Paying on the margin for medical care:
Evidence from breast cancer treatments[*]

Liran Einav	Amy Finkelstein	Heidi Williams
Stanford and NBER	MIT and NBER	MIT and NBER

August 4, 2014

Abstract

We present a simple framework to illustrate the welfare consequences of a "top up" health insurance policy that allows patients to pay the incremental price for more expensive treatment options. We contrast it with common alternative policies that require essentially no incremental payments for more expensive treatments (as in the United States), or require patients to pay the full costs of more expensive treatments (as in the United Kingdom). We provide an empirical illustration of this welfare analysis in the context of treatment choices among breast cancer patients, where lumpectomy with radiation therapy is a more expensive treatment than mastectomy, with similar average health benefits. We use variation in distance to the nearest radiation facility to estimate the relative demand for lumpectomy and mastectomy. Extrapolating the resultant demand curve (grossly) out of sample, our estimates suggest that the "top-up" policy, which achieves the efficient treatment decision, increases total welfare by $700-2,500 per patient relative to the current US "full coverage" policy, and by $700-1,800 per patient relative to the UK "no top up" policy. While we caution against putting much weight on our specific estimates, the analysis illustrates the potential welfare gains from more efficient reimbursement policies for medical treatments. We also briefly discuss additional tradeoffs that arise from the top-up and UK-style policies, which both lead to additional (ex-ante) risk exposure.

1 Introduction

Medical expenditures in the US are high and increasing. Policy and academic discussions of strategies to reduce health care spending have largely focused on increasing cost sensitivity either on the demand side through consumer cost-sharing, or on the provider side by making providers the residual claimant on cost savings. A natural economic solution which has not received much attention is a "top-up" design in which health insurance contracts would cover the cost of a baseline treatment, and patients could choose to pay the incremental cost of more expensive treatments out of pocket.

This type of "top-up" design contrasts with the standard "full coverage" insurance design in the United States, in which consumers face essentially no incremental cost of choosing a more expensive treatment (other than perhaps some minimal consumer cost-sharing). Other high-income countries have taken an alternative approach: individual medical treatments deemed "cost-effective" are fully covered, and treatments deemed not to be cost effective are not covered at all. In the UK, for example, the National Institute for Health and Care Excellence (NICE) determines which medical technologies will be covered by the National Health Service (NHS), using – in recent years – a threshold of around \$50,000 per quality-adjusted year of life saved (McCabe, Claxton and Culyer, 2008). This threshold rule results in the NHS not covering some medical treatments: for example, in 2010 NICE refused coverage for the drug Avastin as a treatment for metastatic colorectal cancer on the basis that the drug improved life expectancy by six weeks (relative to the preexisting standard of care) at a cost of around \$115,000 per quality-adjusted year of life saved.[1] As a result, patients in the UK who want to choose a treatment like Avastin must pay the full cost of that treatment. Such UK-style "no top-up" designs have recently been introduced in Australia, France, and Germany (Chalkidou and Anderson, 2009), and received a great deal of negative publicity in the US under the name of "death panels" during the debate over the 2010 Affordable Care Act.[2]

Relative to either the US "full coverage" or the UK "no top up" regimes, a "top-up" design in which individuals internalize treatment costs on the margin achieves the efficient allocation of patients across treatments. Conceptually, this simple point is not new. It has been made in other contexts, such as public subsidies for education (Peltzman, 1973), public health insurance subsidies (Cutler and Gruber, 1996), pricing of employer-provided health insurance plans (Enthoven and Kronick, 1989), and incentives for patients to see specific providers within health insurance plans (Robinson and MacPherson, 2012). However, to our knowledge, it has not received much attention in discussions of insurance coverage for

---

[1] See `http://www.nice.org.uk/media/E58/E7/2010182BevacizumabForColorectalCancerFinalGuidance.pdf` and the discussion in Chandra, Jena and Skinner (2011).

[2] Such negative publicity not withstanding, Pollack (2014) describes some groups of US medical specialists (in particular, for cardiology and oncology) who are recommending that costs be taken into account when developing medical guidelines; insurance companies often use medical guidelines to determine reimbursement policies.

different treatments, with the exception of a recent paper by Baicker, Shephard and Skinner (2012) who use a calibrated simulation model to explore this idea. Moreover, quantifying the welfare consequences of a "top-up" design naturally requires empirical estimates.

In this paper, we present a simple graphical framework to illustrate the welfare consequences of alternative insurance designs for reimbursement of different treatment choices, and explore this welfare analysis empirically in the specific context of treatment choices among breast cancer patients. Most patients diagnosed with breast cancer receive surgery as an initial course of treatment. The key treatment choice is between two types of surgery: mastectomy, which removes the cancerous breast, and lumpectomy, which removes the tumor while preserving the breast and is generally followed by a course of radiation therapy. While evidence from randomized clinical trials has suggested no average difference in survival between mastectomy relative to lumpectomy with radiation (Fisher et al., 1985), mastectomy tends to be less expensive (Polsky et al., 2003).

Public and private insurance in the US typically covers the costs of both treatments fully (or nearly fully) so that patients do not internalize the difference in treatment costs. In principle, under comparative effectiveness regulations – where the goal is stated as covering the lowest-cost option attaining the best health outcome (Chandra, Jena and Skinner, 2011) – mastectomy would be covered by insurance whereas lumpectomy with radiation would not. Because the latter is a more costly treatment with no evidence of superior average health outcomes, patients choosing it would face the full cost of the treatment.[3] In contrast, a top-up policy in this context is analogous to an indemnity insurance policy that pays out if a patient is diagnosed with breast cancer, at a fixed sum equal to the cost of a mastectomy.[4] While we focus on the example of a comparative effectiveness regulation which uses a threshold of zero, as long as the incremental value for society is less than the incremental cost, the same qualitative analysis holds for a threshold above zero (such as the UK threshold of $50,000 per quality-adjusted year of life saved).

The key empirical object needed to evaluate the welfare consequences of these three insurance designs is the (relative) demand curve for the more expensive treatment option, which in this case is lumpectomy. We know of no useful variation in the relative price for lumpectomy. Instead, we estimate the relative demand curve for lumpectomy using variation across patients in the distance between their residence at the time of diagnosis and the nearest radiation clinic.[5] A standard course of post-lumpectomy radiation

---

[3] Although the comparative effectiveness literature recognizes the potential that health benefits may be heterogeneous, in principle (if not in practice) that limitation could be solved by randomized clinical trials that are sufficiently powered to detect such heterogeneity. Such approaches could not however, even in principle, address heterogeneity in preferences over non-health aspects of treatment. Recent work in social insurance has emphasized the potential importance of heterogeneous preferences both conceptually (Feldstein, 1995) and empirically (Einav, Finkelstein and Schrimpf, 2010).

[4] Of course, such alternative insurance designs for breast cancer may not be particularly realistic from a political perspective. We focus on breast cancer because of the empirical traction it offers for estimating demand for alternative treatments.

[5] The use of distance between patients and providers as variation dates back in the health economics literature at least

therapy requires 25 round-trips to a radiation facility, spread over 5 weeks. Our key economic assumptions are that travel time can be monetized and that preferences for reduction in travel time are analogous to preferences for any other equivalent price difference. These assumptions allow us to use the variation in distance to the radiation facility as if it were variation in the relative price of lumpectomy, thus identifying the demand curve.

We analyze cancer registry administrative data on the characteristics and treatment choices of over 300,000 breast cancer patients initially diagnosed in California between 1997 and 2009, linked to data on the location of radiation treatment facilities over the same time period. Building on similar results in the medical literature (Schroen et al., 2005), we document that women living further away from radiation facilities at the time of their breast cancer diagnosis are more likely to choose mastectomies rather than lumpectomies. Our key econometric assumption is that there are not omitted patient characteristics correlated with both distance and demand for lumpectomy. We observe a rich set of patient demographic and clinical characteristics and find that while some of these variables vary with distance, the magnitude of the relationship between treatment choice and travel time is not very sensitive to their inclusion. Our baseline estimates imply that a 10 minute increase in one-way travel time (approximately two-thirds of a standard deviation in our data) reduces the probability of a lumpectomy by about 0.7 to 1.1 percentage points, relative to a baseline lumpectomy rate of about 58 percent.

We then use the estimated demand curve to illustrate how the welfare effects of alternative insurance designs can be quantified, albeit highly out of sample relative to our observed variation in "price." We estimate, for example, that the efficient "top-up" policy – in which patients pay $10,000 on the margin for a lumpectomy – increases the lumpectomy rate by 15-25 percentage points relative to the UK-style "no top-up" regime, and decreases the lumpectomy rate by 35-40 percentage points relative to the US-style "full coverage" regime. Our estimates suggest total welfare gains from the "top-up" policy of between $700 and $1,800 per patient relative to a "no top-up" UK-style policy and between $700 and $2,500 per patient relative to a "full coverage" US-style policy.

In the final section of the paper, we briefly consider the additional tradeoffs faced when the welfare analysis is done from an ex-ante perspective, thus accounting for differential risk exposure across regimes. Qualitatively, the top-up policy continues to dominate the UK-style no top-up policy, but the relative ranking of the top-up policy and the US-style full coverage policy is now ambiguous. We present a simple and highly stylized calibration which shows that, for high enough levels of risk aversion, ex-ante welfare can be higher under a US-style full coverage policy than under a top-up policy that produces the ex-post efficient treatment decisions.

---

to the work of McClellan, McNeil and Newhouse (1994).

The paper is organized as follows. Section 2 details our empirical setting – treatments for breast cancer – and describes our data. Section 3 outlines our conceptual framework, and describes our empirical strategy. Section 4 presents our main results. Section 5 briefly discusses implications for ex-ante welfare, and Section 6 concludes.

2 Setting and data

2.1 Treatment choices for breast cancer

Our analysis is focused on the treatment choice made by breast cancer patients.[6] Most patients diagnosed with breast cancer have surgery to attempt to remove the cancer from the breast; in our data, 95 percent of women diagnosed with breast cancer receive surgery as an initial course of treatment. For women receiving surgery, the key treatment choice is between two alternatives, lumpectomy and mastectomy. Lumpectomies are breast-conserving surgeries that remove the cancer but not the breast itself, and are generally followed by a course of radiation therapy. Mastectomies, in contrast, remove the entire cancerous breast and are generally not followed by a course of radiation therapy. Other forms of treatment such as chemotherapy are commonly administered either before or after (or both) either type of surgery.

In terms of clinical effectiveness, the key comparative evidence on these treatments comes from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B06 clinical trial, which enrolled women with stage I and stage II breast tumors. The initial results of this trial were released in 1985 (Fisher et al., 1985), with subsequent follow-up results (with longer-term mortality outcomes) published in 1989 (Fisher et al., 1989), 1995 (Fisher et al., 1995), and 2002 (Fisher et al., 2002). The results of this clinical trial suggested there were no detectable differences in survival outcomes across random assignment to more invasive (total mastectomy) versus less invasive (lumpectomy with or without radiation therapy) treatments.

