CHAPTER SIX

Who Ordered That? The Economics of Treatment Choices in Medical Care

Amitabh Chandra, David Cutler, and Zirui Song[1]
Harvard University and NBER, USA

Contents

1. Introduction 398
   1.1. Overview of Variations in Medical Care 400
   1.2. Gray Area of Medicine 402
   1.3. Demand (Patients) and Supply (Doctors) 403
      1.3.1. Benevolent Physician 403
      1.3.2. Physician as an Imperfect Agent 404
2. Heterogeneity In Demand 406
   2.1. Price and Income Effects: Theory 406
   2.2. Empirical Evidence on Price and Income Effects 407
   2.3. Preferences 410
3. Supply-Side Drivers of Clinical Decisions 411
   3.1. Profits versus Patients 412
   3.2. Payment Systems and Supplier-induced Demand 413
   3.3. Physician Specialization and Training 415
   3.4. Defensive Medicine 418
4. Situational Factors 419
   4.1. Availability Heuristic 420
   4.2. Framing, Choice, and Risk 421
   4.3. Status Quo and Confirmation Bias 422
   4.4. Channel Factors 422
   4.5. Resource/Capability Factors 423
5. Conclusion 425
References 425

Abstract

In the United States, two patients with the same medical condition can receive drastically different treatments. In addition, the *same* patient can walk into two physicians' offices and receive equally disparate treatments. This chapter attempts to understand why. It focuses on three areas: the

[1] The authors are grateful to Thomas McGuire for helpful comments. Zirui Song acknowledges support from a National Institute of Aging Pre-doctoral M.D./Ph.D. National Research Service Award (F30-AG039175) and the National Bureau of Economic Research Pre-doctoral Fellowship in Aging and Health Economics (T32-AG000186).

ISSN: 1574-0064, DOI: 10.1016/B978-0-444-53592-4.00006-2

patient, the physician, and the clinical situation. Specifically, the chapter surveys patient or demand-side factors such as price, income, and preferences; physician or supply-side factors such as specialization, financial incentives, and professionalism; and situational factors including behavioral influences and systems-level factors that play a role in clinical decision making. This chapter reviews theory and evidence, borrowing heavily from the clinical literature.

Keywords: variations; decision making; demand heterogeneity; preferences; income; supply-side incentives; payment systems; specialization; behavioral economics

JEL Codes: D03; I10; I12; J44; L84

1. INTRODUCTION

There were 3,368 men aged 45–89 diagnosed with localized or regional prostate cancer in Seattle in 2007. Prostate cancer is the second most common cancer among men in the United States and the second leading cause of cancer deaths in men, accounting for about 200,000 diagnoses and 30,000 deaths annually (Howlader et al., 2011; Smith, 2011; Jemal et al., 2009). A large literature exists on the many treatments for prostate cancer. There are also a number of clinical guidelines for treatment choice, including recommendations from the American Urological Association, American Joint Committee on Cancer, and the National Comprehensive Cancer Network (Thompson et al., 2007; National Comprehensive Cancer Network, 2010; AJCC, 2010).

Despite this clinical experience and expansive literature, however, the treatment of these patients is wildly diverse (Wilt et al., 2008). Forty-two percent of the patients in Seattle had their prostate removed, 38 percent had radiation, and others had chemotherapy (many patients received a combination of these treatments). All the while, a substantial number decided not to pursue any aggressive treatment, but instead to monitor the progression of the cancer over time.

How were these treatment choices made? Why are they so diverse? That is the subject of this chapter. We characterize the factors influencing treatment into three categories. The first is standard demand-side factors: price, income, and tastes. In most markets, variation in purchasing decisions is explained by one of these three considerations. Ironically, given the large variation in medical treatments, the roles of price and income in medicine are blunted. Substantial insurance coverage means that prices do not vary significantly across people and income has little role to play in clinical decisions. In the prostate cancer example, most men have Medicare coverage, typically with supplemental insurance, so the cost of any treatment is minimal. Tastes can and do vary, and the different treatments play to these

different tastes (Cooperberg et al., 2007; Wei et al., 2002). Prostate removal and radiation therapy can both reduce mortality, but treatment also carries risks to quality of life, such as impotence, urinary incontinence, and disrupted bowel and hormonal function (Bill-Axelson et al., 2005; Hayes et al., 2010; Bolla et al., 1997; Sanda et al., 2008; Litwin et al., 2007; Potosky et al., 2000). Different men care about these outcomes to different degrees.

The supply side is the second factor that may influence treatment decisions. The surgeons at one hospital may be particularly good compared to surgeons at other hospitals, and thus surgical treatment may be preferred to radiation and chemotherapy for men treated at that institution. Even a casual perusal of surgical report cards shows enormous variation in risk-adjusted outcomes across surgeons and institutions. The technology available for treatment may also vary across institutions. The daVinci robotic surgical system is gaining popularity, especially at hospitals that can afford the $1.39 million fixed purchasing cost and $140,000 annual service contract. In 2009, 73,000 men (86 percent of the 85,000 who had prostate cancer surgery) received robot-assisted operations, a dramatic rise from less than 5,000 a decade ago (Kolata, 2010). Through practice or adaptive learning, physicians at different institutions may adopt faith in different guidelines or technologies, steering men along particular courses of therapy.

Finally, situational factors, such as contextual or behavioral influences, may be at work. A particular man may have preferred radiation to surgery, but the urologist on call that day happened to talk the man into surgery. The tragic or salient death of a man during surgery the previous week may lead a primary care physician to refer her next patient in similar condition to a radiation oncologist instead. The lengthy wait for a radiation appointment relative to surgery may lead a patient to choose surgery. These examples illustrate some of the richness and randomness of clinical situations.

