Chapter 4: The effects of household cancer diagnosis on the labor supply of young adults

4.1 Introduction

Individuals with cancer are thought to spend nearly four times the amount of total health care expenditures per year than individuals without cancer (Park and Look, 2019). In fact, cancer can often be so costly to treat that researchers have estimated that 40% of older patients with aggressive cancers have depleted their entire life's savings after only two years (Gilligan et al., 2018). Individuals with cancer are more than 2.5 times more likely to file for bankruptcy than individuals without cancer (Ramsey et al., 2013), 1.7 times more likely to have an adverse financial event[40], and 1.3 times more likely to have a past due credit card payment (Shankaran et al., 2022). Because cancer is time-intensive, costly, and debilitating disease, the labor of patients and their household members can frequently be volatile. Patients with cancer often reduce labor supply in the first two years after their diagnosis (Oberst et al., 2010; Bradley et al., 2006), and the psychological and economic strain of medical disease can be as severe for the caregiver as for the patient himself (Grunfeld et al., 2004;

[40] Shankaran et al. (2022) group the presence of any of the following as an Adverse Financial Event: third-party collections, charge-offs, tax liens, delinquent mortgage payments, foreclosures, and repossessions.

Northouse et al., 2010; Kornblith et al., 1994; Given et al., 1993; Cliff and MacDonagh, 2000; Bishop et al., 2007).

Past work has shown mixed labor responses from caregivers, some showing increases in spousal or caregiver labor supply (Siegel, 2006; Fadlon and Nielsen, 2021) and others showing decreases in spousal or caregiver labor supply (Lee, 2020; Grunfeld et al., 2004). Still others have noted that labor supply may respond differently by gender of the spouse (Berger, 1983). While the literature on spousal labor response to a financial or medical shock was the subject of significant study in past decades (Lundberg, 1985; Juhn and Potter, 2007; Maloney, 1991, 1987; Haurin, 1989; Spletzer, 1997; Gruber and Cullen, 1996), and seems to have re-captured the interest of some contemporary researchers (Fadlon and Nielsen, 2021; Dobkin et al., 2018), little is known about the way in which young adults may respond to a health shock of a parent. There has been interest from scholars about the labor response of parents to children (Gould, 2004; Corcnan et al., 2005; Noonan et al., 2005; Wasi et al., 2012; Powers, 2001, 2003; Eriksen et al., 2021) but the corpus of research related to the effects of parental illness on childhood labor supply seems to be limited to a few articles in Scandinavia that study labor responses of children to lone parents with terminal illness, which find small declines in labor supply in the children with a recent parent death (Fevang et al., 2012; Norén, 2020). These two articles also find conflicting results in differences in responses by gender. The contextual landscape of the Scandinavian social safety nets often make findings less comparable to the United

States; and moreover, neither of these studies focuses on the response young adult children.[41]

In this study, I focus on the labor responses of young adults who have an adult household member diagnosed with cancer. The sample of young adults analyzed in this study are the group of young adults enrolled in and graduating from four-year colleges, so they may face a more pronounced trade off in their time commitments. Having a sick household member exacerbates the time commitments that young adults are managing. Because young adults are constrained by a fixed amount of time, they may rationally adjust their time commitments. Three areas where this could be possible is in their college enrollment status, in their labor supply, and in their leisure. In Chapter 3, I found little difference in enrollment of young adults affected by household cancer diagnosis. In this chapter, I study the labor supply responses, which adds additional understanding into the effects of household cancer diagnosis in the lives of young adults, and the inter-generational effects of chronic illness.

I draw from several literatures to motivate this question, and to anticipate a response. In particular, I integrate the theory on household production functions from Becker (1981), which rationalizes the economic benefit of a household, with the theory of the Added Worker Effect (Lundberg, 1985; Mincer, 1958), which conceptualizes the spousal response to an income shock. Furthermore, I incorporate the theory of human capital investment in higher education (Mincer, 1958; Becker, 1962; Ben-Porath, 1967; Keane and Wolpin, 1997) that relates college enrollment and labor supply to rationalize the higher education decisions of young adults. Putting these areas of

[41] The average age of the adult children whose labor responses are studied in the cited works is more than 50 years old.

research together allows me to anticipate a response that household cancer diagnosis may have on the labor supply of young adults.

In order to understand the labor responses of young adults to household cancer diagnosis, I integrate four administrative datasets that allows for novel insight into the economic lives of young adults. Using higher education enrollment and course outcomes data from the Ohio Department of Higher Education, I take young adults who are enrolled at four year public colleges in Ohio, and identify the individuals residing in their household at the beginning of college, using the Ohio Consumer Credit panel. I then use the Ohio Cancer Incidence Surveillance System (OCISS), which is the cancer registry for the state of Ohio, to identify the group of young adults who are affected by household cancer diagnosis. Finally, I integrate labor data from the Ohio Department of Job and Family Services (ODJFS) to analyze differences in labor supply in the several years after diagnosis between the treated and comparison groups. To do this, I estimate a series of stacked difference in differences models that match young adults who began at the same university in the same year. Finally, because labor force decisions may differ by household characteristics of the young adults, I estimate models to consider heterogeneity in the aggregate outcomes.

I find that labor supply generally decreases by about two percentage points in the first year for young adults children who are affected by household cancer diagnosis, relative to the counterfactual. For young adults affected by severe cancer diagnosis, I find a sustained decrease of about five percentage points in the first three years after cancer diagnosis compared to what they would have otherwise supplied. This decrease in employment can also be viewed as decreases in the number of weeks worked, which

I also explore. I find that, in aggregate, young adults reduce the number of weeks worked in a year by about one to two weeks per year for at least the first five years of cancer diagnosis. For those affected by severe cancer, this result is more pronounced. I find that young adults from severely-affected households reduce labor by about two weeks in the first year, and by about five weeks in the fourth year. I explore heterogeneity in these results, and find that young adults who are disadvantaged— who might be less able to adjust their labor supply— generally do not, while young adults who are relatively more advantaged are the ones reducing the number of weeks worked per year. This is especially true for young adults where cancer is especially severe. I also find that young adults who attend college close to home are more likely to reduce labor supply, whereas young adults who attend college further away are not. This latter result is perhaps an artifact of the sensitivity of caregiving to location.[42]

In the next section, I lay out a conceptual framework that is motivated by the literature referenced earlier. I describe the data used in this analysis, and define the empirical model. I describe the results of the models, as well as heterogeneity in the results. In the conclusion, I piece the results together to understand how they fit with the broader literature on caregiving and household production functions.

4.2 Conceptual motivation

Households have often been considered insurance mechanisms for the individual. Gary Becker's "Household Production Model" (Becker, 1965) and much of his later work

[42]Indeed, AARP noted that in 2020 76% of caregivers lived within 20 minutes of the patient's home, with about half of local caregivers living in the home and the other half living independently (AARP, 2020). These results are also congruent with a broader literature on caregiving, which finds positive associations between the distance of a caregivers and the emotional and financial burdens (Chou et al., 2001).

gave rational justifications for living together as a household, rather than as separate individuals. Just as with David Ricardo's theory of comparative advantage in international trade, households members can specialize in the chores for which they had the comparative advantage, and thereby realize more than the sum of their individual production. The household can minimize risk by diversifying assets and investments, and by sharing *extra muros* and *intra muros* labor responsibilities. In a related theory of the Added Worker Effect, the household self-insures against income or financial shocks by flexibly reallocating time towards the labor market (Lundberg, 1985). Decisions to alter time commitments are no longer purely altruistic, but instead have a rational motive that inspires them.

Becker suggests that an individual's utility function is compromised in some part by the outcomes of his family members instead of purely by individual consumption and leisure. Because the psychological transition from "child" to "adult" is a process that spans decades (indeed, only a minority of college students view themselves as "adults") (Arnett, 1994), an important component of young adult's utility function, and hence how he will behave, will depend on the health of his family.

As noted, cancer diagnosis has been shown to decrease the labor supply of the patient (Oberst et al., 2010; Bradley et al., 2006). In aggregate, the Added Worker Effect suggests that for young adults this provides incentive to work more to supplement the missed income from the labor market. For young adults enrolled in college, the cost of maintaining the pre-diagnosis lifestyle exacerbates this incentive even more. A 2023 report from Sallie Mae found that parents or relatives paid for about 55% of the cost of tuition for four year public colleges (Ipsos, 2023). Yet on the other hand, cancer is the second leading cause of death in the United States.[43] And, moreover, the perception of the seriousness of a cancer diagnosis—particularly to children (Michielutte and Diseker, 1982)—gives rationale for wanting to be available to assist a loved one or offer care. Indeed, a survey from AARP found that the largest group of caregivers are household members of the individual receiving care (AARP, 2020). However, time invested in caregiving has an opportunity cost. The decision of how much labor to supply to the labor force is a complex balance of these competing incentives. As a result, on aggregate, this question of the labor response of a young adult to a household member's cancer diagnosis is perhaps best understood empirically.

