UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2014

The Effects of Cancer in the English Labour Market

David Candon, University College Dublin

WP14/09

May 2014

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4

The 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. I use data from the English Longitudinal Study of Ageing (ELSA) and find that, not only does cancer have a negative impact in the first 6-month period following diagnosis, but also in the second 6-month period. I estimate that, in the second 6-month period after diagnosis, respondents with cancer are 20.7 percentage points less likely to work and work 24% less hours a week when compared to matched, healthy controls. This suggests 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. These results have implications for government policy and employers, because it increases both the length of time that survivors may be on government supported sick pay and the expected time that workers will be absent from work due to illness.

JEL classification: I10, J21, J22

Keywords: Labour market; 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 work interruptions. When cancer survivors regain their health, they are faced with decisions regarding whether to work, retire or change their working hours.

There are many areas that 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 in the short term on employment and hours of work (up to 6 months post-diagnosis), and smaller, sometimes insignificant results in the long term (from 2 to 6 years post-diagnosis). This suggests that by exploring the first 2 year period after diagnosis in detail, it could allow us to identify at what stage cancer survivors are ready to return to their original working pattern.

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

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

schedules, and rest time and the reductions in employment and hours of work that the literature shows are to be expected. However, 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, such as the health or social security system. 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 is important for formulating government policy. In England, in particular, employers must provide statutory sick pay (SSP) for up to 28 weeks (just over 6 months), while the state is responsible for paying Employment and Support Allowance (ESA), which people may claim if they are still unable to work, but are no longer eligible for SSP. If cancer survivors are not returning to work after 6 months, it puts a greater strain on government finances. Knowing when the survivors return to work also allows employers to 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. 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 returning to the labour force

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 14.3 percentage points ($p<0.01$) less likely to work than their non-cancer counterparts. This rises to 15.8 percentage points and 20.7 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 16.3% ($p<0.01$) less hours per week. Again, these numbers rise to 21% ($p<0.05$) in the first 6-month period and 24% ($p<0.01$) in the second 6-month period, respectively. These results are robust to numerous different types of specifications and appear insensitive to potentially omitted confounders. 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 robustness and sensitivity in sections 6 and 7. Finally, I conclude in Section 8.

Literature review: the effects of cancer on employment

In the last 20 years, numerous studies in Europe have tried to estimate the effects of cancer on work life. Amir *et al*. (2007), Johnsson *et al*. (2009) and Spelten *et al*. (2003) have used data from hospitals and registries to examine the proportion of survivors who return to work after treatment. In order to separate the effect of cancer on labour market outcomes from other non-cancer related covariates such as comorbidities or labour market conditions, some studies

have also used a non-cancer control group (Gudbergson *et al*., 2007 and Taskila-Åbrandt *et al*., 2004). However, these studies did not measure changes over time which, according to Steiner *et al*. (2004), should be taken into account when studying the effects of cancer on work outcomes. 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 month period following diagnosis, the probability of employment (conditional on prior employment) was the same as the control group, though changes in hours of work 12 months after diagnosis were still statistically different between the two groups (Bradley *et al*., 2007). Men with prostate cancer, 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 post-diagnosis (Bradley *et al*., 2005b and Bradley *et al*., 2007). 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. The mean ages of the survivors were 51 in Bradley *et al*. (2005a) , 54 in Bradley *et al*. (2005b) and between 55 – 65 in Short *et al*. (2008). Though 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, men and women between the ages of 28 and 54 (mean age = 46) 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.

More recently, there has been an increase in the number of cancer studies coming from Scandinavian countries. The establishment of personal identification numbers in the 1960's in these countries has meant that econometricians can now link patients with diagnoses from cancer registries to other population-based surveys which the patients would have completed throughout their lifetime. Not only does this provide the econometrician with demographic information which would otherwise be unavailable, but it also provides them with a large number of cancer observations (n = 3,000 – 6,000) and a large control sample. Using this method of analysis with Danish data, Heinesen and Kolodziejczyk (2013) find that, for men with colorectal cancer, and women with colorectal or breast cancer, there is a significant reduction in the probability of being employed for the first three years after diagnosis. Also, Torp *et al*. (2013), find that women are significantly less likely to employed 5 years after diagnosis but there is no change in the employment for male survivors when compared with respective control groups in Norway.

Both of these Scandinavian studies are similar to Moran *et al*. (2011) in that the average age of the cancer group was 47 – 48. Because the younger survivors' peer group are in the middle of their working life, changes in the labour supply of cancer survivors may be easily picked up. The peer group of older survivors, however, are likely to be retiring or reducing their labour supply anyway, which may mask any negative effects from cancer. Because the respondents in my sample are near the end of their working life, it is important to examine the changes over short intervals, such as 6-month periods, as it will allow these negative effects to become apparent. 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).

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. Because I wish to observe pre-cancer behaviour, I need information on the respondents in 2 different waves to construct 1 observation. For example, if a respondent has no cancer in the first wave and no cancer in the second wave, this is included as a non-cancer observation. Similarly, if they have no cancer in the second wave but do have cancer in the third wave, this is included as a cancer observation. I then stack these different waves into two time periods: T_1 and T_2. T_1 is a period 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[2].

[2] Due to limitations with the sample size I am unable to look at the effects of cancer over a longer period of time.

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. I combine these for a total sample of 91,236. Table 1 provides more details on how I select the sample. I first restrict the data to those respondents with complete survey records for the variables that I am interested in. 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. Also, because I am interested in how cancer affects the respondents who are working, I remove people who are unemployed or looking after the family home[3]. This means that all the respondents are working in T_1. In order to reduce the effect of very extreme values on the results, I restrict the sample to respondents who, in T_1, are between the ages of 49 and 67, who work a maximum of 100 hours per week, who do not triple their hours of work in T_2, and whose income is £5000 or less per week. 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. 2,046 respondents in the non-cancer group and 20 respondents in the cancer group reported their hours of work as missing, despite signalling that they work. These respondents will be included in the analysis on the probability of employment but not on the hours of work. The people who remain in the study are ones who never get cancer or do not have cancer in one wave but have it in a subsequent wave. This leaves a final sample of 10,152 respondents, 129 of which will be cancer survivors in T_2.

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. The ELSA data also contain information on the month when the respondents are interviewed and the

[3] This applies only to T_1. Respondents are allowed become retired, permanently sick or disabled, unemployed, or look after the family home in T_2. This is detailed in the Methodology section.

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 first 6 months, the second 6 months, the third 6 months or the fourth 6 months). There are 7 respondents for which the cancer diagnosis date is missing so these respondents are not included in the subgroup analysis.

The main advantage of using survey data rather than registry data is that we have information on the respondents before they are 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. Asking cancer patients for retrospective information on their working behaviour is likely to be problematic. 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 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[4].

Another benefit with this type of data is that the cancer group and the non-cancer group are drawn from the same dataset, 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),

[4] While the Scandinavian cancer registries can be linked to other demographic information using personal identification numbers, I do not know if this includes information on hours of work. However, both studies previously mentioned do not include information on hours of work either as a dependent or independent variable leading me to believe that it is not possible.

