The medium-term effects of cancer in the English labour market

David Candon*

School of Economics and Geary Institute, University College Dublin, Ireland

Abstract

The continued rise in overall cancer survival rates has ignited a field of research which examines the effect that cancer has on survivors' employment. Previous estimates of the effect of cancer on labour market outcomes, using U.S. data, show a significant reduction in employment and hours of work in the first 6 months after diagnosis. However, this impact has been found to dissipate after 2 years. With data from the English Longitudinal Study of Ageing (ELSA), I estimate the effect of cancer on workers throughout the 2 year period after diagnosis to fully understand the change in working patterns during this time period. Using propensity score matching, I find that in the 2 years after diagnosis, respondents with cancer are 17 percentage points less likely to work and work 5.6 less hours a week when compared to matched, healthy controls. The effects are even larger when restricted to the first 6-month period after diagnosis where respondents with cancer are 20 percentage points less likely to work and work 7.3 less hours. Finally, I find that respondents are less likely to work and work fewer hours in the second 6-month period following diagnosis, suggesting that the negative effects from cancer can persist for longer than the 6 months identified in previous studies. Results are significant at the 1% level. This is the first study to show that, in England at least, the negative effect of cancer on employment persists for as long as 1 year after diagnosis. These results have implications for both employers and employees as it increases the expected time that workers will be absent from work due to illness.

JEL classification: I10, J21, J22

Keywords: Cancer; Employment; Hours worked

*Corresponding author:
Mr D. Candon
Geary Institute
University College Dublin
Belfield
Dublin 4
Ireland
Tel: 00 353 1 7164615
Fax: 00 353 1 7161108
Email: david.candon@ucdconnect.ie

Introduction

Improvements in the screening and treatment of cancer have seen a steady decline in cancer mortality rates over the last decade. In the U.S., there was an annual change of -1.6% in the cancer death rate between 2001 and 2008 (Eheman et al., 2012). From 2001 to 2010, the three year average mortality rates for male and female cancers in the UK fell by 12% and 9% respectively (Cancer Research UK, 2013)[1]. With mortality rates falling, the number of people surviving cancer is increasing. Consequently, this has led to an increase in research into the role that cancer plays in the labour market, since one of the primary areas in which the survivor's life may be affected is in their employment. Doctor and hospital appointments, treatment schedules, and rest and recuperation time often translate into interruptions. When cancer survivors regain their health, they are faced with decisions regarding whether to work, retire or change their working hours if they do return.

There are many areas which have yet to be explored in assessing the effect of cancer on employment and hours of work. So far, analysis has tended to focus on either the short term or the long term. The main results have shown large, negative, statistically significant effects up to 6 months, and smaller, sometimes insignificant results from 2 to 6 years. This suggests that the first two years after diagnosis, or the medium term, could offer a valuable insight into a potential change in behaviour from cancer survivors. By exploring this period in detail, it allows us to identify at what stage cancer survivors are ready to return to their original working pattern.

[1] The three year periods are 1999-2001 and 2008-2010.

According to Bradley et al. (2005a), the first 6 months after diagnosis captures the “shock” of cancer. As mentioned above, this usually involves hospital appointments, treatment schedules, and rest time and, ex ante, the reductions in employment and hours of work that the literature shows are to be expected. Ex ante, it is more difficult to postulate whether labour outcomes would still be affected by this shock at 12, 18 and 24 months after the initial diagnosis. Results from the literature show that after 2 years this “shock” has worn off (Short et al., 2008). Furthermore, there is little evidence to suggest that it is still present at 12 and 18 months (Bradley et al., 2005b, Bradley et al., 2007). It is possible that the length of this “shock” may be dependent on the characteristics of the country. The bulk of the literature on this topic comes from the U.S. Countries with more generous health or social security systems may find that this shock persists longer than 6 months, increasing the need for more research to be conducted in this area. I analyse a sample from England to tackle this question.

Understanding the stage at which workers are ready to return to work benefits numerous parties. Employers can plan around the survivors’ absence from work, reduction in hours of work or complete retirement; survivors can factor in the length of recuperation time before returning to work to prevent further illness or absenteeism. In addition, survivors may need to know how long they must live without their regular income, particularly if their sick/disability payments deviate greatly from this. Governments can also use this information to help decide on the appropriate level of sick/disability payments that survivors should receive. In this paper, I examine the effect of cancer on labour market outcomes up until 2 years after diagnosis as this is the period when treatment and rest is typically completed and survivors are taking their first steps back into the workplace.

In order to examine these effects I use data from the English Longitudinal Study of Ageing (ELSA). Using propensity score matching (PSM), I compare respondents who are diagnosed with cancer to a healthy control group and calculate the effect of cancer on both employment and hours of work. I find that respondents who have been diagnosed at any stage in the last 2 years are 17 percentage points (p <0.01) less likely to work than their non-cancer counterparts. This rises to 20 percentage points (p<0.1) and 23 percentage points (p<0.01) when I focus on respondents in the first or second 6-month period post-diagnosis, respectively. Similarly, respondents diagnosed in the last 2 years work 5.3 (p<0.01) and 5.6 (p<0.01) hours less per week depending on the specification. Again, these numbers rise to 6.8 and 7.3 (p<0.1) in the first 6-month period and to 8.3 (p<0.01) and 7.9 (p<0.01) in the second 6-month period, respectively. These results are robust to numerous different types of matching estimators. The rest of the paper is organised as follows. In Section 2, I discuss the previous work which has been conducted in the area. Sections 3 and 4 contain information on the data and the methodology used. In Section 5, I present the results and discuss their implications in Section 6.

