Is More Care Better?
Economic Returns to Treatment Intensity

Preliminary: Do Not Cite

Lauren Hersch Nicholas*
University of Michigan

Abstract

This paper examines the effect of elective surgery receipt on paid and unpaid labor supply amongst older adults. I find that osteoarthritic patients receiving elective hip or knee replacements in hospital referral regions with high rates of elective procedure use, where the marginal surgery is believed to be more discretionary, increase participation in paid work and volunteering after surgery. The response to surgery occurs primarily on the extensive margin, a 5.9 percentage point increase in work and an 11.6 percentage point increase in volunteering. I show that high and low intensity regions perform surgery on patients with similar levels of disability, but amongst patients who do not receive surgical intervention, levels of disability are higher in low intensity regions. I also find returns to medical management for paid work in low intensity regions only, evidence that regions specialize in particular treatment styles. Findings suggest that there are important economic effects of medical treatment intensity.

*lnichola@umich.edu. I am grateful for funding from the National Institute on Aging through grant AG000221-17 and grant UL1RR024986 from the National Center for Research Resources. This content does not represent the official views of the National Institutes of Health and all errors are my own. Morris Hamilton and Meaghan Hafner provided excellent research assistance. I appreciate helpful comments from Charlie Brown, Ken Langa, Courtney van Houtven, and David Weir as well as seminar participants at RAND, the University of Michigan, and Umea University.

1 Introduction

One of the most contentious topics in health care policy is whether high levels of spending and treatment intensity improve patient outcomes. Although health care spending in the United States averages nearly double spending in other OECD nations, we lag behind on health outcomes including life expectancy and infant mortality (Aaron and Ginsburg, 2009). Within the U.S., researchers at the Dartmouth Atlas Project identify significant geographic variation in use of potentially discretionary and expensive services without corresponding gains in longevity or patient satisfaction using Medicare claims data (Fisher et al., 2003a,b). Cutler et al. (2006) also question the returns to Medicare spending, estimating that the cost per additional year of life gained by 65 year olds between 1990 and 2000 totaled $145,000 in additional medical spending.

These findings have sparked a policy debate about whether treatment intensity, such as use of surgical intervention relative to medical management, should be reduced. This question is particularly salient for Medicare, the Federal health insurance program for the elderly, which currently accounts for more than 3 percent of GDP, and is projected to reach 6 percent by 2050 (Social Security and Medicare Trustees, 2010). Driven by the observation that Medicare beneficiaries in high utilization hospital referral regions do not experience better outcomes than those in low-utilization regions, researchers estimate that up to 30% of health care provided through the Medicare program provides no medical benefit to patients (Fisher et al., 2003a,b).

Wide variation in rates of numerous treatments ranging from tonsillectomy to back surgery to end-of-life care have been consistently documented over the past 3 decades. This variation cannot be explained by patient characteristics or measures of underlying health differences (Garber and Skinner, 2008; Phelps, 2000; Weinstein et al., 2006). Similar patterns have been observed in counties with national health care systems and salaried physicians, which has been interpreted as evidence that physician-induced demand does not account for the variation (Phelps, 2000). Instead, given the similarities in mortality across regions and countries with different levels of spending and utilization, these patterns have been interpreted as flat-of-the-curve medical practice- providing treatment past the point of additional health benefit.

However, several recent studies suggest that there are gains to treatment intensity, with the flat-of-the-curve approach failing to capture important differences between high and low intensity regions. Economists have used identification strategies to disentangle bias that could result from patients in high-intensity settings being unobservably sicker than those in low-intensity regions. Doyle (2007) finds reductions in acute myocardial infarction (heart attack, or AMI) mortality amongst patients vacationing in high-intensity regions of Florida. Chandra and Staiger also find improvements in AMI mortality for patients receiving surgical intervention in high intensity regions, arguing that these regions specialize in surgical treatment, making the treatment appropriate for patients treated in these regions who would not benefit from procedures performed in low-intensity regions which are less skilled at the intensive treatment. Almond et al. (2010) use the discontinuity in neonatal intensive care unit use for newborns around the low birthweight cutoff to identify an improvement in survival rates amongst infants receiving more intensive treatment.

A common weakness of studies examining the returns to treatment intensity is a reliance on administrative data and an almost exclusive focus on the relationship between treatment and mortality. Administrative data provide rich detail about intensity and quality of care, but lack information about many important post-treatment patient outcomes. While greater treatment intensity may not extend patients' lives, there are likely important and currently unappreciated consequences of additional treatments for patient wellbeing. While studies of geographic variation typically assume that high-utilization areas reflect patterns of overuse, "optimal" rates of surgical intervention and elective treatment are unknown. Many of the discretionary surgical interventions characterizing regional variation in utilization and spending in the Medicare program, such as hip and knee replacement surgery, should not necessarily be expected to directly affect patient mortality. However, surgical treatment for underlying chronic conditions (i.e. arthritis), is likely to influence patient outcomes such as mobility, quality of life, and productive engagement.

A large literature has shown the importance of health status and disability on work and retirement decisions of older workers (see for example, Jones et al., 2009; McGarry, 2004; Bound et al., 1999; Kapteyn et al, 2008 and Currie and Madrian, 1999 for a review of the earlier literature). Work disability is not necessarily a permanent state; Kapteyn et al. (2008) note that half of all Health and Retirement Study respondents surveyed in 1992 who report a disability that limits work in one wave report no disability in subsequent waves. Surgery is one channel through which these disabilities could be overcome.

This paper contributes to the literature assessing the returns to health care treatment intensity by examining a new set of outcomes. I use data from the Health and Retirement Study linked to regional intensity measures to examine the effects of elective hip and knee replacement on paid and unpaid labor supply amongst older adults for arthritis. I follow a large panel of older adults with arthritis from a nationally representative study who do and do not receive elective joint replacement post diagnosis. The return to treatment intensity is identified by comparing patients living in hospital referral regions in the highest and lowest quartiles of elective hip replacement rates for Medicare beneficiaries. I find increased participation in paid work and volunteering by older adults in high-intensity regions, where elective surgeries are likely to be most discretionary. I show that while high intensity regions do not perform surgery at lower levels of patient disability, average levels of disability amongst the surgery non-recipients is slightly higher in low intensity regions. Findings suggest that high intensity regions may be better at targeting surgery to appropriate patients, while low intensity regions specialize in medical management.