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 groups before diagnosis, strengthening the conclusion that any differences in outcomes observed after diagnosis is due to surviving cancer.

A drawback from using this data is that the size of the cancer sample is not of the same magnitude as the data taken from MDCSS or PSCSS. This restricts the amount of subgroup analysis which can be done. However, other papers which use PSM to evaluate labour market programs such as Dejehia and Wabha (2002) and Ichino *et al*. (2008) also use a small number of treated observations relative to a large control group.

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. The first outcome variable that I am interested in is employment status, E_{it} , where E_{i1} = 1 for all respondents. 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 the percent change in hours of work, which is defined as $(H_{i2} - H_{i1})/H_{i1}$. This allows for individual time invariant unobservable characteristics to be differenced out.

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

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 surviving cancer[6]. Because only $E(Y_0|D_0)$ is observed in the data, the (weak) conditional independence assumption and overlap condition is used to construct the counterfactual for the treatment group. If the following assumptions hold,

$$Y_0 \perp D|X, \quad (2)$$

$$\Pr(D = 1|X) < 1, \quad (3)$$

then $Y_0|X, D_1 = Y_0|X, D_0$. Instead of using the entire vector of covariates, PSM is used to balance the observable characteristics between the samples and eliminate the impact of the

[6] As Bradley *et al*. (2002) point out, we cannot estimate the total effect of cancer since the effects on those who died are not measured.

observables as confounding factors (Rosenbaum and Rubin, 1983; D'agostino, 1998). The propensity score can be defined by the following equation,

$$p(x) \equiv \Pr(D = 1 | X = x), \quad (4)$$

where the probability of getting the treatment (surviving cancer) is based on the vector of X variables. In this case, the vector of variables includes age, equivalised household income and binary variables indicating whether the respondents are female, are of non-white ethnicity, have a third level education, are married, have a child living at home, are in a state of poor health (poor and fair or very poor, poor, and fair depending on which Likert scale was used), whether they smoke, whether they drink alcohol more than two days a week, whether they have had a heart attack, whether they have had a stroke, whether they are self-employed, level of local area unemployment, and their waves in the ELSA, all measured at baseline[7]. 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). PSM has been routinely used in this type of research and can be found in Bradley *et al*. (2005a), Short *et al*. (2008) and Moran *et al*. (2011). The construction of the outcome and control variables is similar to Bradley *et al*. (2005a) but for the estimation procedure, 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

[7] Data on local area unemployment is unavailable in the ELSA and is obtained from UK National Statistics (2013).

it involves discarding all other potentially valuable observations) (Steiner and Cook, 2013). In order to use more of the large non-cancer control group, I estimate the model using a kernel smoothed matching estimator. In situations like this, where there are numerous control observations per treated observation, kernel matching has been shown to work well (Frölich, 2004). Kernel matching is a form of weighted matching where a kernel 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 order to reduce the standardisation bias to below 5%, a bandwidth of 0.001 is selected. In the matched estimates, 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. The probit models used to construct the propensity scores for both models are available in the appendix.

In addition to the matching estimates, I also perform multiple regressions for comparison. Clustered standard errors (clustered by person) 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. The first panel contains the descriptive statistics for the respondents' T_1 information. In columns 1 and 2, the cancer sample is compared to the non-cancer sample in the employment model. The significant differences between the cancer and non-cancer group indicate that the cancer group are older, more likely to be in poor health, a higher proportion are female and they are more likely to have had a

heart attack. Columns 3 and 4 compare the cancer group and the non-cancer group for the hours of work model. Again, the only significant differences between the groups is that the cancer group is older, more likely to be in poor health and more likely to have had a heart attack. 100% of respondents in both groups work but there is a significant difference in terms of hours worked with the cancer group working 31.96 hours per week in T_1 compared to 34.98 in the non-cancer group.

In T_2, there is a significant difference in the probability of being employed in the employment model with 82.55% of the non-cancer group being employed compared to 61.24% of the cancer group. In the hours of work model, the non-cancer respondents who worked in T_2 worked 28.99 hours per week whereas the cancer group worked 18.46 hours per week. Basic bivariate analysis shows that the probability of being employed in T_2 is significantly lower in the cancer group relative to the non-cancer group in both models. Also, the cancer group work 3 hours less than the non-cancer group ($p<0.01$) in T_1 and this difference becomes almost 10 hours ($p<0.01$) once the respondents are diagnosed with cancer in T_2.

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 in either the employment or hours of work model. The re-estimated *t*-tests for both the employment and hours of work model are available in the appendix.

Effects of cancer less than 2 years post-diagnosis

Table 3 shows the results from the estimation of the effect of cancer on the probability of employment, and the percentage change 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. Finally, the third column gives the results from kernel matching. 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 21.3 percentage points and hours of work by 23.6% in the second period. When baseline characteristics are included as controls, the effect of cancer on employment is reduced, but is still quite large at -14.7 percentage points. Similar results are found for the hours of work model, where cancer changes hours of work by -16.8%. The results from the kernel matching are of a comparable magnitude to the adjusted regression model. All results are significant at the 1% level.

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 6-month period 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, I also estimate the models using the diagnosis period for cancer[8]. Looking at the unadjusted mean

[8] For the PSM, separate propensity scores are created for each new cancer variable.

difference, the respondents who are in the first 6-month period since diagnosis are 24.2 percentage points less likely to work compared to their non-cancer controls and they also work 32.7% hours less per week. As expected, smaller coefficients are found in the adjusted and matching models. All the results are statistically significant at the 10% level, except the kernel estimate for employment in the matched model.

Similar results are found when I look at respondents in the second 6-month period after diagnosis, though the effect on employment and on hours is larger. An interesting finding from this model is that the coefficients are still statistically significant and quite large. If the "shock" from cancer was confined to the first 6-month period, we would expect to not find any significant results the second 6-month period. However, the results suggest that cancer is having a comparable effect in both the first and second 6-month periods.

When the respondents are in the third 6-month period following diagnosis, the unadjusted models still show large, significant coefficients for the effect of cancer on employment and hours of work. However, the adjusted model regression and matching models give smaller, insignificant coefficients. In the fourth 6-month period following diagnosis, cancer still has a negative effect on the outcomes, but the effect sizes are now smaller than any other specification, and none of the results are statistically significant.

Specification checks

Employment model with non-missing hours

As stated in the Data section, the hours of work models contain 2,246 less respondents than the employment models. Reasons for this occurring could be respondents being coded as working when they do not work or self-employed respondents who do not record how many

hours a week they work. In order to remove the impact that these respondents could have on the results, I re-estimate the employment models using only respondents who provide information on hours of work in both periods. In the first panel of Table 4, the unadjusted mean difference shows that being diagnosed with cancer at any stage in the last 2 years reduces the probability of employment by 23.2 percentage points. This falls to 16.1 and 15.2 percentage points when the effect is estimated with the adjusted and matching models respectively. All results are significant at the 1% level.