Literature review: the effects of cancer on employment

In the last 20 years, numerous studies have tried to estimate the effects of cancer on work life. While some studies in Europe have tried to examine the effect of cancer on employment by using a non-cancer control group (van Tulder et al., 1994; Joly et al., 1996; Joly et al., 2002; Taskila-Åbrandt et al., 2004), they did not measure changes over time which, according to Steiner et al. (2004), is one the main areas which should be taken into account when studying the effects of cancer on work outcomes. More recently, Torp et al. (2013) used longitudinal data from Statistics Norway to show that there was a significant decline in the employment

rate for female survivors but not for male survivors in Norway. In this case, though, they were examining outcomes 5 years after diagnosis, which is outside the remit of this paper. The strongest and best designed studies, which have used a non-cancer control group as well as longitudinal data, tend to have been carried out in the U.S, where patients from cancer registries are matched with respondents from national or regional surveys.

In the U.S., studies have shown that in the 6 months following diagnosis, breast cancer has a negative effect on the probability of women being employed (Bradley et al., 2005a). The women who remained working also worked almost 7 hours less per week. However, in the 12 and 18 months following diagnosis, the probability of employment was the same as the control group, though hours of work 12 months after diagnosis were significantly less at the 5% level (Bradley et al., 2007). For men with prostate cancer, they were 10 percentage points less likely to work 6 months after diagnosis than men without prostate cancer, but no significant difference in hours of work or employment was observed at 12 and 18 months (Bradley et al., 2005b). Looking at longer-term outcomes, the literature shows no significant differences between cancer survivors and non-cancer controls up to 4-5 years post-diagnosis. For men and women aged 55 to 65, who were on average 46 months post-diagnosis, survivors with no new cancers showed no significant reduction in probability of working or in hours worked (Short et al., 2008). Another long term study by Chirikos et al. (2002), found that women with breast cancer suffered no significant drop in earnings or withdrawals from the labour market 5 years after diagnosis, though they did report they needed special arrangements to keep working. It should be noted that these results have been found for older workers who have the same age profile as the respondents in the ELSA. Although not the focus of this study, it has been shown that younger workers may suffer more from cancer in the long-term than their older counterparts. For example, the aforementioned Norwegian

study by Torp et al. (2013) used respondents with a mean age of 47. Also, men and women between the ages of 28 and 54 who have survived cancer for between 2 and 6 years were less likely to work when compared to a non-cancer control group (Moran et al., 2011). These effects were more pronounced for survivors who had been diagnosed with new cancers. For survivors who had no new cancers, they were still less likely to work but the effects sizes were smaller and the results were not all statistically significant. There was a significant reduction in hours worked, however. It is the aim of this paper to establish if, in England, the negative employment effect from cancer persists past the 6-month period originally reported by Bradley et al. (2005a).

In previous studies, cancer samples have tended to be quite small. Even when using specific cancer registry data, the number of people with cancer included in the analysis is between 50 and 600. Bradley et al. (2002) used data from the Health and Retirement Study (HRS) in the U.S. to investigate the impact of breast cancer survival on work and earnings with a breast cancer sample of 156, though the longitudinal part of the analysis only included 68 of the breast cancer survivors. Chirikos et al. (2002) identified 105 breast cancer survivors from a cancer registry for their study on the long term impact of breast cancer survivorship. Data from the Metropolitan Detroit Cancer Surveillance System (MDCSS) was used by Bradley et al. (2005a) in assessing the short term impact of breast cancer on labour market attachment. This specific cancer registry data provided information on 496 breast cancer survivors. The MDCSS was used again by Bradley et al. (2005b) in their analysis of men with prostate cancer and contained information on 267 patients. Data from the Penn State Cancer Survivor Survey (PSCSS) has been used in both Short et al. (2008) and Moran et al. (2011) where the number of survivors for all cancers were 504 (but only 405 were still cancer free) and 676

(but only 565 were still cancer free) respectively. Issues regarding the types of cancer and the number of observations used in this analysis are addressed in the next section.

Data

For this analysis, I use the English Longitudinal Study of Ageing (ELSA) which is a large, biennial, longitudinal dataset that contains information on the respondents' health, wealth and employment, comparable to the HRS in the U.S. The ELSA, which was started in 2002, is made up of 5 waves, the latest of which was released in October 2012. I first restrict the data to those respondents with complete survey records. People who are retired or permanently sick or disabled are removed from the sample because of the low probability of returning to work in future waves[2]. In order to reduce the effect of extreme values on the results, I restrict the sample to respondents who are between the ages of 48 and 64 in their first wave, who work a maximum of 65 hours per week and whose income is £4000 or less per week.

