QUANTITATIVE ANALYSIS UNIT

QAU

# DEMONSTRATION EFFECTS IN PREVENTIVE CARE

Working Paper No. QAU07-7

Ritesh Banerjee
*Mayo Clinic College of Medicine*

Ethan Cohen-Cole
*Federal Reserve Bank of Boston*

Giulio Zanella
*University of Siena*

FEDERAL RESERVE BANK OF BOSTON™

QAU

WPS No. QAU07-7

# Demonstration Effects in Preventive Care*

Ritesh Banerjee†, Ethan Cohen-Cole‡, and Giulio Zanella§

September 5, 2007

**Abstract**

Using a unique dataset composed of female employees at a large medical organization, this paper explores the role of social interactions among female co-workers and neighbors in the decision to obtain breast cancer screening exams. In our theoretical framework, the experience of other women is salient because it alters the tolerance for ambiguity about their own vulnerability, via a *comparative ignorance* effect. We find that the social multiplier ranges from 2 to 3: the equilibrium effect of an exogenous shock that impacts the probability of performing a mammogram is two to three times the shock itself. We perform a number of checks: among other things, these reveal (in agreement with the model and our intuition) that such a social effect is stronger for women whose job (according to the O*NET dictionary of occupations) offers more opportunities for social interaction, and weaker for individuals directly involved in health care, such as doctors and nurses.

**JEL codes**: I12, Z13

**Keywords**: preventive care, social interactions, health risk, ambiguity, comparative ignorance, demonstration effect.

*We thank Juliette Liesinger, James Moriarty, Kari Ruud and Jonathan Larson for outstanding research assistance. The data used in this paper contain highly sensitive personal information, obtained under a special agreement. As such they cannot be made publicly available. More information about the data, within the limit of the aforementioned agreement, can be obtained from the authors upon request.

†Mayo Clinic 200 1st St SW Rochester, MN 55905. (507) 284-9975. Email: banerjee.ritesh@mayo.edu. The views expressed in this paper are solely those of the authors and do not reflect official positions of the Mayo Clinic.

‡Federal Reserve Bank of Boston 600 Atlantic Avenue Boston, MA 02210. (617) 973-3294 Email: ethan.cohen-cole@bos.frb.org. The views expressed in this paper are solely those of the authors and do not reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System.

§University of Siena, Department of Economics. Piazza S. Francesco 7, 53100 Siena Italy. Email: zanella@unisi.it.

# 1 Introduction

The potential benefits of cancer screening exams are well known: early detection increases the probability of successful treatment and survival. This is particularly important for breast cancer, which is relatively common, has high mortality rates, and can impact fully half the population.[1] The medical community has reached an effective consensus on the need for widespread mammograms for women over 50 years of age, and some agreement on the need for women aged 40-50.[2] As the benefits of screening have become increasingly well known and the recommendations from the medical community for regular exams more common, the percentage of women who do so is surprisingly low and declining: according to recent estimates by the National Cancer Institute (NCI), 66% of eligible women in the US had a mammogram in the previous two years, compared with 75% in 2000.[3] Several reasons can explain this pattern, after taking into account the fact that most women in the eligible age range are often reminded to perform an annual mammogram. These include: (1) more uninsured women, (2) higher co-payments, (3) fewer places and specialized personnel to perform a mammogram, (4) less confidence in the effectiveness of a mammogram relative to its cost in terms of pain and (5) less fear of breast cancer. Access to a unique dataset from a large medical organization, the universe of employees, enables us to rule out the first three reasons: all women in our dataset are insured employees of a large medical provider, have no co-payment for mammograms, and are able to be screened at the workplace. Despite this, and consistent with the figure reported by the NCI, only 59% of eligible women performed a mammogram in both 2003 and 2004. This figure is remarkable since women in our sample are employees of a health organization, making them perhaps the most likely individuals to be informed about the benefits of screening. Given that the cost of a screening (e.g. the hassle of scheduling an appointment and actually performing the screening) is negligible, as is the embarrassment, pain, or shock associated with a mammogram when

[1] Breast cancer is the most common non-skin malignancy among women in the United States and second only to lung cancer as a cause of cancer-related death. In 2001, an estimated 192,200 new cases of breast cancer were diagnosed in American women, and 40,200 women died of the disease. The risk of developing breast cancer increases with age beginning in the fourth decade of life. The probability of developing invasive breast cancer over the next 10 years is 0.4 percent for women aged 30 to 39, 1.5 percent for women aged 40 to 49, 2.8 percent for women aged 50 to 59, and 3.6 percent for women aged 60 to 69 (American Cancer Society 2001, 2002).

[2] See, for instance, Alexander et al. (1993), and Miller et al. (1992)

[3] "Mammograms in Decline", *The New York Times*, May 15, 2007.

traded-off with a higher probability of survival, a logical explanation may be that these women tend to have a biased perception of the actual health risks they face.

In this paper we relate such biased perception to the behavior of co-workers and colleagues living in close proximity, who we refer to as "neighbors," by modifying a standard model in health economics, that is the human capital model (Grossman, 2000). We assume, quite reasonably we think, that one's health stock is not accurately observable unless a screening is performed. Consistent with evidence summarized below, we assume not only that individuals have a biased perception of health risk, but also that such bias is unsystematic, so that they effectively choose under ambiguity whether to perform a costly screening or not. We then incorporate into the model well-established experimental evidence, initiated by Fox and Tversky (1995), suggesting that tolerance for ambiguity varies when people learn about the behavior of others who might be more knowledgeable than they are. This phenomenon, known as *comparative ignorance*, leads to changes in behavior when other individuals change theirs. In the case under study, a woman can choose among a safe bet (perform a mammogram and know her actual health) and a risky one (do not perform a mammogram and not know) without objectively assessing risk. The safe bet is more likely to be chosen if the woman is less tolerant towards ambiguity. In turn, tolerance for ambiguity decreases if the behavior of other women, co-workers and colleagues living in close proximity, better reveals the actual underlying risk. In this way, social interactions or neighborhood effects are introduced into the model, resulting in the individual probability of getting screened being positively related with the frequency of screenings among "neighbors" — an implication we can test. We refer to this as a *demonstration effect* in preventive care, because it is similar to Duesenberry's (1949) demonstration effect in consumption. According to Duesenberry, in the context of consumption theory,

> mere knowledge of the existence of superior goods is not a very effective habit breaker. Frequent contact with them may be [...]. What kind of reaction is produced by looking at a friend's new car or looking at houses or apartments better than one's own? The response is likely to be a feeling of dissatisfaction with one's own house or car. If this feeling is produced often enough it will lead to action which eliminates it, that is, to an increase in expenditure (p. 27).

Translated to our context, knowing that a co-worker performed a mammogram may not be enough to induce a woman to do the same. However, constantly interacting with female co-workers who have been screened may be sufficient. The reaction is a feeling of anxiety about one's health, which lowers the tolerance for ambiguity. Such a mechanism naturally generates frictions: it takes time to find out about enough co-workers who have performed a screening, a fact we exploit for identification.

The idea that people have a biased perception of the actual health risks they face has some empirical support. In fact, despite the existence of statistics to assess health risk, as well as education about such risks and free online software to estimate personal risk conditional on many individual covariates,[4] both physicians and patients appear to be subject to perception bias in unsystematic ways. For instance, studying the perception of skin cancer risk, Bränström et al. (2006) find that people are aware of the fact that *in general* , sun exposure is a major risk factor, but when they judge *their own* vulnerability they do not use that information in a consistent way. Sloan et al. (2004) report the same finding for the risks associated with smoking. More importantly for our study, women seem to have an inaccurate (either upward or downward) perception of breast cancer risk (see Bottorff et al. 2004).

Exploiting the aforementioned dataset, we focus on women aged 40 to 75 — the recommended screening range — and explore the relation between the individual probability of getting screened and the frequency of screenings, during the previous year, among co-workers and colleagues living in close proximity. The one-year lag is suggested directly by the working of the demonstration effect, and allows straightforward identification. We find that the social multiplier associated with the decision to perform a mammogram, that is the ratio between the cumulative response to a shock, net of feedbacks, and the sum of initial individual responses, is over 2. This means that the initial effect of any exogenous increase in the percentage of women performing a screening—for instance, because of a new information campaign—is more than doubled in equilibrium. The reason is that social interactions amplify the initial variation by inducing more women to get screened as they find out other women in their reference group underwent a mammogram. Such a social multiplier mechanism is important for cost-benefit analysis when evaluating health policy. For instance, in the previous example, a given target in terms of percentage of women up-to-date with a mammogram might be reached with about half the expenditure that

[4] See the free Breast Cancer Risk Assessment Tool provided by the NCI at http://www.cancer.gov/bcrisktool/

would be required in a world without feedbacks, *ceteris paribus*. Such an effect turns out to be larger for individuals whose job offers more opportunities for social interactions, and weaker for individuals such as doctors and nurses who are likely to better assess actual risk. Both these findings are in agreement with intuition.

The fact that social multipliers can be exploited for health policy is the goal of a growing recent literature at the intersection of the social sciences and health literatures: recent empirical studies include the work of Deri (2006) on utilization of health services in Canada, Aizer and Currie (2004) on public maternity care, Burke and Heiland (2006) as well as Christakis and Fowler (2007) on obesity, and Fletcher (2006) on television viewing.

