Tilburg School of Economics and Management
Tilburg University

Employment effects of surviving cancer as adolescent and young adult and the spillover effect to partners: a Dutch population based study

By:
Polle Dankers
SNR: 2013741

Supervised by:
dr. Bettina Siflinger
Silvie Janssen, MSc
dr. Olga Husson

A thesis submitted in partial fulfillment of the requirements for the degree of Master in Econometrics and Mathematical Economics

August 15, 2022

Abstract

This thesis focuses on the effects of a cancer diagnosis on work outcomes of adolescent and young adult (AYA) cancer patients and their partners. To estimate this effect, difference-in-differences methods have been implemented on a treatment and control group.

For patients, significant effects on different work related variables are found. Both employment and hours worked when employed decrease significantly when diagnosed with cancer. Patients are also less likely to be self-employed. In the long term they are less likely to have a permanent contract if they are employed.
The effects are larger for lower educated, female and married patients, as well as for patients with a migration background. The magnitude of effects depends on the type of tumor as well. Especially patients with Central Nervous System tumors are at risk of losing their job. A higher disease stage at diagnosis and more cancer treatments are associated with bigger effects.

The effects on partners of patients are smaller and often insignificant. A cancer diagnosis seems to slightly increase the employment probability, though this effect is insignificant. Employed partners do significantly decrease their number of hours worked.
The effect on employment for partners is mostly found for female partners. Further, the effects strongly differ by the number of hours worked prior to the cancer diagnosis. For partners that were unemployed or worked less than 400 hours in the last year before treatment, the diagnosis is associated with a significant increase in employment probability, while no employment effect is found for partners that worked more than 400 hours in the last year before treatment.

Concluding, both patients as well as partners are affected, although the magnitude of the effects are smaller for partners. Effects on partners differ highly between groups, which is mostly notable for partners who worked a different number of hours prior to diagnosis.

Note that for his master thesis, Ties Siebinga has previously used similar data at the Netherlands Cancer Institute. This study focuses on the effects on partners of AYA's diagnosed with cancer, on employment related outcomes and implements another difference in differences estimator. However, some sections may have some similarity to his previous work due to the usage of similar data and comparable type and goal of the research.

Acknowledgements

I would like to thank my thesis supervisor Bettina Siflinger for her guidance and feedback while writing this thesis. I would also like to thank Silvie Janssen and Olga Husson from the Netherlands Cancer Institute for their guidance, as well as their expert knowledge provided.
Furthermore, I would like to thank the Netherlands Cancer Institute in general for providing the opportunity to write a thesis on this subject and making the data available to me.

Contents

1 Introduction

Over the past decades, the focus of cancer research has shifted more to the group of adolescents and young adults (AYAs), a group into which little research had been done before. This group consists of patients that were diagnosed with cancer when they were between 18 and 39 years old[1]. Due to the recently performed research, the survivor rate has improved steadily by 1 percent annually between 2008 and 2017. Nowadays over 80 percent of AYAs diagnosed with cancer survive at least 5 years. Besides that, the number of AYAs diagnosed with cancer has grown over the past decades. Together, this implies the group of AYA cancer survivors has grown steadily (S. H. Janssen, van der Graaf, van der Meer, Manten-Horst, & Husson, 2021).

Due to the increasing number of AYAs that survived cancer, research into issues these survivors face is increasingly important. AYA cancer survivors face health issues, with the percentage of individuals suffering from a disability and poor physical health being at least twice as big for AYA cancer survivors compared to those that were not diagnosed with cancer. These health issues can affect survivors in the long term (Tai et al., 2012). AYA survivors also experience psychological challenges. AYA cancer survivors are three times more likely to suffer from posttraumatic stress symptoms. Female survivors are more likely to suffer from anxiety and depression as well (Seitz et al., 2010). In general AYA cancer survivors report a worse quality of life than the general population (Quinn, Gonçalves, Sehovic, Bowman, & Reed, 2015).
Surviving cancer, when not specifically targeting AYAs, is also associated with reduced income, employment and hours worked (Zajacova, Dowd, Schoeni, & Wallace, 2015). The employment and hours worked of cacner survivors are affected in the long term (Moran, Short, & Hollenbeak, 2011)[2]. Employment related issues are reduced when survivors are provided with flexible working arrangements, counseling, training and rehabilitation services (Mehnert, 2011). Other research has connected reduced employment for cancer survivors not only to economic problems, but to psychological problems as well (e.g. Peteet (2000), De Boer et al. (2008) and Blinder and Gany (2020)). This stresses the importance of research into employment effects of surviving cancer.

AYA patients likely suffer from employment related issues as well. However, only a small number of studies investigates the effects of surviving cancer on employment specifically for AYAs. Telles (2021) finds that only 6.2 percent of research into AYAs diagnosed with cancer addressed social, economic and demographic

[1] In the Netherlands AYA cancer patients are defined as those that are diagnosed with cancer when 18-39 years old. In other countries ages sometimes differ slightly, with 15-39 year old often being used to define AYAs.

[2] A more comprehensive overview of literature on work related issues is provided in Section 2.

problems[3]. A search for articles concerning economic consequences of surviving cancer for AYAs reveals only few articles in both medical and economics focused journals. The small number of studies that address employment related consequences of surviving cancer for AYAs are survey based. Further they each suffer from other limitations, such as not using a control group.

While there has been little research into the socioeconomic effects of being diagnosed with cancer as AYA, even less research has been conducted on the family of the patients. Only 3.7 percent of research into AYA cancer patients focused on both the AYAs and their family, with 1.2 percent of research only considering the patient's family (Telles, 2021). To my knowledge there is no literature concerning the effect of surviving cancer as AYA on partner employment. However, the partners of patients can face difficulties as well. They may want to compensate for lost income by becoming employed or increasing work hours. On the contrary they could also prefer spending more time with their ill partner and having more time to take care of the partner. Multiple studies find a negative effect of a health shock on spousal labor supply (Jeon and Pohl (2017), García-Gómez, Van Kippersluis, O'Donnell, and Van Doorslaer (2013), Shen, Zheng, and Tan (2019) and Nahum (2007)), but not all literature finds the same effects. A more detailed overview of literature is given in Section 2.

This thesis aims to provide a better insight into the effect of a cancer diagnosis on employment and the number of hours worked when employed, for both AYA cancer survivors and their partners. Since the previously conducted thesis[4] already focused on income effects on the patient itself, the main focus in this thesis will be the effects on partners of patients.
Oncological data of 2456 AYA cancer survivors is linked to longitudinal population data, which includes information on employment and hours worked. From the population data a control group is constructed. Using difference-in-differences (DiD), the data of both the survivors or their partners and the controls is used to find the effect of the cancer diagnosis on work related outcomes. Heterogeneity of the effect with respect to individual characteristics and oncological data is also explored, as well as the dynamics of the effect.
For patients, significant reductions on both employment and the number of hours worked when employed are found. On partners effects are smaller and often insignificant. Partners do seem to slightly increase employment, while decreasing the number of hours worked.

The findings for the cancer survivors can provide a better insight into their employment related issues since it uses longitudinal population data, allowing the

[3]Telles (2021) provides a scoping review of literature about AYAs with cancer between 2015 and 2020. It aims to give an overview of articles on AYAs with cancer, but does not specifically focus on health economic journals

[4]For more information on this thesis, refer to the footnote in the abstract

estimation of the causal effect by comparing the survivors with individuals that did not have cancer. Further it provides insights in what characteristics result in a higher risk of reduced employment and working hours. The results obtained for partners give an unique insight in employment issues for partners of AYA cancer patients.
These insights can be used to determine the magnitude of issues faced by both patients and their partners. It provides an overview of which type of patients are most at risk and could hence profit the most from help, as well as whether partners should be focused on as well.
As mentioned, Mehnert (2011) found that flexible working arrangements, counseling, training and rehabilitation services reduce employment related issues for cancer survivors. These type of services could be offered to survivors, with a more specific focus on those groups that are identified to be at a higher risk of reducing employment and hours worked. Services could be offered to partners as well, but should mainly focus on the patients themselves.

In the next section, a more detailed overview of previous literature is provided. Thereafter the data used is introduced and described. Next the methodologies used are explained and the results are presented along with robustness checks. Lastly the conclusion is presented as well as a discussion including ideas for future research.

2 Literature overview

As mentioned in the previous section only a small number of studies focus on finding how a cancer diagnosis affects employment and hours worked for AYAs[5]. The research that has been conducted previously did however find a correlation between a cancer diagnosis and reduced employment and number of hours worked for AYA patients (Leuteritz et al. (2021) and Parsons et al. (2012)).
While only a small number of papers focus on AYA patients, more research has been conducted into employment effects of surviving cancer without focussing on this specific group. Moran et al. (2011), Heinesen, Imai, and Maruyama (2018) and Jeon (2017) all find similar effects, with surviving cancer being associated with a 5-7 percentage point decrease in employment. Bradley, Bednarek, and Neumark (2002) and Heinesen and Kolodziejczyk (2013) find a similar effect when focusing only on breast or breast and colorectal cancer. Interestingly Moran et al. (2011) find a negative effect on the number of hours worked as well, while Bradley et al. (2002) find that those survivors that are employed work more hours than comparable employed individuals that were not diagnosed with cancer.
The magnitude of effects differs substantially by characteristics of both the patient's background and the type of cancer (Teckle et al. (2018), Gunnes et al. (2016), Syse, Tretli, and Kravdal (2008) and Mehnert (2011)). Most literature finds that the effect is especially large for patients suffering from Central Nervous Cancer (CNS) and soft tissue sarcoma. Being female and having a lower education level are linked to a bigger negative effect as well.

The cancer diagnosis of an AYA patient could affect their partner as well. Different types of effects could have an impact on employment related decisions of their partners. Firstly Lundberg (1985) states that when someone becomes unemployed, their spouse may compensate by starting to work. This is the added worker effect (Nahum, 2007). Nahum adds that spouse may work less in order to take care of their sick partner, known as the caregiver effect. The spouse of a patient may also change preferences in leisure time to spend more time with the patient. This effect is referred to as the joint-leisure effect (Riekhoff & Vaalavuo, 2021). Both can offset the added worker effect. The added worker effect is likely to be smaller in countries with good social benefits, such as the Netherlands. This is due to unemployment benefits mitigating the income loss, shrinking the need to compensate for lost income. Gruber and Cullen (1996) find that indeed the added worker effect is smaller if unemployment insurance is used.

While a spillover effect of the cancer diagnosis on employment and hours worked of partners is likely, only a small number of studies assess such spillover effects of

[5]Since the main focus is effects of cancer on partners of AYA cancer survivors, the overview of literature concerning patients is only described shortly. For a more detailed overview of previous literature on employment effects of surviving cancer and the importance of this area of research, please refer to Appendix A.

health shocks on partners. To my knowledge, no literature is available about the spillover effect of a cancer diagnosis for AYAs on partners.
The literature on the effect of health shocks on spousal labor supply shows different effects. García-Gómez et al. (2013) find that acute hospitalization decreases spousal employment in the Netherlands, focusing on 18-64 year old individuals. Two years after hospitalization, the employment probability of spouses decreased significantly by 0.9 percentage point. The effect is bigger for male spouses compared to female spouses, with a decrease of 1.6 and 0.6 percentage point respectively. Shen et al. (2019) also find that spouse of chronically ill patients significantly decrease their working hours by approximately 4 hours per week. While husbands seem to reduce working hours slightly more than wives, differences between the two are small. Their sample contains 25-64 year old individuals in China. Jeon and Pohl (2017) similarly find a negative effect of cancer on spousal employment using Canadian population data, restricting the sample to individuals aged 25-59. In the first 4 years after diagnosis, both husbands and wives of people diagnosed with cancer incur a decreased employment probability by 2-3 percentage point. In the fifth year after diagnosis, this effect is not significant anymore for husbands, while it is still comparable to the earlier years for wives. Bradley and Dahman (2013) study employment effects of breast cancer on husbands using a survey. Two months after the start of cancer treatment significant reductions in both employment and hours worked are observed. When the same is tested nine months after the start of treatment, no significant effect is found. Nahum (2007) focuses on income effects of spousal sickness absence, while considering people aged between 30 and 58. She finds a negative effect as well, but contrary to García-Gómez et al. (2013) the effect is bigger for female spouse. All of these papers find a negative spousal effect. Hence for these the caregiver and joint-leisure effects seem to outweigh the added worker effect, even though they were conducted in countries with different regulations. Sizes of effects differ which can be due to both the regulations within countries, characteristics of individuals in the used data and the type of health shocks.
Fadlon and Nielsen (2015) on the other hand did not find any effects of a severe non-fatal health shock on spousal labor supply. Their research uses Danish data and focuses on people aged over 45 years old. As a reason for the absence of an effect, they state that insurances accounting for income losses due to the acute illness probably decrease the need to compensate for lost income.

Altogether health shocks are usually associated with small reductions in employment and numbers of hours worked. However the magnitude of effects strongly differs between studies. In some studies the effects differ by gender. It is also likely that effects differ by the type and severity of health shocks, which could partly explain the differences.

3 Data

In this research, data from two sources is used. Data from the Dutch Cancer Registry (NKR) is used for information related to the cancer diagnosis. The NKR includes information on the cancer type, the stage of the cancer and the type of cancer treatment used. From this registry, all individuals aged 18 up to and including 39 who were diagnosed in 2013, had not been diagnosed with cancer before 2013, are not diagnosed with a second tumor up to 2018 and have survived at least 5 years after the diagnosis are selected.

The second source of data used is Statistics Netherlands (CBS) which contains data on multiple personal characteristics, as well as data on households, for the whole Dutch population. From the CBS general information on birth year, gender, migration, marriage status and households is used together with work related outcomes such as the number of hours worked and being self employed. The characteristics of individuals and households are based on the municipal administration. The data on the number of hours worked and type of contract is based on wages declarations to the tax authorities and available in all time periods. Data on self employment comes from tax records as well, but is only available from 2011 onwards. Educational data is also used, which will be elaborated on in section 3.3.
The data from the NKR is linked to the CBS data. Note that only patients that likely stayed in the Netherlands for the whole observed period are used in the analysis. This selection is implemented since work related data for people living outside of the Netherlands is not available. Therefore an employed person living abroad would be considered unemployed due to the unavailability of this data. This removal is based on CBS data that should include the whole Dutch population in a given year, which is available from 2011. If a patient is not in this data set for any year between 2011-2018, the individual is excluded from the analysis. Patients that did survive 5 years but died in 2018 were also removed from the data, since they obviously could not work for the whole of 2018. Therefore their employment and hours worked would decrease substantially due to dying from cancer instead of the difficulties incurred when surviving cancer, while the effects when surviving cancer are of interest. After this selection 2456 patients are left.

3.1 Data about patients

To find the treatment effect of surviving cancer on work, a comparison between a treatment and control group is required. Using both these groups difference-in-differences (DiD) can be used to estimate this effect. Later in Section 4: Methodology this will be elaborated upon further. For this purpose people included in the CBS data aged 18 up to and including 39 are selected and compared to the patients. The patients will from now on be referred to as the treatment group.

There are some clear differences between the treatment group and the whole population of 18-39 year old people in the CBS database. The average birth year of the treatment group is significantly higher compared to other observations. This is as expected since the likelihood of being diagnosed with cancer increases with age (DeSantis et al., 2014). Further, the treatment group is less likely to have a migration background and females are more likely diagnosed with cancer compared to males.
Since these variables are likely to influence work related variables (Hartog & Salverda, 2018) as well as the trend in labor participation, using the whole population aged 18-39 will probably violate the parallel trends assumption[6]. Therefore a control group is created by matching observations based on these characteristics with exact matching. This matching method finds individuals that have the exact same characteristics for birth year, gender and migration background. Using this each patient is matched with 10 controls, similarly to Teckle et al. (2018). This is done without replacement, hence when a control is matched to a patient he/she cannot be matched to another patient as well. After matching the control group consists of 24560 individuals, with exactly the same characteristics of birth year, gender and migration background as the treatment group.

In Table 1 the characteristics of the treatment and control group are tabulated. Due to the matching procedure, the distribution in gender, migration background and age are exactly the same for the treatment and control group. AYA cancer survivors are more often female than male. Only 18.9 percent of patients has a migration background, which is lower than the 25 percent found when all individuals between 25 and 35 are considered (Statistics Netherlands, 2022). Within the group of AYAs, patients are relatively old with 54.4 percent of patients being aged 33 or older. This can be explained by an increasing probability of suffering from cancer when age increases. Patients mostly achieved secondary or higher education, with only 5.5 percent of individuals with known education level having achieved primary education as the highest level.

[6]This is an important assumption made for analysis conducted which will be elaborated on in the methodology section

Table 1: comparison treatment and control group

	Treatment		Control	
	Frequency	Percentage	Frequency	Percentage
Male	1036	42.2%	10360	42.2%
Female	1420	57.8%	14200	57.8%
migration background				
No	1992	81.1%	19920	81.1%
First generation	213	8.7%	2130	8.7%
Second generation	251	10.2%	2510	10.2%
Age < 26	393	16.0%	3930	16.0%
$26 \leq$ Age < 33	727	29.6%	7270	29.6%
Age ≥ 33	1336	54.4%	13360	54.4%
Education level in 2012				
Primary	112	4.6%	1166	4.7%
Secondary	1216	49.5%	12117	49.3%
Higher	722	29.4%	7195	29.3%
Missing	406	16.5%	4082	16.6%

Note: Age refers to the age of an individual at the end of 2013. If someone turns 26 in 2013 he/she is therefore considered 26 in 2013, even if he/she only became 26 years old on the 31st of December 2013.

In Table 2 the clinical data is described by the frequencies and proportions of cancer types, tumor stages at diagnosis and types of cancer treatment used. Note that the stage at diagnosis is constructed by concatenating stages from multiple sources into one variable. This approach is similar to the approach used in De Rooij et al. (2018) and Husson et al. (2020). TNM and Ann Arbor stages are used, which are often used for different types of tumors. Both define the severity of a tumor. The defined stages using these types are reduced to 4 classes, from least severe to most severe. TNM stages are used first. When the TNM stage is missing Ann Arbor stages are used. Patients mostly have low stages at diagnosis, which can be explained by the selection on patients that survived at least 5 years after diagnosis. Higher stage tumors lead to a lower survival rate.
The most occurring tumor types are skin, breast and male genital. In later analysis, effects on types of cancer will be investigated separately for all types with at least 75 treated individuals. This is chosen since fewer observations lead to less reliable estimates, while CNS tumors are likely to result in big effects[7]. Effects for CNS are therefore still estimated and its number of observations is used as threshold.
The percentages of types of treatment add up to more that 100 percent. This can be explained by patients receiving multiple types of treatment.

[7]In previous literature CNS was associated with the biggest employment effects. See Appendix A for more details.

Table 2: Oncological data on patients

	Frequency	Percentage
Type of tumor		
Breast	493	20.1%
Bone, cartilage and soft tissue	67	2.7%
Central Nervous System	75	3.1%
Endocrine	123	5.0%
Hematologic	321	13.1%
Head and neck	44	1.8%
Skin	510	20.8%
Lung	22	0.9%
Male genitals	433	17.6%
Female genitals	193	7.9%
Digestive system	115	4.7%
Urinary tract	50	2.0%
Other	10	0.4%
Stage		
I	1351	55.0%
II	539	21.9%
III	233	9.5%
IV	106	4.3%
Missing	227	9.2%
Type of treatment		
Organ surgery	1340	54.6%
Local surgery	709	28.9%
Radiotherapy	720	29.3%
Chemotherapy	1014	41.3%
Hormonal therapy	1014	41.3%

In Table 3 the frequencies of different tumor types are given, together with stages for these types. Here stages 2-4 are concatenated due to CBS regulations[8]. Clear differences in stages between tumor types can be found. While for endocrine, skin and genital tumors are predominantly stage 1, breast, hematological and digestive tumors are mostly found in higher stages. This can have multiple reasons. For some types of tumors higher stages can lead to a low survival rate. Some types may also be more likely to be diagnosed at a low stage. The stages in which tumors are found can impact the magnitude of effects found. Differences between effects for different types of tumors can partly be due to the distribution of stages.

[8]The CBS does not allow output to include very small frequencies due to privacy concerns. A part of the tumor types had very small amounts of individuals with a certain stage. Therefore these stages are concatenated such that the frequencies of most combinations of types of tumors and stages are big enough.

Table 3: Oncological stages for different types of cancer

		Stage		
		I	II-IV	Missing
Type of cancer	Breast	160	332	1
	Bone, cartilage and soft tissue	36	21	10
	Central Nervous System	x	x	x
	Endocrine	120	3	0
	Hematologic	33	183	105
	Head and neck	17	26	1
	Skin	434	67	9
	Lung	x	x	x
	Male genitals	317	107	9
	Female genitals	158	35	0
	Digestive system	34	76	5
	Urinary tract	29	18	3
	Other	x	x	x

Note: Due to output regulations from the CBS, the information has been removed for tumor types with few observations in certain stages. These are indicated with an x.