4.2.1 Heterogeneity

There are particular groups of young adults, though, that might respond in a more predictable way. In particular, young adults who come from households that have more financial flexibility may be able to reallocate time to caregiving instead of work. Gupta et al. (2018) finds that individuals who are diagnosed with cancer who are able to access liquidity from their homes, all else equal, have better health outcomes than patients who are credit constrained. Similarly, a young adult whose household is able to access wealth from its home (or other assets) may have greater capacity to reduce labor supply to take care of a sick family member.

By contrast, a young adult from a household that is financially constrained may have less capacity to reduce labor supply. These young adults are likely more dependent on their own labor for consumption, and hence they may be unable to adjust labor to offer care. Young adults from these situations may still seek to provide assistance,

[43] https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm

but may be forced to reallocate time from other aspects of their lives, like time spent on leisure.

4.3 Data

Details of the important inclusion and exclusion restrictions that are used in the sample definition in this essay are discussed in depth in Chapter 3 and in Table 3.2. *In summa*, the sample is comprised of the universe of young adults who began in Ohio public four year colleges between academic years 2015 and 2020. I exclude individuals who did not graduate from an Ohio high school, since the cancer registry only captures diagnoses in Ohio. To minimize the possibility of defining dorm mates as household members, I exclude individuals who do not match to a household with at least one parent or grandparent-aged individual or who have more than six parent/grandparent-aged individuals. Amongst a few other smaller limitations, described in more depth in the previous chapter, I also exclude young adults whose households were affected by cancer before college and young adults who themselves are affected by cancer.

Higher education information system

The source of data from which the sample of students is defined is the Higher Education Information System (HEI) from the Ohio Department of Higher Education.[44] At its most granular level, it provides course level enrollment and outcomes data for the universe of students enrolled in Ohio public colleges between academic years 2015 and 2020. It also provides some demographic information on the students that I take advantage of to define the sample, which is discussed in depth in Chapter 3. In particular, it provides information on the year of high school graduation and age. Using

[44] https://highered.ohio.gov/data-reports/hei-system

a secured hashing process, I am able to match the young adults in HEI to their credit data, which allows me to observe household members residing at the same address, as discussed more next.

Experian consumer credit data

I integrate consumer credit records from the Ohio Consumer Credit Panel (CCP). This dataset is comprised of the universe of consumer credit records from the state of Ohio quarterly from Q4 2015 to Q4 2021, approximately 8.8 million individuals or about 95% of the Ohio adult population.[45] While the CCP allows me to see a wide array of important financial characteristics at the quarterly level such as debt levels, measures of financial delinquency, public filings like bankruptcies, and credit score, in this analysis it is primarily used to define the baseline household members for the young adult. In particular, because I have the universe of consumers in Ohio, and because the dataset includes a household identifier that groups individuals together by the address listed in their credit records, I am also able to identify household members of the young adults. When I explore heterogeneity in the results by student loan balance, I also retrieve this information from the CCP.

Ohio cancer incidence surveillance system

For this research, I obtained access to the Ohio Cancer Incidence Surveillance System (OCISS) for 2015-2022, which is the state cancer registry for the state of Ohio, collected by the Ohio Department of Health. By Ohio law, all cancer diagnosis and treatments are required to be submitted to the OCISS, and thus this registry captures

[45] Prior studies estimate that approximately 11 percent of adults in the U.S. do not have a credit file (Brevoort et al., 2016). However, coverage in credit data has expanded over the past few years. Further, Ohio has a very small immigrant population and thus fewer people who have not yet established a credit file compared to states like California, Texas, Florida, and New York.

the universe of cancer diagnoses in Ohio. In addition to date of diagnosis, this registry provides information on the type of cancer, grade, laterality, site of tumor, and a few basic demographic characteristics of the patient. It provides an indicator for the mortality of the patient. Through a secure and anonymous hashing process that is congruent to the hashing process used to connect the HEI data to the other administrative datasets in this essay, the OCISS is linked to the Experian credit panel, and ultimately to the young adult via the young adults's household as defined in Experian and described above. Additional detail on this dataset are discussed in Chapter 3.

Employment data

Finally, I also integrate employment data about young adults and household members, which is provided by the Ohio Department of Job and Family Services (ODJFS). This employment data from ODJFS is quarterly data on all employed individuals in the state of Ohio, and includes quarterly wages and weeks employed. It is compiled as part of the Ohio Longitudinal Data Archive (OLDA), which is a project of the Ohio Education Research Center (oerc.osu.edu) and provides researchers with centralized access to administrative data. The OLDA is managed by The Ohio State University's Center for Human Resource Research (chrr.osu.edu) in collaboration with Ohio's state workforce and education agencies (ohioanalytics.gov), with those agencies providing oversight and funding.[46]

I collapse this data to the academic year (rather than calendar year) by aggregating from July to June (rather than January to December), which both allows for comparison with Chapter 3 and which avoids some of the challenges of modeling quarterly

[46]For information on OLDA sponsors, see http://chrr.osu.edu/projects/ohio-longitudinal-data-archive .

outcomes with high degrees of seasonality. Because I have access to wage data for individuals from 2014 to 2022, the models allow for the possibility of examining individuals who have already graduated from university. This is a deviation from the models that censor individuals after graduation in Chapter 3. To correct for the possibility that individuals may leave the state of Ohio for employment, and hence we would incorrectly observe no wage earnings, I have integrated credit records, and censored individuals who are no longer in Ohio credit data. With this caveat noted, the years for which I model data is consistent with the previous chapter, namely I incorporate wage data for academic years 2015 to 2020.

4.4 Empirical motivation

Congruent to past chapters of this dissertation is the caveat that financial shocks to the household that are the results of chronic medical illness are contextually different than non-medical income shocks, in that they are not transitory income shocks[47] and that they affect the household's labor supply directly. The context of chronic illness results in competing incentives for households.

On one hand, medical costs of disease management are substantial. As noted earlier, dramatically increased expenditures for cancer patients has resulted in higher rates of adverse financial events, bankruptcy, past due credit card payments, and decreased savings (Park and Look, 2019; Gilligan et al., 2018; Ramsey et al., 2013; Shankaran et al., 2022). Gupta et al. (2018) shows that a lack of credit access leads to an increase in mortality for cancer patients. That is, all else equal, patients who are

[47]The analysis in Dobkin et al. (2018) is certainly secondary to their broader inquiry into the financial effects of hospital admissions (i.e. a transitory shock), but the authors explore spousal labor response as one outcome. The authors find no evidence that spouses change labor supply in response to hospital admission.

unable to access credit, in particular from the equity in their home, are more likely to be distressed, less likely to take up recommended therapies, and more likely to die after cancer diagnosis than individuals who are more easily able to finance the disease treatment. Because an ability to manage a disease is critically dependent on an ability to pay for it, the rational justification (particularly for those who are credit constrained) is to increase labor supply. However, the time and physical costs have been shown to reduce patient labor supply (Oberst et al., 2010; Bradley et al., 2006), potentially exacerbating the situation. To the extent that the labor supply of patients is impacted by the disease, it may be incumbent on other household members to offset this loss of income through their own labor. [48]

Using the labor data described above, I visualize changes in labor supply for both patient and household members of the patient relative to the quarter proceeding diagnosis. Figure 4.1 visualizes the patient-half of the story. It shows the extent to which the medical context of the income shock is an important detail. Whereas financial shocks generally add incentive to increase labor supply, Figure 4.1 shows moderate declines in labor supply, wages, and employer count for patients following diagnosis. These descriptive trends are consistent with Oberst et al. (2010) and Bradley et al. (2006) that find increases in physical work-related disabilities and decreases in labor force attachment in at least the first year after cancer diagnosis. In cases of severe cancer, the magnitude of the decrease labor supply is understandably larger.

[48] Fadlon and Nielsen (2021) use Danish administrative records to explore household labor response to fatality and severe medical events (heart attack and stroke). The authors find that fatal events lead to notable increases in spousal labor, while non-fatal but severe events do not lead to changes in labor supply. The major caveat to this analysis is that the context is in Denmark, where the healthcare and social insurance environments are quite different than in the United States.

