



ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Experiencing breast cancer at the workplace

Giulio Zanella
Ritesh Banerjee

Quaderni - Working Paper DSE N°938



Experiencing breast cancer at the workplace*

Giulio Zanella[†] Ritesh Banerjee[‡]

March 14, 2014

We study unique data from a dynamic natural experiment involving more than 7,000 American women to understand how a woman’s propensity to perform an annual mammography changes over time after a co-worker is diagnosed with breast cancer. We find that in the year this event occurs the probability that a woman performs a mammography drops by about 8 percentage points, off a base level of about 70%. This impact effect is persistent during at least the following 2 years, is driven by cases of breast cancer diagnosed at non-early stages, and by the behavior of individuals who are less knowledgeable about health issues. This negative effect is confirmed when we allow for serial correlation in screening behavior and when we estimate the effect of the treatment on the hazard of not screening, at the daily frequency. However, the effect vanishes in placebo experiments.

Keywords: breast cancer, social interactions, information aversion.

JEL codes: I10, C31, D03, Z10.

*We are grateful to Francesca Barigozzi, Sam Bowles, Gary Charness, Ethan Cohen-Cole, Francesca Cornaglia, Jason Fletcher, Margherita Fort, Zach Grossman, George Loewenstein, Marco Manacorda, Paolo Pinotti, Adriaan Soetevent, and seminar participants at Bocconi University, Catholic University in Milan, Queen Mary University of London, UCSB, and the Universities of Milan–Bicocca, Naples–Federico II, and Siena for insightful comments. The project was ongoing while Zanella was visiting the University of California–Santa Barbara, whose hospitality is gratefully acknowledged. This research was conducted with the approval of an Institutional Review Board at the organization that provided the data. All data used in the study are proprietary and confidential in consideration of their highly sensitive nature.

[†]Corresponding author. University of Bologna, Italy. E-mail: giulio.zanella@unibo.it

[‡]Well Traversed, USA. E-mail: ritesh.banerjee@gmail.com

1 Introduction

This paper studies how a woman’s propensity to perform an annual mammography changes over time after a co-worker is diagnosed with breast cancer. As a benchmark, there is no effect to expect for a fully informed, rational decision-maker with given constraints. Breast cancer is a non-infectious disease, and the underlying risk factors are largely independent among unrelated co-workers; therefore, the event that a colleague is diagnosed with the disease brings no new information to such a decision maker. However, effects are possible if we depart from this benchmark. We gained access to employer records from a large not-for-profit medical organization in the U.S., which we use to construct a panel data set spanning three years (2002–2004) and containing demographic, professional, socioeconomic, and high-quality health information, including mammography use and breast cancer occurrences. The data set also contains detailed information about an employee’s location in the workplace, which allows us to construct reference groups where social interactions plausibly occur on a daily basis.¹ All individuals in our sample are fully covered under a comprehensive employer-provided health plan which grants, in particular, free access to mammograms at the workplace. We focus on the screening behavior of women 50 years of age or older: women in this group are subject to unambiguous screening guidelines—yearly mammograms.

We find a negative effect at impact: in the year in which a woman is diagnosed with breast cancer, those women who are spatially closer to her in the workplace (“treatment group”) become about 8 percentage points less likely to perform a mammography relative to all other women (“control group”), off a baseline screening rate of about 70%. This average impact response is persistent. Placebo experiments show that such dynamic effects are absent for future diagnoses, for tumors other than breast cancer, and when actual breast cancer occurrences are randomly reallocated across reference groups. We explore the presence of heterogeneity along several margins. The results suggest that the negative treatment effect is stronger for women exposed to more severe cases of breast cancer, and is both stronger and more persistent for women who are neither medical doctors nor nurses. These findings are important because while the economics of test avoidance has been extensively investigated in individual settings (e.g., Botond Köszegi, 2003, Andrew Caplin and Kfir Eliaz, 2003, and, more recently Emily Oster, Ira Shoulson, and Ray Dorsey, 2013), little is known about avoidance behavior in social contexts. Behavioral and non-behavioral mechanisms rationalize these findings. We illustrate them at the end of the paper, and we weight the evidence in favor and against them in the light of our own empirical results and of existing “mechanism experiments”. Although this remains an open question for future research, two compelling interpretations are, respectively, information aversion, and preventing the employer from learning information that may lead to negative consequences for the employee.

¹This makes our setting similar to the field experiment of Esther Duflo and Emmanuel Saez (2003), who find that providing an information shock about retirement plans to randomly selected employees increases the enrollment rate of their (non-selected) co-workers as well.

The reason why with such rich data set at hand we focus on screening for breast cancer is that this disease has clear prevention guidelines which allow us to unambiguously define “normal” screening behavior. Because it is typically recommended to women 50 years of age or older to perform an annual screening mammography, we can look at the propensity to screen for breast cancer every year after this age as an indicator of the propensity to take “normal” preventive actions against the disease. This is much harder to do for other types of tumors, given the longitudinal limitation of our data set. Consider, for instance, prostate cancer. Like breast cancer for women, this is the most common non-skin cancer affecting men. A recent statement of the US Preventive Services Task Force says that “the current evidence is insufficient to assess the balance of benefits and harms of prostate cancer screening in men younger than age 75 years”, and “recommends against screening for prostate cancer in men age 75 years or older” (USPSTF, 2008). In this case it would be impossible to determine what “normal” screening behavior is. Or consider cervical cancer, where the guideline is that women should have a Pap test every three years. With three years of data and staggered screening it would be impossible, in our data set, to tell who is complying with this guideline and who is not.

In addition to their interest for research on decision-making in social contexts, our findings have potentially important implications for understanding the determinants of screening avoidance. Breast cancer is the most common non-skin malignancy. The National Cancer Institute (NCI) estimates that an American woman has, at birth, a 12.4% probability of developing breast cancer at some point during her life. According to the NCI more than 230,000 new cases have been diagnosed in the US in 2013, and almost 40,000 women died in 2013 because of breast cancer—the most common cause of cancer-related death after lung cancer. The good news is that breast cancer can be treated effectively. According to the NCI the survival rate is about 90%, and this figure increases to more than 98% for cases diagnosed at the very first stage of cancer development. That is, early detection can greatly reduce the mortality rate. The most effective screening tool for early detection of breast cancer is a mammography, a relatively simple and inexpensive preventive measure.² As a consequence, the medical community has reached an effective consensus on the need for mammograms every year for women over 50.³ Yet only two thirds of American women 40 years of age

²Although women can perform non-clinical breast exams (such as a breast self-exam) the NCI stresses that these should not take the place of routine clinical breast exams

³There is some controversy over when women should begin breast cancer screening. The American Cancer Society, American College of Radiology, American Medical Association, National Cancer Institute, and the American College of Obstetrics and Gynecology recommend annual screening starting at age 40. The American College of Physicians recommends that women in the age range 40–49 make decisions about mammography together with their provider, based on the individual risk profile and the potential costs and benefits of performing a mammogram. In November 2009 the US Preventive Services Task Force advised women below 50 years old not to get routine mammograms, but to discuss the pros and cons with their provider and decide together when to start screening. This advice—whose rationale is the fact that mammograms may generate false alarms and unnecessary treatment—has generated a lot of debate

or older perform mammograms at least every other year, and this rate has been declining relative to the past decades (Nancy Breen *et al.*, 2007). The above NCI estimates imply that every year about 3 American women out of 1,000 in the age group 40 and older develop breast cancer. A uniform screening rate of 70% implies that 0.9 of these cases would go undetected during the year they occur, every year—i.e., they would be detected with a delay of at least one year. Therefore, the 8 percentage-point reduction in the propensity to screen we identify in this paper implies, if applied to every woman, that the number of undetected cases at early stages would increase to 1.14 per 1,000 cases, an increase in excess of 1/4.⁴ The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 illustrates the econometric design and discusses the identifying assumptions. Section 4 reports the results. Section 5 discusses mechanisms. Section 6 concludes.

2 Data set

Our data come from a large not-for-profit medical organization (medical practice and medical research) in the U.S. that has over 20,000 employees (70% of whom are women) at its main location. All employees are eligible for a health plan run by the organization itself, and given the plan’s comprehensive nature and the lack of substitutes of comparable quality, participation is virtually universal. This insurance plan includes, in particular, free annual mammograms for all women 40 years of age or older. The facility for screening is located on campus and is well enough staffed so as not to keep patients waiting for an appointment. These women are reminded about their recommended annual mammogram (if they have not already had one) each time they visit their primary care physician. The physician is provided a copy of their electronic medical record, which lists the last time they had a mammogram conducted. If they are due for a mammogram, the physician generally schedules an appointment during the visit (or asks a nurse to do so) unless the patient refuses an appointment. Therefore, many of the barriers typically pointed at when interpreting national rates and trends of mammogram use (e.g., lack of insurance, high copay, lack of reminders, etc.) are absent in our study.⁵

The final data set consists of a three-year panel (2002–2004) and results from the combination of a variety of employer’s records. Demographic information were gathered from

in the US. Moss *et al.* (2006) perform a large randomized controlled trial in the UK and find that annual mammographies in the age range 39–48 do not lead to a significantly lower mortality rate relative to control women who did not screen regularly.

⁴To be clear, this is not necessarily suboptimal; Archie Bleyer and Gilbert Welch (2012) argue that there was a substantial amount of breast overscreening (i.e., mammographies detecting tumors that would not have led to clinical symptoms) in the U.S. during the past 30 years. They estimate that almost 1/3 of all breast cancer diagnoses over this period resulted from overdiagnoses. We will return on this point at the end of the paper.

⁵See, for instance “Mammograms in Decline”, *The New York Times*, May 15, 2007.

electronic administrative records maintained for all patients. We obtained job and career information from data bases maintained by the Human Resources department. A crucial piece of information for the purposes of our research question is an employee’s physical location in the workplace. This is crucial because it is used to define reference groups (who interacts with whom), and was obtained from the Information Technology department, which maintains an employee telephone and internal address directory. Information on health care utilization were obtained from administrative billing records that are maintained for all patients. We used an electronic data base that was developed to track patient visits to their primary care physician along with routine preventive care to create a variable that tracks whether an individual performed a mammography in a given year, as well as high quality health indicators. We measure an individual’s health status using the classification scheme in Elixhauser *et al.* (1998), who use health utilization data to create a set of indicator variables revealing whether an individual has a history of medical claims for a certain health condition (a “comorbidity”). Finally, we used a locally maintained Tumor registry that keeps records of all patients who are diagnosed with a malignant tumor of any type at the organization, or who are diagnosed elsewhere but treated at the organization, to create a cancer diagnosis variable for all employees. This information is available for all years between 2000 and 2004. A high-frequency portion of these records allows us to observe the exact date a woman was diagnosed with breast cancer, as well as the date a woman last performed a mammogram, as of 2004. We will use this higher frequency information to refine the definition of treatment and to perform an auxiliary duration analysis at the daily frequency. The common element in all of these data was a patient ID variable that allowed us to link records. Data obtained from the Human Resources department were pre-linked with the patient ID before being made available to us in order to protect confidential personal information.

Our data set has two main limitations. First, unfortunately demographic and employment information were made available to us only for year 2004. Therefore, apart from age, we must assume in the longitudinal analysis that they do not change between 2002 and 2004. For most socioeconomic characteristics, this is a good approximation. Our inability to observe moves within the organization is more problematic, because we must assume that the reference groups in 2002 and 2003 are the same as in 2004. However, because such moves are not very frequent and because the turnover in our data is limited (only 1% of employees in 2004 are new hires at the organization), the resulting misclassification of reference groups is most likely negligible. Second, we do not observe women who died from breast cancer (or otherwise left the organization) before the end of 2003, because they were removed from the administrative records that were made available to us. We discuss below in greater detail these two limitations and their consequences. Briefly put, they both “go against us” because they bias the effect of interest towards zero. Furthermore, we show later that we can perform a compelling robustness check based on a clean duration analysis at the daily frequency on the 2004 portion of the data set. Such duration analysis completely overcomes these and other potential selection problems.