As previously denoted, the outcomes of interest are work related. The first outcome of interest is employment. For employment a definition from Eurostat is used, which considers an individual employed when working at least one hour per week. This definition is similar to the definition used by the U.S. bureau of labor statistics for people working as a paid employee. As yearly data is used, a threshold of 52 hours is implemented for being employed.
The other main outcome of interest is the number of hours worked. When the number of hours worked are analysed, only those individuals that are employed are considered. This is done since otherwise a large proportion of the reduced number of hours is be due to people that stop working. Hence it would partly measure the same effect as the analysis of employment. Further, the logarithm of the number of hours worked is used in the analyses. Without the logarithm, changes in hours of a similar proportion would show a much bigger effect for people that previously worked many hours compared to those that did not. This would make the estimated coefficients less insightful. Using the logarithm, the percentage change can be estimated and changes for those individuals that did not work many hours before are taken into account as well.
Besides these main outcomes of interest, the effect of a cancer diagnosis on the probability of having a permanent contract is also considered for individuals that

are employed. Lastly effects on self employment are briefly considered[9].

In Table 4 on overview of these outcome variables be found for both the treatment and control group in the last year before treatment. Before treatment, the treatment groups is slightly more likely to be employed compared to the control group, with an employment of 82.6 percent compared to 80.6 percent for the control group. In the results section this level difference will be assessed further, as well as an explanation on whether it is a problem for the analysis. Individuals in the treatment group are also slightly more likely to have a permanent contract when employed, with 64.0 percent of them having a permanent contract compared to 62.6 percent in the control group. The number of hours worked when employed and the probability of being self employed is very similar for the groups. Both groups work an average of approximately 1470 hours when employed, while for both groups slightly below 7 percent of individuals is self employed.

Table 4: Employment, hours worked, self-employment and contract type in 2012

	Treatment group	Control group
Employment	0.826	0.806
Hours worked	1474	1470
Self-employed	0.069	0.068
Permanent contract	0.640	0.626

3.2 Data about partners of patients

To consider the effect on the partners of patients, similar data is used as previously discussed. Using household data, households can be identified as well as the position of individuals within the household. However the household numbers do not refer to a combination of two partners uniquely. When something changes within the household, such as a change of the address or composition of the household, a new observation is added which may get a new household number. Therefore alternative household numbers are created by using the combinations of individual identifying numbers, such that partners for which the household changes but who do not split up are still considered to be part of a continuing household.
Using this modified household data set partners are identified. Here the definition of a partner is someone who is a member of a household as partner. For this the individual does not necessarily have to be married or have a registered partnership.

[9]Data on self employment comes from a different source within CBS data than the other mentioned outcome variables. This data source is only available from 2011. In the analysis considering self employment, 2010 is therefore not considered.

Partners are included in the data when their relationship started before 2012[10]. The final year included for a partner is either the last year partners are together if they broke up before 2018 and 2018 otherwise. If the relationship is broken up before 2013, he or she is also excluded from the data. Partners who are at least 60 years old in 2013 are excluded as they will be at pension age in 2018. For these partners changes in work related variables are likely due to pension related changes rather than the diagnosis of their partner. This procedure is used for both the treatment group and potential control group. The partner treatment group consists of the partners of patients. The resulting partner treatment group consists of 1252 individuals.

Next the control group is matched to these treated partners using a combination of exact matching and nearest neighbor matching. The pool of potential controls identified before is used. Controls are matched to have the exact same characteristics for migration background, gender and birth year as treatment individuals, in line with the matching procedure of the patient control group. This is combined with nearest neighbor matching on the number of years partners have formed a household before the year of treatment and the birth year of the partner. The nearest neighbour is chosen based on the Mahalanobis distance[11]. To each partner of a patient five controls are matched.
Thus for a treated individual, the matching procedure first identifies individuals with the same migration background, gender and birth year from the pool of potential controls. Within this group of potential controls, the five individuals with the most similar duration of their relationship before the year of treatment and birth year of the partner are selected as controls. This is done without replacement, hence when a control is matched to a partner of a patient, he/she cannot be matched to another partner as well. The five matches are less than in the patient analysis, since the population of people with a partner is smaller. Using exact matching on some of the covariates it is not always possible to find 10 similar people for a given treated partner.
It is possible that a matched control partner has been part of a household with a patient before, in which case the control is deleted from the data to prevent a spillover effect in the control group. This only happened in 4 cases.

Table 5 shows the characteristics of individuals in both the treatment and control group. Proportions in both groups are comparable for all characteristics matched on, as should be the case. The partners are predominantly male, which could be expected since females were represented more in the patient group. Another

[10]For the breakup comparison at the end of this chapter, analysis is conducted without the exclusion of people that started their relationship in 2012. The reason behind this is that the exclusion of 2012 is only needed to assess the parallel trends assumption, which is not relevant for the breakup comparison.

[11]Mahalanobis distance is used instead of Euclidean distance since it takes into account correlation within the data and is scale invariant.

difference with the patients is that even a larger proportion of individuals belongs to the oldest age group, with 72 percent of individuals being at least 34 years old. This is as expected as well, since older people are more likely to cohabit with a partner. Notable is that both groups have a very similar educational background and similar number of children in 2013 while this was not used in matching. This strengthens the hypothesis that matching on aforementioned variables probably removes the main differences between groups.

Table 5: comparison characteristics treatment and control group partners

	Treatment		Control	
	Frequency	Percentage	Frequency	Percentage
Male	826	66.0%	4102	65.8%
Female	426	34.0%	2129	34.2%
Migration background				
No	1034	82.6%	5143	82.5%
First generation	119	9.5%	600	9.6%
Second generation	99	7.9%	488	7.8%
Age < 26	26	2.1%	140	2.2%
$26 \leq$ Age < 33	323	25.8%	1594	25.6%
Age ≥ 34	903	72.1%	4497	72.2%
Education level in 2012				
Primary	19	1.5%	96	1.5%
Secondary	409	32.7%	2026	32.5%
Higher	385	30.8%	1936	31.1%
Missing	439	35.1%	2173	34.9%
Average relationship spell before 2013	4,581	-	4,583	-
Average number of children	1,326	-	1,325	-

Note: Age refers to the age of an individual at the end of 2013. If someone turns 26 in 2013 he/she is therefore considered 26 in 2013, even if he/she only became 26 years old on the 31st of December 2013.

In the appendix Table B.1 can be found with an overview of the types of cancer as well as cancer stages of the patients with a partner. As may be expected, ratios mostly resemble those of the whole group of patients.

In Table 6, an overview is given of the outcome variables in the last year before treatment. Partners and their controls are more likely to be employed and when employed they work more hours on average compared to the patients and their respective controls, which is discussed in Table 4. This can be caused by the partners being older, while living together with a partner itself may affect work decisions as well. Further both the treatment and control group have a similar

employment probability, with 84.6 and 84.1 percent employed respectively. The number of hours worked when employed is comparable as well, with the treatment group working 1731 hours on average and the control group working 1721 hours on average.

Table 6: Employment and hours worked in 2012

	Treatment group	Control group
Employment	0.846	0.841
Hours worked	1731	1721

3.2.1 Breakup rate comparison

It is possible that a diagnosis with cancer impacts the breakup rate. As partners are only used in time periods where they are still together, this may introduce a selection bias since dropout is not equal between the treatment and control group if breakup rates differ.
Carlsen, Dalton, Frederiksen, Diderichsen, and Johansen (2007) show that no clear difference in the divorce rate can be found for cancer patients compared to the rest of the population. Syse and Kravdal (2007) similarly do not find a clear effect. In the first 2.5 years divorce rates for people diagnosed with cancer are lower, but this is mostly found for patients diagnosed with cancer types that have a bad prognosis, or cancer that has already spread to other organs. Since the data used consists only of patients that survived at least 5 years after diagnosis, these types of cancer are underrepresented within the data. The effects on divorce found for them are less relevant in this research.
Hence it is unlikely that the breakup rate differs significantly between both groups. To assess this assumption the survival of relations is analyzed. As shown in (Cameron & Trivedi, 2005), this is done by estimating the discrete-time survivor function shown in equation (1):

$$S^d(t) = P(T \geq t) = \prod_{j:t_j \leq t} (1 - \lambda_j), \quad \lambda_j = P(T = t_j | T \geq t_j) \tag{1}$$

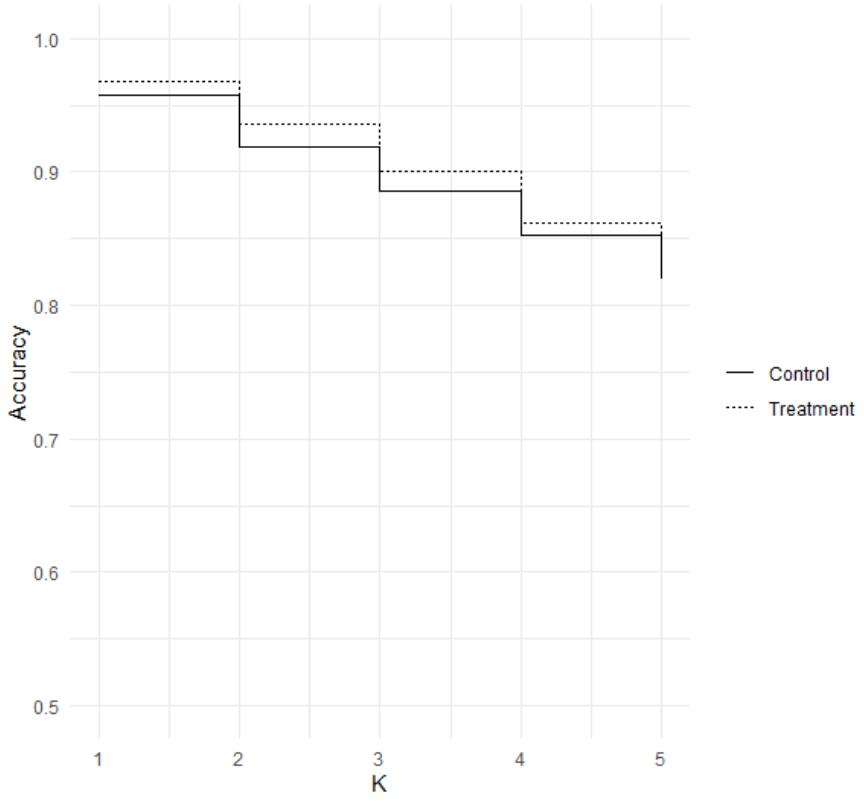
This is the probability that a spell will last at least t years. Note that a spell is the number of years the partners stay together after 2012, which is denoted by T. This survivor function can be estimated using the Kaplan-Meier estimator (Kaplan & Meier, 1958), which is given in equation (2):

$$\hat{S}_T(t) = \prod_{j:t_j \leq t} \frac{\eta_j - \delta_j}{\eta_j} \tag{2}$$

Where η_j denotes the number of spells that have not yet ended before t_j and δ_j the number of spells that end at t_j. This is a nonparametric estimation of $\Pi_{j:t_j \leq t}(1 - \lambda_j)$ using $\hat{\lambda}_j = \frac{\delta_j}{\eta_j}$.

The survivor function is estimated for both the treatment and control group using the Kaplan-Meier estimator. In Figure 1, these estimated survivor functions are plotted. The survival rate of relationships in the treatment group is slightly higher, but the differences are very small.

Figure 1: Kaplan-Meier estimated survival function for relationships of the treatment and control group



To asses if there is a significant difference between the survival function of the relationships of both groups a Log-Rank test is performed (Bland & Altman, 2004). The p-value obtained is 0.53, so there is no significant difference in the survival rate of relationships between the control and treatment group.

3.3 Missing values education

The data used for education comes from two types of sources. The first type of source is different registries on education, from which the majority of education data is obtained. This data is only available for people that received their education from a publicly funded educational institution. The data is unavailable for people that were educated at a private institution or abroad. The registry data is more complete for younger people. The percentage of missing values increases when birth year decreases.

The other source used is the labor force survey (EBB) conducted by Statistics Netherlands. This consists of a random sample of everyone aged 15 through 90 which are not living in institutions such as elderly homes. For this data source, the number of missing values does not depend on whether someone received private education and age.

To gain more insight on the magnitude of the problem with missing values, the frequency of missing values for both the treatment and control group is compared. Tables 7 and 8 show the fraction of observations with missing values jointly with the fraction of each type of education achieved for both the treatment and control group in 2010 and 2013. It can be noted that missing values are less frequent in 2013, which is mostly due to data from another registry being added. This has the biggest impact on older people in the population.
Differences between the treatment and control group seem negligible. To test formally if the educational data is independent of the control and treatment group, a Pearson's Chi-squared test for independence is conducted (McHugh, 2013). No clear dependence between the two could be found, as the p-value of the chi-squared test is 0.706 in 2010 and 0.887 in 2013. Hence the educational data is independent from the group allocation[12].

Table 7: comparison education treatment and control group in 2010

	primary	secondary	higher	missing
control	0.037	0.424	0.258	0.281
treatment	0.037	0.436	0.252	0.275

[12]The tests here are conducted on patients instead of partners since their lower average age means they are more likely to still follow education. Therefore differences are more likely to occur in this group. For the partners, educational data seems independent as well and is more time invariant.

Table 8: comparison education treatment and control group in 2013

	primary	secondary	higher	missing
control	0.039	0.450	0.317	0.194
treatment	0.037	0.450	0.315	0.199

Even though the educational data does not depend on group allocation, the missing values cannot be ignored. As previously explained, the frequency of missing values depends on variables such as birth year and migration background. Simply excluding observations with a missing value would lead to biased results, as people with certain characteristics are more likely to be selected into the data leading to a selection bias. Hence the missing data must be imputed.

3.4 Imputation of missing values

The first method used to impute missing values is simply carrying observed values in other time periods backward and forward within individuals. E.g. when a value is known in 2017 but unknown in 2016 and 2015, then the value of 2017 is imputed in the earlier time periods. The same is done when only previous data is known. When data is e.g. known up to 2015, the educational data of 2015 will be imputed at later time periods. This imputes reasonable estimates of what the true education level is, since the education level does not vary much over time and hence observations at other time periods are likely the same. This is especially the case since most missing values occur for the older people within the population. At older age, people are less likely to have their education vary over time, so this method is appropriate for these people. Carrying backwards was implemented before forwards since the CBS data includes a new registry from 2013 which solves some of the missing values.

The leftover missing values, which are missing for an individual in all observed time periods, are imputed using k-nearest neighbor (kNN) imputation. Troyanskaya et al. (2001) find that kNN imputation yields better results than Single Value Decomposition (SVD) based imputation methods. Jadhav, Pramod, and Ramanathan (2019) show that kNN imputation outperformed all other tested methods, including four multiple imputation methods. Ribeiro and Freitas (2021) found that the accuracy of their random forest model was the highest after imputing missing values in longitudinal data using kNN compared to other methods. Further, kNN imputation is a form of single imputation, which makes analysis more convenient since only one dataset needs to be dealt with. Due to the high performance and ease of use, kNN imputation is implemented.

Imputation with kNN chooses the k observations which are closest to an observation with a missing value and then imputes the missing value with the value that

is most frequent within those k observations. The value of k used will be introduced later. The kNN imputation algorithm is implemented such as described in Kowarik and Templ (2016).
To find which observations are the closest, Gower distance is used as described by Kowarik and Templ. The Gower distance can combine the distance of different types of data. This makes it well applicable for this implementation since both categorical and numerical data are used. To calculate the distance, ξ_{ijm} is defined as the distance contribution for the mth variable between observations i and j. Here $i = 1, ..., n$, $j = 1, ..., n$ and $m = 1, ..., p$, where n is the number of observations and p is the number of variables used for imputation. For numerical data ξ_{ijm} is defined as:

$$\xi_{ijm} = |x_{im} - x_{jm}|/r_m \tag{3}$$

Where x_{im} is the mth variable of the ith observation. r_m denotes the range of the mth variable. The division by r_m ensures that a variable with small absolute differences but big relative differences are not underrepresented in the distance calculations.
The distance for categorical variables is defined as:

$$\xi_{ijm} = \begin{cases} 0 \text{ if } x_{im} = x_{jm} \\ 1 \text{ if } x_{im} \neq x_{jm} \end{cases} \tag{4}$$

Where x_{im} is defined as before. This formulation of ξ_{ijm} can be used for binary data as well.
The total distance between two observation is calculated by a weighted average:

$$\Xi_{ij} = \frac{\sum_{m=1}^{p} w_m \xi_{ijm}}{\sum_{m=1}^{p} w_m} \tag{5}$$

Where w_m is the weight given to each variable. This can be used to make the distance for certain variables more important than others, which is useful if some variables are more informative for the education level than others. These weights can be neglected if all variables should be accounted for similarly. In that case $w_m = 1 \ \forall m$. The importance of variables can be estimated by using a random forest regression. The ranger package in R by Wright and Ziegler (2017) is used to implement the random forest, which is included in the kNN package by Kowarik and Templ (2016). Which of these types of weights works better is tested later in this paragraph.

Let now i be an observation with a missing value for education. The kNN algorithm picks the k observations with the smallest Ξ_{ij} without education missing. The the education level which is most occurring within the k nearest neighbors is imputed for observation i.

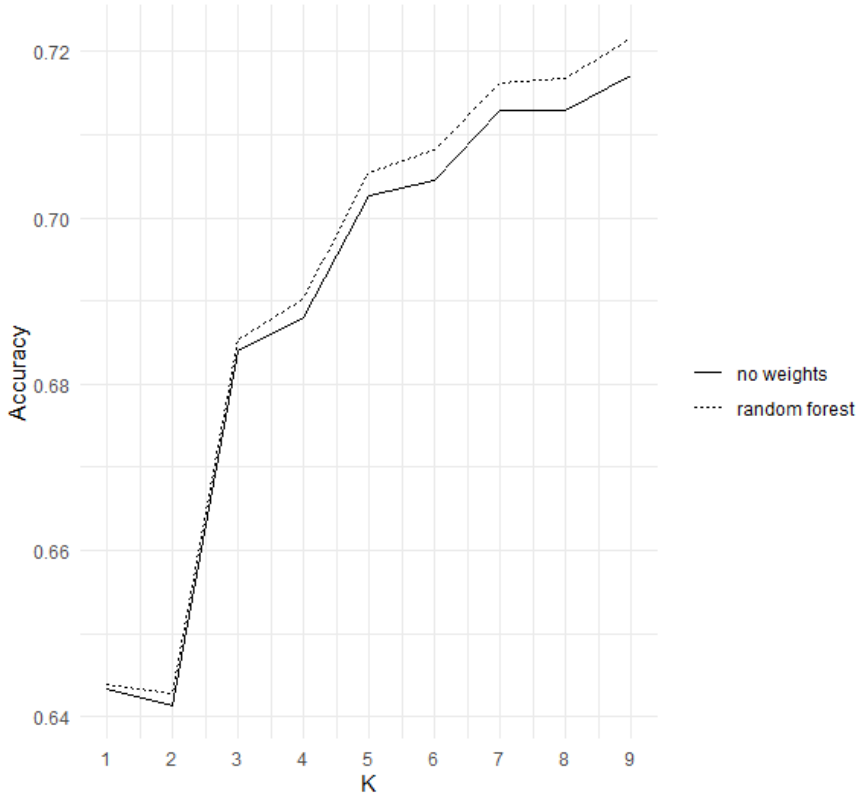
To test the performance of the algorithm in different settings, data sets are created consisting of only individuals with known education. Education is set to

missing for the same proportion of individuals as the proportion of missing education values within the whole data set. This creates test data on which the performance can be tested. Five data sets are created using this methodology to take the average over these five evaluations, such that results are more robust to the selection of data and missing values.

As previously denoted, the weight w_m can be neglected or estimated using a random forest regression. Both these methods are implemented, together with different values of k: $k = 1, 2, 3, 4, 5, 6, 7, 8, 9$. These values of k are chosen as Beretta and Santaniello (2016) note that k should be smaller than 10 to avoid distortion in the data.
The variables used for imputation are the same covariates as used in the analysis along with the hourly wage of an individual in the given year, constructed by dividing the number of hours worked by the salary received. This is of course strongly related to the outcome variable employment, since people that are not employed do not have an hourly wage. However, it does increase the accuracy strongly. Moons, Donders, Stijnen, and Harrell (2006) show that including the outcome variable in imputation should be preferred.

In Figure 2 the accuracy for these different settings can be found. The highest accuracy is found when weights are chosen using a random forest regression and $k = 9$. Therefore the data is imputed using the kNN algorithm with the random forest weighting and using the 9 closest neighbors. Thus for a given observation with missing educational data in the original data, the 9 complete observations with the smallest weighted sum of Gower distances are selected. This is based on distances for birth year, migration background, gender, marital status and hourly wage. The weight of the distance corresponding to each of these variables is determined using a random forest. The missing education level is then imputed with the most occurring level within these 9 closest observations.