At the most basic level, this chapter considers whether the demand side, the supply side, or what we refer to as situational factors is most important in explaining the treatment differences that we observe. Of course, the factors may be more or less appropriate at different levels of aggregation. Demand-side factors may be important in explaining treatment differences between rich and poor regions, but they may be less important within rich or poor regions. Supply-side factors may explain a good share of the differences between one region and another, or even between the United States and other developed nations. For example, prostate cancer guidelines vary within and between countries (Dahm et al., 2008). Situational factors may explain a good deal about why one man gets surgery and another gets radiation, but are likely to be a much smaller part of regional variations in care. Yet they are evident in conceivably every decision to prescribe or receive health care, and understanding them may offer the greatest potential improvements for the productivity of health care spending.

1.1. Overview of Variations in Medical Care

Variation in medical treatments is profound.[2] We noted above the variability in treatment rates for men with localized prostate cancer of the same age and living in the same geographic location. However, variations in utilization extend farther.

Jack Wennberg and colleagues pioneered the study of regional variations in health care in the United States, decades after Glover (1938) first uncovered variations in England and Wales. Wennberg and Gittelsohn (1973) showed that tonsillectomy rates varied significantly across small towns in Vermont. They also showed that treatment patterns were different in Boston, Massachusetts than in New Haven, Connecticut. So-called "small area variations" demonstrate clear differences in demand, supply, or situation across settings (Wennberg and Gittelsohn, 1982; Wennberg et al., 1987). But small area variations are less impressive in some sense. That beliefs about appropriate practice differ across individual physicians and patients is clear. The real issue is whether these differences in beliefs affect care decisions writ large.

In the past decade, there has been enormous emphasis on "large area variations"—differences in medical practice across areas where individual heterogeneity in beliefs is more likely to be averaged out and be a much less important explanation of differences (Fisher et al., 2003a and b; Baicker and Chandra, 2004). We illustrate this in the case of prostate cancer, noted above. There are 17 regions in the *Surveillance, Epidemiology and End Results (SEER)* database over the 2000–2008 time period. We selected all of the cases for local/regional prostate cancer in 2008 and analyzed the treatment that men received. We selected men who were aged 45–89, the bulk of those with prostate cancer. We coded treatment into radical prostatectomy, radiation therapy, or watchful waiting; the SEER data do not indicate if chemotherapy was provided. To control for observable differences across patients, we relate dummy variables for surgery or radiation to age (dummies for the exact year), tumor grade and size, and SEER region. The region dummy variables give the adjusted probability of receiving each treatment in each area.

Figure 6.1 plots the adjusted rate of surgery and radiation in each area. It is clear that treatment proclivities vary enormously by area. In some areas, surgery rates are very high (Seattle being one of them) and in other areas rates of radiation are high (Hawaii is the highest).

Some variation in treatment rates would be expected for natural reasons, such as case mix and patient preferences. However, the variation here is greater than random fluctuations alone would suggest. The Dartmouth Atlas documents such variation along a number of dimensions. Generally, the largest variation is in "supply-sensitive" care or "gray area medicine," in which clinical judgment plays a key role. The use of imaging services and the frequency of specialist office visits are leading examples of

[2]For extended discussion of variations, see Skinner (2011, this Handbook).

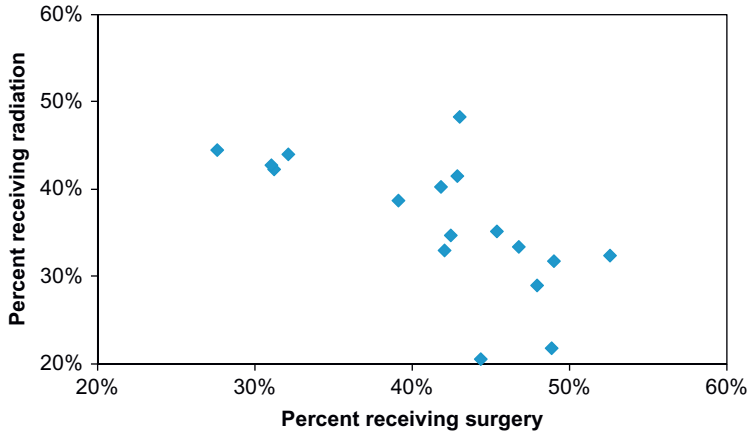

Figure 6.1 Surgery and radiation for prostate cancer in the United States, SEER regions.

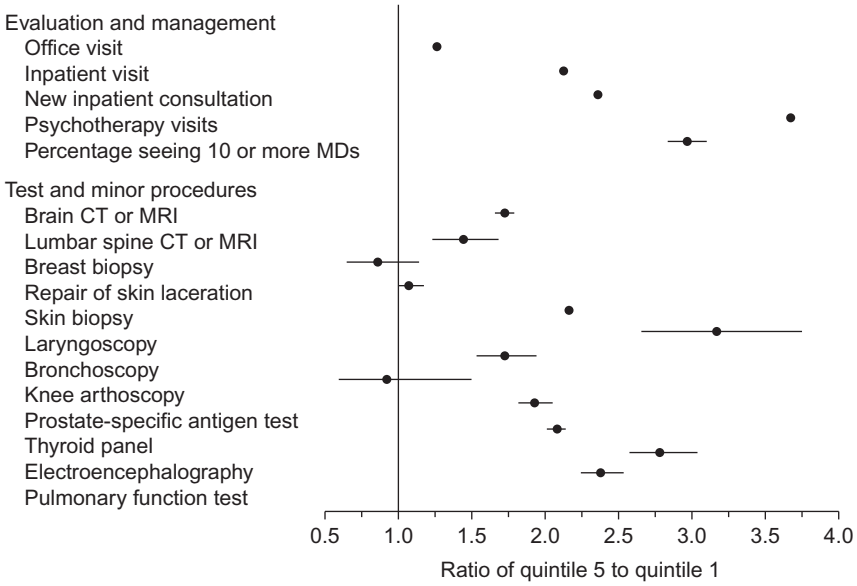

Figure 6.2 Relative rates of utilization for selected clinical services in the united states. relative rate and 95% confidence intervals of specific services provided to cohort members residing in the highest quintile of medicare spending compared with those residing in the lowest quintile for the three chronic disease cohorts combined. (Authors' adaptation of Figure 5 from Fisher, E. S. et al. (2003). *Ann. Intern. Med.*, *138*, 273−287.)