The analysis in this paper focuses primarily on the age group 50 and older (about 3,000 women), for two reasons. First, mammograms are not universally recommended before age 50 (see footnote 3). Therefore, both the individual screening recommendation from one’s physician and the baseline perception of “normal” or “necessary” preventive screening in our sample may be ambiguous in the age range 40–49. Second, we lack screening information in the age range 40–49 for year 2002. Table 1 summarizes what type of information is available for each year. Summary statistics are reported in Tables 2 and 3.

Table 1. Longitudinal availability of data

	2000	2001	2002	2003	2004
Demographic	x	x	x	x	✓
Job	x	x	x	x	✓
Screening status*	x	x	✓	✓	✓
Comorbidities	x	x	✓	✓	✓
Cancer history	✓	✓	✓	✓	✓

Notes: The table summarizes which type of information is available for each year: ✓=available; x=not available; *=not available for 2002 in the age group 40-49. Except for age, missing demographic and job information in 2002 and 2003 is assumed to be the same as in 2004.

Table 2. Summary statistics: demographic and job information

Variable	Mean	Std. Dev	Min	Max
Age	54.55	3.56	50	64
Married	0.71	0.45	0	1
Divorced	0.15	0.36	0	1
Widowed	0.01	0.09	0	1
Separated	0.03	0.17	0	1
Family size	2.11	1.08	1	7
Has children below 18	0.09	0.29	0	1
Whether hourly pay	0.97	0.18	0	1
Medical doctor or nurse	0.25	0.43	0	1
% of full time employee	91.37	14.22	50	100
Job tenure, years	19.35	10.61	0	46
New employee (tenure=0)	0.01	0.11	0	1
Salary range (median)	0.76	0.52	–	–
Chose “High” plan in 2004	0.75	0.43	0	1
Charges for health services	7.92	13.8	0	204.98

Notes: The median of the salary range an individual belongs to is expressed as a % of the mean of all employees at the organization, upon an explicit request of the organization itself to protect the confidentiality of this information. Charges for health services are expressed in thousands of 2004 US dollars. “High” plan is a high premium-low copay health plan. See text below for details.

Table 3. Summary statistics: health information (comorbidities)

Variable	Mean	Std. Dev	Min	Max
Neurological disorders	0.013	0.115	0	1
Chronic lung disease	0.052	0.221	0	1
Diabetes	0.046	0.209	0	1
Diabetes w/complications	0.014	0.117	0	1
Hypothyroidism	0.087	0.282	0	1
Renal failure	0.001	0.037	0	1
Congestive heart failure	0.001	0.037	0	1
Arrhythmia	0.035	0.185	0	1
Valvular disorders	0.025	0.155	0	1
Pulmonary circular disorders	0.002	0.043	0	1
Perivascular disorders	0.009	0.092	0	1
Parathyroid disorders	0.001	0.025	0	1
Liver disorders	0.008	0.087	0	1
Ulcers	0.002	0.044	0	1
Tumor	0.061	0.240	0	1
Arthritis	0.021	0.145	0	1
Coagulation	0.004	0.061	0	1
Extreme obesity	0.088	0.283	0	1
Significant weight loss	0.006	0.075	0	1
Electrolyte disorders	0.013	0.114	0	1
Blood loss	0.001	0.025	0	1
Anemic Deficiency	0.045	0.207	0	1
Alcoholism	0.005	0.069	0	1
Psychoanalytic Disorders	0.059	0.235	0	1
Depression	0.031	0.173	0	1
Lymphatic disorders	0.002	0.047	0	1
Metastatic cancer	0.006	0.078	0	1
Hypertension	0.231	0.422	0	1
Number of comorbidities	0.824	1.110	0	11

Notes: All means refer to observations pooled over the 2002–2004 period. Cancer information comes from the Tumor registry maintained by the organization. Comorbidity indicators are constructed from individual administrative health records using the classification scheme of Elixhauser et al. (1998); see text for details. The age group is 50 and older.

Table 4 reports annual screening rates in the balanced portion of our sample, by age range. These screening rates are larger than corresponding national rates, as one would expect for a group of women employed in the health industry and with full health coverage. To illustrate how screening rates in our sample compare to national rates, in our data the fraction of women who performed a mammogram in 2003 or 2004 is 77.7% in the age group 40-49, and 86.3% in the age group 50-64. The closest corresponding national rates (fraction of women who performed a mammogram in 2004 or 2005) are 63.5% and 71.8%, respectively (Nancy Breen *et al.*, 2007). Gross of cohort effects (which we do not attempt to net out), in our sample screening rates tend to increase with age, consistently with the fact that age is the most important risk factor for breast cancer. The table shows a decline between 2002 and 2003 (except for the age group 55–59), as well as between 2003 and 2004, in line with the national trend in the period under investigation (Nancy Breen *et al.*, 2007).

Table 4. Screening rates by year, age range, and health plan in 2004

Year	Age 41-49	Age 50+	Age 50–54	Age 55–59	Age 60+	“High” in ‘04	“Low” in ‘04
2002	—	0.762	0.747	0.782	0.807	0.764	0.754
2003	0.617	0.746	0.721	0.792	0.786	0.754	0.719
2004	0.513	0.638	0.622	0.683	0.630	0.634	0.649
	n=3794	n=2451	n=1568	n=1253	n=384	n=1889	n=562

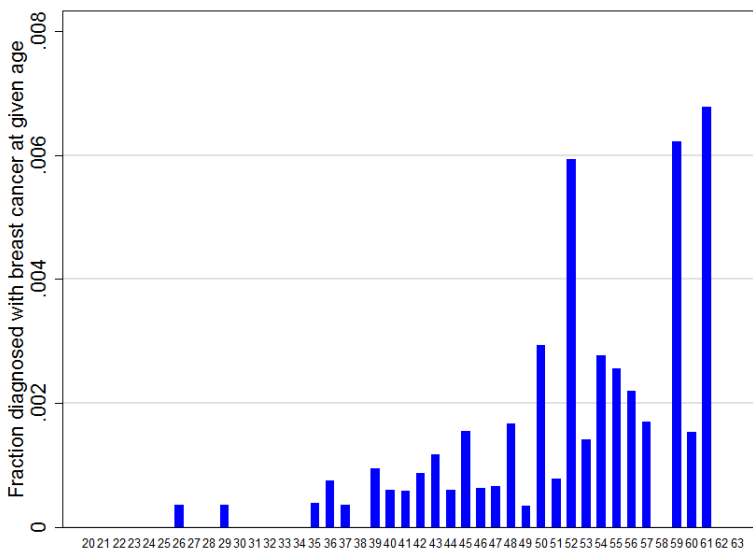
Notes: The table reports the average, by year and age range, of the individual screening indicator, which takes value 1 if a woman performed a mammogram in a given year and 0 otherwise; n is sample size. These rates are computed on the balanced portion of the sample to net out composition effects. “High” in ‘04 and “Low” in ‘04 indicate women in the age group 50+ who have chosen, in 2004, the high premium–low copay and low premium–high copay combinations, respectively.

The drop observed in our data between 2003 and 2004 is quite large, though. The only relevant event in this period we are aware of is a change in the health insurance options that took place at the organization in 2004. Employees were required to choose between two combinations: high premium–low copay (“High” plan) and low premium–high copay (“Low” plan). This may have driven down overall utilization of health services, even those—like mammograms—that remained free (i.e., with zero copay): if women didn’t see their doctor as often as before, they were not reminded about their mammogram as regularly as before. This change poses no particular challenge as long as we control for year effects, and (most important) as long as choices over the two plans are balanced across treatment and control groups (which is clearly the case, as we show later). Furthermore, we see in the last two columns of Table 4 that screening rates of women who choose either the “High” or the “Low” plans do not differ substantially.

In our data, there are 88 cases of breast cancer diagnosed between 2000 and 2004: 11 in 2000, 23 in 2001, 14 in 2002, 20 in 2003, and 20 in 2004. These numbers correspond to an

annual incidence ranging between 0.24% and 0.28%, in line with national estimates.⁶ The age frequency distribution of cancer diagnoses is plotted in Figure 1. Occurrences below age 40 are about 10% of the total, and those above 49 are almost 60% of the total. These cases vary in severity. We observe two measures of severity: stage and grade. Cancer stage depends on the size of the primary tumor and on the extent it has spread in the body. The larger and the more spread the tumor, the higher its stage. Cancer grade (also known as differentiation), instead, depends on the extent to which the tumor is different from the tissue it comes from, and on how quickly it is likely to grow and spread. The less differentiated and the quicker the tumor is likely to spread, the higher its grade. The cross tabulation of stage and grade of the 88 cases we observe are reported in Table 5. This table shows that most cases are diagnosed before the tumor has reached stage 2 or grade 3. Both the age and severity distribution of breast cancer occurrences in our sample are in line with national estimates for the U.S.

Figure 1. Fraction diagnosed with breast cancer, by age, 2000–2004



Notes: The histogram represents the frequency distribution of breast cancer diagnoses, by age, pooling all female employees at the organization in the 2000-2004 period. That is, the height of a bar indicates the fraction of women diagnosed with breast cancer in a given age bin.

⁶The 232,000 new cases of breast cancer estimated by the National Cancer Institute in the U.S. for 2013 are equivalent to an incidence of 0.27% for women above age 35, which are about 86 million in the U.S.

Table 5. Severity of breast cancer cases, 2000-2004

	Stage								Total	
	0	1	2A	2B	3A	3B	3C	4		
Grade	1	0	5	4	0	1	0	0	0	10
	2	3	18	4	7	0	1	2	0	35
	3	4	4	5	3	0	0	0	1	17
	4	5	2	0	2	1	0	0	0	10
	undecided	12	2	1	0	0	0	0	1	16
Total	24	31	14	12	2	1	2	2	88	

Notes: The table reports the cross-tabulation of grade and stage (two measures of severity, see text for details) of the 88 cases diagnosed in our data between 2000 and 2004.

It is worth noticing that the screening history of women later diagnosed with breast cancer is statistically indistinguishable from the corresponding history of those who were not. Table 6 illustrates this fact. The upper portion of the table reports the fraction of women who performed a mammogram in 2002 and 2003, in two distinct groups: those with no breast cancer as of 2004 and those diagnosed with the malignancy in that year. Similarly, the lower portion of the table reports the fraction who screened in 2002, by diagnosis status in 2003. The p -value refers to the null hypothesis that the means are equal in the two groups. This hypothesis is never rejected—although, because of the size of the two groups, only large differences can be rejected.

Table 6. Screening history by diagnosis status

Variable	No BC in 2004		BC in 2004		p-value
	Mean	s.d.	Mean	s.d.	
Screened in 2002	0.725	0.447	0.727	0.467	0.99
Screened in 2003	0.677	0.468	0.647	0.493	0.79

Variable	No BC in 2003		BC in 2003		p-value
	Mean	s.d.	Mean	s.d.	
Screened in 2002	0.724	0.447	0.900	0.316	0.21

Notes: Fraction of women who performed a mammogram in 2002 and 2003 in groups with no breast cancer (BC) in 2004 (upper part) or 2003 (lower part) and in groups with. Women diagnosed with breast cancer prior to 2004 or 2003 are excluded from the calculations in the upper and lower portions of the table, respectively. The p -value refers to the null hypothesis that the means are equal in the two groups.

How is this information transmitted to co-workers? Of course we don't know whether a woman who develops breast cancer in our sample shares such information directly with her colleagues or not. However, the available empirical evidence suggests that it is very unlikely that her closest co-workers remain unaware of her condition for a long time, even if she survives cancer. Michael Hassett *et al.* (2009) document that a woman diagnosed with breast cancer typically leaves the workplace for some time, either to complete therapy or because of a decreased desire or ability to work. Cathy Bradley *et al.* (2005) have analyzed the effects of a new breast cancer diagnosis on labor supply of women who were employed before the diagnosis. They find that six months after the diagnosis, sick women are much less likely to be employed (about 20 percentage points less than the control group) and if they are employed their weekly hours of work are 12% to 28% depressed relative to the control group, depending on the stage of the malignancy. It is very unlikely that such substantial changes in labor supply and prolonged absences go unnoticed by the closest co-workers. It is then plausible that this information further spreads via workplace social interactions.