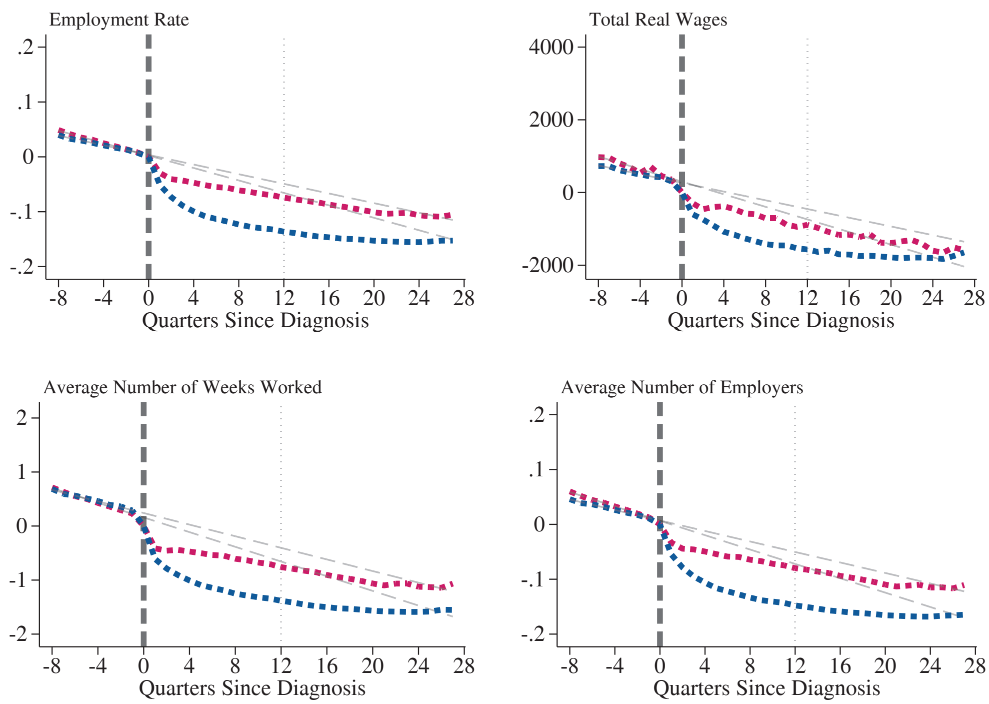


Figure 4.1: Labor supply of cancer patient relative to pre-diagnosis quarter

Note: Measures of labor supply for the patient are shown relative to the quarter prior to diagnosis (*t-1*), where *t=0* is the quarter of cancer diagnosis. The cyclamen line represents all cancer types. The blue line is cancer diagnoses that result in fatality. Line of best fit is generated with data from eight proceeding quarters.

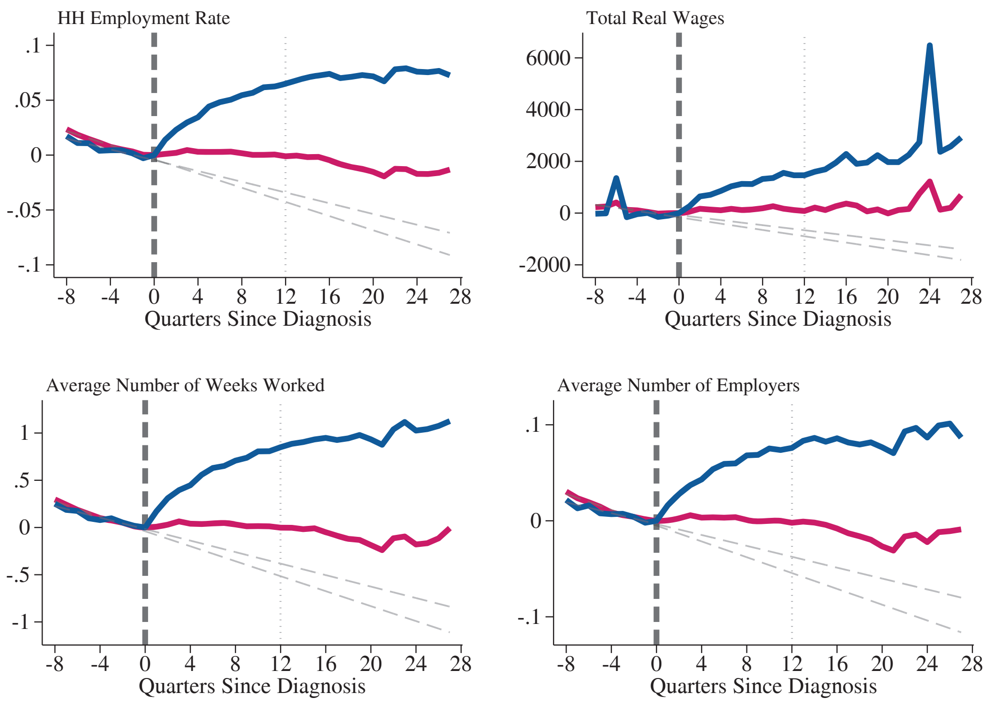


Figure 4.2: Labor supply of (non-patient) household members relative to pre-diagnosis quarter

Note: Measures of labor supply for the non-patient household members are shown relative to the quarter prior to diagnosis (*t-1*), where *t=0* is the quarter of cancer diagnosis. The cyclamen line represents all cancer types. The blue line is cancer diagnoses that result in fatality. Line of best fit is generated with data from eight proceeding quarters.

Figure 4.2 plots the same four labor measures in event time for adult household members of a patient diagnosed with cancer, relative to the quarter that proceeds diagnosis. It appears that diagnosis prompts an increase in labor supply for non-patient household members. We see a marked increase in the number of weeks worked in a quarter, number of employers, and employment rate immediately after cancer diagnosis. This is especially true for household members where cancer diagnosis ultimately ends in fatality, but is true in aggregate, too. Wages in aggregate appear to remain constant with pre-diagnosis levels, though this is a deviation from pre-diagnosis trends. In contrast, especially severe cases see relatively large increases in wages post-diagnosis, consistent with an increase in the other three measures. While past work has shown mixed labor responses from caregivers of cancer patients, the increasing trends in Figure 4.2 support findings documented in (Siegel, 2006; Fadlon and Nielsen, 2021).

These patient and household trends, paired with the conceptual motivation in Section 4.2 that suggests an ambiguous response from young adults, form the basis for the empirical analysis in the next section.

4.5 Empirical specification

I use a comparable empirical specification to Chapter 3. In particular, I estimate a regression of the following form:

$$Y_{i,n,s} = \alpha + \sum_{k=T_0}^{-2} \beta_k diag_{i,n,s} + \sum_{k=0}^{T_T} \beta_k diag_{i,n,s} + \theta_i + \sigma_s + \nu_n + \epsilon_{i,n,s}$$

The model estimates changes in outcome Y (e.g. number of employed weeks) of young adult i in year since beginning college n relative to the omitted period α by incorporating information about the counterfactual group of young adults never treated (i.e. for whom $diag_{i,n,s} = 0$).

Importantly, there are differences in the likelihood of being employed depending on how many years it has been since an individual began college. There is an increasing relationship between one's time since beginning college and his likelihood of being employed. As a result, as in Chapter 3, the primary time unit is a normalized year measure, n, which equates the number of years since the start of college, for young adult i in stack s. I include a fixed effect for normalized time, which captures typical changes in employment behavior over the span of collegiate and post-collegiate time. I also include a fixed effect for the individual, θ, which captures any unobservable time-invariant characteristics of the individual. Finally, given the modelling strategy, described in Chapter 3, I include a stack fixed effect, σ, which isolates the proper comparison of treated young adults to other young adults who begin in the same year at the same college. This ensures that young adults enrolled at a college where employment trends are unique are compared to young adults at this same college.

In the event study difference in differences model, event time dummies are captured in β_k, where the parallel trends assumption suggests that $\beta_{k<0}$ is statistically zero, and the treatment effect relative to pre-treatment is captured in $\beta_{k\geq 0}$. The stacking strategy allows me to make the proper comparisons between individuals, and also solves the critique of the two-way fixed effects estimator that has come under scrutiny lately (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; De Chaisemartin and d'Haultfoeuille, 2020), because I only compare young adults who are treated to young adults who will never be treated, and never to young adults who are already treated. It is worth noting, however, that because household cancer diagnosis is a relatively rare occurrence in this population, only a small share of the population will ever be treated, and so it is unlikely that this stacked strategy and a standard two-way fixed effects strategy will yield significantly different results since the relative share of "bad" comparisons (that is, comparisons of newly-treated to already-treated observations) is low.

One important note about the interpretation of the set of coefficients β_k: the intuitive interpretation of β_k is the difference in outcome, Y, relative to the baseline period $t-1$. For example, if β_0 is -0.1, then the interpretation is that in the first year after household cancer diagnosis, the decrease in employment is 10%. The event study difference in differences design allows this interpretation to have some inferential meaning, i.e. relative to a counterfactual group. However, one important point here is that due to the stark trends of some outcomes over the course of college and post-college *per se*, as is apparent below, it is more proper to understand β_k as the difference in outcome Y relative to the baseline period, conditional on the time fixed effect. In this case, conditional on the years since beginning college. This is a subtle difference in interpretation, but can be helpful in understanding the coefficients in light of pronounced trends in employment by age.

4.5.1 Heterogeneity

In this essay, I analyze heterogeneity in the outcomes by four characteristics of the young adult, as measured at baseline. First, I consider heterogeneity by household wealth advantage, as measured by wage earnings. In particular, I suggest that there could be differences in labor responses for young adults whose families have greater wage earnings compared to young adults who come from relatively disadvantaged households. Empirically, I define compare individuals from households where the maximum wage earnings are in the bottom 25 percentile of household members to those in the top 75 percent (i.e. not the bottom 25 percentile).

Second, I consider heterogeneity by flexibility in household labor supply. The intuition behind this dimension of heterogeneity is that households that do not have two (or more) workers in the labor force may be ones that are financially constrained, and hence the ability to adjust labor supply may be limited. Conversely, a household that already has two individuals in the labor force can also be viewed as having less labor flexibility, and hence a limited ability to respond (Fisher et al., 2019). This dimension is explored empirically. I define disadvantaged in this dimension in the data as being from a household where no more than one individual is in the labor force.

