Cancer Diagnoses and Household Debt Overhang*

Arpit Gupta† Edward R. Morrison‡

Catherine R. Fedorenko and Scott Ramsey§

April 11, 2017

Abstract

This paper explores the effects of unanticipated health shocks on financial outcomes. We draw on data linking individual cancer records to administrative data on personal mortgages, bankruptcies, foreclosures, and credit reports. We present three findings: First, cancer diagnoses are financially destabilizing—as measured by defaults, foreclosure, and bankruptcy filing rates—even among households with public or private health insurance. The instability is caused by out of pocket costs arising from work loss, transportation, and incomplete coverage of medical expenditures. Second, cancer diagnoses are destabilizing only for households that have high levels of debt, preventing them from using their assets to smooth consumption. By contrast, individuals with positive equity extract this equity, and appear to use the funds consistent in a manner which leads to additional treatments and longer longevity. Third, a patient's financial response to a health shock depends on expected mortality. Default and foreclosure are chosen by patients who have received news that they have few years to live; bankruptcy and refinancing are chosen by patients with relatively long expected lifespans. This finding is consistent with the notion that both adverse shocks and strategic behavior explain why households exhibit particular financial outcomes.

JEL classification: I10, I13, G21
Keywords: cancer, bankruptcy, foreclosure

*We are grateful to our discussants Jialan Wang, Crystal Yang, Oren Sussman, and for comments from workshop participants at the University of Amsterdam (Business School), the American Law & Economics Association Annual Meeting, the University of Chicago (Economics and Law School), Oxford University, American College of Bankruptcy, AALS Annual Meeting, Columbia University (Economics, GSB, and Law School), NBER Law & Economics, NYU Stern, Stanford GSB, and the Fred Hutchinson Cancer Research Center for helpful comments. We also thank Equifax, DataQuick, BlackBox Logic, Zillow, the SNR Denton Fund, and the Fred Hutchinson Cancer Research Center for research support. Albert Chang provided excellent research assistance.

†NYU Stern School of Business

‡Columbia Law School

§Fred Hutchinson Cancer Research Center

1 Introduction

This paper explores the effects of unanticipated health shocks on financial outcomes. A long line of scholarship has explored the effects of shocks to health, mortality, and morbidity on consumption and investment decisions.[1] A household's financial response to these shocks—whether to default, experience foreclosure, file for bankruptcy, take on new debt, refinance old debt—is another important margin of adjustment. It is important because a household's financial response helps policymakers evaluate the extent to which households are underinsured ("financially fragile") and the causes of the insurance incompleteness. It is also important because legal institutions may affect the way households respond to health shocks.

We explore the effects of a class of health shocks—cancer diagnoses—that are among the most common and severe shocks in the United States. Cancer is the leading cause of death for women and the second most common cause of death among men. Out-of-pocket costs are substantial, even for those with health insurance. Among Medicare beneficiaries, for example, these costs average $4,727 annually (Davidoff et al. (2013)). Among non-elderly cancer patients, Bernard, Farr and Fang (2011) find that 13% of individuals incurred out of pocket costs exceeding 20% of annual income.

We draw on a comprehensive database of all cancer diagnoses in western Washington State Using and then link link this individual-level database to property transaction records (from DataQuick), mortgage payment histories (from BlackBox), and consumer credit reports (from Equifax). These linkages allow us to study how credit decisions and defaults evolve before and after a cancer diagnosis. Our observation window includes the five years before and after diagnosis. Importantly, the cancer data include information about patients' insurance status and the property transaction records allow us to calculate household leverage. We can therefore test whether a household's response to health shocks varies with its capital structure.

In the context of a standard event-study framework, which censors patient data upon death, we find that cancer diagnoses generate a long-term increase in foreclosure probabilities.

[1] See Oster, Shoulson and Dorsey (2013) for a recent contribution.

Across all cancers, the three-year (cumulative) probability of foreclosure increases from 0.69% to about 0.93% (a 35% increase). The five-year probability of foreclosure increases from 1% to 1.65% (a 65% increase). The foreclosure probabilities are highest for the most advanced cancers ("distant" and "unstaged"), which increase by .0.89% in magnitude, an 89% increase relative to the foreclosure rate during the five years prior to diagnosis.

These effects are driven almost entirely by households with high pre-diagnosis leverage, as measured by home mortgage loan-to-value ratios. Patients with relatively low loan-to-value ratios (measured at origination) do not experience an increase in foreclosure rates (indeed, in some specifications, the rate declines post-diagnosis). The opposite is true for those with loan-to-value ratios exceeding 100%. This suggests that, although cancer is a large financial shock, a household's ability to cope with the resulting financial stress (as measured by foreclosure) depends on its access to home equity. Importantly, these results persist when we focus on patients we believe to be well-insured (through either private or public insurance programs), and they continue for at least the five years subsequent to diagnosis. Consistent with this interpretation, we find that low-leverage patients substantially increase their leverage following a medical shock (by taking on a second mortgage or refinancing an existing one). High-leverage patients are substantially less likely to take on new credit following a shock.

Our empirical analysis contributes to the literature on household financial fragility by highlighting the importance of personal leverage as an important driver of household default decisions. Several related papers examine the financial impact of idiosyncratic health shocks. Hubbard, Skinner and Zeldes (1995) was an early attempt to understand the effect of health shocks on financial outcomes, particularly among the elderly. French and Jones (2004) estimate that 0.1% of households experience a health shock that costs over $125,000 in present value. Our results also echo findings in the household finance literature. We find that a combination of negative shocks and high leverage best explain default patterns, similar to the "double-trigger" theory of mortgage default (see Bhutta, Dokko and Shan (2010). We also highlight the trade-off between risk management and financing current investments in durable goods, such as housing and autos, as analyzed by Rampini and Viswanathan (2016). That trade-off persists even when households carry health insurance.

One interpretation of these findings is that default, equity, and formal medical insurance all serve as different ways in which individuals smooth consumption in the event of an idiosyncratic shock such as a cancer diagnosis. In that light, our work serves as a reminder that housing assets are an important buffer and tool for individuals to manage shocks, as they can both serve as collateral to secure additional financing, or reflect debts which can be discharged in the case of borrowers with negative equity.

However, we also emphasize real consequences of household leverage which has real effects. Consistent with the idea that borrowers with equity use part of the funds in medically relevant ways, we find evidence that borrowers with negative equity are more likely to refuse treatment (in an economically substantial, though not statistically significant way) and have lower longevities than borrowers with positive equity. To try to isolate the role of leverage, we control for a variety of cohort, region, and time effects and find comparable results, suggesting that access to home equity is driving this result. This aspect of the paper relates to a growing literature examining the negative health effects of adverse financial conditions, such as Currie and Tekin (2015), Argys, I.Friedson and Pitts (2016), Pollack and Lynch (2009), and Himmelstein et al. (2009).

This paper also builds on our prior work. Ramsey et al. (2013) find that cancer patients are at higher risk of bankruptcy than those without a cancer diagnosis. Morrison et al. (2013) investigated the causal relationship of car accidents on bankruptcy filings. The latter paper found little evidence that car accidents elevate bankruptcy filings, perhaps because car accidents typically represent smaller shocks than the cancer diagnoses investigated in this paper.

This paper is organized as follows: Section 2 develops the theoretical framework for our empirical analysis. Section 3 describes our data and empirical strategy. Section 4 discusses the implications of our findings and concludes.

2 Theoretical Framework

We are not the first to explore the relationship between health shocks and financial outcomes. Prior work has focused primarily on estimating the elasticity of bankruptcy filing rates (or other financial outcomes) with respect to health shocks of different magnitudes or with respect to policy interventions that expand access to health insurance. Examples include Himmelstein et al. (2005), Gross and Notowidigdo (2009), Ramsey et al. (2013), and Mazumder and Miller (2014). We contribute to this literature by asking two questions that are critical to public policy, but underexplored in prior work:

1. To what extent does household capital structure, independent of health insurance, drive observed financial outcomes, such as foreclosure and bankruptcy?

2. How does a sudden increase in mortality risk—triggered by a cancer diagnosis—affect a household's choice between different legal and economic responses to a health shock?

2.1 Capital Structure

The first question brings a central concern of corporate finance to household finance. Like corporations, a household's response to shocks almost certainly depends on its capital structure, including its ratio of debt to assets (leverage) and the maturity structure of its debt contracts (short versus long term). Capital markets provide an important source of liquidity and consumption-smoothing. As obvious as that proposition seems, we know relatively little about the extent to which access to capital markets matters for the typical household, or how household capital structure affects their responses to shocks. For example, does access to capital markets matter less for households that carry private or public insurance against commonly occurring shocks, such as health problems and auto accidents?

This question is highly relevant to public policy. A large and growing literature on "household financial fragility" has prompted a number of policy proposals, including subsidies to the formation of emergency savings accounts (Lusardi, Schneider and Tufano (2011)), improving household financial literacy (Lusardi and Mitchell (2014)), and expanding insurance coverage for important sources of shocks, such as medical care (Mazumder and Miller (2014)). All of these proposals assume (implicitly) that households do not "leverage up" in response to the reforms. Emergency savings accounts, for example, are useful buffers against shocks only if households have not incurred substantial debt. With high leverage, the household may have effectively (or explicitly) pledged the accounts to creditors.[2] Alternatively, if households already carry debts at high interest rates (for instance on credit cards), a forced savings plan earning a lower interest rate may be welfare-reducing. The implication is that understanding how households manage the credit instruments available to them is essential to understanding how households respond to shocks and to evaluating public policy programs aimed at relieving financial distress.

2.2 Mortality Risk and Financial Management

The second question arises naturally from the vast literature on life-cycle models, which considers the effect of uncertain horizons, health shocks, and mortality risks on investment and consumption (see, e.g., Stoler and Meltzer (2012)). An unanticipated contraction in an individual's time horizon will reduce incentives to invest and increase consumption. Individuals diagnosed with Huntington's Disease, for example, are substantially less likely to invest in education, undertake costly behaviors that reduce other health risks (cancer screening, avoiding smoking), or make other human capital investments, as Oster, Shoulson and Dorsey (2013) show.

A contraction in an individual's time horizon can also affect financial management decisions, such as default, foreclosure, and bankruptcy. Because debt absorbs cash flow available for consumption, a sudden increase in mortality risk can reduce incentives to repay debt. Of course, there are significant costs to default: Creditors can seize assets and the individual's access to capital markets will decline, both of which will be costly if the individual is uncertain about longevity or wants to leave wealth to others (family) after death. This trade-off could, for some individuals, weigh in favor of default, particularly default on a home mortgage. The gains from default can be substantial: Mortgage payments typically consume a large fraction of monthly income, the lender will not pursue foreclosure until the homeowner has missed multiple payments, and the foreclosure process often takes a year to complete. The costs of default can be low, particularly for individuals who have no home equity and whose non-housing wealth is largely protected by state exemption laws. Moreover, many

[2] Indirect evidence of a "leveraging up" phenomenon has been observed in related contexts. Hsu, Matsa and Melzer (2013) find that credit costs decline as unemployment insurance benefits increase.

households view their homes as a combination of investment and consumption good. The mortgage, therefore, is partly funding future investment. When an individual experiences a contraction in time horizon, the incentive to invest declines. By defaulting on the mortgage, the individual can curtail investment and, due to long delays in foreclosure, not reduce consumption of housing services for a substantial period, perhaps more than a year.

These observations imply that the incentive to default and experience foreclosure will be strongest when (a) the individual expects to die within the next few years, (b) default will not put other assets at risk because the individual has no home equity and other assets are shielded by exemption laws, and (c) the individual is either unconcerned about leaving bequests or has already set aside funds for bequests and these funds will be unaffected by default and foreclosure.

Health shocks could have a very different effect on the incentive to file for bankruptcy. A core function of a bankruptcy filing is to discharge debt and either (i) protect future income or (ii) protect assets from creditor collection efforts. The first function is served by a Chapter 7 filing: The filer gives up some assets today in exchange for a discharge of unsecured debts that could be applied against future income. The latter is served by a Chapter 13 filing: The filer agrees to a tax on future income in exchange for a discharge of debts that could be applied against assets in the future. In either case, therefore, a bankruptcy filer uses bankruptcy to conserve future cash flow (or utility) derived from human capital or physical assets. A Chapter 13 filing, for example, is an important device for homeowners to retain their homes when faced with foreclosure, as White and Zhu (2010) show. Chapter 7 is also used to renegotiate with mortgage lenders while discharging unsecured debt (Morrison (2014)).

Seen this way, a bankruptcy filing is analogous an investment decision: An individual renegotiates or discharges debt by exchanging value today (income or assets) for value (income or assets) in the future. Because a contraction in an individual's time horizon will reduce the incentive to invest, it will also reduce the incentive to file for bankruptcy. Similar logic can be applied to refinancing, which is equivalent to renegotiating current debt in order to increase future cash flows. A refinancing is an investment decision, which will be less attractive to individuals with relatively high mortality risk.