Figure 2: kNN imputation accuracy using different k and weighting methods



4 Methodology

In the following section all methodologies used to analyse the effects of being diagnosed with cancer as AYA, or being a partner of an AYA diagnosed with cancer, are discussed. Before the methods are explained, some notation used throughout the section is introduced. The notation is mostly taken from Lechner (2011). Let D denote the treatment. The realization d_i for individual i can be defined as:

$$d_i = \begin{cases} 0 \text{ if individual } i \text{ is part of the control group} \\ 1 \text{ if individual } i \text{ is part of the treatment group} \end{cases}$$

Further time is denoted by T, with t the possible realizations. For the first part of this section, t only has two possible values: $t = 0$ before treatment takes place and $t = 1$ if treatment has happened. Therefore:

$$t = \begin{cases} 0 \text{ if year}<2013 \\ 1 \text{ if year}\geq 2013 \end{cases}$$

In later sections, t sometimes has more possible values as more than two time periods are used. In such cases this is clarified.
Y are the outcomes, with realizations y_{it} for individual i in time period t. y_{it}^d denotes the outcome that would be realized for individual i at time t for a certain value of d_i. This results in two potential outcomes: y_{it}^0 and y_{it}^1 of which only one can be observed, with the other one a counterfactual.

4.1 Effect of treatment

As previously explained, this thesis aims to find the effect of a cancer diagnosis on different outcome variables. Let the effect of treatment be defined as the difference between the two potential outcomes:

$$\alpha_i = y_i^1 - y_i^0 \tag{6}$$

In this application, the effect of treatment on those individuals that are treated is of interest, since we want to know the effect of a cancer diagnosis specifically on the group of AYA cancer survivors. Therefore the average treatment effect on the treated (ATT) must be estimated:

$$\begin{aligned} ATT &= E[\alpha_i | d_i = 1] \\ &= E[y_i^1 - y_i^0 | d_i = 1] \\ &= E[y_i^1 | d_i = 1] - E[y_i^0 | d_i = 1] \end{aligned} \tag{7}$$

As previously explained, one of the potential outcomes is unobserved. This means that the ATT cannot be estimated by simply plugging averages into equation (7).

In particular, $E[y_i^0|d_i = 1]$ cannot be directly observed, since it is a counterfactual. Therefore other methodologies must be used to estimate the ATT. When information from before treatment is available, differences-in-differences (DiD) is an attractive method to estimate the ATT (Lechner, 2011). It uses information from before and after treatment of both a control and treatment group to estimate the ATT. Since all this information is available, DiD is an attractive method for estimation of the ATT in this application.

4.2 Difference-in-Differences

In this section a standard DiD framework is discussed, with only one time period before and after treatment. It uses a control and treatment group, which are discussed in Section 3. While $E[y_i^0|d_i = 1]$ is not observed directly, the changes of y_{it} over time can be observed for both the treatment and control group. DiD uses both these changes and uses the difference in the change within both groups to estimate the treatment effect.

A couple of assumptions are needed for DiD, see e.g. Lechner (2011). These are made for the case with only one time period prior to treatment and one time period after treatment. The first of these is the Stable Unit Treatment Value Assumption (SUTVA):

Assumption 1 *SUTVA:* $y_{it} = d_i * y_{it}^1 + (1 - d_i)y_{it}^0 \ \forall i, t$

SUTVA implies that only one of the two potential outcomes is observed. This assumption states that there may not be a spillover effect of the treatment. It is violated when the treatment affects not only the treated individual itself, but also not treated individuals. In this application the treatment effect could spill over to people that are close to the diagnosed person. However, the people close to diagnosed people are a small fraction of the Dutch population, which is used in the research. The same similarly holds when effects on partners are considered. Therefore the effect of this possible minor violation are insignificant. SUTVA still holds.

Next exogeneity is assumed for the covariates, denoted with X:

Assumption 2 *Exogeneity* $X^1 = X^0 = X$

The assumption states that covariates X should not be influenced by the treatment. As previously stated, the treatment and control groups were matched by time invariant variables in order to create a comparable group. For these variables this assumption holds. Previous tests showed that no significant difference can be found in both education level and the breakup rate for both groups. Additional analysis are performed to check this for the number of children as well, which will be discussed in the robustness section.

Thirdly the existence of anticipation effects is ruled out:

Assumption 3 *NEPT: treatment has no effect on the pre-treatment population;* $\alpha_0 = 0$

Where α_0 is the treatment effect in the pre-treatment time period. This is violated when an individual already knows it will be treated in a later time period and anticipates to this. In this setting people cannot know they will be diagnosed with cancer in a later time period, so they cannot anticipate to it[13]. Therefore NEPT can be assumed to hold. This assumption can be tested by estimating treatment effects prior to treatment which is simultaneously used to assess the plausibility of parallel trends, the fifth assumption introduced here.

Common support must be assumed as well:

Assumption 4 *CS:* $0 < P[d_i = 1|X] < 1$

This ensures that for each value of X a treated and untreated individual is observed. The assumption is required since the treatment and control groups are compared given covariate combinations. It is violated when an individual is certainly diagnosed or not diagnosed with cancer considering the information in the covariates. This is not the case, as any individual can either be diagnosed with cancer or not regardless of their other characteristics.

Lastly the main assumption for DiD is the parallel trends assumption, also referred to as the common trend assumption. This assumption states that the both the treatment and control group would have followed the same trend in absence of treatment. This is needed as DiD uses the differences in the trend between both groups to estimate the unobserved counterfactual, $E[y_i^0|d_i = 1]$. Formally, when only considering time period $t = 0$ before treatment and time period $t = 1$ after treatment:

Assumption 5 *PT:*

$$\begin{aligned} &E[y_{i1}^0|d_i = 1] - E[y_{i0}^0|d_i = 1] \\ &= E[y_{i1}^0|d_i = 0] - E[y_{i0}^0|d_i = 0] \\ &= E[y_{i1}^0] - E[y_{i0}^0] \end{aligned} \tag{8}$$

Assessing the plausibility of this assumption is more difficult. The first method which is often is plotting the outcome variable over time for both the treatment and control group. These plots are used to visually assess if trends are parallel. While this method can give an indication of the plausibility of the assumption, it is not very reliable since it is only eyeballing and not a statistical test.
Another common way to assess the plausibility of parallel trends is estimating the

[13]While anticipation is unlikely, it is possible that patients already suffered from symptoms in the year before diagnosis. In the robustness section, estimation will be performed while controlling for anticipation to check whether this results in different estimates.

treatment effects in lead periods. This can be done by estimating the dynamic treatment effect, which will be explained in later sections, for time periods prior to treatment[14]. For the parallel trends to be plausible, the treatment effect must be zero in the pre-treatment time periods (Pischke, 2005). If the estimated effects before treatment significantly differ from zero, it is unlikely that parallel trends hold. If they are not significantly different from zero, the assumption is more plausible.
To test the joint nullity of the estimated effects prior to treatment, a Wald test is conducted (Cameron & Trivedi, 2005). This tests the null hypothesis that the estimated treatment effects are zero in each pre-treatment time period. A large p-value for this test means that parallel trends likely hold.
The mentioned tests are all applied. The results of these tests can be found in Section 5: Results, together with explanations. It must be noted that these tests can only give an indication of the plausibility of parallel trends. Only time periods before treatment can be used, which means that the parallel trends after treatment cannot be tested. As stated by Kahn-Lang and Lang (2020): "parallel trends in the period prior to treatment is suggestive of counterfactual parallel trends, but parallel pre-trends is neither necessary nor sufficient for the parallel counterfactual trends condition to hold. Any available tests of identification should be seen as a complement to, not a substitute for, logical reasoning."
As previously explained, control individuals are chosen by matching on multiple individual characteristics[15]. Further, getting cancer is thought to be close to random for AYA patients (American Cancer Society, 2022). Considering this, in addition to tests results on the pre-treatment trend, it is likely that the parallel trends assumption holds.

Under these assumptions, DiD can estimate the ATT by using outcomes in pre and post-treatment time periods from both the control and treatment group. The ATT can be computed using equation (9):

$$ATT = (E[y_{i1}|d_i = 1] - E[y_{i1}|d_i = 0]) - (E[y_{i0}|d_i = 1] - E[y_{i0}|d_i = 0]) \quad (9)$$

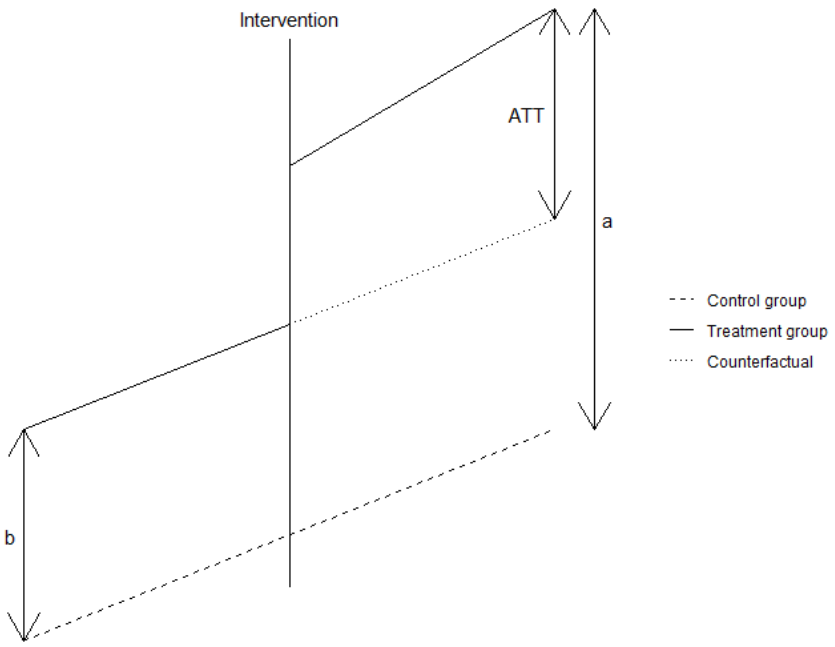
The identification can be straightforwardly proven using given assumptions. In Appendix C.1 this proof can be found. To clarify the DiD framework, equation (9) is showed visually in Figure 3, with for simplicity $a = (E[y_{i1}|d_i = 1] - E[y_{i1}|d_i = 0])$ and $b = (E[y_{i0}|d_i = 1] - E[y_{i0}|d_i = 0])$. It illustrates the observed outcomes for both groups, along with the counterfactual outcome for the treatment group which develops similarly to the control group when the PT assumption holds. Then the difference between treatment outcomes before and after treatment[16] can be subtracted to find the ATT, as formally given in equation (9).

[14]See equation 11 and equation 15 for the estimation of the dynamic effects using different methods.

[15]See Section 3 for more information on the matching procedure

[16]The name difference-in-differences comes from this as well: the ATT is estimated by taking the difference in differences before and after treatment

Figure 3: Visualization Difference-in-Differences



In the next section, an estimation method for DiD is introduced which uses regressions. This method is also expanded to a setting where multiple time periods are available both before and after treatment, as is the case in this application.

4.3 Two-Way Fixed Effects

DiD can be estimated using a regression formulation. Equation (10) shows the framework which can be used. The parameters can be estimated using a simple linear regression.

$$y_{it} = \beta_0 + \beta x_{it} + \gamma d_i + \lambda_t + \alpha(d_i * t) + \epsilon_{it} \tag{10}$$

β_0 is an intercept. x_{it} contains the covariates with β a vector of coefficients corresponding to these covariates. γ is a time invariant group specific effect, with the groups being control and treatment. λ_t is a time varying effect which is the same in both groups. α is the effect of treatment which we are interested in. This can be seen as the average treatment effect on the treated across all post-treatment time periods. In Appendix C.2, a derivation is given which explains the intuition why α equals the ATT. Lastly ϵ_{it} is an iid error term.
Since panel data is available, the group specific effect can be replaced by a individual specific effect. This can be implemented by replacing γd_i by γ_i in equation (10). The main difference with equation (10) is that γ_i is now allowed to differ

across individuals instead of only across the treatment and control group. Using individual specific effects, time invariant covariates are absorbed by the individual fixed effect and thus no coefficients can be estimated for these covariates. The advantage of using individual fixed effects is that it can take into account unobserved confounders (Brüderl & Ludwig, 2015). The resulting regression estimator is also known as the Two-Way Fixed Effects (TWFE) estimator[17], since it incorporates two fixed effects: the individual fixed effect γ_i and the time fixed effect λ_t. TWFE estimations are implemented in R using the feols function from the fixest package (Gaure, 2013).

The regression formulation of DiD can be altered to calculate dynamics of the treatment effect when multiple time periods are available. Instead of being binary, t is now defined to take multiple values, corresponding to the time until treatment happens: $t = -3, -2, ..., 6$. Further, δ_t is a vector of dummy variables indicating the year, with one dummy for each possible value of t.
Then following Autor (2003) and Angrist and Pischke (2008) the formulation of equation (10) can be changed, resulting in equation (11). This formulation is known as the dynamic TWFE or event study regression (De Chaisemartin & D'Haultfoeuille, 2022):

$$y_{it} = \beta_0 + \beta x_{it} + \gamma_i + \lambda_t + \sum_{\tau=t_{start}}^{t_{end}} \alpha_\tau (d_i * \delta_\tau) + \epsilon_{it} \tag{11}$$

Here t_{start} equals -3 and t_{end} equals 5. $t = -1$ is used as baseline year and therefore no coefficient is estimated for that year. α_τ is the DiD estimate in different time periods. The estimates can be used for two main results. α_τ with $\tau < 0$ are the DiD estimates before treatment has happened. These estimates can be used to assess the plausibility of the parallel trends assumption, as well as anticipation. For these time periods the estimated coefficient should be zero. For the time periods after treatment, the differences in α_τ can be used to determine how the treatment effect evolves over time. In particular it can be used to assess whether the effect of a cancer diagnosis fades over time or is persistent.

4.4 Problems with TWFE

Last couple of years, many papers have been published on problems with TWFE estimation of DiD in settings with multiple time periods. These include Callaway and Sant'Anna (2021), Goodman-Bacon (2021), (Sun & Abraham, 2021) and De Chaisemartin and D'Haultfoeuille (2020).

[17] In literature both the regression formulation with group specific and individual specific effects are often referred to as TWFE estimation. In this thesis, TWFE will refer to the formulation containing individual specific effects.

De Chaisemartin and D'Haultfoeuille (2022) provide an elaborate overview of literature concerning the issue and the possible solutions. In their review, the main problems discussed occur when treatment timing is varying by groups. That is when e.g. one group is diagnosed in 2013 and another in 2014. In such cases TWFE estimation could yield very biased results, with possibly even the direction of effects switching. However in the application considered here, treatment timing does not vary and hence most problems considered in the aforementioned literature are not an issue for this application. In particular, De Chaisemartin and D'Haultfoeuille (2022) denote that the TWFE estimation is unbiased if treatment is staggered and binary, and all individuals receive treatment at the same time. The setting discussed in this thesis fulfills all these requirements. The TWFE regression may therefore present fairly good estimates. They state that it is not yet clear if TWFE estimators should be abandoned in general at this point. TWFE estimates often have a lower variance compared to more heterogeneity robust estimates. Comparing TWFE and other estimants can provide interesting insights in the reliability of TWFE estimation in practice.

While the treatment heterogeneity problems that are elaborately discussed in aforementioned literature are not an issue here, there are other issues that can even occur in the simplest two time period DiD design. Caetano, Callaway, Payne, and Rodrigues (2022) present four main problems that could result in poor estimation of the ATT using TWFE:

1. The presence of time varying covariates that are affected by treatment
2. The ATT or parallel trend assumption depends not only on changes in covariates over time but also on the level of time varying covariates before treatment
3. The ATT or the trend of Y_{it}^0 depends on time invariant covariates
4. Violations of strong functional form assumptions

In applications these problems often occur. It is likely that some of these problems occur in this application as well.
The second problem can be present in this application, since the effect of treatment may be different for people with a different education level, different number of children or marital status which are time varying covariates. The same holds for the third problem. It is likely that the trend differs for people born in different age cohorts as well as for people with a different gender. It is also likely that treatment is heterogeneous. Cancer could affect people in their twenties differently than people who are in their late thirties, while the effect of treatment can differ based on gender and other covariates as well[18].

[18]In Section 5 graphs are showed to highlight the presence of the third problem.

The first issue is not important for partner or marriage status. Using Kaplan-Meier estimation for the survival of relationships (see Figure 1), no significant differences were found in the breakup rates of partners in the control and treatment group.
It can expected that education is influenced by treatment, since a student diagnosed with cancer may have a different probability of dropping out compared to other students. However, since most people in the sample are relatively old AYAs, they are likely to have achieved the highest education they intended already. Hence this effect is probably small. Adding to that, S. Janssen et al. (2022) find that the impact of cancer differs by education level, but there is no clear difference in education achieved after treatment between the control and treatment group. This is the case for both one and five year after treatment as well as when comparing differences before and after treatment. This research is done using the same patient population and registries as in this thesis. Therefore it can be assumed that education is not significantly affected by treatment and hence that the first issue does not occur for education levels.
It is likely that the number of childbirths is affected by treatment, while this probably affects employment and hours worked as well. This possible issue is further explored in the robustness section. Caetano et al. (2022) advise using a Doubly Robust estimation method to solve the aforementioned problems when time varying covariates evolve exogeneously with respect to treatment. While most time varying covariates are unaffected by treatment, this may not be true for the number of children. This possible problem explored further in the robustness section.

Next a Doubly Robust estimation method by Sant'Anna and Zhao (2020) is discussed. They focus on a DiD setup where the parallel trends assumption only holds after conditioning on covariates. Similarly to problem 3 discussed above, they show that the TWFE generally is not an unbiased estimator for the ATT when either the treatment effects are heterogeneous in X or the trend differs based on covariates X. This covariate specific trend must be ruled out for both the control and treatment group. Their method resolves these problems. The proposed method is first explained in a simple two time period setting, which is thereafter expanded to allow for estimation in multiple time periods. It combines two types of estimators to make the estimation more robust to model misspecification[19]. Both of these estimators impose three main assumptions:

1. The data are *iid*
2. Conditional Parallel Trends Assumption; when only two time period are considered: $E(Y_{i1}^0 - Y_{i0}^0 | d_i = 1, X) = E(Y_{i1}^0 - Y_{i0}^0 | d_i = 0, X)$
3. Common support: $0 < P[d_i = 1|X] < 1$

[19]Please note that the explanation of the Doubly Robust DiD estimation is based on the paper by Sant'Anna and Zhao (2020) and formulations are largely taken directly from their paper.

The same assumptions are used for the DR estimation in Sant'Anna and Zhao (2020). The first and third assumptions hold. For the conditional parallel trends assumption, tests will be conducted in the results section similarly to the tests explained for the PT assumption in Section 4.2. The imposed conditional parallel trend assumption is a less strong assumption than the parallel trends assumption necessary for TWFE estimation.

The first method used in DR estimation is an Outcome Regression (OR), which is introduced in (Heckman, Ichimura, & Todd, 1997) and (Heckman, Ichimura, & Todd, 1998). This approach estimates the ATT as given in equation (12):

$$\hat{\alpha} = \bar{Y}_{1,1} - (\bar{Y}_{1,0} + n_{treat}^{-1} \sum_{i|d_i=1} (\hat{\mu}_{0,1}(X_i) - \hat{\mu}_{0,0}(X_i))) \tag{12}$$

Where $\bar{Y}_{d,t} = \sum_{i|d_i=r} y_{it}/n_{dt}$ is the average of all individuals belonging to group d at time t, where the possible groups are control or treatment. $\hat{\mu}_{dt}(X_i)$ is an estimator of $m_{dt}(X_i) = E(Y_{it}|d_i = d, X = X_i)$, which is equivalent to the expectation of the potential outcome for individual i conditional on X_i, when treatment status would be d. This is unobserved when $d = 0$ and individual i is treated. The OR consistently estimates the ATT when $\hat{\mu}_{rt}(X_i)$ is correctly specified, or more specifically when $\hat{\mu}_{0,1}(X) - \hat{\mu}_{0,0}(X) = m_{0,1}(X) - m_{0,0}(X)$.

The second method used is the Inverse Probability Weighting (IPW) approach proposed by Abadie (2005). This approach does not use a direct model for the outcome evolution. It uses a formulation of the ATT using the propensity score $p(X)$[20]. Using an estimator of the propensity score, the ATT can be estimated using equation (13):

$$\hat{\alpha} = \frac{1}{E_n(D)} E_n \left(\frac{D - \hat{\pi}(X)}{1 - \hat{\pi}(X)} (Y_1 - Y_0) \right) \tag{13}$$

$\hat{\pi}(X)$ is an estimator of the propensity score $p(X)$. E_n is the average of inserted variables, e.g. $E_n(D) = \frac{1}{n} \sum_{i=1}^{n} d_i$ with n the total number of observations.

Both approaches rely on different models to be estimated. For the OR method $\hat{\mu}_{dt}(X_i)$ needs to be correctly specified to consistently estimate the ATT. The same holds for $\hat{\pi}(X)$ in the IPW approach. Sant'Anna and Zhao (2020) propose a Doubly Robust (DR) estimator which combines both methods to create an estimator which is more robust to model misspecifications.
Let $\pi(X)$ be a model for the propensity score and $\mu_{dt}(X)$ a model for $m_dt(X)$. Then the Doubly Robust estimand is:

$$\alpha_{dr} = E \left[\left(\frac{D}{E_n(D)} - \frac{\frac{\pi(X)(1-D)}{1-\pi(X)}}{E_n(\frac{\pi(X)(1-D)}{1-\pi(X)})} \right) ((Y_1 - Y_0) - (\mu_{0,1}(X) - \mu_{0,0}(X)) \right] \tag{14}$$

[20]The propensity score is the probability of being treated given characteristics X: $p(X) = P(d_i = 1|X)$

Sant'Anna and Zhao (2020) show that equation (14) consistently estimates the ATT if either $\pi(X) = p(X)$ or $\mu_{0,1}(X) - \mu_{0,0}(X) = m_{0,1}(X) - m_{0,0}(X)$.
The estimation of the DR model is done in multiple steps. First $\mu_{rt}(X)$ and $\pi(X)$ are estimated using a parametric model. These estimates are then used jointly in equation (14) to estimate the ATT.