Findings have important implications for policymakers considering changes to the Medicare program. Reductions in availability of discretionary treatments designed to save money may be costly in the long-term if they reduce labor supply from older adults or their adult children whose labor force decisions are tied to parents' provision of informal care and grandchild care. Payers and policymakers interested in adopting new approaches to reduce use of services with limited medical benefit may adversely affect patients if reforms fail to account for the social costs and benefits of surgical intervention on outcomes that cannot be measured in claims.

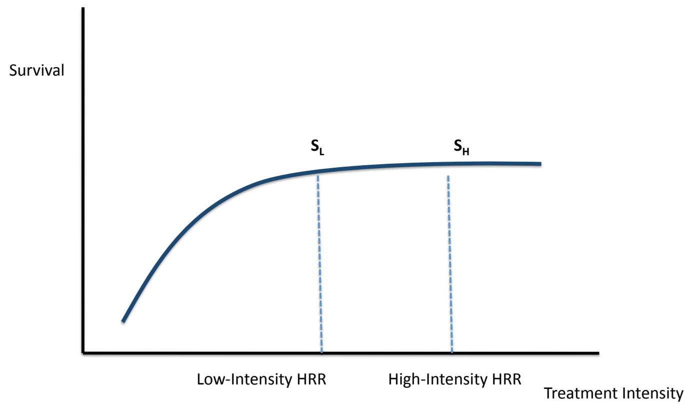
2 Background

2.1 Flat-of-the-Curve Medicine

Figure 1 illustrates the standard flat of the curve medicine argument articulated by Enthoven (1978) and Fuchs (2004). Average survival rates are observed for two groups, those treated in low-intensity regions where T_L units of health care produce survival S_L and those treated in high-intensity regions with $T_H > T_L$ producing survival rates $S_H = S_L$.

In the United States, regional variation in health care utilization is typically classified across 306 hospital referral regions (HRRs). HRRs are units of health geography developed by the Dartmouth Atlas of Healthcare and based on patient commuting patterns to hospitals. HRRs cross state lines and nearby HRRs can exhibit surprisingly different practice patterns. Figure 2 shows the variation in age, sex and race adjusted Medicare spending per capita across HRRs in 2005 calculated by the Congressional Budget Office (CBO, 2008).

Figure 1: Flat of the Curve Medicine

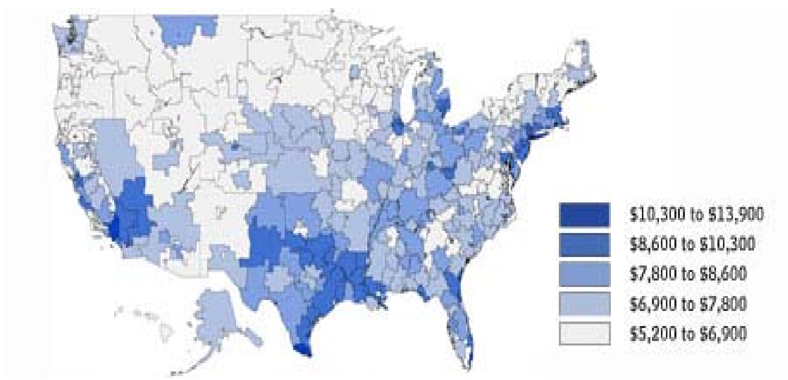


2.2 Joint Replacements

Joint replacement surgeries are common orthopedic procedures performed to alleviate pain. Procedures involve surgically removing diseased or damaged cartilage and bone from the hip or knee joint and replacing them with an artificial joint typically made of metal, plastics and polymers. The medical literature has found that replacements are an effective way of treating arthritis pain and improve mobility for many patients, though there is less consensus about which patients will benefit from surgical intervention (NIH Consensus Development Panel on Total Hip Replacement, 1995; Kane et al., 2003). Medicare spent $3.2 billion in 2000 for hip and knee replacements and demand for joint replacements is projected to increase by 60 percent over the next 30 years in response to the aging of the Baby Boom population (Kane et al., 2003; AAPM&R, 2010). Replacement surgeries are safe, low-mortality procedures where risk is derived from surgical complications or hospital-acquired conditions. Operations are generally performed on the "young old"; in the HRS data 11 percent of elective replacements occur before age 56, 35 percent before age 65 and 75 percent by age 75.

Improvements in pain and mobility are likely to facilitate increases in recipients' functional status and ability to participate in paid and unpaid work, though little is known about the extent to which the hypothesized recoveries occur. Nevitt et al. (1984) assess work disability before and after hip replacement surgery in a small, retrospective study

Figure 2: Average Medicare Per Capita Spending, 2005



of patients from a single hospital. 20 of the 58 patients in the study who report a disability prior to surgery were working one year after the replacement. More recent studies conducted in Canada (Bohm, 2008) and Britain (Mobasheri et al., 2006) also find that some previously disabled patients resume work after surgery. None of these studies compare surgery recipients to non-recipients and all are limited by small samples from single hospitals.

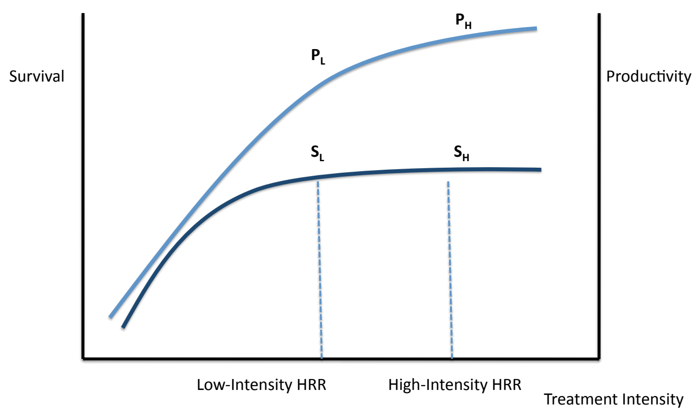
3 Theoretical Framework

3.1 Treatment Decisions

I assume that physicians chose the high-intensity treatment, elective hip or knee replacement in this case, over medical management when $E(U_H) > E(U_L)$. The optimal treatment decision depends on both the perceived treatment efficacy for the patient and the relative costs of treatment. Since the high-intensity treatment is more costly, expected returns from treatments for the high-intensity treatment must also be greater in order to satisfy $[E(U_H) - C(T_H)] > [E(U_L) - C(T_L)]$. If the benefits of the high-intensity treatment include difficult to track outcomes such as reduced disability, the flat of the curve inference based on $S_H \approx S_L$, for example in a comparative effectiveness regime, could suggest *under* rather than *over* provision.