When we separate the results into the 4 diagnosis periods, the relationship between cancer diagnosis and the probability of employment becomes monotonic, where the closer you are to your diagnosis, the larger the effect is. The reductions in the probability of employment given by the matching models are 23 percentage points ($p<0.1$), 18.8 percentage points ($p<0.05$) and 17.1 percentage points ($p<0.1$) for the first, second and third 6-month period after diagnosis respectively. This is different from the main results, where the second 6-month period following diagnosis showed the largest reduction in the probability of being employed.

Working full time or part time in first period

In the middle panels of Table 4, I restrict the analysis to respondents who were either working full time or part time in the first period. In this case, I define working full time as working at least 35 hours per week. Theoretically, it is hard to argue which group should be greater impacted by cancer diagnosis. In one respect, people who are only working part time may not need to reduce their labour supply when they are diagnosed with cancer, since they are not faced with the same time constraints as full time workers. However, part time workers may not be as attached to their jobs as their full time counterparts, and may be more likely to leave the labour force once diagnosed. The results suggest that the first effect dominates, with large,

statistically significant, reductions in employment and hours of work for the full time group. The results for the part time group are smaller, and the adjusted and matching employment models are not statistically significant.

Respondents employed in both periods

So far, I have examined the changes that are happening at the extensive margin; the decrease in the probability of employment once respondents are diagnosed with cancer and the effect that it has on hours of work. It is also possible that changes could be taking place at the intensive margin, where respondents are still working, but that are working more or less than they were before. In the final panel of Table 4, I examine the effect that cancer has on the respondents who work in both periods. I find that while there is a 5% reduction in hours of work in the unadjusted model ($p<0.05$), it is only 4.3% and 3.9% in the adjusted and matching models respectively, and neither are statistically significant. Given that I cannot control for the severity of the disease, it could be that the respondents who return are the ones who have been affected the least by their cancer.

Pre-existing hours of work trend

Another benefit from using the ELSA data is that, not only do I get to observe all of the respondents' work behaviour in the period before they are diagnosed with cancer, but I also get to observe some respondents' behaviour up to 2, 3 or 4 periods before they are diagnosed with cancer. I use this information to plot the hours of work of the respondents before and after they are diagnosed. In Figure 1, we can see the average hours of work per week for the cancer and non-cancer respondents and *p*-values from the *t*-tests of differences between the groups. The only statistically significant difference occurs in the first period after diagnosis

(p=0.001). Because the work behaviour of the cancer and non-cancer groups is the same until diagnosis, it reinforces the argument that the change is due to surviving cancer.

Sensitivity checks

In this section, I hypothesize the presence of an unobserved confounding variable, *U*, which may pose a threat to the validity of the previously estimated ATT. For simplicity, let us assume that *U* is binary. In this scenario, the conditional independence assumption (2) is not met if *U* is not taken into account, but it does hold if *U* is controlled for. In the case of this analysis, the ATT estimates could be affected if *U* not only selects respondents into the treatment (*s*), but also leads to worse labour market outcomes for those respondents who are not in the treatment group (*d*). In this situation, it can be thought of as a variable such as obesity, which is a risk factor for certain cancers. If the respondent is obese, it increases the chances they will be in the cancer group, but it also means that if they are in the non-cancer group, they will have worse labour market outcomes. If

$$p_{ij} \equiv Pr(U = 1 | D_i, Y_j, X) = Pr(U = 1 | D_i, Y_j), \qquad (5)$$

with $i,j \in \{0, 1\}$, then this can be expressed more precisely as $s = p_{1.} - p_{0.} > 0$ (the probability of being in the treated group is greater than the probability of the untreated group) and $d = p_{01} - p_{00} < 0$ (the probability of a negative work outcome when in the untreated group is greater than the probability of a positive work outcome in the untreated group). Ichino *et al*. (2008) and Nannicini (2007) developed the SENSATT command for Stata; a sensitivity test which

allows the econometrician to vary the probability of both *d* and *s* to see how strong of an effect is required to drive the ATT to 0[9].

In Table 5, the unobserved confounder, *U*, is calibrated to mimic the *d* and *s* of the other observed binary variables in the model. The first panel of Table 5 gives the values of *d* and *s* that these confounders produce, as well as the estimated ATT. The values for *d* and *s* that these binary variables produce are quite small and have almost no effect on the ATT. This means that if the distribution of *U* is similar to the distribution of any other binary variable in the model, then it will have a negligible effect on the ATT.

In the second panel of Table 5, I vary the values of both *d* and *s* manually and examine how the ATT changes in response[10]. Even when *d* = -0.3 and *s* = 0.3 the ATT in the employment model is -0.064 and the ATT in the hours model is -0.086. In both models, *s* had a stronger effect of driving the ATT toward 0 than *d*. This could mean that in this case, it may be variables that predict selection into the treatment, rather than variables which are associated with the outcome, that provide the largest threat to the ATT estimates. Given the relatively small values that are produced for *d* and *s* in the first panel, we can see that *U* would have to predict selection into the treatment and negative outcomes for the untreated far more than any of these variables in order to 'kill' the ATT. The values of *d* and *s* that would be required to drive the ATT to zero seem implausibly large when compared to these values, which makes it more likely that we are observing a causal effect.

[9] See both Ichino *et al*. (2008) and Nannicini (2007) for a more rigorous explanation of the tests.
[10] For this analysis, I assume that $Pr(U = 1) = 0.2$ (the prevalence of the confounder in the whole sample is 20%) and $p_{11} - p_{10} = 0$ (the probability of a positive labour market outcome and a negative labour market outcome are the same in the presence of the treatment).

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 14.7 percentage points ($p<0.01$) less likely to work and work 16.8% ($p<0.01$) less hours per week than their non-cancer counterparts. When the 2 year period is broken up into the 4 consecutive 6-month periods, the magnitude of the effect of cancer on work outcomes is larger for the first 6-month period, larger again in the second 6-month period and then disappears in the second year. In this case of the second 6-month period, cancer reduces the probability of employment by 20.7 percentage points ($p<0.01$) and hours of work by 21% ($p<0.01$).

The results are robust to the restriction of the sample to respondents with complete information on hours of work. Separating the sample into full time and part time workers suggests that cancer has a larger effect on full time workers who have been diagnosed anytime in the past 2 years. Their probability of employment is reduced by 20.4 ($p<0.01$) percentage points and their hours of work reduced by 19.1% ($p<0.01$), which falls to 11.5 percentage points and 17.5% ($p<0.01$) when looking at part time workers. When I examine

respondents who are still working in the second period, I find that there is no significant difference in working hours. The sensitivity of the estimates to a potentially unobserved confounder is also examined. It would require an omitted variable that predicts both selection into the treatment group and a negative outcome for the untreated group far greater than any of the variables already included in the model to affect the ATT.