In terms of overall financial costs (shared by both patients and insurers), there is a consensus that lumpectomy with radiation is more expensive than mastectomy. Mastectomy is cheaper primarily because of the add-on cost of the radiation therapy which accompanies lumpectomies. A common argument is that the more relevant costs are not those associated with the initial course of treatment, but rather the total, subsequent "lifetime" costs over the following years. While comparing these longer-run costs of the two treatments is more difficult due to various selection concerns – for example, co-morbidities that vary with treatment choice – the evidence on lifetime costs also strongly suggests that mastectomy is cheaper, even

---

[6]Many of the clinical details in this section are drawn from the National Cancer Institute's guide to the treatment of breast cancer; see `http://www.cancer.gov/cancertopics/pdq/treatment/breast/Patient/`.

after attempting to correct for potential selection bias. For example, Polsky et al. (2003) compare five-year total Medicare payments across Medicare patients with breast cancer receiving lumpectomy relative to mastectomy: the unadjusted difference between mastectomy and lumpectomy was \$8,389 (relative to a baseline of \$40,130 for mastectomy), the risk-adjusted difference was \$13,775 (relative to a baseline of \$38,623 for mastectomy), and the propensity score-adjusted difference was \$14,054 (relative to a baseline of \$38,664). In all three versions of their analysis, the 95-percent confidence intervals can reject cost differences smaller than \$4,500.[7]

An additional difference in cost between the two treatments, which our empirical exercise will focus on, is the time cost of traveling to receive post-surgery radiation therapy associated with lumpectomy.[8] A standard course of radiation therapy requires 25 treatments spread over 5 weeks.[9] Motivated by this substantial time commitment required for radiation therapy, several papers in the medical literature have explored whether women living further away from radiation facilities are more likely to choose to have mastectomies rather than lumpectomies with radiation. For example, Schroen et al. (2005) use data from the Virginia cancer registry from 1996-2000 ($\sim$20,000 patients) and document that the probability of patients choosing mastectomy increases with distance from the radiation facility: in their full sample, 43% of women choose mastectomy if they live within 10 miles of a radiation facility, whereas among women living more than 50 miles from a radiation facility the share is 58%.[10] Our empirical work will build on this medical literature by confirming a relationship between distance from radiation facilities and treatment choice in a much larger sample of over 300,000 patients in California, and using this variation to estimate a demand curve for lumpectomy. This estimated demand curve will in turn be the key input into policy counterfactual exercises investigating how patients might respond to changes in the financial costs of treatments induced by different health insurance contract designs.

---

[7]In our illustrative quantitative welfare analyses below, we assume that the higher payments for lumpectomy relative to mastectomy reflect differences in underlying resource costs (i.e. social marginal cost). To the extent that prices paid are distorted relative to social marginal costs, our results below would naturally need to be adjusted.

[8]While our cancer registry data offers an incomplete set of information on cancer treatments, as observed in our data more than 60 percent of patients choosing lumpectomy receive radiation therapy during their initial course of treatment, compared to less than 20 percent of mastectomy patients.

[9]See this US National Cancer Institute Cancer Bulletin from 2010: `http://www.cancer.gov/aboutnci/ncicancerbulletin/archive/2010/022310/page2`.

[10]Nattinger et al. (2001) and Celaya et al. (2006) document similar patterns in the SEER cancer registry (1991-1992, $\sim$17,000 patients) and the New Hampshire cancer registry data (1998-2000, $\sim$3,000 patients), respectively. Athas et al. (2000) investigate this relationship in the New Mexico tumor registry data (1994-1995, $\sim$1,000 patients) and do not find evidence of a similar relationship, although they do find that the probability of receiving radiation post-lumpectomy falls with distance from a radiation facility.

2.2 Data

Our empirical analysis uses two datasets from the state of California: a patient-level cancer registry dataset, and data on radiation treatment facility locations.

Our patient-level data is drawn from the California Cancer Registry (CCR), a program of the California Department of Public Health. The CCR was established in 1985, and every cancer diagnosis made in California from 1988 forward is required by law to be reported to the CCR.[11] Data are collected directly from cancer patients' medical records at the time of the cancer diagnosis, rather than by interviewing patients. Available variables in the CCR research database include demographic covariates such as age, race, and sex; diagnostic information such as cancer type and stage of disease; and treatment information on the first course of treatment received by the patient (if any). A key advantage of the CCR research database relative to other cancer registry databases such as the SEER cancer registry is that the CCR data include patients' exact address of residence at the time of diagnosis, which enables our empirical analysis to rely on a more precise measure of how far patients live from radiation treatment facilities than would be possible if we only observed county of residence, as is available in the SEER data.

Our data on radiation treatment facility locations comes from the private firm IMV. IMV aims to identify all hospital and non-hospital sites in the US performing radiation therapy, and queries these sites with a telephone survey. The sampling frame for the telephone survey is constructed from several sources, including federal and state nuclear licensing lists.[12] We obtained data on the full sampling frame of California sites, including exact street address for all institutions, for all available survey years (1996 to 2011).[13] Because survey response rates vary across years (ranging from 45 to 87 percent), we use all institutions in the sampling frame as our set of facilities of interest, regardless of whether the institution responded to the survey.

We restrict the CCR data to female breast cancer patients diagnosed between 1997 and 2009, which covers all years after the first IMV survey (1996) and until the last year of the CCR cancer registry data (2009). Following sample restrictions used in National Cancer Institute analyses of breast cancer

---

[11]See http://www.ccrcal.org/pdf/Reports/Physicians.pdf for more details on these reporting requirements, which fall under California Health and Safety Code 103885.

[12]Specifically, IMV reports that it identifies candidate sites from nuclear licensing lists compiled by the Nuclear Regulatory Commission, lists from state licensing agencies, the American Hospital Association *Guide to the Health Care Field* database, internal IMV lists, previous IMV data collections, and internal IMV internet research. Because many radiation facilities are based at non-hospital sites, the IMV data are preferable to relying solely on hospital-based datasets such as the American Hospital Association annual survey data. The IMV data have been used in several previous papers, such as Baker (2001), Baker and Atlas (2004), and Baker, Atlas and Afendulis (2008).

[13]Surveys are conducted approximately every one to two years over this time period. Specifically, IMV's data collection periods are 1996 (covering 2/96 to 1/97), 1998 (covering 3/98 to 9/98), 2000 (covering 3/00 to 10/00), 2001 (covering 11/01 to 7/02), 2003 (covering 11/02 to 1/04), 2004 (covering 8/04 to 12/05), 2006/08 (covering 9/06 to 10/08), and 2010/11 (covering 12/09 to 11/11).

registry data (Gloeckler Ries and Eisner, 2007), we exclude cases identified through autopsy and death certificate only (<0.5% of observations), and cases where the age at diagnosis was less than 20 (<0.1% of observations). Given our empirical strategy, we also limit the sample to patients with non-missing data on treatment (<0.3% of observations) and non-missing data on residence at the time of diagnosis (<0.2% of observations).

For ease of presentation, because our primary analysis is focused on the choice between lumpectomy and mastectomy, we also omit from the baseline sample the 5.8 percent of the cases in which the patient chose neither of these two surgical treatments. Our results are not sensitive to this sample selection, and the appendix presents analogous results for the full sample (Appendix Table A.1).

Finally, combining these two datasets, for each patient we computed (using Google Maps in summer of 2012) the distance between her residence at the time of diagnosis and the nearest facility offering radiation treatment as recorded in the IMV sampling frame as of January 1st in the year of diagnosis. Our baseline analysis uses driving time, while the appendix reports results that are based on driving distance and spherical distance; our results are not sensitive to the choice of distance measure (see Appendix Figures A.1 and A.2, and Appendix Tables A.2 and A.3).

To summarize, our baseline sample covers 323,612 breast cancer patients that were diagnosed between 1997 and 2009 and chose either mastectomy or lumpectomy as their initial treatment, with each patient matched to her nearest radiation facility.

3 Conceptual framework and empirical strategy

Consider a woman i recently diagnosed with breast cancer, facing a binary choice between receiving a mastectomy (M) or receiving a lumpectomy together with radiation therapy (L). The key input into the analysis of the welfare effect of alternative reimbursement policies is the relative valuation (or willingness to pay) for L, given by

$$v_i \equiv v_{i,L} - v_{i,M}, \tag{1}$$

and its distribution across cancer patients, which is given by the cumulative distribution function $F(v_i)$. That is, our main empirical object of interest can be summarized by the demand curve for lumpectomy, which is illustrated in Figure 1, and is given by $F^{-1}(\cdot)$. Variation in the relative valuation of L across patients may reflect heterogeneity in relative health benefits, or heterogeneity in relative valuation of non-health attributes of the two treatments, or both.

Abstracting from income effects, this demand curve $F^{-1}(\cdot)$ is sufficient to evaluate the welfare effects

of alternative policy structures. Note that our discussion of welfare and efficiency here refers to efficiency of treatment choice, or ex-post efficiency, a point we return to in Section 5. With that in mind, the efficient policy is to cover both treatments, but to require cancer patients to incur the incremental costs associated with lumpectomy. We denote this policy by "top-up" in Figure 1, which denotes the incremental (social) cost of L by c, and the incremental price the patient faces for L by p. The top-up policy sets $p = c$. The resultant allocation, point E in Figure 1, occurs when $F^{-1}(\cdot) = c$. Since patients internalize the social marginal cost of treatment, the allocation is efficient; all patients whose incremental willingness to pay for lumpectomy (v_i) is above the social marginal cost of the lumpectomy (c) would choose it, and all those for whom $v_i < c$ would not.

The US-style "full coverage" policy is given by point D in Figure 1; cancer patients can choose between M and L and do not face any of the incremental financial cost associated with L $(p = 0)$. Because lumpectomy is more expensive than mastectomy, this policy produces inefficient treatment decisions. A set of cancer patients whose relative valuation for L is lower than c inefficiently choose L because they do not pay the cost c (paid by the insurer) associated with it. This welfare loss is summarized in Figure 1 by the triangle CDE.

Finally, we consider a third possible policy (denoted "No top-up" in Figure 1) in which the insurance policy covers only the choice of the cheaper treatment M. In this situation, cancer patients may still choose L, but if they do so will have to pay its entire cost out of pocket. As detailed in the introduction, the UK uses this type of "no top-up" insurance regime, and other high-income countries have been moving towards similar frameworks. In our setting, such a "no top-up" policy would cover the patient's full cost of mastectomy, but would not reimburse any costs if the patient chooses lumpectomy. This would lead to a welfare loss that is summarized in Figure 1 by triangle ABE: a set of cancer patients who prefer lumpectomy would opt into mastectomy in order to avoid the financial cost, despite the fact that $v_i > c$, thus implying that a choice of L would have been socially efficient. This is analogous to the classic welfare analyses of requiring individuals who opt out of the public schools to pay the full cost of private schooling (Peltzman, 1973) or requiring individuals who opt out of public insurance to pay the full cost of private insurance (Cutler and Gruber, 1996).

In the rest of the paper we explore these tradeoffs quantitatively. Figure 1 makes clear that the key empirical object for welfare analysis is the demand curve for lumpectomy (relative to mastectomy). The demand curve is derived from v_i, individuals' incremental willingness to pay for L. Our empirical strategy is therefore focused on estimating this demand curve.

Estimation of the demand curve requires identifying variation in the relative price of lumpectomy. Motivated by the medical literature discussed in Section 2.1, which has documented an empirical rela-

tionship between distance from radiation facilities and breast cancer treatment choices, we use variation across patients in the distance to the nearest radiation facility in order to estimate the demand curve. Specifically, we normalize patient i's utility from mastectomy to zero and assume that her (relative) utility from lumpectomy is given by

$$u_i \equiv \alpha_i - \beta_i(\theta_i d_i + p), \tag{2}$$

where α_i and β_i are the (potentially patient-specific) preference parameters, d_i is the distance of patient i to the nearest radiation facility, and p is the incremental price she would need to pay for lumpectomy (relative to mastectomy). Distance is denominated in miles or travel hours while price is denominated in dollars, so the parameter θ_i captures the opportunity cost of time, and thus serves as a simple conversion factor that allows us to monetize distance/time. The patient would choose lumpectomy if and only if $u_i > 0$. From the econometrician's perspective, we obtain

$$Pr(\text{Lumpectomy}) = Pr(u_i > 0). \tag{3}$$

An important assumption in this specification, and one that is crucial for our empirical strategy, is that θ_i can be calibrated using external information so that, conditional on θ_i, price and (monetized) distance have the same effect on individual utility. This assumption allows us to estimate the distribution of α_i and β_i in the data using variation in distance only, but then use the estimated distributions to assess the impact of counterfactual policy designs that change price. Because the out-of-pocket price from either treatment is effectively zero in our California data, this assumption or the choice of θ_i does not affect estimation; it only becomes relevant in the counterfactual exercises. To see this, replace $p = 0$ in equation (2), and define $\beta_i' = \beta_i \theta_i$, to obtain

$$u_i \equiv \alpha_i - \beta_i' d_i, \tag{4}$$

which has a familiar form. We assume that $\alpha_i = x_i' \gamma_i^\alpha + \varepsilon_i$ and that $\beta_i = x_i' \gamma_i^\beta + u_i^\beta$ and that ε_i follows a type I extreme value distribution. These assumptions lend themselves to a standard logit regression when $u_i^\beta = 0$ and to a random-coefficient logit model otherwise.

