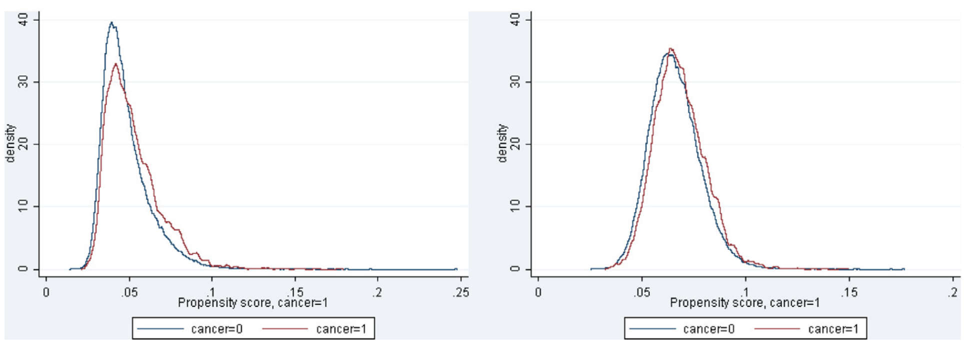

Fig. A2. Density functions for estimated propensity scores (the probabilities of being diagnosed with cancer) by treatment status: Males (left panel) and females (right panel).

O*NET variables as our own variables for (respectively) analytical skills, interpersonal skills and physical strength/fine motor skills; the two sets of variables are highly correlated. We confirm that our results would not be significantly different if we used the three variables of Ottaviano et al. (2013) instead .

A3. Variables for occupational and industry mobility

The occupation code provided by Statistics Denmark indicates the main occupation in that year (i.e., the occupation at the workplace at which the person received the most salaried income during the year; from 2009 onwards, it was the workplace where he/she worked the most hours). Statistics Denmark uses the DISCO classification, which is the Danish version of the International Standard Classification of Occupations (ISCO). There are 594 unique DISCO codes at the 6-digit level. Only codes at the 4-digit level are comparable over time, however, and "the same occupation in *t*+4 as in *t*-2" is therefore defined at the 4-digit level, where there are 377 different values. DISCO codes are reported to Statistics Denmark by firms (or assigned by Statistics Denmark wherever possible if they are not reported). There are no legal consequences for firms that do not report DISCO codes, and often they fail to report, especially for the first year(s) of a worker-firm-spell. As a result, we replace a missing DISCO code for a worker in year *t* by the DISCO code for the worker in *t*+1 if the worker was at the same plant in *t* and *t*+1.

The variable "same industry in *t*+4 as in *t*-2" is based on the 111-grouping of industries provided by Statistics Denmark. From 2008 onwards, Statistics Denmark used a 127-grouping instead, which affects observations for the last two of our six base years (2004 and 2005). For the majority of industries, there is a unique link from the 127-grouping to the 111-grouping. For the last two base years, we define "the same industry in *t*+4 as in *t*-2" to mean that the industry of the 127-grouping in *t*+4 is connected to the industry of the 111-grouping in *t*-2 according to the general crosswalk between the two groupings. This will result in a small downward bias in industry mobility (or upward bias in immobility), presumably to the same extent for both the cancer and control groups.

A4. Control variables

Appendix B. Additional estimation results

Table B1
Effects of cancer on the probability of full-time employment and non-participation in t+4.

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full−time employment							
Cancer	−0.071***	−0.070***	−0.104***	−0.095***	−0.061***	−0.063***	−0.070***	−0.066***
	(0.005)	(0.006)	(0.012)	(0.013)	(0.004)	(0.005)	(0.009)	(0.010)
Cancer*cognitive		0.021*		0.016^{+}		0.005		0.004
		(0.009)		(0.009)		(0.006)		(0.007)
Cancer*manual		−0.003		−0.001		−0.011^{+}		−0.011^{+}
		(0.008)		(0.008)		(0.006)		(0.006)
Cancer*vocational			0.032*	0.028^{+}			0.009	0.003
			(0.014)	(0.014)			(0.011)	(0.012)
Cancer*further			0.058***	0.038*			0.015	0.005
			(0.015)	(0.017)			(0.011)	(0.014)
N	120,072	120,072	120,072	120,072	140,803	140,803	140,803	140,803
	Non−participation (out of labour force)							
Cancer	0.077***	0.076***	0.105***	0.098***	0.066***	0.067***	0.084***	0.078***
	(0.005)	(0.005)	(0.011)	(0.012)	(0.004)	(0.004)	(0.008)	(0.009)
Cancer*cognitive		−0.016^{+}		−0.011		−0.012*		−0.005
		(0.008)		(0.009)		(0.005)		(0.006)
Cancer*manual		0.005		0.004		0.014**		0.015**
		(0.008)		(0.008)		(0.005)		(0.005)
Cancer*vocational			−0.027*	−0.024^{+}			−0.012	−0.005
			(0.014)	(0.014)			(0.010)	(0.011)
Cancer*further			−0.050***	−0.034*			−0.038***	−0.024*
			(0.014)	(0.016)			(0.010)	(0.012)
N	120,072	120,072	120,072	120,072	140,803	140,804	140,803	140,803