Our approach, by focusing on the presence of a social multiplier in health-related decision making, complements a vast empirical literature that focuses mostly on individual effects. Researchers have found that income (Blackman et al. 1999) and access to and/or type of insurance (Kenkel 1994) are both positive predictors of screening rates. Similarly, education levels (Blackman et al. 1999; Muurinen 1982) have been shown to correlate with preventive tests. Age (Blackman 1999, Grunfeld et al. 2003, Blustien and Weiss 1998b, Walter et al. 2004) has a mixed effect. It appears that women over 50 have the test more often than average, but the oldest women (over 70) have it less often. Of course, a number of authors (Rimer 1997, Burack et al. 2000) have studied the effect of cost and found that it is negatively correlated with preventive care. To proxy for individuals' concern for their own health, a number of studies look for correlations between smoking and breast cancer screening. These consistently find a positive correlation (e.g. Beaulieau et al. 1996, Fredman et al. 1999). Yet others have found that the time cost of screening is negatively correlated with screening itself (Acton 1975, Coffey 1983). Though many authors have found that the greatest predictor of screening is the recommendation of an individual's physician (Stoddard et al. 1998, Maxwell et al. 2001) and prior screens of other types are also correlated with breast cancer screening (Frazier et al. 1996, Rakowski et al. 1993), there is a growing realization in the medical community that these indicators are not sufficient to characterize differences in behavior. To fill the gap, researchers have drawn on psychological, sociological and behavioral economic literatures to find additional factors in determining screening participation. Among the reasons that have been posited are the perceived risk of cancer (Aiken et al. 1995), expectations of pain, beliefs about efficacy of tests (Maxwell et al. 1997, Lagerlund et al. 2001, Silverman et al. 2001), fear of potential diagnoses, embarrassment and discrimination (Facione and Katopodi 2000) and others. For our purposes, the most pertinent of these potential additional factors is the nature of an individual's social setting or group. Many researchers to date have confirmed that the presence of a social network appears to be a positive indicator of screening (Bloom 1984, Lichtman 1987, Spiegel 1989, Langlie 1977, Suarez et al. 2000, McCance et al. 1996, Gotay et al. 1998, Fite et al. 1996, Pearlman et al. 1995, Kang et al. 1994, and many others). The unifying feature of these studies for our needs is that they look at whether an individual's exposure to a large social network of a given type leads to greater participation. However, they do not attempt to identify the causal role of social interactions in screening decisions: within the medical literature, the exploration of social or behavioral factors in health decision making has principally involved simple data exploration. While these are broadly useful in their descriptions and characterizations, there is an outstanding debate about finding causal linkages that these types of studies are unable to address.[5] Our goal is to contribute to this debate from an economic perspective.

The remainder of the paper is organized as follows. Sections 2 and 3 present, respectively, the theoretical and econometric frameworks for our study. Section 4 describes the data and Section 5 discusses our results. Section 6 concludes.

## 2 Theoretical framework

In this section, we lay out a motivating model–in the sense that we do not aim at estimating structural parameters–that illustrates how ambiguity surrounding health and, consequently, interdependent decision making can be introduced in the human capital approach to health (Grossman, 2000).[6] The model implies dependence of the probability of getting screened on the frequency of screening in the reference group, that is a role for social interactions. We then test this implication using a simple linear model.

In each period individuals allocate their usable time to work ($\tau^w$) for pay at wage rate $w$, production of health capital ($\tau^h$), and leisure ($\tau^l$). Total usable time is the fixed time endowment (normalized to one) net of time lost because of illness and not usable for any other purpose ($\tau^n$). The

[5] An interesting debate on the claims of causation in these models is presented in Roux (2004), Subramanian (2004), Oakes (2004a,b), and Oakes (2006).

[6] Such approach treats health as a stock in the same way as (but distinct from) human capital. While the return on human capital consists of higher income, the return on health capital consists of higher quality and longer usable life. This straightforward interpretation allows us to conceptualize the demand for medical care — an input into the production of health capital — in a standard fashion.

consumption good is the numeraire. Health capital, $h$, is produced using time $(\tau^h)$ and medical care, $m$, which is priced at $p$ units of consumption and whose supply is infinitely elastic. That is, investment in health capital is $x\left(m, \tau^h\right)$. Health depreciates at a time-varying rate $\delta_t$, and has two effects. First, the time horizon is potentially infinite but death occurs as health hits a critical lower bound, $\underline{h}$. One can rule out that an individual can buy everlasting life by assuming that $\delta_t$ converges to 1 fast enough. Second, healthy time $(1-\tau^n)$ is a constant fraction $\phi$ of health capital. We depart from the standard model and we do not assume that individuals know their health capital. Different from physical and human capital, whose value is explicit in accounting ledgers and labor contracts, health capital is unobservable until a screening is performed. We assume that actual health capital in any period, $h_t$, is perfectly observed if a screening was performed during that period, otherwise it is imperfectly observed:

$$\widetilde{h}_t \equiv h_t + (1-s_t)\,\varepsilon_t \tag{1}$$

where $\widetilde{h}_t$ is *perceived* health capital (which is what people take into account when making choices), $s_t$ is an indicator variable that takes value one if a screening was performed in period $t$ and zero otherwise, and $\varepsilon_t$ is the period $t$ bias in perceived health for an individual who did not undergo a screening. Of course, screening comes at a cost. This cost may be an out-of-pocket direct medical cost, a time cost or inconvenience or some combination of the three. We also assume for simplicity, that screening is the only input into the production of health that requires time, and that screening allows an individual to "see" actual health capital but does not add anything to it (that is, it does not enter the investment equation). All other inputs, represented by $m$, cost money but do not consume additional time. Think, for instance, of buying organic rather than genetically modified food, or eating a few strawberries rather than smoking a cigarette. Furthermore, assume that the out-of-pocket cost of a screening is $q$.

The medical literature is concerned with understanding the distribution of the noise term $\varepsilon_t$ in (1) conditional on observable individual characteristics, that is in assessing the bias in health risk perception. The true distribution of $\varepsilon_t$, as the evidence reported in the introduction suggests, is effectively unknown to patients. This suggests modeling the choice problem in terms of choice under uncertainty or ambiguity (meaning that the distribution of $\varepsilon_t$ is unknown) rather than in terms of choice under risk (meaning that the distribution of $\varepsilon_t$ is known). It is well known, at least since the Ellsberg paradox, that in the presence of ambiguity people do not represent uncertainty using a single (objective or subjective) probability measure that integrates to one. We assume that individuals have preferences defined over streams of consumption ($c$) and leisure and solve:[7]

$$\max_{c_t, m_t, \tau_t^w, s_t} \mathcal{E}_t \sum_{t=0}^{T} \beta^t u\left(c_t, \tau_t^l\right), \tag{2}$$

subject to

$$\begin{aligned}
\widetilde{h}_t &= h_t + (1 - s_t)\,\varepsilon_t & (3)\\
h_{t+1} &= x(m_t) + (1 - \delta_t)\,h_t & (4)\\
c_t + pm_t + qs_t &= \tau_t^w w_t & (5)\\
\tau_t^w + s_t \tau_t^h + \tau_t^l &= \phi \widetilde{h}_t & (6)\\
T &= \min\left\{t : h_t \leq \underline{h}\right\}. & (7)
\end{aligned}$$

Here $\mathcal{E}_t$ denotes a conditional (on information at time $t$) expectation operator appropriate in the context of choice under ambiguity, for instance an integral with respect to a probability kernel correspondence (Epstein and Wang, 1994) or a Choquet integral, assuming such a representation is possible. Finally, $\beta \in (0,1)$ is the discount factor, and $h_0$ is given, with $h_0 > \underline{h} > 0$. Replacing constraints (3), (5) and (6) into the objective function, the problem reduces to the following:

$$\max_{m_t, \tau_t^w, s_t} \mathcal{E} \sum_{t=0}^{T} \beta^t u\left(\tau_t^w w_t - pm_t - qs_t\ ,\ \phi\left(h_t - (s_t - 1)\,\varepsilon_t\right) - \tau_t^w - s_t \tau_t^h\right), \tag{8}$$

subject to

$$\begin{aligned}
h_{t+1} &= x(m_t) + (1 - \delta_t)\,h_t & (9)\\
T &= \min\left\{t : h_t \leq \underline{h}\right\}. & (10)
\end{aligned}$$

Notice that when people decide to undergo a screening and find out their actual health capital, they experience a utility loss if actual health is less than they would have perceived without the screening, because actual leisure

[7] The utility function has all the standard regularity properties. Also, we write the constraints directly as binding for the sake of brevity.

is less than perceived leisure (think of a shocking news announcement regarding your health) and a utility gain otherwise. Definition (1) implies that such a loss or gain is equal to $h_t - \widetilde{h}_t = (s_t - 1)\,\varepsilon_t$, i.e. the negative of the perception bias. Therefore, the decision to perform a screening is costly in terms of consumption and leisure, but it allows one to eliminate the ambiguity associated with not knowing the distribution of $\varepsilon_t$. An ambiguity-averse individual tends to avoid a "bet" characterized by unknown odds. When considering whether to undergo a screening or not, an individual can choose among a "safe bet" (perform the screening, pay the cost, and eliminate the surprise component of utility) and an "ambiguous bet" (do not perform the screening, save on costs, and accept the surprise component of utility). Choosing $s_t = 1$ shuts off the random term in the choice problem, at the cost of $\tau_t^h$ units of time and $q$ units of consumption. This can be regarded as the price of ambiguity. We are not concerned with fully characterizing the solution to the problem defined by equations (8) to (10), which can be done under standard assumptions.[8] It suffices to notice, by inspection of the problem, that in equilibrium the decision to perform a screening must depend on the degree of tolerance for ambiguity. The key assumption of this paper is that ambiguity aversion is endogenously determined by social interactions. This assumption is an implication of the *comparative ignorance hypothesis* formulated by Fox and Tversky (1995), and corroborated by substantial experimental evidence (Fox and Tversky themselves, and Fox and Weber, 2002, among others, but see also Arlo-Costa and Helzner, 2005, for a critique). The version of the hypothesis we are interested in is that individuals who are willing to accept ambiguity when making decisions in isolation become ambiguity averse when they compare themselves with individuals who are making the same decision and who might be more knowledgeable than they are. As Fox and Weber (2002) put it, such comparisons make the notion of competence more salient. This effect is similar to, though distinct from, the main source of ambiguity avoidance identified by Curley et al. (1986), labeled the *other-evaluation hypothesis*. This hypothesis states that if a decision maker anticipates that others will evaluate his or her decision, the decision maker will make the choice that is perceived to be most justifiable to others, i.e. the less ambiguous one. In our context, it doesn't matter which hypothesis is invoked, although we find comparative ignorance much more palatable. What matters is that the social dimension of ambiguity

[8] See Grossman (2000) for the solution of the human capital model, and Epstein and Wang (1994) for the characterization of a dynamic optimization problem in the presence of ambiguity.

aversion implies a positive relationship between the likelihood of performing a screening and the frequency for screening of individuals with whom one interacts more closely. Suppose a woman has not performed a breast cancer test yet i.e. she is choosing the ambiguous bet. She then finds out that a colleague, a neighbor or a friend got tested and eliminated their own ambiguity by choosing the safe bet. This may trigger a feeling of anxiety about living with the ambiguous bet. Consider the woman wondering why her colleague performed the screening: was she better informed about the actual risk? If this feeling is repeated long enough, that is if the woman finds out about more and more colleagues, neighbors or friends who got tested, she will eventually abandon the ambiguous bet and opt for getting tested. This is what we call the *demonstration effect* in preventive care. The goal of our empirical exercise is to test this implication.