Figure 3 illustrates the case where there are additional productivity benefits to treatment so that $P_H > P_L$. In this case, productivity can include standard economic measures of productivity including paid and unpaid work, as well as greater functioning and mobility to facilitate more productive or enjoyable leisure time.

Figure 3: Total Returns to Treatment Intensity



The proposition that treatment intensity improves productivity without changing mortality is compatible with other models, including the Grossman human capital model. In this model, the stock of health in period t represents the health of the previous period less health depreciation and increased by any investments in health. Surgery can be viewed as a greater investment in health than medical management that potentially increases the rate of health depreciation, δ_t by enabling joint replacement recipients to use their bodies more aggressively post-surgery.

3.2 Productivity Spillovers and Specialization

Chandra and Staiger (2007) note that different treatments in high and low intensity regions could produce the same average outcomes if regions specialize in different types of treatments. Thus, high-intensity regions specialize in surgical treatment, while low-intensity regions specialize in medical management. If there are spillovers, returns to intensity increase as more patients are appropriate for treatment. Treatment decisions are then based on individuals' expected returns for treatment in the region accounting for both individual appropriateness and the proportion of other patients receiving the intensive treatment. Patients choosing surgery rather than medical management in high intensity regions may experience differential returns for example because they are more likely to be treated by high-volume surgeons with greater skill and because the region's delivery system is designed to provide high quality surgical care.

Longitudinal HRS data allow me to test for an overall effect of elective surgery receipt on arthritis-related disability levels, paid and unpaid work. I compare the effects of elective surgery and medical management in high and low intensity regions to identify effects of each treatment.

4 Data

4.1 Health and Retirement Study

I use survey data from the 2002 - 2006 waves of the Health and Retirement Study (HRS), a nationally representative longitudinal study of more than 20,000 older Americans and their spouses. The HRS is designed to capture demographic, physical and cognitive functioning, work, and family structure variables related to health and retirement through biennial interviews (Juster and Suzman, 1995). The HRS data include detailed information about paid and unpaid work as well as health status (including disability) and health care utilization.

HRS respondents are asked in each wave about whether a doctor has diagnosed them with arthritis. Those who say yes are also asked whether they have had joint surgery to relieve the arthritis. This study is based on 29,733 person-waves of data from older adults who self-report as arthritic. Conditional on reporting a diagnosis of arthritis, respondents are asked a number of questions about treatments, including surgical replacement of specific joints. Respondents are asked about hip fractures as well as replacements, allowing me to focus this analysis on the effects of *elective* replacements, which replace non-broken joints. The sample includes 28,388 observations from patients who never undergo hip or knee surgery to treat their arthritis and 1,345 observations from eventual surgery recipients.

The primary outcomes are 4 categories of of paid and unpaid work. The first is a work status measure, whether the respondent is currently working for pay. Conditional on working, I also observe weekly hours. Respondents are also asked about whether and for how many hours (in a two year year period) they engage in unpaid activities including volunteering at an outside organization, informal caregiving of friends or relatives, and taking care of grandchildren.

Table 1 describes the sample. Respondents in the ever-surgery group are more likely to be working (32% vs. 28%), caring for grandchildren (30% vs. 28%), volunteering (34% vs. 31%), and providing informal care (55% vs. 49%). This is consistent either with surgery facilitating productive engagement or with those with preferences for paid or unpaid work selecting into surgery to improve physical functioning and sustain activity levels. Respondents in the ever-surgery group report lower levels of paid and unpaid work in post-surgery interviews, although this breakdown confounds effects of surgery with respondents aging.

The HRS collects a variety of socioeconomic, demographic, and health characteristics likely related to labor supply and treatment preferences. Members of the surgery sample average 0.7 years more education than the non-surgery group, and report considerably higher levels of household income, wealth, and employer-sponsored health insurance (including retiree coverage) (Table 1).

HRS respondents also provide detailed information about functional limitations. I construct an index of arthritis-related disability following Steel et al. (2008), a 0 - 10 scale with 10 indicating higher levels of disability. The index is the sum of

- Large muscle index: reported difficulties in sitting, bending, pushing (5 points)
- Mobility index: walking short and long distances and stair climbing (4 points)
- Whether arthritis limits activity (1 point).

On average, the surgery sample reports greater average disability (4.6 points vs. 3.5) as well as greater difficulty with specific components of mobility including the large muscle index, which sums reported difficulties in sitting for 2 hours, getting out of a chair, stooping, kneeling, and pushing or pulling large objects; a mobility index of walking 1 block, walking several blocks, crossing the room, climbing 1 flight of stairs and climbing several flights of stairs; the fine motor index of picking up a dime, eating and dressing activities; and the gross motor index, which combines elements of walking, stair-climbing, eating and getting out of bed. In the descriptive statistics, surgery recipients appear to become more disabled after surgery, this reflects aging since recipients are necessarily older when we observe them post-surgery.

4.2 Medicare Claims

I use 100% national Medicare claims and denominator file to calculate hospital referral region (HRR) level rates of elective hip replacements per 1,000 Fee-for-Service enrollees for the two-year periods corresponding to HRS waves.[1] HRRs are classified into quartiles of utilization for each wave of data. HRRs are ranked by quartile of utilization, with high-intensity regions exhibiting the highest rates of elective surgery per 1,000. HRS respondents are linked by zipcode and year to intensity of the HRR in which they live.

