Disparities In The Intensity Of Breast Cancer Treatment

Emma Bojinova

Abstract

Disparities in the survival of breast cancer patients have been observed between African-American and Caucasian women in the United States. The reasons for this differential are still unclear – part of it can be attributed to differences in biology and genetics, to social, economic, and cultural factors, but also it can be due to racial discrimination. The goal of this paper is to investigate if there are racial differences in the intensity of treatment of breast cancer patients. Based on cross-sectional data from the Healthcare Cost and Utilization Project (HCUP), various specifications are used to estimate if African-Americans and other racial groups are treated less intensively in comparison with white Americans when they are admitted to hospitals in 35 states. This study provides some evidence for the presence of such disparities.

I. Introduction

Cancer has been the second leading cause of death in the United States for a number of years. Breast cancer has the highest mortality rate among females diagnosed with cancer. According to the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute the age-adjusted breast cancer incidence rate for the period 2002-2006 was 123.8 per 100,000 women per year. The breakdown by race is as follows: the incidence rate is highest for white females (127.8 per 100,000 women), followed by black females (117.7 per 100,000), Asian/Pacific Islanders (89.5 per 100,000), and Hispanics (88.3 per 100,000 women). The lowest incidence rate is recorded for American Indian/Alaska Native females (74.4 per 100,000). In contrast, black females have the highest age-adjusted mortality rate (33 per 100,000) as compared to other races for the same period. The breast cancer mortality rate is approximately 38.1 percent lower for white women (23.9 per 100,000) and between 87.5 percent and 264 percent lower for the other three racial groups mentioned above. Expenditures on diagnosis and treatment of breast cancer cases are very high. The National Cancer Institute estimates that approximately $13.9 billion is spent on breast cancer treatment every year in the United States (Cancer Trends Progress Report – 2011/2012 update).

The early detection of breast cancer is very important for a patient's survival. Mammography is especially useful for identifying breast cancer at an early stage even before physical symptoms develop. Early detection increases treatment options and thus decreases mortality. However, the decision to undergo breast cancer screening depends on whether the person has health insurance, as well as on education level and awareness of breast cancer symptoms. For instance, women who lack

*Department of Economics & Finance, Canisius College, 2001 Main St., Buffalo, NY 14208. Phone: (716) 888-2721.
E-mail: bojinove@canisius.edu

health insurance, are poor, less educated, or live without a husband tend to have the lowest mammography use due to their limited access to health care. According to the National Center for Health statistics, African-American, Hispanic, and American Indian women are more likely to be diagnosed with breast cancer at a later stage of disease development, which will affect their hospital expenditures.

Furthermore, these expenditures depend not only on the clinical status of patients, but also on the duration of stay, reason for admission, and whether this is a first admission or a re-admission to a hospital. For example, a woman in an advanced stage of breast cancer during her first admission to a hospital will undergo more diagnostic and therapeutic procedures and as a result her expenditures will be high as compared to, say, her third re-admission when she can have chemotherapy or radiation therapy only. A number of studies have shown that there is a direct relationship between length of stay and hospital charges but as the length of stay decreases expenditures decrease less than proportionately because the latter is associated with higher intensity of treatment during an early hospital stay. Also, for cancer patients it was found that the "cost of treatment may decrease with severity because of the futility of any further active intervention, while at the same time mortality rate goes up for each stage and substage" (Medstat Disease Staging Software Reference Guide, Healthcare Cost and Utilization Project 2002, p. 15). Total expenditures are a good proxy for the intensity of treatment because as mentioned above they capture the cost of treatment of the received health care which varies with the number of performed procedures at the hospital, duration of stay, reason for admission, and health status of the patient.

Racial discrimination, both individual and institutional, along with the feelings of inferiority of the minority groups, can adversely affect health. Socioeconomic status (SES) can explain part of the observed racial disparities in health (see for instance Bradley et al., 2002; Cross et al., 2002; Newman et al., 2002). However, racial differences can still be observed even after controlling for SES.

The purpose of this paper is to investigate whether there are racial disparities in the intensity of treatment of breast cancer patients, which can explain the higher mortality rate for African-Americans. Hospital expenditures recorded for an inpatient claim are used as a proxy for the intensity of treatment. Based on cross-sectional data from the Healthcare Cost and Utilization Project (HCUP) for the year 2002, different specifications are employed to estimate if black and other racial groups such as Hispanics, Asians, Native Americans, and others are treated less intensively in comparison with white Americans when they are admitted to hospitals in 35 states. The results suggest that there is some evidence for the presence of racial disparities in this particular year.

The paper is structured as follows. A brief literature review is presented in the next section followed by a description of the data. Then, the econometric models and estimation techniques are introduced. Next, the estimation results are presented. The final section provides a short discussion, summarizes the findings, and concludes.