this type of care. Figure 6.2 shows the rates of utilization for certain services in the highest spending US regions (quintile 5) relative to rates in the lowest spending US regions (quintile 1), demonstrating the extensive variation.

There is significant debate about the drivers of these regional differences. Some researchers argue that variation is accounted for by population disease burden (Zuckerman et al., 2010), but other authors argue that prevalence of diagnoses itself is endogenous across areas (Song et al., 2010; Welch et al., 2011). Most of the literature agrees that patient characteristics and preferences do not explain much of the differences across areas, and that substantial variations in treatment practices remain after controlling for patient characteristics (Anthony et al., 2009; O'Hare et al., 2010; Baicker et al., 2004).

1.2. Gray Area of Medicine

The gray area of medicine, in which economic incentives are likely to have the largest impact, accounts for a significant portion of clinical practice. In many clinical situations, there are no authoritative guidelines or consensus treatment recommendations. In areas of medicine where clinical trials compare alternative treatments, we learn about average benefits. Yet the average benefit tells physicians little about a particular patient's potential marginal benefit. Moreover, average benefits say nothing about a particular physician's own impact on potential marginal benefits, owing to skill or other factors. The result is that when the same patient walks into two physicians' offices, what that patient receives may differ widely. An example of this is the frequency of follow-up for medical care. In regions of the country that spend more on health care, patients are seen back more frequently and are more likely to receive screening tests and discretionary interventions of unproven benefit relative to low-spending regions (Sirovich et al., 2008).

Another example is the use of percutaneous coronary interventions (PCI) for stable coronary disease (chest pain and associated symptoms caused by strenuous activity), where successive trials have found no survival benefit or quality of life benefit over optimal medical therapy (Boden et al., 2007; Weintraub et al., 2008). Yet despite this, some patients will surely benefit from this procedure.

PCI remains one of the most studied procedures in medicine; matters are substantially less clear for things leading up to this intervention. For a person going to the doctor with chest pain, there are over 7,000 cardiology guidelines for individual clinical decisions. Only 11 percent are based on randomized controlled trials, and 48 percent are from expert opinions, case studies, or prior standards of care (Tricoci et al., 2009). For a simple cough, there are over 4,000 infectious disease guidelines, of which 14 percent are based on randomized controlled trials and 55 percent are from opinions or case series (Lee and Vielemeyer, 2011). In addition, many of the guidelines are based on studies excluding people with multiple chronic conditions (often present in the elderly) or based on other non-random samples. Indeed, clinical guidelines in most fields of medicine suffer from poor adherence to methodological standards (Shaneyfelt et al., 1999; Atkins et al., 2004; Dahm et al., 2008).

The gray area of medicine is characterized by at least three attributes. First, as the examples above illustrate, clinical guidance is scarce. Second, the scope of marginal

harm is small, as otherwise physicians would learn very quickly to refrain from harmful treatments. Third, benefit is idiosyncratic to the patient. Such elements as pain, nausea, and quality of life enter individual utility functions in idiosyncratic ways. Given this, it is possible for physicians to claim, despite clinical trial evidence, that a particular patient has a particular preference which befits the treatment. In the typology of Chandra and Skinner (forthcoming), these are type II and type III technologies: the former denoting treatments highly effective for some but not for all (e.g. cardiac stents), and the latter referring to treatments with uncertain clinical value (e.g. ICU days among chronically ill patients).

1.3. Demand (Patients) and Supply (Doctors)

The difficulty in understanding treatment choices is not just empirical; it is conceptual as well. What the patient wants to receive (demand) may not be the same as what the physician wants to deliver (supply), and given insurance and price regulation in health care, normal forces of market equilibrium are unlikely to bring these into balance. Consider a simple model of medical treatments based on patient demand and physician assessment of the optimal treatment for the patient (Chandra and Staiger, 2007). We assume a decision is being made between two treatments for a particular disease (where "doing nothing" can be modeled as one of the choices). Denote the treatments $i = \{1,2\}$. Each treatment produces benefits to the patient of $B_i(\sigma)$, where σ is the disease severity. In addition to health, people get utility from non-medical consumption, denoted $Y - P_i$, where P_i is the price the consumer pays for the treatment and Y is income net of any insurance premium. The utility to the patient from each treatment is therefore:

$$\mathbf{U_k}(1) = B_1(\sigma_k) + V(Y_k - P_{k1}) + \theta_{k1} \tag{6.1}$$

$$\mathbf{U_k}(2) = B_2(\sigma_k) + V(Y_k - P_{k2}) + \theta_{k2} \tag{6.2}$$

where k denotes individuals. θ_{k1} and θ_{k2} are person-specific error terms that capture heterogeneity in the benefits of each treatment to that patient—for example, the part due to preferences about side-effects.