Callaway and Sant'Anna (2021) show how DR estimation can be extended to multiple time periods. Their method calculates $ATT(g,t)$, which is the average treatment effect at time t for group g. Group g indicates groups that are treated at a different time. In this application, only one group is considered since each patient is diagnosed in the same year. Therefore from now on, g will be left out from the notation used in their paper. Callaway and Sant'Anna (2021) add some assumptions to those made by Sant'Anna and Zhao (2020):

- Treatment is staggered: no unit is treated in the first time period and once a unit is treated, it will be treated in all following time periods.
- No anticipation

Staggered treatment holds in this application. The possibility of anticipation interfering with results is discussed in the robustness section.

The Doubly Robust estimator in different time periods t can be written equivalently to equation (14) as[21]:

$$ATT(t) = E\left[\left(\frac{D}{E_n(D)} - \frac{\frac{\pi(X)(1-D)}{1-\pi(X)}}{E_n(\frac{\pi(X)(1-D)}{1-\pi(X)})}\right)((Y_t - Y_{\mathcal{T}-1}) - (\mu_{0,t}(X) - \mu_{0,\mathcal{T}-1}(X)))\right] \tag{15}$$

Where $\mathcal{T}$ is the time of treatment. The main difference with equation (14) is that instead of only considering time periods 0 and 1, any time period t can be compared to the last period before treatment $(\mathcal{T}-1)$[22].

Having obtained the $ATT(t)$ for all time periods, the average effect of participating in treatment over all post-treatment periods is calculated by simply averaging the $ATT(t)$'s from after treatment periods as shown in equation (15):

$$ATT = \frac{1}{T_{end} - \mathcal{T} + 1}\sum_{t=\mathcal{T}}^{T_{end}} ATT(t) \tag{16}$$

Where T_{end} is the last observed time period.
Bootstrapped standard errors are calculated as advised in Callaway and Sant'Anna (2021).

[21]Note: X contains the covariate values in the base period, before treatment. The DR method explained in Callaway and Sant'Anna (2021) does not take into account how time-varying covariates change over time.

[22]Some minor notation differences are used compared to Callaway and Sant'Anna (2021), since only one group is considered and no anticipation is assumed. Adding to that, $\pi(X)$ and $(\mu_{0,t}(X) - \mu_{0,\mathcal{T}-1}(X))$ are used to keep consistent with previous notation.

4.5 Quantifying treatment effect heterogeneity

While the Doubly Robust estimator improves estimation under treatment heterogeneity, it does not provide insight into how differences in characteristics between individuals have impact on the treatment effect. These insights can be helpful, since they show which type of patients suffers the biggest impact from cancer on employment and the number of hours worked. Using this information, help with employment could be targeted more specifically to patients that are likely to need it the most.

The heterogeneity can be quantified using a regression. While previously some problems concerning the two-way fixed effects estimation are given, such regression type estimators have the benefit that the impact of different factors can be interpreted using the regression coefficients. To estimate the heterogeneity in the treatment effect due to covariates, interaction terms are added to the regression model introduced previously for TWFE using two time periods (equation (10)).

The inclusion of these interaction terms is similar to the Triple Difference (TD) estimation introduced by Gruber (1994) and explained in more detail by Olden and Møen (2020), but then including individual fixed effects[23]. The first additional term is an interaction between the covariate of interest (e.g. gender), the dummy for treatment group and the dummy denoting if the time of treatment has passed. Let the covariate of interest be denoted as $q_{it} \in x_{it}$. Then this interaction term can be written as $d_i * t * q_{it}$, with d_i and t as previously defined. This term captures the heterogeneity in the treatment effect for this covariate, which is the term of interest. The second interaction term added is the covariate of interest with the dummy denoting if time of treatment has passed: $t * q_{it}$. This term captures differences in the employment trend by the different groups in the covariate of interest. Lastly an interaction term combining the treatment group and covariate of interest is added as $d_i * q_{it}$. Using fixed effects, this term is incorporated into the individual specific when q_{it} is time invariant, which is the case for some of the considered covariates.
Including these interaction terms results in the formulation in equation (17), with ϕ the coefficient capturing the heterogeneity in the treatment effect for the considered covariate and α the base treatment effect[24].

$$y_{it} = \beta_0 + \beta x_{it} + \gamma_i + \lambda_t + \zeta(t * q_{it}) + \alpha(d_i * t) + \phi(d_i * t * q_{it}) + \epsilon_{it} \tag{17}$$

Next the treatment heterogeneity due to the type of cancer, stage at diagnosis and type of cancer treatment are considered. While previously treatment has always

[23] for the remainder of this thesis, the estimation method used here will be referred to as the triple differences estimator

[24] When e.g. gender is the covariate of interest, with a value of 0 for males and 1 for females, the baseline effect α is the effect on male individuals and ϕ the additional effect when an individual is female.

been considered to be a diagnosis with cancer for either the patient or partner of patient, these factors likely have a big impact on the magnitude of the effects. To quantify differences in the magnitude of effects based on these factors, estimation is done separately for different types of cancer, stages at diagnosis and types of cancer treatment. This is done by simply estimating the treatment effects for subsets of only patients belonging to a certain group or their partners, together with the controls that were previously matched to them.

5 Results

In this section the results are described. First the results are given for the patients. After that the same is given for the partners of patients. For both patients and partners, the trends for outcome variables over time are showed. These are also split by some individual characteristics, to highlight the need of DR estimation. After these graphical illustrations, the treatment effects estimated using the methodologies introduced in Section 4 are explained. All treatment effects considered are average treatment effects on the treated and ceteris paribus.

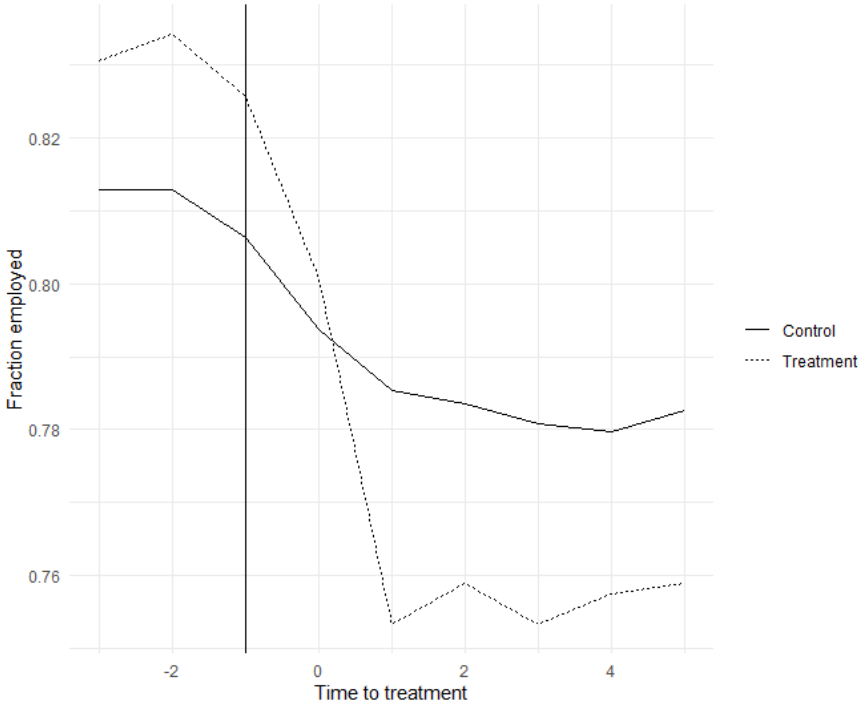
5.1 Results Patients

Results on patients will be discussed more concisely than those of partners. Some results are given in appendix D, to which will be referred throughout this section.

5.1.1 Trend employment and hours worked

In Figure 4, a trend plot can be found for the fraction of individuals that were employed over the observed years, split by treatment and control group for patients. Before treatment, which occurs at $t = 0$, the trends of both groups look parallel. There is a slight level difference between the two groups, which is significant at a five percent significance level when using a two-sample test for equality of proportions. No clear reason for this difference can be found, but since trends are parallel it is not a large issue for DiD estimation. After treatment, both groups clearly show a different trend. The employment of the treatment group declines compared to the control group, with the effect persistent over time. In the plot it is also visible that the employment of the control group decreases as well. This resembles the trend found by Eurostat (2022), which is explained in more detail in Appendix D.1 with Figure D.1.

Figure 4: Employment control and treatment group (patients)

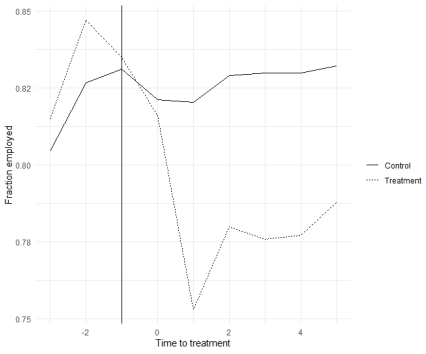


The employment trend is shown separated by age groups and gender in Figures 5a up to 6b. The age groups used are younger than 30 and 30 or older[25]. Trends of both age groups are different. The employment of controls increases for young controls, while it decreases in the control group with older people. The employment of young patients recovers slightly over the years while it stays stable for the older patients. Trends are different by gender as well. While employment for male patients recovers over years, the effect is persistent over time for females. Long term effects may thus be bigger for female subjects.

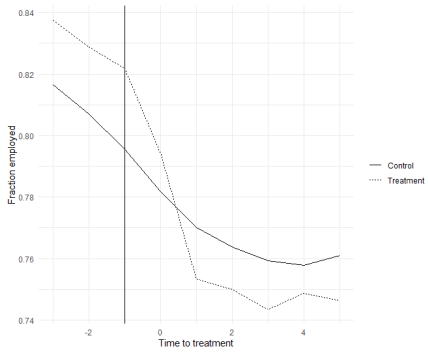
Trends in these figures thus differ by the characteristics of individuals. Further the treatment effect seems to differ based on these characteristics as well, which is most clear in the long term. Sant'Anna and Zhao (2020) showed that both different trends and different treatment effects by individual characteristics X can lead to biased estimation using TWFE. This illustrates why DR estimation should be used in this application.

In Figure D.2 in Appendix D.1, similarly trends of the log number of hours worked for those individuals that are employed are given. Again trends are similar before treatment, while the number of hours worked decreases for the treatment group compared to the control group after treatment. Thus again it is likely that there is a treatment effect of the cancer diagnosis on the number of hours worked.

[25]Age is determined by the age of individuals on the 31st of December 2013

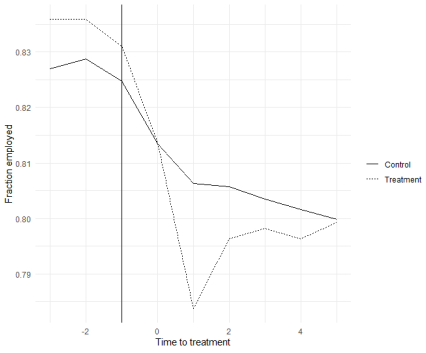


(a) Age <30 in 2013

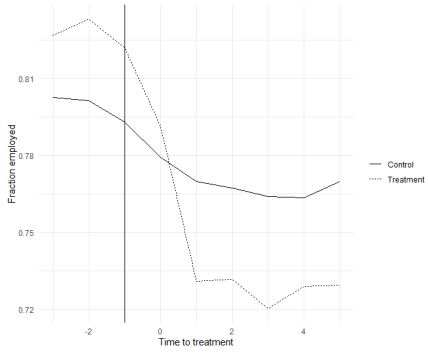


(b) Age $\geq$ 30 in 2013

Figure 5: Employment trend for different age cohorts



(a) Male



(b) Female

Figure 6: Employment trend for different genders

While it is shown that some violations occur due to which TWFE estimation may yield biased results, estimates by TWFE are still reported. These can be compared to DR estimates, to show how possible violations can affect TWFE estimation.

5.1.2 Overall treatment effects patients

In Table 9 the overall ATT's can be found for employment, log hours worked, self employment and having a permanent contract using both the TWFE and DR estimation[26]. Clearly treatment has a negative effect on employment using both models, with only small differences between the models. Using the TWFE regression, a cancer diagnosis is associated with a significant 3.9 percentage point decrease in the probability of employment. Using DR this is 3.8 percentage point. The decrease by approximately 3.8 percent point in employment may seem small. However when looking at the percentage of individuals that was not employed, which was 17.4 percent in the last period before treatment for the patients, this decrease in employment corresponds to a 21.8 percent increase in the probability of not being employed. While a decrease in employment is in line with other studies, the decrease in employment is smaller than that found by Moran et al. (2011), Heinesen et al. (2018) and Jeon (2017). They found decreases of 5-8 percentage point, but did not focus on AYA patients.
Given someone is employed, the number of hours worked on average decreases by 6.4 percent using TWFE and 3.8 percent using DR[27]. For both the effect is significant, but they clearly differ in magnitude. Hence the estimates produced using TWFE estimation are likely biased due to violations mentioned in Sant'Anna and Zhao (2020). The decreasing number of hours worked is in line with Moran et al. (2011).
A cancer diagnosis also decreases the probability of being self employed by 1.4 percentage point using TWFE and 1.3 using DR. Both estimates are significant. For those that are employed, a diagnosis with cancer seems to decrease the probability of having a permanent contract, though this effect is not significant.

[26]The overall effect is calculated using all post-treatment time periods. While treatment occurs in 2013, patients can be diagnosed in any moment of that year. Hence, the effect in 2013 is smaller than the real effect of treatment and the overall ATT over all years is probably underestimated slightly.

[27]For all coefficients on log hours worked, the percentage change is defined as $100*(exp(\hat{\beta})-1)$, with $\hat{\beta}$ the estimated coefficient. Hence the percentage may slightly deviate from the coefficient.

Table 9: Overall outcomes treatment effect of cancer diagnosis on patients

	Employment ($n = 2456$)		Log hours worked ($n = 1936$)	
	TWFE Estimate (Std. Error)	DR Estimate (Std. Error)	TWFE Estimate (Std. Error)	DR Estimate (Std. Error)
Post treatment	-0.031*** (0.002)		0.067*** (0.003)	
Education				
Secondary	0.178*** (0.014)		0.305*** (0.039)	
Higher	0.224*** (0.015)		0.626*** (0.040)	
Partner or Married	-0.020*** (0.003)		-0.051*** (0.006)	
Treatment effect	-0.039*** (0.006)	-0.038*** (0.007)	-0.066*** (0.011)	-0.039*** (0.014)

	Self employed ($n = 2456$)		Permanent contract ($n = 1936$)	
	TWFE Estimate (Std. Error)	DR Estimate (Std. Error)	TWFE Estimate (Std. Error)	DR Estimate (Std. Error)
Post treatment	0.025*** (0.001)		0.036*** (0.003)	
Education				
Secondary	-0.027*** (0.007)		0.011 (0.023)	
Higher	-0.025*** (0.008)		0.005 (0.024)	
Partner or Married	0.017*** (0.003)		0.057*** (0.005)	
Treatment effect	-0.014*** (0.004)	-0.013 (0.004)	-0.011 (0.009)	-0.010 (0.009)

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, education level, direct migration background, indirect migration background, registered partnership & marital status and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on, but does not calculate coefficients for them. Therefore the DR estimation only includes the estimated ATT.
Reported sample size n refers to the number of treated individuals considered. The number of controls is approximately 10 times the reported number of treated individuals.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

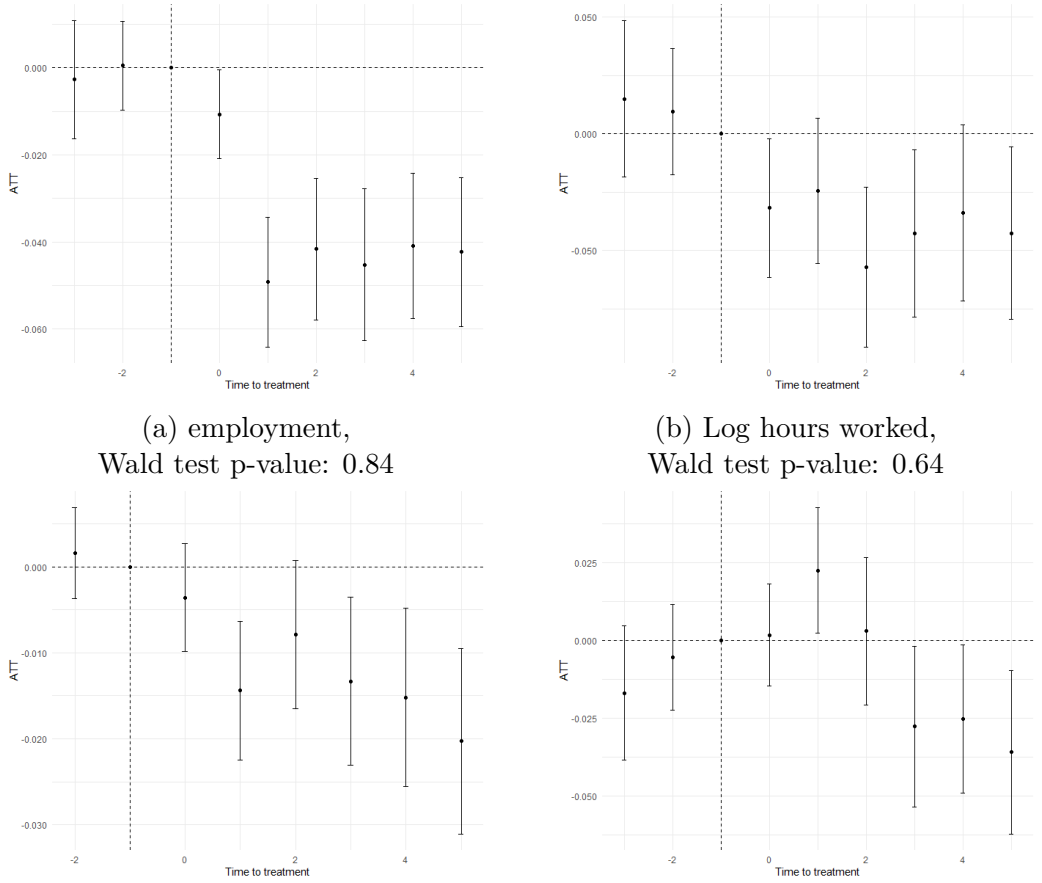
5.1.3 Dynamic effects

Figure 7 shows the dynamic effects of employment, log hours worked, self employment and the type of contract estimated using DR estimation. In Appendix Figure D.3 this can similarly be found using TWFE. For each pre-treatment period, the ATT estimated is insignificant. This strengthens the plausibility of parallel trends. This plausibility is further assessed using a Wald test on the estimated ATT's before treatment being jointly 0. The p-value of this Wald test has been included in the caption of the Figures in the appendix. For none of these tests, the pre period ATT's are found to be significantly different from 0 at any of the usual significance levels. While this does not guarantee the parallel trends assumption to hold, it is likely that the assumption holds.

Effects in 2013, which is time zero in the graphs, are much smaller than in other years. This is probably due to people receiving the diagnosis at an arbitrary moment in 2013, so it is possible for an individual to only be diagnosed in December 2013. This would of course mean that the measured effect in 2013 is very small. The difference between the effect in 2013 and later years is bigger for employment compared to hours worked. This can indicate that patients are likely to decrease working hours shortly after diagnosis, while the decision to completely stop working is made slightly later.

A cancer diagnosis increases the probability of having a permanent contract at first. This may be due to people on a temporary contract being more likely to lose their job, resulting in an increase of the proportion of people having a permanent contract. When time since treatment increases, the effect becomes negative. Thus in the long term surviving cancer decreases the probability of having a permanent contract.

For the other outcome variables results are similar in the second up to sixth post-treatment time period. Therefore the effects of a cancer diagnosis on employment, hours worked and self employment are all persistent over time.



(a) employment,
Wald test p-value: 0.84

(b) Log hours worked,
Wald test p-value: 0.64

(c) Self employment,
Wald test p-value: 0.58

(d) Contract type,
Wald test p-value: 0.32

Figure 7: Dynamic treatment effects patients with 95 percent confidence interval, using DR estimation

5.1.4 Heterogeneity of effects

In Appendix D.3, tables can be found considering the heterogeneity of the effects on both employment and hours worked by both individual characteristics in X and oncological data. Only the most notable findings will be discussed here. Heterogeneity of the treatment effect on employment and log hours worked with respect to individual characteristics that are not cancer related can be found in Table D.1 and D.2. The effects on both employment and hours worked when employed are significantly bigger for patients with a migration background. The effects are also bigger for female patients compared to male patients, while a higher education level reduces the effects found[28]. Patients with a higher education level probably have more options to find work that matches their new

[28]The heterogeneity of effects with respect to X are computed by interacting the treatment effect with these characteristics, as explained in detail in Section 4.5.

situation, while those with a lower education level usually have fewer job choices. The heterogeneity of the ATT by X again shows that DR estimation should be preferred over TWFE.