Note: Regressions include the full set of control variables; see Section 3.4 and the note to Table 4. Heteroskedasticity-robust standard errors in parentheses.
$^{+}$ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B2
Effects of cancer on the probability of being employed in t+4 by gender for the subgroup aged 30–54 in the base year.

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer	−0.074***	−0.071***	−0.111***	−0.100***	−0.068***	−0.070***	−0.093***	−0.085***
	(0.006)	(0.006)	(0.014)	(0.015)	(0.004)	(0.004)	(0.010)	(0.011)
Cancer*cognitive		0.026**		0.021*		0.016**		0.007
		(0.009)		(0.010)		(0.006)		(0.007)
Cancer*manual		−0.003		−0.002		−0.016**		−0.016**
		(0.008)		(0.008)		(0.006)		(0.006)
Cancer*vocational			0.037*	0.032^{+}			0.019	0.011
			(0.017)	(0.017)			(0.012)	(0.013)
Cancer*further			0.066***	0.042*			0.046***	0.029*
			(0.017)	(0.019)			(0.012)	(0.014)
N	65,713	65,713	65,713	65,713	95,840	95,840	95,840	95,840

Note: The regressions include the full set of control variables; see Section 3.4 and the note to Table 4.
Heteroskedasticity-robust standard errors in parentheses.
$^{+}$ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B3
Effects of cancer on the probability of being employed in t+4 for females with breast cancer and for females with other cancers.

	Breast cancer				Other cancers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer	−0.059***	−0.062***	−0.077***	−0.069***	−0.078***	−0.077***	−0.083***	−0.076***
	(0.005)	(0.005)	(0.011)	(0.012)	(0.006)	(0.006)	(0.013)	(0.014)
Cancer*cognitive		0.012^{+}		0.011		0.014^{+}		0.006
		(0.007)		(0.008)		(0.008)		(0.010)
Cancer*manual		−0.019**		−0.019**		−0.006		−0.010
		(0.007)		(0.007)		(0.009)		(0.009)
Cancer*vocational			0.017	0.008			−0.013	−0.020
			(0.014)	(0.014)			(0.016)	(0.017)

Table B3 (*Continued*)

	Breast cancer				Other cancers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer*higher/further			0.030*	0.010			0.031+	0.018
			(0.014)	(0.016)			(0.016)	(0.020)
N	136,770	136,770	136,770	136,770	135,554	135,554	135,554	135,554

Note: The regressions include the full set of control variables; see Section 3.4 and the note to Table 4.
Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B4
Effects of cancer on the probability of being employed in *t*+4 by gender for the subgroup with localised cancer at diagnosis.

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer	−0.055***	−0.053***	−0.068***	−0.060***	−0.044***	−0.042***	−0.056***	−0.045***
	(0.007)	(0.007)	(0.016)	(0.016)	(0.005)	(0.005)	(0.011)	(0.012)
Cancer*cognitive		0.017		0.015		0.020**		0.017*
		(0.012)		(0.012)		(0.007)		(0.008)
Cancer*manual		−0.002		−0.001		0.003		0.003
		(0.011)		(0.011)		(0.007)		(0.007)
Cancer*vocational			0.009	0.007			0.004	−0.001
			(0.019)	(0.019)			(0.014)	(0.015)
Cancer*higher/further			0.031	0.012			0.028*	0.009
			(0.020)	(0.022)			(0.014)	(0.017)
N	117,279	117,279	117,279	117,279	136,435	136,435	136,435	136,435

Note: The regressions include the full set of control variables; see Section 3.4 and the note to Table 4. Here the cancer group is restricted to cancer survivors with solid tumours, for which stage can be assessed using the TNM classification (i.e., we exclude tumours of the central nervous system, lymphomas and leukaemia) and for whom the stage at diagnosis is localised cancer; the number of cancer survivors is 2957 males and 4914 females. Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B5
Effects of cancer on the probability of being employed in *t*+4 by gender for the subgroup with non-localised cancer at diagnosis.