Table 2 describes the patterns of geographic variation in elective surgery. 20 percent of the arthritis sample lives in a high-intensity HRR and an additional quarter live in a low-intensity region; equal proportions of non-surgery recipients (23 percent) live in each type of region. Rates of procedure vary twofold between high-intensity and low-intensity HRRs (6.52 vs. 3.08), though there is little difference in rates of hip fracture, a measure average "hip health" across regions. Average arthritis disability scores are lower in high-intensity regions, potentially reflecting population-level gains from surgery.

5 Methods

I analyze the effect of elective surgery on several relevant economic outcomes for older adults in the years following initial diagnosis. I first test for overall effects of surgery on each of the paid and unpaid work measures W_{it} (including a binary indicator for any work and number of hours per week), I estimate equations of the form

$$W_{it} = \alpha Post * Surg_{it} + \gamma Surg_i + \delta X_{it} + Y + T + \epsilon_{it} \quad (1)$$

where $Post * Surg_{it}$ indicates responses after surgery, $Surg_i$ indicates respondents who will ever have surgery during the study period, X is a vector of time-varying demographic characteristics and non-arthritis health conditions that could affect work, Y is a linear time trend and T is a vector of dummy variables indicating years since initial arthritis

[1] Rates of elective surgery are calculated using discharges with primary procedure of total or partial hip replacement (8151 or 8152), primary diagnosis of osteoarthritis (715.9) and no hip fracture diagnosis (820.xx).

diagnosis. Respondents who report never working for pay are excluded from the paid work regressions. Hours equations are limited to the sample of respondents participating in each type of work. To account for repeated observations on individuals, standard errors are clustered at the respondent level.

Under Equation (1), the coefficient on $\alpha Post * Surg_{it}$ is likely to be biased towards 0 because patients who receive surgical interventions may have more work-limiting consequences of their arthritis than non-recipients. To aid identification of the effect of joint replacement surgery, I use a difference-in-difference type approach to compare post-diagnosis treatment of patients in areas with high rates of surgical intervention, where treatment reflects both condition acuity and treatment style to patients in low-intervention areas. I modify Equation (1) by adding a vector of indicators of HRR intensity and whether surgery in a high or low intensity HRR:

$$\begin{aligned} W_{it} &= \alpha Post * Surg_{it} + \gamma Surg_i + \tau High * PostSurg_{it} + \phi Low * PostSurg_{it} \\ &+ \zeta Low_{it} + \eta High_{it} + \delta X_{it} + Y + T + \epsilon_{it} \end{aligned} \tag{2}$$

Comparing patients in high-utilization and low-utilization areas allows inference about the effects of elective surgery for patients with different clinical thresholds for treatment. I assume that conditional on a rich set of observable respondent characteristics, the average surgical patient in a low-intensity region has greater clinical need for intervention. These sicker patients may have less likely to resume productive engagement following surgery because the intervention cannot reduce disability enough. The assumption that physicians treat patients with greater expected benefit of treatment is consistent with other models of physician decision-making (i.e. Chandra and Staiger, 2008).

I directly test whether treatment thresholds differ by regressing the arthritis disability score and years since diagnosis on HRR type and respondent characteristics amongst surgery recipients in the wave prior to surgery.

$$D_{s-1} = \zeta Low_{it} + \eta High_{it} + \delta X_{it} + Y + \epsilon_{it} \tag{3}$$

I also test for differences in disability scores across patients who *never* receive surgery. High intensity regions may perform more procedures because physicians recommend surgery at lower clinical thresholds of patient need or because criteria are more consistently applied.

$$D_{it} = \zeta Low_{it} + \eta High_{it} + \delta X_{it} + Y + \epsilon_{it} \quad (4)$$

6 Results

Table 3 presents estimates the naive version of Equation 1, which omits the geographic variation measures. There is no statistically significant relationship between receipt of surgery and any of the measures of paid and unpaid work, though eventual surgery recipients are 6.9 percentage points less likely to work for pay during the study period ($\gamma Surg_i = 0.069$, s.e. = 0.033). Health and demographic characteristics generally have the expected effects on paid and unpaid work; participation in all types of work decreases with age and indicators of disease or worsening health. Women are less likely to work and surprisingly, to provide informal care, though more likely to spend time volunteering and providing grandchild care. Blacks are more likely to provide paid and unpaid labor of all types.

Table 4 presents results from the full specification of Equation 2 including the geographic variables. These results indicate that elective surgery receipt increases participation in non-family work amongst patients treated in high intensity regions only. $\tau High * PostSurg_{it}$ is positive and statistically significant for paid work and volunteering, and positive but insignificant for informal and grandchild care. The effects are also large in magnitude, the paid work coefficient is 0.095 (s.e. = 0.048), which translates to a 5.9 percentage point increase in work force participation amongst surgery patients treated in high intensity areas holding years since diagnosis constant. However, weekly hours worked declines on average by 6.4 amongst the high intensity group after surgery.

Volunteering also increases after surgery by 11.6 percentage points in high intensity regions ($\tau High * PostSurg_{it} = 0.018, s.e. = 0.01$). In contrast, volunteering declines by 9.3 percentage points following surgery in low intensity regions.

I next consider several explanations for the returns to surgery in high but not low intensity areas. Table 5 presents results from regressions testing for differential treatment thresholds in high and low intensity areas. The first two panels show that there is no difference in level of arthritis disability or duration of time with arthritis prior to surgery across HRRs. However, at the 10% level of significance I find suggestive evidence that patients who never receive surgery in low intensity regions experience slightly greater arthritis disability burdens than those who do not receive surgery in high intensity regions ($\zeta Low = 0.085, s.e. = 0.045$). Table 6 confirms that patients in low intensity regions are 2.3 percentage points less likely to receive an elective joint replacement conditional on demographic characteristics and a comprehensive set of arthritis disability measures.

6.1 Arthritis Disability After Surgery

Elective surgery may facilitate greater productive engagement by relieving pain or arthritis disability. To test this mechanism, I reestimate Equations (1) and (2) replacing the W_{it} with the arthritis disability score. As Table 7 shows, while patients who eventually receive surgery report higher levels of disability on average, there are no statistically significant improvements in disability following surgery overall or in high or low intensity regions separately.