A simple model can formalize most of these intuitions. Consider a two-period model of a risk-neutral patient who receives a cancer diagnosis in period 1 and learns that she will survive with probability p to period 2. She incurs medical costs equal to M in period 1 only. Her income in each period is $y < M$. She has one asset, a house, which has market value A and delivers housing services equal to γA per period. The home is subject to a mortgage that has face value D and requires periodic payments equal to δD. Assume, for simplicity, that D is sufficiently large relative to A that the patient cannot borrow additional funds to pay her medical expenses (i.e., she cannot access credit markets to smooth consumption). The discount rate is zero.

Because M exceeds the patient's income y in period 1, she will choose between foreclosure and bankruptcy. If the patient chooses bankruptcy, she must pay costs equal to f. Although she will discharge her medical debt (M), she will continue to service her housing debt (mortgage debts are not dischargeable in bankruptcy unless a homeowner abandons her home). Period 1 consumption will therefore equal income (y) plus housing services (γA) minus debt service (αD): $y + \gamma A - \alpha D$. At the end of period 1, she will survive to the next period with probability p. If she survives, she will receive income y and housing services δA and pay debt service (δD). Because it is the final period, she will also consumer her net wealth, $max[\Delta, 0]$, where $\Delta = A - D$. For convenience, we assume the mortgage is non-recourse. That is, if $A < D$, the lender cannot sue the patient for the difference. Conditional on survival, then, period 2 consumption is $y + \gamma A + max[\Delta, 0]$. Because the discount rate is zero, expected consumption from bankruptcy is:

$$C_B = y + \gamma A - \delta D - f + p(y + \gamma A - \delta D + max[\Delta, 0]) \quad (1)$$

If the patient instead chooses foreclosure in period 1, she will default on her mortgage, not pay her medical expenses, and consume her income and housing services. Total period 1 consumption will therefore be $y + \gamma A$. If she survives to period 2, her home will be liquidated in foreclosure. The net recovery to the patient from foreclosure is $max[\Delta, 0]$. She will lose her home, but her debt will be satisfied. The patient will still owe medical expenses M, which exceed her income. She can therefore file for bankruptcy in period 2. By paying costs f, she will keep her income y and the net value from foreclosure (which I assume is protected by state exemption laws). Her expected consumption from foreclosure is therefore:

$$C_F = y + \gamma A + p(y - f + max[\Delta, 0]) \tag{2}$$

The patient will choose foreclosure if $C_F > C_B$, which will be true when:

$$(1 + p)\delta D > p\gamma A - (1 - p)f \tag{3}$$

The left-hand side of the inequality captures the gains from foreclosure relative to bankruptcy: Foreclosure allows the patient to avoid debt service (δD) in periods 1 and 2. The right-hand side captures the next costs of foreclosure relative to bankruptcy: Foreclosure forces the patient to give up consumption services (γA) in period 2. Under either choice, bankruptcy costs (f) will be incurred, but they occur only probabilistically when the patient submits to foreclosure. Thus, the net costs of foreclosure are reduced by the lower expected costs of bankruptcy.

This inequality captures the idea that foreclosure is more attractive as mortality risk increases: When the patient is certain to die during period 1 ($p = 0$), the inequality is always satisfied. Additionally, foreclosure becomes more attractive as debt (D) increases and as bankruptcy filing costs (f) rise.

This simple model illustrates how mortality risk can affect financial policy. The issue is important to public policy because it points to a strategic element in financial management among individuals who experience health shocks. Because of these shocks, the individuals are financially stressed, but can respond to the stress in various ways. Strategic considerations may explain why some people choose foreclosure while others choose bankruptcy.

3 Data and Empirical Strategy

Cancer represents one of the most common and costly health shocks. Roughly 40% of Americans can expect to face a cancer diagnosis over their lifetimes, and 20% of Americans will die due to cancer-related complications (Society (2013)). Cancer diagnosis rates are projected to increase both internationally and domestically over time due to medical progress in other fields, leaving individuals more susceptible to cancer risk. The cost of treating cancer has also been rising over time even faster then overall healthcare inflation, which in turn has been growing faster than economy-wide prices (See Mariotto et al. (2011) and Trogdon et al. (2012)).

Cancer severity is often measured using "stages." A cancer is "localized" if malignant cells are limited to the organ of origin (e.g., liver). "Regional" and "distant" cancers describe tumors that have extended beyond the organ of origin. A cancer is regional if the primary tumor has grown into other organs of the body; it is distant if the primary tumor has produced new tumors that have begun to grow at new locations in the body. Because of this subtlety, it is well known that the coding of these diagnoses is inconsistent (SEER Training Module 2014); the two categories may describe comparably severe cancers. "Unstaged" cancers are those that were not given a formal staging by the investigating physicians. This often occurs when the cancer has spread so extensively through the patient's body that formal staging is not an informative exercise.

Cancer diagnoses generate direct and indirect costs. Direct cancer costs relate to the cost of treatment and typically represent substantial expenses relative to household income. Cancer treatments typically involve some combination of drugs, surgery, radiation, and hormonal therapy. Formal health insurance should cover many of these treatments, but individuals are also exposed to out-of-pocket costs such as co-pays and deductibles. Prior to 2006, for example, older patients (over 65) often had limited insurance coverage of cancer drugs unless they purchased supplemental Medicare plans (in 2006, this situation changed with the enactment of Medicare Part D). Indirect costs include the time required to undergo screening and therapy, transportation to hospitals and clinics, and child or nursing care. Evidence suggests that 6.5% of cancer expenses among the non-elderly ($1.3 billion) are paid out-of-pocket (Howard, Molinari and Thorpe (2004)). Over 40% of cancer patients stop working after initial treatment (A.G. et al. (2009)).

Costs are substantial even among individuals with public or private insurance. Among Medicare beneficiaries, for example, out-of-pocket costs average $4,727 annually (Davidoff et al. (2013)). Among non-elderly cancer patients, Bernard, Farr and Fang (2011) found that 13% of individuals incurred out-of-pocket costs exceeding 20% of annual income. The percentage is much higher among individuals with public insurance (24% of income) and those with health insurance not provided by their employer (43%).[3]

3.1 Data Construction

We link cancer diagnosis data from Washington State to bankruptcy filings, property records, mortgage payment data, and credit reports. Our cancer data are provided by the Cancer Surveillance System of Western Washington, which collects information about all cancer diagnoses in 11 counties in the western side of the state. These data are a subset of the National Cancer Institute's Surveillance Epidemiology and End Results (SEER) program. Our data include about 270,000 diagnoses occurring during calendar years 1996 through 2009. About 110,000 of these diagnoses involved patients between ages 24 and 64.

The cancer data were linked to a dataset on federal bankruptcy records by the Fred Hutchinson Cancer Research Center via a probabilistic algorithm based on the patient's name, sex, address, and last four Social Security Number digits (see Ramsey et al. (2013)). The bankruptcy records include any individual bankruptcy filing under chapters 7, 11, or 13 of the Bankruptcy Code.

We further link the cancer data to property records maintained by DataQuick to create a "Property Database." The DataQuick records are transaction-based and provide information about every sale, mortgage, foreclosure, or other transaction affecting a property address during calendar years 2000 through 2011. We link these property records to the cancer data based on the patient's property address. This Property Database can be used to study the relationship between cancer diagnoses and foreclosure starts.

We link the Property Database to mortgage payment data and credit reports for patients with privately securitized mortgages. BlackBox LLC provided the mortgage payment data, which includes information about the balance, LTV, borrower FICO, and other characteristics

[3] The substantial nature of indirect costs with respect to cancer also suggests that our work may have some applicability to countries with more universal health coverage, to the extent that formal insurance mechanisms are insufficient to fully prevent financial distress resulting from cancer diagnosis.

of the mortgage at origination as well as the borrower's post-origination payment history. These data cover the period January 2000 through July 2014, and are restricted to the universe of private-label securitized loans. Equifax provided credit reports, which include monthly information about the borrower's credit score, utilization of revolving lines of credit (mainly credit cards), total debt burden, and other characteristics. These data cover the period from June 2005 through July 2014.[4] We linked the Property Database to the BlackBox and Equinox records using mortgage origination date, origination balance, zip code fields, and other mortgage fields (mortgage type and purpose) that are common to all datasets.

After linking these databases (SEER cancer registry, bankruptcy filings, DataQuick property records, and the BlackBox and Equinox databases), we subset on individuals between ages 21 and 80 at the time of diagnosis. Younger patients are unlikely to file for bankruptcy; older patients have extremely high mortality rates subsequent to diagnosis. Additionally, we exclude cancer diagnoses that involving benign and in situ stage cancer diagnoses (early stage cancers that have not spread to surrounding tissue) as well as diagnoses discovered only upon death or autopsy. The former cancers represent trivial health shocks; the latter confound death and diagnosis, making it impossible to infer the impact of diagnosis on financial stability. Finally, a number of patients have multiple cancer diagnoses. If the diagnoses were "synchronous"—occurring within a three month period—we treat them as a single event and assign a diagnosis date equal to the first-diagnosed cancer. Synchronous cancers are frequently manifestations of one underlying cancer. [5] If a patient suffered multiple, non-synchronous cancers (diagnoses occurring over a period longer than three months), we included in our analysis any cancer diagnosis that was not followed by another diagnosis during the subsequent three years. These restrictions explain while the "Full Sample" we use for base analysis contains fewer observations (220k) than our complete dataset (270k). The Deeds Sample, consisting of data which merge between SEER and property records, contains

[4]Equifax performed the linkage between its records and the BlackBox data. Because this linkage was imperfect, we retained a linkage only if Equifax reported a "high merge confidence" (based on a proprietary algorithm) or if the BlackBox and Equifax records listed the same property zip code (suggesting a common residence between the subject of the credit report and the holder of the mortgage. Additional information about the BlackBox and Equinox databases, and the merge algorithm, can be found in Mayer et al. (2014) and Piskorski, Seru and Witkin (2015).

[5]We assign these cancers the highest stage among the multiple stages present (localized, regional, or distant). We also assign the site of the cancer to the "Other" category if the sites of the synchronous cancers differ.

around 64k observations.

Appendix A provides a more complete description of the data and information about the merge algorithms. Figure I provides a visual description of our data creation process.

3.2 Summary Statistics

Table I presents summary statistics for the cancer patients in our study. The first two columns of this table contain information on the Full Sample (core SEER data with restrictions as outlined in the Data Construction section, merged with bankruptcy information only). The second two columns contain information on the subset of the data which merge into Deeds property records. The mean age is 61, with a wide standard deviation: ages 32 through 80 are within two standard deviations of the mean. About sixty percent of patients are married, roughly half are male, and over a third had health insurance through Medicare or Medicaid. Although Table I indicates that only 9.5 percent of individuals carried private insurance (14.7 percent in the Deeds sample), health insurance information is missing for nearly half of the sample. Most of the individuals with missing information likely had some form of health insurance. Those age 65 and older are covered by Medicare. Among those aged 18 to 64, prior studies indicate that between 8 and 14 percent had no health insurance coverage (Ferguson and Gardner (2008)).

Table I also presents information about the "occupation" of individuals in our sample. This information is included in the SEER database and derived from a hospital intake form that asks patients to describe their occupation, not whether they are currently employed in that occupation. We interpret this information as a proxy for the patient's human capital investment. Using an algorithm supplied by Washington State, we categorized patient responses into broad categories: Professional, Clerical, Laborer, Other, Not Employed, and Missing. The Not Employed category includes individuals who indicated that they lacked employment status at the time they completed the intake form.[6]

[6]We classify individuals as "unemployed" if they fail to indicate an occupation, but do indicate marital status. We assume that, if an individual fails to answer both the occupation and marital status questions, he or she is refusing to complete the form. If the individual indicates marital status, but leaves occupation blank, we think it reasonable to assume that the individual is leaving it blank because he or she is unemployed.

Table 2 shows the annual number of cancer diagnoses by stage at diagnosis. As described above, cancer diagnoses can be staged, from least to most severe, as localized, regional, and distant. We include unstaged cancers in the "distant" category because these cancers tend to have a very high mortality rate. Nearly half of diagnoses are localized; regional and distant cancers account for most of the remaining diagnoses.[7]

3.3 Empirical Strategy

We estimate a standard event-study difference-in-difference (DD) regression, following Almond, Hoynes and Schanzenbach (2011) and Autor (2003):

$$O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it} \tag{4}$$

Here, O_{it} is an outcome measure. In most specifications it will be a binary equal to one if patient i exhibits a measure of distress (e.g., foreclosure) during calendar year t. θ_t is a matrix of calendar year fixed effects.[8] The matrix X_{it} includes a variety of controls, which vary with the database used for the analysis. In all regressions, we include patient age, marital status, gender, race, occupation, health insurance status, indicators for whether the patient suffered synchronous cancers or had a previous cancer diagnosis, and county fixed effects. In analysis using the BlackBox or Equifax data, the controls include time from origination, static information taken at time of origination (balance, CLTV, details about the purpose and type of mortgage), and dynamic information updated monthly (such as credit score, estimated income, and interest rate).