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer	−0.093***	−0.087***	−0.125***	−0.114***	−0.085***	−0.089***	−0.109***	−0.100***
	(0.012)	(0.013)	(0.027)	(0.029)	(0.007)	(0.007)	(0.015)	(0.016)
Cancer*cognitive		0.005		−0.006		0.015+		0.009
		(0.019)		(0.020)		(0.009)		(0.011)
Cancer*manual		−0.023		−0.022		−0.029**		−0.029**
		(0.017)		(0.017)		(0.009)		(0.010)
Cancer*vocational			0.025	0.022			0.019	0.006
			(0.033)	(0.033)			(0.019)	(0.019)
Cancer*higher/further			0.069*	0.054			0.044*	0.021
			(0.033)	(0.039)			(0.018)	(0.022)
N	115,440	115,440	115,440	115,440	134,668	134,668	134,668	134,668

Note: The regressions include the full set of control variables; see Section 3.4 and the note to Table 4. Here the cancer group is restricted to cancer survivors with solid tumours, for which stage can be assessed using the TNM classification (i.e., we exclude tumours of the central nervous system, lymphomas and leukaemia) and for whom the stage at diagnosis is non-localised cancer (i.e. regional or other metastatic spread); the number of cancer survivors is 1118 males and 3147 females. Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B6
Effects of cancer on the probability of being employed in *t*+4 by gender, for the full sample and by educational level, estimated by OLS and IPW (ATT and ATE).

	Males				Females			
	All	Comp.	Voc.	Further	All	Comp.	Voc.	Further
OLS	−0.075***	−0.099***	−0.081***	−0.047***	−0.067***	−0.081***	−0.076***	−0.047***
	(0.005)	(0.012)	(0.008)	(0.008)	(0.004)	(0.009)	(0.007)	(0.006)
IPW, ATT	−0.074***	−0.099***	−0.080***	−0.047***	−0.067***	−0.082***	−0.076***	−0.048***
	(0.005)	(0.012)	(0.008)	(0.008)	(0.004)	(0.009)	(0.007)	(0.006)

Table B6 (*Continued*)

	Males				Females			
	All	Comp.	Voc.	Further	All	Comp.	Voc.	Further
IPW, ATE	−0.078***	−0.098***	−0.083***	−0.050***	−0.067***	−0.078***	−0.076***	−0.045***
	(0.006)	(0.012)	(0.008)	(0.009)	(0.004)	(0.009)	(0.007)	(0.006)
N	120,072	30,895	53,394	35,783	140,803	39,735	53,316	47,752

Note: 'Comp.', 'Voc.' and 'Further' are abbreviations of compulsory, vocational and further/higher education. OLS estimates for the full sample ('All') are identical to the estimates in columns (1) and (5) of Table 4. In the IPW estimations, we use the same set of control variables in the probit propensity score functions as in the OLS regressions. Heteroskedasticity-robust standard errors in parentheses.

+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B7
Effects of cancer on the probability of being employed in *t*+4. OLS regressions with a control group consisting of later cancer patients.

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer	−0.077***	−0.077***	−0.101***	−0.093***	−0.068***	−0.070***	−0.081***	−0.071***
	(0.006)	(0.006)	(0.013)	(0.014)	(0.005)	(0.005)	(0.010)	(0.011)
Cancer*cognitive		0.022*		0.018+		0.013*		0.011
		(0.010)		(0.010)		(0.006)		(0.008)
Cancer*manual		−0.001		0.000		−0.021**		−0.022**
		(0.009)		(0.009)		(0.007)		(0.007)
Cancer*vocational			0.020	0.016			0.008	−0.004
			(0.016)	(0.016)			(0.013)	(0.014)
Cancer*further			0.049**	0.029			0.030*	0.007
			(0.016)	(0.019)			(0.013)	(0.015)
N	18,886	18,886	18,886	18,886	22,575	22,575	22,575	22,575
N treatment group	5750	5750	5750	5750	9282	9282	9282	9282
N control group	13,136	13,136	13,136	13,136	13,293	13,293	13,293	13,293

Note: The regressions include the full set of control variables; see Section 3.4 and the note to Table 4.
Heteroskedasticity-robust standard errors in parentheses.