Although the results are robust to the various specifications and sensitivity checks it is important to consider the following limitations. 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. 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 (Bradley *et al*., 2007). 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. Respondents who have distant metastasis, the most severe form of cancer, 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. Also, if surviving cancer alters the marginal rate of substitution between work and leisure, what we observe may be a change in the preference for work, rather than an enforced reduction in labour supply due to cancer. However, given that the effects on employment and hours of work reduce to insignificant levels at 18-24 months, this is unlikely to be the case.

Many of the findings in this paper are consistent with other estimates of working and hours worked found in the U.S. 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. The main finding from this paper that is not consistent with the U.S. studies is the significant negative impact of cancer in the second 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 first 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 of work. There is no reduction in hours for the workers who do return to work though, which may suggest that those who return are not impaired by their cancer. If this is the case, employers need to be aware that may need to restructure their work conditions in order to provide employees who do not return with longer convalescent times or different working arrangements, 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 SSP for up to 28 weeks (equivalent to the first 6-month period). These welfare payments may result in the "shock" being pushed out by 6 months. This work has important policy implications as governments could 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 ESA, which is provided for by the state. Or, if employers are willing to change working conditions to accommodate survivors, they could be supplemented by the ESA. 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 similar labour markets.

Acknowledgements

This research was supported by an Irish Research Council grant. I also wish to thank Paul Devereux, Orla Doyle, Kanika Kapur, David Madden and discussants at the 2013 European Doctoral Group in Economics and the 2013 Irish Society for New Economists conferences for helpful comments.

References

Amir Z, Moran T, Walsh L, *et al*. 2007. Return to paid work after cancer: a British experience. *J Cancer Surviv* **1**: 129-136.

Bradley CJ, Bednarek HL and Neumark D. 2002. Breast Cancer and Women's Labor Supply. *Health Services Research* **37**: 1309-1328.

Bradley CJ, Neumark D, Bednarek HL, *et al*. 2005. Short-Term Effects of Breast Cancer on Labor Market Attachment: Results from a Longitudinal Study. *Journal of Health Economics* **24**: 137-160.

Bradley CJ, Neumark D, Zhehui L, *et al*. 2005. Employment Outcomes of Men Treated for Prostate Cancer. *JNCI: Journal of the National Cancer Institute* **97**: 958-965.

Bradley CJ, Neumark D, Zhehui L, *et al*. 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.* Available from: http://publications.cancerresearchuk.org/downloads/product/CS_CS_EUROPE.pdf [May 31, 2013].

Chirikos TN, Russell-Jacobs A and Jacobsen PB. 2002. Functional Impairment and the Economic Consequences of Female Breast Cancer. *Women & Health* **36**: 1 - 20.

D'Agostino RB. 1998. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* **17**: 2265-2281.

Dehejia R and Wahba S. 2002. Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics* **84**: 151-161.

Eheman C, Henley SJ, Ballard-Barbash R, *et al*. 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**: 2338-2366.

Frölich M. 2004. Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators. *Review of Economics and Statistics* **86**: 77-90.

Gudbergsson SB, Fosså SD, Sanne B, *et al*. 2007. A controlled study of job strain in primary-treated cancer patients without metastases. *Acta Oncol* **46**: 534-544.

Heckman J, Ichimura H, Smith J, *et al*. 1998. Characterizing selection bias using experimental data. *Econometrica* **66**: 1017-1098.

Heckman JJ, Lalonde RJ and Smith JA. 1999. The economics and econometrics of active labor market programs. In Handbook of Labor Economics, Ashenfelter O and Card D (eds.), North Holland: Amsterdam; 1865-2097.

Heinesen E and Kolodziejczyk C. 2013. Effects of breast and colorectal cancer on labour market outcomes-Average effects and educational gradients. *J Health Econ* **32**: 1028-1042.

Ichino A, Mealli F and Nannicini T. 2008. From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics* **23**: 305 - 327.

Johnsson A, Fornander T, Rutqvist LE, *et al*. 2009. Predictors of return to work ten months after primary breast cancer surgery. *Acta Oncol* **48**: 93-98.

Leuven E and Sianesi B. 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, Bertuccio P, Levi F, *et al*. 2013. European cancer mortality predictions for the year 2013. *Ann Oncol* **24**: 792-800.

Moran JR, Short PF and Hollenbeak CS. 2011. Long-term employment effects of surviving cancer. *Journal of Health Economics* **30**: 505-514.

Nannicini T. 2007. Simulation-based sensitivity analysis for matching estimators. *The Stata Journal* **7**: 334 - 350.

Rosenbaum PR and Rubin DB. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**: 41-55.

Rubin DB and Thomas N. 1996. Matching using estimated propensity scores: relating theory to practice. *Biometrics* **52**: 249-264.

Short PF, Vasey JJ and Moran JR. 2008. Long-Term Effects of Cancer Survivorship on the Employment of Older Workers. *Health Services Research* **43**: 193-210.

Spelten ER, Verbeek JH, Uitterhoeve AL, *et al*. 2003. Cancer, fatigue and the return of patients to work-a prospective cohort study. *Eur J Cancer* **39**: 1562-1567.

Steiner JF, Cavender TA, Main DS, *et al*. 2004. Assessing the impact of cancer on work outcomes: what are the research needs? *Cancer* **101**: 1703-1711.

Steiner PM and Cook D. 2013. Matching and Propensity Scores. In The Oxford Handbook of Quantitative Methods, Little TD (ed.), Oxford University Press: New York, NY; 237-259.

Taskila-Åbrandt T, Martikainen R, Virtanen SV, *et al*. 2004. The impact of education and occupation on the employment status of cancer survivors. *European Journal of Cancer* **40**: 2488-2493.

Torp S, Nielsen RA, Fosså SD, *et al*. 2013. Change in employment status of 5-year cancer survivors. *Eur J Public Health* **23**: 116-122.

UK National Statistics. 2013. *Labour market.* Available from: http://www.statistics.gov.uk/hub/labour-market/job-statistics/jobs/index.html [May 31, 2013].

Table 1. Sample breakdown and exclusion criteria

Exclusion criteria		**Observations**
Unrestricted sample		91,236
Missing or incomplete survey records with regards to sex, age, children at home, household income, marital status, health status, work status, area unemployment, and hours of work[a]		58,835
Cancer in T_1		1,027
Diagnosed with cancer more than 2 years ago		91
Aged below 49 or above 67 in T_1		12,253
Unemployed, looking after the family home, retired or sick and disabled in T_1		8,550
Works more than 100 hours a week in T_1 or triples their number of hours in T_2		19
Equivalised household income greater than £5000 or less than £0 in T_1		180
Restricted sample		10,281
Cancer	**Observations (Hours only)**	**Observations**
Respondents without cancer in T_2	8,106	10,152
Respondents with cancer in T_2	109	129
Type of cancer	**Observations (Hours only)**	**Observations**
Lung	5	5
Breast	31	38
Colon, bowel or rectum	17	19
Lymphoma	2	3
Leukaemia	2	3
Melanoma or other skin cancer	14	17
Somewhere else	38	44
Period of diagnosis	**Observations (Hours only)**	**Observations**
First 6 months	19	24
Second 6 months	36	42
Third 6 months	24	27
Fourth 6 months	25	29
Missing	5	7

Note: [a] This includes answers which such as refusal, don't know and other non-coded responses.