II. Previous Literature

A lot of research has been done in explaining the differences in cancer survival rates among different socioeconomic groups. Some of the studies considered several cancer sites, whereas others concentrated on a single cancer site. A study by Kravdal (2000), based on individual register and census data for the whole Norwegian population for the period 1960-1991, investigated the social differentials in survival from twelve types of cancer (including breast cancer). The author found that excess mortality was about 15 percent lower for patients who had a post-secondary education as compared to those with compulsory schooling after controlling for age, stage at the time of diagnosis, and registered differences in tumor characteristics. Figueroa and Breen (1995) analyzed cases of breast and cervical cancer diagnosed in the period 1989-1990 in San Francisco, Detroit, and Atlanta. They found that 87 percent of the breast cancer cases were diagnosed late, when the tumor was already malignant. A significant part of the variation in diagnostic stage was explained by residence in an underclass area. The likelihood of late-stage diagnosis also increased with age and was higher for females living without a spouse. Katz and Hoffer (1994) found similar results for breast cancer patients living in Ontario, Canada. According to them factors such as knowledge, attitudes, transportation, differential physician advice, and time constraints explain why poor women have a lower likelihood of receiving screening tests.

A majority of breast cancer survival studies found evidence suggesting that the socially advantaged have better survival rates after controlling for possibly earlier detection of the disease among people from higher social classes (see for example LeMarchand et al., 1984; Bassett and Krieger, 1986; Karjalainen and Pukkala, 1990; Gordon et al., 1992; Ansell et al., 1993; Schrijvers et al., 1995).

Health insurance really matters when it comes to breast cancer screening, surgical procedures or other treatment procedures, because it affects treatment and hospital choices. Mitchell and Hadley (1997) analyzed hospital inpatient discharges of nonelderly women diagnosed with breast cancer in 1988 and 1991 for five states (CA, MD, MA, NJ, and NY). The authors found that the probability of breast-conserving surgery is 2.7 percent lower for females enrolled in HMOs, 4.8 percent lower for Medicaid and 6.6 percent lower for self-pay patients as compared to females having private insurance plans. Thorpe and Howard (2003) found substantial differences in cancer spending by insurance status based on the Medical Expenditure Panel Survey for 1996-1999. They considered five big cancer types, among which was breast cancer. Their results showed that “uninsured patients under age sixty-five spent 57 percent as much over a six-month period as privately insured patients spent on their cancer care” (p.189). They concluded that nonelderly cancer patients without health insurance have higher risk of being inadequately treated especially if they are of Hispanic origin.

A number of studies utilized data from the Healthcare Cost and Utilization Project (HCUP) to investigate racial and sex disparities in the treatment of patients diagnosed with various diseases including some types of cancer (see for instance Ball and Elixhauser, 1996; Harris et al, 1997; Andrews and Elixhauser, 2000; Shenn, 2002; Dowell et al., 2004). Andrews and Elixhauser (2000)

examined whether there is difference in the rate of receiving therapeutic procedures between Hispanic and white non-Hispanic patients based on 1993 discharge data for California, Florida, and New York (states that account for half of the Hispanic population in the United States). They used logistic regressions to estimate the likelihood of receiving a major therapeutic procedure for 63 conditions controlling for age, gender, disease severity, health insurance, income level of patient's community, and hospital characteristics. Their results showed that Hispanics were undertreated in a sense that they were less likely to receive major therapeutic procedures for 38 percent of the 63 conditions they examined and more likely for six percent of the conditions as compared to non-Hispanic whites. Dowell et al. (2004) found significant racial and sex disparities in the access to health care, lengths of stay, and types of procedures performed for Type 2 diabetes patients above 40 years of age for the period 1994-1997.

This paper also utilizes HCUP data but looks at a different cancer site (i.e. cancer of the female breast). The study also employs various estimation procedures to check for the presence of racial differences in the intensity of treatment of patients suffering from breast cancer who were admitted to a hospital in one of the thirty-five states participating in the HCUP project. The paper sheds more light in this less researched area.

III. Data Description

The source of data is the Healthcare Cost and Utilization Project for the year 2002, Agency for Healthcare Research and Quality. The nationwide inpatient sample (NIS) consists of approximately 7.85 million hospital stays from about 1,000 hospitals in the United States. It covers 35 states and is designed to approximate a twenty-percent sample of U.S. community hospitals, which allows for making inferences for the country as a whole. The advantages of using HCUP data are the availability of a large number of inpatient records, good data on health insurance and hospital characteristics, and different disease diagnoses. It should be noted that the unit of observation in this data set is an inpatient claims record, not the patient. As a result, it is possible for a patient to have more than one hospital stay in a given year and this will be considered as a different observation. The data are also censored because we observe only the individuals that go to a hospital for a treatment and file a claim.