# 3 Econometric framework

Although we employed a dynamic theoretical framework, our data–which is essentially cross-section–does not allow us to estimate a fully dynamic model. However, any dynamic theory has static, i.e. intratemporal, implications one can test. In order to keep the empirical analysis as simple as possible, we assume that, in each period, the intratemporal condition involving the decision to get screened can be approximately expressed as a linear function of individual characteristics, group characteristics, and peers' behavior. In this case, one could employ the following model:

$$s_{it} = b + cX_{it} + dY_{gt} + Js_{gt} + u_{it}, \tag{11}$$

where $s_{it}$ is a binary indicator assuming value 1 if person $i$ performed a screening (a mammogram in this case) at time $t$, and zero otherwise; $X_{it}$ is a set of individual characteristics, $Y_{gt}$ a set of contextual characteristics common to all persons in group $g$ (formally a set), $s_{gt}$ the frequency of testing in the group (which proxies for the intensity of the demonstration effect), and $u_{it}$ an unobservable term. This is the so-called linear-in-means model. As well known, the main cause of concern with such a model when trying to estimate the effect of mean group behavior on individual behavior is the endogeneity of the former (Manski 1993, Brock and Durlauf 2001). Equation (11) implies that in large samples:

$$s_{gt} = E\left(s_{it}|i \in g\right) = \frac{b}{1-J} + \frac{c}{1-J}E\left(X_{it}|g\right) + \frac{d}{1-J}Y_{gt}$$

Replacing this in (11), one is left with a reduced form that cannot identify the effect of mean group behavior $(J)$ if mean individual characteristics, $E(X_{it}|g)$, are linearly related with group characteristics, $Y_{gt}$ as is often the case. This is the well-known reflection problem of Manski (1993). In order to avoid this problem, we follow Aizer and Currie (2004) and consider lagged rather than current relative frequency of screenings, exploiting the fact that in general $s_{gt-1} \neq E(s_{it}|g)$. Technically, this is the simplest way of sidestepping the reflection problem and so achieve identification (See Brock and Durlauf, 2001, for a detailed discussion).[9] The question is whether such an assumption has some justification or is simply *ad-hoc*. In light of our theory, this is actually a very natural assumption, because it is equivalent to assuming that the demonstration effect is somehow sluggish — one lag in this case. That this is the case is suggested directly by the psychological mechanism that defines such an effect: it takes time to find about *enough* peers for such an effect to induce a change in behavior. Based on this argument, our identifying assumption is that the effects of social interactions manifest themselves after one lag — one year in our case. Of course this assumption comes at a cost: one must also assume that mean group behavior has not reached a steady state. The reason is that at a steady state the percentage of testing in the group is stationary, that is $s_{gt-1} = s_{gt} = E(s_{it}|g)$. In this case identification fails, in the absence of restrictions, because using lagged mean behavior would be no different than using the contemporaneous one.

Two additional reasons of concern in social interactions models are self-selection and the presence of group unobservables. These are mitigated by the nature of our data. Our sample is actually the universe of employees in a large medical organization, and groups, in our main specification, are defined by job titles, i.e. who works with whom, as we illustrate in depth below. As for group unobservables, it is reasonable to assume that the relevant unobservables that affect all individuals in a group are determined at the clinic level, e.g. test-reminder letters. Therefore such effects are absorbed by the constant regressor. However, a second class of unobservables arises from selection into the job. For instance, nurses that worked together have something in common for the very reason that they all chose to be nurses. Since skills are predetermined with respect to selection into the job, in order to avoid selection bias one must assume that such skills are unrelated with the concern for health. This is admittedly a strong assumption, but given

[9]This assumption allows Aizer and Currie (2004) to identify social interactions in the use of maternity-related services.

the limitations of the data this is the only thing we can do to identify a causal effect.

In summary, we estimate a linear probability model. In the following, $\theta_{ij}$ is the frequency of contacts, or strength of interaction, between individual $i$ and $j$, while the other notation is as in equation (11). The equation of interest is:

$$s_{it} = b + cX_{it} + dY_{gt} + J \sum_{j \neq i} \theta_{ij} s_{jt-1} + \varepsilon_{it}. \tag{12}$$

The crucial assumption is about who interacts with whom — the reference group — which is the specification of the set of weights, $\{\theta_{ij}\}$. In the basic specification we assume that the reference group is composed of individuals who work together at the clinic, that is have the same job title. The assumption is that, for instance, nurses interact with nurses, accountants with accountants, and so on. We also assume that interactions are uniform within the group — the so-called mean field case, $\theta_{ij} = N_g^{-1}$, where $N_g$ is the size of group $g$, the occupation. We refer to this case as the baseline reference group.

The assumption that people associate within the clinic on the basis of their job title, although palatable, might bias the results if instead people associate across job titles on the basis of the building where they work, for instance. Unfortunately, in the data we cannot see the specific location of an individual within the clinic. The best we can do is to exploit geographic information and expand the reference group to include employees who live close to each other, regardless of whether they have the same job title within the organization or not. We construct a metric that varies inversely with the residential distance between $i$ and $j$, $d_{ij}$, and we consider two individuals as if they were neighbors if they work together, regardless of where they actually live. That is $\theta_{ij} = N^{-1} d_{ij}$, where $N$ is the number of employees, $d_{ij} = 1$ (minimum distance) if $i$ and $j$ have the same job title within the clinic, and $d_{ij}$ less than one if $i$ and $j$ have different job titles. That is, if $\delta_{ij}$ is the actual distance between $i$ and $j$, $d_{ij} = 1 - \frac{\delta_{ij}}{\max(\delta)}$ where $\delta \equiv [\delta_{ij}]$, when $i$ and $j$ have different job titles. Clearly, when using residential information the problem stemming from selection may be exacerbated. To illustrate the use of weights, that is the $\theta_{ij}$ in (12), consider five employees: 1 and 2 have one job title, and 3, 4 and 5 have another job title. Then the basic specification uses the following weighting matrix:[10]

[10] Notice that "loops" — interacting with oneself — are ruled out.

$$[\theta_{ij}] = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix} \tag{13}$$

while the specification that exploits residential information uses this one:

$$[\theta_{ij}] = \begin{bmatrix} 0 & \frac{1}{5} & \frac{1}{5}d_{13} & \frac{1}{5}d_{14} & \frac{1}{5}d_{15} \\ \frac{1}{5} & 0 & \frac{1}{5}d_{23} & \frac{1}{5}d_{24} & \frac{1}{5}d_{25} \\ \frac{1}{5}d_{31} & \frac{1}{5}d_{32} & 0 & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5}d_{41} & \frac{1}{5}d_{42} & \frac{1}{5} & 0 & \frac{1}{5} \\ \frac{1}{5}d_{51} & \frac{1}{5}d_{52} & \frac{1}{5} & \frac{1}{5} & 0 \end{bmatrix} \tag{14}$$

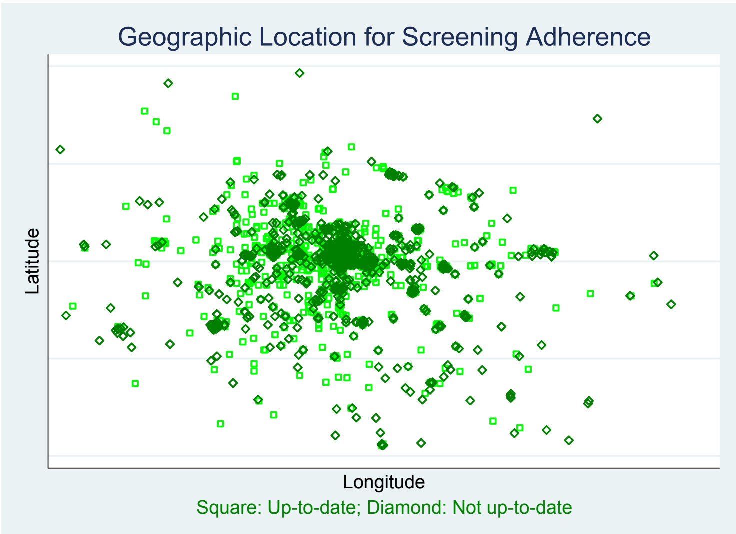
# 4 Data

Our data includes information, among other things, on age, marital status, family size, whether a person works part time and total utilization of health care (measured in dollars) in the previous year. The variables we use are listed in table 1, and summary statistics are reported in table 2. We use job title to create our occupation variable. The titles are listed in table 3. In most cases, these titles are self-explanatory and suggest an easy classification scheme e.g. Administrative Assistant or Nurse Practitioner. In some cases, we grouped titles together based on our own understanding of the jobs in terms of likely interactions e.g. we grouped Custodian with Custodian Supervisor, and Opthalmic Assistant with Opthalmic Assistant Trainee. We limit our analysis to groups that have 20 members or more in order to avoid small-group problems; this includes 4109 individuals classified into 61 distinct occupations (groups). These groups vary substantially in size ranging from 20 to 409 members with a median size of 118. We measure an individual's health status using the classification scheme in Elixhauser et al. (1998) which essentially uses health utilization data to create a set of indicator variables revealing whether an individual has a certain health condition. Examples of categories identified by this classification algorithm are hypertension, diabetes, depression, ulcers and metastatic cancer. Our data contains a variable that keeps track of whether a woman is due for breast cancer screening as defined by national guidelines.[11] Our dependent variable is a binary indicator that takes a value one if a woman is up-to-date on

[11] There is some controversy over when women should begin breast cancer screening. The American Cancer Society, American College of Radiology, American Medical Associ-

her screening requirements and zero otherwise. We also take advantage of a recent change in health plan availability to create a variable that controls for unobserved health-related willingness to take risk. In essence, this variable was set to unity when individuals switched from a high premium, full coverage plan to a low premium, cost sharing plan. We calculated average risk by age for a woman to develop breast cancer over her remaining lifetime using a publicly available risk calculation tool (http://www.cancer.gov/bcrisktool). We lack specific information that would allow us to create a more accurate score so we calculated risk by age based on average values.[12] Note that our results are not sensitive to exactly what information we use to construct our risk variable. However, as they indicate, this measure has high explanatory power in most specifications.