Table 2. Descriptive statistics

Variable	(1) Cancer group	(2) Non-cancer group	(3) Cancer group	(4) Non-cancer group
T_1	Employment model		Hours of work model	
Female	58.14%**	48.95%	58.72%	52.70%
Age	58.30***	56.40	58.27***	56.19
Degree	16.28%	15.72%	15.60%	14.81%
Married	79.07%	77.44%	77.06%	77.60%
Non-white	0.78%	2.82%	0.92%	2.80%
Child at home	39.54%	38.88%	38.53%	40.13%
Poor health	21.71%**	13.87%	21.10%**	13.91%
Heart attack	3.88%**	1.57%	4.59%***	1.47%
Stroke	0.78%	0.83%	0.00%	0.00%
Smoking	18.61%	17.38%	22.02%	17.26%
Alcohol	37.98%	35.67%	37.62%	33.74%
Household income (weekly, £s)	407.19	402.07	373.54	394.70
Self-employed	13.18%	16.26%	1.84%	0.88%
Local area unemployment	5.52%	5.51%	5.43%	5.51%
Wave 1 – 2	27.13%	27.47%	28.44%	27.99%
Wave 2 – 3	21.71%	21.16%	22.94%	21.33%
Wave 3 – 4	17.83%*	24.26%	17.43%	23.96%
Wave 4 – 5	33.33%	27.11%	31.19%	26.68%
Working	100.00%	100.00%	100.00%	100.00%
Hours worked	31.96**	34.98	31.88**	34.97
T_2				
Working	61.24%***	82.55%	58.72%***	81.89%
Hours worked	18.34***	28.99	18.46***	28.99
Observations	129	10,152	109	8,106

Note: *** Significantly different from the non-cancer sample at the 1% level (when comparing (1) with (2) or (3) with (4)).
** Significantly different from the non-cancer sample at the 5% level (when comparing (1) with (2) or (3) with (4)).
* Significantly different from the non-cancer sample at the 10% level (when comparing (1) with (2) or (3) with (4)).

Table 3. Effects of cancer in the first 2 years post-diagnosis

	Unadjusted mean difference	Regression adjusted mean difference	Kernel matching	*N*
		Less than 2 years post-diagnosis		
Working	-0.213***	-0.147***	-0.143***	10,281
	(0.043)	(0.041)	(0.042)	
Hours of work	-0.236***	-0.168***	-0.163***	8,215
	(0.048)	(0.046)	(0.049)	
		First 6 months post-diagnosis		
Working	-0.242**	-0.163*	-0.158	10,176
	(0.101)	(0.088)	(0.101)	
Hours of work	-0.327***	-0.234**	-0.210**	8,125
	(0.103)	(0.094)	(0.106)	
		Second 6 months post-diagnosis		
Working	-0.254***	-0.207***	-0.207***	10,194
	(0.076)	(0.073)	(0.075)	
Hours of work	-0.309***	-0.257***	-0.240***	8,142
	(0.077)	(0.071)	(0.085)	
		Third 6 months post-diagnosis		
Working	-0.196**	-0.107	-0.114	10,179
	(0.093)	(0.087)	(0.095)	
Hours of work	-0.200*	-0.114	-0.158	8,130
	(0.102)	(0.093)	(0.101)	
		Fourth 6 months post-diagnosis		
Working	-0.136	-0.088	-0.104	10,181
	(0.086)	(0.084)	(0.082)	
Hours of work	-0.123	-0.080	-0.075	8,131
	(0.112)	(0.114)	(0.113)	

Note: In the regression adjusted model, the variables included are the respondent's age, equivalised household income and binary variables indicating whether the respondents are female, are of non-white ethnicity, have a third level education, are married, have a child living at home, are in a state of poor health, whether they smoke, whether they drink alcohol more than two days a week, whether they have had a heart attack, whether they have had a stroke, whether they are self-employed, level of local area unemployment, and their waves in the ELSA, all measured at baseline. Clustered standard errors are in parentheses. Bootstrapped standard errors are in parentheses for kernel matching results.

* Result significant at the 10% level.

** Result significant at the 5% level.

*** Result significant at the 1% level.

Table 4. Specification checks

	Unadjusted mean difference	Regression adjusted mean difference	Kernel matching	*N*
	Employment model with non-missing hours (Diagnosis period)			
Working (Less than 2 years)	-0.232*** (0.048)	-0.161*** (0.044)	-0.152*** (0.048)	8,215
Working (First 6 months)	-0.345*** (0.115)	-0.251*** (0.097)	-0.230* (0.118)	8,125
Working (Second 6 months)	-0.263*** (0.083)	-0.204** (0.080)	-0.188** (0.093)	8,142
Working (Third 6 months)	-0.194* (0.099)	-0.113 (0.088)	-0.171* (0.098)	8,130
Working (Fourth 6 months)	-0.179* (0.096)	-0.128 (0.093)	-0.126 (0.091)	8,131
	Working full time in T_1 Less than 2 years post-diagnosis			
Working	-0.280*** (0.066)	-0.230*** (0.060)	-0.204*** (0.067)	5,331
Hours of work	-0.272*** (0.062)	-0.216*** (0.057)	-0.191*** (0.062)	5,331
	Working part time in T_1 Less than 2 years post-diagnosis			
Working	-0.149** (0.069)	-0.086 (0.066)	-0.115 (0.072)	2,884
Hours of work	-0.193*** (0.073)	-0.122* (0.069)	-0.175*** (0.074)	2,884
	Working in T_1 and T_2 Less than 2 years post-diagnosis			
Working	- -	- -	- -	-
Hours of work	-0.050** (0.025)	-0.043 (0.026)	-0.039 (0.027)	6,702

Note: Clustered standard errors are in parentheses. Bootstrapped standard errors are in parentheses for kernel matching results.

* Result significant at the 10% level.