With regards to cancer, respondents are asked "Thinking about what has happened since we last saw you has a doctor ever told you that you have (or had) any of the conditions on this card?" Cancer (excluding minor skin cancers) is included as one of the conditions. Respondents are excluded if they had cancer in the first wave when they joined the survey as it does not allow us to observe their pre-cancer behaviour. Once respondents have been diagnosed with cancer in a particular wave, they are excluded from future waves, as we are not interested in changes once people have a second cancer. The people who remain in the study are ones who do not have cancer in one wave but have it in a subsequent wave, two years later. For example, if a respondent has no cancer in the first wave and no cancer in the second wave, this is included as an observation. Similarly, if they have no cancer in the

[2] This applies only to the first period. Respondents are allowed become retired or permanently sick or disabled in the second period. This is detailed in the Methodology section.

second wave but do have cancer in the third wave, this is also included as an observation. All other information from the third wave on would be excluded as the respondent now has cancer. Because the waves in both data sets take place in different years, we compress them into two time periods: T_1 and T_2. T_1 is a period where no one has cancer where no one has cancer and is referred to as the baseline. T_2 is a period where some people now have cancer 2 years later[3,4]. This leaves a final sample of 8119 respondents, 110 of which will be survivors in T_2. This sample, while not as large as data taken from the MDCSS or PSCSS is still of the same magnitude of the studies of Bradley et al. (2002) and Chirikos et al. (2002). Table 1 provides a breakdown of the types of cancers in the sample.

Table 1. Cancer type

	Number	%
Breast	33	30.00
Colon, bowel or rectum	17	15.45
Leukaemia	2	1.81
Lymphoma	3	2.73
Lung	3	2.73
Melanoma	14	12.73
Other	38	34.55
Total	110	100.00

The main advantage of using survey data rather than registry data is that we have information on the respondents before they were diagnosed with cancer, rather than from a retrospective interview on their background characteristics and employment conditions when they have cancer. In this analysis, we observe the working behaviour of cancer patients *before* they get cancer. On the one hand, people who have just been diagnosed with cancer, and are not working, may feel a particular sense of loss about their job. This may lead them to

[3] Due to limitations with the sample size I am unable to look at the effects of cancer over a longer period of time.
[4] Having 5 waves offers 4 chances to observe a non-cancer to cancer transition: wave 1 to wave 2; wave 2 to wave 3; wave 3 to wave 4 and wave 4 to wave 5.

overestimate how many hours a week they previously worked before they had cancer. On the other hand, people who have returned to work and feel like they have “beaten” cancer may want to think that their life has returned to normal and so may underestimate how many hours a week they previously worked to bring it in line with how much they work now. Interviewing the respondents’ when they do not have cancer removes any potential for this type of bias.

Another benefit with these types of data is that the cancer group and the non-cancer group are drawn from the same datasets, meaning there can be no issues surrounding the comparability of the two groups. In previous studies, such as Bradley et al. (2005) and Short et al. (2008), the survivors in the cancer registries can often have significant differences from their matched controls, which are drawn from other surveys. Heckman et al. (1999) state that once respondents are comparable people, administered the same questionnaire and their histories are known, much of the bias in using non-experimental methods is attenuated. I show in the Results section that there are minimal differences between the cancer and non-cancer group before diagnosis, strengthening the conclusion that any differences in outcomes observed after diagnosis is due to the cancer.

Methods

I estimate the effect of cancer on the probability of working, and on the hours of work, as a function of whether the respondents have cancer, while controlling for a set of other possible covariates and allowing for random shocks. While using data from the same survey increases the chances of the cancer group and non-cancer group being as similar as possible, the non-

cancer group cannot be considered a true counterfactual for the cancer group because of the possibility of selection effects. If the cancer respondents had not been diagnosed with cancer, it is possible that their work outcomes would still be different from the respondents who do not have cancer. In order to estimate the average effect of cancer on those who are diagnosed with cancer (the average effect of the treatment on the treated (ATT)), the unobserved employment outcomes of those who have cancer, if they did not have cancer, are required. This can be expressed in the following equation

$$ATT = E(Y_1 - Y_0)|D_1 = E(Y_1)|D_1 - E(Y_0)|D_1, \quad (1)$$

Where Y_1 is the outcome when receiving the treatment, Y_0 is the outcome when not receiving the treatment and D_1 signifies that the respondent received the treatment. For the purposes of this study, "treatment" is defined as having cancer. PSM is used to balance the observable characteristics between the samples and eliminate the impact of the observables as confounding factors (Rosenbaum and Rubin, 1983; D'agostino, 1998). The propensity score can be defined by the following equation,

$$p(x) \equiv \Pr(D = 1|X = x), \quad (2)$$

where the probability of getting the treatment (cancer) is based on the vector of X variables. In this case, the vector of variables includes age, whether the respondent has a third level education, their marital status (never married, married and divorced, widowed or separated), whether there is a child living at home, whether the respondent is now in a state of bad health (poor and fair or very poor, poor, and fair depending on which Likert scale was used), their wave in the ELSA and the hours the respondents were working, all measured at baseline. I

also include a variable for equivalised household income (which is adjusted for inflation). These variables are included because even though they may be unrelated to the treatment, they may still influence the outcome variable (Rubin and Thomas, 1996). The first outcome variable I am interested in is E_{it} where E_{i1} = 1 when the respondents are employed or self-employed and E_{i1} = 0 when the respondents are unemployed or looking after the family home. E_{i2} = 1 when the respondents are employed or self-employed and E_{i2} = 0 when the respondents are unemployed, looking after the family home, retired or permanently sick or disabled[5]. The second outcome variable is hours of work, H_{i2}, a continuous variable, bounded at 0. For the third outcome, I use a difference-in-differences approach to examine the effect of cancer on ΔH_i, where $\Delta H_i = H_{i2} - H_{i1}$, in order to control for time invariant unobservable characteristics as well as the baseline observables.