This study focuses on breast cancer inpatient stays and restricts the HCUP sample to discharges with a principal “breast cancer” diagnosis based on ICD-9-CM codes. This reduced the sample to 22,678 observations. Observations with missing values for the variables of interest, i.e. race and total charges, were deleted. One hundred and fifty six observations for male breast cancer patients were also dropped from the sample to avoid potential unobserved gender differences with regards to treatment and disease development. Furthermore, the state of Georgia did not report the race of patients in 2002 due to confidentiality of reports. Also missing are race data for some of the other states. As a result, after deleting these observations the sample size decreased by 6,084 inpatient records.

The following variables from the NIS are used in the study: total charges (*totchg*), median household income for the patient's zip code (*zipinc*), length of stay in days (*los*), number of procedures on this record (*npr*), died during hospitalization (*died* – a dummy variable equal to one if the patient died during the hospital stay and zero otherwise), expected primary payer (*pay1*), age in years at admission (*age*), and whether the admission was elective (*elective* – also a categorical variable). The dependent variable *lntotchg* is the natural logarithm of total charges for an inpatient stay. The logarithmic transformation is used to account for possible skewness of the expenditure distribution. Two dummy variables for race, black and other race, are created. Each variable is equal to one if the patient is black or other race respectively, and zero otherwise. The indicated category is white. Hispanics, Asians, Native Americans, and others are combined into one dummy variable called *otherrace* because I am generally interested in the potential disparities between African-Americans and non-Hispanic whites. I also create categorical variables for health insurance status and median household income for the patient's zip code. The expected primary payer variables are Medicaid, Medicare, and private insurance (including HMOs and PPOs). The indicated category is other expected primary payer, which includes self-pay, charity, and the like. The median household income for a patient's zip code is not a continuous variable but instead income is reported in ranges, i.e. from $1 to $24,999, from $25,000 to $34,999, from $35,000 to $44,999, and from $45,000 or more. Thus, the following dummy variables are generated to correspond to these ranges: low income, below median income, and above median income. The indicated category is high income ($45,000 or more).

As mentioned before, hospital expenditures depend on disease severity. Variables such as disease staging and comorbidity measures developed by the Agency for Healthcare Research and Quality (AHRQ) are included to control for this relationship. I merged the national inpatient sample with HCUP severity data by a common code that is uniquely defined in both datasets. The AHRQ comorbidity measures define thirty different coexisting medical conditions, which are likely to be present prior to the hospital stay and are not directly related to the reason for hospital admission or principal diagnosis. All these comorbidity measures are defined as categorical variables (equal to one if the patient has the disease and zero otherwise). The presence of comorbidities can increase the cost of treatment, so it is important to rule out their impact on total expenditures for breast cancer patients. Disease staging criteria, developed by Medstat, define the severity for different medical diseases. They are measured on a scale from 1 to 4 with stage one being a disease with no complications; stage 2 is a disease with local complication; stage 3 corresponds to an increased disease complexity – it involves multiple sites or has systemic complications; and stage 4 is death. The staging variable (*ds_stage*) is measured on a continuous scale (has substages) to better represent the severity of a particular disease.

The summary statistics by race are presented in Table 1. White non-Hispanics seem to be a little older compared to breast cancer patients of African-American origin or belonging to other race. The length of stay in a hospital is somewhat longer for African-Americans than for other racial groups on

average (3.41 vs. 2.62 and 2.41). In addition, in this data set African-Americans have relatively higher hospital expenditures, a lower rate for elective admission, a higher mortality rate, more advanced stage of the disease at admission, and are poorer on average in comparison with whites and patients of other races.

Because hospital expenditures can be correlated with hospital characteristics, categorical variables for hospital ownership, location, region, and size were added as control variables. Such data are provided in the supplemental HCUP hospital data set, which I merged with the NIS and severity data sets. It is important to mention that the perception of a patient about the disease and the social support she receives from her family and friends can influence the timing of hospitalization and the length of stay, which indirectly affect hospital expenditures. However, they cannot be easily measured and proxies are unavailable in the HCUP data set.