The identifying assumption in our model is that, conditional on observables, the *timing* of cancer diagnosis is exogenous. We focus for this reason only on individuals diagnosed with cancer in our sample, and compare individuals diagnosed at different times. Common trends due to time and geographical drivers of financial distress are differenced out in our sample design.

[7] The numbers in this table do not total our sample exactly due to a small number of cancers which lacked staging information.

[8] We do not include individual fixed effects because our dependent variable is binary and we are typically studying the first occurrence of an event (such as foreclosure or bankruptcy). In this setting, with non-repeating events, fixed effect analysis is not feasible (Andress, Golsch and Schmidt (2013).

The coefficients of interest are μ_k, which measure the change in the outcome variable during the s calendar years prior to and following the diagnosis in year T_i, where s is typically 5. Years $[-s, -1]$ reflect the s pre-treatment years, while the interval $[0, s-1]$ is the post-treatment window. These coefficients are measured relative to the (omitted) year prior to the diagnosis. Standard errors are clustered by patient.

If outcome O_{it} occurs during year t, data for that patient is censored in all subsequent years. This censoring renders our framework similar to a discrete time hazard model. Additionally, if patient i dies during year t, data for that patient is also censored in all subsequent years. Finally, the model is only estimated during years for which we are confident that the patient lived in the property in question as determined by sale transactions data.

4 Results

We begin by documenting the average effect of cancer diagnoses on household financial outcomes, including foreclosure, bankruptcy, and missed payments. The overall patterns, however, conceal important heterogeneity with respect to household leverage. Households that have untapped liquidity through home equity or credit cards are better able to withstand cancer diagnoses.

4.1 Average Effects

Figure II plots yearly coefficients from out event-study model using three outcome variables: notice of default, foreclosure, and bankruptcy. The model is estimated using the event-study specification listed in equation 4. The figure plots the coefficients of interest, μ_k, which reported for the five years before and after diagnosis, with the year before as the excluded category. Year zero corresponds to the calendar year of diagnosis. The model is shown separately for stage one cancers, and cancers staged two or higher.

The key coefficients across specifications are generally insignificant prior to the treatment year. The test of the pre-diagnosis coefficients serves as a test of the identifying assumption in our event-study framework: that the cancer diagnosis is indeed an unexpected event for households and not predicted, for instance, by other changes in household variables also driving financial fragility. This might happen, for instance, in the presence of "comorbidities"—other diseases which typically present in conjunction with cancer diagnoses (for instance, emphysema and lung cancer). The existence of comorbidities may drive financial distress independently of the cancer prior to the time of diagnosis. By testing the pre-diagnosis coefficients, we test whether there are pre-existing trends in financial distress prior to diagnosis which might otherwise contaminate our results. Encouragingly, we find little evidence of pre-trends suggestive of financial hardship prior to cancer diagnosis. By contrast, yearly coefficients after diagnosis are frequently positive and significantly different from zero, suggesting that we find evidence of cancer diagnosis on measures of financial defaults.

To provide better quantitative sense of our results, we move to tables. Table III reports the yearly coefficients μ_k for a regression of bankruptcy against time from diagnosis. Coefficients for the controls are suppressed to simplify presentation. At the bottom of the table, we report the cumulative estimated effect for the five years after diagnosis ("Treatment 5 Years"). Again, these estimates are measured relative to the year prior to diagnosis. Additionally, the bottom of the table reports the average bankruptcy rate during the year prior to diagnosis ("Ref. Bankruptcy Prob. 1 Year") and the cumulative probability during the five years prior to diagnosis ("Ref. Bankruptcy Prob. 5 Years").

We find small (and insignificant) effects of cancer diagnoses on bankruptcy filings when measured on the Full Sample. We find larger estimates when we restrict to the Deeds sample (which matches with mortgage records through address), especially among stage one cancers. In column 3, we find that cancer diagnoses lead to a significant cumulative increase in bankruptcy filing of 0.005 percentage points in the five years after diagnosis, which represents about a 24% increase in the relative rate of bankruptcy filing. We find much smaller estimates of bankruptcy filing (a cumulative five year increase of 0.00058) among cancers staged two or higher.

Table IV examines financial defaults as measured by a notices of default (columns 1–2) and foreclosures (columns 3–4). Notices of default correspond to a publicly available statement conveying to borrowers that if they fail to repay money owed, lenders may foreclose on the property. It corresponds therefore to a situation of sizable mortgage delinquency, typically after a borrowers is three or more months behind on payments. More foreclosures in the state of Washington are non-judicial, meaning that they correspond to an event when the lender has seized the property. Defaults capture individual decisions to stop payment; foreclosures capture an additional decision of the lender's actions to seize property in the event of nonpayment of mortgage debt. All of our estimates in this table are measured using the Deeds sample.

In the full Property Database sample, including all cancers, Columns (1) and (2) show a substantial, sustained increase in the probability of foreclosure during the five years following diagnosis. During the five years post-diagnosis, the default rate increases 0.007 percentage points for stage one ("localized") cancers, a 100% increase in the relative frequency of defaults relative to the five year baseline. We find effects of comparable relative magnitude for higher stage cancers (an increase of 0.0081 percentage points relative to a baseline of 0.0091 percent).

Though we observe large effects across all cancer stages, we do find that the timing varies. Among higher stage cancers, we observe an increase in foreclosure rates beginning in the second post-diagnosis year. Among less severe cancers (localized and regional), significant effects appear in the third year following diagnosis. Overall foreclosure rates are large in relative magnitude: representing a relative increase of 156% among stage one cancers, and 96% among higher stage cancers. Note also that all results are censored at mortality.[9]

These findings establish our baseline results: that cancers are financially destabilizing as measured by defaults and foreclosures, and depending on the specification when measured by bankruptcies. One possible reaction to this finding is suggested by Mahoney (2015), which argues for the substitutability of insurance status and bankruptcy. While our measures of insurance status are incomplete (and, in particular, we lack good estimates on truly uninsured people); we can identify subpopulations which we believe are well insured medically: individuals with documented private medical insurance in our data, as well as individuals over 65 (who typically qualify for Medicare). In Table V we restrict on individuals with medical insurance by those criteria.

[9] Results are higher when we do not impose this restriction.

We continue to find quite strong evidence of financial distress induced by cancer diagnoses in this specification. For instance, our estimate of the impact of diagnosis on cumulative five-year effects reflect a 92% increase in the relative probability of experiencing severe mortgage default among stage one cancers. Other estimates are quite similar in magnitude whether or not we condition on insurance status. Because we do not measure uninsurance status well, these numbers cannot be interpreted to suggest that insurance status is unimportant in determining default rates. Rather, we interpret our results to suggest that *even* medically insured individuals appear to respond to cancer diagnoses by defaulting on debts, particularly on their mortgages.

4.2 Financial Fragility and Household Leverage

The analysis thus far conceals important heterogeneity across patients. Cancer diagnoses are destabilizing as measured through mortgage default primarily for households with high levels of pre-diagnosis leverage.

Table VI reexamines the effect of cancer diagnoses on foreclosure, but subsets on patients for whom we can verify the origination date and balance of a mortgage in the Deeds database.[10] Although the sample here is smaller than in Table IV, the estimated effects are comparable. Column 1 restricts on individuals for which a combined loan to value ratio (CLTV) can be measured. CLTV is equal to total mortgage debt, including both first and second mortgages, divided by the purchase price of the home. This restriction establishes a benchmark to verify our results on a sample with mortgage information. Column 1 suggests that default rate increases by .0084 percentage points after diagnosis, a 44 percent increase. The magnitude of the effect here is comparable as in our base table, though the underlying rate of defaults is substantially higher when we subset on individuals with CLTV information.

Columns 3 and 5 of this Table establish that the effects we see on cancer diagnoses driving financial default are driven by highly levered borrowers. Column 3 uses a measure of CLTV taken at origination; Column 5 uses an estimate of the current CLTV (CCLTV) at the time of diagnosis. Cancer is destabilizing only for patients who have no home equity (CLTV $\geq$

[10] We cannot observe the origination date and balance of a mortgage originated prior to around 2000. Our data track transactions after that date.

100) at mortgage origination. Among these patients we observe a very large increase—2 percent—in the foreclosure probability during the five years following diagnosis, over a 200 percent increase relative to the baseline (.01). The foreclosure rate declines among patients with home equity at origination (CLTV<100). Default and bankruptcy rates are also higher among highly levered individuals relative to those with equity. We find comparable, and in some case stronger, results when subsetting on medically insured individuals and cutting across equity in Table VII. In this sample, we also find evidence of statistically significant bankruptcy effects (an increase of 3.4 percent, relative to a five-year average of 4.4 percent) among individuals with high leverage at origination.

These estimates suggest that home equity—and access to liquidity generally—is an important channel through which patients cope with the financial stress of health shocks. This is true regardless of whether the patient carries health insurance. We can study this channel more directly by looking at patients' use of credit following cancer diagnosis. Panel D of Table VI predicts the annual probability that a patient refinances a first mortgage or takes on a second mortgage as the dependent outcome. Although we see a decline in credit use by the average patient during the years following a diagnosis (Column 1), the decline is driven entirely by patients with high levels of leverage (Column 3). We observe comparable patterns when we subset on patients with health insurance, as Table VII shows.

By contrast, we observe a substantial rise in equity extraction among the population with positive equity in their homes. Our effects are quire large, suggesting cancer diagnosis leads to as many as 18% of affected individuals with positive equity to extract some of it.

Together, these results highlight the importance of home equity as a source of insurance. We find that individuals with negative equity respond strongly to the distress induced by a cancer diagnosis to default on their homes, experience an ultimate foreclosure, and in some specifications declare bankruptcy. By contrast, individuals with positive equity tap into the value of equity in their home, possibly to help manage their cancer diagnosis.

To add further robustness to these leverage results, we examine in Figure III how the yearly coefficients of the results change under alternate specifications. Motivated by Struyven (2014) and Bernstein (2016), we examine our baseline leverage results adding additional controls. Under specification 1) Loan Age controls are added; under specification 2) region (zip code) $\times$ cohort controls are added, and under 3) cohort $\times$ time controls are added. The purpose of these controls is to constrain the variation driving current CLTV in different ways. Under specification 1), the role of loan age in mortgage amortization is accounted for. In specification 2), we examine the variation within buyers in a particular area and purchase period, with the variation entirely coming from across time variation in home prices. Under specification 3, we look within cohort and time; and focus on variation across geographical region in home prices. We find the estimates are very comparable across all three specifications. Though these results do not conclusively establish the causal role of leverage in driving our results, they do suggest that variations along the dimensions we are able to control for do not appear be driving the relationship between leverage and financial distress subsequent to cancer diagnosis.

4.3 Credit Bureau Panel Data

The richness of our data allow us to go further in examining the implications of cancer diagnoses on household financial outcomes. In Table VIII, we focus on a sample of loans which merge into BlackBox (a dataset containing close to the full universe of private-label securitized loans), which has also been linked with Equifax credit report data. Panel A of this Table confirms prior findings that cancer patients are more likely to default on their mortgage; in this specification we isolate the impacts on borrowers who have missed three or more payments on their mortgage. Effects are negligible prior to diagnosis, but exceed 2 percentage points for years two and three subsequent to diagnosis.

The other two columns in this panel examine defaults on other debts as measured from credit bureau data. Column 2 examines borrowers' choice to default on their installment accounts (which include student and auto loans). Borrowers are over one percent more likely to default on this type of debt, though not at a statistically significant level. We do find statistical significance when examining defaults on revolving debts, which include credit cards.

Panel B of this table examines other credit outcomes taken from the Equifax data. Column 2 establishes that borrowers take a hit on their credit score after defaulting, with effects around -14 points three years after diagnosis. This effect is unsurprising given previous results that borrowers are delinquent on a variety of debts. Surprisingly, we also find negative effects on owning auto debt, which provides suggestive evidence that patients may avoid auto purchases. Though this effect is statistically insignificant, we can rule out substantially positive auto consumption responses to cancer diagnoses.

Also intriguing are responses on credit limits. We find that borrowing limits on credit cards go up by over $1,000 the calendar year of diagnosis, and continue rising to an economically large $1664 in the third year after diagnosis. Though we lack some statistical precision in these estimates due to the small sample, these estimates are consistent with home equity data in suggesting that cancer patients appear to have strong desire for accessing credit markets. Column 3 of this table suggests that this increase in credit limits is not matched by an increase in card balances, suggesting that this motive may be precautionary, while column 5 indicates that the expansion in credit is coming from the extensive margin: borrowers apply for more revolving accounts (typically credit cards), and are accepted at a rate of an additional 0.5 lines of credit in the year of diagnosis.

4.4 Robustness

To further explore the heterogeneity of responses, we examine in Table IX default responses across occupational status. As discussed under Summary Statistics, we impute occupational status using written responses under the occupational field in our data. We find that professional workers appear relatively well-insured against foreclosures and defaults in our sample, relative to laborers and clerical workers. We also observe statistically significant bankruptcy responses among laborers in our sample.