6.2 Specialization

Results from Tables 5 and 6 indicate the differences in use of elective surgery across high and low intensity regions is not driven by overall patient acuity. If low intensity regions perform fewer joint replacements because of specialization in medical management, we should see improved outcomes for patients receiving the low intensity treatment in low intensity regions relative to patients receiving the low intensity treatment in high intensity regions. Table 8 provides slight support for the specialization hypothesis by estimating a version of Equation (2) that replaces the surgery terms with an indicator for medical management.

I find little evidence of an average effect of medical management on any of the work variables, however $\tau Med * Low$ is positive and statistically significant for paid work, and approximately equal in magnitude to the surgery coefficient in high intensity regions ($\tau Med * Low = 0.097, s.e. = 0.04$). Since arthritic older adults in low intensity regions are less likely to work overall, the net effect of medical management in low intensity regions is a 0.05 percentage point increase. Under specialization, the average returns to surgery and medical management across regions differ because the treatments are used on different marginal populations and because the treatments themselves differ so that patients who are appropriate for surgical intervention experience greater productivity returns from treatment in regions that specialize in the high intensity treatment. Although there are productivity benefits to surgery in high-surgery regions and medical management in low-surgery regions, I do not find adverse productivity effects of receiving the non-dominant treatment in either high or low surgery regions.

7 Conclusions

A large body of literature documents the lack of correlation between health care treatment intensity and patient outcomes has motivated policy debate about whether utilization rates should be reduced in high intensity areas. While theoretical arguments suggest that additional care provided in high intensity regions may provide little value, in practice little is known about the optimal rate of surgical intervention. This paper analyzes elective joint replacements amongst older adults with arthritis, a surgical treatment that is known to be effective, though physicians lack guidance about which patients will especially benefit from intervention relative to medical management. I study an underappreciated but important consequence of health care treatment for older adults; the effects of surgery and medical management on paid and unpaid work.

I show that there are significant economic returns to elective joint replacement in high intensity hospital referral regions. Following surgery, older adults with arthritis increase participation in paid work by 5.9 percentage points relative to non-surgery recipients in other regions. Public policy efforts to reduce variation in Medicare spending and utilization by targeting high intensity regions could have adverse effects on both patient welfare and total Federal revenue. With older adults facing longer lifespans and nearing retirement with lower levels of savings than expected as a result of the recent economic downturn, there is a large role for medical intervention to facilitate longer workspans.

This paper is closely related to Chandra and Staiger's (2007) work on productivity spillovers and specialization across health care markets. I show that their finding that patients appropriate for high intensity treatments experience a greater return to the high intensity treatment in areas with higher proportions of eligible patients receiving the high intensity treatment. While their work focused on mortality following surgical versus medical management for heart attack patients experiencing an acute health event, I find similar patterns amongst patients seeking elective treatment for arthritis, a chronic condition. The implication of this result is that in the short term, efforts to equalize rates of procedure use across high and low utilization regions would reduce patient returns to treatment either if rates of surgery were reduced in high intensity regions or increased in low intensity regions due to specialization.

A recent Agency for Healthcare Research and Quality evidence report for total knee replacement notes that “The current state of empirical work does not provide a strong basis for making clinical recommendations regarding indications for outcomes from TKA. As pressures mount for more discrimination in identifying subjects for elective surgery, better information will be needed.” The authors note that a randomized trial is unlikely to succeed given and call for “a major component of research into the effectiveness of joint replacement and the patient characteristics associated with better outcomes will be well done observational studies,” (Kane et al., 2003). As the results of this paper highlight, observational studies face the additional challenge of determining whether surgery is the same treatment across all sites. Failure to compare treatments in high and low intensity regions, for example, could lead comparative effectiveness researchers to over or understate the benefits of surgical intervention.

As the newly formed Patient-Centered Outcomes Research Institute establishes the national comparative effectiveness research agenda, efforts should be made to include outcomes besides mortality as indicators of treatment efficacy. Two treatments for arthritis with little effect on mortality other than iatrogenic complications have different implications for patient productivity during the remaining years of life despite similar reports of functional status.

8 References

- Aaron HJ, Ginsburg PB. 2009. “Is Health Spending Excessive? If So, What Can We Do About It?” Health Affairs, 28(5):1260-1275.
- Almond D, Doyle JJ, Kowalski A, Williams H. 2010. “Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborns,” The Quarterly Journal of Economics 125(2): 591-634.
- America Academy of Physical Medicine and Rehabilitation. 2010. Maximize the Benefits of Hip Replacement with Rehabilitation. Available at http://www.aapmr.org/condtreat/rehab/hip.htm.
- Bohm ER. 2010. “The Effect of Total Hip Arthoplasty on Employment.” The Journal of Arthoplasty, 25(1):15-18.
- Bound J, Schoenbaum M, Stinebrinker TR, Waidmann T. 1999. ”The dynamic effects of health on the labor force transitions of older workers,” Labour Economics, 6(2) 179 - 202.
- Chandra A, Staiger DO. 2008. ”Identifying provider prejudice in healthcare,” Available at http://www.hks.harvard.edu/fs/achandr/Provider%20Prejudice%204March%202008.pdf.
- Chandra A, Staiger DO. 2008. “Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks.” Journal of Political Economy 115(1): 103-140.
- Currie J, Madrian BC. 1999. Health, Health Insurance and the Labor Market. *Handbook of Labor Economics.* Elesvier.
- Congressional Budget Office (2008). Geographic Variation in Health Care Spending. Washington, DC: The Congress of the United States: Congressional Budget Office.
- Doyle, J. 2007. Returns to local-area spending: Using health shocks to patients far from home. MIT mimeo.
- Enthoven AC. 1978. “Shattuck Lecture– Cutting Cost Without Cutting the Quality of Care,” New England Journal of Medicine, 298(22): 1229-1238.

- Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL."Implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care," Annals of Internal Medicine. 138(4): 273-287.
- Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. 2003a. "The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care," Annals of Internal Medicine. 138, 4, 288-298.
- Fuchs VR. 2004. "More Variation, More Flat of the Curve Medicine." Health Affairs Web Exclusive, VAR104-VAR107.
- Garber AM, Skinner J. 2008. "Is American Health Care Uniquely Inefficient?" Journal of Economic Perspectives, 22(4): 27-50.
- Grossman M. 2000. The Human Capital Model. *Handbook of Health Economics, Volume 1.* Elesvier.
- Jones AM, Rice N, Roberts J. 2009. "Sick of work or too sick to work?" Evidence on self-reported health shocks and early retirement from the BHPS, Economic Modeling, in press.
- Kane RL Saleh KJ, Wilt TJ, Bershadsky B, Cross WW III, MacDonald RM, Rutks I. 2003. Total Knee Replacement. Evidence Report/Technology Assessment No. 86 (Prepared by Minnesota Evidence-based Practice Center, Minneapolis, Minnesota). AHRQ Publication No. 04-E006-1. Rockville, MD: Agency for Healthcare Research and Quality.
- Kapteyn A, Smith JP, van Soest A. 2008. "Dynamics of Work Disability and Pain," Journal of Health Economics, 27(2): 496-509.
- McGarry, K. 2004. "Do changes in health affect retirement expectations?" Journal of Human Resources, 39(3): 624-648.
- Nevitt MC, Epstein WV, Masem M, Murray WR. 1984. "Work Disability Before and After Total Hip Arthoplasty." Arthritis and Rheumitism, 27(4): 410-421.
- NIH Consensus Development Panel on Total Hip Replacement. 1995. "Total Hip Replacement," Journal of the American Medical Association, 273(24): 1950-1956.

- Phelps CE. 2000. “Information Diffusion and Best Practice Adoption.” *Handbook of Health Economics, Volume 1.* Elesvier.
- Social Security and Medicare Boards of Trustees (2010). Status of the Social Security and Medicare Programs.
- Steel N, Clark A, Lang IA, Wallace RB, Melzer D. (2008). “Racial Disparities in Receipt of Hip and Knee Joint Replacements are Not Explained by Need: The Health and Retirement Study 1998 - 2004.” Journal of Gerontology: Medical Sciences, 63A(6)629-634.

Table 1: Respondent Characteristics: Never Surgery and Elective Joint Replacement Patients

	Never Surgery	Surgery	Pre-Surgery	Post-Surgery
Currently Working	0.28	0.32	0.35	0.31
	(0.45)	(0.47)	(0.48)	(0.46)
Hours| Working	33.9	35.3	35.2	35.3
	(15.7)	(15.2)	(14.4)	(15.5)
Weeks| Working	47.5	47.4	48.0	47.2
	(10.2)	(10.1)	(9.0)	(10.4)
Grandkid Care	0.28	0.30	0.32	0.29
	(0.45)	(0.46)	(0.47)	(0.46)
Hrs Grandkid| Any	1259.8	1378.6	1742.9	1262.1
	(1907.4)	(2155.6)	(2662.2)	(1957.9)
Currently Volunteering	0.31	0.34	0.32	0.34
	(0.46)	(0.47)	(0.47)	(0.48)
Hours| Volunteering	141.1	115.3	149.1	106.1
	(218.7)	(126.2)	(206.2)	(91.8)
Provide Informal Care	0.49	0.55	0.57	0.54
	(0.50)	(0.50)	(0.50)	(0.50)
Hours| Informal Care	112.6	92.5	111.9	86.5
	(202.4)	(110.0)	(163.1)	(86.9)
Disability Index	3.5	4.6	4.3	4.6
	(2.9)	(2.9)	(2.8)	(2.8)
Years Education	11.9	12.6	12.6	12.6
	(3.3)	(2.8)	(2.7)	(2.9)
Wealth	326,133	395,884	394905	396169
	(1,391,359)	(834,417)	(846593)	(831254)
Income	48,403	62,028	62432	61910
	(95,954)	(96,005)	(72623)	(101831)
Age	69.6	67.0	65.0	67.6
	(10.6)	(9.9)	(9.6)	(9.9)
Female	0.64	0.64	0.64	0.64
	(0.48)	(0.48)	(0.48)	(0.48)
Black	0.15	0.11	0.09	0.11
	(0.36)	(0.31)	(0.29)	(0.31)
Hispanic	0.08	0.07	0.09	0.06
	(0.28)	(0.25)	(0.29)	(0.24)
Married	0.61	0.68	0.69	0.68
	(0.49)	(0.47)	(0.46)	(0.47)
Hypertension	0.61	0.62	0.58	0.63

Table 1 con't

	Never Surgery	Surgery	Pre-Surgery	Post-Surgery
	(0.49)	(0.48)	(0.49)	(0.48)
Cancer	0.16	0.13	0.07	0.15
	(0.36)	(0.34)	(0.25)	(0.36)
Diabetes	0.22	0.18	0.16	0.19
	(0.41)	(0.38)	(0.36)	(0.39)
Lung Disease	0.12	0.13	0.10	0.14
	(0.33)	(0.34)	(0.30)	(0.35)
Heart Condition	0.30	0.28	0.25	0.29
	(0.46)	(0.45)	(0.43)	(0.45)
Stroke	0.1	0.05	0.03	0.06
	(0.3)	(0.23)	(0.18)	(0.24)
Years with Arthritis	4.53	4.2	1.78	4.94
	(2.53)	(2.8)	(1.55)	(2.66)
Medicare	0.69	0.66	0.57	0.68
	(0.46)	(0.48)	(0.50)	(0.47)
Medicaid	0.11	0.08	0.05	0.09
	(0.31)	(0.27)	(0.22)	(0.29)
VA coverage	0.05	0.07	0.03	0.08
	(0.23)	(0.25)	(0.18)	(0.27)
Employer-Sponsored HI	0.44	0.52	0.56	0.51
	(0.50)	(0.50)	(0.50)	(0.50)
Other HI	0.19	0.20	0.19	0.20
	(0.39)	(0.40)	(0.40)	(0.40)
Arthritis limits activity	0.41	0.59	0.54	0.61
	(0.49)	(0.49)	(0.50)	(0.49)
Gross Motor Index	0.75	0.90	0.82	0.92
	(1.30)	(1.35)	(1.29)	(1.36)
Mobility Index	1.46	1.80	1.75	1.82
	(1.59)	(1.64)	(1.61)	(1.65)
Large Muscle Scale	1.66	2.01	1.85	2.05
	(1.33)	(1.32)	(1.34)	(1.31)
ADL count	0.49	0.53	0.46	0.55
	(1.12)	(1.08)	(0.99)	(1.10)
IADL count	0.21	0.13	0.14	0.13
	(0.62)	(0.47)	(0.50)	(0.47)
Observations	28,388	1,345	303	1,042

Standard deviations in parentheses.