In Table D.3 treatment effects on employment are presented for patients with different oncological characteristics. As can be expected, a higher stage at diagnosis and a larger number of different types of cancer treatments increase the magnitude of effects found. The effects are smaller when organ or local surgery is used to treat a patient compared to chemotherapy, radiotherapy and hormonal therapy. CNS tumors reduce employment levels the most, which is similarly found by Teckle et al. (2018) and Gunnes et al. (2016). For these patients the employment probability decreased by 22.9 percentage point. Endocrine tumors barely affect employment. This may be due to stage at diagnosis being 1 for almost all patients with endocrine tumors (see Table 3 in the Data section)[29]. Further, breast cancer and hematologic cancer have a relatively big negative effect.
In Table D.4 similarly the treatment effect on the log number of hours worked for different oncological characteristics can be found. Hematologic and CNS tumors are associated with the biggest decrease in hours worked[30]. They are closely followed by endocrine tumors. Breast cancer has a significant negative effect on hours worked as well.

5.2 Results Partners

In this section the results obtained for partners of the patients are discussed[31]. First an explanation of trends found for partners is given, after which the effects found using methodologies from Section 4 are discussed. The plausibility of the parallel trends assumption will be tested when the dynamics of effects are considered.

5.2.1 Trend employment and hours worked

In Figures 8 and 9 the employment and log hours worked trends of the partners can be found, split by the control and treatment group. The employment trend before treatment is similarly decreasing for both groups. However, the biggest

[29] In general, the effects found for different types of cancer can partly be due to the stage at diagnosis of surviving patients

[30] While the estimated effect for CNS is large, it is not significant. This probably has to do with the small sample size available.

[31] Effects are only estimated for partners of patients that survived for at least 5 years after diagnosis. Therefore the effects found should only be considered to for partners of this type of patient. Fadlon and Nielsen (2015) found clear differences in effects of acute illness between partners of surviving individuals and those that did not survive the illness, showing the importance of this distinction. Further the added worker, caregiver and joint leisure effects likely depend regulations of the country people live in, so results found for partners in the Netherlands cannot simply be considered to hold in other countries.

decrease is found in year -3 for the treatment group, while this is in year -2 for the control group. The overall trend is still quite similar.
The log hours worked in the treatment group increases in year -3 while it decreases in year -2. For the control group the log hours worked decreases both periods. No clear explanation can be found for this difference. However considering the scale, differences are still very small and could have to do with randomness.
The parallel trends seem more questionable for partners than for the patients which were considered before. The plausibility of the parallel trends will be assessed further when the dynamics of the treatment effect are considered.
Interestingly employment slightly increases for the treatment group compared to the control group after treatment, while the log hours worked decreases.

Figure 8: employment control and treatment group partners

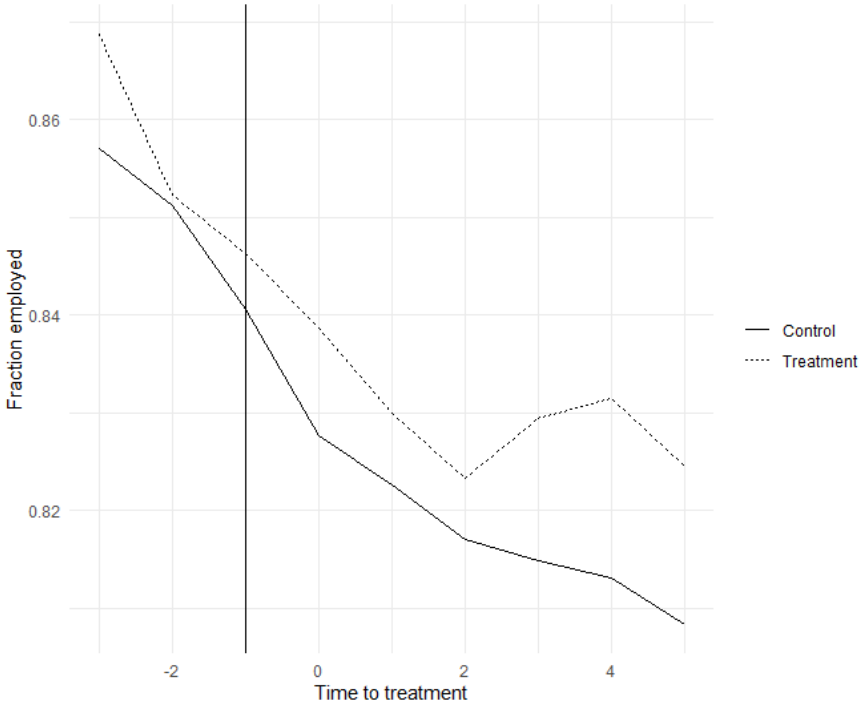
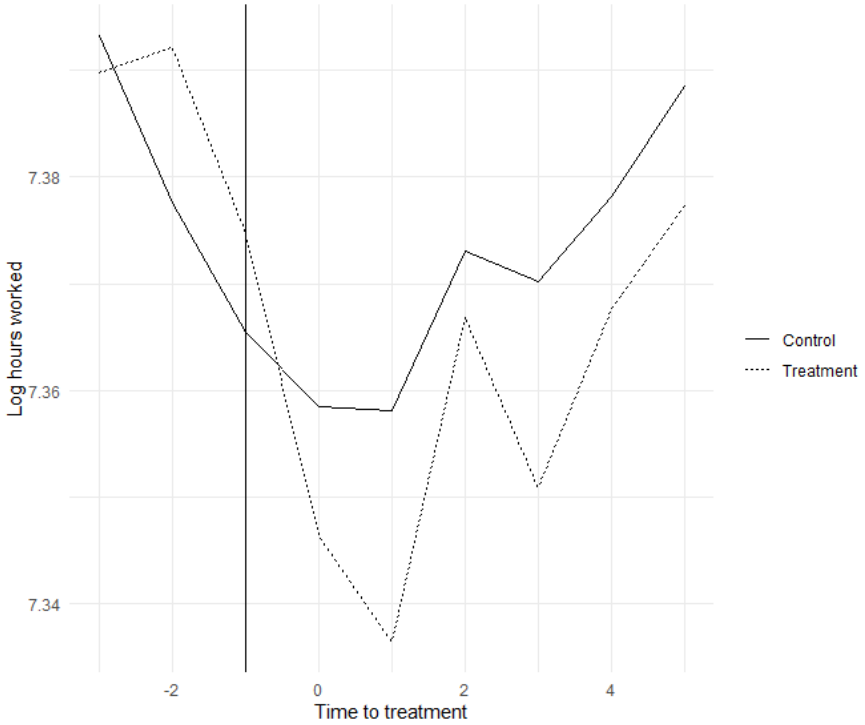
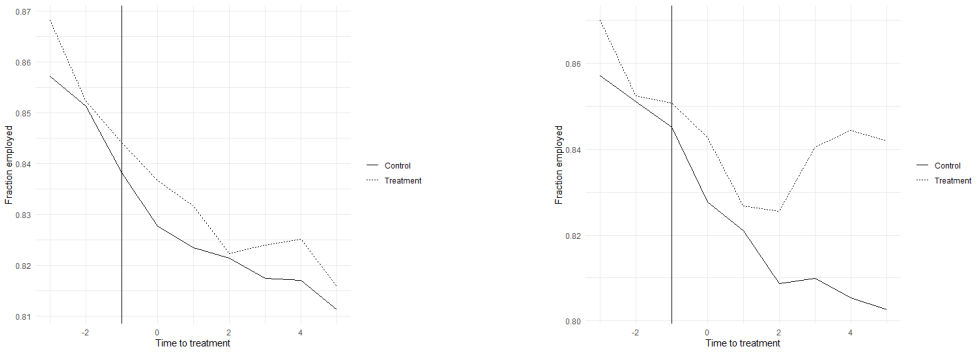


Figure 9: log hours worked control and treatment group partners



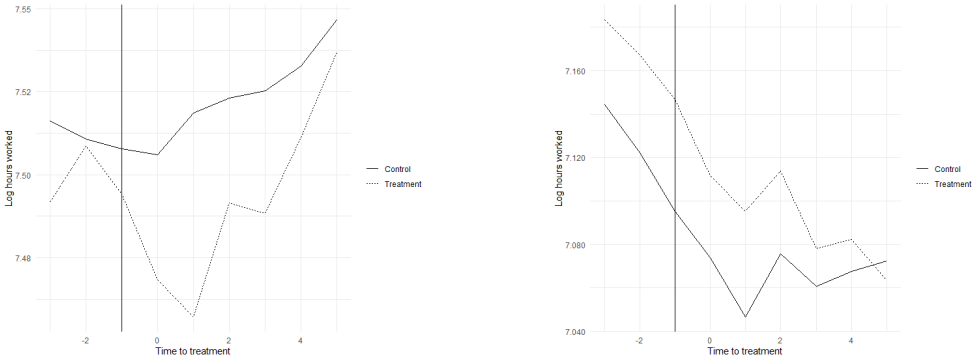
Next in Figure 10 the employment trend can be found split by gender[32]. The diagnosis increases employment for female partners compared to their controls, with the effect mostly clearly observed in the long term. For male partners no clear effect can be observed. In Figure 11 the same is given for the log number of hours worked. Again differences are visible between genders. For males the log number of hours worked first decreases after treatment, after which it recovers to be comparable to the control group 5 years after treatment. For females effects are more clear in the long term, with the log hours worked first following the same trend as controls, but then decreasing from the third period after treatment. For females the parallel trends assumption looks more plausible, since the trend before treatment is similar to the trend of controls.
Similarly to effects on the patients themselves, the treatment effect on partners differs based on individual characteristics. The trend in the log hours worked of the control group differs strongly by the gender of partners as well. This again shows that TWFE may find biased results, as shown by Sant'Anna and Zhao (2020). DR estimation is therefore the preferred estimation method. Though DR is preferred, TWFE estimates will be reported as well to compare with the DR estimates.

[32]For partners these plots are given for different age groups as well. This is not presented for partners because proportion of partners below 30 is very small.



(a) Male (b) Female

Figure 10: Employment trend for partners with different genders



(a) Male (b) Female

Figure 11: Log hours worked trend for partners with different genders

5.2.2 Overall Treatment Effect

In Table 10 the results of both TWFE and DR estimation for partners can be found[33]. A positive effect on employment is found using both estimation methods, with TWFE and DR finding an increased employment probability of 0.7 and 0.8 percentage point respectively. These effects are insignificant. This is different from studies by García-Gómez et al. (2013), Shen et al. (2019) and Jeon and Pohl (2017), which all find a negative effect of a health shock on spousal employment. The finding is similar to (Fadlon & Nielsen, 2015), who did not find a significant effect of a health shock on spousal employment when the patient survives.

[33]As previously mentioned TWFE estimation cannot estimate coefficients for covariates that are time invariant. These are excluded from the table. DR estimation uses covariates to condition on, but does not calculate coefficients for them. Therefore the DR estimation only includes the estimated ATT. Education level barely varies for partners. This can be explained by these individuals being relatively old, due to which they are less likely to obtain a degree in a new education level. This results in education being removed in the TWFE estimation, since variance over time is too small.

The effect is reversed on the number of hours worked. Both the TWFE and DR find a negative effect. The magnitude of the effect found strongly differs between estimation methods, indication that TWFE estimates are biased. TWFE finds a 0.6 percent decrease in hours worked, while this is 2.4 percent when DR is used. The effect found using DR is significant at a 90 percent level. Shen et al. (2019) similarly found a reduction in the number of hours worked by spouse following a health shock, though the effect found is bigger with a decrease of 4 hours per week[34].

Table 10: Overall outcomes treatment effect of cancer diagnosis on partners of AYA patients

	Employment ($n = 1252$)		Log hours worked ($n = 1050$)	
	TWFE	DR	TWFE	DR
	Estimate (Std. Error)	Estimate (Std. Error)	Estimate (Std. Error)	Estimate (Std. Error)
Post	-0.024*** (0.003)		-0.007 (0.005)	
married	-0.011* (0.006)		-0.017* (0.009)	
Number of children	-0.017*** (0.003)		-0.059*** (0.005)	
Treatment effect	0.007 (0.007)	0.008 (0.008)	-0.006 (0.011)	-0.024* (0.014)

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, number of children, duration relationship before 2013, direct migration background, indirect migration background, marital status, education level and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on, but does not calculate coefficients for them. Therefore the DR estimation only includes the estimated ATT.
Reported sample size n refers to the number of treated individuals considered. The number of controls is approximately 5 times the reported number of treated individuals.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The difference in the direction of the effect on employment and hours worked is interesting. It seems that people that were previously unemployed are more likely to become employed when their partner is diagnosed with cancer. This can be explained by the added worker effect. For those that were previously unemployed getting a job for a limited number of hours could compensate for loss of income, while keeping a substantial number of hours to take care of their ill partner and to spend time with them. Similarly partners of patients that only worked few hours before the diagnosis may be more likely to keep their job compared to those whose partners are not diagnosed.
The number of hours worked is decreased. This reversed direction compared to employment could be explained by people that do work a substantial number of hours shifting preferences to spend more time with their spouse and to take care of them. For these individuals the caregiver and joint leisure-effects outweighs the added worker effect. Further, the individuals that become employed may be likely to only work a small number of hours.

[34]Before treatment the subjects in this study worked an average of 50 hours per week. Hence the 4 hour reduction corresponds to a decrease by eight percent.

To test if indeed the employment effect is mostly due to people who did not work before treatment, or worked a small amount of hours, estimations of treatment effects are conducted for groups with a different number of hours worked in the last year before treatment. Similarly, treatment effects on the number of hours worked are estimated for these groups to test if the effect on the number of hours worked is mostly found for those partners that worked a large amount of hours prior to treatment.

The groups considered are partners who[35]:

- Worked less than 400 hours or unemployed
- Worked between 400 and 1600 hours
- Worked more than 1600 hours

In Table 11 the results for these subgroups can be found, estimated using DR estimation. Indeed for partners that were unemployed or worked a small number of hours in the last year before treatment, treatment does result in an increased employment probability. For this group having their partner diagnosed with cancer was associated with a significant 5.5 percentage point increase in employment. This can be explained by them still having a substantial amount of time to spend with, and care for, their partner even when starting to work some hours each week. For groups that worked more than 400 hours the year prior to treatment, no clear effect could be found.
An insignificant 7.0 percent decrease in the number of hours worked is found for the group that worked only a small number of hours prior to diagnosis. The number of observations used is very small as few people worked between 52 and 400 hours (only 39 partners of patients), which is probably the reason for the large standard error. No clear conclusion can be made due to this.
For people working between 400 and 1600 hours no clear effect can be found, with an estimated effect close to zero. For the group that worked at least 1600 hours in the last year prior to diagnosis a decrease of 1.3 percent is found, which is not significant. The difference between the second and third group is in line with the hypothesis that people that worked a large number of hours are more likely to reduce number of hours worked. However, the difference between groups is more clear for employment than for the number of hours worked.

[35] 400 hours is used as a threshold, since an individual working 400 hours in a year worked slightly less than 8 weekly hours on a weekly basis. Thus individuals working below 400 hours work less than a day on average, given that there are 52 weeks in a year. Similarly 1600 hours is used as the next threshold, since people working less than 1600 hours a year worked less than 4 full days each week. People working more than 1600 hours in a year work close to full time.

Table 11: treatment effects estimated with DR by hours worked before treatment

	Employment		Log hours worked	
	Estimate	Std. Error	Estimate	Std. Error
hours < 400	0.055**	0.028	-0.070	0.173
$400 \leq$ hours ≤ 1600	0.014	0.016	-0.005	0.026
hours > 1600	-0.002	0.006	-0.013	0.009

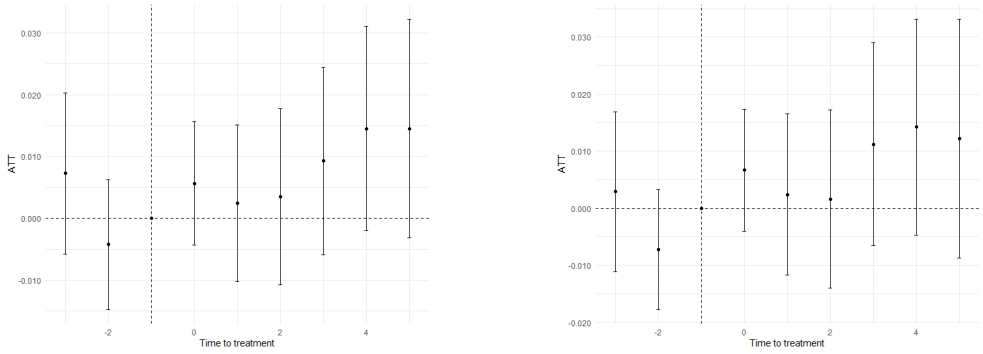
Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, number of children, duration relationship before 2013, direct migration background, indirect migration background, marital status, education level and gender.
Sample sizes employment: hours < 400: $n = 221$, $400 \leq$ hours ≤ 1600: $n = 281$, hours > 1600: $n = 750$. Sample sizes hours worked: hours < 400: $n = 39$, $400 \leq$ hours ≤ 1600: $n = 267$, hours > 1600: $n = 744$
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5.2.3 Dynamic Treatment Effect

In Figure 12 plots can be found with the dynamics of the treatment effect for partners of patients on employment. It shows the estimated ATT in different time periods, together with a 90 percent confidence interval[36].
No clear effects can be found before treatment. This is confirmed by a joint wald test on pre-treatment ATTs being zero. Using TWFE the p-value is 0.223, with it being 0.265 for DR estimation. Hence the parallel trends assumption is plausible, even though p-values are lower than those previously found for patients. More caution should be taken regarding the possibility of parallel trends not holding.
For both the TWFE and DR estimation the effect found is positive in each of the post-treatment periods, but never significant. The magnitude of the effect increases over the years. This could indicate that unemployed partners are more likely to start looking for a job when problems of their partners do not fade over time[37].
Considering the effect is positive and comparable in all post-treatment periods, it is likely that indeed the diagnosis with cancer increases partner employment slightly. The confidence interval is big which may be caused by the smaller sample size. Using a bigger sample size, effects could become more clear. That said, effects are still insignificant and therefore an increase in employment cannot be concluded with certainty.

[36]For partners a 90 percent confidence interval is used to show significance since this can more clearly depict how close estimates are to significance. For patients themselves a 95 percent confidence interval was used since those estimates were clearly significant. Thus when comparing effects on both groups, this difference should be kept in mind.

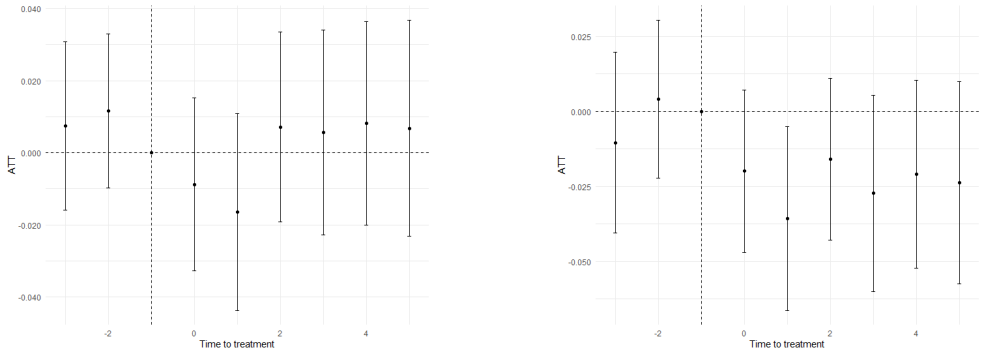
[37]The treatment effects on patients do not fade over time, as shown by estimating dynamic effects in the previous section.



(a) TWFE (b) DR

Figure 12: Dynamic effects employment partners with 90 percent confidence interval

In Figure 13 similarly the pattern of treatment effects can be found for the number of hours worked. Here TWFE and DR find different results. Using TWFE estimation treatment is associated with a decrease in the number of hours worked in the first two periods after treatment. This switches to an increase in the third time period after treatment and the effect stays positive in later time periods. Using DR estimation treatment is associated with a decrease in the number of hours worked in all time periods after treatment. As mentioned TWFE probably finds biased ATT's and thus DR estimates are preferred. Using DR estimation, the number of hours worked decreases for individuals whose partners are diagnosed with cancer, with the magnitude of the effect fairly stable over the years. Again effects are not significant in most years.
The parallel trends assumption is plausible considering no significant effect can be found before treatment. This is indicated by the Wald test as well, for which p-values of 0.673 and 0.652 are found using TWFE and DR respectively



(a) TWFE (b) DR

Figure 13: Dynamic treatment effects log hours worked partners with 90 percent confidence interval

5.2.4 Heterogeneity Treatment Effect

In this section the heterogeneity of treatment effects are considered. Tables for this section can be found in Appendix E[38]. Table E.1 contains the fixed effects triple differences estimates on employment for partners of patients. These give insight into the heterogeneity of effects by different covariates.
Female partners experience a bigger increase in employment compared to male partners. The baseline treatment effect for male partners is an insignificant 0.1 percentage point increase. An additional 1.9 percentage point increase is found for female partners, leading to a total increase of of 2.0 percentage point for females. Part of the difference between genders can be explained by female partners being more likely to work a small number of hours compared to male partners. Previously it was shown that the effect on employment depends on the number of hours worked before diagnosis. Both the baseline estimate and the added effect for females are not significant. To check whether a significant effect can be found when only considering female partners, TWFE and DR estimation have been implemented on a subset with only female partners and their matched controls. The estimated ATT is comparable to the 2.0 percentage point increase found by adding up the baseline and female specific effects, but are still insignificant at any widely used significance level.
Effects are mainly found for partners with a primary education level, with an estimated baseline employment increase of 8.6 percentage point. Having obtained a secondary or tertiary education degree leads to a decrease of the effect by 8.2 and 8.1 percentage point respectively. Adding this to the baseline effect of 8.6 percentage point results in only a 0.4-0.5 percentage point increase for these groups. The effect on individuals that only obtained primary education is still insignificant, which probably has to do with the small number of individuals in this category[39]. Hence this finding is not very reliable.
Migration background, number of children and being married do not have a considerable impact on the treatment effect[40]. The coefficients for interaction terms between these variables and the treatment effect are all small and insignificant.