4 Results

4.1 Summary statistics and initial evidence

Table 1 presents some summary statistics for the baseline sample. We aggregate the detailed treatment information in the cancer registry data into indicator variables for whether the patient received a lumpectomy or a mastectomy as their initial course of treatment: 58% of women receive lumpectomy and 42% receive mastectomy.[14] A standard course of radiation consists of 25 round trips. Patients would need, on average, to drive eight miles (11 minutes) each way to the nearest radiation facility. However, there is a fair amount of variation along this dimension, with many patients living within a few minutes of a facility, while others would be required to drive more than half an hour each way for each treatment.

An attraction of our data is the availability of a relatively rich set of covariates measuring patient demographics and clinical characteristics. Specifically, we observe patient age, race, marital status, and some information on type of insurance coverage. While the data lack individual-level measures of covariates such as income or educational attainment, the cancer registry data matches on these and other characteristics at the census-block level. In terms of clinical covariates measuring cancer severity, we observe data on the two primary measures of the extent of the disease at the time of diagnosis which are used by physicians to guide patient treatment decisions: stage and grade. In situ corresponds to an early stage (sometimes called "pre-cancer"), and local, regional, and remote correspond to increasing extents of disease.[15] Grade is an alternative measure of the extent of disease at the time of diagnosis, with higher values corresponding to increasing extents of disease.[16]

Table 2 splits the sample by above and below median travel time from a radiation facility. Women who live further from a radiation facility are more likely to receive mastectomies and less likely to receive lumpectomies. However, as in Schroen et al. (2005) and other previous work, women closer to and further from radiation facilities also appear to differ on observable characteristics. In terms of demographic characteristics, women living closer to radiation facilities tend to be older, less likely to be white, less likely to be married, and more likely to have Medicaid as a primary payment source.[17] In terms of neighborhood characteristics, women living closer to radiation facilities tend to live in slightly poorer neighborhoods (as measured by income), but also in areas with slightly higher median home values. In terms of clinical characteristics, the above- and below-median distance samples appear relatively more balanced on stage

---

[14]Only the initial course of treatment is recorded in the cancer registry data. Our coding of lumpectomy and mastectomy follows Roetzheim et al. (2008).

[15]For more details, see the SEER training website: `http://training.seer.cancer.gov/ss2k/staging/review.html`.

[16]For more details, see the SEER instructions for coding grade: `http://seer.cancer.gov/tools/grade/`.

[17]The California cancer registry data also includes information on non-Medicaid payment sources: around 16% of women are covered by Medicaid, 26% by Medicare, 56% by private payers, and the small remainder (<2%) self-pay or other sources.

and grade of disease: while several of the differences in grade indicators are statistically significant, no clear pattern emerges in terms of one group being in "better" health in terms of cancer progression at the time of diagnosis. Overall, Table 2 provides preliminary support for the idea that distance may affect treatment choices among breast cancer patients, but also highlights the need to examine the robustness of this relationship to conditioning on demographic, neighborhood, and clinical characteristics, as distance to radiation facility is clearly correlated with other patient characteristics which may themselves affect treatment choice.

To explore how travel distance to a radiation facility affects the treatment choice of breast cancer patients in our sample, Figure 2(a) plots the probability of mastectomy and the probability of lumpectomy by travel time to the nearest radiation facility. The histogram in light gray displays the number of observations (patients) in each travel time bin in our sample. These raw data on treatment choices display the expected pattern: women who live further from radiation facilities are more likely to receive mastectomies and less likely to receive lumpectomies.

Figure 2(b) investigates whether this relationship between distance and treatment choice can be explained by the differences in demographic or health characteristics of the patients that we saw in Table 2. It presents a series of plots which residualize the y-axis (lumpectomy) for various covariates; to retain comparability with Figure 2(a) we do not residualize the x-axis (distance), but our regression specifications below which condition out these covariates paint a similar picture. The first line (dashed and square denoted) presents the de-meaned lumpectomy rates for each travel time bin as a point of comparison. The other two lines show the residualized lumpectomy estimates after sequentially adding more covariates. The second line (solid and triangle denoted) conditions out patient characteristics, and the third (dashed and circle denoted) adds neighborhood-level covariates and clinical characteristics. Consistent with what we will document in the regression specifications below, Figure 2(b) suggests that the overall relationship between treatment choice and travel time is not very sensitive to the inclusion of these covariates.

4.2 Treatment choices by distance

Table 3 quantifies the treatment-distance relationship, estimating different specifications of the logit regression in equation (4). For ease of interpretation, Table 3 reports our estimate of the average marginal effect on lumpectomy probability of a ten-minute increase in (one way) travel time, with bootstrapped standard errors clustered at the county level (50 iterations); recall from Table 1 that this ten-minute increment is about two-thirds of a standard deviation of travel time in our sample. We also report the standard deviation of this effect across patients. In the appendix, we report the parameter estimates (and

their standard errors) that give rise to these average marginal effects (see Appendix Table A.4).

Column (1) of Table 3 reports the simplest specification, where we estimate a logit model of whether the patient chose lumpectomy on travel time with no controls. Using the notation of Section 3, column (1) assumes that neither α_i nor β_i are affected by any patient-specific variables. Columns (2), (3), and (4) retain the same (homogenous logit) assumption for β_i but increasingly add covariates that shift α_i, thus affecting the mean utility from lumpectomy. In column (5) we also allow these observables to change β_i by adding interaction terms between these covariates and distance. Finally, column (6) reports results from a specification that allows random coefficients on distance: it assumes that β_i follows a lognormal distribution, thus relaxing the assumption of $u_i^\beta = 0$ that is assumed in all other specifications.

The effect of distance is statistically significant and is quantitatively reasonably stable across all specifications, as would be expected given the patterns we documented in Figure 2. The specification of column (6) leads to the largest effect, but this estimate is quite noisy (as the estimated parameters that govern the distribution of the random coefficient are imprecisely estimated; see Table A.4). Overall, we find that having the nearest radiation facility ten minutes further from the patient's residence makes her less likely to choose lumpectomy by about 0.7 to 1.1 percentage points (or about 1.2 to 1.9 percent relative to the mean lumpectomy probability of 58 percent). Observable characteristics do not appear to have an important effect on this distance estimate, as can be seen from the fact that the heterogeneity in this estimate does not change much in column (5) relative to the preceding columns.

Taken together, these estimates suggest a reasonably robust relationship between the distance from womens' place of residence at the time of their diagnosis with breast cancer to the nearest radiation treatment facility, and their choice of cancer treatments. In the appendix, we show that these basic results are robust to a variety of alternative specifications, including alternative (mileage-based rather than time-based) measures of distance (Appendix Figures A.1 and A.2, and Appendix Tables A.2 and A.3), non-linear parameterizations of distance (Appendix Table A.5), and estimation with linear probability models (Appendix Table A.6). In the next section, we build on these estimates to use this distance measure as a shifter in the effective relative price women face for breast cancer treatments in order to investigate how breast cancer patients would respond to various (out of sample) policy counterfactuals.

4.3 Policy counterfactuals: estimating treatment choices and ex-post welfare

Following the conceptual framework depicted in Figure 1, our estimated demand curve for lumpectomy allows us to perform quantitative exercises of the impact of alternative insurance designs which vary the price the consumer faces for L. .To see what the exercise is, one can think of each demand specification

as estimating a distribution of the willingness to pay for lumpectomy. To do so, we can use equation (2), which defines patient i's willingness to pay for lumpectomy to be

$$v_i \equiv \frac{\alpha_i}{\beta_i} - \theta_i d_i, \tag{5}$$

and, as before, denote the estimated demand by the distribution $F(v_i)$. A given demand specification provides estimates for the joint distribution of α_i and β_i, and d_i is observed. We assume throughout this section that θ_i is equal to \$1,150 for all patients. To arrive at this estimate for θ_i, we rely on the fact (see Section 2.1) that a typical course of radiation therapy treatment involves 25 round trips to the radiation facility, and that the average hourly wage as reported by the Bureau of Labor Statistics is just over 23 dollars (so $23 \cdot 25 \cdot 2 = \$1,150$).[18] Of course, one could (and should) raise plausible concerns that the opportunity cost of time may be heterogeneous across patients, or that the opportunity cost of time of a breast cancer patient may be higher or lower than that of a healthy working individual. Fortunately, the transformation is sufficiently simple and transparent that one could fairly easily use our results to obtain quantitative estimates that rely on alternative values of θ_i.

Equipped with an estimate of $F(v_i)$ and given an (incremental) price of lumpectomy p defined by the insurance design, the share of patients choosing lumpectomy is given by $1 - F(p)$ and consumer surplus (per patient, relative to everyone being forced to choose mastectomy) is given by $(1 - F(p))E(v_i|v_i > p)$. The total incremental cost is given by $(1 - F(p))c$.

Figure 3 illustrates the nature of this exercise. In Figure 3(a) we plot the implied demand system for lumpectomy using the simplest specification - column 1 of Table 3 - and in Figure 3(b) we plot the implied demand system for lumpectomy using our richest specification - column (6) of Table 3. These figures are the empirical analogs to our conceptual Figure 1, and we indicate the analogous points along them. We also use Figure 3 to illustrate the variation (in distance) used to estimate the demand function by plotting the empirical distribution of the monetized distance (that is, distance in hours multiplied by $\theta = \$1,150$), illustrating the point we emphasized in the introduction: our key counterfactual exercises are quite far out of sample, and therefore should be treated with caution. For this illustrative purpose, we use the information discussed in Section 2.1 to approximate the incremental cost of lumpectomy (c) at \$10,000 and the total cost of lumpectomy at \$50,000, which is the incremental cost together with the baseline cost of \$40,000 for mastectomy.

Figure 3(a) is based on specification (1) of Table 3, which does not include any controls. This speci-

---

[18]Specifically, this Bureau of Labor Statistics figure is for average hourly earnings in October 2012: `http://www.bls.gov/news.release/empsit.t19.htm`.

fication indicates that the US-style "full coverage" policy (given by point D) in which consumers do not pay on the margin for lumpectomy raises the lumpectomy rate by about 37 percentage points relative to the efficient level (given by point E) of about 21 percent; the associated welfare cost is about $2,000 per patient relative to the efficient allocation. Figure 3(b) is instead based on the richest model (column (6) of Table 3), which allows heterogeneity in the response to price and thus make the demand curvature much greater, leading to more elastic demand for small price changes but to a much lower elasticity for large changes. This specification suggests that the US-style "full coverage" policy raises the lumpectomy rate by only about 10 percentage points relative to the efficient level of 48 percent, with a resultant welfare cost of about $710 per patient.