In Table **??** we also examine differential responses across categories of cancer. In each of these specifications, we compare individuals with, say, Lung cancer against *other* patients who are also diagnosed with lung cancer, but at different times. This allows us to flexibly account for ways in which patients of different cancers have different trends in the background rate of financial default; the identifying assumption throughout is that the timing of the diagnosis, conditional on having a cancer of a certain type, is exogenous. We find relatively stronger default and foreclosure outcomes among patients diagnosed with lung and thyroid cancer, and lower responses among individuals diagnosed with Skin or Colon cancer.

Following our model, which leads us to expect impacts of longevity shocks on default rates, we cut our sample by expected longevity in Table XIV. To calculate that measure, we perform a survival analysis among all cancers with longevity as an outcome variable against a broad range of controls (including age, stage interacted with type of cancer). We use the resulting estimates to divide the sample into two groups: those with greater than average expected survival, and those with less than average survival. We find in Table XIV that foreclosure and bankruptcy tend to be more common outcomes after diagnosis among individuals with low survival rates, though bankruptcy tends to be a more preferred outcome among individuals with relatively high survival rates.

5 Impact of Financial Situation on Health

In this section, we reverse the analysis: instead of asking how cancer diagnoses impact consumer financial decisions, we now investigate whether background household financial situations impact the standard of care or longevity of patients. In the previous sections, we have shown the various tools households have to manage the idiosyncratic shocks induced by a cancer diagnosis, and document how borrowers appear to extract mortgage equity to the extent they can. Here, we emphasize that households are not neutral across these outcomes, that there appear to be real consequences to borrower leverage outcomes on medical status.

Panel A of Table XIV conducts a survival analysis using a Cox hazard regression against leverage. To account for the potentially endogenous assignment of current loan to value, we control (as in the previous Household Leverage section) for 1) loan age, 2) region (zip code) $\times$ cohort, and 3) cohort $\times$ time effects. Across all three specifications, we find that individuals with negative equity exhibit higher mortality rates with a relative hazard of around 17%. Panel B of this Table finds that a possible reason is neglected treatment choice: we find that individuals with negative equity are 0.0084 percentage points more likely to refuse treatments. Though this effect is not statistically significant, it is sizable in magnitude and consistent across specifications controlling for leverage. Taken as a whole, our results provide suggestive evidence that mortality may be worsened among patients who are not as able to access financial markets to borrow and provide for medical care for reasons of negative equity.

6 Conclusion

Our results point to the central importance of credit markets as a buffer against health events and other adverse financial shocks. Even households with health insurance face sizable out-of-pocket costs after a cancer diagnosis. These costs are destabilizing when a household has taken on high pre-diagnosis leverage. The household is effectively priced out of the credit market.

Our research is, however, subject to several caveats. First, we document the patterns of financial distress surrounding severe medical events, but do not make claims about the strategic nature of those defaults. Nor do we make any normative claims about the desirability of foreclosure among affected households. Bankruptcy, default, and foreclosure are commonly viewed as manifestations of severe financial distress, with adverse consequences for debtors and creditors alike. An alternative view might see these outcomes as manifestations of strategical calculations by households. Because a cancer diagnosis reduces a patient's life expectancy, for example, a rational household might strategically default on long-term debts such as mortgages. Under this interpretation, our results on leverage form an analogue to the "double trigger" theory of household default: Default may be the result of both (i) an adverse shock (cancer diagnosis) that reduces ability to pay and (ii) an adverse financial position (negative equity) that limits the household's desire to repay.

Our analysis also leaves unexplored the question of how and why households acquire the capital structures they have. To draw the parallel with corporate finance: We know much more about the overall determinants of corporate leverage decisions than household leverage.

Finally, we look exclusively at the effects of cancer diagnoses on financial management (defaults, foreclosures, bankruptcies). We are unable to test whether cancer diagnoses affect broader wealth and consumption choices.

Our results present a sharp contrast with much of the prevailing literature on household financial fragility and health insurance because we find a limited role of formal insurance in fully preventing financial default. Highly levered individuals face a higher probability of financial default even in the presence of medical insurance. While medical insurance is clearly an important buffer for households facing severe medical shocks, our results show that household financial fragility depends on much more than the existence of such insurance. Many individuals with insurance file for bankruptcy or experience foreclosure (particularly if they are heavily levered); many individuals without insurance never file for bankruptcy or foreclosure (particularly if they have equity). Household capital structure is, at the very least, an additional, important, and underemphasized driver of default decisions among medically distressed households.

Consistent with the idea that real estate assets serve as an important buffer for individuals faced with idiosyncratic shocks, we find that borrowers with positive equity are likely to extract this equity after diagnosis, and appear to be more likely to undergo treatment and live longer as a result. These results provide evidence of the real effects of financial markets on an important tangible household outcome: life expectancy.

A potential implication of our work is that public policy should focus on household asset-building, both by limiting leverage or by raising savings. Unlike efforts to increase medical coverage, efforts to build household assets have the advantage that accumulated savings may be used to deal with any sort of shock, not just medical ones. Laws that limit household leverage, such as restrictions on recourse mortgages, may also help households preserve assets that can fund out-of-pocket costs.

References

A.G., De Boer, T. Taskila, A. Ojajarvi, F.J. van Dijk, and J.H. Verbeek. 2009. "Cancer Survivors and Unemployment: A Meta-Analysis and Meta-Regression" *JAMA*, 301(7): 753–762.

Almond, Douglas, Hilary W Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the war on poverty: The impact of food stamps on birth outcomes" *The Review of Economics and Statistics*, 93(2): 387–403.

Andress, Hans-Jurgen, Katrin Golsch, and Alexander W. Schmidt. 2013. *Applied Panel Data Analysis for Economic and Social Surveys* Springer.

Argys, Laura M., Andrew I.Friedson, and M. Melinda Pitts. 2016. "Killer Debt: The Impact of Debt on MOrtality" *Working Paper.*

Autor, David 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.

Bernard, Didem SM, Stacy L Farr, and Zhengyi Fang. 2011. "National estimates of out-of-pocket health care expenditure burdens among nonelderly adults with cancer: 2001 to 2008" *Journal of clinical oncology*, 29(20): 2821–2826.

Bernstein, Asaf. 2016. "Household Debt Overhang and Labor Supply" *Working Paper.*

Bhutta, Neil, Jane Dokko, and Hui Shan. 2010. "The Depth of Negative Equity and Mortgage Default Decisions" *Forthcoming, Journal of Finance.*

Currie, Janet, and Erdal Tekin. 2015. "Is there a link between foreclosure and health?" *American Economic Journal: Economic Policy*, 7(1): 63–94.

Davidoff, A.J., M. Erten, T.Shaffer, J.S. Shoemaker, I.H. Zuckerman, N. Pandya, M.-H. Tai, X. Ke, and B. Stuart. 2013. "Out-of-pocket health care expenditure burden for Medicare beneficaries with cancer" *Cancer*, 119(6): 1257–1265.

Ferguson, Deron, and Erica Gardner. 2008. "Estimating Health Insurance Coverage Using Hospital Discharge Data and Other Sources" *Health Care Research Brief No. 48.*

French, Jeric, and John B. Jones. 2004. "On the Distribution and Dynamics of Health Care Costs" *Journal of Applied Econometrics*, 19(6): 705–721.

Gross, Tal, and Matthrew J. Notowidigdo. 2009. "Health Insurance and the Consumer Bankruptcy Decision: Evidence from Expansions of Medicaid" *Journal of public Economics*, 7-8(767-778).

Himmelstein, David U, Deborah Thorne, Elizabeth Warren, and Steffie Woolhandler. 2009. "Medical bankruptcy in the United States, 2007: results of a national study" *The American journal of medicine*, 122(8): 741–746.

Himmelstein, David U., Elizabeth Warren, Deborah Thorne, and Steffie Woolhandler. 2005. "Market Watch: Illness and Injury as Contributors to Bankruptcy" *Health Affairs*, 63.

Howard, David H, Noelle-Angelique Molinari, and Kenneth E Thorpe. 2004. "National estimates of medical costs incurred by nonelderly cancer patients" *Cancer*, 100(5): 883–891.

Hsu, Joanne W, David A Matsa, and Brian T Melzer. 2013. "Unemployment insurance and consumer credit" *Working Paper*.

Hubbard, R Glenn, Jonathan Skinner, and Stephen P Zeldes. 1995. "Precautionary saving and social insurance" *Journal of political Economy*, 103(2): 360–399.

Lusardi, Annamaria, and Olivia S Mitchell. 2014. "The economic importance of financial literacy: Theory and evidence" *Journal of Economic Literature*, 52(1): 5–44.

Lusardi, Annamaria, Daniel J Schneider, and Peter Tufano. 2011. "Financially fragile households: Evidence and implications" *Brookings Papers on Economic Activity*, 83–134.

Mahoney, Neale. 2015. "Bankruptcy as Implicit Health Insurance" *American Economic Review*, 105(2): 710–46.

Mariotto, Angela B, K Robin Yabroff, Yongwu Shao, Eric J Feuer, and Martin L Brown. 2011. "Projections of the cost of cancer care in the United States: 2010–2020" *Journal of the National Cancer Institute.*

Mayer, Christopher, Edward Morrison, Tomasz Piskorski, and Arpit Gupta. 2014. "Mortgage Modification and Strategic Behavior: Evidence from a Legal Settlement with Countrywide" *American Economic Review*, 104(9): 2830–57.

Mazumder, Bhashkar, and Sarah Miller. 2014. "The effects of the Massachusetts health reform on financial distress" Working Paper, Federal Reserve Bank of Chicago No. 2014-01.

Morrison, Edward. 2014. "Coasean Bargaining in Consumer Bankruptcy" University of Chicago Law School.

Morrison, Edward R., Arpit Gupta, Lenora Olsen, Larry Cook, and Heather Keenan. 2013. "Health and Financial Fragility: Evidence From Car Crashes and Consumer Bankruptcy" *University of Chicago Coase-Sandor Institute for Law & Economics Research Paper No. 6565.*

Oster, Emily, Ira Shoulson, and E Dorsey. 2013. "Limited life expectancy, human capital and health investments" *The American Economic Review*, 103(5): 1977–2002.

Piskorski, Tomasz, Amit Seru, and James Witkin. 2015. "Asset Quality Misrepresentation by Financial Intermediaries: Evidence from RMBS Market" *Journal of Finance*, 70(6).

Pollack, Craig Evan, and Julia Lynch. 2009. "Health status of people undergoing foreclosure in the Philadelphia region" *American journal of public health*, 99(10): 1833–1839.

Rampini, Adriano A, and S Viswanathan. 2016. "Household risk management" National Bureau of Economic Research.

Ramsey, Scott, David Blough, Anne Kirchhoff, Karma Kreizenbeck, Catherine Fedorenko, Kyle Snell, Polly Newcomb, William Hollingworth, and Karen Overstreet. 2013. "Washington State cancer patients found to be at greater risk for bankruptcy than people without a cancer diagnosis" *Health affairs*, 32(6): 1143–1152.

Society, American Cancer. 2013. "Lifetime Risk of Developing or Dying From Cancer"

Stoler, Avraham, and David Meltzer. 2012. "Mortatlity and Morbidity Risks and Economic Behavior" *Health Economics*, 22(2): 132–143.

Struyven, Daan. 2014. "Housing Lock: Dutch Evidence on the Impact of Negative Home Equity on Household Mobility" *Working Paper*.

Trogdon, Justin G, Florence KL Tangka, Donatus U Ekwueme, Gery P Guy Jr, Isaac Nwaise, and Diane Orenstein. 2012. "State-level projections of cancer-related medical care costs: 2010 to 2020" *The American journal of managed care*, 18(9): 525.

White, Michelle J., and Ning Zhu. 2010. "Saving Your Home in Chapter 13 Bankruprtcy" *Journal of Legal Studies*, 39(1): 33–61.