** Result significant at the 5% level.

*** Result significant at the 1% level.

Table 5. Sensitivity of the kernel ATT estimates to an unobserved confounder

Calibrated confounder

Working (original ATT = -0.143)

	ATT	*d*	*s*
Confounder like:			
Poor health	-0.137	-0.06	0.08
Female	-0.138	-0.06	0.09
Child at home	-0.145	0.16	0.01
Degree	-0.144	0.01	0.00
Married	-0.143	0.01	0.02
Heart	-0.140	-0.02	0.02
Smoking	-0.143	0.02	0.02
Alcohol	-0.143	0.01	0.02
Non-white	-0.142	0.00	-0.02
Self-employed	-0.142	0.03	-0.03

Hours of work (original ATT = -0.163)

	ATT	*d*	*s*
Confounder like:			
Poor health	-0.162	-0.03	0.07
Female	-0.161	-0.05	0.07
Child at home	-0.162	0.13	-0.01
Degree	-0.162	-0.02	0.01
Married	-0.163	0.02	-0.01
Heart	-0.160	-0.01	0.03
Smoking	-0.163	0.01	0.05
Alcohol	-0.163	0.00	0.04
Non-white	-0.160	0.01	-0.02
Self-employed	-0.162	0.00	0.01

'Killer' confounder

Working (original ATT = -0.143)

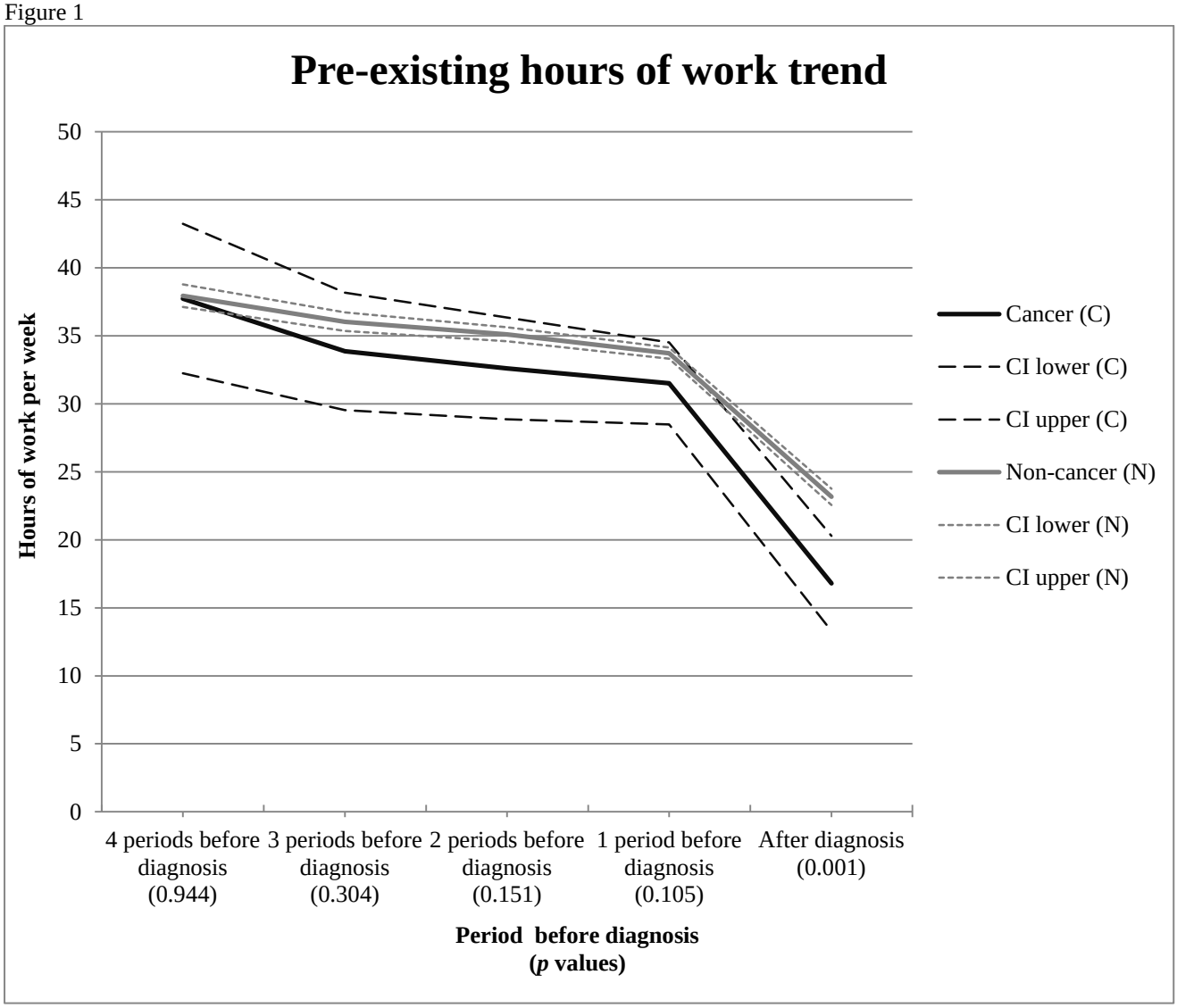
	s = 0.1	*s* = 0.2	*s* = 0.3
d = -0.1	-0.134	-0.126	-0.116
d = -0.2	-0.126	-0.108	-0.090
d = -0.3	-0.121	-0.096	-0.064

Hours of work (original ATT = -0.163)

	s = 0.1	*s* = 0.2	*s* = 0.3
d = -0.1	-0.154	-0.147	-0.136
d = -0.2	-0.146	-0.131	-0.111
d = -0.3	-0.140	-0.114	-0.086

Note: All the ATT estimates are averaged over 50 replications.

Figure 1


Pre-existing hours of work trend
Hours of work per week
50
45
40
35
30
25
20
15
10
5
0
4 periods before diagnosis (0.944)
3 periods before diagnosis (0.304)
2 periods before diagnosis (0.151)
1 period before diagnosis (0.105)
After diagnosis (0.001)
Period before diagnosis (p values)
Cancer (C)
CI lower (C)
CI upper (C)
Non-cancer (N)
CI lower (N)
CI upper (N)

Appendix

Table 6. Raw coefficients from probit models for the probability of getting cancer: Employment model

Variables	(1) Cancer	(2) Cancer 1	(3) Cancer 2	(4) Cancer 3	(5) Cancer 4
Female	0.227***	0.303**	0.264**	-0.057	0.319**
	(0.074)	(0.140)	(0.110)	(0.132)	(0.125)
Age	0.048***	0.052***	0.035***	0.057***	0.023*
	(0.009)	(0.017)	(0.012)	(0.018)	(0.012)
Degree	0.060	0.185	-0.000	0.228	-0.479
	(0.096)	(0.171)	(0.158)	(0.149)	(0.303)
Married	0.060	0.277	0.090	-0.110	0.006
	(0.084)	(0.190)	(0.131)	(0.139)	(0.147)
Non-white	-0.501	-	-	0.143	-
	(0.316)	-	-	(0.288)	-
Child at home	0.128*	0.011	0.225**	0.040	0.149
	(0.076)	(0.154)	(0.112)	(0.141)	(0.123)
Poor health	0.241***	0.040	0.163	0.187	0.299**
	(0.088)	(0.197)	(0.136)	(0.154)	(0.141)
Heart	0.345*	-	0.391	0.104	0.613**
	(0.205)	-	(0.318)	(0.339)	(0.280)
Stroke	-0.220	-	0.251	-	-
	(0.417)	-	(0.431)	-	-
Smoking	0.047	-0.193	0.088	0.215	0.018
	(0.091)	(0.229)	(0.134)	(0.151)	(0.158)
Alcohol	0.081	0.243**	0.004	-0.173	0.327***
	(0.072)	(0.123)	(0.114)	(0.138)	(0.120)
Household income	0.000	-0.000**	-0.000	0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	-0.098	0.054	0.006	-0.402*	-0.271
	(0.105)	(0.180)	(0.154)	(0.232)	(0.239)
Self-employed	-0.034	-0.056	0.022	-0.115*	0.034
	(0.033)	(0.055)	(0.054)	(0.066)	(0.046)
Wave 2 – 3	-0.063	-0.069	-0.064	0.066	-0.031
	(0.096)	(0.211)	(0.128)	(0.182)	(0.178)
Wave 3 – 4	-0.120	-0.130	-0.469***	0.284	-0.082
	(0.105)	(0.227)	(0.180)	(0.190)	(0.177)
Wave 4 – 5	0.092	0.371*	-0.328	0.282	-0.062
	(0.123)	(0.206)	(0.207)	(0.257)	(0.185)
Constant	-5.057***	-5.992***	-4.948***	-5.643***	-4.482***
	(0.559)	(1.049)	(0.750)	(1.170)	(0.730)
Observations	10,281	9,672	9,908	10,095	9,814