Likewise, the estimates without controls in Figure 3(a) suggest that a UK-style "no top-up" policy (given by point A) in which insurance only covers mastectomies and patients must pay the full cost of a lumpectomy reduces the lumpectomy rate from about 21 percent under the efficient top-up policy to nearly zero, and reduces welfare by about $1,400 per patient, relative to the efficient outcome. When we use the richest model, Figure 3(b), however, a large fraction of the patients are estimated to be not very price sensitive, so lumpectomy rates fall by only about 4.5 percentage points from the efficient level of 48 percent, with a welfare cost of about $800 per patient relative to the efficient level.

For completeness, Table 4 reports additional estimates for a variety of counterfactuals for each of the six demand specifications reported in Table 3. The top panel reports the observed outcome (for $p = 0$), which corresponds to our US-style full coverage benchmark. Each of the subsequent panels report a set of estimates for a different (counterfactual) price for lumpectomy (retaining a zero price for mastectomy). The first row of each panel reports the demand response; that is, the reduction in lumpectomy share (relative to the observed level) from the increase in price. The second row of each panel reports the (per patient) reduction in consumer surplus (relative to the observed level). The change in surplus consists of two components: a set of "marginal" patients change their choice to mastectomy as a result of the price change, and their change in consumer surplus is given by integrating under this portion of the demand curve; and a set of "inframarginal" patients, who have high willingness to pay for lumpectomy, do not change their choices, but now face a higher price which reduces their surplus. The third row of each panel uses a cost of $10,000 for the incremental costs of lumpectomy to report the change in insurer profit, which consist of not paying for lumpectomies for the "marginal" patients and (except for the case of mandate) from charging an incremental price for lumpectomies from the "inframarginal" patients. Finally, the fourth row of each panel reports the overall change in welfare by adding up the change in consumer surplus and the change in insurer profits. We chose the counterfactuals to illustrate price changes of $5,000 as a figure that is close to the lower 95% confidence interval of cost differences from Polsky et al. (2003), and is a

smaller change that is less out of sample; \$10,000 as a figure that is close to the midpoint of the cost difference range from Polsky et al. (2003); and \$50,000 as a figure that is close to the "full" (no top-up) cost based on the figures from Polsky et al. (2003).

5 Ex-ante efficiency: A simple calibration

Our analysis thus far has focused on the impact of alternative insurance designs for the (ex-post) efficiency of treatment choice, taking as given the extent of the patient's risk exposure. We would be remiss, however, to analyze the welfare consequences of insurance designs without considering their impact on risk exposure and hence ex-ante utility. We briefly do so here.

The different insurance designs have different implications for ex-ante risk exposure, making the qualitative ranking of ex-ante efficiency between the US policy and the top-up policy a priori ambiguous. This is because the "top up" policy – which produces the (ex-post) efficient treatment decision – leaves the consumer exposed to risk ex ante. To see this, note that under a "top up" policy in which individuals can pay on the margin for L, risk exposure is increasing in v_i (up to c).

To evaluate ex-ante utility of a given individual, the key empirical object will once again be the individual's willingness to pay for lumpectomy relative to mastectomy v_i, which we assume is known to the individual at the beginning of the coverage year. In addition, we assume individuals are expected-utility maximizers with CARA utility $w(x) = -exp(-rx)$ with a (homogeneous) coefficient of absolute risk aversion r, and (homogeneous) annual probability of illness ρ. Assuming mastectomy is fully covered and the lumpectomy price is p, the individual is faced with a risk of losing $\min(p, v_i)$ with probability ρ. She will either choose lumpectomy and face a financial risk of p, or choose mastectomy and incur a monetized risk of v_i, whichever is smaller.

The individual's ex-ante utility is given by π_i, which is the solution to: $w(x - \pi_i) = (1 - \rho)w(x) + \rho w(x - \min(p, v_i))$. For $v_i > 0$[19] this yields:

$$\pi_i = \frac{1}{r}\log[\rho \cdot \exp(r \cdot \min(v_i; p)) + 1 - \rho]. \tag{6}$$

The price p depends on the insurance design. Under the "top up" policy $p = c$, and equation (6) indicates that this leaves the individual exposed to risk ($\pi_i > 0$).

The "first-best" policy – which involves no ex-ante risk exposure and achieves the ex-post efficient treatment choice – would be to offer a continuum of indemnity insurance policies, which pay each patient a lump sum of $\min(v_i, c)$ in the event of illness and then allows individuals to pay the incremental cost of L

---

[19]For $v_i < 0$ the individual will always choose M and be fully insured under any of the insurance arrangements we consider.

out of pocket; in a competitive equilibrium a policy that pays $\min(v_i, c)$ would be priced at $\rho \min(v_i, c)$ and will be bought by individuals with willingness to pay $v_i > 0$, leaving them unexposed to risk. Such a policy would eliminate both ex-ante and ex-post efficiency losses. From an ex-post perspective, patients would obtain the lump sum, but only those patients with $v_i > c$ would choose lumpectomy, thus replicating the treatment efficiency described in the previous section. From an ex-ante perspective, individuals would be fully insured and would not be exposed to any risk, which is the efficient outcome (assuming, as is typical, that the insurance provider is risk neutral).

Given the impracticality of this policy, we consider the ex-ante utility properties of the three policies we have explored.[20] The "full coverage" (US) policy sets $p = 0$; it removes ex-ante risk exposure ($\pi_i = 0$) and maximizes consumer surplus ($-\pi_i$) but, as discussed, produces ex-post socially inefficient treatment choices. The "no top up" UK policy sets $p = TC$, the total cost of L; it therefore not only produces ex-post inefficient treatment choices but also exposes the individual to ex-ante risk. Moreover, since $TC > c$ – by definition the total cost of L is greater than the incremental cost – it is clear from equation (6) that consumer surplus is lower (π_i is higher) under the UK policy than the top-up policy. The higher social welfare ranking of the top-up policy relative to the UK policy is thus preserved when ex-ante utility is considered. The relative social welfare ranking of the US policy and the top-up policy is a priori unclear as is the relative ranking of the US and UK policies; all else equal, the US policy's relative ranking is increasing in risk aversion r.

To assess ex-ante social welfare under the alternative insurance designs, equation (6) makes clear that in addition to the demand curve $F(v_i)$ that we have already estimated, we also need values for the risk of breast cancer (ρ) and the coefficient of absolute risk aversion (r). Table 5 reports the results from an illustrative calibration exercise. We assume a homogeneous annual risk of breast cancer for a 60 year old female of $\rho = 0.48\%$.[21] We calibrate r based on a range of estimates reported in Table 6 of Cohen and Einav (2007). We use the estimates of $F(v_i)$ from the simplest empirical specification reported in column (1) of Table 3.

The first column reports the share of breast cancer patients choosing lumpectomy, which does not depend on the level of risk aversion. The second column reports the cost savings to insurers relative to the

---

[20]The first best seems impractical primarily because typical markets offer discrete rather than continuous coverage and coverage tends to be in the form of payment for treatment options, rather than lump sum cash payments. One might naturally consider therefore the possibility of a (break-even) insurance policy that would cover the social incremental cost of L in the event of illness. Note however that this does not achieve the first best. Any individual with $v_i > c$ would purchase the policy, face no ex-ante risk exposure and make efficient ex-post treatment decisions. However, some individuals with sufficiently high $v_i < c$ would also purchase the insurance to avoid financial risk, and would therefore (inefficiently) choose L ex-post.

[21]In our California cancer registry data in 2000, there were 585 60-year-old females diagnosed with breast cancer. Dividing this number by the total population of 60-year-old females in California as of 1-July 1999 (120,668; `http://www.census.gov/popest/data/state/asrh/1990s/tables/st-99-10.txt`) gives $p = 0.48\%$.

US-style full coverage case. For the no top-up case, these insurer cost savings are \$50,000 for the small share of people who still choose L, and \$10,000 for anyone who previously chose L but now chooses M. For the top-up case these insurer cost savings are \$10,000 times the share of people who chose L under full coverage. The third column reports the loss in consumer surplus relative to full coverage; we measure this by the change in π_i, as defined in equation (6). Note that this loss in consumer surplus is associated with all individuals in the population, not only with those who are subsequently diagnosed with breast cancer, so even a small magnitude of π could be magnified once it affects the entire population of potential breast cancer patients. Finally, the fourth column reports the total social cost relative to full coverage, by adding up the second and third columns.

The results indicate how the (total) efficiency ranking of the top-up policy relative to the US-style full coverage policy depends on risk aversion. For the lowest value of risk aversion we consider, social welfare is higher under the top-up policy, but for higher values of risk aversion it is higher under the US-style full coverage policy. The full-coverage policy always delivers higher total welfare than the UK-style "no top up" policy for our calibrated values. This illustrative analysis suggests that focusing solely on ex-post efficiency analysis could miss an important part of the picture, and that the ex-ante risk exposure generated by top-up policies could be much more costly than the allocative efficiencies these policies may provide. Of course, this calibration exercise is extremely stylized; evaluating this trade-off more systematically would require, among other things, better estimates of r and ρ – as well as potentially heterogeneity in them and selection on them – for our population.

6 Conclusion

We present a simple framework to illustrate the welfare gains from a health insurance policy that allows patients to pay the incremental price for more expensive treatment options. Such a policy efficiently sorts low willingness-to-pay patients to the cheaper treatment option, in contrast with the current status quo in the US where the incentives for such sorting are minimal. At the same time, this policy does not "over price" the more expensive treatment, as is common in the UK and several other high-income countries, which allocates too many patients to the cheaper treatment. Our analysis of the choice between lumpectomy and mastectomy for breast cancer patients provides an empirical illustration as to how such a policy could be evaluated, and what the quantitative welfare gains to such a policy might be.

Most of our analysis (with the exception of Section 5) focused on analyzing the efficiency of sorting patients to treatments, taking the overall level of risk exposure as given. In this analysis, our "top up" policy resembles a form of indemnity insurance, which pays a fixed amount for a given medical "event,"

such as the diagnosis of breast cancer. Cutler and Zeckhauser (2000) describe how health insurance in the US started off as a quasi-indemnity policy – in most cases paying a fixed cash amount per day in the hospital. Ma and McGuire (1997) argue that pure indemnity payments dissolved over time because of the difficulty of verifying – in a contractually feasible way – what the adverse health event was. Thus, even absent political economy barriers, implementing such an indemnity-type policy could be challenging in many – although perhaps not all – contexts.

References

Athas, William F., Meg Adams-Cameron, William C. Hunt, Andrew Amir-Fazli, and Charles R. Key, "Travel Distance to Radiation Therapy and Receipt of Radiotherapy Following Breast-Conserving Surgery," *Journal of the National Cancer Institute*, 2000, *92* (3), 269–271.

Baicker, Katherine, Mark Shephard, and Jonathan Skinner, "Public Financing Of The Medicare Program Will Make Its Uniform Structure Increasingly Costly To Sustain," *Health Affairs*, 2012, *32* (5), 882–890.

Baker, Laurence C., "Managed Care and Technology Adoption in Health Care: Evidence from Magnetic Resonance Imaging," *Journal of Health Economics*, 2001, *20* (3), 395 – 421.

_ and Scott W. Atlas, "Relationship Between HMO Market Share and the Diffusion and Use of Advanced MRI Technologies," *Journal of the American College of Radiology*, 2004, *1* (7), 478 – 487.

_ , _ , and Christopher C. Afendulis, "Expanded Use Of Imaging Technology And The Challenge Of Measuring Value," *Health Affairs*, 2008, *27* (6), 1467–1478.

Celaya, Maria O., Judy R. Rees, Jennifer J. Gibson, Bruce L. Riddle, and E. Robert Greenberg, "Travel Distance and Season of Diagnosis Affect Treatment Choices for Women with Early-Stage Breast Cancer in a Predominantly Rural Population (United States)," *Cancer Causes Control*, 2006, *17*, 851–856.

Chalkidou, Kalipson and Gerard Anderson, "Comparative Effectiveness Research: International Experiences and Implications for the United States," `http://www.academyhealth.org/files/publications/CER_International_Experience_09%20%283%29.pdf` 2009. Academy Health.