PSM has been routinely used in this type of research and can be found in Bradley et al. (2005), Short et al. (2008) and Moran et al. (2011). In this paper, I follow the method of Moran et al. (2011) and use the PSMATCH2 commands developed by Leuven and Sianesi (2003) for Stata. The most basic form of PSM is 1-to-1 matching where a single treated unit is matched to a single control unit. The benefit of this method is that it minimises bias in the analysis (only the best control unit is matched to the treated unit) but it implies a loss of efficiency (because it involves discarding all other potentially valuable observations) (Steiner and Cook, 2013).

In addition to simple 1-to-1 matching, I also estimate 3 smoothed matching estimators: *k*-nearest neighbour, radius and kernel. They are referred to as smoothing estimators because they use information from adjacent control observations with similar propensity scores and

[5] The reason for not including the retired or permanently sick or disabled in E_{i1} is described in the Data section.

discard others to make a new weighted value for the outcome variables. In *k*-nearest neighbour matching, *k* observations are selected which have a similar propensity score to a treated observation and all observations used are weighted equally. Radius matching is comparable to nearest neighbour, where observations are included if they fall within a given radius of the propensity score. Again, all observations used are weighted equally. For the purposes of this paper, $k = 10$ and the radius is 0.001. Kernel matching is a slightly different form of weighted matching where a distribution is placed around the propensity score under consideration and the weight attached to any observation used is inversely related to the distance from the propensity score (Heckman et al., 1998). With kernel matching, Stata automatically uses the Epanechnikov kernel and, in this case, a bandwidth of 0.001 is selected. In the matched estimates, matching is done without replacement and the max-min common support condition is used, so no observations outside of the maximum and minimum propensity score of the control group are included.

In addition to the matching estimates, I also perform bivariate and multivariate regressions for the use of comparison. Robust standard errors are used in the regression estimates and bootstrapped standard errors, using 500 replications, are calculated for the matched estimates.

Results

Descriptive statistics

The descriptive statistics for the cancer group and non-cancer group, as well as the *t*-tests for the equality of means, are presented in Table 2. In columns 1 and 2, the overall cancer sample is compared to the overall non-cancer sample in the first period of the model. In terms of

employment, there is no significant difference between the groups but there is a significant difference in terms of hours worked (for those who were employed) with the cancer group working 31.38 hours per week in the first period compared to 35.15 in the non-cancer group. In the second period, there is a significant difference in the probability of being employed with 71.69% of the non-cancer group being employed compared to 48.18% of the cancer group. For the respondents who worked in the second period, the non-cancer group worked 29.36 hours per week whereas the cancer group worked 19.42 hours per week. The other significant differences between the cancer and non-cancer group indicate that the cancer group were older, more likely to be in poor health and a higher proportion were female. Columns 3 and 4 compare the cancer group and the non-cancer group for respondents who were employed in the first period. Again, aside from the probability of working and hours worked, the only other significant difference between the groups was that the cancer group were older and more likely in poor health.

Basic univariate analysis shows that the probability of being employed in the second period is significantly lower for both the entire cancer sample and the cancer sample who are employed in the first period. Also, the cancer group work over 3 hours less than the non-cancer group ($p<0.01$) and this difference becomes almost 10 hours ($p<0.01$) once the respondents are diagnosed with cancer.

As I mentioned, using a longitudinal survey which contains the cancer and non-cancer respondents reduces the differences between the two samples due to consistency with relation to the methodology of variable construction, geographic location, the type of respondent being sampled, etc. Unfortunately, as outlined in the previous paragraph, some differences still remain. I use propensity score matching to combat this problem and make the control

sample more comparable. When I match the samples using propensity scores and re-estimate the *t*-tests, there are no significant differences between the two groups.

Table 2. Descriptive statistics

Variable	(1) Cancer group	(2) Non-cancer group	(3) Cancer group, working first period	(4) Non-cancer group, working first period
Female	68.18%**	56.60%	62.50%*	52.07%
Degree	12.73%	13.68%	12.50%	14.93%
Poor health	25.45%***	16.12%	21.59%**	14.05%
Single	4.55%	5.48%	5.68%	5.62%
Married	79.09%	78.05%	75.00%	77.63%
DSW	16.36%	16.47%	19.32%	16.75%
Child at home	33.64%	39.89%	37.50%	40.06%
Household income (weekly, £s)	367.70	369.19	388.46	389.15
Age	57.50***	56.26	57.40***	56.15
Working, first period	80.00%	86.45%	100.00%	100.00%
Hours worked, first period	31.38***	35.05	31.38***	35.05
Working, second period	48.18%***	71.69%	60.23%***	82.93%
Hours worked, second period	19.42***	29.36	19.42***	29.36
Observations	110	8009	88	6924