Similarly in Table E.2 these results can be found for the log number of hours worked. Results are similar, but the coefficient for the interaction term with gender is far from significant as well. It does switch the sign however, with a positive effect found for females, while the effect is negative for males. The effect is positive for both primary and tertiary education, while being negative for those that achieved secondary education as the highest level. All these effects are insignificant.

[38]Tables are presented in the appendix since they take multiple pages.

[39]For only 19 partners of patients, the highest obtained education level was primary (see Table 5)

[40]All considered individuals are partners. Therefore it makes sense that being married does not have a big effects, especially considering that a registered partnership is equivalent to marriage in the Netherlands (Government of the Netherlands, 2022)

Next the heterogeneity of effects for different types of tumors and stages at diagnosis are discussed. Treatment effects are estimated by both TWFE and DR estimation, with separate estimations used for each subgroup. Results can be found in Tables E.3 and E.4. The number of partners of patients with the considered characteristics are reported in the tables as well. Group sizes are generally small, which is especially the case for endocrine, digestive system and CNS cancer, as well as stage four tumors. For these groups the estimated effects are not very reliable and can only give an indication of possible effects within the group. Results of both the effects on employment and the number of hours worked seem to vary by different tumor stages. The effect fluctuates, instead of it being consistently increasing or decreasing by the stage. This is unexpected since the stage is ordinal, so it would be more logical when effects would steadily increase or decrease by higher stages. Standard errors are generally big compared to estimated treatment effects, probably partly caused by small group sizes. No clear conclusion can be made on the effect of the stage on treatment effects.
The type of tumor influences effects on employment, see Table E.3. The treatment effect on employment is relatively large on partners of patients diagnosed with male genital tumors, but still insignificant. For partners of patients diagnosed with endocrine cancer the treatment effect on employment is significantly positive when using DR estimation, while the effect is smaller and insignificant when TWFE is used. On these partners a 8.1 percentage point increase in employment is found using DR, while this is only 1.7 when TWFE is used. The big difference can partly be explained by modelling differences with regards to the base year to which after treatment periods are compared[41]. The magnitude of the effect found for endocrine tumors is not very reliable given the group size and the effect differing highly by the base years used. Still it is likely that it has a positive effect on employment. Partners of patients with endocrine or male genital cancer were mostly female, which could account for a part of found effects.
The effect on employment is significantly negative for partners of patients suffering from breast cancer when DR is used, with it again being insignificant using TWFE. This likely has to do with the issues with TWFE as discussed by (Sant'Anna & Zhao, 2020) and thus DR is more reliable[42]. Having a partner diagnosed with breast cancer is associated with a 2.7 percentage point decrease in employment using DR estimation. This is different from effects found by Bradley and Dahman (2013). While they did find a negative effect on employment for hus-

[41]In the computation of the overall ATT using TWFE, time periods after treatment are compared to all pre-treatment periods. The overall ATT of DR is computed by averaging the ATT's found by comparing time periods after treatment with only the last pre-treatment period. This has a substantial impact for this specific estimated effect, while the impact on other estimates is generally negligible. In the robustness section this difference be explained further and clarifications will be given on where such problem could have impacted the differences between TWFE and DR estimates.

[42]For breast cancer, the difference in base year between TWFE and DR does not substantially affect found treatment effects.

bands of women diagnosed with breast cancer in the short term, this effect was insignificant when it was assessed nine months after the start of cancer treatment.

The effects on the number of hours worked in Table E.4 also differ by tumor type. Digestive system tumors are associated with a 14.2 percent decrease in the number of hours worked using DR and 8.7 percent decrease using TWFE. The difference between both is partly due to differences in the base year used, with more explanation on this in the robustness section. The estimated ATT for this type of cancer is significant, but only based on a small number of observations ($n = 50$). Female genital cancer is associated with a relatively large 7.2 percent decrease in the number of hours worked, but this effect is not significant. For breast cancer the estimated effect is negative, but insignificant. This is in line with Bradley and Dahman (2013), who found no significant effect of breast cancer on spousal number of hours worked in the long term.

As stated before, the number of observations used is small. This partly causes the insignificance of effects. The reliability of effects found for cancer types with a small number of observations is doubtful and they should be taken with caution. Treatment effects found can give an indication of possible effects, but for more reliable estimates on specific types of cancer, research with more observations is necessary. Corresponding to this it is important to note that an insignificance of effects does not necessarily mean the absence of an effect, since this can also be due to a high standard error caused by the small number of observations.

6 Robustness of estimates

Education has been imputed. As previously explained, it is not likely that this imputation results in biased outcomes. To check again whether it had any effects, estimations on employment and hours worked have been conducted while leaving out education as control variable. In Table F.1, the results of this can be found. When patients are considered, differences are very small. For the partners differences are almost non existing, which is logical since education varied barely for this group and using TWFE it was therefore even left out of estimation. The imputation of education thus does not have a considerable impact.

Further, the definition of employment may affect outcomes. In this research a definition of at least 52 hours per year is used[43]. To check for differences with different definitions of employed, estimations are also conducted while considering someone employed when only one hour has been worked in a year. This decreases the effects on employment slightly, while effects on hours worked increase. This could be expected. In the estimation where a minimum of 52 hours is used to be considered employed, people that almost stop working and decrease hours to e.g. 10 hours in a year are seen as unemployed. These people are not taken into account for the effect on hours worked. When a minimum of 1 hours is used, these same people will be considered employed. Their reduction in work hours is then captured in the analysis on the amount of hours worked.
Differences in the treatment effects are negligible when considering partners. Using one hour as the minimum did however decrease the p-value of pre trend tests. There may thus have been a difference in the trend of people working less than 52 hours between the treatment and control group. Considering the small number of hours worked in this group and the small group size, it is likely that this has to do with randomness.

It is also possible that patients are already ill before the diagnosis. There can be some time between the first symptoms and the diagnosis. Thus it is possible that the first effects are already found a year before diagnosis. This would violate the assumption of no anticipation. Since no significant treatment effects are found prior to diagnosis, it is unlikely that such a problem affects estimates. To check this again estimation is performed while accounting for possible anticipation, which can be done using DR estimation. The base year, to which post treatment outcomes are compared, is changed to one year earlier, while still only considering years after diagnosis for the computation of the overall ATT. Estimates are thus based on the differences between treated years and the time period two years before treatment, which is 2011 or year -2 in this case.
For patients the effect on employment does not change when using one year of anticipation. The effect on the number of hours worked is more negative when one year anticipation is used, with an estimated reduction in employment of 4.88

[43]See the data section for explanations on this choice

percentage point. This was 3.88 percentage point when no anticipation is used. It is possible that patients already reduce hours worked before diagnosis due to illness. The difference between both estimates is within the magnitude of the standard error. Anticipation thus does not have a big impact and differences are possibly just due to randomness.

Significant differences in child births between patients and their partners and the control group may impact estimates as well. To assess this difference analyses are conducted on the number of children living with a couple based on household data. The partner group has been used for this. In the appendix, Figure F.1 can be found which plots the number of children living with couples for both the treatment and control group. Before treatment the trend is parallel, but after treatment the average number of children increases faster for controls. Using DiD it is found that a cancer diagnosis leads to a decrease in the number of children living with a couple of 0.062. While the effect is small, it is significant even at a 99 percent significance level. Since childbirth can lead to decisions such as stopping with work or reducing hours worked, this difference could have an impact on the estimates. In particular, including the number of children in the estimation may result in biased estimation of the ATT (Zeldow & Hatfield, 2021). In the supplemental material to this paper Zeldow and Hatfield (2021) further elaborate onto the problems with a covariate which is affected by treatment and affects the outcome. They show that both adjusting and not adjusting for this type of covariate may yield biased results. Not adjusting for it mainly causes problems with the parallel trends, while adjusting for it can absorb a part of the treatment effects.
To find out in what extend the decision whether or not to include the number of children as a covariate affects outcomes, estimations have been performed excluding this covariate as well. Differences between estimates with and without the number of children are very small and insignificant. Even for genital tumors the differences are negligible, while these types of tumors have a big impact on fertility (Vakalopoulos, Dimou, Anagnostou, and Zeginiadou (2015) and Wilmoth and Spinelli (2000)).

For partners sometimes large differences between effects estimated using TWFE and DR are found. This can be due to modelling choices on how overall effects are constructed, which differ for TWFE and DR[44].
The overall effect estimated using TWFE is constructed by a regression comparing all periods before treatment with all periods after treatment, see equation (10). Using DR, the overall effect is estimated by averaging over all ATT's estimated for each post treatment time period using the last year before treatment as base year. This has two main implications that can influence estimates:

[44]For the estimated effects on which this applies, a reference has previously be made to this section. These are estimates on partner employment for endocrine and on partner hours worked for digestive system cancer

1. When estimated ATT's before treatment are different from 0, using all pre-treatment time periods (as in TWFE) or using only the last year before treatment as base year (as in DR) can give different estimates.

2. If the panel is unbalanced and treatment effects differ over time, differences in post treatment aggregation of effects can lead to different estimates[45].

The first implication accounts for the largest part of the differences between TWFE and DR estimated effects on employment for partners of patients with endocrine cancer[46]. While clearly insignificant, the estimated pre-treatment effects are positive in year -2 and -3. When these time periods are taken into account instead of only the base year, the estimated overall ATT becomes smaller. It is possible to calculate the TWFE estimate similarly to DR. For this the average of the post treatment ATTs, estimated using dynamic effects with TWFE, is calculated. By using this method, the overall effect estimated by TWFE increases strongly and is closer to that estimated using DR.
Thus for endocrine cancer, found effects highly depend on the type of base year(s) used in the calculation of effects. This implies that these results are not reliable, since these modelling choices should not have such an impact. The same is the case when effects on the number of hours worked for digestive system tumors on partners is calculated. However, the effect found for that tumor type is significant regardless of which base year is used.
For other estimated effects, differences due to the choice of base years are small and negligible. The problems only occur when small sample sizes are used, which can result in relatively large but still insignificant estimated pre-treatment effects. The larger but insignificant pre-treatment effect can still cause estimates to highly depend on the choice of the base year definition. Difference-in-differences methods can thus be very sensitive to the choice of the base year when small samples are used. In such instances it is important to check how differences in the base year affects estimation. Substantial differences in effects found can indicate that the estimation is unreliable.

[45]This difference can be most clearly explained by a simple example. Let there be two post treatment time periods, where half of the individuals in the sample in the first period have dropped out in the second. Let the ATT estimated in the first post treatment time period be 5 and 20 in the second post treatment time period. Then using the TWFE method described above, where simply every post treatment observation is considered separately in aggregation the overall ATT is $\frac{2*5+1*20}{3} = 10$. Using the DR method described above, the ATT's estimated in both time periods are given the same wight, independent of the number of individuals leading to an overall ATT of $\frac{5+20}{2} = 12.5$. Hence the aggregation method can significantly impact the overall effect when treatment effects and the number of individuals available both differ over time

[46]Treatment is estimated to increase partner employment by 1.7 percentage point using TWFE, while this is 8.1 percentage point using DR

7 Conclusion

This thesis focused on identifying the effects of a cancer diagnosis on employment for AYA patients and the spillover effect to partners. Mainly large effects were found on patients. For the partners, some effects could be found but these were often insignificant.

On patients, a cancer diagnosis is associated with a 3.8 percentage point reduction in employment probability, a 3.8 percent decrease in the number of hours worked when employed and a 1.3 percentage point decrease in self employment. In the long term, surviving cancer as an AYA reduces the probability of having a permanent contract when employed. While the magnitude of effects differ slightly, these findings are similar to previous studies.
The results differ strongly by differences between patients. Both having a migration background and being female are associated with a larger treatment effect, while having completed higher education reduces the effect. CNS tumors are associated with a very big treatment effect, while skin cancer does not affect work outcomes[47]. Both a higher stage at diagnosis and a higher number of different cancer treatments increases the magnitude of effects as would be expected. Lastly effects are smaller when local or organ surgery is used.
Chan et al. (2008) found that vocational rehabilitation services increases the employment of survivors, with 57 percent of unemployed survivors finding a job after receiving these services. This research was conducted on survivors aged 16 or older, with a mean of 40 years old. Considering the results found in this thesis, these types of programs could be more specifically targeted at those sub groups which are most in danger of becoming unemployed. This could not only positively impact the income of patients, but also their psychological well being (Peteet (2000), De Boer et al. (2008) and Blinder and Gany (2020)).

While clear effects could be found for patients, the spillover effects to partners is limited. Effects found for partners were generally small and in most cases insignificant. Having a partner which is diagnosed with cancer did increase employment by 0.8 percentage point, though this effect was not significant. The number of hours worked is decreased significantly by 2.4 percent. The minor effects that are found are persistent over time. When only partners are considered that were unemployed in the last year before diagnosis, or worked less than 400 hours, the employment probability is increased significantly by 5.5 percentage point. For them the added worker effect seems to outweigh the caregiver and joint leisure effects. For groups that worked more than 400 hours, no effect is found. Hence the effect on employment differs strongly based on employment characteristics before treatment. The found treatment effects differ from findings in most studies

[47]These are only the two biggest outliers. Other types of cancer had different effects as well. For more in depth information on these effects, refer to the results section and Appendix D

concerning health shocks[48]. This may indicate that partners of young people are affected differently than older partners, since the other studies did not specifically target partners of AYAs

The effect on partner employment depends on gender. For male partners no effects is found. For females an increased employment of 2.0 percentage point is found, though the effect is still insignificant. Other covariates did not have a clear impact on the treatment effect, which may sometimes be due to few observations within categories.
While effects seem different for different stages of tumors, no clear conclusion can be made on how stages affect the treatment effect. The type of tumor affects the treatment effect. These differences may partly be due to gender differences of partners for certain types of tumors. Endocrine and male genital tumors are associated with a relatively large positive treatment effect on the employment of partners, while the effect was negative for partners of patients diagnosed with breast cancer. Further, digestive system tumors are associated with the biggest decrease in the number of hours worked by partners of patients. For digestive system and endocrine cancer sample sizes are small, due to which the reported effects for them may be unreliable.

Altogether effects on partners are small, though employment does seem to slightly increase while the number of hours worked when employed decreases. These effects depend on the work situation prior to diagnosis.
Help with employment related issues, such as vocational rehabilitation services, should therefore primarily target the patients themselves. The diagnosis has a bigger impact on their work than on the work of partners.

[48]These differences are discussed in more detail in the results section

8 Discussion

In this section some possible problems with the validity of findings are discussed. Further, some areas are given in which future research could add on findings in this thesis.

The main assumption that is needed for valid estimates of the treatment effect on the treated is that parallel trends hold. Tests indicate that this assumption is likely to hold. However there are only few time periods prior to treatment, which means the trend cannot be assessed over a longer time period. This would give more insights in the plausibility of the assumption. For partners of patients the trends visually did not look very similar. While tests indicate that parallel trends are plausible, the findings should be taken with some more caution.
Even though the parallel trends cannot be validated with certainty, I believe the results are reasonable. The construction of the control group guarantees that patients and their partners are compared to similar people.

One threat that must always be considered when analysing treatment effects, which could invalidate parallel trends, is a selection into treatment bias. As previously mentioned, cancer is thought to be mostly random at a young age and does not depend on lifestyle choices (American Cancer Society, 2022). The probability of contracting a tumor does increase based on age, but this is taken account for by the matching procedure. Hence it is unlikely that this type selection bias is a problem.

It is always possible that there are unobserved confounders which are not controlled for. These can e.g. be time invariant variables that have a time varying effect on the outcome, while there is a level difference of this covariate between the treatment and control group (Zeldow & Hatfield, 2021). As demographic covariates are used to match upon and controlled for in estimation, it is unlikely that this plays an important role but the possibility still exists. Other possible confounders are time varying variables that either develop differently in the treatment or control group, or have a time varying effect while there is a level difference for this covariate between both groups (Zeldow & Hatfield, 2021). Multiple time varying covariates have been considered, with most of these not developing differently between the two groups. These are not confounders by the aforementioned definition. Only the number of children is likely to be a confounder, though it can be seen as a mediator[49]. In the robustness section this possible confounder has been investigated further. It does not have a substantial impact. Still the existence of other time varying confounders cannot be completely ruled out.

Next some possible research areas are given which could add upon findings from this thesis. Data has been used of cancer patients that survived for at least 5

[49] A time varying confounder in which differences are caused by treatment itself are referred to as mediators, which are variables through which treatment indirectly affects the outcome

years after diagnosis. Since the survival probability is directly associated with the severity of a tumor, this means that patients with more severe types of cancer are underrepresented in this study. The effects found on partners are also specifically for partners of this subgroup of patients. Fadlon and Nielsen (2021) previously focused on the spousal labor effects of both fatal and non fatal health shocks. They found no effect for spouse of surviving patients, while large effects were found for partners of deceased patients. Hence it is likely that results differ strongly between both groups when partners of AYA cancer patients are considered. Investigation into spousal labor effects when patients do not survive 5 years could give important insights into the spillover effects of a cancer diagnosis for this different group.

The different effects that probably play an important role in the spillover effects into the labor of patients are likely to differ substantially between countries. The added worker effect probably depends on the social security laws within a country, while the caregiver effect could depend on the accessibility and affordability of healthcare in a country. Therefore these results found using Dutch data do not necessarily give a good representation of effects in countries with other social security and healthcare regulations. Similar research in different countries could add on the knowledge of spousal labor effects and their dependency on countries.

When more specific sub groups were considered, such as only partners of patients diagnosed with a certain type of cancer, the group size was often very small. This can partly explain why these results are insignificant, while reliability of treatment effects found is smaller due to the small sample size as well. Research using larger groups likely finds more reliable estimates.

Lastly partners of patients may face issues in different areas than employment, while employment related issues themselves can go beyond those measured in objective data on employment and hours worked. Takeuchi, Ichikura, Amano, Takeshita, and Hisamura (2018) find that partners of AYA patients face comparable social difficulties to patients in many categories. Amano et al. (2019) similarly find psychological distress and a lower quality of life for the spouse of cancer patients, which was especially notable for young spouses. Further research into AYA spousal effects is important, especially considering social issues they could face. This research could be conducted in a more qualitative manner, e.g. by conducting interviews to identify the biggest issues faced.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1–19.

Amano, K., Ichikura, K., Hisamura, K., Motomatsu, Y., Matsushima, E., et al. (2019). Factors associated with social support needs of spouses of patients with cancer: Online survey. *International Journal of Clinical Medicine*, *10*(04), 270.

American Cancer Society. (2022). *Risk factors and causes of cancers in young adults.* Retrieved 2022-07-20, from `https://www.cancer.org/cancer/cancer-in-young-adults/risk-factors-and-causes.html`

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics*. Princeton university press.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics*, *21*(1), 1–42.

Beretta, L., & Santaniello, A. (2016). Nearest neighbor imputation algorithms: a critical evaluation. *BMC medical informatics and decision making*, *16*(3), 197–208.

Bland, J. M., & Altman, D. G. (2004). The logrank test. *Bmj*, *328*(7447), 1073.

Blinder, V. S., & Gany, F. M. (2020). Impact of cancer on employment. *Journal of Clinical Oncology*, *38*(4), 302.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In *Handbook of econometrics* (Vol. 5, pp. 3705–3843). Elsevier.

Bradley, C. J., Bednarek, H. L., & Neumark, D. (2002). Breast cancer survival, work, and earnings. *Journal of health economics*, *21*(5), 757–779.

Bradley, C. J., & Dahman, B. (2013). Time away from work: employed husbands of women treated for breast cancer. *Journal of Cancer Survivorship*, *7*(2), 227–236.

Bradley, C. J., Neumark, D., Bednarek, H. L., & Schenk, M. (2005). Short-term effects of breast cancer on labor market attachment: results from a longitudinal study. *Journal of Health Economics*, *24*(1), 137–160.

Brüderl, J., & Ludwig, V. (2015). Fixed-effects panel regression. *The Sage handbook of regression analysis and causal inference*, *327*, 357.