Figure 1. Spatial dispersion of up-to-date vs. not up-to-date women




We also have detailed information on employees' home addresses including street address and 9-digit zip code. We used freely available geocoding data files available from http://www.census.gov/geo/www/tiger and

ation, National Cancer Institute and the American College of Obstetrics and Gynecology recommend annual screening starting at age 40. Most of these also recommend screening through the age of 75.

[12] In particular, we assumed that women had menarche at age 12-13, they have had no live childbirths, no first-degree relatives have had breast cancer, they have never had a breast biopsy or any history of breast cancer, and are Caucasian.

geocoding macros available in SAS (%GCBATCH) to geocode addresses into precise latitude-longitude couples. We then calculated actual distance in miles based on an approximation of the spherical law of cosines. Spot checking of distances with actual distances calculated using an online mapping tool revealed inaccuracies in our measure of no more than mere fractions of a mile. We exclude individuals who live greater than 80 miles from work in order to have a distance index that spans the unit interval but is not highly skewed by very distant employees. Figure 1 graphs screening information based on where employees live. Being up-to-date or not on screening requirements does not seem to be related to distance from work, which is at the center of the cluster in the middle of the graph. Also, individual location seems to be random with respect to the up-to-date status (what appear to be clusters are actually population centers).

# 5 Results

In the following subsections, we describe our basic specification and then elaborate on various specification checks we have conducted.

## 5.1 Basic specification

In this section, we provide results and robustness analyses from specification (12) using the baseline distance matrix (13), i.e. reference groups are an individual's occupation category, as indexed by job title. To start, we estimate the equation by OLS. Table 4 reports results for different combination of controls. Each of the four yield coefficient estimates on the mean group screening variable between .59 and .67, and highly significant. This means that a uniform exogenous increase of, say, 1 percentage point in the individual probability of getting screened—for instance because of a new information campaign—causes, initially, a .6 percentage point individual response. This further increases the screening rate, and so on in a cumulative way. The limit of this process, the social multiplier, is $1/(1-0.6)$ or about 2.4. Thus the equilibrium increase is more than twice the initial shock. The four columns include progressively more control variables. Column 1 controls for age and family size as well as the mean of each within each group. Adding variables for individual and group-level marriage variables in Column 2 has little impact on the mean-group effect. Columns 3 and 4 add controls for the type of employment. First, we include whether an individual works full time for the health organization; since our groups are based on occupation, one might expect that individuals that work full time would be more impacted by the demonstration effect that those that do not. In Columns 3 and 4, the coefficient is positive, but not significant. Column 4 adds hourly employee status, which we would expect to have a negative coefficient: we find the expected direction, but without significance. As for the other controls, the Table shows broadly reasonable results, although we do not have strong priors as to the relevance or sign of any of them. We find that age, unsurprisingly, is a strong predictor of screening. As well, having been married and the size of one's family are similarly strong predictors. In particular, family size is negatively related to the propensity for testing. We interpret this as a time-cost issue: as the number of children increases, available time decreases and the associated cost of going for a screening increases.

## 5.2 Prior illnesses

Among the possible correlates of the decision to be screened for cancer is the presence of other prior illnesses. The pathway could be in a number of forms. Though we don't model this explicitly, once an individual is aware of her own illness, the tradeoffs proposed above could change; now, the decision to obtain a screening becomes more desirable. Of course, for the very ill, such a screening may be simply a waste of time. Having other illnesses can also reduce the cost of obtaining the screening from a time perspective if it's possible to schedule screens at the same time as other procedures. Finally, more exposure to doctors due to other illnesses can effectively add an additional, possibly influential, person to the reference set. Table 5 includes the full specification from Table 4, column 4 plus a number of indicators for the presence of other illnesses. Column 1 include other cancer-related illnesses, column 3 mental health related illnesses and column 2 others that are included in the dataset. The column 2 list includes such conditions as the presence of ulcers, renal failure, and many others. Finally, column 4 includes all indicators. Each of the variables appears in the regressions as a dummy variable with a '1' indicating that the condition has been previously diagnosed for the individual. All columns again include the group-level average prevalence of the condition. The result is that the mean testing rate now has a coefficient of between .49 and .59: when controlling for prior illnesses, the estimated social effect is weaker, as one would expect. As reported in the Table, a number of comorbidity indicators are significant and positive. As argued above, this is expected, both because frequency of other medical visits may increase the ease with which testing is accomplished, provide additional reminders of the need and potentially provide additional information on mortality overall. Of particular note, we find that the prior presence of a tumor has a positive coefficient.

### 5.3 Risk attitudes

Personal attitudes toward risk are obviously relevant to the decision of being up-to-date with mammograms. We explore in Table 6 individual perceptions of risk in two distinct ways. Column 1 repeats the baseline specification of table 5, column 4. This baseline has our coefficient of interest, mean group testing rate, and a number of relevant controls. Column 2 of Table 6 includes a measure of average breast cancer risk for an individual over their remaining life. As we noted above, these rates are publicly available to individuals over the internet. While many may not have accessed the information, we explore the role of additional available information on the role of social interactions in the presence of ambiguity. Given that under our model, realizing lower health than previously perceived leads to a utility loss, the availability of additional information can reduce this ex-post utility loss and thus should lead individuals to be more likely to test. We find a relatively large and highly significant coefficient of .184 on the risk indicator (column 2) The social interaction coefficient is not significantly altered.

Column 3 looks at the role of individual risk preferences. Recent changes in the available health plan allowed us to follow individuals' choices regarding the 'riskiness' of their plan. One provides low monthly premiums and high co-insurance and the other a high monthly premium and lower co-pay. Prior to the change, most employees were covered by a full coverage plan. Recall that individuals do not pay for the screening itself; however, they do of course bear their share of the cost of treatment if diagnosed. We include an indicator variable set at one when individuals switch to a higher risk plan. Column 3 includes this new variable as well as its group level mean. The variable itself is negative and insignificant. The coefficient on its group-level mean is positive, large and significant, suggesting that being surrounded by individuals with low risk aversion leads to greater testing probability. The key coefficients are essentially unchanged across these specifications.

### 5.4 Knowledge

Since all the individuals in our dataset are employees of a medical organization, one may suspect that the population being considered would be particularly aware of the benefits of health screenings. To verify the relevance of this possibility, we decompose the dataset into direct health service and other professions. Though all the individuals are employees of a medical organization, the hypothesis is that doctors and nurses directly involved in service provision may be more aware of the benefits than those involved in support (custodial, administration, etc.) services. Table 7 reports estimates based only on the subsamples of employees either directly involved in medical care or not involved. Column 1 is again the baseline specification. Columns 2 and 3 show the subsample of employees that are involved in patient care jobs (doctors, nurses, etc.).[13] Column 3 includes the risk preferences variable from table 6 to ensure that in subdividing the sample, we are not simply capturing differences in risk preferences rather than differences in knowledge. Columns 4 and 5 repeat for non-patient care jobs (computer technicians, janitors, etc.). While the magnitude of the key coefficients are largely in line with the prior tables, we note here principally that the coefficient on the key variable is significantly lower for patient care specialities. That is, as expected, social interactions explain less about the individual propensity to be screened.

### 5.5 Geography

We exploit residential information in our dataset to evaluate our initial assumption of networks being formed within the workplace on the basis of job titles. Our results are in fact robust to a location-based reference group as well. We performed this check in two ways.

First, we consider residential information alone. What if people interact, for instance, not with colleagues with the same job title but with colleagues who live nearby? This is clearly also a plausible reference group. We define a group for each individual in the sample as a circle with a 2-mile radius. The coefficient of interest is again in the 0.5-0.6 range across the columns in Table 9a. This is robust to increasing the radius, probably because of the decreasing social weight of progressively distant colleagues. The similarity of the coefficient across these different reference groups might be due to residential choice and job titles being related. To gain insight into this issue, we look at the relationship between location and job classification. Table 8 below illustrates this with a few regressions. These show the coefficient of a regression of actual distance between households and an indicator for whether two individuals are in the same job classification. Notice that this produces $n(n-1)$ observations; each individual has an observation for her

[13] Note, because of the reduced sample sizes for the divided population, we are unable to include all comorbidity and mean comorbidity information. For the purposes of this table, we exclude the mean comorbidities from the analysis.

pairing with every other individual. Column 1 shows results from this simple regression. The constant term suggests that individuals live, on average, 18 miles apart. The coefficient of 'groupflag' indicates that individuals in the same group live on average 2 miles further apart than individuals that do not work together. Both the constant and the group variable are highly significant. This provides some evidence that individuals do not choose where to live based on their work classification. Column 2 repeats the exercise including controls for various observable individual characteristics that might influence the decision of where to live. This has a minor effort on the constant term, increasing it to 24, but does not impact the magnitude of the groupflag variable. Finally, column 3 looks at location decisions after isolation into small areas, defined by a 2-mile radius. Perhaps individual select neighborhoods based on some unobserved characteristic and then locally sort based on job classification. This regression finds that within a 2 mile radius, individuals live on average 1.3 miles apart. The groupflag variable turns negative, but only very slightly so. The magnitude suggests that individuals within a job classification live 1/100 of a mile closer together.