+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B8
DID estimates of the effect of cancer on earnings and income variables. The dependent variables are the changes from *t*-2 to *t*+4.

Dependent variable	Males			Females		
(change from *t*-2 to *t*+4)	(1)	(2)	(3)	(4)	(5)	(6)
Earnings	−18.140***	−27.602***	−27.062***	−17.805***	−22.732***	−21.990***
	(3.434)	(3.282)	(2.383)	(1.458)	(1.332)	(1.235)
Wages	−19.802***	−27.650***	−26.152***	−15.853***	−20.422***	−20.334***
	(2.675)	(2.516)	(2.156)	(1.324)	(1.199)	(1.198)
Income	1.528	−3.732	−6.633**	−10.726***	−12.623***	−8.398***
	(6.664)	(6.662)	(2.057)	(2.232)	(2.125)	(0.944)
Disposable income	2.743	1.293	−0.974	−5.265***	−6.204***	−3.959***
	(3.851)	(3.857)	(1.477)	(1.344)	(1.284)	(0.623)
Conditional on employment t+4:						
Earnings	−1.434	−8.395*	−8.655***	−8.525***	−10.503***	−9.653***
	(3.936)	(3.901)	(2.512)	(1.308)	(1.286)	(1.119)
Wages	−4.698+	−9.977***	−9.542***	−6.384***	−8.145***	−8.022***
	(2.725)	(2.689)	(2.153)	(1.074)	(1.055)	(1.053)
Income	11.684	7.508	0.637	−10.060***	−10.524***	−5.076***
	(8.964)	(8.986)	(2.464)	(2.823)	(2.513)	(1.070)
Disposable income	8.168	7.913	3.877*	−5.023**	−5.056***	−1.999**
	(5.198)	(5.210)	(1.874)	(1.696)	(1.526)	(0.716)
Hourly wage rate	0.613	−2.080	−1.986	−2.063***	−2.428***	−1.971***
	(1.621)	(1.610)	(1.377)	(0.621)	(0.619)	(0.594)

(1), (4): Unconditional DID estimate.
(2), (5): Control for age and base year.
(3), (6): Control for age and base year, and exclusion of outliers. The excluded outliers are observations with values of the dependent variable exceeding 1000 numerically (i.e. above DKK 1 m (USD 167,000) for earnings, wages, income and disposable income; and above DKK 1000 (USD 167) for the hourly wage rate).
Heteroskedasticity-robust standard errors in parentheses.

+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B9
DID estimates for the effect of cancer for job characteristic variables (conditional on being employed in *t*+4 and non-missing job skills in *t*-2 and *t*+4).

Dependent variable	Males		Females	
(Change from *t*-2 to *t*+4)	(1)	(2)	(3)	(4)
Analytical skills	0.010	−0.001	−0.006	−0.007
	(0.012)	(0.012)	(0.008)	(0.008)
Interpersonal skills	0.012	−0.000	−0.009	−0.011
	(0.011)	(0.011)	(0.007)	(0.007)
Strength	−0.004	0.004	−0.002	−0.001
	(0.011)	(0.011)	(0.007)	(0.007)
Fine motor skills	−0.001	0.005	0.007	0.008
	(0.011)	(0.011)	(0.007)	(0.007)
Visual skills	0.009	0.009	−0.002	−0.002
	(0.013)	(0.013)	(0.006)	(0.006)
Customer contact	−0.010	−0.016	−0.008	−0.009
	(0.013)	(0.013)	(0.008)	(0.008)
Cognitive skills	0.017	0.004	−0.011	−0.012
	(0.011)	(0.011)	(0.007)	(0.007)
Manual skills	−0.009	−0.002	0.006	0.008
	(0.011)	(0.011)	(0.006)	(0.006)

(1), (3): Unconditional DID; the change in skills is regressed on only a constant and the cancer dummy.
(2), (4): DID controlling for age and base year dummy variables.
Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Table B10
Effects of cancer diagnosed in year *t* on employment, earnings and wages in *t*-2, *t*-1 and *t*+4. Extended sample including non-employed in *t*-2.