Third, I consider differences in outcomes by young adults who have relatively high levels of student loans. Empirically, I define this as having student loans in the top 25 percentile of young adults. Importantly, this groups young adults who have zero student loans due to financial assistance, scholarships, or grants with young adults who have zero student loans due to a financially supportive household. While this is a counter intuitive grouping, the idea here is that it focuses on the group of young adults for whom the financial effects of the cancer diagnosis might be most severe. As noted above, the intuition of these three dimensions of household heterogeneity is to understand how household financial constraints may influence a young adults's capacity to adjust his labor supply.

The fourth dimension of heterogeneity that I exploit is geographic. While the travel (and hence financial and time) costs of a young adult who attends college far away is greater than for a local young adult, another important difference between these two groups of individuals is the extent to which caregiving may be feasible. I define the disadvantaged group in this context as those who live at or above the 75 percentile of distance from home to college.[49] Empirically this equates to about 100 miles (160 km), or roughly the distance from Cincinnati to Columbus. As noted, the majority of caregiving is done locally, and so we might expect that the demand to assist in offering care for this group of young adults is reduced compared to the group who is local.

4.6 Main results

Figure 4.3 provides analysis into the extensive margin of labor supply. Because a young adult is counted as being employed if he earns any wage earnings at all during the year, the share of young adults working is actually fairly constant over time. This is presumably due to summer jobs, work study, and/or employment outside of studying. As a result, I consider a young adult to be employed as full time if he works for 13 weeks in any quarter of an academic year. Admittedly, this is an imperfect measure, because it relies on the number of weeks worked rather than on the number of hours worked in a week, which is unobserved in this data. Figures 4.4 and 4.5

[49] I use the Haversine formula to calculate the distance between the geographic coordinates of the centroid of the home zip code and college campus zip code of each young adult.

may provide complementary information on weeks of employment and wage earnings, respectively.

A second note in the results, is that the age of this sample is unique in the sense, *a priori*, we probably expect that employment and wage earnings will increase over time for the entire sample. Wage earnings are a reflection of productivity, so as young adults age, their wage earnings will increase, too. As a result, when discussing results in this section, it is critical to understand these results as being deviations from the counterfactual reality (i.e. one in which a household member was not affected by cancer) or age adjusted differences, rather than necessarily as increases or decreases in absolute labor supply.

Panel (a) of Figure 4.3 shows us that in all three groups, those whose household members are unaffected (grey), those whose household members are affected by non-fatal cancer (cyclamen), and those whose household members are affected by severe cancer (blue), the likelihood of being employed increases over time since beginning college. Unsurprisingly this is also true if we change the outcome of interest from employment to weeks of employment (Figure 4.4) to wage earnings (Figure 4.5). We observe that in all three representations of labor supply, the mean from the first year of college to the fourth year after beginning college is greatest for those whose household members will ever be affected by severe cancer, and least for those with non-severe cancer. This is remarkably consistent with unconditional enrollment behavior, shown in Chapter 3 and thus lends some credence to the idea that college and employment are competing time demands.

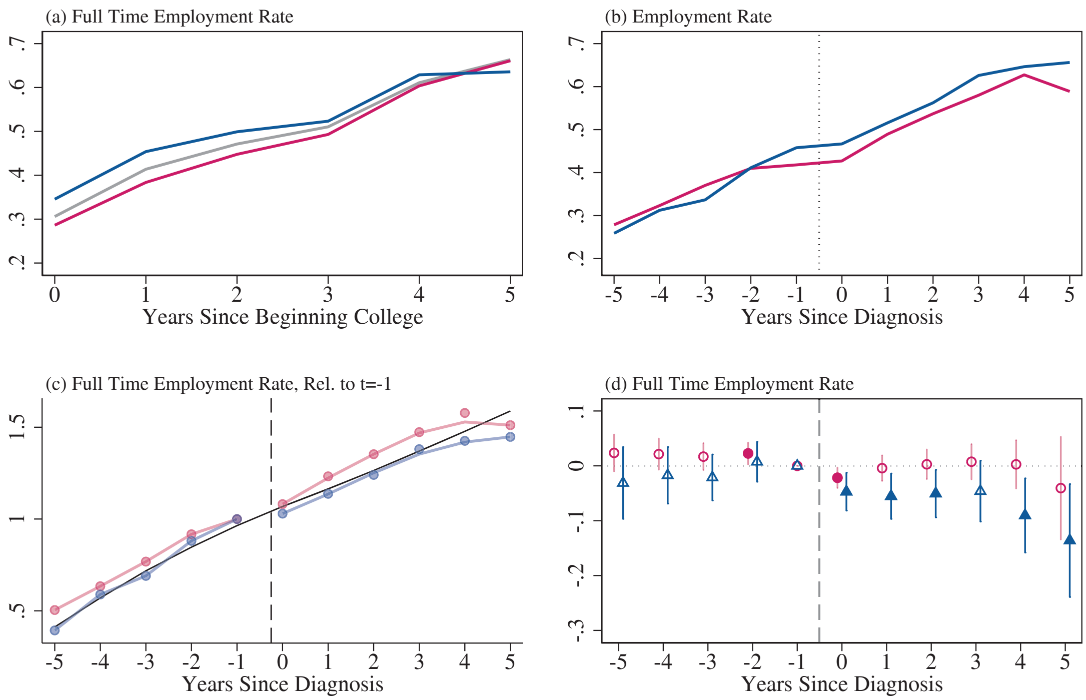


Figure 4.3: Difference in differences results - labor participation rates

Note: This figure presents four insights into understanding the relationship of cancer diagnosis and full time employment for the young adult. Full time employment is define as having 13 weeks of wage earnings from an employer during one quarter of an academic year. In panel (a), the full time employment rates for three groups of young adults are plotted over time since beginning college, agnostic to cancer timing. The grey line depicts the group of young adults who will never have a household member that is diagnosed with cancer. The cyclamen and blue lines represent the groups of young adults who will experience non-fatal and fatal household cancer diagnoses, respectively. In panel (b), trends in full time employment rates are depicted in event time, and are limited to the group of young adults who experience non-fatal household cancer diagnosis (cyclamen), or fatal household cancer diagnosis (blue). Panel (c) plots the ratio of full time employment at each period in event time relative to t-1. The black line is the local regression for the non-treated counterfactual. The cyclamen line in panel (c) represents all cancer diagnoses, and the blue line represents only fatal cancer diagnoses. The cyclamen and blue lines are the local linear regression for the respectively-colored points. Panel (d) is the event study difference in differences model, as outlined in the Empirical Specification section. The results come from a stacked event study difference in differences, where stacks are created for each cohort of student, 2015-2020, at each public college in Ohio.

Panel (b) plots the two groups whose household members are affected by cancer in event time, and presents the results without a true counterfactual. Panel (b) does not control for any age or year-related secular trends, and so changes globally are a mix of treatment effect and age effect, which are parsed out ultimately in panel (d). We can see that wage earnings and weeks worked, as well as employment, all increase pre-diagnosis, as expected, but then stagnate temporarily before increasing again. Because labor supply increases globally due to secular age trends, the stagnation may be the result of the treatment effect. This gives suggestive reason to believe that cancer may have an effect on labor supply; however, since there is no counterfactual, this figure relies on assumptions that the reader might make about the outcome based solely on the prior-trends.

Panel (c) begins to introduce a more meaningful visualization. In panel (c), I plot the ratio of weeks worked in event time relative to weeks worked in the academic year prior to treatment. In panel (c), the cyclamen line now represents the full sample of young adults whose household members are affected by cancer, while the blue line still represents just those whose household members are affected by severe cancer. This panel also incorporates elements of matching outlined in the empirical specifications section of Chapter 3. The black line is the local linear approximation of the outcome variable relative to the baseline period for the group of young adults that will not be treated. Intuitively, then, this black line represents the group of young adults at the same college, who began in the same year, and traces their labor supply in comparison to the treated group. Panel (c) presents important intuition both on the relative levels of the colored dots to the black line, but also on the trends pre-diagnosis relative to post-diagnosis.

Panel (c) is also where we can begin to see the effect that household cancer diagnosis might have on young adult labor supply. In Figure 4.3, we can see that in the first three years of diagnosis that labor supply decreases relative to the counterfactual group, for those young adults whose household members are affected with severe household cancer. This is true also in Figures 4.4 and 4.5, for weeks worked and wages earned, respectively. We also see a stronger break in the trends for the treated groups than for the counter-factual group. Both pieces of information give us intuition for the results that are presented in panel (d).

Panel (d) expands on panel (c), and now presents the results of the event study difference in differences. Considering employment first, panel (d) of Figure 4.3 suggests that the effect of cancer diagnosis on employment in young adults in aggregate is a decrease of 2.17 percentage points. This effect disappears in the second year and beyond, suggesting no sustained effect on employment in aggregate. For the young adults whose household members are affected by severe cancer, we see a sustained decrease in employment by about five percentage points in the first three years of diagnosis, relative to what would have otherwise occurred if the household member of the young adult had not been diagnosed.