Table 1: Summary Statistics by Race

	Black			***Other Race***			***White***		
Variable	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**
Age	1672	57.31	14.42	1634	56.75	14.15	12053	62.79	14.45
Length of stay (los)	1672	3.41	4.73	1634	2.62	3.80	12053	2.41	2.77
Number of procedures (npr)	1672	1.85	1.22	1634	1.91	1.26	12053	2.03	1.24
Total charges (totchg)	1672	18162.04	22945.78	1634	17930.95	18109.24	12053	15409.90	14147.13
Log of total chages (lntotchg)	1672	9.47	0.78	1634	9.51	0.73	12053	9.39	0.69
Died	1670	0.03	0.17	1633	0.02	0.14	12039	0.01	0.12
Elective	1664	0.72	0.45	1628	0.79	0.41	12025	0.84	0.37
Low income	1672	0.12	0.33	1634	0.07	0.26	12053	0.02	0.13
Below median	1672	0.28	0.45	1634	0.17	0.37	12053	0.18	0.38
Above median	1672	0.26	0.44	1634	0.21	0.41	12053	0.25	0.43
Disease stage (ds_stage)	1580	1.32	0.71	1614	1.20	0.58	10936	1.16	0.51

IV. Econometric Model and Estimation Techniques

The incurred hospital expenditures serve as a proxy for the intensity of breast cancer treatment. Using expenditures, however, should be done with caution because there may be issues with co-insurance and health insurance reimbursement to doctors for government provided insurance such as Medicaid and Medicare. In addition, there is a possibility that some of these expenditures may be due to defensive medicine. Therefore, various specifications are employed to estimate the possibility of racial disparities in the intensity of treatment of breast cancer patients controlling for health insurance status. To avoid the possibility of highly skewed expenditures or having big outliers that dramatically change the mean, I use the logarithm of expenditures as the dependent variable and estimate several model specifications via ordinary least squares (OLS). Since there can be omitted variable bias that causes heteroskedasticity, I employ the Huber-White correction to the OLS regressions. Thus, the

standard errors of the estimates will be consistent and inferences can be made. The basic model can be specified as follows:

$$\text{lntotchg} = \alpha + \beta X + \gamma Race + \theta HI + \varepsilon$$

where *X* is a vector of inpatient claims' characteristics such as age in years at admission, length of stay in the hospital, number of procedures on record, dummy variables for patient's median income, whether the person died in the hospital and whether the admission was elective. The coefficients on the race dummy variables (*black* and *otherrace*) show the difference between the respective base category and the indicated category, white female patients with breast cancer, in terms of log of total expenditures. The differences among patients with regard to their health insurance (HI) status are controlled for with three dummy variables (*Medicaid, Medicare, and Private Insurance*).

In the next specifications, I include controls for hospital characteristics such as location (urban or rural – specifications 5-7), ownership/control (public, private for profit, and private not-for-profit – specification 7), size approximated by the number of beds (small, medium, and large – specifications 6 and 7), and region (Northeast, Midwest, South, and West – specification 6). Furthermore, to take into account that expenditures depend on disease severity, I include dummy variables for disease staging and comorbidities. In addition, I expand the model by adding an interaction term between age and race to account for possible differences in hospital expenditures for women of different ethnicities at different ages. Thus, the expanded model is:

$$\text{lntotchg} = \alpha + \beta X + \gamma Race + \theta HI + \varphi Hospital + \delta Severity + \lambda(XRace) + \varepsilon$$

I also estimate the expanded model via OLS for Medicaid and Medicare claims separately, which solves the problem with having various out-of-pocket expenditures and prices for given procedures provided to patients with different types of health insurance.

Finally, I estimate quantile regression models introduced by Koenker and Bassett (1978) that focus on the median, the 25th, and 75th percentiles of hospital stays. In these regressions, the dependent variable is the number of procedures, which is a reasonable proxy for the health care and attention a patient receives when controlling for all the other variables that affect hospital stays as mentioned above (length of stay, HI, income level, disease stage, hospital characteristics, comorbidities, and race). As a result, inferences for possible racial disparities at more similar hospital stays in terms of the number of procedures can be made.

V. Estimation Results

The estimation results for the different specifications (one through seven) are provided in Table 2. The signs of the regression coefficients make sense except for the ones on race, private insurance, and elective admission in the specifications where these variables are statistically insignificant. As expected the coefficients on *npr* and *los* are positive and statistically significant at the 1 percent level

Table 2: OLS Estimation Results (Including Controls for Severity and Hospital Characteristics)