However, this same evidence also points to a potential selection problem, which relates to the second data limitation mentioned above: if a woman diagnosed with breast cancer between 2000 and 2003 is no longer employed at the beginning of 2004 (either because she quit, or because she was fired, or because she died), then this occurrence is not observed. This is a data limitation which we cannot correct. The consequence, however, is the usual attenuation bias arising when treated individuals are incorrectly assigned to the control group. Because the close colleagues of a woman diagnosed with breast cancer in the past but no longer employed at the time we observe the employer records would be classified as not having had a co-worker diagnosed with breast cancer, the estimated treated effect would be biased towards zero. It follows that our estimates should be interpreted as a lower bound (in absolute value) of the true effect. There are three reasons to believe this potential selection problem is mild. First, the high survival rate and the fact that most cases are detected early enough in our data imply that most women diagnosed with breast cancer between 2000 and 2003 are still alive at the end of 2003. Second, a sick woman with a premium employer-provided health plan is very unlikely to quit. Third, under the Americans With Disabilities Act it is illegal to discriminate (e.g., fire) a woman because she was diagnosed with cancer. While possible, it is unlikely that a large not-for-profit medical organization would infringe such provision.

A crucial piece of information is the physical location of employees in the workplace. The organization is administratively divided into departments. Smaller departments tend to be physically located in one place, while larger departments may have multiple physical locations across campus. The campus consists of about 60 buildings, most of which are multi-storied; large departments will often occupy multiple contiguous floors. Each floor has a mail room where the internal mail service regularly delivers campus and external mail. In large buildings, a floor may have more than one mail room, whereas in small buildings, different floors may share a mail room. Each employee is assigned a mail room code based

on their main office location. Such mail rooms define as our first notion of reference groups, labeled “Mail” groups. That is, two individuals belong to the same Mail reference group if and only if they collect their mail in the same mail room. This defines groups of people who, presumably, interact closely on a daily basis, because they are very likely to belong to the same department—regardless of whether and how they collect their mail. A concern with this definition is that because of the relatively small dimension of Mail groups as well as the particular setting of our study, the Stable-Unit-Treatment-Value assumption (SUTVA) may be violated: people may well interact with other employees outside the group defined by mailbox locations, which would also result in attenuation bias. To address this concern, we employ two additional, alternative definitions of a reference group. First, we define groups at the floor level, labeled “Floor” groups. Second, we define groups at the level of entire buildings, labeled “Building” groups. On average, Floor groups are more than three times larger than Mail groups, and Building groups are almost five times as large as Floor groups. These larger reference groups extend the possible spatial range of social interactions and allow for broader patterns of information sharing. Table 7 summarizes the number and size (counting both men and women and excluding singletons) of Mail, Floor, and Building groups. As a comparison, Esther Duflo and Emmanuel Saez (2003) define reference groups as departments, whose average size in terms of staff is about 30 in their data.

Table 7. Reference groups

Group	Number	Mean size	Std. dev.	Min	Max
Mail	881	22.7	26.1	2	214
Floor	282	71.4	69.7	2	446
Building	59	341.2	555.1	2	2,913

Notes: The Table reports the number and average size of the three alternative reference groups used in this study—groups defined by mailbox locations (Mail), floors within buildings (Floor), and entire buildings (Building)—as well as the standard deviation and min and max of group size. Groups with less than 2 individuals are excluded.

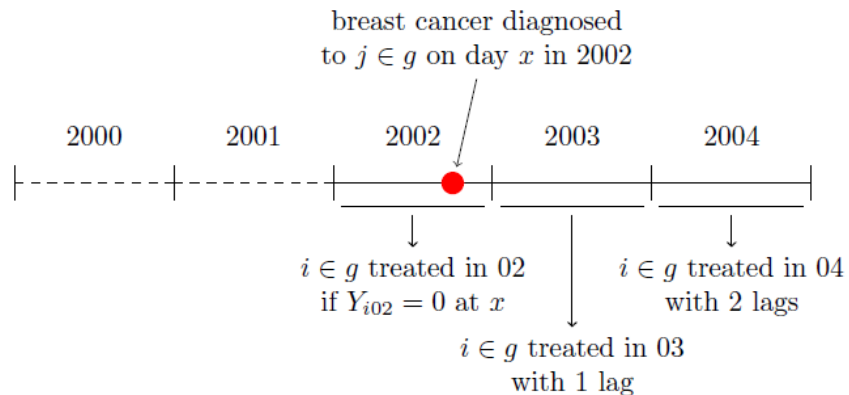
3 Econometrics

3.1 Outcome and treatment

The outcome of interest is whether a woman performs a mammogram in a given year or not, denoted Y_t . If a woman screens during year t , then $Y_t = 1$. Otherwise $Y_t = 0$. The dynamic treatment is defined as follows. A woman i in group g is classified as treated in year t if a colleague $j \in g$ was diagnosed with breast cancer during year t and if in that year i was not

up-to-date with breast screening the day that j was diagnosed with the malignancy. In this case we set the contemporaneous treatment indicator $T_{g,i}^t = 1$. Otherwise $T_{g,i}^t = 0$. Notice that this indicator depends on i because of the qualification about the timing of screening in year t . Such qualification is important: if we defined treatment in year t as just “belonging to a group where a woman was diagnosed with breast cancer during that year” then we would incorrectly classify as “treated” women who are not—if one performed a mammogram in year t before a co-worker was diagnosed with breast cancer then she cannot respond to this event that year because she is already up to date. Of course such qualification also introduces an endogeneity problem with the contemporaneous treatment indicator, $T_{g,i}^t$, because the timing of screening within the year is not random. We postpone the illustration of how we deal with this problem until we discuss identification. Lagged treatment is defined in a similar way, and is denoted T_g^{t-k} , where k is the lag. Thus, if during the past year a co-worker j of individual i in group g was diagnosed breast cancer, then $T_g^{t-1} = 1$. Otherwise, $T_g^{t-1} = 0$. In this lagged case, the complication associated with the timing of screening within the year a co-worker is diagnosed with breast cancer does not arise. Given the longitudinal limitations of our data set we restrict up front to treatment histories of the type $\mathbf{T}_g^t = (T_{g,i}^t, T_g^{t-1}, T_g^{t-2})$. It is useful to illustrate the dynamic treatment using the timeline in Figure 2. Suppose that $j \in g$ is diagnosed with breast cancer on September 1st, 2002 (“day x ”, in the figure). If in 2002 j ’s co-worker $i \in g$ had not performed a mammogram by September 1st, then she is classified as treated in 2002. Furthermore, she is classified as treated with 1 and 2 lags in 2003 and 2004, respectively. If instead $i \in g$ had performed a mammogram on, say, June 1st, 2002 then $i \in g$ is classified as not treated in 2002, although she is still classified as treated with 1 and 2 lags in 2003 and 2004, respectively.

Figure 2. Timeline.



Given this definition of treatment, we report in Table 8 the number of breast cancer occurrences in a given year of the panel and in the two preceding years, as well as how

many women aged 50 and over belong to such groups. Of course women who are diagnosed with breast cancer are throughout excluded from the analysis, after using them to identify groups with breast cancer occurrences. We see from this table that a fraction in the range 11%–16% of women above 50 experience breast cancer at the workplace in Mail groups at one of the lags we are considering. This fraction, of course, increases with the spatial scope of the reference group, ranging between 29% and 39% in Floor groups, and between 77% and 82% in Building groups. The table also shows that even at the Mail group level, some groups are treated more “intensely” than others because they comprise more than one case of breast cancer. We will check whether such “intensity” of treatment matters.

Table 8. Groups affected by breast cancer and individuals therein

	Year	Breast cancer cases at t, t-1, t-2	Groups affected	Share of total	Women 50+ in affected groups	Share of total
Mail	2002	48	43	0.05	366	0.11
	2003	57	52	0.06	522	0.16
	2004	54	53	0.06	502	0.16
Floor	2002	48	37	0.13	947	0.29
	2003	57	45	0.16	1118	0.35
	2004	54	46	0.16	1246	0.39
Build.	2002	48	18	0.31	2502	0.77
	2003	57	21	0.36	2657	0.82
	2004	54	20	0.33	2582	0.80

Notes: The Table reports the number of groups where someone is diagnosed with breast cancer in 2002, 2003, and 2004, as well as in the two preceding years. That is, rows labeled “2002” refer to new cases diagnosed in 2000, 2001, and 2002, rows labeled “2003” refer to new cases diagnosed in 2001, 2002, and 2003, and rows labeled “2004” refer to new cases diagnosed in 2002-2004.

Table 9 compares the observable characteristics of women in groups with a new breast cancer diagnosis in a given year or the two preceding years, and those without, at the Mail group level. The p -value refers to the null hypothesis that the means are equal across the two types of groups. Once the longitudinal dimension is taken into account, there are no systematic differences across the two types of groups. The one exception is group size—the table reveals that breast cancer occurrences are concentrated in larger groups. This fact has a simple explanation: the larger the group the higher the probability of finding a sick woman in that group. In the interest of space we do not report the analogs of Table 9 for Floor and Building groups. These show the same pattern, and are available from the authors upon request.

Table 9. Characteristics of women across cancer and no-cancer groups

Variable	Year: 2002			Year: 2003			Year: 2004		
	no BC	BC	p-val	no BC	BC	p-val	no BC	BC	p-val
Group size	46.3	78.6	0.00	41.2	97.2	0.00	42.3	90.8	0.00
Age	53.9	53.9	0.96	54.2	54.1	0.63	54.5	54.5	0.89
Married	0.70	0.73	0.32	0.71	0.70	0.88	0.71	0.70	0.53
Divorced	0.16	0.16	0.88	0.15	0.18	0.14	0.15	0.16	0.38
Widowed	0.01	0.01	0.79	0.01	0.02	0.08	0.01	0.01	0.23
Separated	0.03	0.03	0.96	0.03	0.03	0.65	0.03	0.03	0.81
Family size	1.98	2.06	0.23	2.04	2.09	0.44	2.11	2.14	0.52
Children < 18	0.06	0.04	0.18	0.08	0.07	0.64	0.09	0.11	0.20
Hourly pay	0.96	0.97	0.64	0.96	0.97	0.43	0.96	0.98	0.05
MD or nurse	0.24	0.25	0.57	0.24	0.26	0.22	0.25	0.25	0.79
% full time	0.91	0.90	0.27	0.91	0.91	0.39	0.91	0.91	0.18
Job tenure	19.7	21.0	0.06	19.7	18.7	0.09	19.4	18.8	0.23
New employee	0.01	0.02	0.09	0.01	0.02	0.01	0.02	0.03	0.03
Salary range	0.76	0.78	0.65	0.76	0.75	0.68	0.77	0.72	0.04
“High” plan in 04	0.77	0.74	0.26	0.76	0.76	0.97	0.75	0.77	0.42
Charges	8.21	7.67	0.58	7.70	8.43	0.32	7.56	8.73	0.09
Comorbidities	0.79	0.69	0.14	0.89	0.87	0.78	0.78	0.83	0.38
Neurological disord.	0.014	0.017	0.79	0.011	0.022	0.06	0.012	0.022	0.08
Chronic lung disease	0.055	0.042	0.38	0.053	0.054	0.89	0.047	0.062	0.16
Diabetes	0.046	0.042	0.74	0.046	0.039	0.53	0.044	0.060	0.14
Diabetes w/complic.	0.015	0.004	0.15	0.013	0.015	0.73	0.014	0.011	0.61
Hypothyroidism	0.085	0.054	0.10	0.086	0.086	0.98	0.087	0.084	0.83
Renal failure	0.001	0.000	0.62	0.001	0.002	0.58	0.001	0.002	0.74
Congestive heart fail.	0.001	0.000	0.55	0.001	0.002	0.76	0.001	0.000	0.47
Arrhythmia	0.036	0.025	0.38	0.035	0.044	0.36	0.037	0.022	0.10
Valvular disorders	0.015	0.038	0.01	0.029	0.032	0.73	0.025	0.024	0.92
Pulmonary c. disord.	0.001	0.000	0.55	0.002	0.000	0.35	0.001	0.002	0.56
Perivascular disord.	0.005	0.013	0.19	0.008	0.012	0.43	0.009	0.009	0.89
Parathyroid disord.	0.000	0.000	0.73	0.000	0.000	0.67	0.001	0.000	0.56
Liver disorders	0.008	0.000	0.15	0.007	0.000	0.07	0.009	0.009	0.98
Ulcers	0.000	0.000	0.73	0.003	0.000	0.26	0.003	0.000	0.24
Tumor	0.053	0.088	0.03	0.053	0.051	0.94	0.050	0.042	0.49
Arthritis	0.016	0.013	0.68	0.027	0.025	0.81	0.023	0.016	0.32
Coagulation	0.002	0.008	0.13	0.005	0.005	0.98	0.003	0.002	0.86
Extreme obesity	0.086	0.067	0.31	0.106	0.103	0.86	0.074	0.080	0.65
Signific. weigh tloss	0.002	0.000	0.44	0.009	0.002	0.17	0.007	0.002	0.24
Electrolyte disord.	0.011	0.017	0.44	0.014	0.012	0.74	0.012	0.020	0.19
Blood loss	0.001	0.000	0.62	0.001	0.000	0.55	0.000	0.000	0.68
Anemic Deficiency	0.049	0.008	0.00	0.056	0.037	0.11	0.035	0.035	0.97
Alcoholism	0.005	0.000	0.25	0.005	0.007	0.51	0.004	0.007	0.039
Psycho. disord.	0.057	0.029	0.07	0.057	0.084	0.04	0.057	0.064	0.56
Depression	0.031	0.038	0.58	0.032	0.032	0.99	0.029	0.024	0.65
Lymphatic disord.	0.002	0.004	0.50	0.003	0.002	0.84	0.002	0.000	0.35
Metastatic cancer	0.001	0.008	0.03	0.002	0.010	0.01	0.004	0.006	0.39
Hypertension	0.222	0.217	0.87	0.246	0.232	0.53	0.219	0.250	0.13