1.3.1. Benevolent Physician

In the simplest case, let us assume that physicians are perfect agents for their patients, know everything about benefits, and observe everything about preferences. In this case, there is no additional role for physician utility in the model. The physician chooses treatment 1 over treatment 2 provided $\mathbf{U_k}(1) > \mathbf{U_k}(2)$. This can be expressed in terms of the difference in the two error terms. In particular, treatment 1 will be chosen provided that:

$$\theta_{k1} - \theta_{k2} > [B_2(\sigma_k) - B_1(\sigma_k)] + [V(Y_k - P_{k2}) - V(Y_k - P_{k1})] \tag{6.3}$$

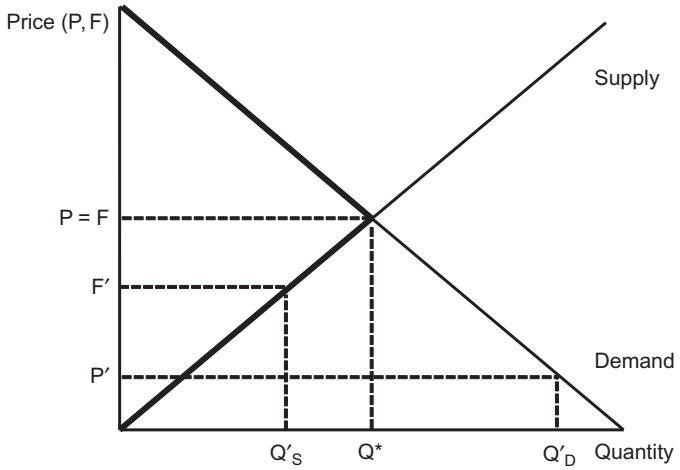

Figure 6.3 Supply and demand in standard market equilibrium. Note: P is the consumer price, F is the physician fee, Q* denotes a standard market equilibrium where P = F, and Q'_D and Q'_S denote quantity demanded and supplied, respectively.

With restrictions on the distributional form of error terms (for example, a normality assumption), this probability can be solved for directly:

$$\begin{aligned}\Pr[\text{treatment } 1] &= \Pr\{i=1\} = \Pr\{\mathbf{U_k}(1) > \mathbf{U_k}(2)\} \\ &= \Pr\{[B_1(\sigma_k) - B_2(\sigma_k)] + [V(Y_k - P_{k1}) - V(Y_k - P_{k2})] > \theta_{k2} - \theta_{k1}\} \\ &= \Pr\{\Delta B(\sigma_k) + \Delta V(Y_k - P_k) > \Delta\theta_k\}\end{aligned} \tag{6.4}$$

Equations (6.1) and (6.2) give a downward sloping demand curve for each treatment as a function of its price, holding the price of the other treatment constant (see below). This demand curve is shown in Figure 6.3. Integrating equation (6.4) over the distribution of severities (σ) in a population produces the market demand curve for treatment 1. Taking $f(\sigma)$ as the distribution of σ in the population, the following must hold in equilibrium:

$$D_1 = \int_\sigma \Pr(\Delta B(\sigma_k) + \Delta V(Y_k - P_k) > \Delta\theta_k) f(\sigma) \mathrm{d}(\sigma) \tag{6.5}$$

If providers mistakenly assess idiosyncratic patient preferences, this will result in allocative inefficiency with a potentially large utility loss. The utility loss would be further exacerbated by miscalculations of average benefits and average costs.

1.3.2. Physician as an Imperfect Agent

The supply side is given by physician desires. Modeling physician utility is complex, as we discuss below. Following Ellis and McGuire (1986), we assume that physician

utility depends on three components: the benefit to the patient (B_k), the fee the physician earns net of the costs of providing the service (F_i), and other factors (ε_j). These may include the hassles from malpractice (direct malpractice costs would be netted out in F_i), the availability of a particular service in the area or at that particular time, or the guidelines the physician chooses to follow. For simplicity, we assume that utility is linear in these three factors.

$$\mathbf{W_j}(1) = \beta_j B_1(\sigma_k) + V(F_1) + \varepsilon_{j1} \tag{6.6}$$

$$\mathbf{W_j}(2) = \beta_j B_2(\sigma_k) + V(F_2) + \varepsilon_{j2} \tag{6.7}$$

The coefficient β reflects the relative importance of patient health to physicians in comparison to profits; one interpretation is altruism, or professionalism. It may differ across physicians, as denoted by the j subscript. Similarly, ε may differ across physicians because of differing preferences or constraints. In this imperfect agency model, the provider chooses treatment 1 if:

$$\varepsilon_{j1} - \varepsilon_{j2} > \beta_j[B_2(\sigma_k) - B_1(\sigma_k)] + [V(F_2) - V(F_1)] \tag{6.8}$$

Equation (6.8) differs from equation (6.3) in several notable ways that reflect differences in treatment choice between the two models. First, physicians and patients may differ in their person-specific error terms, which would drive differences in treatment choice. Second, the physician coefficient β may be less than 1, in which case the imperfect agent undervalues patient benefit relative to the benevolent physician. In our simple model, we take professionalism to be exogenous, where a fully altruistic physician has $\beta = 1$. As β approaches 0, the relative importance of the physician's own economic motives increases. One consideration that we leave out of this simple model is what happens when β is a function of physician income or reimbursement incentives, $\beta_j = \beta_j(F_i)$. Alternatively, β may depend on the physician's place of training or organizational incentive contracts. Such endogenous versions of β would generate further variation in treatment decisions. Third, physicians may not be reimbursed on a fee-for-service basis, such as in various capitated or other managed care contracts (this could correspond to a negative net margin); moreover, the fee (F_i) need not be equal to the full marginal cost of care faced by the patient (P_{ki}) due to insurance.

Whether fees (F_i) increase or decrease, the amount of care provided is, of course, unknown; there are income and substitution effects. For now, we assume that supply is upward sloping in fees, which our empirical review will demonstrate is generally the case. Figure 6.3 shows an upward sloping supply curve.