Panel (d) of Figure 4.4 reports a decrease between one to three weeks in the total number of weeks worked during an academic year in aggregate, relative to the counterfactual group. This effect is larger for the group whose household members are affected by severe cancer, where we see a decrease of about two weeks in the first year and up to four weeks in the fourth year after diagnosis, relative to what the young adult would have otherwise worked.

Finally, Panel (d) of Figure 4.5 reports a decrease of about 30% in wage earnings in the first year both in aggregate and in the group whose household members are affected by severe cancer, relative to the counterfactual. This decrease in log real wage earnings continues over time, too, for both groups.

4.7 Heterogeneity

I explore heterogeneity in the results by four dimensions.

For young adults from disadvantaged households, defined both by household wage earnings (Figure 4.6) and by the flexibility of household labor supply (Figure 4.7), I find decreases in labor supply for the advantaged group but generally not for the disadvantaged group. This is especially true for instances of severe cancer, where the diagnosis ultimately ends in fatality. We see decreases in employment and in weeks worked in an academic year for young adults who do not come from financially disadvantaged households (at least relative to those from the bottom 25% or from households with less than two employed household members), while we see no change in employment or labor supply for those who are disadvantaged.

A similar consideration could be made for young adults with high levels of student loans. As discussed earlier, here young adults are considered disadvantaged if they have high levels of student loans, which may ultimately reflect the middle class who receive less familial support and no government support. Figure 4.8 provides similar results to Figures 4.6 and 4.7, namely that we see decreases in labor supply, relative to the counter factual group, for those who are relatively advantaged (i.e. less strapped by student loan debt). We see no change in labor supply for those who are financially

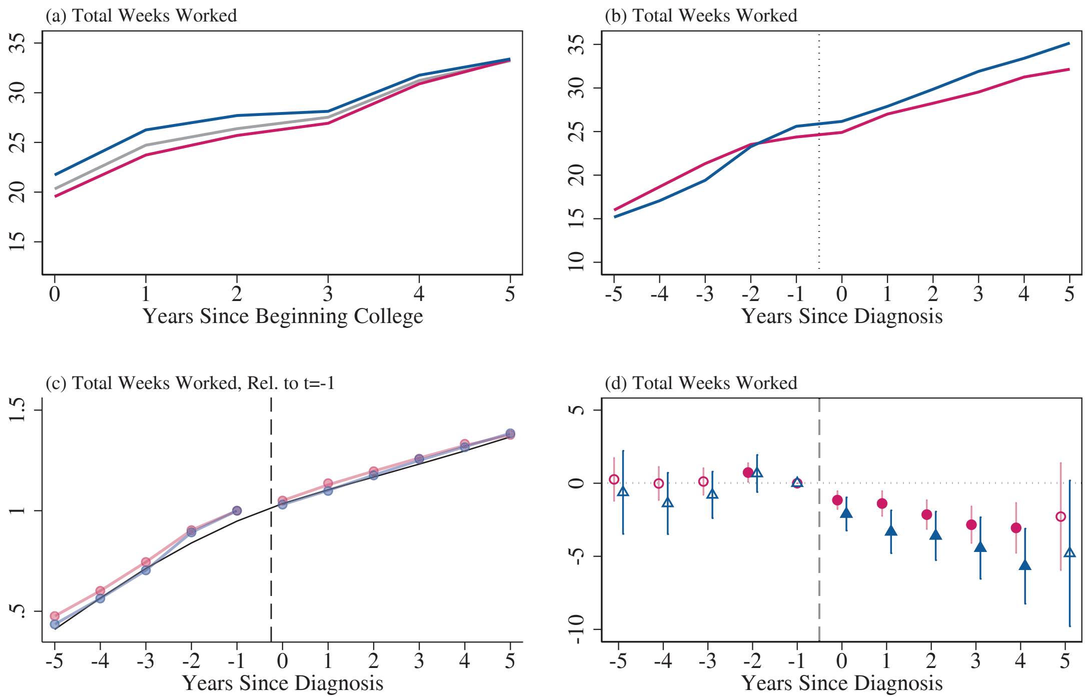


Figure 4.4: Difference in differences results - weeks of labor

Note: This figure presents four insights into understanding the relationship of cancer diagnosis and weeks worked for the young adult. The number of weeks worked per year is the sum of the maximum number of weeks worked in each quarter of the academic year. In panel (a), the weeks worked for three groups of young adults are plotted over time since beginning college. The grey line depicts the group of young adults who will never have a household member that is diagnosed with cancer. The cyclamen and blue lines represent the groups of young adults who will experience non-fatal and fatal household cancer diagnoses, respectively. In panel (b), trends in the weeks worked are depicted in event time, and are limited to the group of young adults who experience non-fatal household cancer diagnosis (cyclamen), or fatal household cancer diagnosis (blue). Panel (c) plots the ratio of the weeks worked at each period in event time relative to t-1. The black line is the local regression for the non-treated counterfactual. The cyclamen line in panel (c) represents all cancer diagnoses, and the blue line represents only fatal cancer diagnoses. The cyclamen and blue lines are the local linear regression for the respectively-colored points. Panel (d) is the event study difference in differences model, as outlined in the Empirical Specification section. The results come from a stacked event study difference in differences, where stacks are created for each cohort of student, 2015-2020, at each public college in Ohio.

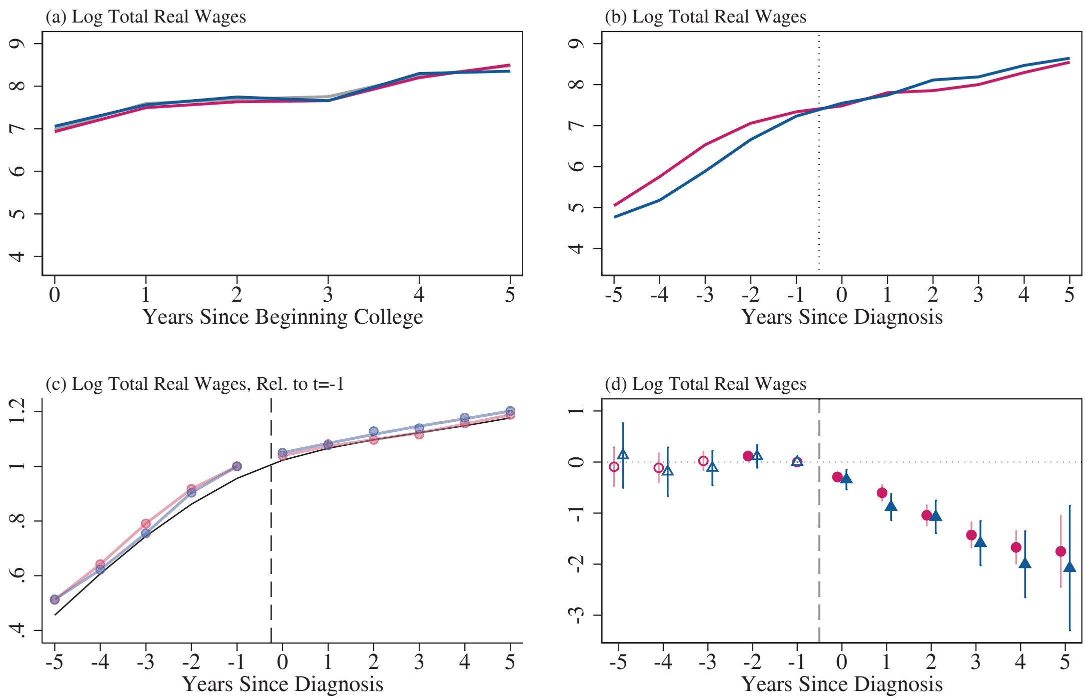
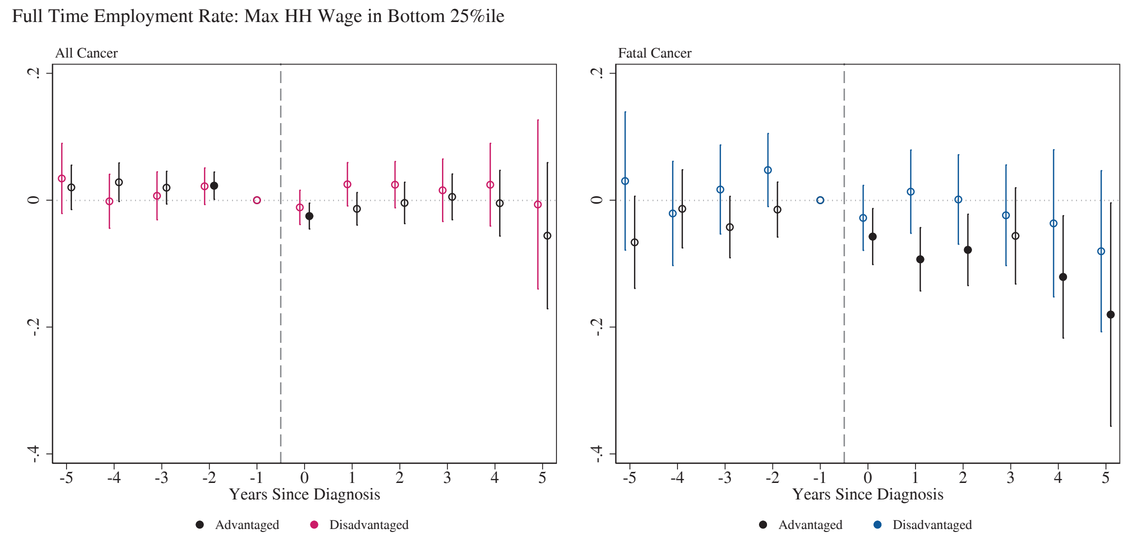
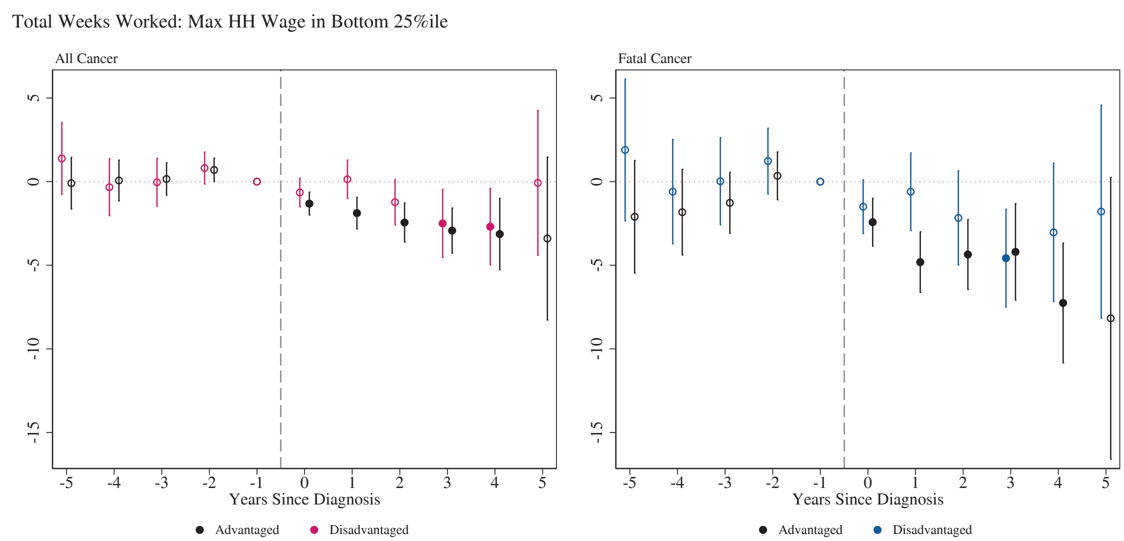