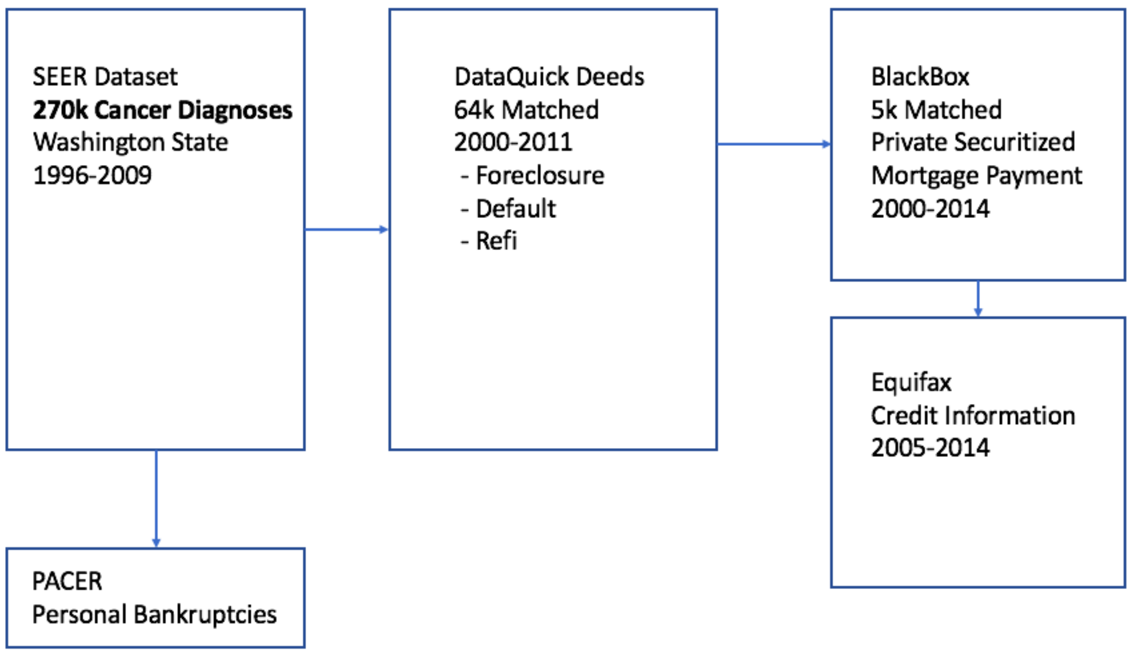
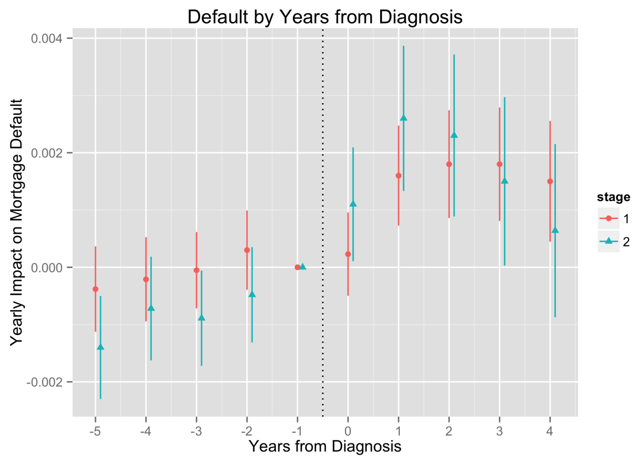
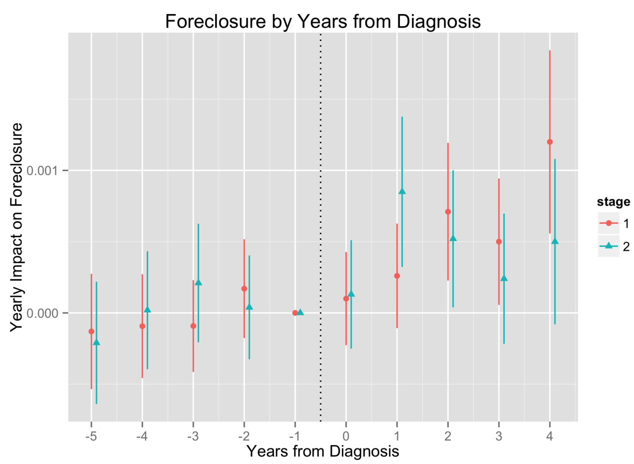


FIGURE I Illustration of Merged Datasets

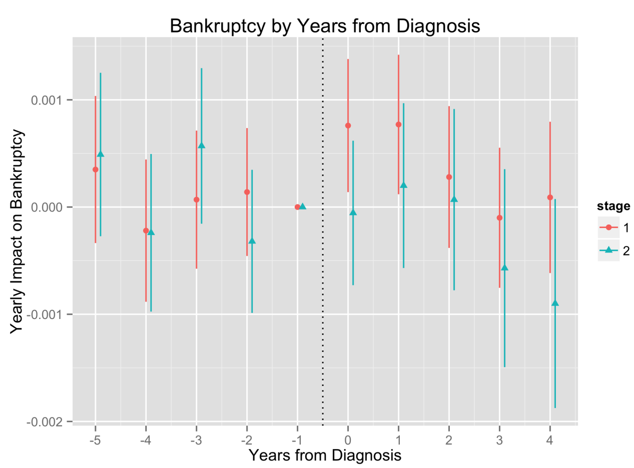
This figure illustrates the connections between the datasets used in this study. The core dataset is the SEER dataset containing diagnosis and treatment information on cancer patients in Western Washington State. This dataset is combined with individual bankruptcy information to produce the Full Sample. This composite dataset is also merged with Deeds data using home address, which provides information on household leverage as well as default and foreclosure information. Deeds data are also linked for some observations to BlackBox and Equifax, which contain information on defaults on private-label mortgages, as well as associated credit bureau information



Panel A: Notice of Default



Panel B: Foreclosure



Panel C: Bankruptcy

FIGURE II Yearly Coefficients from Panel Event Study

These graphs plot the yearly coefficients from the event study regressions as described in the study methodology section.

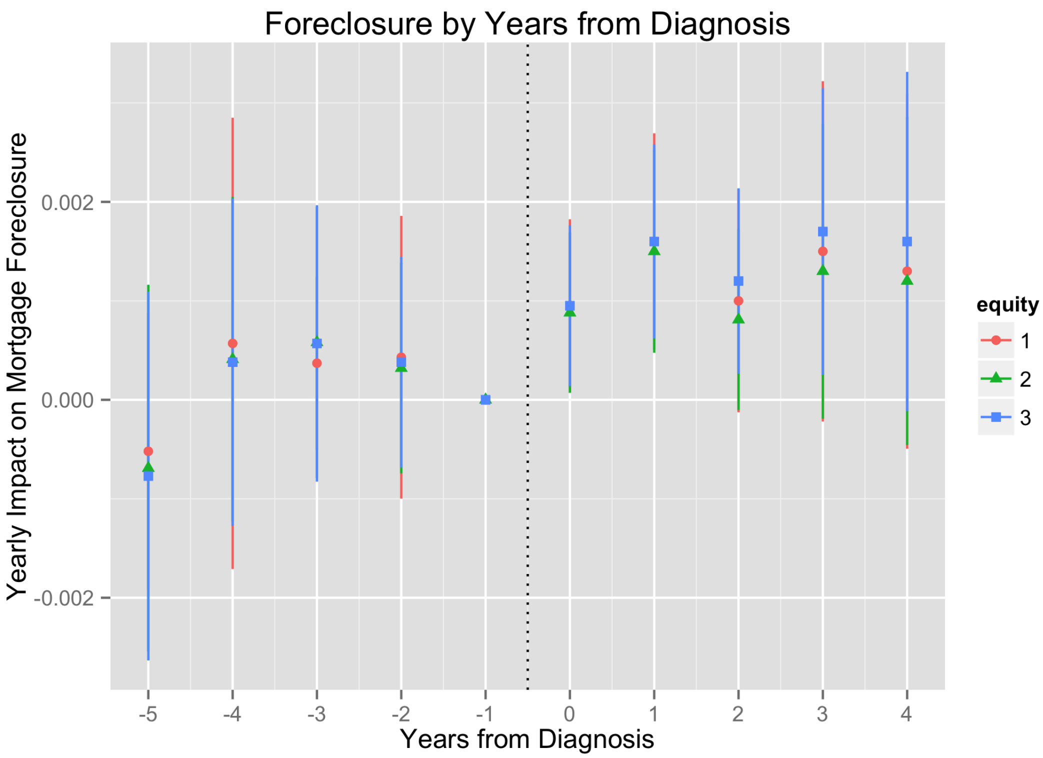


FIGURE III Comparison of Results Across Mortgage Equity Specifications

This figure illustrates yearly coefficients of diagnosis on foreclosure under a variety of specifications which constrain the variation in mortgage equity. Specification one controls in addition for loan age; specification two also controls for region × cohort, specification three also controls for cohort × time.

TABLE I Summary Statistics

This table illustrates sample statistics for our two samples: the Full Sample and the Deeds Sample. The Full Sample contains information from the SEER Cancer dataset matched with bankruptcy information for all patients. The Deeds sample contains information on the subset of the data for which we were able to merge into Deeds records (using address). A full description of the merge process can be found in Appendix A.

	Full Sample		Deeds Sample	
	Mean	SD	Mean	SD
Age	60.926	12.8	58.086	12.8
Married	0.604	0.49	0.650	0.48
Marriage Missing	0.091	0.29	0.096	0.29
Male	0.505	0.50	0.497	0.50
Non-White	0.118	0.32	0.141	0.35
Synchronous Cancer	0.020	0.14	0.019	0.14
Occupation				
- Professional	0.184	0.39	0.211	0.41
- Clerical	0.169	0.37	0.186	0.39
- Laborer	0.256	0.44	0.236	0.42
- Other	0.064	0.25	0.056	0.23
- Not Employed	0.061	0.24	0.065	0.25
Insurance				
- Self-Pay	0.003	0.052	0.003	0.051
- Private Insured	0.095	0.29	0.147	0.35
- Medicare	0.449	0.50	0.341	0.47
- Medicaid	0.012	0.11	0.011	0.10
- Other	0.009	0.093	0.008	0.089
- Missing	0.432	0.50	0.491	0.50
Previous Cancer	0.059	0.24	0.058	0.23
Has Mortgage			0.221	0.41
Origination CLTV			94.127	48.9
Current CLTV			78.263	51.1
Sample Size	220117		64281	

TABLE II Staging Frequency by Year

	Localized	Regional	Distant	Unstaged	Total
1996	1460	600	634	208	2902
1997	1644	660	702	222	3228
1998	1719	666	743	213	3341
1999	1870	757	791	197	3615
2000	2013	832	793	151	3789
2001	2171	991	953	123	4238
2002	2348	1098	1055	87	4588
2003	2464	1137	1086	112	4799
2004	2599	1208	1100	87	4994
2005	2640	1169	1222	113	5144
2006	2784	1135	1209	126	5254
2007	2989	1355	1299	138	5781
2008	3116	1386	1270	92	5864
2009	3269	1394	1336	264	6263
Total	33086	14388	14193	2133	63800
Observations	63800				

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE III Bankruptcy Default Impacts

This table analyzes the impact of cancer diagnoses on bankruptcy filings. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where O_{it} is one if the individual files for bankruptcy in in the calendar year, measured in years from diagnosis. Columns 1 and 3 subset on stage one cancers, columns 2 and 4 subset on cancers staged two and above. Columns 1–2 focus on the whole sample, while columns 3–4 subset on the Deeds sample for which mortgage information is known. The statistic "Treatment 5 Years" captures the linear combination of the treatment effects for five calendar years after the initial diagnosis, inclusive of the year of diagnosis itself. The Reference Probability captures the base rate of foreclosure or default for the year prior to diagnosis (which is excluded in the regression), or the five years prior to establish the baseline. Standard errors are clustered at the patient level.

Dep Var:	Bankruptcy			
	Stage 1	Stage 2+	Stage 1	Stage 2+
Year 5 Before Diagnosis	0.00035	0.00049	-0.00014	-0.00024
	(1.00)	(1.26)	(-0.22)	(-0.34)
Year 4 Before Diagnosis	-0.00022	-0.00024	-0.00060	-0.00099
	(-0.65)	(-0.64)	(-1.00)	(-1.49)
Year 3 Before Diagnosis	0.000069	0.00057	-0.0013*	0.00022
	(0.21)	(1.54)	(-2.51)	(0.34)
Year 2 Before Diagnosis	0.00014	-0.00032	-0.00070	-0.00088
	(0.46)	(-0.94)	(-1.34)	(-1.48)
Year 1 After Diagnosis	0.00076*	-0.000055	0.00088	0.000025
	(2.40)	(-0.16)	(1.57)	(0.04)
Year 2 After Diagnosis	0.00077*	0.00020	0.0014*	0.00088
	(2.32)	(0.51)	(2.24)	(1.21)
Year 3 After Diagnosis	0.00028	0.000069	0.0013*	-0.00038
	(0.83)	(0.16)	(2.06)	(-0.48)
Year 4 After Diagnosis	-0.00010	-0.00057	0.00071	0.000086
	(-0.30)	(-1.21)	(1.09)	(0.10)
Year 5 After Diagnosis	0.000090	-0.00090	0.00069	-0.000036
	(0.25)	(-1.81)	(1.01)	(-0.04)
Sample:	Full Sample		Deeds Sample	
Treatment 5 Years	0.0018	-0.0013	0.0050	0.00058
S.E.	0.0013	0.0015	0.0023	0.0028
Ref. Prob. 1 Year	0.0045	0.0056	0.0046	0.0057
Ref. Prob. 5 Years	0.022	0.027	0.021	0.027
N	857745	747067	264973	221465

Marginal effects; t statistics in parentheses

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$

TABLE IV Financial Defaults on Mortgage Debt

This table analyzes the impact of cancer diagnoses on mortgage outcomes on the Deeds Sample, for which mortgage information is known. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where the outcome in columns 1–2 is notice of default, and foreclosure in Columns 3–4. Columns 1 and 3 subset on stage one cancers, columns 2 and 4 subset on cancers staged two and above. The statistic "Treatment 5 Years" captures the linear combination of the treatment effects for five calendar years after the initial diagnosis, inclusive of the year of diagnosis itself. The Reference Probability captures the base rate of foreclosure or default for the year prior to diagnosis (which is excluded in the regression), or the five years prior to establish the baseline. Standard errors are clustered at the patient level.