In standard markets (markets without insurance and where buyers are perfectly informed), prices adjust so that demand and supply are equal. This is shown by the point Q^* in the figure, where $Q_D = Q_S = Q^*$ and P equals F. In the presence of insurance, however, P' and F' are unequal and thus standard price equilibration

cannot take place. For example, patients might face generous cost sharing and so demand $Q'_D > Q^*$. Physicians might not receive a high fee for each service and thus are only willing to supply $Q'_S < Q^*$.

We need to add to the story to describe equilibirum. One possible equilibrium is given by the short-side principle: $Q_E = \min(Q'_D, Q'_S)$. The set of possible equilibria would correspond to the area of demand and supply to the left of Q^*. A second possible view is that physicians are (imperfect) agents for patients and patients implicitly trust their physicians. Thus, $Q_E \approx Q'_S$ provided Q'_S is not too far from Q'_D. Such a model would magnify the role of physician-specific characteristics such as fees, clinical judgment, and beliefs about appropriate care on equilibrium outcomes relative to role of patient preferences and prices.

Further, this model suggests that not all services provided should be interpreted as reflecting demand. That is, one should not necessarily assume that the area of the country where radical prostatectomy rates are 15 percent above the national average (Utah) has patients who are much more likely to believe that radical prostatectomy is the best treatment for localized prostate cancer.

Our goal in the remainder of this chapter is to understand how treatment decisions are actually made. This chapter is not a review in the traditional sense; we do not attempt a definitive summary of all the papers on the determinants of treatment choice—if one could even be conducted. Rather, we selectively summarize the theories and literature that have the most bearing on understanding treatment choices across individuals and areas. Our aim is to provoke, so we shall give our conclusion here: *None of the theories for which there is a lot of evidence can be shown to explain a major part of cross-individual or cross-area variation in treatments.* As a result, we suspect that new theory and empirical work will be needed to address this issue.

The chapter proceeds as follows. Section 2 examines the theory and evidence on the role of demand factors. Section 3 reviews the theory and evidence behind supply factors. Section 4 considers a set of situational factors and systems-level factors that have been shown to influence decision making, but have yet to be incorporated into economic models of treatment choices. Throughout, we focus on how such factors can help to understand treatment decisions.

2. HETEROGENEITY IN DEMAND

2.1. Price and Income Effects: Theory

It is straightforward to show that the demand functions above have a positive income effect and a negative own-price effect. Imagine that treatment 1 is more expensive than treatment 2—for example, treatment 1 might be a stent for stable angina, while

treatment 2 might be medical management of the same condition. Recall equation (6.3) for choosing treatment 1 from above, where utility from non-medical consumption is $V(Y-P_i)$, P_i is the price the consumer pays for the treatment, and Y is income net of insurance premiums. Let τ be the right-hand side of equation (6.3). Then $d\tau/dY$ is given by (ignoring the k subscript):

$$d\tau/dY = V'(Y-P_2) - V'(Y-P_1) \tag{6.9}$$

With diminishing marginal utility ($V'>0$ and $V''<0$); given that $P_1>P_2$, it follows that $V'(Y-P_1)<V'(Y-P_2)$, and therefore $d\tau/dY<0$. With a lower threshold for τ, the probability that $\theta_1-\theta_2>\tau$ increases, and thus higher income patients are more likely to desire treatment 1.

The effect of price on treatment choice is more complex, since one or both prices may change. In general:

$$d\tau = V'(Y-P_1)dP_1 - V'(Y-P_2)dP_2 \tag{6.10}$$

A change in a single price will have an unambiguous effect. For example, an increase in P_1 alone will increase τ (because $V'>0$) and thus make consumers less likely to desire treatment 1. The opposite is true about an increase in P_2. An equal increase in both P_1 and P_2 has an ambiguous effect on demand for treatment 1, depending on which price is greater. This is akin to an income reduction. In a general setting where one "treatment" option is to do nothing, the consumer may choose to forego care as the price of all treatments increases.

The translation of patient demands into equilibrium outcomes is complex, as noted above. In the classic moral hazard models (Pauly, 1968; Zeckhauser, 1970), this is assumed away. Implicitly in those models, the supply of medical services is competitive: fees equal marginal cost and physicians are willing to supply anything demanded. Without a financial interest in the treatment choice, the physician acts as a perfect agent for the patient and equilibrium outcomes are equal to demand.

2.2. Empirical Evidence on Price and Income Effects

The theory about how prices and income affect treatment decisions makes generally clear predictions. The empirical question asks how large these effects are. There is a lengthy empirical literature on the income and price elasticities for medical care, beginning in the 1960s. The difficulty in estimating demand elasticities in observational data is that moral hazard and adverse selection both imply that better insurance is associated with use of more health care. Moral hazard posits that people exogenously assigned to lower cost sharing plans will use more services. Adverse selection denotes the tendency of people who expect they will use more medical care to choose insurance plans with lower cost sharing. Just observing that low cost sharing in insurance is

correlated with high spending does not differentiate between the moral hazard and adverse selection explanations.

The importance of distinguishing moral hazard from adverse selection warranted a randomized experiment. The RAND Health Insurance Experiment (HIE), begun in the 1970s and lasting three to five years, was a controlled trial to empirically examine moral hazard. The HIE randomized approximately 2,750 families into plans with different levels of cost sharing: from essentially no cost sharing to roughly a catastrophic policy (albeit with only a moderate out-of-pocket deductible). Spending was recorded and related to the exogenous plan assignment. There were three important results from the HIE (Manning et al., 1987; Newhouse et al., 1993). First, spending is related to out-of-pocket price; the overall demand elasticity is about −0.2. This effect is statistically significant but modest in magnitude. Second, there is a small income elasticity for care, conditional on being insured; the elasticity is about 0.1. This elasticity is much smaller than the elasticity one gets by looking at area-level spending or cross-country evidence, likely reflecting the endogeneity of medical technology at the national level (Newhouse, 1992). For our purposes, the elasticity holding technology constant is the value of interest. Third, cost sharing affects whether a person gets into the system, but not what happens once a person is in the system. To take an example, cost sharing might influence whether a person with chest pain sees a cardiologist, but not what services the cardiologist performs once care has been initiated. This latter finding contradicts the predictions of the simple demand model, and thus demonstrates that the supply side is important as well.