Caetano, C., Callaway, B., Payne, S., & Rodrigues, H. S. (2022). Difference in differences with time-varying covariates. *arXiv preprint arXiv:2202.02903*.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Carlsen, K., Dalton, S. O., Frederiksen, K., Diderichsen, F., & Johansen, C. (2007). Are cancer survivors at an increased risk for divorce? a danish cohort study. *European Journal of Cancer*, *43*(14), 2093–2099.

Chan, F., Strauser, D., da Silva Cardoso, E., Xi Zheng, L., Chan, J. Y., &

Feuerstein, M. (2008). State vocational services and employment in cancer survivors. *Journal of Cancer Survivorship*, *2*(3), 169–178.

De Boer, A., Verbeek, J., Spelten, E., Uitterhoeve, A., Ansink, A., De Reijke, T., ... Van Dijk, F. (2008). Work ability and return-to-work in cancer patients. *British journal of cancer*, *98*(8), 1342–1347.

De Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–96.

De Chaisemartin, C., & D'Haultfoeuille, X. (2022). *Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey* (Tech. Rep.). National Bureau of Economic Research.

De Rooij, B. H., Ezendam, N. P., Mols, F., Vissers, P. A., Thong, M. S., Vlooswijk, C. C., ... van de Poll-Franse, L. V. (2018). Cancer survivors not participating in observational patient-reported outcome studies have a lower survival compared to participants: the population-based profiles registry. *Quality of Life Research*, *27*(12), 3313–3324.

DeSantis, C. E., Lin, C. C., Mariotto, A. B., Siegel, R. L., Stein, K. D., Kramer, J. L., ... Jemal, A. (2014). Cancer treatment and survivorship statistics, 2014. *CA: a cancer journal for clinicians*, *64*(4), 252–271.

Eurostat. (2022). *Employment rates by sex, age and country of birth (%).* Retrieved 2022-06-17, from `appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do`

Fadlon, I., & Nielsen, T. H. (2015). *Family labor supply responses to severe health shocks* (Tech. Rep.). National Bureau of Economic Research.

Fadlon, I., & Nielsen, T. H. (2021). Family labor supply responses to severe health shocks: Evidence from danish administrative records. *American Economic Journal: Applied Economics*, *13*(3), 1–30.

García-Gómez, P., Van Kippersluis, H., O'Donnell, O., & Van Doorslaer, E. (2013). Long-term and spillover effects of health shocks on employment and income. *Journal of Human Resources*, *48*(4), 873–909.

Gaure, S. (2013). Ols with multiple high dimensional category variables. *Computational Statistics & Data Analysis*, *66*, 8–18.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Government of the Netherlands. (2022). *Marriage, registered partnership and cohabitation agreements.* Retrieved 2022-07-20, from `https://www.government.nl/topics/marriage-cohabitation-agreement-registered-partnership/marriage-registered-partnership-and-cohabitation-agreements`

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American economic review*, 622–641.

Gruber, J., & Cullen, J. B. (1996). *Spousal labor supply as insurance: Does unemployment insurance crowd outthe added worker effect?* National Bureau

of Economic Research Cambridge, Mass., USA.

Gunnes, M. W., Lie, R. T., Bjørge, T., Syse, A., Ruud, E., Wesenberg, F., & Moster, D. (2016). Economic independence in survivors of cancer diagnosed at a young age: a norwegian national cohort study. *Cancer*, *122*(24), 3873–3882.

Hartog, J., & Salverda, W. (2018). The labor market in the netherlands, 2001-2016. *IZA World of Labor*(418).

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The review of economic studies*, *65*(2), 261–294.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605–654.

Heinesen, E., Imai, S., & Maruyama, S. (2018). Employment, job skills and occupational mobility of cancer survivors. *Journal of health economics*, *58*, 151–175.

Heinesen, E., & Kolodziejczyk, C. (2013). Effects of breast and colorectal cancer on labour market outcomes—average effects and educational gradients. *Journal of health economics*, *32*(6), 1028–1042.

Husson, O., de Rooij, B. H., Kieffer, J., Oerlemans, S., Mols, F., Aaronson, N. K., ... van de Poll-Franse, L. V. (2020). The eortc qlq-c30 summary score as prognostic factor for survival of patients with cancer in the "real-world": Results from the population-based profiles registry. *The Oncologist*, *25*(4), e722–e732.

Jadhav, A., Pramod, D., & Ramanathan, K. (2019). Comparison of performance of data imputation methods for numeric dataset. *Applied Artificial Intelligence*, *33*(10), 913–933.

Janssen, S., van der Meer, D., van Eenbergen, M., Manten-Horst, E., van der Graaf, W., & Husson, O. (2022). Short-and long-term impact of cancer on employment and financial outcomes of adolescents and young adults (ayas): a large population-based case-control registry study in the netherlands. *ESMO open*, *7*(4), 100521.

Janssen, S. H., van der Graaf, W. T., van der Meer, D. J., Manten-Horst, E., & Husson, O. (2021). Adolescent and young adult (aya) cancer survivorship practices: An overview. *Cancers*, *13*(19), 4847.

Jeon, S.-H. (2017). The long-term effects of cancer on employment and earnings. *Health economics*, *26*(5), 671–684.

Jeon, S. H., & Pohl, R. V. (2017). Health and work in the family: Evidence from spouses' cancer diagnoses. *Journal of health economics*, *52*, 1–18.

Kahn-Lang, A., & Lang, K. (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, *38*(3), 613–620.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53*(282), 457–

481.

Kowarik, A., & Templ, M. (2016). Imputation with the r package vim. *Journal of Statistical Software*, *74*, 1–16.

Kreiner, C. T., Lassen, D. D., & Leth-Petersen, S. (2013). *Measuring the accuracy of survey responses using administrative register data: Evidence from denmark* (Tech. Rep.). National Bureau of Economic Research.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methodsestimation of spatial panels. *Foundations and Trends in Econometrics*, *4*, 165-224.

Leuteritz, K., Friedrich, M., Sender, A., Richter, D., Mehnert-Theuerkauf, A., Sauter, S., & Geue, K. (2021). Return to work and employment situation of young adult cancer survivors: results from the adolescent and young adult-leipzig study. *Journal of Adolescent and Young Adult Oncology*, *10*(2), 226–233.

Lundberg, S. (1985). The added worker effect. *Journal of Labor Economics*, *3*(1, Part 1), 11–37.

Mahar, K. K., BrintzenhofeSzoc, K., & Shields, J. J. (2008). The impact of changes in employment status on psychosocial well-being: a study of breast cancer survivors. *Journal of psychosocial oncology*, *26*(3), 1–17.

McHugh, M. L. (2013). The chi-square test of independence. *Biochemia medica*, *23*(2), 143–149.

Mehnert, A. (2011). Employment and work-related issues in cancer survivors. *Critical reviews in oncology/hematology*, *77*(2), 109–130.

Moons, K. G., Donders, R. A., Stijnen, T., & Harrell, F. E. (2006). Using the outcome for imputation of missing predictor values was preferred. *Journal of clinical epidemiology*, *59*(10), 1092–1101.

Moran, J. R., Short, P. F., & Hollenbeak, C. S. (2011). Long-term employment effects of surviving cancer. *Journal of health economics*, *30*(3), 505–514.

Nahum, R. A. (2007). *Labour supply response to spousal sickness absence*. Institutet för framtidsstudier.

Olden, A., & Møen, J. (2020). The triple difference estimator. *NHH Dept. of Business and Management Science Discussion Paper*(2020/1).

Parsons, H. M., Harlan, L. C., Lynch, C. F., Hamilton, A. S., Wu, X.-C., Kato, I., ... Keegan, T. H. (2012). Impact of cancer on work and education among adolescent and young adult cancer survivors. *Journal of clinical oncology*, *30*(19), 2393.

Peteet, J. R. (2000). Cancer and the meaning of work. *General hospital psychiatry*, *22*(3), 200–205.

Pischke, J. S. (2005). Empirical methods in applied economics: Lecture notes. *Downloaded July*, *24*, 2015.

Quinn, G. P., Gonçalves, V., Sehovic, I., Bowman, M. L., & Reed, D. R. (2015). Quality of life in adolescent and young adult cancer patients: a systematic review of the literature. *Patient related outcome measures*, *6*, 19.

Ribeiro, C., & Freitas, A. A. (2021). A data-driven missing value imputation approach for longitudinal datasets. *Artificial Intelligence Review*, *54*(8), 6277–6307.

Riekhoff, A.-J., & Vaalavuo, M. (2021). Health shocks and couples' labor market participation: A turning point or stuck in the trajectory? *Social Science & Medicine*, *276*, 113843.

Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122.

Seitz, D. C., Besier, T., Debatin, K.-M., Grabow, D., Dieluweit, U., Hinz, A., ... Goldbeck, L. (2010). Posttraumatic stress, depression and anxiety among adult long-term survivors of cancer in adolescence. *European Journal of cancer*, *46*(9), 1596–1606.

Shen, Z., Zheng, X., & Tan, Y. (2019). The spillover effects of spousal chronic diseases on married couples' labour supply: evidence from china. *International Journal of Environmental Research and Public Health*, *16*(21), 4214.

Spelten, E. R., Sprangers, M. A., & Verbeek, J. H. (2002). Factors reported to influence the return to work of cancer survivors: a literature review. *Psycho-Oncology: Journal of the Psychological, Social and Behavioral Dimensions of Cancer*, *11*(2), 124–131.

Statistics Netherlands. (2022). *Bevolking; onderwijsniveau en migratieachtergrond 2003-2021.* Retrieved 2022-08-12, from `opendata.cbs.nl`

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Syse, A., & Kravdal, Ø. (2007). Does cancer affect the divorce rate? *Demographic Research*, *16*, 469–492.

Syse, A., Tretli, S., & Kravdal, Ø. (2008). Cancer's impact on employment and earnings—a population-based study from norway. *Journal of Cancer Survivorship*, *2*(3), 149–158.

Tai, E., Buchanan, N., Townsend, J., Fairley, T., Moore, A., & Richardson, L. C. (2012). Health status of adolescent and young adult cancer survivors. *Cancer*, *118*(19), 4884–4891.

Takeuchi, T., Ichikura, K., Amano, K., Takeshita, W., & Hisamura, K. (2018). The degree of social difficulties experienced by cancer patients and their spouses. *BMC palliative care*, *17*(1), 1–9.

Teckle, P., Peacock, S., McBride, M. L., Bentley, C., Goddard, K., & Rogers, P. (2018). Long-term effects of cancer on earnings of childhood, adolescent and young adult cancer survivors–a population-based study from british columbia, canada. *BMC health services research*, *18*(1), 1–10.

Telles, C. M. (2021). A scoping review of literature: What has been studied about adolescents and young adults (ayas) with cancer? *Cancer Treatment and Research Communications*, 100316.

Theloudis, A. (2021, September). *Lecture notes in microeconometrics: Lecture 7:*

Difference in differences estimation. Tilburg University.

Timperi, A. W., Ergas, I. J., Rehkopf, D. H., Roh, J. M., Kwan, M. L., & Kushi, L. H. (2013). Employment status and quality of life in recently diagnosed breast cancer survivors. *Psycho-Oncology*, *22*(6), 1411–1420.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., ... Altman, R. B. (2001). Missing value estimation methods for dna microarrays. *Bioinformatics*, *17*(6), 520–525.

Vaalavuo, M. (2021). The unequal impact of ill health: Earnings, employment, and mental health among breast cancer survivors in finland. *Labour Economics*, *69*, 101967.

Vakalopoulos, I., Dimou, P., Anagnostou, I., & Zeginiadou, T. (2015). Impact of cancer and cancer treatment on male fertility. *Hormones*, *14*(4), 579–589.

Warner, E. L., Kent, E. E., Trevino, K. M., Parsons, H. M., Zebrack, B. J., & Kirchhoff, A. C. (2016). Social well-being among adolescents and young adults with cancer: a systematic review. *Cancer*, *122*(7), 1029–1037.

Wilmoth, M. C., & Spinelli, A. (2000). Sexual implications of gynecologic cancer treatments. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, *29*(4), 413–421.

Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), 1–17. doi: 10.18637/jss.v077.i01

Zajacova, A., Dowd, J. B., Schoeni, R. F., & Wallace, R. B. (2015). Employment and income losses among cancer survivors: estimates from a national longitudinal survey of american families. *Cancer*, *121*(24), 4425–4432.

Zeldow, B., & Hatfield, L. A. (2021). Confounding and regression adjustment in difference-in-differences studies. *Health services research*, *56*(5), 932–941.

Appendices

A Literature overview effect cancer diagnosis on survivors

As stated in the introduction, a cancer diagnosis has a substantial impact on employment of a patient. While many researches do not focus on AYA patients, do not use population data or have no control group, previous literature may give a good indication of effects that can be found. Below an overview is given of previous findings, such as stated in the introduction.

Leuteritz et al. (2021) and Parsons et al. (2012) did look into employment related issues. They both find that most survivors do return to work. The percentage that returns to work depends on the type of cancer as well as treatment intensity, insurance and quitting the current job/school directly after diagnosis. Further, Leuteritz et al. (2021) find that survivors often work reduced hours.
Both papers are based on survey data, which could make data less reliable (Bound, Brown, & Mathiowetz, 2001). This reliability issue can be shown by an example from research in Denmark. In a comparison between survey and register data on gross income in Denmark, the mean of the survey data was biased while the data was noisy as well (Kreiner, Lassen, & Leth-Petersen, 2013). On top of this neither of the papers uses a control group. Hence no reliable comparison with the whole population can be made. In a literature overview, Warner et al. (2016) concluded that new research into the social well-being of AYAs with cancer should among others be focused on employment reintegration and should be conducted using comparison groups.

Teckle et al. (2018) and Gunnes et al. (2016) did use population data to investigate the effects of surviving cancer at a young age on income and employment. Teckle et al. (2018) finds that survivors earned significantly less. The magnitude of this effect differs based on the type of therapy and cancer. The effect is the biggest for survivors that only received radiotherapy or were diagnosed with central nervous system (CNS) and soft tissue sarcoma cancer. Gunnes et al. (2016) focus on the effect on employment and similarly find a negative effect of surviving cancer at a young age. Again the effect is the biggest for survivors of CNS cancer, while the effect is bigger for female survivors as well. Note that both papers only focus on effects for survivors diagnosed before age 25. Similar effects are found for adults (Syse et al. (2008), Zajacova et al. (2015) and Vaalavuo (2021)). Further, Mehnert (2011) shows that flexible working arrangements, counseling, training and rehabilitation services, higher levels of education, male gender, less physical symptoms, shorter sick leave and continuity of care increase the the probability of returning to work for survivors. None of these researches focuses specifically on AYAs.
While most mentioned papers are from medical journals, similar effects are found in papers published in journals focused on health economics. Moran et al. (2011) find a reduction of employment by 7-8 percentage point for cancer survivors.

Cancer survivors are also found to reduce hours, with a reduction of 3.5 hours per week for females and 5.5 hours per week for males. the effect considering the number of hours worked included those survivors that stopped working. Heinesen et al. (2018) similarly finds a 7 percentage point decrease in employment, with the effect being bigger when survivors are lower educated. Jeon (2017) finds a alightly smaller negative effect of 5 percentage point reduction in employment. Heinesen and Kolodziejczyk (2013) also finds negative effects for both breast and colorectal cancer. Bradley et al. (2002) find that breast cancer patients are 10 percentage point less likely to be employed in the first 2 years after diagnosis, with this dropping to 6 percent from 3 years after diagnosis. Interestingly those patients that are employed work 1 hour more per week in the first 2 years after diagnosis, with this being increased to 4 hours per week from 3 years after diagnosis.
Bradley, Neumark, Bednarek, and Schenk (2005) find a reduction of employment by 17 percentage point 6 months after diagnosis. Part of this high percentage compared to other research may be due to the inclusion of all patients, instead of only those are still alive after multiple years. This probably results in a sample containing more severe cases.

While research into returning to work for AYAs has not been done extensively, it is important to gain more knowledge on the subject. Only 24 percent of AYA cancer patients that returned to work reported cancer related financial distress, while this is 68 percent for patients that did not return to work (Leuteritz et al., 2021). However the positive effects go beyond only financial related effects. Returning to work can improve the self esteem of cancer survivors (Peteet, 2000). For breast cancer patients, being employed improved the quality of life in general, while a positive effect is found for physical well being, functional well being and social well being (Timperi et al., 2013). Similarly breast cancer patients who continued working during and after treatment were less likely to suffer from psychosocial distress such as depression, had a better physical functioning and again a higher quality of life (Mahar, BrintzenhofeSzoc, & Shields, 2008). Not being able to return to work after illness can lead to social isolation and reduction of self-esteem (Spelten, Sprangers, & Verbeek, 2002). Further, patients often see returning to work as a symbol of regaining a normal life and full recovery (De Boer et al., 2008), while ongoing employment and return to work are linked to a sense of normalcy and control for patients as well (Blinder & Gany, 2020). This stresses the importance of work for the well-being of survivors.

B Appendix to Section 3

Table B.1: Oncological data on patients with partners

	Frequency	Percentage
Type of tumor		
Breast	325	26.0%
Bone, cartilage and soft tissue	25	2.0%
Central Nervous System	35	2.8%
Endocrine	60	4.8%
Hematologic	136	10.9%
Head and neck	19	1.5%
Skin	283	22.6%
Lung	11	0.9%
Male genitals	160	12.8%
Female genitals	102	8.1%
Digestive system	57	4.6%
Urinary tract	32	2.6%
Other	7	0.6%
Stage		
I	684	54.6%
II	288	23.0%
III	133	10.6%
IV	51	4.1%
Missing	96	7.7%

C Appendix to Section 4

C.1 Appendix to Subsection 4.2

Below a proof is given of equation (9) identifying the ATT under the assumptions mentioned in Section 4.2 (Lechner, 2011):

$$
\begin{aligned}
ATT &= E[y_{i1}^1 - y_{i1}^0 | d_i = 1] \\
&\overset{SUTVA}{=} E[y_{i1}|d_i = 1] - E[y_{i1}^0|d_i = 1] \\
&\overset{PT}{=} E[y_{i1}|d_i = 1] - (E[y_{i1}^0|d_i = 0] - E[y_{i0}^0|d_i = 0] + E[y_{i0}^0|d_i = 1]) \\
&\overset{SUTVA}{=} E[y_{i1}|d_i = 1] - (E[y_{i1}|d_i = 0] - E[y_{i0}|d_i = 0] + E[y_{i0}^0|d_i = 1]) \\
&\overset{NEPT}{=} E[y_{i1}|d_i = 1] - (E[y_{i1}|d_i = 0] - E[y_{i0}|d_i = 0] + E[y_{i0}^1|d_i = 1]) \\
&\overset{SUTVA}{=} E[y_{i1}|d_i = 1] - (E[y_{i1}|d_i = 0] - E[y_{i0}|d_i = 0] + E[y_{i0}|d_i = 1]) \\
&= (E[y_{i1}|d_i = 1] - E[y_{i1}|d_i = 0]) - (E[y_{i0}|d_i = 1] - E[y_{i0}|d_i = 0])
\end{aligned}
$$

C.2 Appendix to Subsection 4.3

In equation 10 noted α corresponds to the effect of interest, the ATT. To provide some intuition behind this claim, equation (9) can be used by substituting y_{it} from equation (10) (Theloudis, 2021). Note that using Assumption 2 it is known that $E[x_{it}|d_i = 0] = E[x_{it}|d_i = 1] = E[x_{it}]\ \forall t$, which will be used below. Let two time periods be considered again: $t = 0$ before treatment and $t = 1$ after treatment. Then taking the expectations in equation (9) we get:

- $E[y_{i0}|d_i = 0] = \beta_0 + \lambda_0 + \beta E[x_{i0}]$
- $E[y_{i0}|d_i = 1] = \beta_0 + \lambda_0 + \gamma + \beta E[x_{i0}]$
- $E[y_{i1}|d_i = 0] = \beta_0 + \lambda_1 + \beta E[x_{i1}]$
- $E[y_{i1}|d_i = 1] = \beta_0 + \gamma + \alpha + \lambda_1 + \beta E[x_{i1}]$

And thus equation (9) yields:

$$
\begin{aligned}
ATT &= (E[y_{i1}|d_i = 1] - E[y_{i1}|d_i = 0]) - (E[y_{i0}|d_i = 1] - E[y_{i0}|d_i = 0]) \\
&= ((\beta_0 + \gamma + \alpha + \lambda_1 + \beta E[x_{i1}]) - (\beta_0 + \lambda_1 + \beta E[x_{i1}]) - \\
&\qquad ((\beta_0 + \lambda_0 + \gamma + \beta E[x_{i0}]) - (\beta_0 + \lambda_0 + \beta E[x_{i0}])) \\
&= \gamma + \alpha - \gamma \\
&= \alpha
\end{aligned}
\tag{18}
$$

D Detailed results patients

D.1 Appendix to Section 5.2.1

The trend in Figure 4 resembles data from Eurostat (2022) based on The European Union Labour Force Survey for the same time period. In the data the Dutch population of 15 through 39 year old are included. Figure D.1 shows the employment over the years based on these data. The age composition of people considered in the Eurostat data differs slightly. In the control group the same people are followed over years, while the Eurostat considers people that are in certain age groups in the year observed. Therefore the average age of people in the control group increases over the years while this hardly changes in the Eurostat data. Altogether due to these reasons it is logical that there are differences between the figures, while the trend is somewhat similar.