Chandra, Amitabh, Anupam Jena, and Jonathan Skinner, "A Pragmatist's Guide to Comparative Effectiveness Research," *Journal of Economic Perspectives*, 2011, *25* (2), 27–46.

Cohen, Alma and Liran Einav, "Estimating Risk Preferences From Deductible Choice," *American Economic Review*, 2007, *97* (3), 745–788.

Cutler, David and Richard Zeckhauser, "Handbook of Health Economics, Volume 1," in Anthony Culyer and Joseph Newhouse, eds., *Anthony Culyer and Joseph Newhouse, eds.*, Elsevier, 2000, chapter The Anatomy of Health Insurance, pp. 564–643.

Cutler, David M. and Jonathan Gruber, "Does Public Insurance Crowd Out Private Insurance?," *Quarterly Journal of Economics*, 1996, *111* (2), 391–430.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf, "Optimal Mandates and The Welfare Cost of Asymmetric Information: Evidence from The U.K. Annuity Market," *Econometrica*, 2010, *78* (3), 1031–1092.

Enthoven, Alain and Richard Kronick, "A Consumer-Choice Health Plan for the 1990s," *New England Journal of Medicine*, 1989, *320* (1), 29–37.

Feldstein, Martin, "The Economics of Health and Health Care: What Have We Learned? What Have I Learned?," *American Economic Review Papers and Proceedings*, 1995, *85* (2), 28–31.

Fisher, Bernard, Carol Redmond, Roger Poisson, Richard Margolese, Norman Wolmark, Lawrence Wickerham, Edwin Fisher, Melvin Deutsch, Richard Caplan, Yosef Pilch, Andrew Glass, Henry Shibata, Harvey Lerner, Jose Terz, and Lynne Sidorovich, "Eight-Year

Results of a Randomized Clinical Trial Comparing Total Mastectomy and Lumpectomy with or without Irradiation in the Treatment of Breast Cancer," *New England Journal of Medicine*, 1989, *320* (13), 822–828.

＿ , **Madeline Bauer, Richard Margolese, Roger Poisson, Yosef Pilch, Carol Redmond, Edwin Fisher, Norman Wolmark, Melvin Deutsch, Eleanor Montague, Elizabeth Saffer, Lawrence Wickerham, Harvey Lerner, Andrew Glass, Henry Shibata, Peter Deckers, Alfred Ketcham, Robert Oishi, and Ian Russell**, "Five-Year Results of a Randomized Clinical Trial Comparing Total Mastectomy and Segmental Mastectomy with or without Radiation in the Treatment of Breast Cancer," *New England Journal of Medicine*, 1985, *312* (11), 665–673.

＿ , **Stewart Anderson, Carol K. Redmond, Norman Wolmark, D. Lawrence Wickerham, and Walter M. Cronin**, "Reanalysis and Results after 12 Years of Follow-up in a Randomized Clinical Trial Comparing Total Mastectomy with Lumpectomy with or without Irradiation in the Treatment of Breast Cancer," *New England Journal of Medicine*, 1995, *333* (22), 1456–1461.

＿ , ＿ , **John Bryant, Richard G. Margolese, Melvin Deutsch, Edwin R. Fisher, Jong-Hyeon Jeong, and Norman Wolmark**, "Twenty-Year Follow-up of a Randomized Trial Comparing Total Mastectomy, Lumpectomy, and Lumpectomy plus Irradiation for the Treatment of Invasive Breast Cancer," *New England Journal of Medicine*, 2002, *347* (16), 1233–1241.

Gloeckler Ries, Lynn A. and Milton P. Eisner, "Cancer Survival Among Adults: U.S. SEER Program, 1988-2001, Patient and Tumor Characteristics," in Lynn A. Gloeckler Ries, John L. Young, Gretchen Keel, Milton P. Eisner, Yi Dan Lin, and Mari-Jospehe D. Horner, eds., *Lynn A. Gloeckler Ries, John L. Young, Gretchen Keel, Milton P. Eisner, Yi Dan Lin, and Mari-Jospehe D. Horner, eds.*, National Cancer Institute, SEER Program, NIH Pub. No. 07-6215, Bethesda, MD, 2007, chapter Chapter 13: Cancer of the Female Breast, pp. 101–110.

Ma, Albert and Thomas McGuire, "Optimal Health Insurance and Provider Payment," *American Economic Review*, 1997, *87* (4), 685–704.

McCabe, Christopher, Karl Claxton, and Anthony Culyer, "The NICE Cost-Effectiveness Threshold: What It Is and What That Means," *Pharmacoeconomics*, 2008, *26* (9), 733–744.

McClellan, Mark, Barbara J. McNeil, and Joseph P. Newhouse, "Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables," *Journal of the American Medical Association*, 1994, *272* (11), 859–859.

Nattinger, Ann Butler, Ronald T. Kneusal, Raymond G. Hoffmann, and Mary Ann Gilligan, "Relationship of Distance from a Radiotherapy Facility and Initial Breast Cancer Treatment," *Journal of the National Cancer Institute*, 2001, *93* (17), 1344–1346.

Peltzman, Sam, "The Effect of Government Subsidies-in-Kind on Private Expenditures: The Case of Higher Education," *Journal of Political Economy*, 1973, *81* (1), 1–27.

Pollack, Andrew, "Cost of Treatment May Influence Doctors," *New York Times*, 2014, *17 April.*

Polsky, Daniel, Jeanne S. Mandelblatt, Jane C. Weeks, Laura Venditti, Yi-Ting Hwang, Henry A. Glick, Jack Hadley, and Kevin A. Schulman, "Economic Evaluation of Breast Cancer Treatment: Considering the Value of Patient Choice," *Journal of Clinical Oncology*, 2003, *21* (3), 1139–1146.

Robinson, James C. and Kimberly MacPherson, "Payers Test Reference Pricing and Centers of Excellence to Steer Patients to Low-Price and High-Quality Providers," *Health Affairs*, 2012, *31* (9), 2028–2036.

Roetzheim, Richard G., Thomas N. Chirikos, Kristen J. Wells, Ellen P. McCarthy, Long H. Ngo, Donglin Li, Reed E. Drews, and Lisa I. Iezzoni, "Managed Care and Cancer Outcomes for Medicare Beneficiaries With Disabilities," *American Journal of Managed Care*, 2008, *14* (5), 287–296.

Schroen, Anneke T., David R. Brenin, Maria D. Kelly, William A. Knaus, and Craig L. Slingluff Jr, "Impact of Patient Distance to Radiation Therapy on Mastectomy Use in Early-Stage Breast Cancer Patients," *Journal of Clinical Oncology*, 2005, *23* (28), 7074–7080.

Figure 1: **Conceptual framework: Treatment choice**
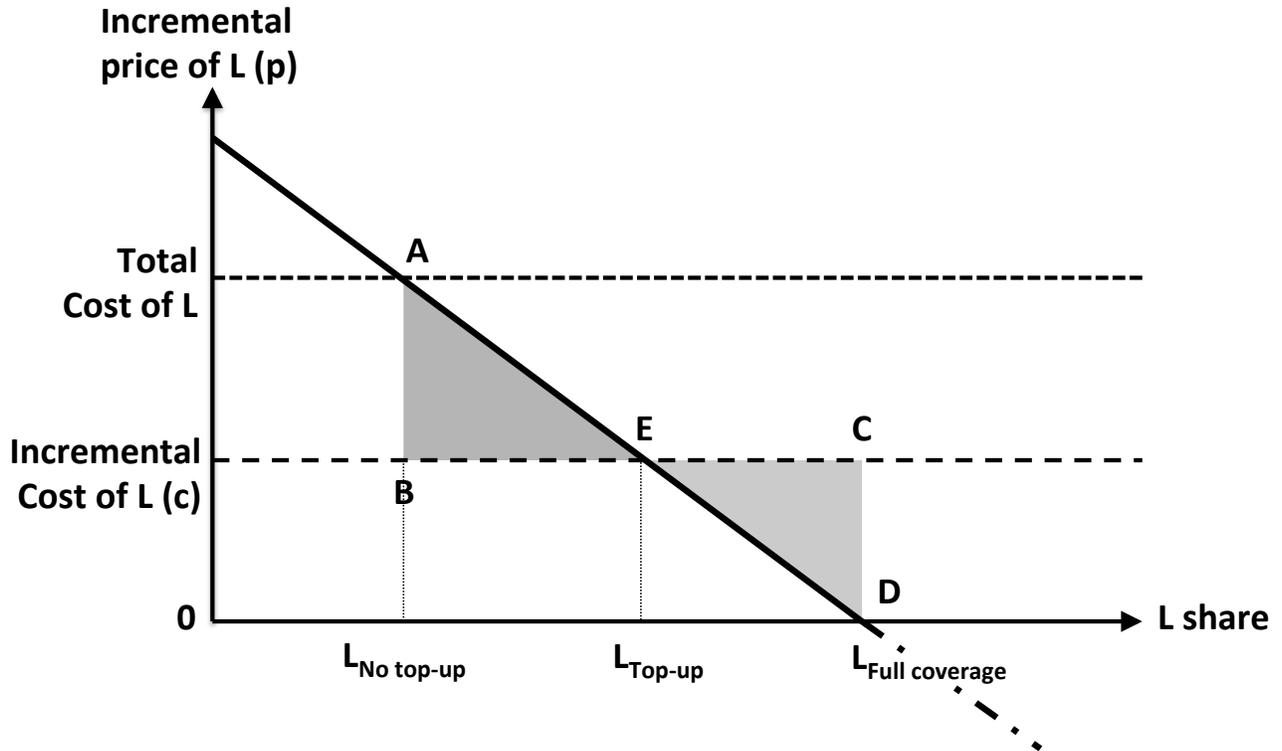
Notes: This figure illustrates conceptually the efficiency consequences of alternative insurance designs (i.e. prices for lumpectomy (L) relative to mastectomy (M)). The efficient allocation is given by point E and the "top-up" insurance design under which patients pay the incremental cost of L relative to M, and fraction $L_{\text{top-up}}$ choose L. Equilibrium under a US-style "full coverage" insurance design in which individuals do not pay on the margin for L relative to M is given by point D, where fraction $L_{\text{full coverage}}$ choose L. The welfare loss from this outcome relative to the efficient outcome is given by triangle CDE. Equilibrium under a UK-style "no top-up" alternative insurance design in which only M is covered by insurance and patients must pay the total cost for L is given by point A, where $L_{\text{no top-up}}$ choose L and the welfare loss relative to the efficient outcome is given by triangle ABE.