At this point, the HIE is nearly 30 years old. Because the health care system has changed dramatically—there were few expensive prescription drugs when the HIE was conducted, for example—it is possible that the elasticities uncovered in the HIE are no longer applicable. In fact, research on the demand elasticity for care has continued since the RAND study. Researchers have pursued a variety of natural experiments. For example, some firms increase cost sharing in a year and other firms do not. If one believes these cost-sharing changes are orthogonal to other reasons why people might seek care, comparing patterns of care received between employees of firms that increased cost sharing and employees of firms that did not would allow us to learn about the demand elasticity for care. For example, three recent papers on the role of prices have found "offset" effects, a cross-price elasticity of demand. Chandra et al. (2010) found that while own-price elasticities for physician visits and prescription drug usage in an elderly population were similar to those of the RAND HIE, savings from increased cost sharing were offset by increased costs to Medicare for increased subsequent hospital care. Hsu et al. (2006) also found that savings derived from capping drug benefits were offset by increased spending on hospitalizations and emergency department visits. Similarly, Trivedi et al. (2010) found that increased cost sharing for ambulatory care among Medicare beneficiaries led to increased spending

on hospital care. We provide several observations about this literature below; McGuire (2011, this Handbook) discusses it as well.

By and large, the HIE results hold up well. Most of the demand elasticities from these natural experiments are along the lines of the HIE. In some areas, the RAND results have been refined with newer data. In particular, recent studies show that:

- Use of prescription drugs is very price sensitive (Hsu et al., 2006; Huskamp et al., 2003; Joyce et al., 2002). In general, the literature finds elasticities of about −0.2 to −0.6 (Goldman and Joyce, this Handbook).
- People seem to cut back on both necessary and unnecessary care. When cost sharing increases, people use fewer services, but the services foregone are neither uniformly valuable nor wasteful (Buntin et al., 2011; Chandra et al., 2010).
- Higher cost sharing deters recommended preventive and chronic care, which may lead to undesirable "offsets" in greater use and spending on other services, such as hospital care (Trivedi et al., 2010; Chandra et al., 2010; Hsu et al., 2006). The economic theory of offsets is discussed in Goldman and Philipson (2007) and Newhouse (2006).
- There are complementarities across types of care (Buntin et al., 2011). Raising costs for prescription drugs increases hospital costs, and lowering costs for preventive care has only a modest effect on utilization if people need to see their primary care physician before accessing preventive care.

Because the price and income elasticities from the HIE and more recent studies are low, these standard demand factors are unlikely to explain the variation in treatment across individuals or over space. To put it simply: Among insured people with the same condition, cost sharing differs little, yet treatments vary a good deal. Several studies provide direct evidence for this. These studies use micro data to estimate models of medical spending as a function of health status, regional supply characteristics, insurance coverage, and income. They then evaluate the importance of demand factors relative to other characteristics. Typically, this analysis is done with samples of Medicare beneficiaries. This is valuable in part because everyone in the sample has the same base level of coverage, with no quantity or other non-price restrictions on use. Price varies to the extent that people have supplemental insurance, through Medicaid or private insurers.

As noted above, regional variation is extensive in the Medicare population. People living in areas with the highest quintile of spending use, on average, 50 percent more care than people living in areas with the lowest quintile of spending, and it is tempting to believe that area income is an important factor. Sutherland et al. (2009) show that virtually none of these regional differences is accounted for by differences in area income or poverty rates. Health status differences explain about 18 percent of the difference in spending across areas, and the rest is unaccounted for. The lack of a strong income effect is also shown by McClellan and Skinner (2006), who estimate that

Medicare spending is approximately equal across deciles of zip code income. However, there is some disagreement in the literature. In the Health and Retirement Study, for example, Marshall et al. (2010) find that out-of-pocket expenditures vary by wealth, as the richest 20 percent of households spend an average of $18,232 per year, versus $7,173 for the poorest 20 percent of households. These differences were driven by nursing home and home health care.

Zuckerman et al. (2010) estimate a similar model, including supplementary insurance coverage along with income and health status. Supplemental Medicaid coverage is associated with higher spending, while supplemental private insurance coverage is associated with lower spending—in each case perhaps reflecting selection unaccounted for by the health status controls. Still, the combined impact of supplemental insurance on regional differences in spending is small. They estimate that income and supplementary insurance together explain 1 percent of the higher spending in high-cost areas compared to low-cost areas. Thus, while price and income matter for spending, they are unlikely to explain a large part of why spending differs so much across people or areas.

2.3. Preferences

Preferences for different health outcomes (e.g. length vs. quality of life) differ across people. In the case of localized prostate cancer, for example, some men prefer to live with symptom-free cancer under conservative management, while others prefer treatment given its potential associated complications (Stewart et al., 2005). While age and disease severity may explain some of this variation (e.g. selection effects), variation remains (Shappley et al., 2009). Some of the remaining variation is owing to idiosyncratic factors. For example, prostate cancer patients report wildly unstable preferences for different health states that violate basic utility rank order assumptions (Dale et al., 2011). Moreover, preferences conveyed through subjective measures of well-being are subject to a host of psychological influences that often render them unreliable as a tool for comparisons across people (Smith et al., 2006, 2008; Ubel et al., 2005). Similarly, the literature on shared decision-making demonstrates that patients have different preferences, which can lead to different optimal treatments (Barry et al., 1995; Sepucha and Mulley, 2009). One interpretation of this is that preferences for health states vary across people, and the random variation from study to study reflects this. Yet the variation can also be driven by factors outside of the patient's demand. Sommers et al. (2008b) surveyed men who have been diagnosed with localized prostate cancer and are being treated using either surgery or radiation. Preferences for the side-effects of those treatments differ across men, but the care the men received does not correlate highly with those preferences.