Notes: Comparison, at the Mail group level, of the observable characteristics of women in groups with a new breast cancer diagnosis in a given year or during the previous two years (BC) with the corresponding characteristics of women in groups with no breast cancer diagnosis (no BC). The p -value refers to the test of the hypothesis that the difference between the two means is zero.

Before turning to a formal discussion of identification, we give a visual sense of the effect of interest by comparing the screening rates of women 50 years of age and older in groups where someone was diagnosed with breast cancer at some point in time with those in groups where no one was, for the three years of the panel (2002, 2003, and 2004). For the moment, no correction is made for the timing of screening during the year of the diagnosis. Furthermore, to ensure we compare the same groups of women over time, we restrict the sample to women who are 50 or older in 2002 and with a complete (2002–2004) breast screening history. These screening rates are illustrated in Figures 3–6. The vertical segments around the point estimates are the 95% confidence intervals. The first comparison is reported in Figure 3, where we consider Mail groups, breast cancer cases diagnosed in 2004 (hence the vertical line, which divides the “pre”, 2002 and 2003, and the “post”, 2004) and four different categories. The first category are all women (upper-left). The screening rates in the two types of groups are similar in 2002 and 2003 (ranging between 73% and 77%) and a common trend hypothesis cannot be rejected. In 2004 the screening rates decline in both types of groups, but the drop is larger in groups where someone is diagnosed with breast cancer that year than in the other groups—the screening rates become 59% and 65%, respectively. The second category are women who performed a mammogram in 2002 (upper-right). We see a similar pattern—the screening rates are higher because we are now considering women with a higher propensity to screen. The third category are women who performed a mammogram in 2003 (lower-left). In this case the pre-2004 trends are perfectly overlapped, and we see again that in 2004 the screening rate is much lower in groups with a new breast cancer occurrence (63%) than in those without (72%). The fourth and last category are women who screened both in 2002 and 2003—“always screeners”. The figure shows that among these women the decline in the screening rate in 2004 is much more pronounced in groups with a new breast cancer case than in other groups (the rates are 65% and 75%, respectively). We also see in this picture that always screeners respond in a way similar to more irregular screeners. Figures 4–5 show these same data at the Floor and Building groups levels. In these cases there are no visible effects of new breast cancer occurrences. Therefore, these pictures suggest that if there is an impact effect (i.e., at lag 0), this is likely present at the Mail group level only—i.e., among the spatially closer co-workers. In order to also evaluate the dynamic effect in this same, preliminary fashion we repeat the visual exercise considering breast cancer occurrences in 2003. In this case, the “pre” period is 2002, and the “post” period is 2003–2004. Figure 6 also suggest a negative effect at the Mail group level, at least in the “all women” category, as well as some persistence.

Figure 3. Screening rates by breast cancer occurrences in 2004: Mail groups

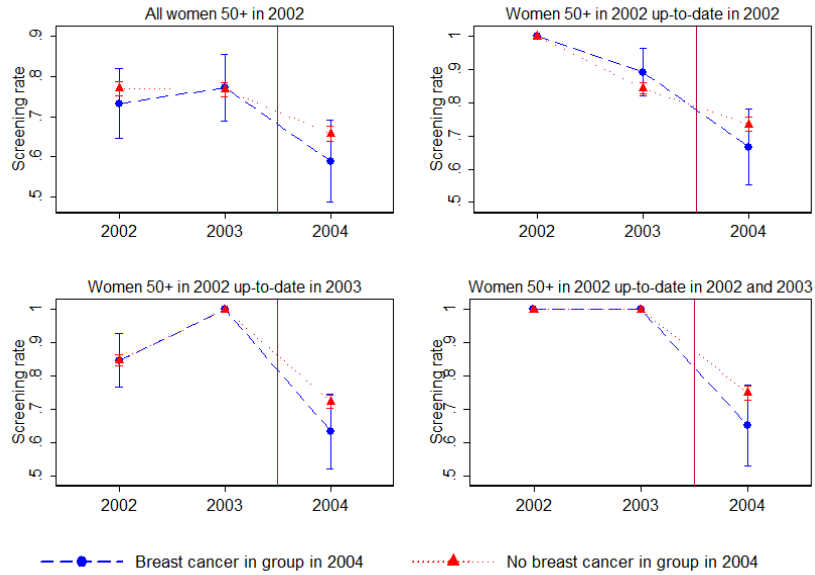


Figure 4. Screening rates by breast cancer occurrences in 2004: Floor groups

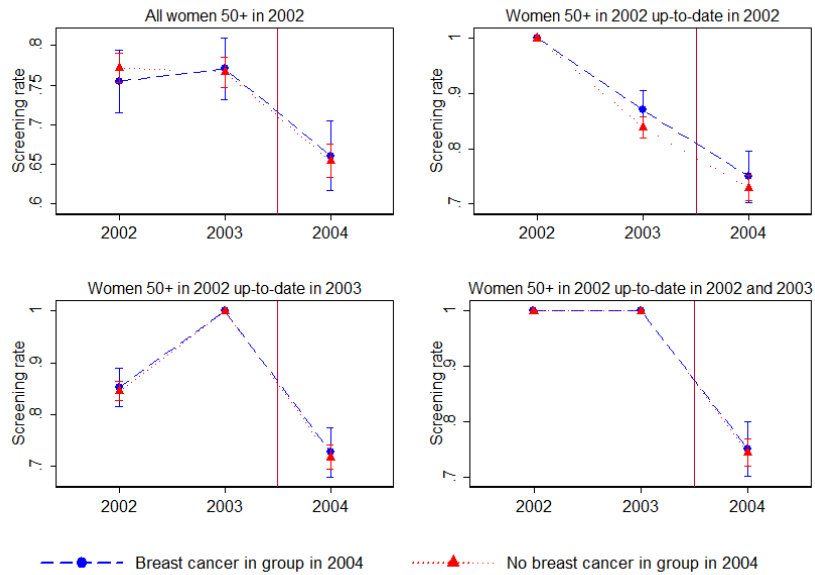


Figure 5. Screening rates by breast cancer occurrences in 2004: Building groups

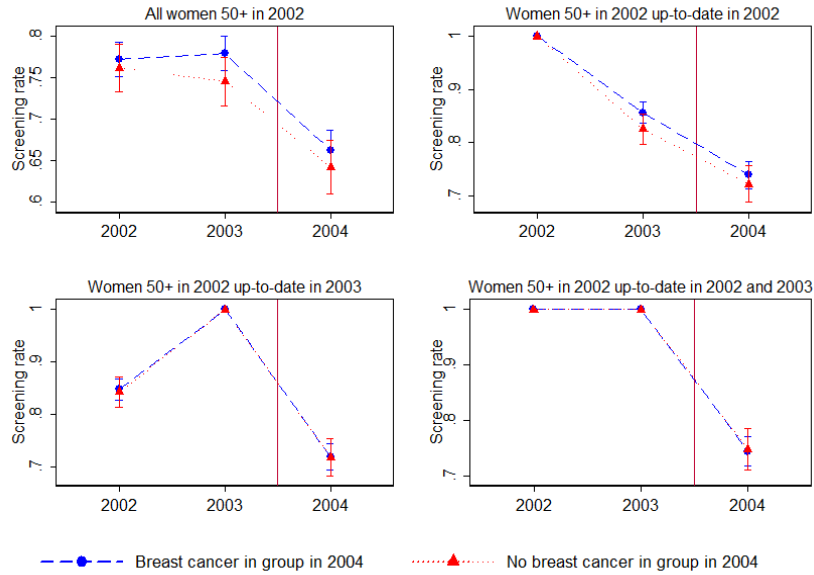
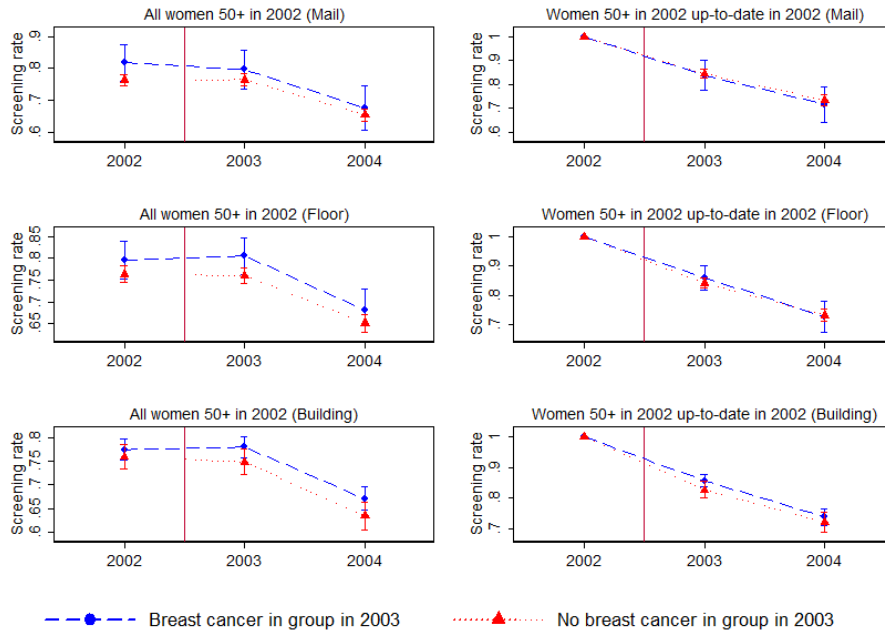


Figure 6. Screening rates by breast cancer occurrences in 2003: all groups



3.2 Identification

The goal of our empirical analysis is to identify the dynamic average treatment effect of experiencing breast cancer in the group of co-workers (i.e., the effect of treatment history \mathbf{T}_g^t) on the conditional probability of screening,

$$\Pr(Y_t = 1 | \mathbf{T}_g^t, \mathbf{X}) = \mathbb{E}[Y_t | \mathbf{T}_g^t, \mathbf{X}], \tag{1}$$

where \mathbf{X} are the conditioning variables. For any individual i we observe one of eight counterfactual outcomes at each point in time:

$$Y_{it} = Y_{it}(\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2) \quad \text{if } T_{g,i}^t = \mathcal{T}_0 \quad \text{and } T_g^{t-1} = \mathcal{T}_1 \quad \text{and } T_g^{t-2} = \mathcal{T}_2, \tag{2}$$

where $(\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2) \in \{0, 1\} \times \{0, 1\} \times \{0, 1\}$. Given the way we have defined the treatment, the two lagged treatment indicators, T_g^{t-1} and T_g^{t-2} , are as good as randomly assigned conditional on observables—this is ensured by the plausible fact that breast cancer arises randomly across otherwise identical groups, notably groups of the same size (Table 9). However, the current treatment indicator, $T_{g,i}^t$, is not. There are two sources of endogeneity. First, as illustrated above, $T_{g,i}^t$ depends on the timing of screening within year t . For instance, women who have a stronger unobserved preference for screening may tend to screen very early in the year. In those groups where someone is diagnosed with breast cancer, these women are systematically less likely to be treated in that year, and would differ in unobserved ways from those who are instead more likely to be treated because they screen later in the year. Second, we can perform only a partial adjustment for exact calendar dates of screening in 2002 and 2003. The reason is that although we observe diagnosis date for all breast cancer occurrences, we only observe the date a woman last screened as of December 31st, 2004, not all *screening dates* associated with one’s screening history (although the screening history itself in terms of *screening dummies*, Y_{2002} , Y_{2003} , and Y_{2004} , is of course observed). So, for instance, in 2002 and 2003 we cannot adjust the treatment indicator for those women who perform a mammogram every year (these are about 1/3 of the total), because for these women we know the exact day they screened in 2004 but not in 2002 and 2003. This is a non-classical measurement error that would bias naïve OLS estimates in an unpredictable direction.⁷

We address this endogeneity problem and the possible selection problems arising from the other limitations of the data discussed in Section 2 by employing a threefold empirical strategy. First, a baseline instrumental variable strategy; second, a robustness check on the baseline that allows for serial correlation in the decision to screen; third, an additional robustness check in the form of a duration analysis at the daily frequency on the 2004 portion of the sample.