	(1) Employed *t*-2	(2) Employed *t*-1	(3) Employed *t*+4	(4) Earnings *t*-2	(5) Earnings *t*-1	(6) Earnings *t*+4	(7) Wages *t*-2	(8) Wages *t*−1	(9) Wages *t*+4
	Males								
Cancer	−0.001	−0.001	−0.071***	−0.730	−0.856	−24.524***	0.014	−0.147	−21.316***
	(0.002)	(0.003)	(0.005)	(1.182)	(1.304)	(1.889)	(0.967)	(1.147)	(1.763)
N	156013	156013	156013	152982	152912	152372	155111	155105	155141
N cancer	6843	6843	6843	6695	6685	6672	6786	6786	6806
	Females								
Cancer	0.000	0.001	−0.059***	1.286*	0.774	−16.899***	1.157*	0.893	−15.239***
	(0.002)	(0.002)	(0.004)	(0.604)	(0.704)	(1.006)	(0.572)	(0.681)	(1.000)
N	187909	187909	187909	187105	187111	186899	187862	187850	187823
N cancer	11,775	11,775	11,775	11,731	11,730	11,705	11,772	11,771	11,769

Note: All regressions include the following controls: Education (3 dummies), age (30 dummies), base year (5 dummies), county of residence (14 dummies), family type (2 dummies), hospitalisation in *t*-5 to *t*-3 by type of diagnosis (16 dummies), consumption of selected categories of prescription drugs in *t*-5 to *t*-3 by type of drug (20 dummies), number of contacts with primary health care sector in *t*-3 (3 variables: GPs, specialists and dentists), and the following variables for labour market status and earnings in *t*-5, *t*-4 and *t*-3: not employed (3 dummies), out of labour force (3 dummies), degree of unemployment, some unemployment (3 dummies), log earnings, hourly wage rate, missing wage rate (3 dummies), income, a dummy for no child at age 30 (for females only), and a constant term. The numbers of observations are slightly smaller in the earnings and wage regressions because outliers with values of the dependent variable below zero or above DKK 1 m (USD 167,000) are excluded.
Heteroskedasticity-robust standard errors in parentheses.
+ $p < 0.10$.
** $p < 0.01$.
* $p < 0.05$.
*** $p < 0.001$.

References

Bacolod, M.P., Blum, B.S., 2010. Two sides of the same coin: U.S. 'residual' inequality and the gender gap. J. Hum. Resour. 45 (1), 197–242.
Bradley, C.J., Bednarek, H.L., Neumark, D., 2002a. Breast cancer survival, work, and earnings. J. Health Econ. 21, 757–779.
Bradley, C.J., Bednarek, H.L., Neumark, D., 2002b. Breast cancer and women's labor supply. Health Serv. Res. 37 (5), 1309–1328.
Bradley, C.J., Neumark, D., Bednarek, H.L., Schenk, M., 2005. Short-term effects of breast cancer on labor market attachment: results from a longitudinal study. J. Health Econ. 24, 137–160.
Bradley, C.J., Neumark, D., Luo, Z., Bednarek, H.L., 2007. Employment-contingent health insurance, illness, and labor supply of women: evidence from married women with breast cancer. Health Econ. 16, 719–737.
Cameron, A.C., Trivedi, P.K., 2009. Microeconometrics Using Stata. Stata Press, College Station, Texas.
Candon, D., 2015. The effects of cancer on older workers in the English labour market. Econ. Hum. Biol. 18, 74–84.
Cutler, D.M., 2008. Are we finally winning the war on cancer? J. Econ. Perspect. 22 (4), 3–26.
Datta Gupta, N., Kleinjans, K.J., Larsen, M., 2011. The Effect of an Acute Health Shock on Work Behaviour: Evidence from Different Health Care Regimes. IZA Discussion Paper No. 5843.
Foged, M., Peri, G., 2016. Immigrants' effect on native workers: new analysis on longitudinal data. Am. Econ. J.: Appl. Econ. 8 (2), 1–34.
García-Gómez, P., van Kippersluis, H., O'Donnell, O., van Doorslaer, E., 2013. Long-Term and spillover effects of health shocks on employment and income. J. Hum. Resour. 48 (4), 873–909.
Gjerstorff, M.L., 2011. The Danish cancer registry. Scand. J. Public Health 39 (7), 42–45.
Heinesen, E., Kolodziejczyk, C., 2013. Effects of breast and colorectal cancer on labour market outcomes—average effects and educational gradients. J. Health Econ. 32, 1028–1042.
Hirano, K., Imbens, G.W., 2001. Estimation of causal effects using propensity score weighting: an application to data on right heart catheterization. Health Serv. Outcomes Res. Methodol. 2, 259–278.
Hirano, K., Imbens, G.W., Ridder, G., 2003. Efficient estimation of average treatment effects using the estimated propensity score. Econometrica 71, 1161–1189.
Imai, S., Stacey, D., Warman, C., 2014. From Engineer to Taxi Driver? Language Proficiency and the Occupational Skills of Immigrants Working Paper. University of Technology Sydney.