Figure 4.5: Difference in differences results - log total real wage earnings

Note: This figure presents four insights into understanding the relationship of cancer diagnosis and wage earnings for the young adult. Log total real wage earnings is defined as the natural log (plus 1) of the total real wage earnings earned by a young adult in an academic year. In panel (a), the log total real wage earnings for three groups of young adults are plotted over time since beginning college, agnostic to cancer timing. The grey line depicts the group of young adults who will never have a household member that is diagnosed with cancer. The cyclamen and blue lines represent the groups of young adults who will experience non-fatal and fatal household cancer diagnoses, respectively. In panel (b), trends in log total real wage earnings are depicted in event time, and are limited to the group of young adults who experience non-fatal household cancer diagnosis (cyclamen), or fatal household cancer diagnosis (blue). Panel (c) plots the ratio of log total real wage earnings at each period in event time relative to t-1. The black line is the local regression for the non-treated counterfactual. The cyclamen line in panel (c) represents all cancer diagnoses, and the blue line represents only fatal cancer diagnoses. The cyclamen and blue lines are the local linear regression for the respectively-colored points. Panel (d) is the event study difference in differences model, as outlined in the Empirical Specification section. The results come from a stacked event study difference in differences, where stacks are created for each cohort of student, 2015-2020, at each public college in Ohio.



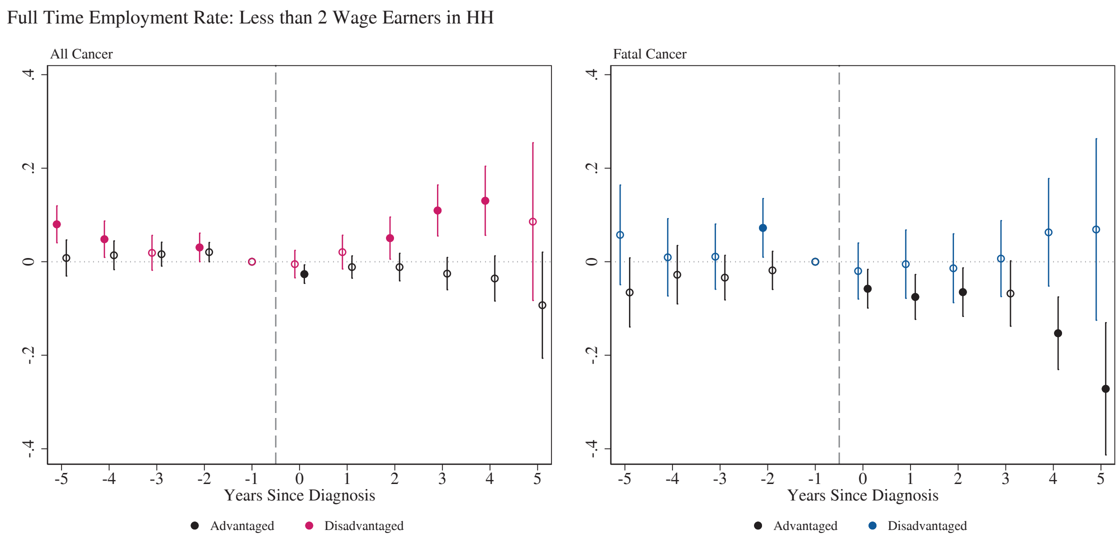
(a) Employment by household wage earnings



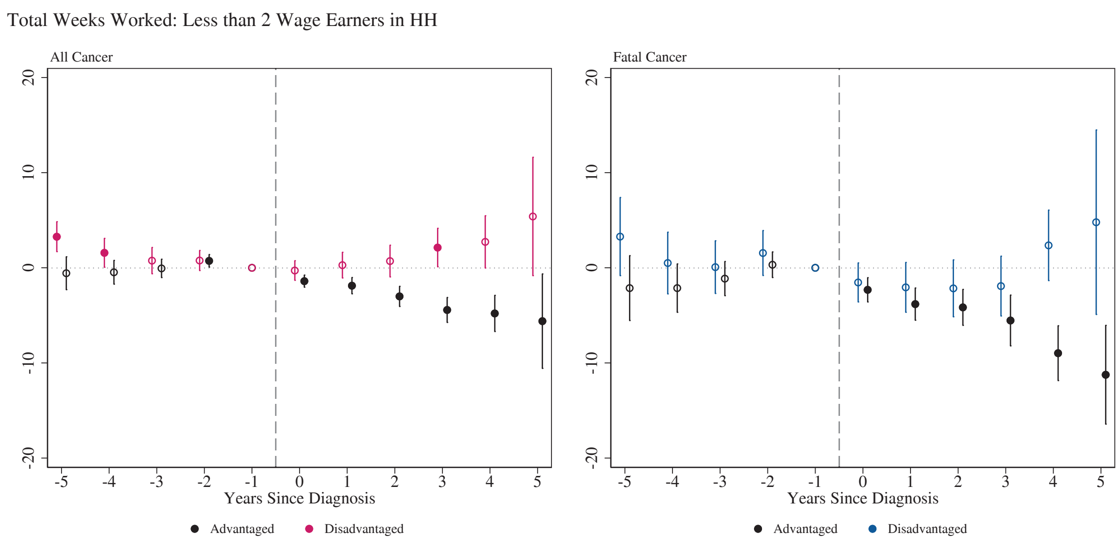
(b) Weeks worked by household wealth

Figure 4.6: Young adult labor supply by household wage earnings

Note: This figure presents heterogeneity in the main results by household wage earnings. A young adult from a disadvantaged household in this figure is defined by one who comes from a household where the maximum household wage earner is in the bottom 25% of wage earnings across the full sample of household members of the analytic sample. The results come from the stacked event study difference in differences model described in the empirical methods section, and includes an interaction for the group. Results may be interpreted in relation to the other group.



(a) Employment by household labor supply



(b) Weeks worked by household labor supply

Figure 4.7: Young adult labor supply by household labor supply

Note: This figure presents heterogeneity in the main results by household labor supply. A young adult from a disadvantaged household in this figure is defined by one who fewer than two workers in his baseline household. This may include himself as the only worker. The results come from the stacked event study difference in differences model described in the empirical methods section, and includes an interaction for the group. Results may be interpreted in relation to the other group.