⁷When there are multiple breast cancer occurrences in a group—which happens at the Floor and Building levels—the relevant calendar date for the correction is the day the first case was diagnosed.

3.2.1 Instrumental variable model

The baseline model relies on a “natural” instrumental variable strategy. Conceptually, the problem we are facing is similar to what happens when eligible individuals do not comply with a treatment. In our data there are women who are “eligible” for the contemporaneous treatment (i.e., have a co-worker diagnosed with breast cancer during the current year) and who do not “comply” with such treatment (i.e., screened before the co-worker was diagnosed). It is standard in these cases to use eligibility (“initial treatment”, or “intention to treat”) as an instrument for actual treatment (Joshua Angrist, Guido Imbens, and Donald Rubin, 1996). Specifically, we instrument $T_{g,i}^t$ (the joint event that a co-worker of i was diagnosed with breast cancer during year t and in that year i had not performed a mammogram by the day the co-worker was diagnosed with the malignancy) with a dummy equal to 1 if a co-worker of i was diagnosed with breast cancer during year t , and zero otherwise. Such dummy is denoted T_g^t . The properties of this instrument are such that it corrects both the correlation between timing of screening within the year and unobservable preferences for screening on the one hand, and the non-classical measurement error in timing of screening on the other. We will assume constant treatment effects, and estimate the following baseline linear probability model:

$$Y_{it} = \alpha \mathbf{t} + \beta \mathbf{x}_{it} + \gamma \mathbf{x}_{gt} + \tau_0 T_{g,i}^t + \tau_1 T_g^{t-1} + \tau_2 T_g^{t-2} + \theta_i + \varepsilon_{igt}, \quad (3)$$

where Y_{it} is the individual screening indicator, \mathbf{t} is a vector of year dummies, \mathbf{x}_{it} and \mathbf{x}_{gt} are time-varying individual and group covariates, respectively, τ_0 is the contemporaneous treatment effect, τ_1 and τ_2 are the lagged effects at 1 and 2 years distance, respectively, θ_i is an individual fixed effect, and ε_{igt} contains residual unobservables. In estimating this equation, we include in \mathbf{x}_{it} the time-varying comorbidity indicators, and in \mathbf{x}_{gt} their group-level averages. Furthermore, because the treatment variable is perfectly correlated within reference groups, we cluster standard errors within such groups.

3.2.2 Lagged dependent variable model

Even if the instrument corrects the endogeneity problem arising from the definition of the contemporaneous treatment indicator, $T_{g,i}^t$, one may worry that because of possible serial correlation in the decision to screen, the entire treatment history—not just $T_{g,i}^t$ —is actually endogenous: if the lagged dependent variable, Y_{it-1} , is part of the error term, then the latter is correlated with both T_g^{t-1} and T_g^{t-2} under specification (3). We will address this concern by including the lagged dependent variable in the RHS of (3) and then employing the Arellano-Bond estimator to identify the parameters. Therefore, we will also estimate the first-differenced equation,

$$\Delta Y_{it} = a + \beta \Delta \mathbf{x}_{it} + \gamma \Delta \mathbf{x}_{gt} + \delta \Delta Y_{it-1} + \tau_0 \Delta T_{g,i}^t + \tau_1 \Delta T_g^{t-1} + \tau_2 \Delta T_g^{t-2} + e_{igt}, \quad (4)$$

via instrumental variables, using Y_{it-2} and higher-order lags as instruments for ΔY_{it-1} . As for the other endogenous regressor, $\Delta T_{g,i}^t$, the previous instrumental variable strategy continues to apply so that ΔT_g^t is also employed as an instrument. Unfortunately, our panel has only three periods, and feasibility requires us to use only one lag, Y_{it-2} , as an instrument and to give-up two years of out of three (one year is lost because of first-differencing, the other because of the use of Y_{it-2} as instrument)⁸. Nonetheless, this is a useful robustness check.

3.2.3 Duration model

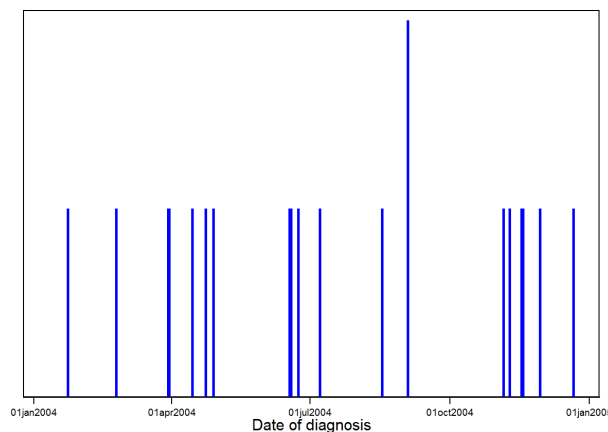
A cleaner way to circumvent the identification problems stemming from the definition of the contemporaneous treatment indicator $T_{g,i}^t$ and from the impossibility of fully adjusting for exact calendar dates of screening in 2002 and 2003 is to conduct a duration analysis at the daily frequency using only the 2004 portion of the sample—for this year, we have both calendar dates of breast cancer diagnosis and calendar dates of screening. We stress that, in addition, such a duration analysis addresses all of the potential selection problems arising from the unobservability of women who died or otherwise left the organization before 2004, and from the unobservability of the actual composition of Mail, Floor, and Building groups in 2002 and 2003, as discussed in Section 2: there are no selection issues (after the treatment is redefined at the daily frequency) in the 2004 portion of our sample. Therefore, we regard this duration analysis as a particularly important robustness check. To perform this check we fix the starting date at January 1st, 2004 and we estimate the hazard function of *not* screening during that year conditional on the daily treatment history. That is, approximately, the probability that a woman leaves between days d and $d + 1$ of 2004 the state where she is not up to date with her annual mammogram in that year, conditional on not having performed a mammogram by day d of that year, $h(d, T_g^d) \simeq \Pr(d \leq D_i < d + 1 | T_g^d)$, where $h(\cdot)$ is the hazard function, and T_g^d is the treatment indicator defined at the daily (d) rather than annual (t) frequency—i.e., $T_g^d = 1$ if on day d or earlier in 2004 a co-worker was diagnosed with breast cancer, and $T_g^d = 0$ otherwise. Figure 7 shows the calendar distribution of the 20 breast cancer diagnoses in 2004. These occurred in 20 different days of the year (except for September 3, 2004, when 2 cases were diagnosed) and are roughly uniformly distributed; 10 cases are diagnosed during the first half of the year, and the remaining 10 during the second half. Because we are only interested in how the event of a colleague diagnosed with breast cancer in one’s group on day d in 2004 shifts this hazard function, we employ the Cox estimator—i.e., we implement a proportional hazard model without estimation of the baseline hazard—with the daily treatment status T_g^d (the only covariate in our data set that changes at the daily frequency) as the only explanatory variable:

$$h(d, T_g^d) = \exp(\tau T_g^d) h_0(d), \tag{5}$$

⁸Therefore, we are actually employing the Anderson and Hsiao (1981) estimator.

where $h_0(d)$ is the baseline hazard, and τ is the treatment effect of interest. Parameter τ is, approximately, the variation of the hazard function for treated individuals relative to the baseline, i.e., $\tau \simeq [h(d, 1) - h(d, 0)]/h_0(d)$. The data is right-censored at December 31, 2004, because we obtained data only up to that point. We deal with this problem by assuming that the unobserved duration for those women not screening by December 31, 2004 is independent of this censoring point. This is not an ideal assumption, because the dates co-workers are diagnosed with breast cancer vary across treated individuals, but is the most convenient way of performing a robustness check in a duration framework.

Figure 7. Calendar distribution of breast cancer diagnoses in 2004



Notes: The histogram shows the calendar distribution of the 20 cases of breast cancer diagnosed in our sample in 2004. The shorter bars correspond to a single occurrence, the taller bar corresponds to a double occurrence on September 3, 2004.

4 Results

4.1 Baseline: instrumental variables estimates

The baseline results from 2SLS estimation of equation (3) are reported in Table 10. Column 1 shows a negative and significant impact response at the Mail group level:⁹ women with a co-worker diagnosed with breast cancer in a given year become 8 percentage points less likely to screen during that year. The coefficients on the lagged treatment indicators at the Mail group level are also negative, and indicate that the effect of the treatment is persistent: after 2 years treated women are still about 6 percentage points less likely to screen than women in

⁹The first-stage coefficient on the excluded instrument, T_g^t , is 0.86, with a robust standard error of 0.05. All first-stages in the following regressions are equally good and are not reported in the interest of space.

the control groups. Moving up at the Floor level (Column 2), the impact effect is essentially zero. We then see an increasing (in absolute value) negative lagged effect, which becomes significant after two years, although still smaller than at the Mail level. These different impact and lagged effects indicate that there is no attenuation bias when using the smaller Mail groups. On the contrary, the effect is weaker in the larger Floor groups, where distance from the woman diagnosed with cancer is, on average, higher. The fact that after two years we detect a negative effect at the Floor level too seems consistent with a diffusion process whereby information about the new diagnosis first reaches the closest co-workers in Mail groups and then those farther away in the larger Floor groups. Therefore, some attenuation bias may be present for the lagged response, and so the impact effect may be even more persistent than Table 10 indicates. Moving further up at the Building level (Column 3) we find a small positive effect, which is statistically significant at lag 1. This is driven by 2003 occurrences, as Figure 6 suggests. Such positive effect is puzzling, for two reasons. First, the treatment is relatively large and negative at the Mail level and relatively small and negative at the Floor level. While one can envision stories that rationalize a reversal at the higher, Building level of spatial aggregation, such a reversion seems hard to rationalize in a plausible way. Second, the effect at the Building group level for breast cancer occurrences in 2004 seems to be zero (see Figure 5). A reversal in time seems even less plausible. We believe this positive effect at the Building group level is a manifestation of the misclassification of treatment and control groups in a finite sample when expanding the spatial scope of reference groups. As Table 8 indicates, because of the large share of treated women in Building groups, such misclassification can be severe at this level of spatial aggregation.

Table 10. Results: baseline IV

	1	2	3
	Y_t	Y_t	Y_t
$T_{g,i}^t$	-0.082*	-0.009	0.024
	(0.033)	(0.023)	(0.027)
T_g^{t-1}	-0.069*	-0.029	0.051**
	(0.029)	(0.024)	(0.021)
T_g^{t-2}	-0.058	-0.047*	0.035
	(0.032)	(0.021)	(0.024)
Reference Group	Mail	Floor	Building
Age	50+	50+	50+
Observations	7,578	7,578	7,578
Individuals	2,710	2,710	2,710

Notes: Linear probability models, two-stage within estimator. In addition to individual fixed effects all regressions include year dummies, comorbidity indicators, and the group-level average of these. Robust standard errors clustered at the group level are given in parentheses; * significant at 5%; ** significant at 1%.