Note: ***Significantly different from the non-cancer sample at the 1% level (p <0.01 when comparing (1) with (2) and (3) with (4)).
**Significantly different from the non-cancer sample at the 5% level (p <0.05 when comparing (1) with (2) and (3) with (4)).
*Significantly different from the non-cancer sample at the 5% level (p <0.10 when comparing (1) with (2) and (3) with (4)).
DSW: divorced, separated, widowed

Table 3. Effects of cancer in the first 2 years post-diagnosis: Regression and 1-to-1 matching estimates

< 2 years post-diagnosis	Unadjusted mean difference	Regression adjusted mean difference	1-to-1 matching
Working ^	-0.235***	-0.170***	-0.182**
	(0.048)	(0.048)	(0.072)
Hours worked	-9.940***	-5.513***	-4.784*
	(1.849)	(1.565)	(2.455)
Difference in hours worked	-6.268***	-5.513***	-5.023**
	(1.755)	(1.565)	(2.460)
N	7,012	7,012	174
1st 6 months post-diagnosis			
Working ^	-0.262**	-0.205**	-0.333**
	(0.106)	(0.100)	(0.156)
Hours worked	-11.194***	-7.134**	-4.500
	(3.833)	(3.348)	(4.909)
Difference in hours worked	-9.147**	-7.134**	-11.333**
	(4.063)	(3.348)	(5.421)
N	6,942	6,942	36
2nd 6 months post-diagnosis			
Working ^	-0.281***	-0.226***	-0.161
	(0.080)	(0.085)	(0.145)
Hours worked	-12.458***	-7.711***	-9.129**
	(3.116)	(2.561)	(4.212)
Difference in hours worked	-8.281***	-7.711***	-9.097**
	(2.781)	(2.561)	(3.829)
N	6,955	6,955	61
3rd 6 months post-diagnosis			
Working ^	-0.152	-0.067	0.000
	(0.104)	(0.090)	(0.187)
Hours worked	-2.139	0.169	-2.333
	(4.089)	(2.731)	(7.245)
Difference in hours worked	-0.813	0.169	-0.722
	(2.784)	(2.731)	(5.322)
N	6,942	6,942	36
4th 6 months post-diagnosis			
Working ^	-0.217**	-0.148	-0.286**
	(0.098)	(0.103)	(0.144)
Hours worked	-11.837***	-5.765	-12.381**
	(3.484)	(3.572)	(5.054)
Difference in hours worked	-5.504	-5.765	-8.571*
	(4.149)	(3.572)	(4.839)
N	6,945	6,945	42

Note: In the regression adjusted model, the variables included are the respondent's age, average household weekly income, whether there is a child living at home, whether the respondent is in poor health, whether the respondent has a degree, marital status and the time period. For regression estimates, robust standard errors are in parentheses. For matching estimates, bootstrapped standard errors are in parentheses.
*** p <0.01, ** p <0.05, * p <0.10.
^Sample size for Working (unadjusted mean difference): 8,119; 8,031; 8,048; 8,032; 8,035.

Table 4. Effects of cancer in the first 2 years post-diagnosis: Smoothed matching estimates

< 2 years post-diagnosis	Nearest neighbour-10	Radius	Kernel
Working	-0.145**	-0.170***	-0.171***
	(0.057)	(0.053)	(0.053)
Hours worked	-4.977***	-5.371***	-5.334***
	(1.77)	(1.58)	(1.585)
Difference in hours worked	-4.620**	-5.531***	-5.577***
	(1.975)	(1.859)	(1.862)
N	7,012	7,012	7,012
1st 6 months post-diagnosis			
Working	-0.217	-0.201*	-0.201*
	(0.135)	(0.116)	(0.116)
Hours worked	-6.75	-6.847*	-6.825
	(4.754)	(4.144)	(4.173)
Difference in hours worked	-6.467	-7.175*	-7.306*
	(4.395)	(4.088)	(4.037)
N	6,942	6,942	6,942
2nd 6 months post-diagnosis			
Working	-0.194*	-0.234***	-0.231***
	(0.099)	(0.088)	(0.088)
Hours worked	-7.810**	-8.286***	-8.252***
	(3.091)	(2.918)	(2.915)
Difference in hours worked	-7.977***	-7.895***	-7.883***
	(2.975)	(2.785)	(2.771)
N	6,955	6,955	6,955
3rd 6 months post-diagnosis			
Working	-0.050	-0.076	-0.075
	(0.131)	(0.111)	(0.113)
Hours worked	-0.750	-0.562	-0.438
	(5.118)	(4.338)	(4.386)
Difference in hours worked	0.400	-0.432	-0.381
	(3.444)	(2.791)	(2.814)
N	6,942	6,942	6,942
4th 6 months post-diagnosis			
Working	-0.133	-0.149	-0.166
	(0.119)	(0.109)	(0.112)
Hours worked	-6.481	-6.059*	-6.531*
	(3.983)	(3.409)	(3.384)
Difference in hours worked	-6.152	-5.583	-5.969
	(4.615)	(4.165)	(4.157)
N	6,945	6,945	6,945

Note: For matching estimates, bootstrapped standard errors are in parentheses.
*** p <0.01, ** p <0.05, * p <0.10.