Imbens, G.W., 2015. Matching methods in practice—three examples. J. Hum. Resour. 50 (2), 373–419.
Ingram, B.F., Neumann, G.R., 2006. The returns to skill. Labour Econ. 13, 35–59.
Jensen, L.F., Pedersen, A.F., Andersen, B., Vedsted, P., 2012. Identifying specific non-attending groups in breast cancer screening—population-based registry study of participation and socio-demography. BMC Cancer 12, 518.
Keane, M.P., 1992. A note on identification in the multinomial probit model. J. Bus. Econ. Stat. 10 (2), 193–200.
Lee, J., Kim, H., 2008. A longitudinal analysis of the impact of health shocks on the wealth of elders. J. Popul. Econ. 21 (1), 217–230.
Long, J.S., Freese, J., 2014. Regression Models for Categorical Dependent Variables Using Stata, Third edition. Stata Press, College Station, Texas.
Lundborg, P., Nilsson, M., Vikström, J., 2015. Heterogeneity in the impact of health shocks on labour outcomes: evidence from Swedish workers. Oxf. Econ. Pap. 67 (3), 715–739.
Lynge, E., Sandegaard, J.L., Rebolj, M., 2011. The Danish national patient Register. Scand. J. Public Health 39 (Suppl. 7), 30–33.
Moran, J.R., Short, P.F., 2014. Does cancer reduce labor market entry? Evidence for prime-age females. Med. Care Res. Rev. 71 (3), 224–242.
Moran, J.R., Short, P.F., Hollenbeak, C.S., 2011. Long-term employment effects of surviving cancer. J. Health Econ. 30, 505–514.
Ottaviano, G.I.P., Peri, G., Wright, G.C., 2013. Immigration, offshoring, and American jobs. Am. Econ. Rev. 103 (5), 1925–1959.
Poletaev, M., Robinson, C., 2008. Human capital specificity: evidence from the dictionary of occupational titles and displaced worker surveys, 1984–2000. J. Labor Econ. 26 (3), 387–420.
Roodman, D., 2011. Fitting fully observed recursive mixed-process models with cmp. Stata J. 11 (2), 159–206.
Short, P.F., Vasey, J.J., Moran, J.R., 2008. Long-term effects of cancer survivorship on the employment of older workers. Health Serv. Res. 43 (1), 193–210.
Spelten, E.R., Sprangers, M.A.G., Verbeek, J.H.A.M., 2002. Factors reported to influence the return to work of cancer survivors: a literature review. Psychooncology 11, 124–131.
Steiner, J.F., Cavender, T.A., Main, D.S., Bradley, C.J., 2004. Assessing the impact of cancer on work outcomes—what are the research needs? Cancer 101 (8), 1703–1711.
Thielen, K., Kolodziejczyk, C., Andersen, I., Heinesen, E., Diderichsen, F., 2015. The impact of cancer stage on socioeconomic differences in employment consequences. PLoS One 10 (6), e0128621.
Tippins, N., Hilton, M.L. (Eds.), 2010. A Database for a Changing Economy: Review of the Occupational Information Network (O*NET). National Research Council, National Academies Press, Washington, D.C.
van de Mheen, H., Stronks, K., Schrijvers, C.T.M., Mackenbach, J.P., 1999. The influence of adult ill health on occupational class mobility and mobility out of and into employment in the Netherlands. Soc. Sci. Med. 49, 509–518.
Williams, R., 2006. Generalized ordered logit/partial proportional odds models for ordinary dependent variables. Stata J. 6 (1), 58–82.
Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data, Second edition. MIT Press, Cambridge, Massachusetts.
Yamaguchi, S., 2012. Tasks and heterogeneous human capital. J. Labor Econ. 30 (1), 1–53.