Table 11 reports the results of the baseline estimation using the age group 41–49 only. When focusing on this younger group we can employ only two years of data because screening information for the 41–49 group is missing in 2002. The coefficients are now generally positive, although the estimated standard errors are quite large, despite sample size is comparable to the age group 50 and older. This result begs the questions of why estimates in the age group 41–49 are so imprecise relative to the age group 50 and older, and why they are possibly positive. Apart from the fact that shrinking the longitudinal dimension of the panel from three to two periods reduces the variability exploited by the within estimator, thus leading to more imprecise estimates, we believe this happens because, as discussed in Section 2, the age group 41–49 is not subject to the clear-cut guidelines of the age group 50 and older. This is reflected in the substantially lower screening rates reported in Table 4. It is also possible that young women respond in a different way than older ones for other reasons, an open question for future research.¹⁰

Table 11. Results: baseline IV, age group 40–49

	1	2	3
	Y_t	Y_t	Y_t
$T_{g,i}^t$	0.061 (0.057)	0.042 (0.038)	0.030 (0.054)
T_g^{t-1}	0.044 (0.049)	0.024 (0.034)	0.059 (0.036)
T_g^{t-2}	-0.018 (0.045)	0.020 (0.032)	0.004 (0.041)
Reference Group	Mail	Floor	Building
Age	40–49	40–49	40–49
Observations	7,453	7,453	7,453
Individuals	3,546	3,546	3,546

Notes: Linear probability models, two-stage within estimator. In addition to individual fixed effects all regressions include year dummies, comorbidity indicators, and the group-level average of these. Robust standard errors clustered at the group level are given in parentheses; * significant at 5%; ** significant at 1%.

4.2 Robustness checks: lagged dependent variable and duration

The Arellano-Bond estimates from equation (4) are reported in Table 12a. At the Mail and Floor group levels, this table reproduces the pattern of Table 10, with slightly larger coeffi-

¹⁰A positive effect on young women would be consistent with the “Jade Goody effect”. After Jade Goody (a British reality TV celebrity) was diagnosed with cervix cancer in 2007, newspapers reported an *increase* in smear tests take-up rates among young British women. We are grateful to Michele Belot for pointing us to this story.

cients. At the Building group level, point estimates are now generally negative and smaller, in line with what we see moving farther away from Mail groups. Therefore, this alternative specification and estimation method confirm, by and large, the baseline results, despite the non-trivial differences between specifications (3) and (4) and the associated instruments. In particular, it is remarkable that despite the large drop in sample size we obtain at the Mail group level point estimates of the same sign and comparable magnitude that are still significantly different from zero at the 10% confidence level. The fact that the coefficient on the ΔY_{t-1} term is insignificant, of course, explains why the two sets of estimates agree: after netting out individual fixed effects there is no significant residual serial correlation in the dependent variable.

Table 12a. Robustness: Arellano-Bond

	1	2	3
	ΔY_t	ΔY_t	ΔY_t
$\Delta T_{g,i}^t$	-0.103 (0.063)	-0.033 (0.044)	-0.033 (0.088)
ΔT_g^{t-1}	-0.099 (0.055)	-0.052 (0.040)	0.010 (0.054)
ΔT_g^{t-2}	-0.089* (0.046)	-0.079* (0.034)	-0.044 (0.060)
ΔY_{t-1}	-0.021 (0.038)	-0.020 (0.038)	-0.009 (0.065)
Reference Group	Mail	Floor	Building
Age	50+	50+	50+
Observations	2,288	2,288	2,288

Notes: Linear probability models, two-stage least squares estimator. All regressions include year dummies, individual comorbidity indicators, and the group-level average of these. Robust standard errors clustered at the group level are given in parentheses; * significant at 5%; ** significant at 1%.

The duration estimates from equation (5), i.e., estimates of parameter τ , are reported in Table 12b. Although this procedure leads to very imprecise point estimates, these confirm the baseline results. In particular, having a co-worker diagnosed with breast cancer at a given date in 2004 reduces, at the Mail group level, the conditional probability of testing by 15.7% relative to the baseline. Given a baseline screening rate of 63.8% in 2004 (Table 4), this means $.157 \times .638 = 10$ percentage points at the Mail group level, roughly the baseline instrumental variable estimate (and exactly the Arellano-Bond estimate). At the Floor and Building levels, instead, we see virtually no effects. This, too, is consistent with the absence of significant contemporaneous effects at these higher levels of aggregation in the baseline

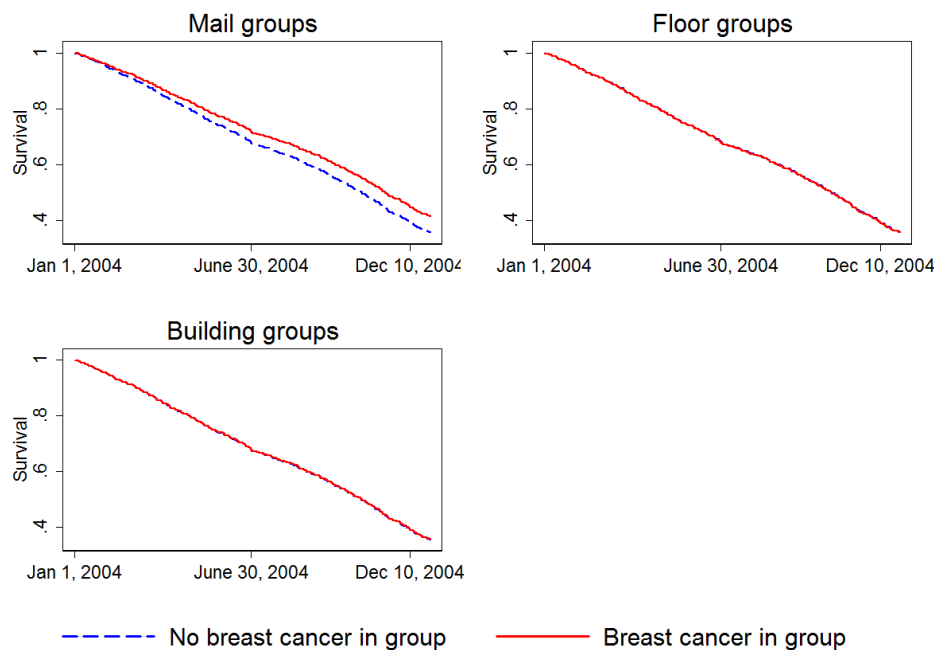
and Arellano-Bond estimates. Figure 8 shows the associated conditional survival functions by treatment status: only at the Mail group level we see that the two functions diverge: at this level, treated individuals are more likely to remain in the “untested” state at any calendar date; at the Floor and Building levels, instead, the two functions are effectively overlapped.

Table 12b. Robustness: duration model

	1	2	3
T_g^d	$h(d, T_g^d)$	$h(d, T_g^d)$	$h(d, T_g^d)$
	-0.157	0.005	-0.003
	(0.171)	(0.077)	(0.052)
Reference Group	Mail	Floor	Building
Years	2004	2004	2004
Age	50+	50+	50+
Individuals	3,136	3,136	3,136

Notes: Cox proportional hazard model at the daily frequency, on the 2004 portion of the sample.

Figure 8. Survival function in 2004, by treatment status that year



Notes: The figure shows the survival (probability of not screening) functions at the daily frequency in groups with and without breast cancer occurrences, based on the duration model in (5) and the associated estimates in Table 12b.

4.3 Heterogeneity

We investigated the presence of heterogenous responses. First, we used information on breast cancer severity. Table 13 reports the results at the Mail group level, using two alternative definitions of “mild cases”: stage 0 and grades 1–2, respectively (see Table 5). The point estimates are greatly reduced when a colleague is diagnosed with stage 0 breast cancer, while for grades 1–2 they become noisy. On the other hand, for non-mild cases we see point estimates of a magnitude similar to the average effect, although generally not significant. This suggests that the average effect we identify is driven by these non-mild cases. These may be thought of as the “real” treatment, or a stronger treatment: non-mild cases are unlikely to go unnoticed—they have more prominent effects, such as longer absences from work, aesthetic effects of chemotherapy, etc. Next, we used professional information. A group of interest are medical doctors and nurses. This group should be more knowledgeable about health issues—because of training and experience—and so may be expected to respond differently compared to other employees. Table 14 reports estimates from the baseline model (3) in two different samples: women of age 50 and older who are either medical doctors or nurses (MD/nurse) and all other women in the same age group. We continue to focus on Mail groups. The table shows that a negative effect is present (and actually reinforcing over time) only for women who do not belong to the MD/nurse category. In this category, instead, no significant effect is found, and the point estimates are actually positive after one year. This difference is not a reflection of differences in income and education: regressions by salary classes (not reported) yield no significantly different treatment effects. Although the MD/nurse vs. the rest heterogeneity may reflect a reduced sample size (medical doctor and nurses are just 1/4 of the total) it is potentially important as a source of information about underlying mechanisms—we will return on this in Section 5.

We have explored additional margins of possible heterogeneity. First, we have looked at how younger and older women in the 50+ age group respond to the treatment (do younger women in the age range 50-55 respond differently from those in the 56-65 range?). Second, we have investigated the presence of differences associated with the age of the woman diagnosed with breast cancer (do women respond differently when a co-worker is diagnosed before age 40 compared to when she is diagnosed after age 60?). Third, the interaction of these two (do women respond differently to cases of breast cancer involving a co-worker of similar age?). Fourth, we have exploited information on tenure to understand whether having being co-workers for a longer time affects the response (do women respond differently when both them and the woman diagnosed with breast cancer have been at the organization for, say, more than 10 years?). None of these yields statistically different treatment effects or systematic informative patterns in the point estimates like in Tables 13 and 14. Finally, we have investigated whether multiple occurrences in the same group (“more intense” treatment, see Table 8) make a difference. We did this by adding to the RHS of equation (3) indicators for whether there is more than one occurrence in that group, at the usual three lags. The

coefficients on these additional indicators are never statistically different from zero, and the coefficients on $T_{g,i}^t$, T_g^{t-1} , and T_g^{t-2} are unaffected.

Table 13. Results: severity of cancer

	1	2	3	4
	Y_t	Y_t	Y_t	Y_t
$T_{g,i}^t$	-0.013 (0.074)	0.347 (0.394)	-0.034 (0.041)	-0.069 (0.037)
T_g^{t-1}	0.014 (0.065)	0.017 (0.176)	-0.081* (0.033)	-0.050 (0.032)
T_g^{t-2}	0.008 (0.070)	0.163 (0.095)	-0.062 (0.038)	-0.059 (0.034)
Severity	Mild	Mild	Non-mild	Non-mild
Mild defined as	stage 0	grades 1-2	stage 0	grades 1-2
Observations	6,686	6,383	7,271	7,369
Individuals	2,408	2,312	2,613	2,636

Notes: Linear probability models, two-stage within estimator. Estimates are for women aged 50+ when reference groups are Mail groups. See Section 2 for the definitions of “stage” and “grade”. Model specification, standard errors, and significance level like in previous tables.

Table 14. Results: doctor and nurses vs. the rest

	1	2
	Y_t	Y_t
$T_{g,i}^t$	-0.036 (0.069)	-0.075* (0.038)
T_g^{t-1}	0.038 (0.059)	-0.087* (0.037)
T_g^{t-2}	0.117 (0.070)	-0.101** (0.037)
Job title	MD/nurse	Other
Observations	1,890	5,768
Individuals	671	2,066

Notes: Linear probability models, two-stage within estimator. Estimates are for women aged 50+ when reference groups are Mail groups. “MD/nurse” indicates employees who are either medical doctors or nurses. Model specification, standard errors, and significance level like in previous tables.

4.4 Placebo tests

Finally, we perform three placebo tests to check against the possibility that our IV estimates reflect spurious correlations. First, we try to predict the probability of screening today using *future* breast cancer occurrences in the reference group. Second, we change the definition of treatment from “occurrences of breast cancer in the reference group” to “occurrences of *non*-breast cancer in the reference group, involving both women and men.” Third, we randomly reallocate actual breast cancer occurrences across reference groups.