Second, we mix association within occupational titles and within residential neighborhoods, i.e. use weight matrix (14). In this case the coefficient of interest increases from the basic specification to about 0.75, as reported in Table 9b. This is a large increase, which is likely due to the fact that by using information about on-the-job *and* spatial association we are effectively working with a pan-reference-group. That is, every individual's behavior becomes relevant—although with varying weights—while in the previous specifications only the behavior of individuals with the same job title *or* living nearby was included. This is clear from the fact that matrix (13) has many off-diagonal zeroes, while matrix (14) has none. The implication is that one should not directly compare the coefficients across measures, and that, as always with empirical studies of social interactions, our group categories are approximations to true social networks.

### 5.6 Sociability

Our empirical model centers on the idea that women working at the clinic talk to each other or simply observe each other's behavior (if they belong to the same reference group), in particular about cancer screening. If the job title within the organization is the appropriate reference group, then one can perform a minimal test that this assumption is correct, by exploiting the fact that the intensity of on-the-job social interactions differs across occupation. We augmented our dataset with information about social interactions implied by one's occupation. Such information comes from the O*NET Dictionary of Occupational Titles, an online dictionary that associates to each occupational title a variety of information, including the degree of on-the-job opportunities for social contact. From these data, we construct four categories of jobs based on their degree of social contact. Column 1 of Table 10 reports results for the entire sample, categorizing jobs with a dummy variable for each value of the index. The least 'social' job category is excluded. After controlling this way for sociability, the estimated social effect declines slightly with respect to the basic specification. Furthermore, we split the sample into individuals with high- and low-sociability jobs, and estimated the model separately for the two groups. Columns 2 and 3 report results for jobs belonging to categories 1 to 3 and category 4, respectively.[14] In line with the implications of our hypothesis, we find a significantly higher social interactions coefficient for the group of individuals with high sociability jobs.

### 5.7 Nonlinearities

A linear model is admittedly a strong choice in many respects: not only does it imply that probabilities can be misrepresented at the extremes, but it neglects possible nonlinearities in the response to certain variables, over their whole range. We chose the linear model because it provides robustness with respect to the distribution of unobservables, which is unrestricted. However, to validate that our results are not a product of forcing a linear response, we include a replica of table 4 using a probit model with the individual's test/no test decision as the dependent variable. Results are reported in Table 11. All of the coefficients are translated into marginal effects at the mean of the independent variables. One can see easily that the magnitude and significance of these are very close to that found in table 4.

## 6 Conclusions

Despite a recent controversy surrounding mammograms, as evidenced by the recent NYT treatment cited in the Introduction, there is abundant evidence that cancer screening is an effective tool, and one that is widely available. Yet, many women do not obtain cancer screening, even when it is free, easily available and doctor encouraged — as per our sample. Though we are not able to rule out pain-avoidance as a factor, we conjecture that this is unlikely to explain the large rates of non-participation observed, since the

[14] This defines two subsamples of 2,136 and 1,973 individuals, respectively.

payoff from a screening is potentially high. In this paper we have suggested and tested an explanation based on peer decision making. Specifically, we find that behavior appears to mirror the implications of a health capital model with ambiguity aversion under a 'comparative ignorance' effect. Using information from a population working at a large health services provider, we have found that individuals make cancer screening decisions in part based on the behavior of their colleagues and neighbors. This implies the existence of a social multiplier that amplifies the response of screening rates to individual shocks. Our results are confirmed under a wide array of sensitivity and robustness checks. What makes this finding remarkable is that the women in our sample are employees of a health organization, making them perhaps the most likely individuals to be informed about the benefits of screening. Then again perhaps the cobbler's children are simply going barefoot...

Table 1. Variable Descriptions

| VARIABLE | Description |
|---|---|
| meantestMG03: | percent up-to-date in 2003 |
| separated: | whether separate |
| age: | age |
| famsize: | family size |
| full: | whether full time employed |
| hourly: | hourly-basis employee |
| new: | whether new employee |
| divorced: | whether divorced |
| married: | whether married |
| unmarried: | whether not married |
| BC Risk: | average risk of cancer in demographic |
| meanseparated: | within group average of (separated) |
| meanage: | within group average of (age) |
| meanfamsize: | within group average of (famsize) |
| meanfull: | within group average of (full) |
| meanhourly: | within group average of (hourly) |
| meannew: | within group average of (new) |
| meandivorced: | within group average of (divorced) |
| meanmarried: | within group average of (married) |
| meanunmarried: | within group average of (unmarried) |
| meanbcriskless30: | within group average of (BC Risk) |
| Congestive Heart Failure: | whether has CHF |
| Diabetes Mellitus: | whether has DM |
| Hypothyroidism: | whether has Hypothyroidism |
| Previous Tumor: | whether has previous tumor |
| Obese: | whether obese |
| Depression: | whether has depression |
| Metastatic Cancer: | whether has metastatic cancer |
| Hypertension: | whether has hypertension |
| Mean Group Comorbidity: | within group average of morbidity indicators |
| Hitolow: | whether changed from hi premium health plan to low |
| newmeantestMG03: | interaction of new employee and meantestMG03 |

Table 2. Summary Statistics

| VARIABLE | Mean | Std. Dev. |
|---|---|---|
| meantestMG03 | 0.6915 | 0.0860 |
| separated | 0.0034 | 0.0583 |
| age | 49.6990 | 5.3873 |
| famsize | 2.6834 | 1.4189 |
| full | 0.8776 | 0.3278 |
| hourly | 0.9718 | 0.1657 |
| new | 0.0066 | 0.0808 |
| divorced | 0.1343 | 0.3411 |
| married | 0.7289 | 0.4446 |
| unmarried | 0.1039 | 0.3052 |
| BC Risk | 9.8590 | 0.8462 |
| meanseparated | 0.0034 | 0.0073 |
| meanage | 49.6987 | 0.9844 |
| meanfamsize | 2.6847 | 0.2889 |
| meanfull | 0.8778 | 0.0868 |
| meanhourly | 0.9718 | 0.1562 |
| meannew | 0.0073 | 0.0109 |
| meandivorced | 0.1341 | 0.0491 |
| meanmarried | 0.7292 | 0.0764 |
| meanunmarried | 0.1036 | 0.0576 |
| meanbcriskless30 | 9.8591 | 0.1565 |
| Congestive Heart Failure | 0.0007 | 0.0270 |
| Diabetes Mellitus | 0.0314 | 0.1744 |
| Hypothyroidism | 0.0781 | 0.2684 |
| Previous Tumor | 0.0392 | 0.1941 |
| Obese | 0.0657 | 0.2478 |
| Depression | 0.0302 | 0.1711 |
| Metastatic Cancer | 0.0046 | 0.0679 |
| Hypertension | 0.1562 | 0.3631 |
| Mean Group Comorbidity | 0.6817 | 0.1424 |
| Hitolow | 0.1504 | 0.3575 |
| newmeantestMG03 | 0.0044 | 0.0551 |
| Observations | 4109 | |

Table 3. Occupation Groups

| TITLE | Size | Up2Date03 | Up2Date04 |
|---|---|---|---|
| Administrative Assistant | 239 | 0.7699 | 0.7406 |
| Administrative Office Manager | 21 | 0.7619 | 0.8095 |
| Anesthesia Assistant | 22 | 0.8182 | 0.7273 |
| Assistant Financial Representative | 62 | 0.6774 | 0.6774 |
| Billing Representative | 47 | 0.7447 | 0.7447 |
| Central Appointment Desk | 20 | 0.9500 | 0.6500 |
| Certified Clinical Nursing Specialist | 24 | 0.8333 | 0.8333 |
| Clinical Assistant | 144 | 0.7083 | 0.6319 |
| Clinical Lab Technician | 25 | 0.6400 | 0.5600 |
| Clinical Lab Technician Limited | 24 | 0.6667 | 0.5417 |
| Clinical Lab Technologist | 66 | 0.7273 | 0.6364 |
| Consultant | 104 | 0.6923 | 0.6538 |
| Custodian | 22 | 0.5455 | 0.4091 |
| Education Coordinator | 20 | 0.9000 | 0.8500 |
| Electrocardiographer | 25 | 0.6800 | 0.6800 |
| Finance Specialist | 61 | 0.7377 | 0.7705 |
| Financial Representative | 97 | 0.8041 | 0.7113 |
| General Services Transporter | 29 | 0.5862 | 0.6897 |
| Health Information Management Clerk | 20 | 0.7500 | 0.8000 |
| Hospital Pharmacist | 60 | 0.6500 | 0.4167 |
| Human Resources | 20 | 0.6500 | 0.6000 |
| Lab Assistant | 73 | 0.6301 | 0.5479 |
| Laboratory Service Technician II | 40 | 0.6250 | 0.4000 |
| Librarian | 34 | 0.6471 | 0.7059 |
| Licensed Practical Nurse-Ambulatory | 118 | 0.7627 | 0.7288 |
| Licensed Practical Nurse-Nursing Services | 52 | 0.5385 | 0.6346 |
| Medical Secretary | 364 | 0.7115 | 0.7280 |
| Medical Secretary Supervisor | 21 | 0.8095 | 0.8571 |
| Medical/Surgical Transcriptionist | 120 | 0.6750 | 0.6583 |
| Nurse Anesthetist | 25 | 0.6800 | 0.5600 |
| Nurse Manager | 27 | 0.6667 | 0.6296 |
| Nurse Practitioner | 31 | 0.7742 | 0.8065 |
| Nursing Education Specialist | 29 | 0.7586 | 0.7586 |
| Office Assistant II | 21 | 0.7143 | 0.6667 |
| Patient Appointment Coordinator | 147 | 0.7551 | 0.6395 |