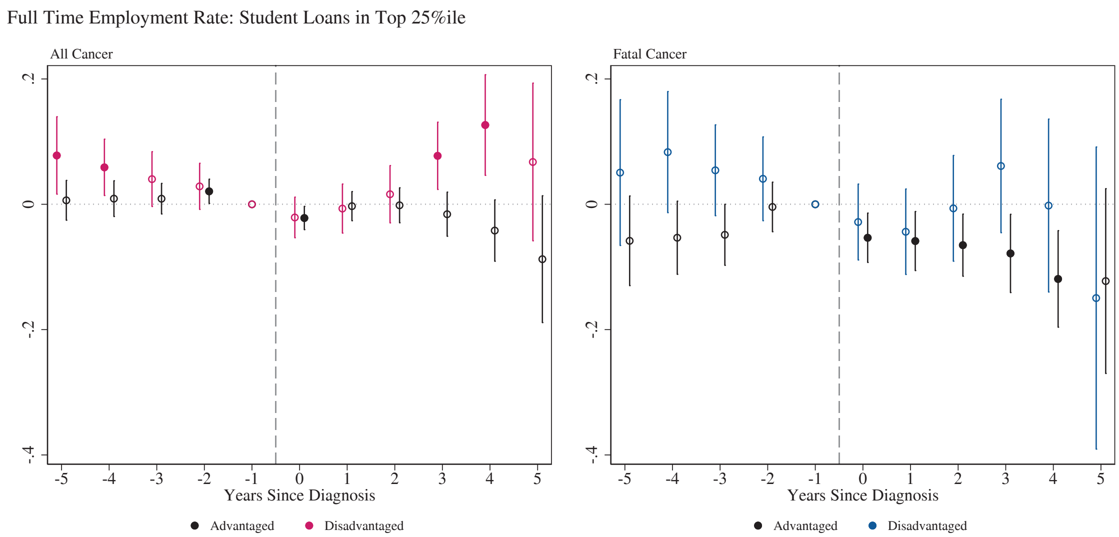
disadvantaged in terms of student loans. Again, this is especially true for young adults from households with severe cancer diagnosis.

The fourth dimension that I explore is geographic. For young adults who attend college far from home, it may be unlikely that they can offer much physical caregiving, and so, as I discussed earlier, it may be unlikely that we see much change in labor supply for this group. By comparison, we might expect that a young adult who lives relatively close to home might be more inclined to offer care to his loved family member, and thus may substitute away from the labor market due to time constraints. Indeed, the results of Figure 4.9 match this interpretation. We see a decrease in the number of weeks worked, both in aggregate and for those with severe household cancer, and no change for those who live far away (160km+).

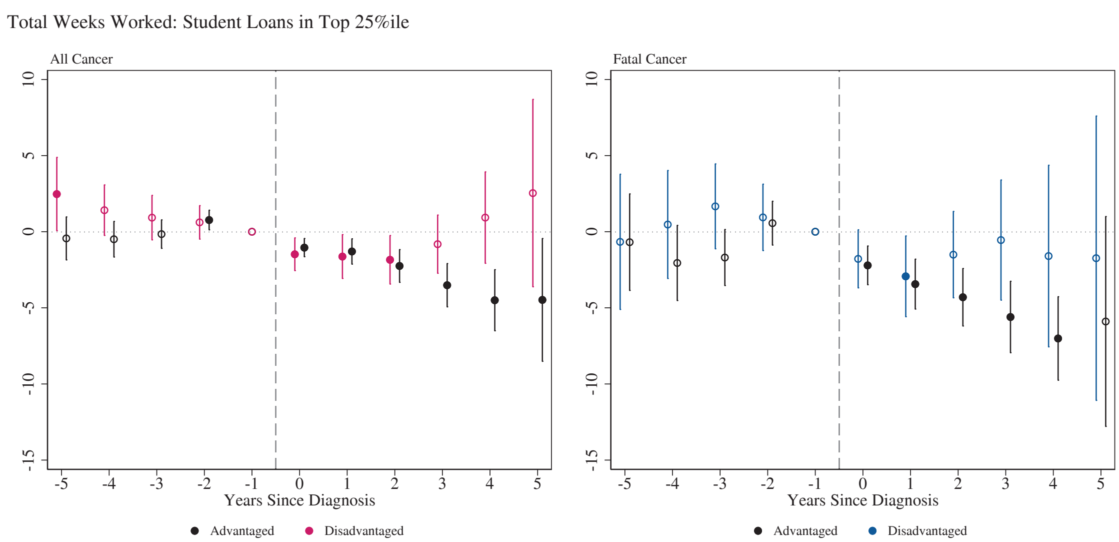
4.8 Conclusion

This chapter explores the labor responses of young adults to household cancer diagnosis. To consider this, I take young adults who began college in public four year colleges in Ohio that were affected by household cancer diagnosis, and consider changes in labor supply of subsequent years relative to peers whose households were not affected. At colleges across the state, for example at Ohio State University, the requirement to live on campus for first year students means that the majority of young adults will move out of their household.[50] While the physical residence of young adults is changing, the overwhelming majority of college students do not consider themselves adults,

[50] Indeed, Ohio State University reports that only about 38% of students of all years are commuter students. https://cssl.osu.edu/posts/632320bc-704d-4eef-8bcb-87c83019f2e9/documents/comparing-on-campus-off-campus-and-commuter-students-accessible.pdf



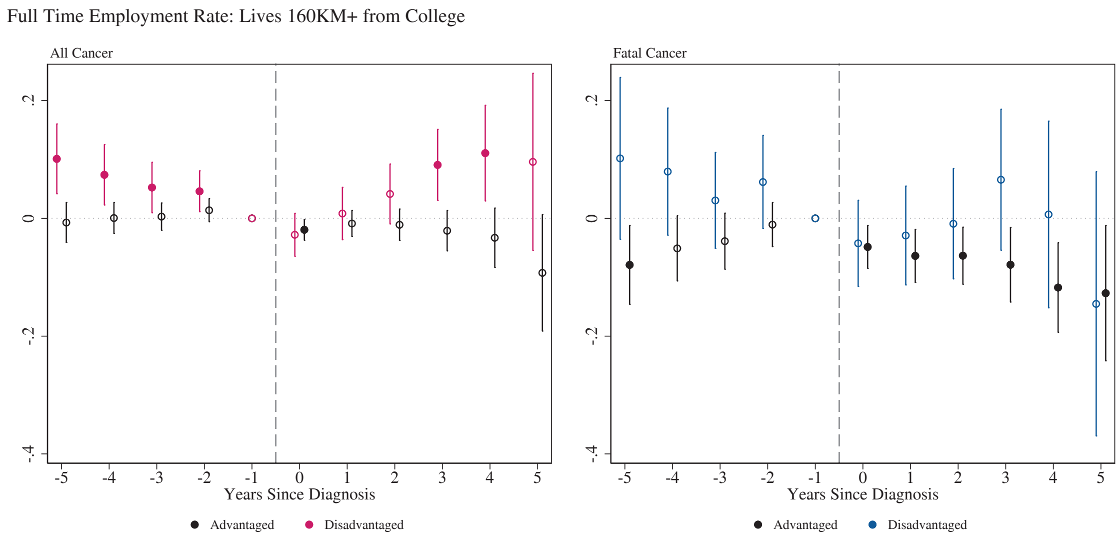
(a) Employment by student loan balance



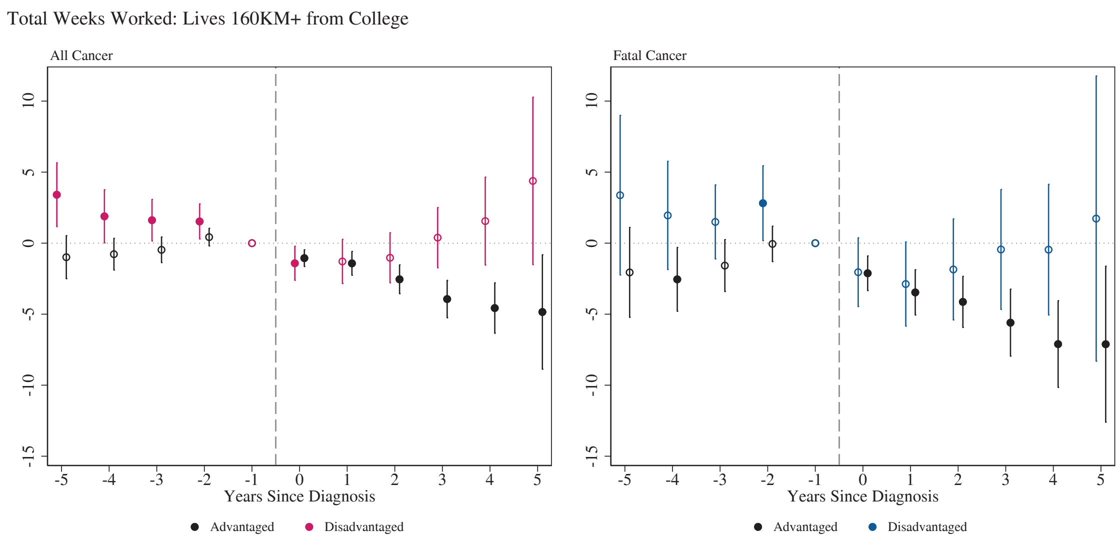
(b) Weeks worked by student loan balance

Figure 4.8: Young adult labor supply by student loan balance

Note: This figure presents heterogeneity in the main results by student loan balance. A young adult from a disadvantaged household in this figure is defined by one who has student loan debt in the top 25% of young adults. The results come from the stacked event study difference in differences model described in the empirical methods section, and includes an interaction for the group. Results may be interpreted in relation to the other group.