Figure 2: **Treatment choice by travel time to nearest radiation facility**

(a) Raw data

(b) Residualized data
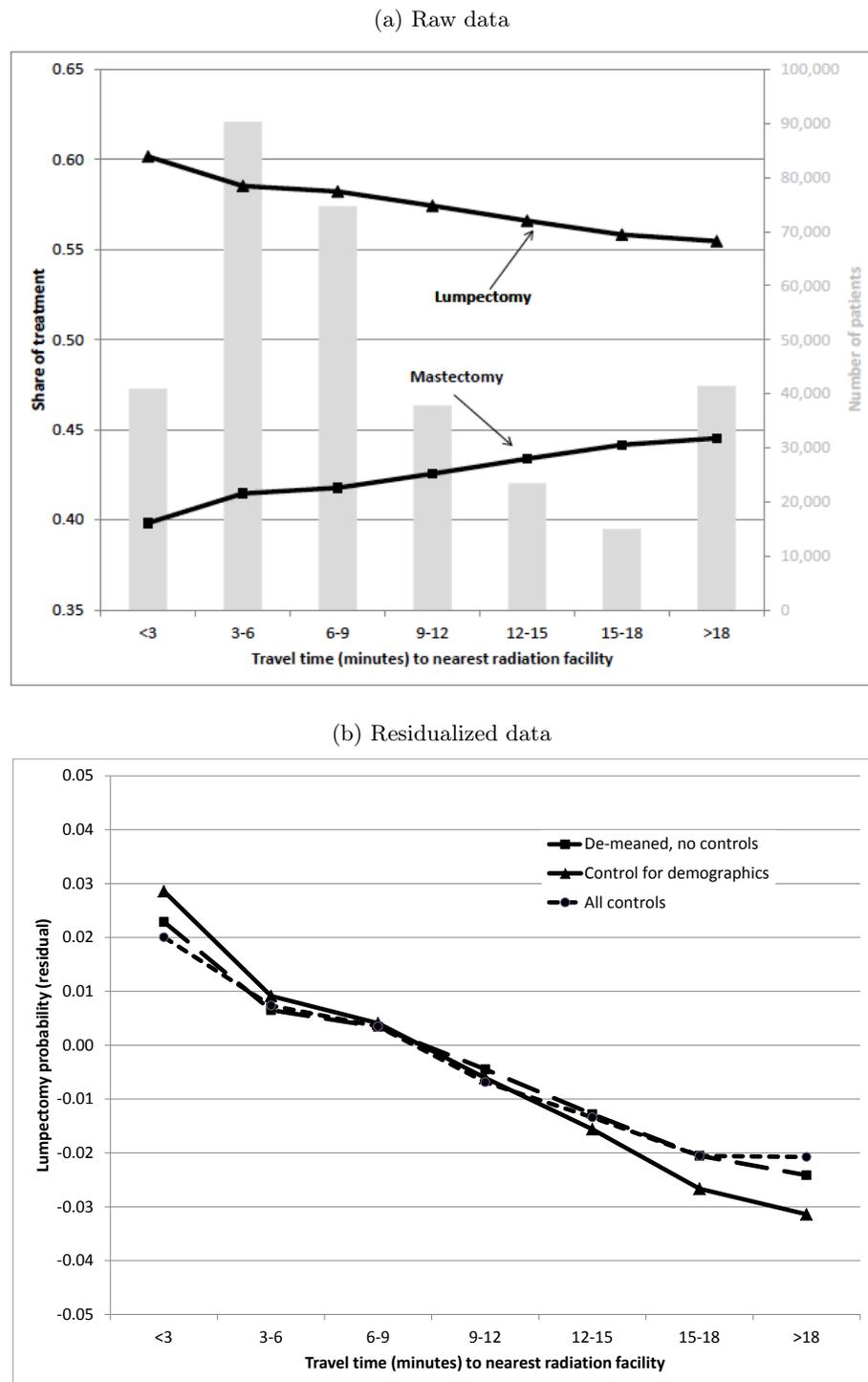
Notes: These figures plot the probability of treatment by travel time in minutes to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). Panel (a) plots the raw data, and a histogram of the number of patients by travel time. Panel (b) plots the de-meaned lumpectomy rate as well as two residualized versions. The first residualizes lumpectomy probability using patient characteristics. The second adds neighborhood-level covariates from the 2000 Census and clinical covariates. All covariates are as described in the notes to Table 1.

Figure 3: **Empirical analog of conceptual framework: Treatment choice**

(a) Homogeneous logit
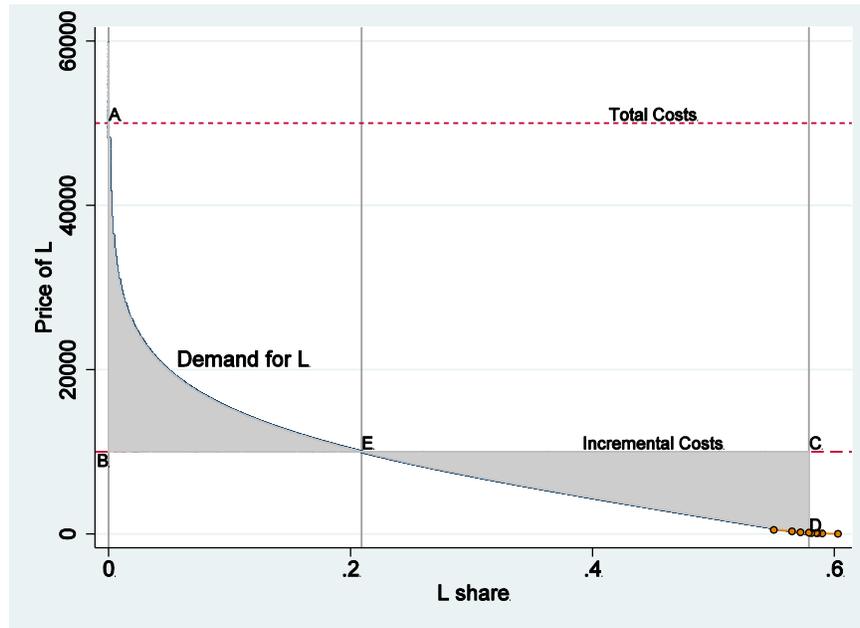
(b) Random coefficient logit
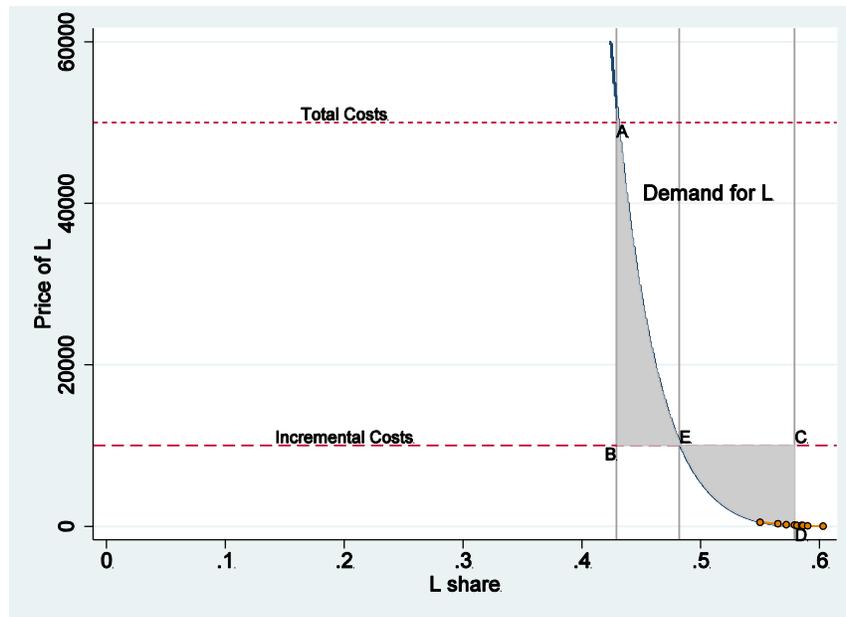
Notes: These figures represent the empirical analog of Figure 1. The curves plot the implied demand system for lumpectomy based on: in Panel (a), the estimates in column (1) of Table 3; and in Panel (b), the estimates in column (6) of Table 3. The scatterplots in the lower right-hand corners of the graphs illustrate the variation (in travel time) used to estimate the demand function by plotting the empirical distribution of the monetized distance (that is, distance in hours multiplied by $\theta = \$1,150$), using the 7 distance "bins" shown in Figure 2; this illustrates the point we emphasize in the text: our key counterfactual exercises are quite far out of sample, and therefore should be taken with caution.

Table 1: **Summary statistics**

	Mean	Std. Dev.	5th Pct.	95th Pct.
Treatment choice:				
Lumpectomy	0.579			
Mastectomy	0.421			
Distance to nearest radiation:				
Driving time (minutes)	10.9	14.3	2	32
Driving distance (miles)	7.9	11.6	1.0	24.9
Spherical distance (miles)	5.3	7.8	0.7	17.6
Demographics (at time of diagnosis):				
Year of diagnosis	2003.2	3.7	1997	2009
Age	60.8	13.8	40	83
Married	0.570			
Medicaid coverage	0.156			
Race: Non-Hispanic White	0.692			
Race: Hispanic	0.141			
Race: Asian or Pacific	0.101			
Race: Other	0.066			
Census-block characteristics (using 2000 census):				
Share below twice poverty line	0.253	0.187	0.041	0.636
Median annual income	58,908	28,525	24,063	110,595
Share with blue collar job	0.108	0.060	0.020	0.213
Share unemployed	0.037	0.029	0.001	0.090
Median monthly rental rate	955.2	404.6	466	1,814
Median home value	266,417	188,861	76,300	660,300
Average years of schooling	13.9	1.3	11.3	15.5
Clinical characteristics (at time of diagnosis):				
Cancer Stage:				
In Situ	0.182			
Local	0.533			
Regional spread	0.264			
Remote spread	0.017			
Tumor grade:				
1	0.190			
2	0.379			
3	0.278			
4	0.047			

Notes: This table displays the mean and - for non-binary variables - the standard deviation, 5th percentile, and 95th percentile in our baseline sample (N=323,612). Distance is one-way distance. Our demographic covariates are year of diagnosis (here continuous; indicator variables in our analyses, with 1997 as the omitted year); age (here continuous; indicator variables in our analyses for below 40, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, and above 80, with below 40 as the omitted group); marital status indicators (married; other/missing is omitted in our analyses); a Medicaid coverage indicator (=1 if Medicaid covered part of the primary source of payment to the hospital); and race/ethnicity indicators (non-Hispanic White, Hispanic, and Asian or Pacific; other/missing is omitted in our analyses). Our census block characteristics are drawn from the 2000 census: percent of population within the census block group that is at or below 200% of the poverty line; median household income within the census block group; proportion of those 16 years of age or over within the census block group with a blue collar job; proportion of those 16 years of age or over within the census block group in the labor force that are unemployed; median gross rent of renter occupied houses within the census block group; median gross home value of owner occupied houses within the census block group; and average years of schooling in the census tract. Our clinical characteristics are two measures of the extent of disease at the time of diagnosis: indicators for cancer stage (in situ, localized, regional, and remote; missing stage is omitted in our analyses), and indicators for cancer grade (1, 2, 3, and 4; missing grade is omitted in our analyses).

Table 2: **Summary statistics by above and below median travel time**

| | Mean | | P-val |
	Above median distance	Below median distance	
Treatment choice:			
Lumpectomy	0.569	0.588	0.000
Mastectomy	0.431	0.412	0.000
Distance to nearest radiation:			
Driving time (minutes)	17.4	4.4	0.000
Driving distance (miles)	13.2	2.6	0.000
Spherical distance (miles)	8.8	1.9	0.000
Demographics (at time of diagnosis):			
Year of diagnosis	2003.2	2003.1	0.000
Age	60.5	61.1	0.000
Married	0.608	0.533	0.000
Medicaid coverage	0.150	0.163	0.000
Race: Non-Hispanic White	0.724	0.660	0.000
Race: Hispanic	0.128	0.155	0.000
Race: Asian or Pacific	0.088	0.114	0.000
Race: Other	0.060	0.071	0.000
Census-block characteristics (using 2000 census):			
Share below twice poverty line	24.004	26.613	0.000
Median annual income	60,245	57,572	0.000
Share with blue collar job	0.111	0.104	0.000
Share unemployed	0.037	0.037	0.007
Median monthly rental rate	952.6	957.9	0.000
Median home value	253,002	279,832	0.000
Average years of schooling	14.0	13.9	0.000
Clinical characteristics (at time of diagnosis):			
Cancer Stage:			
In Situ	0.183	0.182	0.386
Local	0.533	0.533	0.981
Regional spread	0.263	0.265	0.388
Remote spread	0.017	0.017	0.573
Tumor grade:			
1	0.195	0.186	0.000
2	0.376	0.382	0.000
3	0.273	0.284	0.000
4	0.048	0.046	0.068

Notes: This table splits our baseline sample (N=323,612) by above and below median distance from a radiation facility, presents the mean for each sub-sample, and presents the p-value from a test for a difference between these means (with unadjusted standard errors). All covariates are as described in the notes to Table 1.