	Spec.1	Spec.2	Spec.3	Spec.4	Spec.5	Spec.6	Spec.7
Black	0.0311	-0.1851**	-0.1976**	-0.1863**	-0.2156**	-0.1805**	-0.2005
	(0.0169)	(0.0690)	(0.0692)	(0.0683)	(0.0676)	(0.0659)	(0.1209)
Other race	0.1108**	-0.1226	-0.1186	-0.1527*	-0.1658*	-0.1871**	-0.0444
	(0.0164)	(0.0673)	(0.0672)	(0.0667)	(0.0662)	(0.0628)	(0.0942)
Age	-0.0037**	-0.0049**	-0.0050**	-0.0046**	-0.0045**	-0.0046**	-0.0052**
	(0.0005)	(0.0006)	(0.0006)	(0.0005)	(0.0005)	(0.0005)	(0.0008)
Number of procedures	0.1603**	0.1614**	0.1596**	0.1522**	0.1473**	0.1420**	0.1592**
	(0.0047)	(0.0051)	(0.0051)	(0.0050)	(0.0048)	(0.0047)	(0.0073)
Length of stay	0.0963**	0.0937**	0.0932**	0.0982**	0.0983**	0.1000**	0.0996**
	(0.0043)	(0.0046)	(0.0047)	(0.0047)	(0.0047)	(0.0049)	(0.0085)
Medicare	0.0960**	0.1117**	0.1062**	0.0867**	0.0929**	0.0530*	0.0009
	(0.0268)	(0.0278)	(0.0278)	(0.0271)	(0.0269)	(0.0263)	(0.0423)
Medicaid	0.1300**	0.1419**	0.1377**	0.1142**	0.1234**	0.0378	0.0568
	(0.0305)	(0.0312)	(0.0311)	(0.0309)	(0.0305)	(0.0290)	(0.0466)
Private insurance	0.1719**	0.1827**	0.1820**	0.1559**	0.1535**	0.1123**	-0.0031
	(0.0245)	(0.0250)	(0.0249)	(0.0247)	(0.0245)	(0.0239)	(0.0395)
Low income	-0.0906**	-0.0994**	-0.1019**	-0.1130**	-0.0680*	-0.033	-0.0675
	(0.0276)	(0.0268)	(0.0268)	(0.0269)	(0.0268)	(0.0258)	(0.0399)
Below median	-0.1166**	-0.0901**	-0.0919**	-0.1013**	-0.0407**	-0.0281*	-0.0741**
	(0.0125)	(0.0134)	(0.0133)	(0.0132)	(0.0135)	(0.0135)	(0.0207)
Above median	-0.1027**	-0.1055**	-0.1070**	-0.1094**	-0.0655**	-0.0485**	-0.1139**
	(0.0113)	(0.0119)	(0.0119)	(0.0117)	(0.0117)	(0.0114)	(0.0191)
Died in hospital	-0.5872**			-0.5376**	-0.5428**	-0.5239**	-0.7114**
	(0.0596)			(0.0624)	(0.0616)	(0.0621)	(0.1011)
Elective	0.0702**			0.0628**	0.0503**	0.0466**	-0.0105
	(0.0149)			(0.0150)	(0.0148)	(0.0143)	(0.0231)
Age*black		0.0031**	0.0032**	0.0031**	0.0031**	0.0030**	0.0018
		(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0019)
Age*other race		0.0039**	0.0038**	0.0045**	0.0043**	0.0033**	0.0024
		(0.0012)	(0.0012)	(0.0011)	(0.0011)	(0.0011)	(0.0016)
Disease stage		-0.0910**	-0.0944**	-0.0403**	-0.0442**	-0.0461**	-0.0395
		(0.0137)	(0.0136)	(0.0134)	(0.0130)	(0.0131)	(0.0246)
Metastatic cancer			0.1205**				0.1277*
			(0.0413)				(0.0641)
Small hospital				-0.2715**	-0.2859**	-0.2736**	-0.1398**
				(0.0161)	(0.0161)	(0.0151)	(0.0286)
Medium hospital				-0.1033**	-0.1183**	-0.1026**	-0.0551**
				(0.0119)	(0.0118)	(0.0113)	(0.0190)
Urban hospital					0.3029**	0.3017**	0.4182**
					(0.0146)	(0.0148)	(0.0223)
Public hospital							-0.3540**
							(0.0212)
Ptivate not for profit							-0.4430**
							(0.0192)
Northeast hospital						-0.5045**	
						(0.0144)	
Midwest hospital						-0.3780**	
						(0.0144)	
South hospital						-0.383**	
						(0.0134)	
Constant	8.9391**	9.15434**	9.1574**	9.1236**	8.859**	9.248**	9.3365**
	(0.0415)	(0.0437)	(0.0442)	(0.0460)	(0.0469)	(0.0478)	(0.0769)
Observations	15300	14130	14130	14072	14072	14072	5236
R-squared	0.36	0.35	0.35	0.37	0.39	0.45	0.43

Standard errors in parentheses; * significant at 5%; ** significant at 1%

of significance showing that the log of hospital expenditures goes up as the number of procedures and length of stay increase. The estimated coefficient on *age* is negative, which is consistent with the theory that elderly patients are treated less intensively on average. The results show that patients with low, below median, and above median income levels have lower hospital expenditures as compared to those coming from high-income zip codes (all coefficients are significant at the 1 percent or 5 percent level in all specifications excluding the last two for the low income variable).