Table 3. Occupation Groups contd.

| | | | |
|---|---|---|---|
| Patient Services Representative | 130 | 0.7615 | 0.7385 |
| Pharmacist | 20 | 0.6000 | 0.6000 |
| Pharmacy Administrator | 34 | 0.5000 | 0.5294 |
| Physical Therapist | 21 | 0.6190 | 0.6667 |
| Radiologic Technician | 38 | 0.7895 | 0.7895 |
| Registered Nurse Day I | 101 | 0.7921 | 0.7822 |
| Registered Nurse Day II | 75 | 0.7200 | 0.7867 |
| Registered Nurse Extended I | 77 | 0.7403 | 0.7143 |
| Registered Nurse Extended II | 47 | 0.7660 | 0.7234 |
| Registered Nurse Research Coordinator | 22 | 0.7727 | 0.6818 |
| Registered Nurse Supervisor | 33 | 0.4848 | 0.6667 |
| Research Study Coordinator I | 45 | 0.7778 | 0.7333 |
| Research Study Coordinator II | 38 | 0.7632 | 0.7632 |
| Research Technologist | 31 | 0.8387 | 0.5484 |
| Secretary | 28 | 0.7500 | 0.6071 |
| Senior Analyst/Programmer | 21 | 0.7143 | 0.7143 |
| Senior Research Tech I | 66 | 0.7273 | 0.6667 |
| Senior Research Tech II | 20 | 0.8000 | 0.7500 |
| Staff Nurse Day | 23 | 0.6522 | 0.8696 |
| Staff Nurse-Nursing Services I | 409 | 0.5477 | 0.5575 |
| Staff Nurse-Nursing Services II | 283 | 0.6113 | 0.6537 |
| Staff Nurse-Surgical Services I | 77 | 0.6494 | 0.6104 |
| Staff Nurse-Surgical Services II | 36 | 0.5278 | 0.6111 |
| Transitional Programmer | 34 | 0.7059 | 0.6471 |
| Unit Secretary | 131 | 0.6489 | 0.6336 |
| X-Ray Records Clerk | 27 | 0.8148 | 0.6667 |

Table 4. Baseline

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| COEFFICIENT | up2date | up2date | up2date | up2date |
| up2date | 0.667*** | 0.631*** | 0.570*** | 0.586*** |
| | (0.087) | (0.10) | (0.11) | (0.11) |
| age | 0.00972*** | 0.00944*** | 0.00945*** | 0.00944*** |
| | (0.0015) | (0.0015) | (0.0015) | (0.0015) |
| famsize | -0.0190*** | -0.0250*** | -0.0248*** | -0.0248*** |
| | (0.0058) | (0.0067) | (0.0067) | (0.0067) |
| meanage | -0.00529 | -0.00289 | -0.000332 | 0.000299 |
| | (0.0091) | (0.0095) | (0.0096) | (0.0097) |
| meanfamsize | 0.0854*** | 0.136*** | 0.157*** | 0.161*** |
| | (0.031) | (0.039) | (0.042) | (0.042) |
| separated | | -0.0523 | -0.0525 | -0.0525 |
| | | (0.15) | (0.15) | (0.15) |
| divorced | | 0.0340 | 0.0340 | 0.0340 |
| | | (0.084) | (0.084) | (0.084) |
| married | | 0.0952 | 0.0960 | 0.0960 |
| | | (0.082) | (0.082) | (0.082) |
| unmarried | | 0.0777 | 0.0775 | 0.0775 |
| | | (0.085) | (0.085) | (0.085) |
| widowed | | 0.0299 | 0.0313 | 0.0313 |
| | | (0.098) | (0.098) | (0.098) |
| meanseparated | | -0.612 | -0.643 | -0.818 |
| | | (1.27) | (1.27) | (1.29) |
| meandivorced | | 1.718** | 1.714** | 1.490* |
| | | (0.81) | (0.81) | (0.86) |
| meanmarried | | 1.316* | 1.244 | 0.997 |
| | | (0.78) | (0.79) | (0.85) |

Continued on next page...

Table 4. Baseline contd.

| COEFFICIENT | (1) up2date | (2) up2date | (3) up2date | (4) up2date |
|---|---|---|---|---|
| meanunmarried | | 1.427* | 1.377* | 1.198 |
| | | (0.80) | (0.80) | (0.83) |
| meanwidowed | | 0.562 | 0.544 | 0.250 |
| | | (0.89) | (0.89) | (0.97) |
| full | | | 0.0139 | 0.0140 |
| | | | (0.023) | (0.023) |
| meanfull | | | 0.134 | 0.150 |
| | | | (0.11) | (0.11) |
| hourly | | | | -0.0184 |
| | | | | (0.13) |
| meanhourly | | | | 0.0637 |
| | | | | (0.14) |
| Constant | -0.190 | -1.819** | -2.035** | -1.909** |
| | (0.49) | (0.89) | (0.90) | (0.91) |
| Observations | 4109 | 4109 | 4109 | 4109 |
| R-squared | 0.04 | 0.04 | 0.04 | 0.04 |

Groups are defined according to job title. Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with "1" representing that the individual is up to date on recommended screenings. All observations represent a single individual employed by a large medical organization.

Table 5. Prior Illnesses

| COEFFICIENT | (1) up2date | (2) up2date | (3) up2date | (4) up2date |
|---|---|---|---|---|
| up2date | 0.580*** | 0.429** | 0.593*** | 0.535** |
| | (0.12) | (0.20) | (0.11) | (0.22) |
| Previous Tumor | 0.176*** | | | 0.155*** |
| | (0.028) | | | (0.029) |
| Metastatic Cancer | 0.222*** | | | 0.200** |
| | (0.074) | | | (0.079) |
| Congestive Heart Failur | | -0.606** | | -0.570** |
| | | (0.24) | | (0.25) |
| Diabetes Mellitus | | -0.0241 | | -0.0233 |
| | | (0.041) | | (0.041) |
| Hypothryroidism | | 0.0851*** | | 0.0848*** |
| | | (0.025) | | (0.025) |
| Obese | | 0.0596** | | 0.0551** |
| | | (0.027) | | (0.027) |
| Hypertension | | 0.146*** | | 0.144*** |
| | | (0.018) | | (0.018) |
| Depression | | | 0.106*** | 0.0947** |
| | | | (0.040) | (0.041) |
| Constant | -1.689* | -2.695 | -1.813* | -1.751 |
| | (0.93) | (1.74) | (0.94) | (1.84) |
| Observations | 4109 | 4109 | 4109 | 4109 |
| R-squared | 0.05 | 0.07 | 0.04 | 0.08 |

Groups are defined according to job title. Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with “1” representing that the individual is up to date on recommended screenings. All observations represent a single individual employed by a large medical organization. Vector of individual variables (as in Table 4) included but not reported. Other medical conditions along with their mean by group included in regression but not reported.

Table 6. Risk Attitudes

| COEFFICIENT | (1) up2date | (2) up2date | (3) up2date | (4) up2date | (5) up2date |
|---|---|---|---|---|---|
| up2date | 0.535** | 0.552** | 0.515** | 0.502** | 0.675*** |
| | (0.22) | (0.23) | (0.22) | (0.23) | (0.087) |
| bcless30 | | 0.184** | | 0.184** | -0.0713*** |
| | | (0.077) | | (0.077) | (0.0086) |
| meanbcriskless30 | | 0.0718 | | -0.373 | 0.121** |
| | | (1.01) | | (1.02) | (0.047) |
| hitolow | | | -0.00577 | -0.00589 | |
| | | | (0.020) | (0.020) | |
| meanhitolow | | | 0.452** | 0.460** | |
| | | | (0.21) | (0.22) | |
| Constant | -1.751 | -6.251 | -2.470 | 0.838 | -0.292 |
| | (1.84) | (17.7) | (1.88) | (17.9) | (0.47) |
| Observations | 4109 | 4109 | 4109 | 4109 | 4109 |
| R-squared | 0.08 | 0.08 | 0.08 | 0.08 | 0.03 |

Groups are defined according to job title. Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with "1" representing that the individual is up to date on recommended screenings. All observations represent a single individual employed by a large medical organization. Vector of comorbidities and mean by group (as in Table 5), and individual variables (as in Table 4) included but not reported.

Table 7. Knowledge

| COEFFICIENT | (1) up2date | (2) up2date | (3) up2date | (4) up2date | (5) up2date |
|---|---|---|---|---|---|
| up2date | 0.535** | 0.572*** | 0.532*** | 0.633*** | 0.633*** |
| | (0.22) | (0.17) | (0.18) | (0.16) | (0.16) |
| bcless30 | | | 0.178 | | 0.185* |
| | | | (0.12) | | (0.098) |
| meanbcriskless30 | | | 1.323 | | -0.238 |
| | | | (1.51) | | (0.75) |
| Constant | -1.751 | 0.222 | -26.03 | -3.778*** | -2.847 |
| | (1.84) | (1.48) | (26.4) | (1.44) | (12.9) |
| Observations | 4109 | 1806 | 1806 | 2303 | 2303 |
| R-squared | 0.08 | 0.07 | 0.07 | 0.09 | 0.09 |
| Mean of uptodate | | 0.6611 | 0.6611 | 0.6773 | 0.6773 |

Groups are defined according to job title. Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with "1" representing that the individual is up to date on recommended screenings. All observations represent a single individual employed by a large medical organization. Vector of comorbidities (as in Table 5), and individual variables (as in Table 4) included but not reported.