Table 3: **Treatment choice by travel time**

	Homogeneous Logit				Heterogenous Logit	
	No covariates	Column (1) + demographics	Column (2) + Census block characteristics	Column (3) + clinical characteristics (all controls)	All controls, also interacted with distance	Column (4) + random coefficients
	(1)	(2)	(3)	(4)	(5)	(6)
Average effect of a 10-minute increase in travel time (standard error)	-0.0077 (0.00056)	-0.0100 (0.00056)	-0.00704 (0.00049)	-0.00729 (0.00056)	-0.0108 (0.00064)	-0.0254 (0.00263)
Standard deviation (across patients) of the above effect	0.00004	0.00006	0.00003	0.00004	0.00007	0.02442

Notes: This table summarizes estimates from logit regressions estimating the relationship between breast cancer treatment choice of lumpectomy (mean = 0.58) and (one-way) travel time to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). The outcome variable is an indicator variable, equal to 1 if the patient receives a lumpectomy. This table reports the estimated effect with respect to a ten-minute increase in one-way travel time by computing - for each patient in our sample - the predicted probability of a lumpectomy choice (given the estimated coefficients) and the way it would change if we made her drive 10 minutes longer; we then report the average marginal effect in the sample and its standard error (clustered at the county level, and based on 50 replications of the bootstrap), as well as the standard deviation of this effect across patients. Column (1) reports estimates from a logit model with no controls. Columns (2), (3), and (4) add covariates for demographics, Census block controls, and clinical controls successively. Column (5) adds interactions between these covariates and our distance measures, and Column (6) allows random coefficients on distance. All covariates are as described in the notes to Table 1.

Table 4: **Counterfactual analysis**

| | Homogeneous Logit | | | | Heterogenous Logit | |
| | No covariates | Column (1) + demographics | Column (2) + Census block characteristics | Column (3) + clinical characteristics (all controls) | All controls, also interacted with distance | Column (4) + random coefficients |
	(1)	(2)	(3)	(4)	(5)[a]	(6)[b]
Observed choices ("Full coverage" benchmark)						
Lumpectomy share	0.579	0.579	0.579	0.579	0.579	0.579
Consumer surplus	5,255	4,066	5,796	5,899	9,559[c]	6.84E+13
Charge $5,000 for lumpectomy (= 4.35 hours)						
Change in lumpectomy share	-0.203	-0.257	-0.184	-0.177	-0.254	-0.076
Change in consumer surplus	-2,385	-2,242	-2,434	-2,453	-2,125	-2,623
Change in insurer profits	3,910	4,180	3,815	3,780	4,163	3,275
Overall change in welfare	1,525	1,938	1,381	1,327	2,038	652
Charge $10,000 for lumpectomy (= 8.7 hours). ("Top up" benchmark)						
Change in lumpectomy share	-0.369	-0.438	-0.342	-0.335	-0.398	-0.096
Change in consumer surplus	-3,826	-3,357	-3,996	-4,053	-3,234	-5,082
Change in insurer profits	5,790	5,790	5,790	5,790	5,790	5,790
Overall change in welfare	1,964	2,433	1,794	1,737	2,556	708
Charge $50,000 for lumpectomy (= 34.8 hours). ("No-top up" benchmark)						
Change in lumpectomy share	-0.579	-0.579	-0.578	-0.578	-0.579	-0.149
Change in consumer surplus	-5,253	-4,066	-5,791	-5,894	-9,559[c]	-23,099
Change in insurer profits	5,790	5,790	5,790	5,790	5,790	22,990
Overall change in welfare	537	1,724	39	-64	-3,769	-109
Mandate mastectomy						
Change in lumpectomy share	-0.579	-0.579	-0.579	-0.579	-0.579	-0.579
Change in consumer surplus	-5,255	-4,066	-5,796	-5,899	-9,559[c]	-6.84E+13
Change in insurer profits	5,790	5,790	5,790	5,790	5,790	5,790
Overall change in welfare	535	1,724	-6	-109	-3,769	-6.84E+13

Notes: All reported changes (in lumpectomy share, consumer surplus, and total cost saving) are relative to the observed choices, for our baseline sample (N=323,612). Counterfactuals involve various potential charges for lumpectomy (ranging from $5,000 to infinity). For each counterfactual we report in parentheses the equivalent addition to one-way drive time (assuming 25 round trips valued at $23 per hour, as described in the text). The columns are as described in the notes to Table 3. All covariates are as described in the notes to Table 1.

(a): in the specification of column (5), the point estimate of β_i is negative for 4.3% of individuals. The reported results assume that $\beta_i = 0$ for these individuals.

(b): due to the large estimated heterogeneity of β_i in the specification of column (6), the welfare estimates for the "mandate mastectomy" case (bottom set of rows) are extremely large and should be interpreted with a great deal of caution. This does not affect the welfare estimates for finite price changes, in which the individuals with low β_i's are infra-marginal and thus do not affect the estimates.

(c): the welfare estimates for actual choices (top set of rows) and for the "mandate mastectomy" case (bottom set of rows) are computed for the 95.7% of the individuals whose β_i's are strictly positive.

Table 5: **Ex-ante efficiency: Simple calibration**

	Share of patients who choose lumpectomy	Insurer cost difference relative to full coverage	Reduction in consumer surplus relative to full coverage (π_i)	Social cost relative to full coverage
	(1)	(2)	(3)	(4)
Absolute risk aversion = 0.000026 (median estimate in Cohen and Einav, 2007)				
Full coverage	0.535	0	0	0
No top up	0.002	-5,443	11,308	5,865
Top up allowed	0.350	-5,355	4,978	-377
Absolute risk aversion = 0.00029 (75th percentile estimate in Cohen and Einav, 2007)				
Full coverage	0.535	0	0	0
No top up	0.002	-5,443	85,529	80,085
Top up allowed	0.350	-5,355	17,746	12,391
Absolute risk aversion = 0.0027 (90th percentile estimate in Cohen and Einav, 2007)				
Full coverage	0.535	0	0	0
No top up	0.002	-5,443	162,477	157,034
Top up allowed	0.350	-5,355	80,019	74,664

Notes: We analyze three alternative insurance designs for three different levels of risk aversion. The "Full coverage" (US-style policy) sets $p = 0$; the "No top up" (UK-style) policy sets $p = TC$; the "Top up allowed" policy sets $p = c$. The first column reports the share of breast cancer patients choosing lumpectomy. The second column reports the insurer cost savings relative to full coverage. The third column reports the change (relative to full coverage) in average risk premium π_i, as defined in equation (6). The fourth column reports the total social cost relative to full coverage, by adding up the second and third columns.

A   Appendix: Additional figures and tables

Figure A.1: **Treatment choice by spherical distance to nearest radiation facility**

(a) Raw data

(b) Residualized data
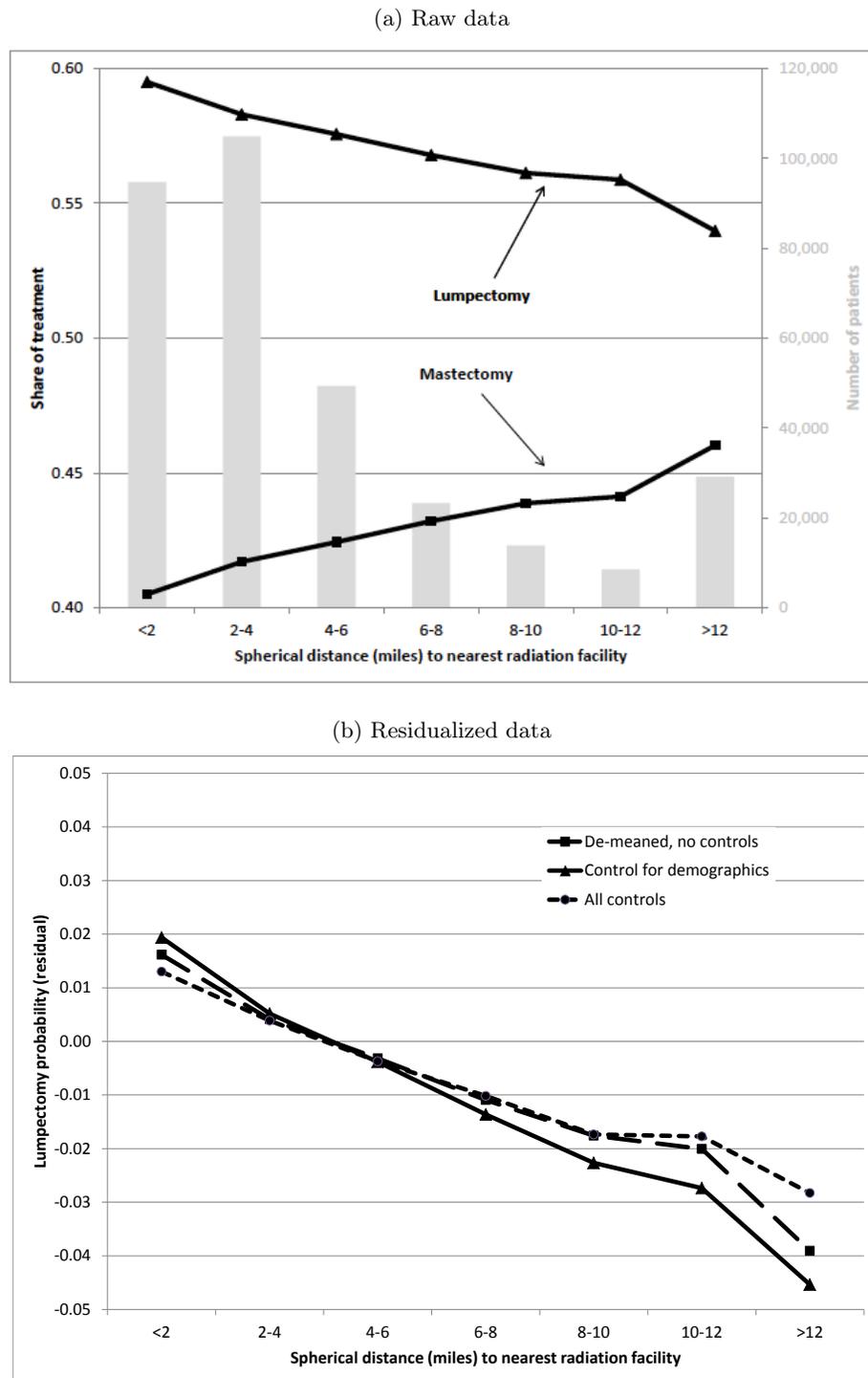
Notes: These figures plot the probability of treatment by spherical distance in miles to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). Panel (a) plots the raw data, and a histogram of the number of patients by spherical distance. Panel (b) plots the de-meaned lumpectomy rate as well as two residualized versions. The first residualizes lumpectomy probability using patient characteristics. The second adds neighborhood-level covariates from the 2000 Census and clinical covariates. All covariates are as described in the notes to Table 1.