The literature has not evaluated, in a general sense, how much patient preferences contribute to treatment differences across individuals or areas. There is some evidence

that suggests it is small, however. For example, most patients prefer to die at home, but most actually die in a hospital (Pritchard et al., 1998). Angus et al. (2004) also found that some patients receive more intensive treatment at the end of life than they had preferred, while others received less. At the end of life, when the trade-off between length and quality of life becomes most salient, people actually have particularly strong preferences (e.g. they want their values to be respected, their symptoms to be well controlled, and their time with loved ones maximized) (Steinhauser et al., 2000). Yet intensive treatments are often delivered in place of palliative care. One explanation is poor communication of patient preferences, as patients are transferred to intensive settings and cared for by providers they have never met (Back et al., 2009).

However, another potential explanation is that physicians do not fully take into account patient preferences in clinical decision-making. Evidence suggests that traditional devices for communicating patient preferences to physicians, such as advance directives and orders to forego resuscitation, do not influence end-of-life treatments (Fagerlin and Schneider, 2004; Teno et al., 1997). Pritchard et al. (1998) found that where patients actually die has much more to do with where they live and supply-side factors than with the patient's preferences or demographic and clinical characteristics. In general, patient preferences for primary care and specialty care do not significantly explain regional variations in health care use (Anthony et al., 2009; Baicker et al., 2004; O'Hare et al., 2010).

Overall, our conjecture is that differences in preferences do not explain a large part of treatment variation—not because preferences do not differ, but because they are frequently not accounted for in actual treatment decisions. Indeed, this literature suggests that patients generally prefer less aggressive care than their physicians would recommend. Still, this is an area where more work would be quite valuable.

3. SUPPLY-SIDE DRIVERS OF CLINICAL DECISIONS

Medical care is ultimately provided or supervised by physicians. Understanding the motivation of physicians is critical. Just as the demand-side model simplified physicians to be perfect agents, supply-side models typically simplify patients to be comatose, uninformed, or unable to go elsewhere. In that setting, physicians maximize their utility, unconstrained by patient demands.

A less extreme assumption is that $B_i(\sigma)$ (a term that reflects the benefit to a patient from a particular treatment) in equations (6.6) and (6.7) reflects a constraint that patients cannot be too unhappy with the care received. For example, the threat of a patient leaving for another physician or bringing a lawsuit against the doctor may be increasing in the patient's level of unhappiness with her medical outcome. In this case,

even physicians who care only about their own income will implicitly care about the benefits of treatment to the patient. Using equations (6.6) and (6.7), the physician will employ treatment 1 if:

$$W_j(1) > W_j(2) \rightarrow \beta_j(B_1(\sigma_k) - B_2(\sigma_k)) + (V(F_1) - V(F_2)) + (\varepsilon_{j1} - \varepsilon_{j2}) > 0 \quad (6.11)$$

If ε has a stochastic component, this will be a probabilistic expression, as in the optimal decision for the patient. If ε is deterministic, this will yield a definitive cutoff.

In this model, physicians will choose one treatment over another for one of four reasons. First, the benefits to the patient may differ. Second, different physicians may place different weight on patient benefits relative to profits β_j. The extent to which this affects medical practice is determined by physician professionalism and competition in the provider market. Third, the fees (net revenue) associated with the two treatments may induce physicians to provide more of one treatment. Assuming substitution effects dominate income effects, fee increases for treatment 1 will make physicians more likely to prescribe that therapy. Finally, there may be other, non-fee factors that influence care, such as capacity constraints, defensive medicine, or the combination of guidelines and experience.

3.1. Profits versus Patients

There is some evidence about how physicians value patient benefits versus their own net revenue, although the data are limited. Campbell et al. (2007b) surveyed approximately 1,600 physicians about norms and behaviors in professional practice. Ninety-six percent of the physicians agreed with the statement that "Physicians should put the patient's welfare above the physician's financial interests." Actual practices are not this pristine, however. When given the following scenario: "You and your partners have invested in a local imaging facility near your suburban practice. When referring patients for imaging studies, would you: 1. Refer your patients to this facility? 2. Refer your patients to this facility and inform patients of your investment? 3. Refer patients to another facility?" Twenty-four percent of physicians chose scenario 1—they would refer without informing the patients of the financial interest.

Further, there is some suggestion that β_j (professionalism) may vary significantly across areas, and that this may explain a reasonable part of the regional variation in care. In Gawande's (2009) description of care in McAllen and El Paso, Texas, he notes that physicians in McAllen seem to be more entrepreneurial than physicians in El Paso. Stories of dramatic overprovision of certain intensive services, such as angioplasties in Elyria, Ohio, which performs three times more angioplasties than neighboring Cleveland or Provo, Utah, where Medicare beneficiaries receive ten times more shoulder replacements than Syracuse, New York, suggest that professionalism and training play a role (Abelson, 2006; Fisher et al., 2010).