Note: Clustered standard errors are in parentheses.
* Result significant at the 10% level.
** Result significant at the 5% level.
*** Result significant at the 1% level.

Table 7. Raw coefficients from probit models for the probability of getting cancer: Hours of work model

	(1)	(2)	(3)	(4)	(5)
Variables	Cancer	Cancer 1	Cancer 2	Cancer 3	Cancer 4
Female	0.194**	0.254*	0.269**	-0.093	0.289**
	(0.081)	(0.154)	(0.122)	(0.138)	(0.129)
Age	0.051***	0.060***	0.047***	0.050***	0.028**
	(0.009)	(0.018)	(0.013)	(0.019)	(0.012)
Degree	0.100	0.232	0.091	0.231	-0.452
	(0.111)	(0.192)	(0.174)	(0.176)	(0.316)
Married	0.018	0.196	0.098	-0.152	-0.091
	(0.089)	(0.191)	(0.139)	(0.143)	(0.151)
Non-white	-0.431	-	-	0.164	-
	(0.321)	-	-	(0.295)	-
Child at home	0.100	-0.020	0.193	-0.014	0.162
	(0.084)	(0.165)	(0.127)	(0.150)	(0.132)
Poor health	0.214**	-0.027	0.141	0.216	0.306**
	(0.098)	(0.228)	(0.155)	(0.161)	(0.155)
Heart	0.426**	-	0.509*	0.136	0.649**
	(0.208)	-	(0.308)	(0.334)	(0.285)
Stroke	-	-	-	-	-
	-	-	-	-	-
Smoking	0.136	-0.100	0.187	0.244	0.090
	(0.096)	(0.235)	(0.138)	(0.157)	(0.163)
Alcohol	0.114	0.243*	0.027	-0.148	0.386***
	(0.079)	(0.140)	(0.125)	(0.152)	(0.129)
Household income	-0.000	-0.000*	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	0.327	0.607	0.510	-0.123*	-
	(0.298)	(0.391)	(0.345)	(0.074)	-
Self-employed	-0.059	-0.065	-0.008	-	0.021
	(0.038)	(0.065)	(0.060)	-	(0.053)
Wave 2 – 3	-0.064	-0.030	-0.072	-0.081	0.004
	(0.104)	(0.216)	(0.136)	(0.205)	(0.185)
Wave 3 – 4	-0.111	-0.195	-0.524**	0.323*	-0.091
	(0.115)	(0.257)	(0.205)	(0.196)	(0.192)
Wave 4 – 5	0.132	0.392*	-0.305	0.319	-0.092
	(0.138)	(0.232)	(0.227)	(0.281)	(0.213)
Constant	-5.024***	-6.255***	-5.478***	-5.034***	-4.738***
	(0.610)	(1.143)	(0.830)	(1.200)	(0.745)
Observations	8,140	7,710	7,842	7,984	7,762

Note: Clustered standard errors are in parentheses.
* Result significant at the 10% level.
** Result significant at the 5% level.
*** Result significant at the 1% level.

Table 8. Balancing covariates in kernel matching for respondents less than 2 years post-diagnosis: Employment model

		Mean			% reduction	*t*-test	
Variable	Status	Treated	Control	% bias	|bias|	*t*	$p > \|t\|$
Female	Unmatched	0.5814	0.48946	18.5		2.08	0.038
	Matched	0.5814	0.58735	-1.2	93.5	-0.1	0.923
Age	Unmatched	58.295	56.402	44.3		5.17	0
	Matched	58.295	58.321	-0.6	98.6	-0.05	0.961
Degree	Unmatched	0.16279	0.15721	1.5		0.17	0.863
	Matched	0.16279	0.16114	0.4	70.4	0.04	0.971
Married	Unmatched	0.7907	0.77443	3.9		0.44	0.66
	Matched	0.7907	0.78676	1	75.8	0.08	0.939
Non-white	Unmatched	0.00775	0.02817	-15.4		-1.4	0.162
	Matched	0.00775	0.0082	-0.3	97.8	-0.04	0.968
Child at home	Unmatched	0.39535	0.38879	1.3		0.15	0.879
	Matched	0.39535	0.40076	-1.1	17.5	-0.09	0.93
Poor health	Unmatched	0.21705	0.13869	20.6		2.55	0.011
	Matched	0.21705	0.22048	-0.9	95.6	-0.07	0.947
Heart attack	Unmatched	0.03876	0.01566	14.2		2.08	0.037
	Matched	0.03876	0.02967	5.6	60.6	0.4	0.689
Stroke	Unmatched	0.00775	0.00827	-0.6		-0.07	0.948
	Matched	0.00775	0.00917	-1.6	-171.5	-0.12	0.902
Smoking	Unmatched	0.18605	0.17376	3.2		0.37	0.714
	Matched	0.18605	0.18347	0.7	79	0.05	0.958
Alcohol	Unmatched	0.37984	0.35668	4.8		0.55	0.585
	Matched	0.37984	0.38145	-0.3	93.1	-0.03	0.979
Household income	Unmatched	407.19	402.07	1.3		0.18	0.858
	Matched	407.19	407.84	-0.2	87.4	-0.01	0.99
Local area unemployment	Unmatched	0.13178	0.16263	-8.7		-0.94	0.345
	Matched	0.13178	0.13575	-1.1	87.1	-0.09	0.926
Self-employed	Unmatched	5.5225	5.5066	1.1		0.12	0.904
	Matched	5.5225	5.5378	-1	4	-0.08	0.935
Wave 2 – 3	Unmatched	0.21705	0.21158	1.3		0.15	0.88
	Matched	0.21705	0.20474	3	-125.1	0.24	0.809
Wave 3 – 4	Unmatched	0.17829	0.24261	-15.8		-1.7	0.09
	Matched	0.17829	0.1858	-1.8	88.3	-0.16	0.876
Wave 4 – 5	Unmatched	0.33333	0.27108	13.6		1.58	0.114
	Matched	0.33333	0.34041	-1.5	88.6	-0.12	0.905
Overall	Unmatched			**10.0**			
	Matched			**1.3**			