Table 8. Location and Job classification

| COEFFICIENT | (1) Distance | (2) Distance | (3) Distance |
|---|---|---|---|
| groupflag | 1.97351 | 1.90479 | -0.01267 |
| | (0.0193) | (0.0192) | (0.0027) |
| full | | -0.74771 | 0.00069094 |
| | | (0.0112) | (0.0016) |
| hourly | | 3.81475 | -0.12613 |
| | | (0.0222) | (0.0028) |
| famsize | | 0.4008 | 0.00026059 |
| | | (0.0033) | (0.0005) |
| age | | -0.12437 | -0.00015982 |
| | | (0.0008) | (0.0001) |
| divorced | | -4.19344 | 0.02972 |
| | | (0.0376) | (0.0057) |
| married | | -2.83876 | 0.00656 |
| | | (0.0365) | (0.0056) |
| unmarried | | -5.3717 | 0.04526 |
| | | (0.0382) | (0.0057) |
| widowed | | -2.8982 | 0.03741 |
| | | (0.0449) | (0.0065) |
| separated | | -6.99372 | 0.02365 |
| | | (0.0724) | (0.0091) |
| Constant | 18.26626 | 23.5936 | 1.30699 |
| | (0.0038) | (0.0608) | (0.0089) |
| Observations | 16879772 | 16879772 | 1036800 |
| R-Squared | 0 | 0.014 | 0.003 |

Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is distance between two employees. Groupflag indicates if two individuals are in same occupation category. Columns 1 and 2 include all individual pairs. Column 3 includes individual pairs that live within 2 miles.

Table 9a. Geography (with weight matrix (13))

| COEFFICIENT | (1) up2date | (2) up2date | (3) up2date | (4) up2date | (5) up2date |
|---|---|---|---|---|---|
| up2date | 0.667*** | 0.631*** | 0.570*** | 0.586*** | 0.535** |
| | (0.087) | (0.10) | (0.11) | (0.11) | (0.22) |
| Constant | -0.190 | -1.819** | -2.035** | -1.909** | -1.751 |
| | (0.49) | (0.89) | (0.90) | (0.91) | (1.84) |
| Observations | 4109 | 4109 | 4109 | 4109 | 4109 |
| R-squared | 0.04 | 0.04 | 0.04 | 0.04 | 0.08 |

Groups are defined by distance. All employees of the medical organization that live within 2 miles of the individual in question are included in the group. Influence is weighted by the inverse of distance, as illustrated in matrix (19). Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with "1" representing that the individual is up to date on recommended screenings. All observations represent a single individual employed by a large medical organization. Vector of comorbidities and mean by group (as in Table 5), and individual variables (as in Table 4) included but not reported. Column 1 includes age, family size, mean age and mean family size as controls. Column 2 includes these variables plus marriage variables. Column 3 includes both plus an indicator for fulltime status. Column 4 adds an hourly wage dummy. Column 5 adds the full vector of comorbidities.

Table 9b. Geography (with weight matrix (14))

| COEFFICIENT | (1) up2date | (2) up2date | (3) up2date | (4) up2date | (5) up2date |
|---|---|---|---|---|---|
| up2date | 0.773*** | 0.776*** | 0.751*** | 0.752*** | 0.751*** |
| | (0.11) | (0.11) | (0.11) | (0.11) | (0.11) |
| Constant | -1.081** | -3.976*** | -4.111*** | -4.095*** | -5.140*** |
| | (0.51) | (0.86) | (0.85) | (0.86) | (1.39) |
| Observations | 4108 | 4108 | 4108 | 4108 | 4108 |
| R-squared | 0.03 | 0.04 | 0.05 | 0.05 | 0.09 |

Groups are defined by a combination of occupational title and distance. All employees of the medical organization that live within 2 miles of the individual as well as individuals of the same occupation code are included in the group. Influence is weighted by the inverse of distance and the occupation code, as illustrated in matrix (20). All employees of the medical organization that live within 2 miles of the individual in question are included in the group. Influence is weighted by the inverse of distance, as illustrated in equation 20. Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with "1" representing that the individual is up to date on recommended screenings. All observations represent a single individual employed by a large medical organization. Vector of comorbidities and mean by group (as in Table 5), and individual variables (as in Table 4) included but not reported. Column 1 includes age, family size, mean age and mean family size as controls. Column 2 includes these variables plus marriage variables. Column 3 includes both plus an indicator for fulltime status. Column 4 adds an hourly wage dummy. Column 5 adds the full vector of comorbidities.

Table 10. Sociability

| COEFFICIENT | (1) up2date | (2) up2date | (3) up2date |
|---|---|---|---|
| sind2 | 0.0587 | | |
| | (0.0400) | | |
| sind3 | 0.110*** | | |
| | (0.0370) | | |
| sind4 | 0.114*** | | |
| | (0.0390) | | |
| up2date | 0.540*** | 0.610*** | 0.675*** |
| | (0.1100) | (0.1800) | (0.1600) |
| Constant | -3.596 | -16.8 | 2.738 |
| | (10.6000) | (13.7000) | (18.5000) |
| Observations | 4109 | 1973 | 2136 |
| R-squared | 0.08 | 0.1 | 0.07 |

Estimates using weight matrix (19) with information on "sociability" of jobs. Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with “1” representing that the individual is up to date on recommended screenings. Vector of comorbidities and mean by group (as in Table 5), and individual variables (as in Table 4) included but not reported. All observations represent a single individual employed by a large medical organization.

Table 11. Nonlinear (Probit) Model

| COEFFICIENT | (1)<br>up2date | (2)<br>up2date | (3)<br>up2date | (4)<br>up2date |
|---|---|---|---|---|
| up2date | 0.672*** | 0.635*** | 0.574*** | 0.590*** |
| | (0.087) | (0.100) | (0.11) | (0.11) |
| age | 0.0101*** | 0.00992*** | 0.00991*** | 0.00990*** |
| | (0.0016) | (0.0016) | (0.0016) | (0.0016) |
| famsize | -0.0185*** | -0.0244*** | -0.0243*** | -0.0243*** |
| | (0.0057) | (0.0066) | (0.0066) | (0.0066) |
| meanage | -0.00504 | -0.00316 | -0.000513 | 0.000147 |
| | (0.0096) | (0.0100) | (0.010) | (0.010) |
| meanfamsize | 0.0863*** | 0.137*** | 0.159*** | 0.163*** |
| | (0.033) | (0.041) | (0.044) | (0.044) |
| separated | | -0.0478 | -0.0479 | -0.0483 |
| | | (0.15) | (0.15) | (0.15) |
| divorced | | 0.0337 | 0.0337 | 0.0335 |
| | | (0.076) | (0.076) | (0.076) |
| married | | 0.0976 | 0.0985 | 0.0983 |
| | | (0.080) | (0.080) | (0.080) |
| unmarried | | 0.0761 | 0.0760 | 0.0758 |
| | | (0.072) | (0.072) | (0.072) |
| widowed | | 0.0260 | 0.0272 | 0.0272 |
| | | (0.090) | (0.090) | (0.089) |
| meanseparated | | -0.560 | -0.599 | -0.768 |
| | | (1.28) | (1.28) | (1.30) |
| meandivorced | | 1.779** | 1.775** | 1.551* |
| | | (0.83) | (0.83) | (0.89) |
| meanmarried | | 1.367* | 1.292 | 1.046 |
| | | (0.80) | (0.80) | (0.87) |

Continued on next page...

Table 11. Nonlinear (Probit) Model contd.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| COEFFICIENT | up2date | up2date | up2date | up2date |
| meanunmarried | | 1.476* | 1.426* | 1.247 |
| | | (0.81) | (0.81) | (0.85) |
| meanwidowed | | 0.595 | 0.576 | 0.280 |
| | | (0.91) | (0.92) | (1.00) |
| full | | | 0.0128 | 0.0129 |
| | | | (0.024) | (0.024) |
| meanfull | | | 0.135 | 0.152 |
| | | | (0.11) | (0.11) |
| hourly | | | | -0.0169 |
| | | | | (0.13) |
| meanhourly | | | | 0.0627 |
| | | | | (0.14) |
| Observations | 4109 | 4109 | 4109 | 4109 |
| R-squared | . | . | . | . |

Estimates obtained using a probit model. Coefficients shown are implied probabilities. Robust standard errors in parentheses. Significance at 10, 5, and 1 percent level denoted by *,**, and *** respectively. Dependent variable is an indicator with "1" representing that the individual is up to date on recommended screenings. All observations represent a single individual employed by a large medical organization. Vector of comorbidities and their mean by group (as in Table 5) included but not reported.

## References


[1] Acton, JP. Nonmonetary factors in the demand for medical services. *Journal of Political Economy*1975; 83; 595-614.

[2] Aiken, LS, Fenaughty, AM, West, SG, Johnson, JJ, Luckett, TL. Perceived determinants of risk for breast cancer and the relations among objective risk, perceived risk, and screening behavior over time. *Women's Health* 1995; 1; 27-50.

[3] Aizer, A, Currie, J. Networks or Neighborhoods? Correlations in the Use of Publicly-funded Maternity Care in California. *Journal of Public Economics* 2004; 88; 2573-2585.

[4] Alexander FE, Chamberlain J, Forrest AP. Screening for breast cancer in the under 50s. *Lancet* 1993; 341(8853); 1149.

[5] Arlo-Costa, H, Helzner, J. (2005). Comparative Ignorance and the Ellsberg Phenomenon. *Proceedings of the 4th International Symposium on Imprecise Probabilities and Their Applications, Pittsburgh PA.*

[6] Beaulieu, MD, Beland, F, Roy, D, Falardeau, M, Hebert, G. Factors determining compliance with screening mammography. *Canadian Medical Association Journal* 1996; 154; 1335-1343.

[7] Blackman, DK, Bennet, EM, Miller, DS. Trends in Self-Reported Use of Mammograms (1989-1997) and Papanicolaou Tests (1991-1997) – Behavioral Risk Factor Surveillance System. *Morbidity and Mortality Weekly Report* 1999; 48(SS06); 1-22.

[8] Bloom, JR, Spiegel, D. The relationship of two dimensions of social support to the psychological well-being and social functioning of women with advanced breast cancer. *Social Science & Medicine* 1984; 19; 831-837.

[9] Blustein, J, Weiss, LJ. The use of mammography by women aged 75 and older: factors related to health, functioning, and age. J*Journal of the American Geriatrics Society*, 1998a; 46(8); 941-946.