(a) Employment by distance from home



(b) Weeks worked by distance from home

Figure 4.9: Young adult labor supply by distance from home

Note: This figure presents heterogeneity in the main results by distance from home. A young adult from a disadvantaged household in this figure is defined by one who comes from a household where the zip code is the 75% of distance (160km) from the campus. Distance is calculated using the Haversine formula. The results come from the stacked event study difference in differences model described in the empirical methods section, and includes an interaction for the group. Results may be interpreted in relation to the other group.

and likely still view themselves as being part of their family's household (Arnett, 1994).

In Gary Becker's Household Production Model, he notes that an individual's utility function is likely comprised of some component of the household's production. A young adult affected by household cancer diagnosis is thus likely to react to the cancer diagnosis of a loved one. Patients with cancer spend four times the amount of total health care expenditures than individuals without cancer Park and Look (2019), and have been shown to experience notable increases in measures of financial distress (Gilligan et al., 2018; Ramsey et al., 2013; Shankaran et al., 2022). Still more, cancer patients are likely to reduce or cease their labor supply in the first years of diagnosis (Oberst et al., 2010; Bradley et al., 2006), which may exacerbate the financial burden that the household already experiences.

A feature of households through an economic perspective is that it insures against risk. One postulated way that households can buffer income shocks is by allocating labor supply collectively rather than individually, and by increasing labor supply amongst other household members when one falls ill (Lundberg, 1985). The demands for caregivers of cancer patients can be immense (Given et al., 2001), and the burdens on caregivers can be as psychologically and financially damaging for caregivers as for the patient himself (Grunfeld et al., 2004; Northouse et al., 2010; Kornblith et al., 1994; Given et al., 1993; Cliff and MacDonagh, 2000; Bishop et al., 2007). In particular, informal caregivers that are employed elsewhere report rates of absenteeism, decreased productivity, and reduced income (Xiang et al., 2022).

For young adults, the financial burden that besets their households may offer incentive to increase labor supply, while the emotional and time costs of the disease may provide an opposing incentive. In this essay, I seek to provide empirical analysis to this conceptual ambiguity. In aggregate I find that, while the employment rates increases across time regardless of whether the household members of a young adult is affected by cancer or not, the increase stagnates for affected young adults. For young adults from households with severe cancer diagnosis (where it is plausible that the caregiving demands are elevated), the decline in employment relative to what it would have been otherwise is sustained. The number of weeks supplied to the labor market per year also decreases both in aggregate and in a more pronounced way for young adults from households with severe cancer.

I also suggested that particular groups of young adults may have household characteristics that constrain their capacities to respond, and hence have more predictable labor force behavior. In particular, I suggested that we would expect to see divergence in the behavior of young adults who have the greatest ability to reduce labor from those who are more advantaged. A hypothetical young adult who engages in the labor force either for extra discretionary income or for work experience seems more likely to adjust his labor supply than a young adult who must work to pay for the necessities of college, like tuition or room and board. I explore heterogeneity in three characteristics of households that may give particular young adults an increased capacity to adjust their labor supply. In particular, I find that across the full sample with cancer, young adults from relatively wealthy advantaged households made no adjustments to their labor, while those from lower income households were actually likely to slightly increase their labor supply. When limiting to young adults from households with severe cancer, where time demands are elevated, we saw that young adults from relatively greater income households were likely to decrease their labor supply, while young adults from disadvantaged households made no changes. This result supports our expectation that elevated time demands may result in decreased labor supply for those who can afford to do so. I noted that labor supply was unperturbed for young adults from single-income households, for young adults where the maximum household earner was in the bottom 25 percentile of earnings, and for young adults who had relatively high levels of student loans. Conversely, young adults from higher income households were more likely to adjust their labor supply downward.

Additionally, the behavior of young adults of varying distances from home is likely to vary, too. Caregiving is exceedingly local. In fact, AARP suggests that as many as 76% of caregivers live within 20 minutes of the patient, and about half of those individuals live in the patient's home. Young adults who attend college at a distance certainly still experience emotional burden, but the capacity to offer care—short of moving—is limited. As a result, I speculated that it was unlikely that young adults living further away would respond in as much of a pronounced way as their local peers. Indeed, I find that young adults who live nearby are more likely to adjust their labor supply downwards, while young adults who attend college further away are no more likely to adjust their labor supply than young adults who are not affected.

This chapter makes at least three contributions. First, it is amongst the first essays to consider how the labor supply decisions of young adults might respond to household illness. While a significant body of work has examined the effect of illness on the labor decisions of spouses, and while a less expansive body of work has examined the effects of parental labor responses to child illness, this is amongst the first essays to consider how the chronic illness of family members may impact young adults. I integrate several economic theories to conceptualize rational responses of young adults in aggregate, which I suggest is ultimately ambiguous, and across subsets of the population that might be particularly affected. Second, in a novel way this essay integrates a number of administrative datasets and econometric techniques to empirically analyze this question. Third, the results of this analysis fit with the conceptual framework that I discuss in the beginning of this chapter. The results from this essay suggest that young adults decrease their labor supply in the wake of household cancer diagnosis, which suggests that the time demands of disease management may outweigh the financial burden of the disease on aggregate. For some groups, this capacity to moderate labor supply is diminished, as the conceptual framework suggests. Finally, this chapter contributes to a broader literature about the inter-generational effects of chronic illness. For young adults from households that are affected by cancer, a decrease in labor supply in the years after diagnosis may suggest difference in long-run productivity and wages. Future work may endeavor to expand on these findings.

Appendix G: Corresponding tables to labor supply event study models, main results, Chapter 4

	Full Time Employment	Total Weeks Worked	Log Total Real Wages	Number of Primary Employers
-5	0.0236	0.263	-0.0931	0.000618
	(1.38)	(0.35)	(-0.48)	(0.14)
-4	0.0214	-0.0179	-0.113	0.00267
	(1.48)	(-0.03)	(-0.77)	(0.76)
-3	0.0168	0.119	0.0225	0.00717*
	(1.35)	(0.25)	(0.24)	(2.17)
-2	0.0227*	0.731*	0.117*	0.00389
	(2.25)	(2.22)	(2.25)	(1.50)
0	-0.0217*	-1.155***	-0.292***	-0.00101
	(-2.29)	(-3.63)	(-6.16)	(-0.43)
1	-0.00415	-1.386**	-0.602***	-0.00424
	(-0.35)	(-3.17)	(-7.60)	(-1.39)
2	0.00285	-2.146***	-1.042***	-0.000522
	(0.21)	(-4.23)	(-10.27)	(-0.13)
3	0.00749	-2.838***	-1.429***	-0.0147**
	(0.46)	(-4.42)	(-11.26)	(-3.13)
4	0.00273	-3.054***	-1.671***	-0.00306
	(0.12)	(-3.49)	(-10.18)	(-0.46)
5	-0.0406	-2.282	-1.749***	-0.00724
	(-0.85)	(-1.22)	(-4.90)	(-0.54)
Obs	6564472	6564472	6656867	6656867
Students	180739	180739	183596	183596

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table G.1: Tabular results of aggregate labor supply models

	Full Time Employment	Total Weeks Worked	Log Total Real Wages	Number of Primary Employers
-5	-0.0312	-0.625	0.130	-0.0272*
	(-0.93)	(-0.43)	(0.40)	(-2.26)
-4	-0.0172	-1.383	-0.189	-0.0140
	(-0.66)	(-1.29)	(-0.78)	(-1.65)
-3	-0.0212	-0.797	-0.115	0.00494
	(-0.99)	(-0.98)	(-0.67)	(0.57)
-2	0.00743	0.665	0.111	-0.000129
	(0.40)	(1.03)	(0.97)	(-0.02)
0	-0.0471**	-2.104***	-0.342***	-0.00318
	(-2.68)	(-3.63)	(-3.48)	(-0.52)
1	-0.0554**	-3.323***	-0.881***	-0.00544
	(-2.63)	(-4.45)	(-6.67)	(-0.81)
2	-0.0506*	-3.602***	-1.074***	0.00131
	(-2.29)	(-4.25)	(-6.57)	(0.14)
3	-0.0461	-4.436***	-1.588***	-0.00749
	(-1.63)	(-4.13)	(-7.14)	(-0.88)
4	-0.0904**	-5.676***	-2.002***	0.0100
	(-2.62)	(-4.33)	(-6.05)	(0.68)
5	-0.136**	-4.799	-2.075***	-0.00997
	(-2.59)	(-1.89)	(-3.33)	(-1.62)
Obs	6529349	6529349	6621295	6621295
Students	174840	174840	177621	177621

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table G.2: Tabular results of severe cancer labor supply models