Effects of cancer < 2 years post-diagnosis

Table 3 shows the results from the estimation of the effect of cancer on the probability of employment, hours worked and the difference in hours of work. The first column shows the results from simple OLS regressions of the outcome variables on whether the respondent has been diagnosed with cancer. The second column shows an adjusted regression model controlling for baseline characteristics. In the final column, 1-to-1 matching estimates are presented. The top panel of Table 3 shows the effect of cancer if the respondents have been diagnosed anytime in the last 2 years. Cancer reduces the probability of being employed by 23.5 percentage points and hours of work by 9.9 hours. When hours of work in the first period are taken into account, cancer still reduces the weekly hours of work by 6.3 hours in the second period. These estimates can also be observed by looking at the difference in outcomes between the groups in Table 2. When baseline characteristics are included as controls, the effect of cancer on employment is reduced, but is still quite large at -17 percentage points. Similar results are found for the hours of work model, where the effect of cancer is now only -5.5 hours per week. Because hours of work in the first period are included as a control in the hours of work model, the effect of cancer is the same as the difference in hours worked model. All results here are significant at the 1% level. Finally, in the matching model, cancer reduces employment by 18.4 percentage points (p <0.05), hours of work per week by 4.8 hours (p <0.1) and the difference in hours by 5 hours (p <0.05).

Effects of cancer in the four 6-month periods post-diagnosis

We should note, however, that being diagnosed with cancer may result in an immediate reduction in labour supply, followed by a return to normal working patterns after this "shock". This is what Bradley et al. (2005a) find when they investigate employment and hours worked of breast cancer survivors in the first six months after diagnosis. It is possible

that the effect that we observe is being driven by an initial change in working patterns which may only last 6 months. In order to examine this issue in more detail, length of time since diagnosis is considered. The ELSA data contain information on the month when the respondents are interviewed and the month in which they report their diagnosis. This information is used to construct new cancer variables which indicate which 6-month period in the last 2 years that the respondents were diagnosed in (the 1st 6 months, the 2nd 6 months, the 3rd 6 months or the 4th 6 months). I estimate the models again using these different variables for cancer. Separate propensity scores were created for each new cancer variable.

Looking at the unadjusted mean difference, the respondents who are in the 1st 6 months since diagnosis are 26 percentage points less likely to work compared to their non-cancer controls. They also work 11.2 hours less per week and 9 hours less when the first period's hours are taken into consideration. Similar results are found when I look at respondents in their 2nd 6-month period after diagnosis. An interesting pattern that arises from these specifications is that the coefficients in the unadjusted and adjusted models are of the same magnitude. If the "shock" from cancer was confined to the 1st 6-month period, we would expect to see larger coefficients here compared to the 2nd 6-month period. However the values for the coefficients are within 1% or 1 hour of work of one another, suggesting that cancer is having the same effect in both the 1st and 2nd 6-month periods. The matching estimates tell a different story with estimates that are inconsistent with the other results. In this model, the reduction in the probability of working after cancer diagnosis is 33 percentage points in the first 6-month period versus 16 percentage points in the second 6-month period, the reduction in hours of work is 4.5 versus 9.1 and the reduction in the difference of hours of work is 11.3 versus 9.1. This will be discussed further in the next section

When the respondents are in their 3rd 6-month period following diagnosis, cancer has a negligible effect across all the specifications. While cancer still has a negative effect, the effect sizes are now smaller than any other specification and none of the results are statistically significant. In the final model, when some respondents are in their 4th 6-month period following diagnosis, the results prove to be inconclusive. The effect of cancer on the outcomes varies from model to model in terms of the size of the coefficient and its statistical significance.

Smooth matching estimators

The problem with 1-to-1 matching, which I mentioned in the Methodology section, is that while it has the minimum bias of all matching estimators, this comes at a loss of efficiency due to the discarded observations. In every specification in Table 3, the standard errors in the matching model were larger than the standard errors in either of the regression models. In order to use more of the available data, I estimate the model again using the 3 smoothed matching estimators discussed earlier: *k*-nearest neighbour, radius and kernel.

In the first column of Table 4, the results from the nearest neighbour matching are presented. In this model, cancer reduces the probability of employment by 14.5 percentage points and the difference in hours of work by 4.6 hours per week if the respondent is in their first 2 year after diagnosis, both of which are significant at the 5% level. In the second column, where the radius matching estimates are presented, cancer reduces employment by 17 percentage points ($p<0.01$) and the difference in hours of work by 5.5 hours ($p<0.01$).

The size of these effects increases when just looking at the respondents who are in the first 6-month period post-diagnosis. While the results for the nearest neighbour and radius matching

are almost identical, the radius matching estimates are significant at the 10% level. The size of the coefficients increase even further when looking at respondents who are in the second 6-month period following diagnosis (with the exception of employment in the nearest neighbour estimator which falls to 19.4 percentage points) and all the results are significant at 10% level.

However, in situations like this, where there are numerous control observations per treated observation, kernel matching has been shown to work well (Frolich, 2004). While it does use more than one observation to calculate the mean outcome, which increases the bias, the observations are weighted, with higher weights given to controls which are closest to the propensity score, which tries to reduce the bias. This allows the estimator to become more efficient while still trying to reduce the bias.