Table 15 presents the results of the first test. In this table T_g^{t+1} is an indicator that takes value 1 at any point in time if someone is diagnosed with breast cancer in group g during the following year, and zero otherwise. Similarly, T_g^{t+2} takes value 1 if someone is diagnosed in two years from now, and zero otherwise. To perform this test we pulled from the tumor registry cancer information for 2005 as well. Therefore, when using only T_g^{t+1} as a regressor, we are still using screening behavior from 2002 to 2004 (i.e., the full panel). When adding T_g^{t+2} we lose year 2004. The table shows that future occurrences have no predictive power on current screening behavior. Coefficients are typically very small (sometimes essentially zero) and never statistically significant.

Table 15. Placebo 1: breast cancer diagnosed in the future

	1	2	3	4	5	6
	Y_t	Y_t	Y_t	Y_t	Y_t	Y_t
T_g^{t+1}	0.035 (0.024)	-0.003 (0.016)	-0.020 (0.015)	0.033 (0.037)	-0.018 (0.028)	0.030 (0.054)
T_g^{t+2}	-	-	-	0.003 (0.041)	-0.020 (0.027)	0.020 (0.029)
Ref. group	Mail	Floor	Build.	Mail	Floor	Build.
Age	50+	50+	50+	50+	50+	50+
Obs.	7,578	7,578	7,578	4,448	4,448	4,448
Individuals	2,710	2,710	2,710	2,224	2,224	2,224

Notes: Linear probability models, two-stage within estimator. In addition to individual fixed effects all regressions include year dummies, comorbidity indicators, and the group-level average of these. Robust standard errors clustered at the group level are given in parentheses.

Table 16 illustrates the second placebo. There are 144 cases of non-breast cancer affecting women, and 79 cases affecting men in our sample between 2000 and 2004. Does the propensity to perform a mammogram responds to these non-breast cancer events like it does for cases of breast cancer? Table 16 shows the answer is “no”. This table reports the results from the baseline specification after using such non-breast cancer occurrences to define the treatment.

A “~” over the treatment indicator denotes the modified treatment.¹¹ The magnitude of the coefficients is greatly reduced, often down to zero, and they are never statistically significant. These results are unchanged if we include breast cancer occurrences affecting only other women in the reference group or if we restrict to female-specific occurrences—i.e., uterus, vagina, vulva, and ovary cancer. Therefore, the significant negative relationship we identify is a unique relationship between the event that a colleague was diagnosed with breast cancer and a woman’s propensity to perform a mammography.

Table 16. Placebo 2: non-breast cancer diagnosed in group

	1	2	3
	Y_t	Y_t	Y_t
\tilde{T}_g^t	-0.019 (0.021)	-0.005 (0.017)	0.029 (0.019)
\tilde{T}_g^{t-1}	-0.003 (0.024)	-0.006 (0.018)	0.005 (0.024)
\tilde{T}_g^{t-2}	-0.026 (0.021)	-0.020 (0.016)	-0.006 (0.020)
Ref. group	Mail	Floor	Building
Age	50+	50+	50+
Observations	7,578	7,578	7,578
Individuals	2,710	2,710	2,710

Notes: Linear probability models, two-stage within estimator. A “~” indicates the treatment is redefined as exposure to non-breast cancer occurrences. In addition to individual fixed effects all regressions include year dummies, comorbidity indicators, and the group-level average of these. Robust standard errors clustered at the group level are given in parentheses.

Finally, Table 17 presents the results of the third and final placebo test—random reallocation of breast cancer occurrences across reference groups. We proceeded as follows. Each woman diagnosed with breast cancer between 2000 and 2004 is randomly assigned to one (and only one) reference group, which may be her own actual group—this is a quite unlikely event, of course. All women in such reallocation groups are then classified as treated at different lags, with the usual adjustment for screening dates and instrument for the contemporaneous treatment—exactly like we do for the actual treatment. In the table, a “~” over the treatment indicator denotes it is a placebo treatment from such random reallocation.

¹¹Because for these non-breast cancer occurrences we don’t know the exact date of the diagnosis, we cannot perform the usual adjustment for timing of screening. Therefore, Table 20 contains fixed effects-OLS estimates, which should be interpreted as produced by the reduced form of baseline model (3). With a first-stage coefficient of almost 0.9 in the baseline model, these are only slightly below the associated 2SLS estimates.

The dynamic treatment effect is then estimated on this partially simulated data set, like in the baseline model. This procedure is repeated 1,000 times at the Mail, Floor, and Building group level. The coefficients and standard errors reported in Table 17 are the averages of the respective 1,000 estimates. The average is essentially zero at all levels of aggregation.

We report in Figure 7 the distribution of the 1,000 underlying point estimates, at the three different lags in the three alternative reference groups. The vertical line in each panel is the corresponding baseline point estimate from the actual treatment—i.e., the point estimates in Table 10.

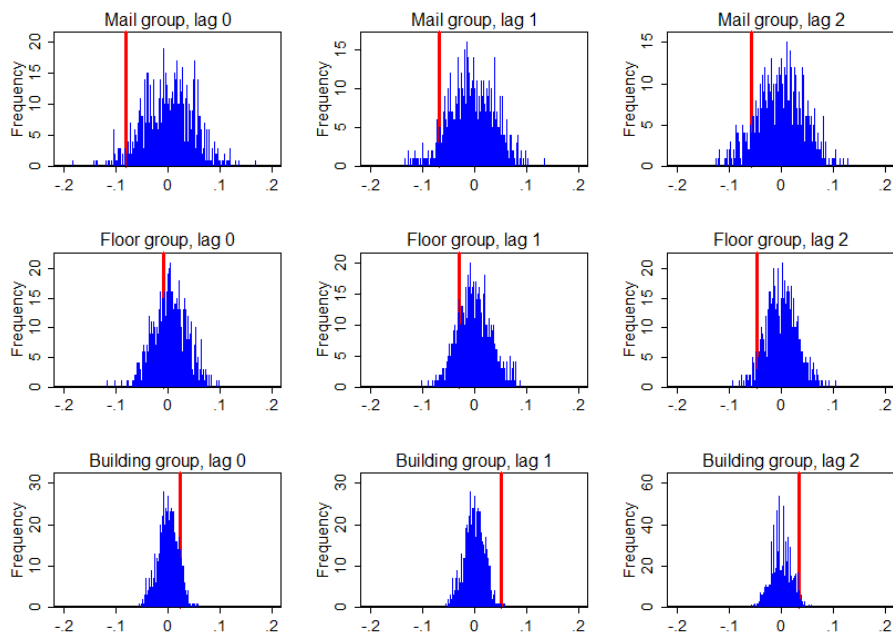
Table 17. Placebo 3: random reallocation of cases across groups

	1	2	3
	Y_t	Y_t	Y_t
$\tilde{T}_{g,i}^t$	0.002 (0.047)	0.004 (0.028)	0.001 (0.020)
\tilde{T}_g^{t-1}	-0.006 (0.046)	-0.001 (0.028)	0.000 (0.019)
\tilde{T}_g^{t-2}	-0.001 (0.047)	-0.002 (0.028)	0.001 (0.019)
Ref. group	Mail	Floor	Building
Observations	7,658	7,658	7,658
Individuals	2,737	2,737	2,737

Notes: Baseline estimates, women 50+ (see Table 10). A “~” indicates the treatment is defined by random reassignment across reference groups of all women diagnosed with breast cancer between 2000 and 2003. Averages of 1,000 replications.

These figures show that our baseline point estimates tend to be located in the tails of the distribution of estimates we would obtain “by chance”. Furthermore, it is quite hard to obtain significant estimates “by chance”. Table 18 reports the fraction of the 1,000 simulated estimates that are significantly different from zero, and Figure 8 shows the distribution of these significant coefficients. The table shows that less than 5% of estimates at the Mail group level are significantly different from zero, and the figure shows that these are polarized at the extremes of the coefficients distribution. This has an obvious explanation: for a given level of variability in the data, random draws generating large (in absolute value) estimated coefficients tend to be associated with significant estimates. The point estimates from the baseline specification tend to be close to those generated by these extreme draws. However, given the very low incidence of these cases reported in Table 18, we should have been extremely “lucky” to obtain our results “by chance”.

Figure 7. Placebo 3: distribution of point estimates



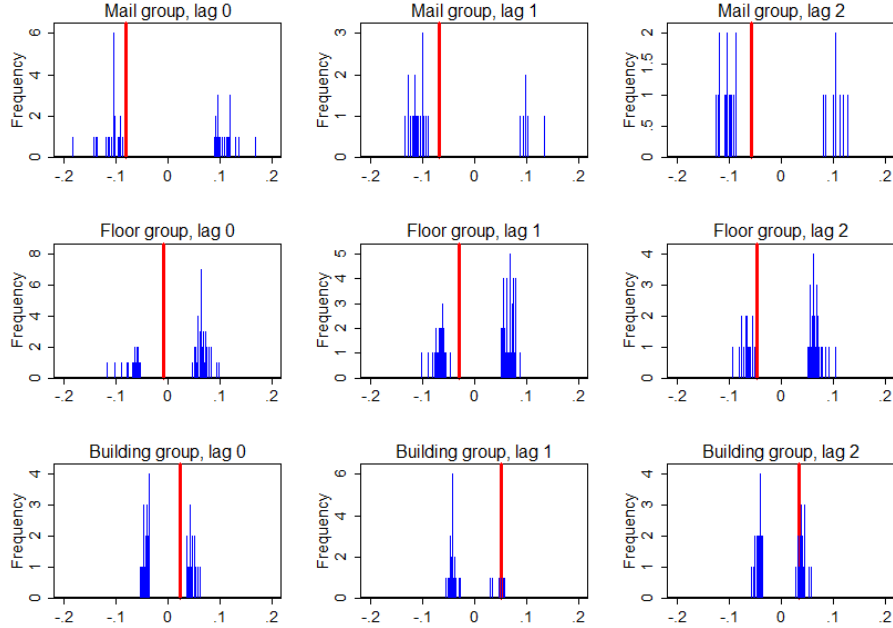
Notes: Frequency distribution of the 1,000 point estimates underlying the averages reported in Table 18. The vertical line in each panel is the corresponding baseline point estimate from the actual treatment—i.e., the point estimates in Table 10.

Table 18. Placebo 3: Fraction significant estimates in simulation

	Mail	Floor	Building
Lag 0	0.047	0.064	0.036
Lag 1	0.028	0.068	0.032
Lag 2	0.023	0.050	0.037

Notes: Fraction of the 1,000 simulated estimates, at a given lag and by reference group, that are significantly different from zero

Figure 8. Placebo 3: distribution of significant point estimates



Notes: Frequency distribution of point the 1,000 estimates underlying the averages reported in Table 18 which are statistically different from zero.

5 Discussion: mechanisms

Our analysis shows that the event that a colleague is diagnosed with breast cancer matters: it generates an updating of one’s information set or changes in one’s constraints (or both) that, in turn, lead to a reduced rate of mammogram uptake through the employer’s health plan. This result is rationalized by a number of different mechanisms, which we briefly discuss here. This discussion is necessarily speculative, because the data does not allow us to pin down tightly the underlying mechanisms. Therefore, this remains to a large extent an open question—an important one, we believe—for future research. However, both the evidence we have produced and existing evidence from “mechanism experiments” (in the jargon of Jens Ludwig, Jeffrey R. Kling, and Sendhil Mullainathan, 2011) conducted in the economic and medical literatures can be used to assess the plausibility of any particular story. To this purpose, an important part of our findings is that women do not learn from non-breast cancer occurrences what they learn from breast cancer cases, or that their constraints are not altered in the same way by these two different events (Table 16). In what follows we refer to this as the “dissimilarity” property of the underlying mechanism(s).

The first interpretation we suggest is motivated by an important caveat: our findings indicate a reduced propensity to use *employer-provided* mammograms. What one may learn when a co-worker is diagnosed with breast cancer is that a breast cancer diagnosis that is known to the employer has adverse consequences for one’s career and, possibly, health coverage. By avoiding screening at the employer’s facility one prevents the employer from learning information that may lead to such adverse consequences. Therefore, it is possible that our results are driven by a shift from regular screening at the workplace to regular screening elsewhere or to regular non-clinical exams¹² rather than to screening avoidance. This is a compelling interpretation. In this case the “dissimilarity” property can be explained by the uninformative nature of a co-worker’s experience with a tumor different from breast cancer with respect to the career and health coverage consequences of a breast cancer diagnosis.