Health and Retirement Study data, 2002 - 2006; respondents with arthritis only.

Table 2: Hospital Referral Region Characteristics

	High-Intensity	Low-Intensity
Surgery Recipients living in HRR	0.20	0.25
	(0.40)	(0.43)
Non-Surgery living in HRR	0.23	0.23
	(0.42)	(0.42)
Elective Hip Replacements/1,000	6.52	3.08
	(0.70)	(0.59)
Hip Fractures/1,000	5.3	5.2
	(2.2)	(2.6)
Proportion of Replacements	0.91	0.86
Elective	(0.05)	(0.04)
Disability Index	3.5	3.9
	(2.8)	(3.9)

Standard deviations in parentheses.

HRS and Medicare data, 2002 - 2006.

Table 3: Effect of Elective Joint Replacement on Paid and Unpaid Labor Supply: Base Specification

	Work		Informal Caregiving		Volunteering		Grandchild Care	
	Any	Hours	Any	Hours	Any	Hours	Any	Hours
After Surgery	0.038	1.587	0.035	-26.48	0.034	-6.8	-0.002	-437.9
	(0.03)	(1.51)	(0.04)	(22.39)	(0.03)	(13.40)	(0.04)	(318.40)
Ever Surgery	-0.069*	-0.159	-0.033	10.711	-0.015	-1.5	0	435.72
	(0.03)	(1.36)	(0.03)	(22.20)	(0.03)	(12.97)	(0.04)	(297.33)
Education	0.009**	-0.210**	0.031**	5.56**	0.026**	-0.44	0.003*	-43.8**
	(0.00)	(0.08)	(0.00)	(0.80)	(0.00)	(0.61)	(0.00)	(9.07)
Age	0.118**	7.9**	-0.059**	8.587	0.02	18.4	0.25**	43.18
	(0.02)	(1.66)	(0.02)	(21.60)	(0.02)	(12.64)	(0.02)	(240.08)
Female	-0.083**	-5.72**	0.044**	-4.06	-0.09**	19.2**	0.05**	466.0**
	(0.01)	(0.46)	(0.01)	(6.63)	(0.01)	(3.68)	(0.01)	(50.50)
Black	0.006	-1.1	0.070**	5.358	-0.01	-10.37*	0.048**	545.6**
	(0.01)	(0.57)	(0.01)	(7.32)	(0.01)	(5.09)	(0.01)	(82.64)
Married	-0.03**	-2.03**	0.06**	-9.7	0.007	-9.5*	0.032**	-235.4**
	(0.01)	(0.49)	(0.01)	(7.50)	(0.01)	(4.30)	(0.01)	(67.1)
Hypertension	-0.027**	-0.55	-0.016*	-0.058	-0.024**	-2.4	0.01	52.2
	(0.01)	(0.42)	(0.01)	(5.07)	(0.01)	(3.64)	(0.01)	(52.33)
Diabetes	-0.046**	-0.17	-0.038**	2.7	-0.061**	0.38	0.013	63.2
	(0.01)	(0.57)	(0.01)	(6.22)	(0.01)	(4.35)	(0.01)	(66.94)
Lung	-0.064**	-0.326	-0.062**	-4.54	-0.038**	0.27	0.002	94.8
	(0.01)	(0.75)	(0.01)	(8.24)	(0.01)	(5.72)	(0.01)	(76.35)
Heart	-0.049**	0.191	-0.017*	0.918	-0.012	1.093	0.003	16.782
	(0.01)	(0.55)	(0.01)	(5.55)	(0.01)	(3.96)	(0.01)	(59.61)
Stroke	-0.071**	0.973	-0.080**	-3.8	-0.10**	-6.24	-0.037**	22.2
	(0.01)	(1.23)	(0.01)	(9.71)	(0.01)	(7.10)	(0.01)	(110.06)
Medicaid	-0.12**	-3.4*	-0.035**	-13.1	-0.087**	-4.54	-0.04**	68.6
	(0.01)	(1.72)	(0.01)	(11.15)	(0.01)	(8.17)	(0.01)	(123.6)
VA coverage	-0.043**	-0.527	0.02	19.4	0.024	-0.28	-0.021	-123.494
	(0.01)	(1.02)	(0.02)	(12.12)	(0.01)	(6.81)	(0.02)	(108.05)
Employer Ins	0.107**	4.02**	0.038**	-11	0.05**	-2.07	0.005	-170.1**
	(0.01)	(0.55)	(0.01)	(6.88)	(0.01)	(4.29)	(0.01)	(60.17)
Other Insurance	0.036**	0.41	0.041**	-7.37	0.050**	2.6	-0.018*	-178.9*
	(0.01)	(0.76)	(0.01)	(7.36)	(0.01)	(5.31)	(0.01)	(71.75)
R^2	0.278	0.193	0.092	0.034	0.139	0.023	0.087	0.081
N	27,028	7,844	29,177	9,050	29,172	14,423	24,947	6,951

Robust standard errors in parentheses. Health and Retirement Study data, 2002 - 2006.
Models include controls for duration of arthritis diagnosis, time trend and a quadratic in age.