[10] Blustein, J, Weiss, LJ. Visits to specialists under medicare: Socioeconomic advantage and access to care. *Journal of Health Care for the Poor and Underserved* 1998b; 2; 153-169.

[11] Bottorff, JL, Richardson, C, Balneaves, LG, McCullum, M, Buxton, JA, Ratner, PA, Hack, T. Unraveling Women's Perceptions of Risk for Breast Cancer. *Health Education Research* 2004; 19; 469-475.

[12] Bränström, R, Kristjansson, S, Ullén, H. Risk Perception, Optimistic Bias, and Readiness to Change Sun Related Behaviour. *The European Journal of Public Health* 2006; 16; 492-497.

[13] Brock, W, Durlauf, S. Discrete Choice with Social Interactions. *The Review of Economic Studies* 2001; 68(2); 235-260.

[14] Brock, W., Durlauf, S. Interactions-Based Models. In J. Heckman and E. Learner, eds., *Handbook of Econometrics, Volume 5.* 2001. Elsevier Science, Chapter 54.

[15] Burack, R, George, J, Gurney, J. Mammography use among women as a function of age and patient involvement in decision-making. *Journal of American Geriatric Society* 2000; 48(7); 817-821.

[16] Burke, M, Heiland, F. Social Dynamics of Obesity. *Economic Inquiry*, forthcoming 2007.

[17] Calle, EE, Miracle-McMahill, HL, Moss, RE, Heath, CW. Personal contact from friends to increase mammography usage. *American Journal of Preventive Medicine* 1994; 10; 361-366.

[18] Clemow, L, Costanza, ME, Haddad, WP, Luckmann, R, White, MJ, Klaus, D, et al. Underutilizers of mammography screening today: characteristics of women planning, undecided about, and not planning a mammogram. *Annals of Behavioral Medicine* 2000; 22; 80-88.

[19] Coffey, RM. "The effect of time price on the demand for medical care services." *Journal of Human Resources* 1983; 18; 407-424.

[20] Christakis N, Fowler, J. The Spread of Obesity in a Large Social Network over 32 Years. *The New England Journal of Medicine* 2007; 357; 370-379.

[21] Curley, S, Yates, J, Abrams, R. Psychological Sources of Ambiguity Avoidance. *Organizational Behavior and Human Decision Processes* 1986; 38; 230-256.

[22] Deri, C. Social networks and health service utilization. *Journal of Health Economics* 2005; 24(6); 1076-1107.

[23] Diez Roux, AV. Estimating neighborhood health effects: The challenges of causal inference in a complex world. *Social Science and Medicine* 2004; 58(10); 1953-1960.

[24] Duesenberry, J.*Income, Saving and the Theory of Consumer Behavior.* Cambridge University Press 1949.

[25] Elixhauser, A, Steiner, C, Harris, D, Coffey, R. Comorbidity Measurs for Use with Administrative Data. *Medical Care* 1998; 36(1); 8-27.

[26] Epstein, L., and T. Wang. Intertemporal Asset Pricing Under Knightian Uncertainty. *Econometrica* 1994; 62(2); 283-322.

[27] Facione, N, Katapodi, M. Cultures as an influence on breast cancer screening and early detection. *Seminars in Oncology Nursing* 2000; 16(3); 238-247.

[28] Fernandez, M., R. Palmer, and C. Leong-Wu. Repeat mammography screening among low-income and minority women: a qualitative study. *Cancer Control* 2005; 12 Suppl. 2; 77-83.

[29] Fite, S, Frank, DI, Curtin, J. The relationship of social support to women's obtaining mammography screening. *Journal of the American Academy of Nurse Practitioners* 1996; 8; 565-569.

[30] Fletcher, J. Social interactions in adolescent television viewing. *Archives of Pediatrics and Adolescent Medicine* 2006; 160(4); 383-386.

[31] Fox, C, Tversky, A. Ambiguity Aversion and Comparative Ignorance. *Quarterly Journal of Economics* 1995; 110; 585-603.

[32] Fox, C, Weber, M. Ambiguity Aversion, Comparative Ignorance, and Decision Context. *Organizational Behavior and Human Decision Processes* 2002; 88; 476-498.

[33] Frazier, E, Jiles, R, Mayberry, R. Use of Screening Mammography and Clinical Breast Examinations among Black, Hispanic, and White Women. *Preventive Medicine* 1996; 25(2); 118-125.

[34] Fredman, L, Sexton, M, Cui, Y, Althius, M. Cigarette smoking, alcohol consumption, and screening mammography among women aged 50 and older. *Preventive Medicine* 1999; 28(4); 407-417.

[35] Gotay, CC, Wilson, ME. Social support and cancer screening in African American, Hispanic, and Native American women. *Cancer Practice* 1998; 6; 31-37.

[36] Grossman, M. The Human Capital Model. In A. Culyer and J. Newouse, eds., *Handbook of Health Economics* Elsevier Science. 2000.

[37] Grunfeld, EA, Hunter, MS, Ramirez, AJ, Richards, MA. Perceptions of breast cancer across the lifespan. *Journal of Psychosomatic Research* 2003; 54(2); 141-146.

[38] Hubbell, F, Chavez, L, Mishra, S, Valdez, R. Differing beliefs about breast cancer among Latinas and Anglo women. *Western Journal of Medicine* 1996; 164(5); 405-409.

[39] Kang, SH, Bloom, JR, Romano, PS. Cancer screening among African-American women: their use of tests and social support. *American Journal of Public Health* 1994; 84; 101-103.

[40] Kenkel, DS. The demand for preventive medical care. *Applied Economics* 1994; 26; 313-325.

[41] Lagerlund, M, Widmark, C, Lambe, M, Tishelman, C. Rationales for attending or not attending mammography screening - a focus group study among women in Sweden. *European Journal of Cancer Prevention* 2001; 10; 439-442.

[42] Lairson, DR, Chan, W, Newmark, GR. Determinants of the demand for breast cancer screening among women veterans in the United States. *Social Science and Medicine* 2005; 61; 1608-1617.

[43] Langlie, J. Social Networks, Health Beliefs, and Preventive Health Behavior. *Journal of Health and Social Behavior* 1977; 18(3); 244-260.

[44] Lichtman R, Taylor S, Wood J. Social support and marital adjustment after breast cancer. *Journal of Psychosocial Oncology* 1989; 5; 47-74.

[45] Manski, C. Identification of Endrogenous Social Effects: The Reflection Problem. *The Review of Economic Studies* 1993; 60(3); 531-542.

[46] Maxwell, CJ, Kozak, JF, Desjardins-Denault, SD, Parboosingh, J. Factors important in promoting mammography screening among Canadian women. *Canadian Journal of Public Health* 1997; 88; 346-350.

[47] McCance, KL, Mooney, KH, Field, R, Smith, KR. Influence of others in motivating women to obtain breast cancer screening. *Cancer Practice* 1996; 4; 141-146.

[48] Miller A, Baines, C, To, T, Wall, C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *Canadian Medical Association Journal* 1992; 147(10); 1459-1476.

[49] Muurinen, J. Demand for health: A generalised Grossman model. *Journal of Health Economics* 1982; 1; 5-28.

[50] Oakes, JM. Causal inference and the relevance of social epidemiology. *Social Science and Medicine* 2004a; 58(10); 1969-1971.

[51] Oakes, JM. The (mis)estimation of neighborhood effects: Causal inference for a practicable social epidemiology. *Social Science and Medicine* 2004b; 58(10); 1929-1952.

[52] Oakes, JM. Commentary: Advancing neighbourhood-effects research—selection, inferential support, and structural confounding. *International Journal of Epidemiology* 2006; 35(3); 643-647

[53] Pearlman, DN, Rakowski, W, Ehrich, B. The information environment of women and mammography screening: assessing reciprocity in social relationships. *Journal of Women's Health* 1995; 4; 541-553.

[54] Rahman, SMM, Dignan, MB, Shelton, BJ. Factors influencing adherence to guidelines for screening mammography among women aged 40 years and older. *Ethnicity and Disease* 2003; 13; 477-484.

[55] Rakowski, W, Pearlman, D, Rimer, BK, Ehrich, B. Correlates of mammography among women with low and high socioeconomic resources. *Preventive Medicine* 1995; 24; 149-158.

[56] Roux, A. Estimating neighborhood health effects: the challenges of causal inference in a complex world. *Social Science & Medicine* 2004; 58; 1953-1960.

[57] Silverman, E, Woloshin, S, Schwartz, L, Byram, S, Welch, H, Fischhoff, B. Women's Views on Breast Cancer Risk and Screening Mammography: A Qualitative Interview Study. *Medical Decision Making* 2001; 21(3); 231-240.

[58] Sloan, F, Smith, VK, Taylor, D. *The Smoking Puzzle: Information, Risk Perception, and Choice*. Harvard University Press. 2004.

[59] Spiegel D, Bloom J, Kraemer H, Gottheil E. Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *Lancet* 1989; 2; 888-89.

[60] Stoddard, A, Rimer, B, Lane, D, Fox, S, Lipkus, I, Luckmann, R, Avrunin, J, Sprachman, S, Costanza, M. Underusers of mammogram screening: Stage of adoption in five U.S. subpopulations. *Preventive Medicine* 1998; 27; (3), 478-487.

[61] Suarez, L, Ramirez, AG, Villarreal, R, Marti, J, McAlister, A, Talavera, G, et al. Social networks and cancer screening in four U.S. Hispanic groups. *American Journal of Preventive Medicine* 2000; 19; 47-52.

[62] Subramanian, SV. The relevance of multilevel statistical methods for identifying causal neighborhood effects. *Social Science and Medicine* 2004; 58(10); 1961-1967.

[63] Walter, LC, Lindquist, K, Covinsky, KE. Relationship between health status and use of screening mammography and Papanicolaou smears among women older than 70 years of age. *Annals of Internal Medicine* 2004; 140(9); 681-688.

[64] Woods, A. Multilevel modeling in primary care research. *British Journal of General Practice* 2004; 54(504); 560-561.Q