The estimated coefficients on *died* and *disease staging* are negative implying that at a more advanced stage of the disease, including dying in the hospital, the patients will not have many opportunities for treatment and as a result their expenditures will tend to be lower. *Died* is statistically significant in all specifications and *ds_stage* is significant in specifications two through six. Regression results also show that elective admissions lead to higher hospital expenditures on average. This variable is only insignificant in specification 7, which has controls for hospital ownership.

According to the estimates of the regressions, government (Medicaid or Medicare) or private insurances are associated with higher expenditures per inpatient record (excluding specification 7 for all three types of insurance and specification 6 for Medicaid) as compared to self-pay, charity, or other types of insurance. This result is plausible considering the fact that uninsured people have to pay all of their healthcare costs out-of-pocket which will affect the number of requested treatment procedures. All estimates of the health insurance variables are significant at the 1 percent level of significance in specifications 1-5 (with controls for hospital size and location). Adding control variables for hospital region and ownership affected the signs and significance of some of the estimated coefficients. In specifications 3 and 7 various comorbidity measures are included. It turns out that the five comorbidities significantly affect expenditures in specification 3 (deficiency anemias, uncomplicated diabetes, metastatic cancer, obesity, and peptic ulcer disease) and only three in specification 7 (metastatic cancer, obesity, and peptic ulcer disease).

The main variables of interest in the regressions are the race variables and the interaction terms of the race variables with age: *black, otherrace, ageblack* (age times black)*, and ageotherrace* (age times other race). The interaction terms are positive and statistically significant in specifications 1 through 6. The coefficients on the two race variables are negative and statistically significant in all but specification 7 which provides some evidence for the presence of racial differences in the intensity of breast cancer treatment of African-Americans and other races compared to white Americans.

Estimating the model for Medicare and Medicaid claims separately with all necessary controls is helpful in avoiding issues of having differences in prices for various procedures and variation in out-of-pocket expenditures that complicate the analysis. The results of the regression based on Medicare claims illustrate that age, number of procedures, length of stay, disease staging, hospital size, location, and ownership are significant predictors of hospital expenditures. Some of the income variables and comorbidities are also statistically significant. The coefficients on black and other race, as well as on the interaction terms between age and race, are not statistically significant. The same

can be said for the race variables estimated via OLS using Medicaid claims. A possible explanation for these results is that patients of different races with such insurance plans are more alike or have similar characteristics and as a result are less likely to be treated differently.

Table 3 presents the results of the three quantile regressions - the median, the 25th percentile, and the 75th percentile. The estimated coefficients on age, length of stay, and disease stage are significant at the 1 percent level in all regressions and have the expected signs. Three of the comorbidity measures (anemia, metastatic cancer, and uncomplicated diabetes) are significant at least at the 5 percent level for the different percentiles. The estimated coefficients on some of the health insurance variables (Medicare or Medicaid) and hospital characteristics (small hospital, medium hospital, and public hospital) in some of the quantile regressions are insignificant. The race variables (black and other race) are negative and statistically significant for the median, 25th, and the 75th percentiles. These results suggest that racial disparities in the intensity of treatment of breasts cancer are present for the inpatient claims in the 75th percentile (claims with high number of procedures per hospital stay or in the top 25th percentile), 50th percentile (claims with a median number of procedures), and 25th percentile (claims at the lowest 25 percent of the sample), which confirms the results found in the OLS regressions (specifications 1-7).

VI. Discussion and Conclusion

According to the SEER cancer statistics, white female Americans have the highest prevalence of breast cancer among different racial groups in the United States but African-Americans have the highest mortality rate. One possible explanation for this outcome is the difference in the socioeconomic status between the two races (see for instance Ward et al., 2008 and the published statistics by the US Census Bureau). The disparities in the educational level, income, health insurance status, marital status, and the like can affect the timing of breast cancer diagnosis and the stage of the cancer at diagnosis which in turn will have impact on the rate of survival of the patient. Another reason for the higher mortality rate of African-Americans can be due to differences in the healthcare treatment.