Table 4: Effect of Elective Joint Replacement on Paid and Unpaid Labor Supply: Geographic Variation

	Work		Informal Caregiving		Volunteering		Grandchild Care	
	Any	Hours	Any	Hours	Any	Hours	Any	Hours
After Surgery	0.035	1.95	0.034	-27.1	0.037	-9.96	-0.006	-449.8
	(0.03)	(1.59)	(0.04)	(22.18)	(0.03)	(13.75)	(0.04)	(319.01)
Ever Surgery	-0.07*	-0.12	-0.044	17.3	-0.02	1.89	-0.004	416.8
	(0.03)	(1.49)	(0.03)	(22.20)	(0.03)	(12.51)	(0.04)	(300.27)
Surgery x High	0.095*	-3.5*	0.11*	-31.1	0.029	21.5	0.093	132.7
	(0.05)	(1.74)	(0.05)	(23.58)	(0.05)	(26.12)	(0.05)	(257)
Surgery x Low	-0.123	6.27	-0.05	-18.9	0.021	-65.2*	-0.117	-117.0
	(0.07)	(4.26)	(0.07)	(27.65)	(0.07)	(29.70)	(0.07)	(371.9)
High Intensity	0.003	-0.92*	0.02*	-6.1	0.014	-2.8	0.007	-132.4*
	(0.01)	(0.42)	(0.01)	(5.48)	(0.01)	(4.39)	(0.01)	(54.32)
Low Intensity	-0.004	0.73	-0.03**	-11.2*	-0.06**	-5.4	-0.006	-2.5
	(0.01)	(0.46)	(0.01)	(5.51)	(0.01)	(3.94)	(0.01)	(61.69)
R^2	0.278	0.194	0.094	0.035	0.141	0.023	0.087	0.081
N	26,905	7,823	29,039	9,013	29,034	14,361	24,844	6,925

Robust standard errors in parentheses. Health and Retirement Study data, 2002 - 2006.
Models control for all variables in Table 3.

Table 5: Disability Treatment Thresholds for Elective Surgery

	Disability Pre-Surgery	Years with Arthritis	Disability Never Surgery
High-Intensity HRR	0.121	0.064	0.003
	(0.21)	(0.14)	(0.04)
Low-Intensity HRR	0.125	-0.072	0.085
	(0.22)	(0.14)	(0.05)
Ever Surgery	Yes	Yes	No

Robust standard errors in parentheses.

Models control for variables included in Table 3.

Table 6: Receipt of Elective Joint Replacement amongst Older Adults with Arthritis

		Elective Surgery
High-Intensity HRR		0.002
		(0.01)
Low-Intensity HRR		-0.023*
		(0.01)
Years Education	0.002	0.002
	(0.002)	(0.002)
Age	-0.001*	-0.001*
	0	0
Female	-0.004	-0.005
	(0.009)	(0.009)
Black	(0.02)	(0.02)
	(0.013)	(0.013)
Married	-0.003	-0.004
	-0.01	-0.01
Good S-R Health	0.005	0.004
	(0.01)	(0.01)
Fair-Poor S-R Health	-0.011	-0.01
	(0.01)	(0.01)
Hypertension	0.011	0.012
	(0.009)	(0.009)
Cancer	-0.02	-0.022
	(0.01)	(0.01)
Diabetes	-0.031**	-0.034**
	(0.01)	(0.01)
Medicaid	-0.003	-0.001
	(0.02)	(0.02)
VA coverage	0.005	0.001
	(0.02)	(0.02)
Employer-Sponsored HI	0.043***	0.041***
	(0.01)	(0.01)
Other HI	0.039**	0.038**
	(0.01)	(0.01)
Arthritis limits activity	0.033***	0.031**
	(0.01)	(0.01)
Gross Motor Index	0.011	0.011
	(0.01)	(0.01)
Mobility Index	0.016*	0.016*
	(0.01)	(0.01)

Table 6 con't

		Elective Surgery
Large Muscle Scale	0.004	0.004
	(0.00)	(0.00)
Fine Motor Index	0.041**	0.037*
	(0.02)	(0.02)
ADL count	(0.019)	-0.019
	(0.01)	(0.01)
IADL count	-0.021*	-0.023*
	(0.01)	(0.01)
R^2	0.028	0.029
N	3,928	3,906

Robust standard errors in parentheses.

Health and Retirement Study Survey data, self-reported arthritic respondents 2002 - 2006.

Models also control for additional health conditions, year and duration of diagnosis.

Table 7: Effect of Elective Joint Replacement on Arthritis Disability Index

	Disability	Disability
After Surgery	-0.106	-0.121
	(0.21)	(0.21)
Ever Surgery	1.26**	1.25**
	(0.20)	(0.20)
Surgery x High		0.146
		(0.29)
Surgery x Low		0.048
		(0.39)
High Intensity		0.016
		(0.04)
Low Intensity		0.095*
		(0.04)
R^2	0.214	0.214
N	29,185	29,047

Robust standard errors in parentheses.
Models control for demographic variables in Table 3.

Table 8: Effects of Medical Management on Paid and Unpaid Labor Supply: Geographic Variation

	Work		Informal Caregiving		Volunteering		Grandchild Care	
	Any	Hours	Any	Hours	Any	Hours	Any	Hours
Medical Management	0.008	-1.09	-0.014	12.04	-0.025	12.0*	0.033	79.8
	(0.02)	(1.14)	(0.02)	(7.67)	(0.02)	(5.49)	(0.02)	(176.0)
Medical Management x Low	0.097**	-3.62	0.053	10.5	0.037	-12.9	-0.063	-300.6
	(0.04)	(2.86)	(0.04)	(15.65)	(0.04)	(11.54)	(0.04)	(346.2)
Medical Management x High	-0.007	0.072	0.002	8	-0.023	-3.7	-0.057	70.3
	(0.03)	(1.67)	(0.04)	(12.19)	(0.04)	(9.41)	(0.04)	(222.2)
Low Intensity	-0.10**	4.29	-0.078*	-21.9	-0.09*	6.53	0.054	285.5
	(0.03)	(2.82)	(0.04)	(14.73)	(0.04)	(10.8)	(0.04)	(340.3)
High Intensity	0.01	-1.0	0.015	-13.7	0.036	0.83	0.062	-200.3
	(0.03)	(1.62)	(0.04)	(10.75)	(0.04)	(8.25)	(0.04)	(214.81)
R^2	0.278	0.194	0.094	0.034	0.141	0.023	0.087	0.081
N	26,905	7,823	29,039	9,013	29,034	14,361	24,844	6,925

Robust standard errors in parentheses. Health and Retirement Study data, 2002 - 2006.
Models include controls for the same variables as Table 3.