The third column of Table 4 shows the results from kernel matching. When respondents are diagnosed in the last 2 years, cancer reduces the probability of employment by 17 percentage points ($p<0.01$) and the difference in hours worked by 5.6 hours ($p<0.01$). In the first 6-month period post-diagnosis, the effect of cancer on the probability of employment and the difference in hours worked is -20 percentage points and -5.6 hours, where both estimates are significant at the 5% level. Like the previous matching estimates, the coefficients are larger if the respondents are in the second 6-month period following diagnosis and they are now significant at the 1% level.

Reassuringly, the size of the coefficients and the standard errors, and therefore the significance levels, are similar across the regression adjusted model and the different types of weighted matching estimators.

Discussion

Advances in cancer research have led to large increases in cancer survival rates. As a result, the question of how survivors' future employment outcomes are affected by the trauma of cancer has become increasingly important. This study provides an investigation into the working patterns of English survivors who are in their first 2 years post-diagnosis. In addition, this study highlights the time periods in which the labour outcomes are most affected. While previous studies have used cancer registry data, this paper uses data taken from a study of old age and retirement, ELSA. This avoids the problems associated with registry data, such as the comparability of the cancer and non-cancer sample and any bias stemming from survivors overestimating or underestimating previous working patterns.

I find that respondents who have been diagnosed in the last 2 years are 17 percentage points less likely to work than their non-cancer counterparts. They also work 5.3 less hours per week, 5.6 hours less per week when I examine the change in their hours, and all the results are significant at the 1% level. When the two year period is broken up into the 4 respective 6-month periods, the magnitude of the effect of cancer on work outcomes is larger for the first 6-month period and larger again in the second 6-month period. In this case, cancer reduces the probability of employment by 23 percentage points ($p<0.01$) and hours of work by 8 hours ($p<0.01$).

Many of the findings in this paper are consistent with other estimates of working and hours worked in the literature. Bradley et al. (2005a), Bradley et al. (2005b) and Bradley et al. (2007) have all found that the first 6-month period following cancer diagnosis is associated with a reduced probability of employment and fewer hours worked and I find similar results in this analysis. Due to restrictions with the sample size I do not estimate propensity scores

separately for men and women like Short et al. (2008) and Moran et al. (2011) but I include sex as a characteristic to generate the propensity score. As I wish to examine how the effect of cancer changes over different diagnoses periods, the propensity score were estimated separately for the different cancer diagnoses periods. While I cannot identify the effect of cancer for men or women separately after 12 months, the fact that an overall effect is present is noteworthy considering it was absent for men and women in analysis of prostate and breast cancer. Also, like Short et al. (2008) again, I cannot control for the severity of the cancers. In Bradley et al (2005a), the authors use the stages of the cancer (in situ, local, and regional/distant) to control for different levels of severity in the cancers. For the respondents who have distant metastasis, the most severe form of cancer, they may not be able to respond to the ELSA questionnaires because they are too sick, in hospital, in convalescence etc. If this type of bias is present it would only serve to strengthen the results that are found as a reduction in the probability of employment or hours worked would likely be more if the severe cases had been included.

The main finding from this paper that is not consistent with the literature is the significant negative impact of cancer in the 2nd 6-month period following diagnosis. Bradley et al. (2007) find that for breast and prostate cancer survivors, the negative impact from cancer has disappeared after 6 months. This study provides evidence that, in England at least, the negative impacts from cancer persist throughout the first year following diagnosis, though I do look at all cancers, not just breast or prostate. The results from this study may have important implications for the labour market. It is possible that the “shock” from cancer to working patterns is not confined to the 1st 6 months after diagnosis and may persist for up to 1 year after diagnosis in terms of the probability of the workers being employed and the hours that they work. As such, employers need to be aware that may need to restructure their work

conditions in order to provide employees with longer convalescent times, if they wish to retain them. In the U.S., there is no federal law requiring employers to provide paid sick leave, but in England, employers must provide Statutory Sick Pay (SSP) for up to 28 weeks. These generous welfare payments may explain why this result is present in this study but not in the U.S. based ones. Governments could also use the results to amend their current practice with relation to sick/disability payments. In England's case, this may involve reviewing the length of time SSP is provided for by employers. This may, in the future, reduce dependence on Employment and Support Allowance, which is provided for by the state. A future study using a larger European dataset would provide an interesting comparison and would allow us to see if the effects of cancer persist past the 6-month post diagnosis period in a different labour market.

Acknowledgements

This research was supported by an Irish Research Council grant.

References

BRADLEY, C. J., H. L. BEDNAREK, and D. NEUMARK (2002): "Breast Cancer Survival, Work, and Earnings," *Journal of Health Economics*, 21, 757-779.

BRADLEY, C. J., D. NEUMARK, H. L. BEDNAREK, and M. SCHENK (2005): "Short-Term Effects of Breast Cancer on Labor Market Attachment: Results from a Longitudinal Study," *Journal of Health Economics*, 24, 137-160.

BRADLEY, C. J., D. NEUMARK, Z. LUO, and H. L. BEDNAREK (2007): "Employment-Contingent Health Insurance, Illness, and Labor Supply of Women: Evidence from Married Women with Breast Cancer," *Health Economics*, 16, 719-737.