Figure A.2: **Treatment choice by travel distance to nearest radiation facility**

(a) Raw data

(b) Residualized data
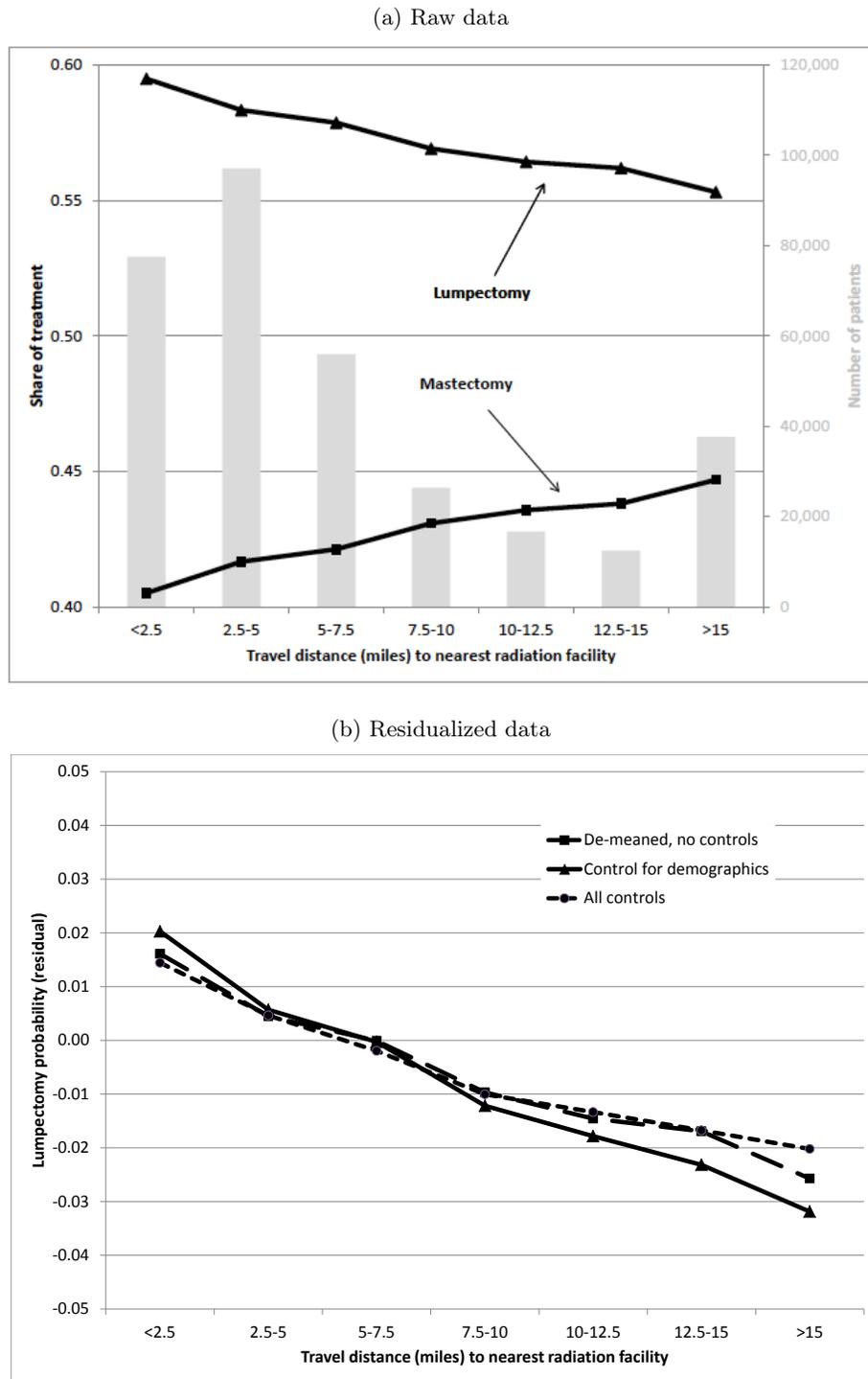
Notes: These figures plot the probability of treatment by travel distance in miles to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). Panel (a) plots the raw data, and a histogram of the number of patients by spherical distance. Panel (b) plots the de-meaned lumpectomy rate as well as two residualized versions. The first residualizes lumpectomy probability using patient characteristics. The second adds neighborhood-level covariates from the 2000 Census and clinical covariates. All covariates are as described in the notes to Table 1.

Table A.1: **Robustness: Treatment choice by travel time, full sample**

	Homogeneous Logit				Heterogenous Logit	
	No covariates	Column (1) + demographics	Column (2) + Census block characteristics	Column (3) + clinical characteristics (all controls)	All controls, also interacted with distance	Column (4) + random coefficients
	(1)	(2)	(3)	(4)	(5)	(6)
Average effect of a 10-minute increase in travel time (standard error)	-0.0068 (0.00057)	-0.0096 (0.00054)	-0.0067 (0.00063)	-0.0060 (0.00056)	-0.0088 (0.00056)	-0.0265 (0.00262)
Standard deviation (across patients) of the above effect	0.00002	0.00018	0.00018	0.00009	0.00598	0.04869

Notes: This table summarizes estimates from logit regressions estimating the relationship between breast cancer treatment choice of lumpectomy (mean = 0.58) and travel time to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our full sample (that is, not limited to women receiving lumpectomy or mastectomy; N=343,399). The outcome variable is an indicator variable, equal to 1 if the patient receives a lumpectomy. This table reports the estimated effect with respect to a ten-minute increase in travel time by computing - for each patient in our sample - the predicted probability of a lumpectomy choice (given the estimated coefficients) and the way it would change if we made her drive 10 minutes longer; we then report the average marginal effect in the sample and its standard error (clustered at the county level, and based on 50 replications of the bootstrap), as well as the standard deviation of this effect across patients. Column (1) reports estimates from a logit model with no controls. Columns (2), (3), and (4) add covariates for demographics, Census block controls, and clinical controls successively. Column (5) adds interactions between these covariates and our distance measures, and Column (6) allows random coefficients on distance. All covariates are as described in the notes to Table 1.

Table A.2: Robustness: Treatment choice by spherical distance

	Homogeneous Logit				Heterogenous Logit	
	No covariates	Column (1) + demographics	Column (2) + Census block characteristics	Column (3) + clinical characteristics (all controls)	All controls, also interacted with distance	Column (4) + random coefficients
	(1)	(2)	(3)	(4)	(5)	(6)
Average effect of a 10-mile increase in spherical distance (standard error)	-0.0177 (0.00116)	-0.0211 (0.00097)	-0.0146 (0.00097)	-0.0151 (0.00101)	-0.0231 (0.00120)	-0.0732 (0.00840)
Standard deviation (across patients) of the above effect	0.00010	0.00052	0.00052	0.00124	0.01470	0.26626

Notes: This table summarizes estimates from logit regressions estimating the relationship between breast cancer treatment choice of lumpectomy (mean = 0.58) and spherical distance to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). The outcome variable is an indicator variable, equal to 1 if the patient receives a lumpectomy. This table reports the estimated effect with respect to a ten mile increase in spherical distance by computing - for each patient in our sample - the predicted probability of a lumpectomy choice (given the estimated coefficients) and the way it would change if we made her travel ten spherical miles longer; we then report the average marginal effect in the sample and its standard error (clustered at the county level, and based on 50 replications of the bootstrap), as well as the standard deviation of this effect across patients. Column (1) reports estimates from a logit model with no controls. Columns (2), (3), and (4) add covariates for demographics, Census block controls, and clinical controls successively. Column (5) adds interactions between these covariates and our distance measures, and Column (6) allows random coefficients on distance. All covariates are as described in the notes to Table 1.

Table A.3: **Robustness: Treatment choice by travel distance**

	Homogeneous Logit				Heterogenous Logit	
	No covariates	Column (1) + demographics	Column (2) + Census block characteristics	Column (3) + clinical characteristics (all controls)	All controls, also interacted with distance	Column (4) + random coefficients
	(1)	(2)	(3)	(4)	(5)	(6)
Average effect of a 10-mile increase in travel distance (standard error)	-0.0091 (0.00072)	-0.0116 (0.00069)	-0.0081 (0.00064)	-0.0084 (0.00066)	-0.0112 (0.00081)	-0.0691 (0.00510)
Standard deviation (across patients) of the above effect	0.00004	0.00007	0.00003	0.00004	0.00006	0.24492

Notes: This table summarizes estimates from logit regressions estimating the relationship between breast cancer treatment choice of lumpectomy (mean = 0.58) and travel distance to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). The outcome variable is an indicator variable, equal to 1 if the patient receives a lumpectomy. This table reports the estimated effect with respect to a ten mile increase in travel distance by computing - for each patient in our sample - the predicted probability of a lumpectomy choice (given the estimated coefficients) and the way it would change if we made her travel ten miles longer; we then report the average marginal effect in the sample and its standard error (clustered at the county level, and based on 50 replications of the bootstrap), as well as the standard deviation of this effect across patients. Column (1) reports estimates from a logit model with no controls. Columns (2), (3), and (4) add covariates for demographics, Census block controls, and clinical controls successively. Column (5) adds interactions between these covariates and our distance measures, and Column (6) allows random coefficients on distance. All covariates are as described in the notes to Table 1.

Table A.4: **Parameter estimates and standard errors underlying Table 3**

		Homogeneous Logit			Heterogenous Logit	
	No covariates	Column (1) + demographics	Column (2) + Census block characteristics	Column (3) + clinical characteristics (all controls)	All controls, also interacted with distance	Column (4) + random coefficients
	(1)	(2)	(3)	(4)	(5)	(6)
Travel time (minutes) (standard error)	-0.0032 (0.00102)	-0.0041 (0.00124)	-0.0029 (0.00112)	-0.0030 (0.00129)	-0.0045 (0.00197)	
Mean of (random) coefficient on travel time (standard error)						-8,910 (51,629)
Std. deviation of (random) coefficient on travel time (minutes) (standard error)						3.5E+12 (4.78E+13)

Notes: This table summarizes the parameter estimates and their standard errors that give rise to the average marginal effects reported in Table 3. All details are as reported in the notes to Table 3.

Table A.5: **Robustness: Treatment choice by non-linear travel time**

	Homogeneous Logit			
	No covariates	Column (1) + demographics	Column (2) + Census block characteristics	Column (3) + clinical characteristics (all controls)
	(1)	(2)	(3)	(4)
Travel time <= 5 minutes	0.153 (0.0439)	0.203 (0.0503)	0.150 (0.0423)	0.165 (0.0475)
Travel time is 5-10 minutes	0.104 (0.0373)	0.140 (0.0426)	0.0998 (0.0377)	0.111 (0.0424)
Travel time is 10-15 minutes	0.0574 (0.0329)	0.0806 (0.0354)	0.0445 (0.0291)	0.0480 (0.0302)
Travel time >= 15 minutes	------------- Omitted category -------------			

Notes: This table summarizes estimates from logit regressions estimating the relationship between breast cancer treatment choice of lumpectomy (mean = 0.58) and travel time to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). Unlike in Table 3, we here parametrize travel time non-linearly: <=5 minutes, 5 to 10 minutes, 10 to 15 minutes, or more than 15 minutes (omitted). All other details are as reported in the notes to Table 3.

Table A.6: **Robustness: Treatment choice by travel time, linear probability models**

	No covariates	Column (1) + demographics	Column (2) + Census block characteristics	Column (3) + clinical characteristics (all controls)	All controls, also interacted with distance
	(1)	(2)	(3)	(4)	(5)
Trvael time (in 10-minute units)	-0.0078	-0.0100	-0.0070	-0.0068	-0.0101
(standard error)	(0.00252)	(0.00303)	(0.00275)	(0.00293)	(0.00440)

Notes: This table summarizes estimates from linear probability models estimating the relationship between breast cancer treatment choice of lumpectomy (mean = 0.58) and travel time to the nearest radiation facility, measured from the patients' address of residence at the time of cancer diagnosis, for our baseline sample (N=323,612). The outcome variable is an indicator variable, equal to 1 if the patient receives a lumpectomy. Column (1) reports estimates from a model with no controls. Columns (2), (3), and (4) add covariates for demographics, Census block controls, and clinical controls successively. Column (5) adds interactions between these covariates and our distance measures. Standard errors are clustered by county. All covariates are as described in the notes to Table 1.