Figure D.1: Employment in the Netherlands for age group 15-39 using data from Eurostat (2022)

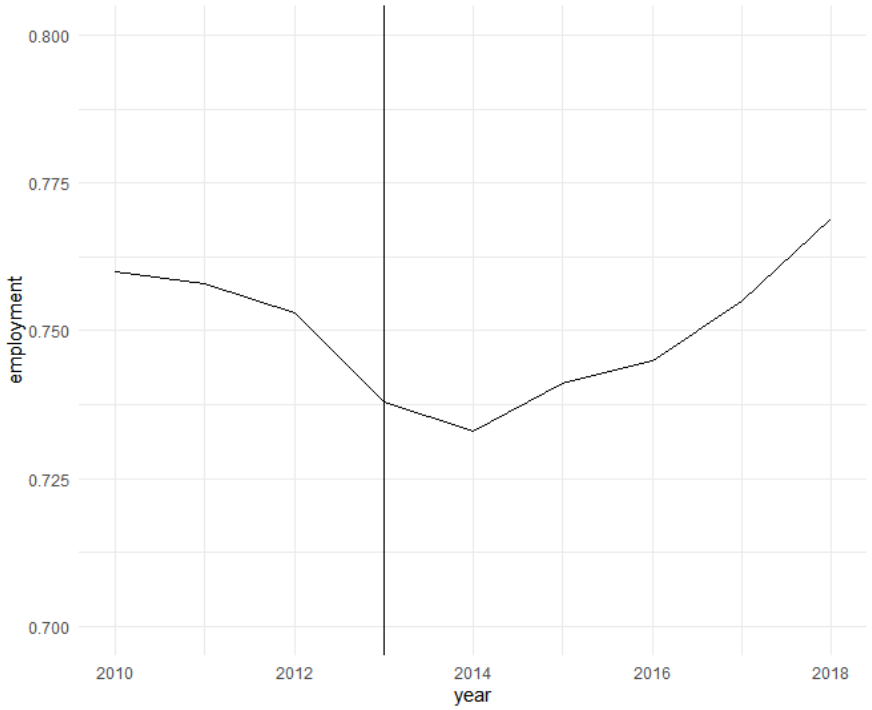
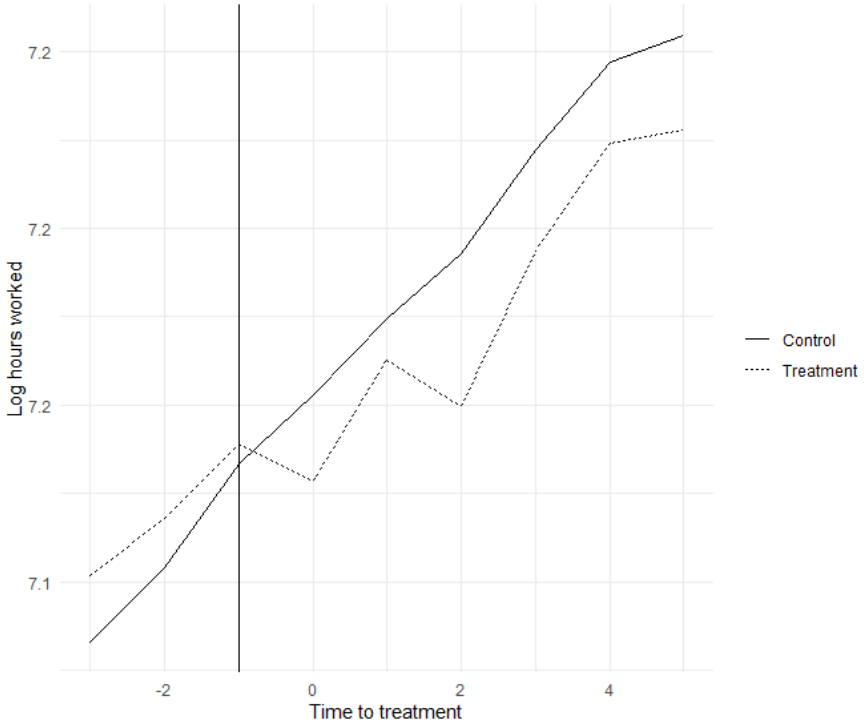
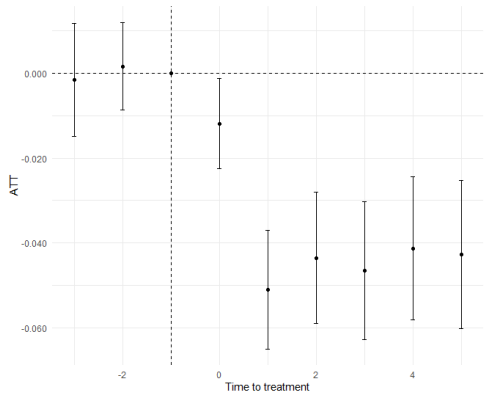


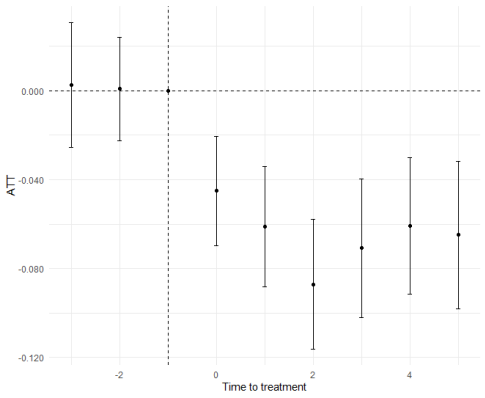
Figure D.2: log hours worked control and treatment group



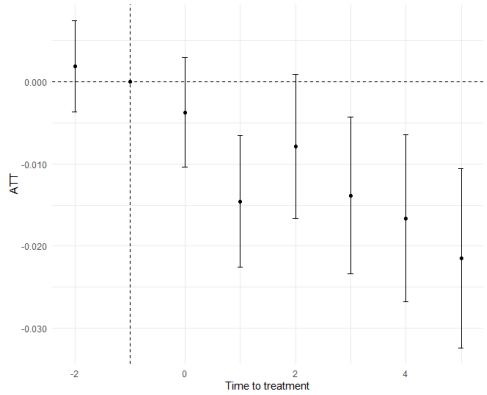
D.2 Dynamics treatment effect



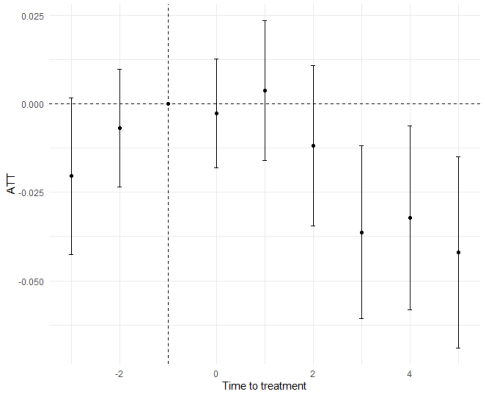
(a) employment,
Wald test p-value: 0.83



(b) Log hours worked,
Wald test p-value: 0.98



(c) Self employment,
Wald test p-value: 0.51



(d) Contract type,
Wald test p-value: 0.18

Figure D.3: Dynamic treatment effects patients with 95 percent confidence interval, using TWFE estimation

D.3 Heterogeneity treatment effect

Table D.1: Heterogeneity effects on employment by individual characteristics for patients ($n = 2456$)

	Education level	Gender	Migration background	Married or partner
	Estimate (Std. Error)	Estimate (Std. Error)	Estimate (Std. Error)	Estimate (Std. Error)
post	-0.023** (0.009)	-0.033*** (0.002)	-0.032*** (0.002)	-0.028*** (0.002)
secondary	0.188*** (0.015)	0.177*** (0.014)	0.177*** (0.014)	-0.014*** (0.004)
higher	0.229*** (0.016)	0.223*** (0.015)	0.223*** (0.015)	0.177*** (0.014)
Partner or married	-0.020*** (0.003)	-0.020*** (0.003)	-0.020*** (0.003)	0.222*** (0.015)
post*treated	-0.058* (0.032)	-0.058*** (0.008)	-0.031*** (0.006)	-0.034*** (0.008)
post*secondary	-0.010 (0.009)			
post*higher	-0.004 (0.010)			
treated*secondary	-0.058 (0.048)			
treated*higher	-0.079 (0.049)			
post*treated*secondary	0.002 (0.033)			
post*treated*higher	0.050 (0.033)			
post*male		0.007** (0.003)		
post*treated*male		0.043*** (0.012)		
post*first generation			0.008 (0.007)	
post*second generation			0.012* (0.006)	
post*treated*first generation			-0.040 (0.025)	
post*treated*second generation			-0.051** (0.022)	
post*Partner or married				-0.007* (0.003)
treated:Partner or married				0.001 (0.015)
post*treated*Partner or married				-0.014 (0.012)

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, education level, direct migration background, indirect migration background, registered partnership & marital status and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on
Further, Post*treated is the baseline treatment effect. Terms with three multiplications are the heterogeneity with respect to that variable.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table D.2: Heterogeneity effects on log hours worked by individual characteristics for patients ($n = 1936$)

	Education level Estimate (Std. Error)	Gender Estimate (Std. Error)	Migration background Estimate (Std. Error)	Married or partner Estimate (Std. Error)
post	0.080*** (0.025)	0.025*** (0.004)	0.058*** (0.003)	0.118*** (0.004)
secondary	0.300*** (0.043)	0.295*** (0.039)	0.303*** (0.039)	0.051*** (0.007)
higher	0.640*** (0.044)	0.615*** (0.040)	0.623*** (0.040)	0.290*** (0.038)
Partner or married	-0.050*** (0.006)	-0.054*** (0.006)	-0.051*** (0.006)	0.598*** (0.040)
post*treated	-0.135* (0.080)	-0.077*** (0.014)	-0.051*** (0.011)	-0.073*** (0.015)
post*secondary	0.000 (0.025)			
post*higher	-0.035 (0.025)			
treated*secondary	-0.019 (0.123)			
treated*higher	-0.016 (0.128)			
post*treated*secondary	0.043 (0.081)			
post*treated*higher	0.105 (0.081)			
post*male		0.097*** (0.007)		
post*treated*male		0.025 (0.022)		
post*first generation			0.004 (0.016)	
post*second generation			0.087*** (0.013)	
post*treated*first generation			-0.087** (0.043)	
post*treated*second generation			-0.106** (0.045)	
post*Partner or married				-0.146*** (0.006)
treated:Partner or married				-0.024 (0.024)
post*treated*Partner or married				0.024 (0.021)

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, education level, direct migration background, indirect migration background, registered partnership & marital status and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on
Further, Post*treated is the baseline treatment effect. Terms with three multiplications are the heterogeneity with respect to that variable.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table D.3: Heterogeneity effects on employment by oncological data for patients

	Treatment effect TWFE		treatment effect DR	
	Estimate	Std. Error	Estimate	Std. Error
Type of tumor				
Breast ($n = 493$)	-0.058***	0.012	-0.056***	0.014
Hematologic ($n = 321$)	-0.045***	0.018	-0.048***	0.018
Skin ($n = 510$)	-0.0029	0.011	0.000	0.011
Male genitals ($n = 433$)	0.002	0.013	-0.009	0.016
Female genitals ($n = 193$)	-0.039	0.024	-0.031	0.025
Endocrine ($n = 123$)	-0.022	0.028	-0.016	0.028
Digestive system ($n = 115$)	-0.024	0.028	0.005	0.030
CNS ($n = 75$)	-0.256***	0.043	-0.229***	0.049
Stage				
I ($n = 1351$)	-0.022	0.017	-0.007	0.008
II ($n = 539$)	-0.134***	0.027	-0.056***	0.015
III ($n = 233$)	-0.128***	0.040	-0.071***	0.022
IV ($n = 106$)	-0.130***	0.058	-0.090***	0.031
Number of different types of treatment				
0 ($n = 60$)	-0.079*	0.046	-0.048	0.045
1 ($n = 1348$)	-0.024***	0.008	-0.023***	0.008
2 ($n = 602$)	-0.050***	0.012	-0.053***	0.014
3 ($n = 245$)	-0.065***	0.020	-0.059***	0.022
4 ($n = 201$)	-0.101***	0.029	-0.090***	0.033
Type of treatment				
Organ surgery ($n = 1340$)	-0.035***	0.008	-0.036***	0.009
Local Surgery ($n = 709$)	-0.025***	0.011	-0.023**	0.011
Chemotherapy ($n = 1014$)	-0.064***	0.010	-0.063***	0.011
Radiotherapy ($n = 720$)	-0.071***	0.011	-0.065***	0.012
Hormonal therapy ($n = 307$)	-0.058***	0.016	-0.060***	0.018

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, education level, direct migration background, indirect migration background, registered partnership & marital status and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on Reported sample size n refers to the number of treated individuals considered. The number of controls is approximately 10 times the reported number of treated individuals.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table D.4: Heterogeneity effects on log hours worked by oncological data for patients

	Treatment effect TWFE		treatment effect DR	
	Estimate	Std. Error	Estimate	Std. Error
Type of tumor				
Breast ($n = 385$)	-0.105***	0.020	-0.083***	0.025
Hematologic ($n = 254$)	-0.188***	0.031	-0.152***	0.036
Skin ($n = 435$)	0.018	0.021	0.007	0.026
Male genitals ($n = 366$)	-0.036	0.027	0.019	0.031
Female genitals ($n = 148$)	0.051	0.045	0.060	0.057
Endocrine ($n = 96$)	-0.118**	0.057	-0.119*	0.068
Digestive system ($n = 86$)	-0.056	0.053	-0.060	0.054
CNS ($n = 48$)	-0.364***	0.072	-0.147	0.112
Stage				
I ($n = 1120$)	-0.011	0.014	-0.009	0.018
II ($n = 451$)	-0.124***	0.022	-0.078***	0.027
III ($n = 197$)	-0.104***	0.036	-0.040	0.043
IV ($n = 85$)	-0.152***	0.048	-0.076	0.063

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, education level, direct migration background, indirect migration background, registered partnership & marital status and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

E Detailed results partners

Table E.1: Heterogeneity effects on employment by individual characteristics for partners ($n = 1252$)

	Education level Estimate (Std. Error)	Number of children Estimate (Std. Error)	Gender Estimate (Std. Error)	Migration Background Estimate (Std. Error)	Married Estimate (Std. Error)
post	-0.054** (0.024)	-0.029*** (0.005)	-0.022*** (0.003)	-0.020*** (0.003)	-0.026*** (0.005)
married	-0.011* (0.006)	-0.009 (0.006)	-0.010 (0.006)	-0.011* (0.006)	-0.014* (0.008)
Children	-0.018*** (0.003)	-0.020*** (0.004)	-0.017*** (0.003)	-0.017*** (0.003)	-0.017*** (0.003)
post*treated	0.086 (0.059)	-0.002 (0.011)	0.001 (0.008)	0.007 (0.007)	0.012 (0.011)
post*secondary	0.025 (0.025)				
post*higher	0.036 (0.025)				
post*treated*secondary	-0.082 (0.059)				
post*treated*higher	-0.081 (0.059)				
post*children		0.003 (0.003)			
treated*children		0.002 (0.009)			
post*treated*children		0.005 (0.007)			
post*female			-0.009 (0.006)		
post*treated*female			0.019 (0.014)		
post*first generation				-0.015 (0.012)	
post*second generation				-0.031** (0.013)	
post*treated*first generation				-0.022 (0.030)	
post*treated*second generation				0.021 (0.029)	
post*married					0.003 (0.006)
treated:married					0.012 (0.015)
post*treated*married					-0.010 (0.014)

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, number of children, duration relationship before 2013, direct migration background, indirect migration background, marital status, education level and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on.
Further, Post*treated is the baseline treatment effect. Terms with three multiplications are the heterogeneity with respect to that variable.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table E.2: Heterogeneity effects on log hours worked by individual characteristics for partners ($n = 1050$)

	Education level Estimate (Std. Error)	Number of children Estimate (Std. Error)	Gender Estimate (Std. Error)	Migration Background Estimate (Std. Error)	Married Estimate (Std. Error)
post	-0.053 (0.050)	-0.026*** (0.008)	0.012** (0.005)	-0.010** (0.005)	0.008 (0.008)
married	-0.017* (0.009)	-0.012 (0.009)	-0.013 (0.009)	-0.016* (0.009)	-0.008 (0.011)
Children	-0.060*** (0.005)	-0.066*** (0.006)	-0.056*** (0.005)	-0.059*** (0.005)	-0.059*** (0.005)
post*treated	0.078 (0.126)	0.005 (0.018)	-0.013 (0.013)	-0.006 (0.011)	-0.023 (0.018)
post*secondary	0.036 (0.050)				
post*higher	0.060 (0.050)				
post*treated*secondary	-0.096 (0.126)				
post*treated*higher	-0.072 (0.127)				
post*children		0.013*** (0.004)			
treated*children		-0.001 (0.014)			
post*treated*children		-0.006 (0.010)			
post*female			-0.063 (0.010)		
post*treated*female			0.020 (0.024)		
post*first generation				0.021 (0.020)	
post*second generation				0.011 (0.018)	
post*treated*first generation				0.003 (0.064)	
post*treated*second generation				-0.006 (0.036)	
post*married					-0.023** (0.009)
treated:married					-0.001 (0.028)
post*treated*married					0.024 (0.023)

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, number of children, duration relationship before 2013, direct migration background, indirect migration background, marital status, education level and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on.
Further, Post*treated is the baseline treatment effect. Terms with three multiplications are the heterogeneity with respect to that variable.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table E.3: Treatment effect employment partner by oncological characteristics

	Treatment effect TWFE		treatment effect DR	
	Estimate	Std. Error	Estimate	Std. Error
Type of tumor				
Breast ($n = 325$)	-0.015	0.013	-0.027**	0.014
Hematologic ($n = 136$)	0.020	0.023	0.014	0.031
Skin ($n = 283$)	0.004	0.013	-0.001	0.015
Male genitals ($n = 160$)	0.030	0.019	0.044	0.028
Female genitals ($n = 102$)	0.026	0.029	0.025	0.032
Endocrine ($n = 60$)	0.017	0.026	0.081**	0.040
Digestive system ($n = 57$)	-0.024	0.030	-0.042	0.037
CNS ($n = 35$)	0.032	0.043	0.039	0.059
Stage				
I ($n = 684$)	0.009	0.009	0.013	0.011
II ($n = 288$)	-0.017	0.014	-0.025	0.017
III ($n = 133$)	0.010	0.020	0.031	0.025
IV ($n = 51$)	0.034	0.046	-0.003	0.053

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, number of children, duration relationship before 2013, direct migration background, indirect migration background, education level, marital status and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on.
Reported sample size n refers to the number of treated individuals considered. The number of controls is approximately 5 times the reported number of treated individuals.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table E.4: Treatment effect log hours worked partner by oncological characteristics

	Treatment effect TWFE		treatment effect DR	
	Estimate	Std. Error	Estimate	Std. Error
Type of tumor				
Breast ($n = 262$)	-0.013	0.021	-0.011	0.030
Hematologic ($n = 107$)	0.051	0.038	0.056	0.050
Skin ($n = 246$)	-0.016	0.019	0.000	0.023
Male genitals ($n = 139$)	-0.003	0.035	-0.011	0.043
Female genitals ($n = 84$)	-0.036	0.049	-0.075	0.060
Endocrine ($n = 51$)	0.071**	0.035	0.006	0.087
Digestive system ($n = 50$)	-0.091*	0.051	-0.153***	0.049
CNS ($n = 29$)	0.028	0.050	0.054	0.078
Stage				
I ($n = 582$)	-0.009	0.015	-0.023	0.018
II ($n = 236$)	0.002	0.023	0.004	0.030
III ($n = 108$)	-0.010	0.027	-0.051	0.032
IV ($n = 41$)	-0.010	0.064	-0.003	0.053

Note: Standard errors are clustered on an individual level. Control variables used: age in 2013, number of children, duration relationship before 2013, direct migration background, indirect migration background, education level, marital status and gender. Note that for TWFE models, time invariant controls are left out of the estimation. For DR estimation, the control variables are used to condition on.
Reported sample size n refers to the number of treated individuals considered. The number of controls is approximately 5 times the reported number of treated individuals.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

F Appendix to Section 6

Table F.1: Estimated effects without education as control

		TWFE		DR	
Group	Outcome	Estimate	Std. Error	Estimate	Std. Error
Patients					
	Employed ($n = 2456$)	-0.040***	0.006	-0.040***	0.006
	Log hours worked ($n = 1936$)	-0.069***	0.011	-0.037***	0.013
Partners					
	Employed ($n = 1252$)	0.007	0.007	0.008	0.008
	Log hours worked ($n = 1050$)	-0.006	0.011	-0.024*	0.014

Note: Standard errors are clustered on an individual level. Control variables used: Same as in other tables for both patients and partners, but without education
Reported sample size n refers to the number of treated individuals considered. The number of controls is approximately 5 times the reported number of treated individuals.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure F.1: Number of children over different years