The goal of this paper is to investigate whether there are disparities in the intensity of breast cancer treatment of patients of different races such as white non-Hispanics, African-Americans, and others (Hispanics, Asians, and Native Americans) when they are admitted to hospitals in 35 states. Based on cross-sectional data from the Healthcare Cost and Utilization Project for year 2002, different models are estimated using ordinary least squares and quantile regressions. The estimated coefficients on the race variables are negative and statistically significant in most of the specifications providing some evidence for racial discrimination. In particular, in the majority of the specifications, African-Americans are found to be treated less intensively (or have lower log of total expenditures for treatment) than Caucasians when controlling for differences in terms of health insurance plans, hospital characteristics, disease severity and various comorbidities, as well as other patient specific

Table 3: Quantile Regressions

VARIABLES	Median	25 Percentile	75 Percentile
Age	-0.015***	-0.000***	-0.022***
	(0.002)	0.000	(0.002)
Length of stay	0.086***	0.002***	0.124***
	(0.007)	0.000	(0.007)
Black	-0.899***	-0.013***	-1.449***
	(0.278)	(0.003)	(0.289)
Other race	-0.480**	-0.017***	-0.488*
	(0.235)	(0.002)	(0.251)
Medicare	0.086	0.001	0.275**
	(0.102)	(0.001)	(0.108)
Medicaid	-0.003	0.002**	0.13
	(0.113)	(0.001)	(0.121)
Private insurance	0.302***	0.005***	0.452***
	(0.094)	(0.001)	(0.101)
Low income	-0.176*	-0.002**	-0.418***
	(0.100)	(0.001)	(0.108)
Below median	-0.182***	-0.001***	-0.276***
	(0.051)	(0.001)	(0.054)
Above median	-0.044	0	-0.160***
	(0.046)	0.000	(0.049)
Age*black	0.011**	0.000***	0.017***
	(0.004)	0.000	(0.005)
Age*other race	0.003	0.000***	0.003
	(0.004)	0.000	(0.004)
Disease stage	-0.351***	-0.497***	-0.306***
	(0.038)	0.000	(0.041)
Small hospital	-0.08	-0.002**	-0.177**
	(0.067)	(0.001)	(0.071)
Medium hospital	-0.136***	-0.001	-0.064
	(0.048)	0.000	(0.051)
Urban hospital	0.254***	0.003***	0.196***
	(0.059)	(0.001)	(0.064)
Public hospital	0.137**	0.001	0.179***
	(0.058)	(0.001)	(0.062)
Private not for profit	0.100**	0.001*	0.099**
	(0.044)	0.000	(0.047)
Elective	0.160***	0.003***	0.169***
	(0.053)	(0.001)	(0.056)
Anemia	0.477***	0.971***	0.350***
	(0.102)	(0.001)	(0.106)
Uncomplicated diabetes	-0.131**	-0.001**	-0.204***
	(0.061)	(0.001)	(0.065)
Metastatic cancer	0.911***	1.005***	0.905***
	(0.188)	(0.002)	(0.197)
Constant	2.384***	1.515***	3.420***
	(0.178)	(0.002)	(0.189)
Observations	5,239	5,239	5,239

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

characteristics. This result is also confirmed by the quantile regression for inpatient claims in the 75th, 50th, and 25th percentiles of hospital stays in terms of number of procedures.

This study contributes to the existing literature on racial disparities in health care and more specifically focuses on breast cancer which is the cancer with the highest mortality rate for American women. Providing more evidence to prove that African-Americans and other races are treated less intensively when admitted to hospitals for breast cancer procedures can help policy makers develop strategies specifically aimed at improving the provision of health care to minority groups. More studies investigating this topic will be helpful in determining whether this undesirable outcome continues to persist or if the racial disparity gap narrows over time. A possibility for future research is to investigate this question using time-series data, as well as analyses by regions or states.

ENDNOTES

1. The regression results for Medicare and Medicaid are available from the author upon request.

REFERENCES

Altekruse S.F., C.L.Kosary, M. Krapcho, N. Neyman, R. Aminou, W. Waldron, J. Ruhl, N. Howlader, Z. Tatalovich, H. Cho, A. Mariotto, M.P. Eisner, D.R. Lewis, K. Cronin, H.S. Chen, E.J. Feuer, D.J. Stinchcomb, B.K. Edwards (eds). SEER Cancer Statistics Review, 1975-2007, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2007/, based on November 2009 SEER data submission, posted to the SEER web site, 2010.

Andrews, R.M. and A. Elixhauser (2000) "Use of Major Therapeutic Procedures: Are Hispanics Treated Differently than Non-Hispanic Whites" *Ethnicity & Disease*, 10(3): 384-394.

Ansell, D, S. Whitrnan, R. Lipton and R. Cooper (1993) "Race, Income, and Survival from Breast Cancer at Two Public Hospitals" *Cancer*, 72(10): 2974-2978.

Ball, J.K. and A. Elixhauser (1996) "Treatment Differences between Blacks and Whites with Colorectal Cancer" *Medical Care*, 34(9): 970-984.