Another story that is strictly economic derives from a time-constraint argument. When someone is diagnosed with breast cancer (thus reducing working time or work effort—see Section 2), his or her colleagues may choose to (or have to) temporarily work more in order to compensate. By working more, they have less time for health care and may delay breast screening too. For instance, it has been shown that during upturns, when people work more, there is a reduction in the propensity to engage in healthy behaviors and an increase in mortality rates (Christopher Ruhm, 2000). However, it should be observed that a screening mammography takes no more than 15–30 minutes, and travel time in our sample is virtually zero, because the mammogram is performed right at the workplace. Furthermore, given the way reminders are sent and appointments are scheduled (see Section 2), it’s not easy to forget about the test even during an unusually busy year. Finally, this interpretation seems to us very hard to reconcile with the “dissimilarity” property—working time and work effort are reduced following any diagnosis of a serious tumor requiring treatment.

A third mechanism is loss of confidence in either the test itself or the treatment of breast cancer. For instance, by experiencing breast cancer at the workplace a woman may learn that it is not screening per se that saves her life (Welch and Frankel, 2011) and that she is in fact overscreening, i.e., screening at a rate that may lead to diagnoses (and associated treatment) of tumors that would not have led to clinical symptoms. Archie Bleyer and Gilbert Welch (2012) estimate that of all breast cancers diagnosed in the U.S. since the end of the 1970s, almost 1/3 resulted from overdiagnoses and so induced unnecessary treatment. In other words, the negative effect we identified may represent an optimal downward adjustment of one’s screening rate—as we have documented in Section 2, women 50 years of age and older in our sample have screening rates 14.5 percentage points above the corresponding U.S. average. In this case the “dissimilarity” property is easily explained, because this channel is specific to breast screening. However, if this is a relevant mechanism then it is puzzling that we do not detect any significant effect for doctors and nurses (Table 14), who plausibly understand the overdiagnosis argument better.

¹² Archie Bleyer and Gilbert Welch (2012) report that only 60% of cases are diagnosed via clinical screening.

More “behavioral” mechanisms may also be at work. For instance, Matthew Rabin (2002) shows that if people observe a binary signal and if they believe such signal is generated by a random draw without replacement but they know the rate at which signals are drawn, then they expect the next signal to be negatively correlated with the one they just observed—a phenomenon labelled “law of small numbers”. Applied to our problem, a woman observing that a colleague was diagnosed with breast cancer in the group would think that it is unlikely that someone else in the group has the disease. This would lead to a lower perceived probability of having breast cancer and thus to a lower propensity to screen in an otherwise standard expected utility model. However, the medical literature consistently finds that increasing the saliency of breast cancer leads, if anything, to a *higher* perceived probability of developing the disease. Montgomery *et al.* (2003) find that having a friend diagnosed with breast cancer increases the perceived risk of the disease. Similarly, Kathy Helzlsouer *et al.* (1994) find that experiencing cancer through family and friends is significantly associated with the perceived risk of cancer in general. Penelope Hopwood (2000) and both Evans *et al.* (1993) and Caryn Lerman *et al.* (1995) report that experience of breast cancer in the family leads to substantial overestimation of the probability of developing the disease.

A fifth, possible story is a different “behavioral” mechanisms labelled “information aversion”. Information aversion is a property of preferences whereby these also depends on beliefs—in addition to outcomes. Such dependence is formalized via an anticipatory component of utility, which allows for utility from “thinking” today about what could happen in the future: a woman has lower anticipatory utility the higher the belief that she is sick, because this means that negative health events such as painful treatment and death, or negative economic events such as a prejudiced career are more likely. This has implications for the propensity to collect information about the likelihood of such events: by screening, one may learn that she is sick and thus experience the worst possible anticipatory outcome. When anticipatory utility is convex in the belief of being sick, an individual is information-averse and may reject health information (Andrew Caplin and John Leahy, 2001; Andrew Caplin and Kfir Eliaz, 2003; Botond Köszegi, 2003, 2010). This happens when the expected utility loss from anticipation is larger than the expected utility gain from receiving medical treatment. There are two ways an information averse woman may screen less in response to the event that a colleague is diagnosed with breast cancer. First, this event makes the disease more salient and induces a higher perceived probability of developing the disease (starting from a sufficiently low prior). Second, experiencing breast cancer at the workplace induces an upward revision of the anticipatory utility loss when breast cancer is diagnosed—that is, a woman realizes that the physical, psychological, or economic consequences of being diagnosed with breast cancer are worse than she thought. The first way, however, seems inconsistent with both the fact that screening rates increase with age (i.e., when the probability of developing the disease increase), as well as additional external evidence. Foster *et al.* (2007), for instance, show that within three years from genetic testing, women who learned that they carry a genetic mutation responsible for breast cancer (i.e., women who

presumably revised upward the probability of developing breast cancer) have a 89% rate of mammogram use. Those who learned they do not carry the mutation, instead, have a 46% rate. The second way, instead is consistent with the evidence reviewed by Botond Kőszegi (2003), which indicates that many high-risk women refuse free information on whether they have a genetic susceptibility to breast cancer, and that a large fraction of women with breast cancer symptoms delay seeking help, the delay being longer the worse the symptom is getting. In this context, the “dissimilarity” property is naturally explained by the tight link between the saliency of breast cancer and the decision to screen for breast cancer. Therefore, a model with information aversion and unfavorable update of one’s beliefs about the anticipatory consequences of a diagnosis seems a compelling interpretation too.

This is far from being an exhaustive list. At the very least, it suggests that our findings can be rationalized within standard economic models, and it sketches an agenda for future research.

6 Conclusions

In this paper we have exploited a natural experiment occurring repeatedly over time in a unique panel data set of employees at a large medical organization in the U.S. to study the reaction of women 50 years of age or older—a demographic group that is subject to unambiguous breast screening guidelines—to the event that a colleague is diagnosed with breast cancer. It is plausible that such events soon become known—through workplace social interactions—to co-workers in one’s department. We have estimated a dynamic treatment effect model at the annual frequency, and we have identified a negative and persistent response. This negative effect is confirmed by alternative specifications based on allowing for serial correlation in screening behavior and on estimating the effect of the treatment on the hazard of not screening at the daily frequency. Furthermore, the effect vanishes in placebo experiments. The negative effect we identify seems driven by non-mild cases of breast cancer and by the behavior of women who are neither medical doctors nor nurses. We have discussed a number of economic mechanisms that rationalize our findings, and we have assessed their plausibility. Information aversion and preventing the employer from learning information that may have adverse career consequences are two compelling interpretations of our findings. We are well aware of the fact that the external validity of our study is limited. Women in our sample, in particular, all work in the health industry and have full coverage. We are also well aware that our inability to discriminate among alternative mechanisms makes the generalization of our findings even harder. Despite these limitations our results point to potentially important sources of screening avoidance in social contexts, and indicate correspondingly important avenues for future research.

References

- Anderson, T.W., and Cheng Hsiao (1981). "Estimation of Dynamic Models with Error Components." *Journal of the American Statistical Association*, 75, 598-606.
- Angrist, Joshua, Guido Imbens, and Donald Rubin (1996). "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434), 444-455.
- Bleyer, Archie, and Gilbert Welch (2012). "Effect of Three Decades of Screening Mammography on Breast-Cancer Incidence." *New England Journal of Medicine*, 367(21), 1998-2005.
- Blume, Lawrence, William Brock, Steven Durlauf, and Yannis Ioannides (2011). "Identification of Social Interactions." In Jesse Benabib, Alberto Bisin, and Matthew Jackson (eds.), *Handbook of Social Economics*. Volume 1B.
- Bradley, Cathy, David Neumark, Heather L. Bednarek, and Maryjean Schenk (2005). "Short-term effects of breast cancer on labor market attachment: results from a longitudinal study." *Journal of Health Economics*, 24, 137-160.
- Breen, Nancy, Kathleen A. Cronin, Helen I. Meissner, Stephen H. Taplin, Florence K. Tangka, Jasmin A. Tiro, Timothy S. McNeel (2007). "Reported drop in mammography. Is this cause for concern?" *Cancer*, 109(12), 2405-2409.
- Caplin, Andrew, and John Leahy (2001). "Psychological Expected Utility Theory and Anticipatory Feelings." *Quarterly Journal of Economics*, 116(1), 55-79.
- Caplin, Andrew, and Kfir Eliaz (2003). "AIDS Policy and Psychology: A Mechanism-Design Approach." *RAND Journal of Economics*, 34(4), 631-646.
- Dufló, Esther, and Emmanuel Saez (2003). "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics*, 815-842.
- Elixhauser, A., Steiner, C., Harris, D., Coffey, R. (1998). "Comorbidity Measures for Use with Administrative Data." *Medical Care*. 36(1), 8-27.
- Evans, D., L. Burnell, P. Hopwood, and A. Howell (1993). "Perception of risk in women with a family history of breast cancer." *British Journal of Cancer*, 67(3), 612-614.
- Foster C., M. Watson, R. Eeles, D. Eccles, S. Ashley, R. Davidson, J. Mackay, P. Morrison, P. Hopwood, D. Evans (2007). "Predictive genetic testing for BRCA1/2 in a UK clinical cohort: three-year follow-up." *British Journal of Cancer*, 96(5), 718-24.

- Hassett, Michael, James O'Malley, and Nancy Keating (2009). "Factors Influencing Changes in Employment Among Women With Newly Diagnosed Breast Cancer." *Cancer*, 115, 2775-2782.
- Helzlsouer, Kathy, D Ford, R. Hayward, M. Midzenski, and H. Perry (1994). "Perceived Risk of Cancer and Practice of Cancer Prevention Behaviors Among Employees in an Oncology Center.", *Preventive Medicine*, 23(3) 302-308.
- Hopwood, Penelope (2000). "Breast cancer risk perception: what do we know and understand?" *Breast Cancer Research*, 2(6), 387-391.
- Köszegi, Botond (2003). "Health Anxiety and Patient Behavior." *Journal of Health Economics*, 22(6), 1073-1084.
- Köszegi, Botond (2010). "Utility from Anticipation and Personal Equilibrium." *Economic Theory*, 44(3), 415-444.
- Lerman, Caryn, Edward Lustbader, Barbara Rimer, Mary Daly, Suzanne Miller, Colleen Sands, and Andrew Balshem (1995). "Effects of Individualized Breast Cancer Risk Counseling: a Randomized Trial." *Journal of the National Cancer Institute*, 87(4), 286-292.
- Ludwig, Jens, Jeffrey R. Kling, and Sendhil Mullainathan (2011). "Mechanism Experiments and Policy Evaluations." *Journal of Economic Perspectives*, 25(3), 17-38.
- Manski, Charles. (1993), "Identifications of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies*, 60, 531-542.
- Montgomery G., J. Erlich J, T. DiLorenzo, and D. Bovbjerg (2003). "Family and friends with disease: their impact on perceived risk." *Preventive Medicine*, 37(3), 242-249.
- Moss S., H. Cuckle, A. Evans, L. Johns, M. Waller, B. Bobrow (2006). "Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial." *Lancet*, 368(9552), 2053-2060.
- Oster, Emily, Ira Shoulson, and Ray Dorsey (2013) "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease." *American Economic Review*, 103(2), 804-830.
- Rabin, Matthew (2002). "Inference by Believers in the Law of Small Numbers." *Quarterly Journal of Economics*, 117(3), 775-816.
- Ruhm, Christopher (2000). "Are Recessions Good for Your Health?" *Quarterly Journal of Economics*, 115 (2), 617-650.

Welch, Gilbert, and Brittney Frankel (2011). “Likelihood That a Woman With Screen-Detected Breast Cancer Has Had Her ‘Life Saved’ by That Screening.” *Archives of Internal Medicine*, forthcoming.

USPSTF (2008). “Screening for Prostate Cancer: U.S. Preventive Services Task Force Recommendation Statement.” *Annals of Internal Medicine*, 149(3):185-191.



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
<http://www.dse.unibo.it>