Table 9. Balancing covariates in kernel matching for respondents less than 2 years post-diagnosis: Hours of work model

		Mean			% reduction	*t*-test	
Variable	Status	Treated	Control	% bias	|bias|	*t*	$p > \|t\|$
Female	Unmatched	0.58716	0.52696	12.1		1.25	0.211
	Matched	0.58716	0.58323	0.8	93.5	0.06	0.953
Age	Unmatched	58.266	56.186	49.4		5.29	0
	Matched	58.266	58.315	-1.2	97.6	-0.08	0.934
Degree	Unmatched	0.15596	0.14805	2.2		0.23	0.817
	Matched	0.15596	0.15305	0.8	63.2	0.06	0.953
Married	Unmatched	0.77064	0.77599	-1.3		-0.13	0.894
	Matched	0.77064	0.76642	1	21.1	0.07	0.941
Non-white	Unmatched	0.00917	0.02802	-14		-1.19	0.234
	Matched	0.00917	0.00905	0.1	99.4	0.01	0.993
Child at home	Unmatched	0.38532	0.40132	-3.3		-0.34	0.735
	Matched	0.38532	0.37296	2.5	22.8	0.19	0.852
Poor health	Unmatched	0.21101	0.13909	19		2.15	0.032
	Matched	0.21101	0.20402	1.8	90.3	0.13	0.899
Heart attack	Unmatched	0.04587	0.01469	18.2		2.65	0.008
	Matched	0.04587	0.05392	-4.7	74.2	-0.27	0.786
Stroke	Unmatched	0	0	.		.	.
	Matched	0	0	.	.	.	.
Smoking	Unmatched	0.22018	0.17258	12		1.3	0.192
	Matched	0.22018	0.22169	-0.4	96.8	-0.03	0.979
Alcohol	Unmatched	0.37615	0.33744	8.1		0.85	0.396
	Matched	0.37615	0.37211	0.8	89.6	0.06	0.951
Household income	Unmatched	373.54	394.7	-8		-0.8	0.421
	Matched	373.54	371.92	0.6	92.3	0.05	0.961
Local area unemployment	Unmatched	0.01835	0.00884	8.2		1.05	0.296
	Matched	0.01835	0.01086	6.5	21.2	0.46	0.647
Self-employed	Unmatched	5.4257	5.5133	-5.9		-0.61	0.541
	Matched	5.4257	5.4172	0.6	90.3	0.04	0.966
Wave 2 – 3	Unmatched	0.22936	0.2133	3.9		0.41	0.684
	Matched	0.22936	0.22474	1.1	71.2	0.08	0.936
Wave 3 – 4	Unmatched	0.17431	0.23957	-16.1		-1.59	0.112
	Matched	0.17431	0.18155	-1.8	88.9	-0.14	0.89
Wave 4 – 5	Unmatched	0.31193	0.26684	9.9		1.06	0.291
	Matched	0.31193	0.31144	0.1	98.9	0.01	0.994
Overall	Unmatched			**12.0**			
	Matched			**1.6**			

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS

WP13/04 Cormac Ó Gráda: 'Great Leap, Great Famine' March 2013
WP13/05 David Madden: 'Health and Wealth on the Roller-Coaster: Ireland, 2003-2011' May 2013
WP13/06 Karl Whelan: 'Ireland's Economic Crisis - The Good, the Bad and the Ugly' July 2013
WP13/07 Neil Cummins, Morgan Kelly, and Cormac Ó Gráda: 'Living Standards and Plague in London, 1560–1665' July 2013
WP13/08 Neil Cummins, Morgan Kelly and Cormac Ó Gráda: 'Living Standards and Plague in London, 1560–1665' July 2013
WP13/09 Karl Whelan: 'Sovereign Default and the Euro' July 2013
WP13/10 Orla Doyle, Colm Harmon, James J Heckman, Caitriona Logue and Seong Hyeok Moon: 'Measuring Investment in Human Capital Formation: An Experimental Analysis of Early Life Outcomes' August 2013
WP13/11 Morgan Kelly, Joel Mokyr and Cormac Ó Gráda: 'Precocious Albion: a New Interpretation of the British Industrial Revolution' September 2013
WP13/12 Morgan Kelly, Joel Mokyr and Cormac Ó Gráda: 'Appendix to "Precocious Albion: a New Interpretation of the British Industrial Revolution"' September 2013
WP13/13 David Madden: 'Born to Win? The Role of Circumstances and Luck in Early Childhood Health Inequalities' September 2013
WP13/14 Ronald B Davies: 'Tariff-induced Transfer Pricing and the CCCTB' September 2013
WP13/15 David Madden: 'Winners and Losers on the Roller-Coaster: Ireland, 2003-2011' September 2013
WP13/16 Sarah Parlane and Ying-Yi Tsai: 'Optimal Contract Orders and Relationship-Specific Investments in Vertical Organizations' October 2013
WP13/17 Olivier Bargain, Eliane El Badaoui, Prudence Kwenda, Eric Strobl and Frank Walsh: 'The Formal Sector Wage Premium and Firm Size for Self-employed Workers' October 2013
WP13/18 Kevin Denny and Cormac Ó Gráda 'Irish Attitudes to Immigration During and After the Boom' December 2013
WP13/19 Cormac Ó Gráda ''Because She Never Let Them In': Irish Immigration a Century Ago and Today' December 2013
WP14/01 Matthew T Cole and Ronald B Davies: 'Foreign Bidders Going Once, Going Twice... Protection in Government Procurement Auctions' February 2014
WP14/02 Eibhlin Hudson, David Madden and Irene Mosca: 'A Formal Investigation of Inequalities in Health Behaviours after age 50 on the Island of Ireland' February 2014
WP14/03 Cormac Ó Gráda: 'Fame e Capitale Umano in Inghilterra prima della Rivoluzione Industriale (Hunger and Human Capital in England before the Industrial Revolution)' February 2014
WP14/04 Martin D O'Brien and Karl Whelan: 'Changes in Bank Leverage: Evidence from US Bank Holding Companies' March 2014
WP14/05 Sandra E Black, Paul J Devereux and Kjell G Salvanes: 'Does Grief Transfer across Generations? - In-Utero Deaths and Child Outcomes' March 2014
WP14/06 Morgan Kelly and Cormac Ó Gráda: 'Debating the Little Ice Age' March 2014
WP14/07 Alan Fernihough, Cormac Ó Gráda and Brendan M Walsh: 'Mixed Marriages in Ireland A Century Ago' March 2014
WP14/08 Doireann Fitzgerald and Stefanie Haller: 'Exporters and Shocks: Dissecting the International Elasticity Puzzle' April 2014