Bassett, M.T. and N. Krieger (1986) "Social Class and Black-White Differences in Breast Cancer Survival" *American Journal of Public Health*, 76(12): 1400-1403.

Bradley, C.J., C.W. Given, and C. Roberts (2002) "Race, Socioeconomic Status, and Breast Cancer Treatment and Survival" *Journal of the National Cancer Institute* 94: 490-496.

Cancer Trends Progress Report – 2011/2012 Update (2012), National Cancer Institute, NIH, DHHS, Bethesda, MD, August, http://progressreport.cancer.gov.

Cross, C.K., J. Harris, and A. Recht (2002) "Race, Socioeconomic Status, and Breast Carcinoma in the U.S: What have We Learned from Clinical Studies" *Cancer* 95: 1988-1999.

Dowell, M.A., B. Rozell, D. Roth, H. Delugach, P. Chaloux, and J. Dowell (2004) "Economic and Clinical Disparities in Hospitalized Patients with Type 2 Diabetes" *Journal of Nursing Scholarship*, 36(1): 66-72.

Figueroa, Janis B. and Nancy Breen (1995) “Significance of Underclass Residence on the Stage of Breast or Cervical Cancer Diagnosis” *The American Economic Review*, 85(2): 112-116.

Gordon, N.H., J.P. Crowe, J. Brumberg, and N.A. Berger (1992) “Socioeconomic Factors and Race in Breast Cancer Recurrence and Survival” *American Journal of Epidemiology,* 135(6): 609-618.

Harris, D.R., R. Andrews, and A. Elixhauser (1997) “Racial and Gender Differences in Use of Procedures for Black and White Hospitalized Adults” *Ethnicity & Disease*, 7(1): 91-105.

HCUP Databases, Healthcare Cost and Utilization Project (HCUP), 2002. Agency for Healthcare Research and Quality, Rockville, MD, www.hcup-us.ahrq.gov/databases.jsp.

Karjalainen *S.* and E. Pukkala (1990) “Social Class as a Prognostic Factor in Breast Cancer Survival” *Cancer*, 66(4): 819–826.

Katz, S.J. and T.P. Hofer (1994) **“**Socioeconomic Disparities in Preventive Care Persist Despite Universal Coverage. Breast and cervical cancer screening in Ontario and the United States**”** *The Journal of the American Medical Association*, 272(7): 530-534.

Koenker, R. and G. Bassett Jr.(1978) ”Regression Quantiles” *Econometrica*, 46 (1): 33-50.

Kravdal, Oystein (2000) “Social Inequalities in Cancer Survival” *Population Studies,* 54(1), 1-18.

LeMarchand, L., L.N. Kolonel, A. Nomura (1984) “Relationship of Ethnicity and Other Prognostic Factors to Breast Cancer Survival Patterns in Hawaii” *Journal of the National Cancer Institute*, 73:1259-65.

Medstat Disease Staging Software Reference Guide, Healhcare Cost and Utilization Project 2002, Thomson Reuters, http://www.hcup-us.ahrq.gov/db/nation/nis/DiseaseStagingV5_25ReferenceGuide.pdf

Mitchell, Jean M. and Jack Hadley (1997) “The Effect of Insurance Coverage on Breast Cancer Patients’ Treatment and Hospital Choices” *The American Economic Review*, 87(2): 448-453.

Newman, L.A., J. Mason, D. Cote, Y. Vin, K. Carolin, D. Bouwman, and G.A. Colditz (2002) “African-American Ethnicity, Socioeconomic Status, and Breast Cancer Survival: A Meta-Analysis of 14 Studies Involving over 10,000 African-American and 40,000 White American Patients with Carcinoma of the Breast” *Cancer,* 94 (11): 2844-54.

Schrijvers, C.T., J.W. Coebergh, L.H. Van der Heijden, and J.P. Mackenbach (1995) “Socioeconomic Status and Breast Cancer Survival in the Southeastern Netherlands, 1980-1989” *European Journal of Cancer*, 31A(10):1660–1664.

Shen, J.J. (2002) “Severity of Illness, Treatment Environments and Outcomes of Treating Acute Myocardial Infarction for Hispanic Americans” *Ethnicity & Disease*, 12(4):488-98.

Thorpe, Kenneth E. and David Howard (2003) “Health Insurance and Spending among Cancer Patients” *Health Affairs*, 189-198. US Census Bureau. http://www.census.gov/population/www/socdemo

Ward, E., M. Halpern, N. Schrag, V. Cokkinides, C. DeSantis, P. Bandi, R. Siegel, A. Stewart, and A. Jemal (2008) “Association of Insurance with Cancer Care Utilization and Outcomes” *CA: A Cancer Journal for Clinicians*, 58(1):9-31.