Dep Var:	Notice of Default		Foreclosure	
	Stage 1	Stage 2+	Stage 1	Stage 2+
Year 5 Before Diagnosis	-0.00038	-0.0014**	-0.00013	-0.00021
	(-1.00)	(-3.05)	(-0.63)	(-0.96)
Year 4 Before Diagnosis	-0.00021	-0.00072	-0.000093	0.000019
	(-0.56)	(-1.56)	(-0.50)	(0.09)
Year 3 Before Diagnosis	-0.000051	-0.00089*	-0.000092	0.00021
	(-0.15)	(-2.10)	(-0.56)	(0.99)
Year 2 Before Diagnosis	0.00030	-0.00048	0.00017	0.000039
	(0.85)	(-1.13)	(0.96)	(0.21)
Year 1 After Diagnosis	0.00023	0.0011*	0.000100	0.00013
	(0.62)	(2.17)	(0.60)	(0.67)
Year 2 After Diagnosis	0.0016**	0.0026**	0.00026	0.00085**
	(3.59)	(4.02)	(1.39)	(3.16)
Year 3 After Diagnosis	0.0018**	0.0023**	0.00071**	0.00052*
	(3.75)	(3.19)	(2.88)	(2.12)
Year 4 After Diagnosis	0.0018**	0.0015*	0.00050*	0.00024
	(3.56)	(2.00)	(2.21)	(1.03)
Year 5 After Diagnosis	0.0015**	0.00064	0.0012**	0.00050
	(2.79)	(0.83)	(3.66)	(1.69)
Sample:	Deeds Sample			
Treatment 5 Years	0.0070	0.0081	0.0028	0.0022
S.E.	0.0016	0.0022	0.00076	0.00083
Ref. Prob. 1 Year	0.0020	0.0032	0.00040	0.00053
Ref. Prob. 5 Years	0.0070	0.0091	0.0018	0.0023
N	241301	202392	246495	227923

Marginal effects; t statistics in parentheses

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$

TABLE V Financial Default by Cancer Stage, Among Insured

This table follows the constructions of Tables III and IV, but restricts on medically insured individuals.

Panel A: By Mortgage Characteristics

Dep Var:	Notice of Default		Foreclosure	
	Stage 1	Stage 2+	Stage 1	Stage 2+
Year 5 Before Diagnosis	-0.000067	-0.00044	0.00022	-0.0000068
	(-0.11)	(-0.74)	(0.71)	(-0.02)
Year 4 Before Diagnosis	0.000022	0.00054	-0.00016	0.00025
	(0.04)	(0.86)	(-0.76)	(0.93)
Year 3 Before Diagnosis	-0.00011	-0.00029	0.00014	0.00057*
	(-0.21)	(-0.53)	(0.61)	(2.09)
Year 2 Before Diagnosis	0.00013	-0.000040	0.00018	0.00030
	(0.26)	(-0.07)	(0.81)	(1.32)
Year 1 After Diagnosis	0.00039	0.0019**	-0.00016	0.00036
	(0.66)	(2.68)	(-0.83)	(1.33)
Year 2 After Diagnosis	0.0025**	0.0031**	0.00021	0.00084*
	(3.16)	(3.23)	(0.80)	(2.29)
Year 3 After Diagnosis	0.0010	0.0018	0.00063	0.00036
	(1.59)	(1.83)	(1.72)	(1.19)
Year 4 After Diagnosis	0.0018*	0.00047	-0.000065	-0.000070
	(2.28)	(0.51)	(-0.44)	(-0.41)
Year 5 After Diagnosis	0.0014	0.0012	0.00097*	0.000091
	(1.82)	(0.99)	(2.01)	(0.33)
Sample:	Deeds Sample			
Treatment 5 Years	0.0071	0.0086	0.0016	0.0016
S.E.	0.0024	0.0030	0.00097	0.00099
Ref. Prob. 1 Year	0.0020	0.0032	0.00040	0.00053
Ref. Prob. 5 Years	0.0077	0.0090	0.0020	0.0022
N	103672	99832	106436	113320

Panel B: By Bankruptcy Filing

Dep Var:	Bankruptcy		Bankruptcy	
	Stage 1	Stage 2+	Stage 1	Stage 2+
Year 5 Before Diagnosis	0.0013**	0.00077	0.00039	-0.00067
	(3.28)	(1.79)	(0.55)	(-0.79)
Year 4 Before Diagnosis	0.00032	0.00020	0.00020	-0.0013
	(0.84)	(0.48)	(0.28)	(-1.58)
Year 3 Before Diagnosis	0.00026	0.00080*	-0.00070	0.00034
	(0.73)	(2.02)	(-1.23)	(0.44)
Year 2 Before Diagnosis	0.00031	-0.000084	0.00051	-0.0010
	(0.95)	(-0.24)	(0.90)	(-1.56)
Year 1 After Diagnosis	0.00061	0.000030	0.0017**	-0.00031
	(1.85)	(0.08)	(2.60)	(-0.44)
Year 2 After Diagnosis	0.000020	0.00011	0.0016*	0.000091
	(0.06)	(0.27)	(2.28)	(0.11)
Year 3 After Diagnosis	0.00014	-0.00013	0.00079	0.000066
	(0.39)	(-0.28)	(1.15)	(0.06)
Year 4 After Diagnosis	-0.00012	-0.0010*	0.0016*	-0.00088
	(-0.33)	(-2.14)	(2.11)	(-0.89)
Year 5 After Diagnosis	-0.000041	-0.0015**	0.0013	-0.00088
	(-0.10)	(-2.98)	(1.69)	(-0.86)
Sample:	Full Sample		Deeds Sample	
Treatment 5 Years	0.00061	-0.0025	0.0070	-0.0019
S.E.	0.0013	0.0015	0.0023	0.0032
Ref. Prob. 1 Year	0.0045	0.0056	0.0046	0.0057
Ref. Prob. 5 Years	0.015	0.020	0.012	0.021
N	438598	409441	113041	108972

TABLE VI Panel Regression, OLS, By Mortgage Equity

This table analyzes the impact of cancer diagnoses on three measures of financial default cutting by pre-existing mortgage leverage. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t-T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where the outcome is notice of default in Panel A, foreclosure in Panel B, bankruptcy in Panel C, and accessing mortgage credit in Panel D (through a refinancing or adding a second lien). All specifications restrict on the Deeds subsample. The statistic "Treatment 5 Years" captures the linear combination of the treatment effects for five calendar years after the initial diagnosis, inclusive of the year of diagnosis itself. Column one restricts on patients having a measured combine loan-to-value (CLTV). Column two restricts on origination CLTV being less than 100; column 2 captures individuals above 100. Columns 3–4 cut above and below a current CLTV (CCLTV) above or below 80.The Reference Probability captures the base rate of foreclosure or default for five five years prior to establish the baseline. Standard errors are clustered at the patient level.

	Has CLTV	CLTV < 100	CLTV >= 100	CCLTV < 80	CCLTV >= 80
			Panel A: Notice of Default		
Notice of Default 5-Year Effect	0.0084	-0.00024	0.039	0.0012	0.028
S.E.	0.0053	0.0053	0.013	0.0044	0.012
Ref. 5-Year Default Probability	0.019	0.013	0.032	0.012	0.029
N	76614	52618	23996	50919	25643
			Panel B: Foreclosure		
Foreclosure 5-Year Effect	0.0026	-0.0038	0.020	0.0015	0.0045
S.E.	0.0026	0.0028	0.0062	0.0019	0.0063
Ref. 5-Year Foreclosure Probability	0.0077	0.0060	0.011	0.0063	0.010
N	80109	55041	25068	52963	27091
			Panel C: Bankruptcy		
Bankruptcy 5-Year Effect	0.0078	0.0076	0.0099	0.0046	0.0092
S.E.	0.0045	0.0047	0.010	0.0049	0.0087
Ref. 5-Year Bankruptcy Probability	0.035	0.027	0.050	0.026	0.050
N	105437	73345	32092	66107	39261
			Panel D: New Credit		
New Credit 5-Year Effect	-0.050	0.18	-0.42	0.11	-0.48
S.E.	0.028	0.032	0.053	0.034	0.048
Ref. 5-Year New Credit Probability	0.67	0.60	0.82	0.66	0.69
N	77811	53159	24652	51502	26257

TABLE VII Panel Regression, OLS, By Mortgage Equity Among Insured

This table performs an identical specification as in Table VI, but subsets on medically insured individuals.

	Has CLTV	CLTV < 100	CLTV >= 100	CCLTV < 80	CCLTV >= 80
			Panel A: Notice of Default		
Notice of Default 5-Year Effect	0.015	-0.0045	0.077	0.0098	0.046
S.E.	0.0082	0.0071	0.027	0.0059	0.022
Ref. 5-Year Default Probability	0.020	0.013	0.032	0.012	0.034
N	37360	23960	13400	25547	11784
			Panel B: Foreclosure		
Foreclosure 5-Year Effect	0.0038	-0.0024	0.020	0.0058	0.0052
S.E.	0.0033	0.0031	0.0098	0.0031	0.0079
Ref. 5-Year Foreclosure Probability	0.0069	0.0051	0.010	0.0054	0.0092
N	39008	25068	13940	26575	12404
			Panel C: Bankruptcy		
Bankruptcy 5-Year Effect	0.0090	0.00014	0.034	0.0088	0.0099
S.E.	0.0059	0.0058	0.016	0.0061	0.012
Ref. 5-Year Bankruptcy Probability	0.030	0.022	0.044	0.022	0.043
N	49756	32664	17092	31548	18173
			Panel D: New Credit		
New Credit 5-Year Effect	-0.0090	0.17	-0.28	0.093	-0.28
S.E.	0.043	0.050	0.088	0.050	0.078
Ref. 5-Year New Credit Probability	0.67	0.57	0.84	0.65	0.70
N	38013	24204	13809	25890	12094

TABLE VIII Mortgage-Credit Bureau Panel

This Table focuses on outcomes measured in the BlackBox-Equifax panel. This dataset comprises private-label securitized mortgages, and associated credit bureau information from Equifax, that linked with Deeds records. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where time here is measured monthly relative to diagnosis, and effects are combined for the three years before and after diagnosis. Panel A examines outcomes including measures of financial default, where 90 DPD refers to missing three or more payments on a mortgage (taken from BlackBox), Installment Delinquency captures missing two or more payments on installment accounts (including student loans, auto loans, etc.) and Revolving Delinquency captures defaults on revolving lines of credit (such as credit cards and other store cards). Panel B includes as outcomes other fields from the Equifax. “Has Auto” refers to the presence of automobile-related debt (as a proxy for car ownership), Credit Score refers to the Vantage Score, Card Balance is the cumulative total of all credit card debt, and Credit Limit combines the available credit on all lines of credit cards.

Panel A: Measures of Financial Default

	90 DPD+	Installment Delinquency	Revolving Deliquency
Year -3	0.0025	0.019	0.016
	(0.51)	(1.58)	(1.09)
Year -2	-0.00014	0.014	0.015
	(-0.04)	(1.63)	(1.49)
Year +1	0.0062	-0.0092	0.012
	(1.27)	(-1.14)	(1.26)
Year +2	0.024**	0.010	0.020
	(3.67)	(1.05)	(1.90)
Year +3	0.020**	0.013	0.025*
	(2.87)	(1.26)	(2.15)
N	1339760		

Panel B: Other Measures from Credit Bureau Data

	Has Auto	Credit Score	Card Balance	Credit Limit	# Revolving Accts
Year -3	-0.0023	-3.07	400.8	209.3	-0.073
	(-0.18)	(-0.98)	(0.73)	(0.08)	(-0.18)
Year -2	-0.0099	0.76	-209.7	189.3	0.14
	(-1.12)	(0.37)	(-0.68)	(0.11)	(0.53)
Year +1	-0.0069	-3.01	152.6	1149.1	0.54
	(-0.89)	(-1.69)	(0.55)	(0.72)	(2.62)
Year +2	-0.016	-11.7**	10.0	1497.0	0.53
	(-1.56)	(-4.36)	(0.03)	(0.73)	(1.93)
Year +3	-0.0099	-13.9**	388.4	1663.6	0.15
	(-0.84)	(-4.55)	(0.98)	(0.71)	(0.50)
N	1339760				

* p¡0.05, ** p¡0.01

TABLE IX Panel Regression, OLS, By Occupation, Aged 26–60

This table analyzes the impact of cancer diagnoses on three measures of financial default cutting by stated Occupation among individuals aged 26–60. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where the outcome is notice of default in Panel A, foreclosure in Panel B, and bankruptcy in Panel C. Panels A and B subset on the Deeds subsample; Panel C uses the Full sample. Occupation status was computed using written occupations matched with coding derived from the Washington State government (details in Appendix A). Non-employment was imputed for individuals without a written response. The Reference Probability captures the base rate of foreclosure or default for five five years prior to establish the baseline. Standard errors are clustered at the patient level.