BRADLEY, C. J., D. NEUMARK, L. ZHEHUI, H. BEDNAREK, and M. SCHENK (2005): "Employment Outcomes of Men Treated for Prostate Cancer," *JNCI: Journal of the National Cancer Institute*, 97, 958-965.

BRADLEY, C. J., D. NEUMARK, L. ZHEHUI, and M. SCHENK (2007): "Employment and Cancer: Findings from a Longitudinal Study of Breast and Prostate Cancer Survivors," *Cancer Investigation*, 25, 47-54.

CANCER RESEARCH UK (2004): "Cancer in the European Union. Http://Publications.Cancerresearchuk.Org/Downloads/Product/Cs_Cs_Europe.Pdf," (accessed on 31.05.2013).

CHIRIKOS, T. N., A. RUSSELL-JACOBS, and P. B. JACOBSEN (2002): "Functional Impairment and the Economic Consequences of Female Breast Cancer," *Women & Health*, 36, 1.

D'AGOSTINO, R. B. (1998): "Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group," *Stat Med*, 17, 2265-81.

EHEMAN, C., S. J. HENLEY, R. BALLARD-BARBASH, E. J. JACOBS, M. J. SCHYMURA, A. M. NOONE, L. PAN, R. N. ANDERSON, J. E. FULTON, B. A. KOHLER, A. JEMAL, E. WARD, M. PLESCIA, L. A. RIES, and B. K. EDWARDS (2012): "Annual Report to the Nation on the Status of Cancer, 1975-2008, Featuring Cancers Associated with Excess Weight and Lack of Sufficient Physical Activity," *Cancer*, 118, 2338-66.

FRÖLICH, M. (2004): "Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators," *Review of Economics and Statistics*, 86, 77-90.

HECKMAN, J., H. ICHIMURA, J. SMITH, and P. TODD (1998): "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66, 1017-1098.

HECKMAN, J. J., R. J. LALONDE, and J. A. SMITH (1999): "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, 1865-2097.

JOLY, F., M. HENRY-AMAR, P. ARVEUX, O. REMAN, A. TANGUY, A. M. PENY, P. LEBAILLY, J. MACÉ-LESEC'H, B. VIÉ, J. Y. GÉNOT, A. BUSSON, X. TROUSSARD, and M. LEPORRIER (1996): "Late Psychosocial Sequelae in Hodgkin's Disease Survivors: A French Population-Based Case-Control Study," *J Clin Oncol*, 14, 2444-53.

JOLY, F., J. F. HÉRON, L. KALUSINSKI, P. BOTTET, D. BRUNE, N. ALLOUACHE, J. MACÉ-LESEC'H, J. E. COUËTTE, J. PÉNY, and M. HENRY-AMAR (2002): "Quality of Life in Long-Term Survivors of Testicular Cancer: A Population-Based Case-Control Study," *J Clin Oncol*, 20, 73-80.

LEUVEN, E., and B. SIANESI (2003): "Psmatch2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing", http://Ideas.Repec.Org/C/Boc/Bocode/S432001.Html/ (Version 4.0.6).

MALVEZZI, M., P. BERTUCCIO, F. LEVI, C. LA VECCHIA, and E. NEGRI (2013): "European Cancer Mortality Predictions for the Year 2013," *Ann Oncol*, 24, 792-800.

MORAN, J. R., P. F. SHORT, and C. S. HOLLENBEAK (2011): "Long-Term Employment Effects of Surviving Cancer," *Journal of Health Economics*, 30, 505-514.

ROSENBAUM, P. R., and D. B. RUBIN (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41-55.

RUBIN, D. B., and N. THOMAS (1996): "Matching Using Estimated Propensity Scores: Relating Theory to Practice," *Biometrics*, 52, 249-64.

SHORT, P. F., J. J. VASEY, and J. R. MORAN (2008): "Long-Term Effects of Cancer Survivorship on the Employment of Older Workers," *Health Services Research*, 43, 193-210.

STEINER, J. F., T. A. CAVENDER, D. S. MAIN, and C. J. BRADLEY (2004): "Assessing the Impact of Cancer on Work Outcomes: What Are the Research Needs?," *Cancer*, 101, 1703-11.

STEINER, P. M., and D. COOK (2013): "Matching and Propensity Scores," in *The Oxford Handbook of Quantitative Methods*, ed. by T. L. Little. New York, NY: Oxford University Press.

TASKILA-ÅBRANDT, T., R. MARTIKAINEN, S. V. VIRTANEN, E. PUKKALA, P. HIETANEN, and M.-L. LINDBOHM (2004): "The Impact of Education and Occupation on the Employment Status of Cancer Survivors," *European Journal of Cancer*, 40, 2488-2493.

TORP, S., R. A. NIELSEN, S. D. FOSSÅ, S. B. GUDBERGSSON, and A. A. DAHL (2013): "Change in Employment Status of 5-Year Cancer Survivors," *Eur J Public Health*, 23, 116-22.

VAN TULDER, M. W., N. K. AARONSON, and P. F. BRUNING (1994): "The Quality of Life of Long-Term Survivors of Hodgkin's Disease," *Ann Oncol*, 5, 153-8.