	Professional	Clerical	Laborer	Non-employed	Other
		Panel A: Notice of Default			
Notice of Default 5-Year Effect	0.0057	0.015	0.014	-0.0047	-0.00087
S.E.	0.0030	0.0042	0.0051	0.0082	0.0048
Ref. 5-Year Default Probability	0.0017	0.0036	0.0042	0.0049	0.0042
N	61903	53369	56450	16963	57737
			Panel B: Foreclosure		
Foreclosure 5-Year Effect	-0.0013	0.0054	0.0065	0.00060	0.0038
S.E.	0.0021	0.0018	0.0020	0.0044	0.0018
Ref. 5-Year Foreclosure Probability	0.00087	0.00078	0.00088	0.0017	0.00094
N	64056	55897	60623	17236	60659
			Panel C: Bankruptcy		
Bankruptcy 5-Year Effect	0.0050	-0.00089	0.0080	0.0081	-0.00086
S.E.	0.0031	0.0043	0.0040	0.0055	0.0036
Ref. 5-Year Bankruptcy Probability	0.022	0.037	0.045	0.032	0.035
N	160084	147277	183786	47541	183381

TABLE X Panel Regression, OLS, By Cancer Site

This table analyzes the impact of cancer diagnoses on three measures of financial default cutting by the category of cancer. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where the outcome is notice of default in Panel A, foreclosure in Panel B, and bankruptcy in Panel C. All specifications restrict on the Deeds subsample. The statistic "Treatment 5 Years" captures the linear combination of the treatment effects for five calendar years after the initial diagnosis, inclusive of the year of diagnosis itself. The Reference Probability captures the base rate of foreclosure or default for five five years prior to establish the baseline. Standard errors are clustered at the patient level.

	Breast	Colon	Lymphoma/Leukemia	Lung	Prostate	Skin	Thyroid	Uterine	Other
				Panel A: Notice of Default					
Notice of Default 5-Year Effect	0.0060	0.0012	0.0061	0.014	0.0041	-0.00012	0.016	0.0055	0.013
S.E.	0.0030	0.0051	0.0044	0.0050	0.0022	0.0046	0.0074	0.0058	0.0033
Ref. 5-Year Default Probability	0.0087	0.0076	0.0078	0.0098	0.0051	0.0079	0.010	0.0099	0.0082
N	87199	35732	42101	42029	77442	29926	14262	13157	101845
				Panel B: Foreclosure					
Foreclosure 5-Year Effect	0.0018	0.0022	0.0031	0.0029	0.00030	0.0023	0.0034	0.0025	0.0044
S.E.	0.0014	0.0024	0.0019	0.0013	0.0014	0.0021	0.0034	0.0020	0.0011
Ref. 5-Year Foreclosure Probability	0.0019	0.0025	0.0021	0.0023	0.0012	0.0029	0.0019	0.0016	0.0022
N	88829	38015	45150	51876	78442	30415	14414	13584	113693
				Panel C: Bankruptcy					
Bankruptcy 5-Year Effect	0.0086	-0.0039	0.0061	0.00040	0.00034	-0.0059	0.030	0.015	-0.0014
S.E.	0.0038	0.0062	0.0060	0.0066	0.0037	0.0066	0.012	0.0093	0.0042
Ref. 5-Year Bankruptcy Probability	0.022	0.022	0.026	0.028	0.016	0.023	0.024	0.026	0.028
N	95583	38894	46372	45164	83662	33569	16292	14368	112534

TABLE XI Panel Regression, OLS, By Survival Status

This table analyzes the impact of cancer diagnoses on three measures of financial default divided by duration of survival. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where the outcome is notice of default in Panel A, foreclosure in Panel B, and bankruptcy in Panel C. Survival duration is predicted using a survival analysis using all covariates (including age, cancer type, and stage) as well as an interaction of cancer type with cancer stage. The sample is divided in half into "High Survival" and "Low Survival" subpopulations.The statistic "Treatment 5 Years" captures the linear combination of the treatment effects for five calendar years after the initial diagnosis, inclusive of the year of diagnosis itself. The Reference Probability captures the base rate of foreclosure or default for five five years prior to establish the baseline. Standard errors are clustered at the patient level.

	Full Sample		Aged 26–60		CLTV > 100	
	High Survival	Low Survival	High Survival	Low Survival	High Survival	Low Survival
			Panel A: Notice of Default			
Notice of Default 5-Year Effect	0.0057	0.010	0.0052	0.013	0.032	0.085
S.E.	0.0016	0.0023	0.0023	0.0047	0.017	0.025
Ref. 5-Year Default Probability	0.0074	0.0087	0.0084	0.012	0.030	0.034
N	264181	179433	176216	70158	14502	9475
			Panel B: Foreclosure			
Foreclosure 5-Year Effec	0.0022	0.0031	0.0019	0.0064	0.020	0.033
S.E.	0.00078	0.00082	0.0012	0.0017	0.0076	0.011
Ref. 5-Year Foreclosure Probability	0.0019	0.0022	0.0023	0.0028	0.011	0.012
N	268339	205997	179073	79347	14744	10305
			Panel C: Bankruptcy			
Bankruptcy 5-Year Effect	0.0046	0.000083	0.0058	-0.0014	0.029	-0.011
S.E.	0.0023	0.0028	0.0033	0.0062	0.014	0.018
Ref. 5-Year Bankruptcy	0.022	0.026	0.027	0.039	0.043	0.061
N	291260	195083	196038	77446	19475	12592

TABLE XII Panel Regression, OLS, among Married by Gender

This table analyzes the impact of cancer diagnoses on three measures of financial default; restricting on married individuals and cutting by gender. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + X_{it} + \theta_t + \varepsilon_{it}$, where the outcome is foreclosure in columns 1–2, notice of default in columns 3–4, and bankruptcy in columns 4–5. Columns 1–4 restrict to the Deeds subsample; Columns 5–6 include all individuals. All columns restrict on married individuals, and alternate between male and female patients. The statistic "Treatment 5 Years" captures the linear combination of the treatment effects for five calendar years after the initial diagnosis, inclusive of the year of diagnosis itself. The Reference Probability captures the base rate of foreclosure or default for five five years prior to establish the baseline. Standard errors are clustered at the patient level.

Dep Var:	Foreclosure		Notice of Default		Bankruptcy	
	Male	Female	Male	Female	Male	Female
Year 5 Before Diagnosis	-0.000041	-0.00020	-0.00028	-0.0011*	0.00064	0.000029
	(-0.22)	(-0.69)	(-0.63)	(-2.09)	(1.51)	(0.06)
Year 4 Before Diagnosis	0.00019	-0.000070	-0.00031	-0.00061	-0.00011	-0.00059
	(0.99)	(-0.24)	(-0.77)	(-1.17)	(-0.27)	(-1.34)
Year 3 Before Diagnosis	0.00013	0.00023	-0.00023	-0.00051	0.00032	0.00011
	(0.78)	(0.82)	(-0.58)	(-1.05)	(0.84)	(0.24)
Year 2 Before Diagnosis	0.000080	-0.000073	-0.000023	-0.00026	-0.00021	0.00065
	(0.54)	(-0.33)	(-0.06)	(-0.54)	(-0.59)	(1.52)
Year 1 After Diagnosis	0.00037	-0.00014	0.00093*	0.00023	0.00099**	0.00053
	(1.83)	(-0.67)	(2.01)	(0.45)	(2.63)	(1.28)
Year 2 After Diagnosis	0.00063*	0.00030	0.0015**	0.0017*	0.00047	0.0016***
	(2.53)	(1.11)	(2.80)	(2.57)	(1.19)	(3.36)
Year 3 After Diagnosis	0.00022	0.0011**	0.0011	0.0021**	0.00040	0.00076
	(1.14)	(2.99)	(1.95)	(3.01)	(0.93)	(1.55)
Year 4 After Diagnosis	0.00014	0.00060	0.0022**	0.0015*	0.00012	0.00028
	(0.69)	(1.96)	(3.14)	(2.14)	(0.27)	(0.55)
Year 5 After Diagnosis	0.00087*	0.00069	0.00096	0.0026**	0.00025	0.00044
	(2.39)	(1.85)	(1.46)	(2.95)	(0.54)	(0.83)
5-Year Treatment Effect	0.0022	0.0025	0.0066	0.0080	0.0022	0.0036
S.E.	0.00078	0.00098	0.0019	0.0023	0.0015	0.0017
Ref. Treatment Effect	0.00036	0.00072	0.0018	0.0024	0.021	0.021
N	169642	142917	157664	135702	543232	439525

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE XIII Panel Regression, OLS, Choice of Bankruptcy Chapter

This table analyzes the impact of cancer diagnoses on bankruptcy choices, divided out by chapter.

three measures of financial default cutting by the category of cancer. The specification is the standard event-study diff-in-diff: $O_{it} = \alpha + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t-T_i)=k] + X_{it} + \theta_t + \varepsilon_{it}$, where the outcome is notice of default in Panel A, foreclosure in Panel B, and bankruptcy in Panel C. All specifications restrict on the Deeds subsample. The statistic "Treatment 5 Years" captures the linear combination of the treatment effects for five calendar years after the initial diagnosis, inclusive of the year of diagnosis itself. The Reference Probability captures the base rate of foreclosure or default for five five years prior to establish the baseline. Standard errors are clustered at the patient level.

Sample:	Full Sample		Deeds Sample		Low Survival		High Survival	
Dep. Var:	Ch. 7	Ch. 13	Ch. 7	Ch. 13	Ch. 7	Ch. 13	Ch. 7	Ch. 13
Year 5 Before Diagnosis	0.00042	0.000026	-0.00011	-0.000071	0.00047	-0.000020	0.00037	0.000067
	(1.64)	(0.23)	(-0.26)	(-0.33)	(1.24)	(-0.12)	(1.05)	(0.42)
Year 4 Before Diagnosis	-0.00019	-0.000100	-0.00061	-0.00041*	-0.00028	-0.00026	-0.00011	0.000053
	(-0.76)	(-0.93)	(-1.43)	(-2.08)	(-0.77)	(-1.73)	(-0.33)	(0.35)
Year 3 Before Diagnosis	0.00038	-0.00010	-0.00049	-0.000020	0.00045	-0.00019	0.00030	-0.000030
	(1.59)	(-1.01)	(-1.21)	(-0.10)	(1.28)	(-1.27)	(0.92)	(-0.20)
Year 2 Before Diagnosis	-0.000030	-0.00014	-0.00066	-0.00022	-0.00042	-0.00036**	0.00034	0.000064
	(-0.13)	(-1.36)	(-1.73)	(-1.15)	(-1.29)	(-2.58)	(1.06)	(0.45)
Year 1 After Diagnosis	0.00038	-0.000071	0.00042	-0.00012	0.00000093	-0.00025	0.00077*	0.000099
	(1.70)	(-0.72)	(1.04)	(-0.63)	(0.00)	(-1.76)	(2.43)	(0.71)
Year 2 After Diagnosis	0.00055*	-0.00015	0.0011*	-0.000050	-0.00019	-0.00029	0.0011***	-0.000033
	(2.23)	(-1.45)	(2.37)	(-0.24)	(-0.53)	(-1.89)	(3.39)	(-0.23)
Year 3 After Diagnosis	0.00032	-0.00014	0.00059	-0.00011	-0.00038	-0.00038*	0.00079*	0.000011
	(1.21)	(-1.25)	(1.24)	(-0.49)	(-0.92)	(-2.22)	(2.27)	(0.08)
Year 4 After Diagnosis	-0.00015	-0.00027*	0.00036	-0.00019	-0.0011**	-0.00048**	0.00041	-0.00014
	(-0.56)	(-2.31)	(0.73)	(-0.79)	(-2.62)	(-2.73)	(1.17)	(-0.92)
Year 5 After Diagnosis	-0.00018	0.000017	0.00011	0.00020	-0.0012**	-0.00023	0.00038	0.00015
	(-0.63)	(0.13)	(0.23)	(0.73)	(-2.58)	(-1.07)	(1.03)	(0.88)
5-Year Treatment Effect	0.00093	-0.00062	0.0026	-0.00026	-0.0029	-0.0016	0.0035	0.000089
S.E.	0.00093	0.00041	0.0017	0.00083	0.0014	0.00062	0.0013	0.00056
Ref. Filing Rate	0.024	0.0043	0.023	0.0052	0.025	0.0041	0.024	0.0045
N	1604812	1604812	486438	486438	728182	728182	876435	876435

Marginal effects; t statistics in parentheses

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE XIV Impact of Leverage on Treatment and Survival

This Table examines how financial leverage impacts the progression of cancer diagnoses. Panel A runs a survival regression where the dependent variable is survival, other controls are included, and the key variables are current CLTV at the time of diagnosis. Additional controls constrain the variation in home equity. Specification one controls in addition for loan age; specification two also controls for region $\times$ cohort, specification three also controls for cohort $\times$ time. Panel B runs an OLS regression with the same independent variables, but examining as an outcome variable the decision to refuse treatment.

Panel A: Hazard Rate of Leverage on Default

Current CLTV $\leq$ 60		Excluded	
60 $<$ Current CLTV $\leq$ 80	0.073	0.069	0.072
	(1.52)	(1.23)	(1.49)
80 $<$ Current CLTV $\leq$ 100	0.100	0.12	0.10
	(1.80)	(1.81)	(1.88)
100 $<$ Current CLTV	0.17**	0.15*	0.18**
	(3.01)	(2.36)	(3.10)
Other Controls	Yes	Yes	Yes
Loan Age	Yes	Yes	Yes
Region $\cdot$ Cohort	No	Yes	No
Cohort $\cdot$ Time	No	No	Yes

Panel B: Refusal of Treatment against Leverage

Current CLTV $\leq$ 60		Excluded	
60 $<$ Current CLTV $\leq$ 80	0.0018	0.0027	0.0018
	(0.50)	(0.69)	(0.47)
80 $<$ Current CLTV $\leq$ 100	0.0040	0.0022	0.0039
	(0.91)	(0.48)	(0.89)
100 $<$ Current CLTV	0.0084	0.0086	0.0084
	(1.75)	(1.65)	(1.74)
Other Controls	Yes	Yes	Yes
Loan Age	Yes	Yes	Yes
Region $\cdot$ Cohort	No	Yes	No
Cohort $\cdot$ Time	No	No	Yes

* *p*¡0.05, ** *p*¡0.01

Appendix A: Data Construction

Data Sources

SEER Data Our data are a subset of the National Cancer Institute's Surveillance Epidemiology and End Results (SEER) program, and comprise the Cancer Surveillance System of Western Washington. The data are intended to be a comprehensive catalog of cancer diagnoses occurring between 1996–2009, totaling over 270,000 cases overall. A unique patient id links records together: patients re-enter the dataset for each separate diagnosis.

The data include a rich set of fields detailing the demographic characteristics of the patient (such as race, age, listed occupation, marital status), the nature of the cancer (its type and staging), as well as select treatment decisions taken by the patient.

Bankruptcy Data Our bankruptcy data comprise all federal bankruptcy records from Western Washington state including chapters 7, 11, and 13. These data are readily accessible through PACER and have been frequently used in prior academic scholarship on bankruptcy.

Deeds Data Our Deeds dataset is provided by DataQuick, a vendor which collects public-use transactions information. The data are organized at a property level and are comprehensive of all mortgage transactions which take place from 2000–2011 (foreclosure transactions typically go back further in time). The data list each mortgage transaction—including sales, transfers, new mortgages (first and second liens), and refinancing—which occur on a given property. We use the timing of the sales information to infer when cancer patients were resident in the property, and follow foreclosures for the duration of the time individuals were resident. We additionally use mortgage information dating to the time of the patient's residence to calculate our key leverage statistics.

BlackBox Data BlackBox LLC is a private vendor which has collected the individual mortgage records related to private label securitized bonds (ie, those not securitized by a government-sponsored entity like Fannie Mac or Freddie Mae). Though private label securitization made up only a fraction of total mortgage origination even at its peak before the crisis; our data contain more than 20 million mortgages in total; which is typically either subprime, Alt-A, or jumbo-prime in credit risk.

The BlackBox data contain static information taken at the time of origination, such as origination balance, credit score (FICO score), interest rate, and contract terms. The data are also updated monthly with dynamic information on fields like interest rates, mortgage payments, and mortgage balances. The mortgage payment field is most critical for our analysis, as it allows us to calculate the precise number of payments the household has made, not just whether or not the household has entered foreclosure.

Equifax Data Equifax is a major credit bureau which maintains detailed dynamic monthly credit information on households concerning their balances on mortgage and other debt, as well as credit scores (Vantage score).

Data Merges

A key innovation our of analysis is the use multiple sources of data on individual behavior to track financial outcomes around cancer diagnosis. This requires us to implement complex merges between many datasets which were not originally intended to be linked. Due to privacy restrictions, we are unable to make these data publicly available. However, the code used for all analysis is available upon request and below we document the document the merge process and linking variables which enable us to construct our dataset.

SEER-Bankruptcy The linkage between the SEER and Bankruptcy datasets was performed by the Fred Hutchinson Cancer Research Center via a probabilistic algorithm based on the patients name, sex, address, and last four Social Security Number digits (Ramsey et al. 2013).

SEER-Deeds Data Three match criteria were used to link SEER and Deeds data based on common text address fields:

- A *tight* match was based on full address, street directional (ie, NW), zip or city, and census tract.

- An *intermediate* match was based on house number, the first three letters of the street name, street end (ie, lane or drive), end number (any number in the last position of the address, such as an apartment number), street directional, zip or city, and census tract.

- A *loose* match was based on house number, the first three letters of the street name, street end, end number, zip or city, and census tract. These are all of the match criteria used in the intermediate match, with the exception of street directional.

The match was conducted by first prioritizing tight matches. Intermediate matches not found using the tight match were added next, and finally any loose matches not found using either of the two other methods were added. The vast majority of matches were achieved using the tight match (63,661 records were matched using the tight match; 7,970 using the intermediate match; and 2,065 using the loose match for a total of 73,696 SEER records which matched into a record in the Deeds data.

Deeds Data-BlackBox Though Deeds and BlackBox data were not designed to be linked, they are both administrative datasets containing reliable information on a variety of mortgage fields. We developed a novel a match method to link the two datasets using a training dataset (for which we knew matches exactly) to develop the algorithm. The merge relies on the following common fields:

1. Exact date matches between origination dates of the mortgage are reported in the two datasets (not used if the origination date was likely imputed; ie the date reflected in BlackBox was the first or end of the month.

2. Zip code matches between the two datasets.

3. Matches based on mortgage purpose (ie, refinancing or purchase).

4. Matches based on mortgage type (ie, adjustable-rate or fixed-rate).

5. Matches based on mortgage origination amount (rounded down to the hundred)

We used a *backward* window of 31 days, in which the mortgage origination date reflected in BlackBox was at most 31 days after the date of the mortgage reflected in Deeds; and a *forward* window of 20 days.

The match algorithm worked by first focusing on 1) zip matches and 2) origination amount matches within the backward window (or the forward window if no matches existed in the backward window).

If only one match was found using those criteria, it was kept. If there were multiple matches, we restricted further by iteratively applying the following the following criteria. We first employed a "tight" match which required that the loan match uniquely on day, or (if there were multiple day matches) uniquely on mortgage purpose or type among those that matched on day.

If this did not uniquely identify a match, we next restricted to "looser" matches where there was 1) only one match uniquely on mortgage type and purpose. If no mortgage matched, we moved on to cases where there was 2) one unique match of either mortgage type or purpose with the other field missing; 3) one unique match on mortgage type, and 4) one unique match on mortgage purpose. The merge algorithm proceeded among all matching cases in the order specified above—if a high quality match was found, the mortgage was kept and the procedure only moved on to the other match cases in the order specified if no match was found.

BlackBox-Equifax BlackBox, a mortgage-level dataset, was linked by Equifax to borrower-level information on a variety of debts, including mortgages. The merge algorithm relied on a proprietary code which we cannot access. The vast majority of accounts in BlackBox were linked to a credit account.

To verify the accuracy of the merge, we imposed a restriction samples which make use of Equifax variables. Specifically, we require that the two entries match either on 1) zip code of the borrower (at least once over the life of the loan); or 2) have a match confidence of at least .85. The zip code restriction compares the zip code of the property as listed in BlackBox matches with the address of the borrower as listed in Equifax. A mismatched zip code is not necessarily indicative of a mismatch in loans—it could also suggest the presence of an investor who does not live in the property in question.

In addition to the zip code measure, Equifax provided a measure of match confidence ranging from 0–0.9. Loans at the top end of the confidence score reflect extremely well matched loans, and we allow for a mismatch in zip code so long as it is accompanied by a match confidence score of at least 0.85. Robustness checking based on other common attributes between the two datasets (such as common measures of default) suggest that the two measures of match accuracy we employ are effective in correctly identifying well-matched loans. For further details of the BlackBox-Equifax merge algorithm; see Piskorski, Seru and Witkin (2015)

Variable Definitions

Occupation The SEER data provide a numerical occupation coding. Using the occupation coding derived from Washington State government at `https://fortress.wa.gov/doh/occmort/docs/OccupationList.pdf`; we classified the following occupation fields: Professional, Clerical, Laborer, Other Occupation, and Occupation Missing.

We impute "Unemployed" individuals as those who: 1) Are listed as "Occupation Missing," and 2) have a marital status at diagnosis which is not missing or listed as "Unknown." We assume that the occupation non-response of such individuals, since it is paired with a response on the martial status form, is indicative of a genuine non-response for occupation (which would have been recorded by the reporting hospital as an occupation had the individual reported an occupation) and is assumed to come from an unemployed individual.

Mortgage Equity For the Property Database, we measure housing equity by estimating the total mortgage amount (of both first and second liens) at origination and comparing with an estimate of house price.

To estimate the house price, we begin with the purchase price if given. Unfortunately, sometimes we lack information on sale prices (but do have data on mortgages if the mortgage was refinanced). In that case, we impute the house price based on other sales on the same property at a different time (including by other owners), and infer the original house price using a zip-level house price index from Zillow.

For the Credit Report Dataset, we use the exact mortgage balances. We combine data on both first liens (data from which is derived from BlackBox) and second liens (from Equifax). We use an estimate of origination house value derived from the reported origination loan-to-value; and adjust the house price at the time of diagnosis using the Zillow index to compute a current loan to value ratio.

Data Cleaning

From the base SEER data, the following cuts were made:

- Benign cancers were dropped.
- Among cancers reported multiple times within the same day, only one cancer entry was kept.
- Synchronous cancers were identified in which multiple cancers presented within a three month interval. Only the first instance of the synchronous cancer was kept; if the stages of the two cancers differed, the maximum stage was taken. If the sites of the two cancers differed, the cancer was classified as "Other."
- In the case of multiple, non-synchronous cancers; the cancer was included if there was at least three years subsequent to diagnosis in which there were no intervening cancer diagnoses. If there was an intervening cancer; the second cancer would be included (provided that there were no subsequent diagnosis in the three years subsequent to that diagnosis), with a dummy variable indicating the presence of a prior cancer.
- We keep patients aged 21–80 at the time of diagnosis.

To connect the SEER data with the DataQuick Deeds records, the DataQuick data were separated on the basis of sale records. If a cancer diagnosis was associated with a record prior to any recorded sale; it is assumed that a real estate transaction took place prior to when the DataQuick records begin (the year 2000) resulting in the move-in of a resident who was subsequently diagnosed with cancer prior to any other sale.

The data were organized in a panel structure based on diagnosis-calendar year. It is possible for the same patient to have multiple cancers and so be repeated in the data for the years surrounding each diagnosis (again, provided a three year window). The panel includes the five calendar years subsequent to diagnosis (counting the year of diagnosis); and five calendar years prior to diagnosis.

Three forms of censoring were applied to the panel data:

- Censoring based on property information. Calendar years prior to the individual moving into the property as reflected in a sale record were excluded, as are calendar years after the person moving out (again as reflecting in a sale record).
- Censoring based on mortality. Our data record the death date of individuals. We censor all calendar year subsequent to death.
- Censoring due to previous episode of financial distress. Given the property-centric nature of out dataset, we can only follow one foreclosure per patient, and so censor all future observations in the calendar year subsequent to financial distress (it is possible for individuals to file for multiple bankruptcies; but such events are more rare due to the statue of limitations imposed after typical bankruptcy filings. We adopt an identical censoring strategy with respect to bankruptcies.

In addition to the other cuts, the Credit Panel Data made the following additional restrictions:

1. We require that the diagnosis take place subsequent to origination.
2. We require sufficient data from our datasets in order to estimate effects. If observations are missing for the entire year of observation, the year is dropped.
3. If more than two BlackBox entries matched a given borrower in the Property Dataset, we dropped the entries. Two were permitted as these frequently coincided with a refinanced mortgage (in which both original and refinanced mortgage were present in the dataset), or a first and second lien.

4. Among entries with two BlackBox entries, entries were dropped if:

    (a) The two BlackBox entries did not share a common id as reported in Equifax. These entries may reflect mismatched loans, rather than different borrowing by the same consumer.

    (b) If the two BlackBox entries were non-overlapping in date (ie, as frequently happens in the case of refinancing), they were kept. If they were overlapping, the entry with the smaller mortgage amount was dropped (frequently, this was a second lien).