Further, it is plausible that physicians with high or low β_j would choose to work in the same area, or that it might be spread through a community. This discussion may make it seem that there are agglomerations in professionalism—and while there certainly are, there is no reason to believe that this is purely a regional phenomenon. To the extent that physicians with low levels of professionalism are scattered throughout all systems, the scope of improving the overall productivity of health care is large. While professionalism might not be immediately apparent, it partially determines provider income, different referral practices, different decisions about care, and so on. Understanding how professionalism—the extent to which patient well-being is traded off against economic interests—is created and destroyed, and how it affects outcomes, is a high priority for future research efforts.

As mentioned in the previous section, β_j may be a function of other physician incentives. Yet even if β_j were exogenous and equal to 1, implying that the physician cares only about patients' well-being and therefore would always choose the appropriate treatment that maximizes patient benefit (B_i), "appropriateness" can still be difficult to define. In fact, appropriateness may itself be endogenous. An illustrative example is the consensus definition of "an appropriate imaging study" determined by the American College of Cardiology Foundation, American Society of Echocardiography, and related medical societies (2011): "An appropriate imaging study is one in which the expected incremental information, combined with clinical judgment, exceeds the expected negative consequences [risks of the procedure (i.e. radiation or contrast exposure), downstream impact of poor test performance such as delay in diagnosis (false-negatives), or inappropriate diagnosis (false-positives)] by a sufficiently wide margin for a specific indication that the procedure is generally considered acceptable care and a reasonable approach for the indication." Such guidelines leave appropriateness open to interpretation. They also suggest that professionalism, if it is endogenous, may be a function of classic uncertainty (Arrow, 1963). For example, physicians may exhibit greater professionalism when the trade-off between benefits and costs—the width of the "sufficiently wide margin" in the definition above—is clear and salient. In clinical scenarios with less certainty, other incentives that offer certainty in physician gains may trump professionalism. The utility maximizing physician may be more likely to choose the better-reimbursed treatment, for example, when the field of treatment options presents no clear-cut winner *a priori*.

3.2. Payment Systems and Supplier-induced Demand

A substantial literature has examined the impact of payment systems on physician behavior. For recent reviews, see McClellan (2011) and McGuire (2000). The literature is clear that providers respond to payments, and that the response can be very large. These responses need not, of course, be referred to as "supplier-induced

demand," which has a very specific meaning whereby the physician increases utilization by changing patients' desired treatments (demand).

The most salient example of the impact of payment systems on supply is the movement from fee-for-service payment to prospective payment. In 1983, the Medicare program moved from a retrospective, fee-for-service payment system for hospital care to the prospective, Diagnostic-Related Group (DRG) payment system. Thus, marginal reimbursement of additional days in the hospital, minor procedures, and tests went from greater than marginal cost to zero. The result was a large reduction in total hospital days (Coulam and Gaumer, 1991): the volume of hospital admissions dropped by 11 percent in the first eight years, while length of stay and total Medicare spending also declined sharply after 1983 (Hodgkin and McGuire, 1994; Ellis and McGuire, 1993).

Some of the inpatient days were substituted to an outpatient setting—beneficiaries who needed rehabilitation after a hip fracture, for example, were often transferred to a skilled nursing facility to receive that care, when it used to be provided in the hospital. For an integrated provider organization, this was beneficial because the skilled nursing facility could bill for rehabilitation services, whereas the hospital could not. In other cases, technology implemented about the time of prospective payment led to a reduction in inpatient days, for example the ability to do cataract surgery on an outpatient basis. Still, some inpatient days were eliminated as hospitals became more efficient. This reduction in inpatient usage was accomplished without significant adverse effects on patient health (Miller and Luft, 1994; Lurie et al., 1994; Cutler, 2004; Berwick, 1996).

Increasing services when fees are high is conceptually similar to reducing services that are underpriced. The literature gives a subset of this increase in services a special name, however: physician-induced demand. Physician-induced demand is defined as a physician providing care that a fully informed patient would not choose for himself. Note that physician-induced demand would not occur in a fully informed market—some degree of information asymmetry (as in Arrow, 1963) or other imperfection is needed to generate care off the demand curve. Similarly, some external constraint is needed to determine when physicians stop providing care, whether that is substantially lower benefits to the patient, fear of being sued, or professional norms. McGuire and Pauly (1991) provide a framework for thinking about induced demand in the context of multiple payers, in which income and substitution effects are important for determining treatment choice.

The literature has explored physician-induced demand in some detail. Early empirical studies tested the physician-induced demand explanation by looking for an "availability" effect: The hypothesis that an increase in the physician−population ratio, which should lead to a decrease in average income, will lead to more services being provided (Fuchs, 1978; Cromwell and Mitchell, 1986; see also Phelps, 1986; Feldman and Sloan, 1988). The general conclusion from these studies is that availability was positively associated with use. The difficulty with these studies is readily apparent,

however: supply may be high because demand is high. Distinguishing whether greater equilibrium quantities are provided because supply is high or demand is high is always difficult, and is virtually impossible when prices are not allowed to adjust.

In more recent years, empirical studies have taken two approaches: (1) examining exogenous demand shocks and (2) studying changes in utilization following cuts in physician payment. In each case, the literature has generally hypothesized a backward bending supply curve, along the lines of a target-income model: if physicians' fees are reduced, do physicians provide more care, as a way of supplementing their income. Notable papers in this literature include Rossiter and Wilensky (1984); Dranove and Wehner (1994); Gruber and Owings (1996); Nguyen and Derrick (1997); Yip (1998); Jacobson et al. (2010); and Rice (1983). Yip (1998), for example, finds that following Medicare fee changes for cardiac bypass surgeries in 1988, physicians whose incomes were cut the most responded by performing greater quantities of these surgeries, both in Medicare and privately insured patients. Similarly, following the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 which reduced payments for certain cancer drugs, physicians responded by raising the use of those drugs and substituting towards higher profit drugs (Jacobson et al., 2010). While none of these studies is entirely convincing, the belief in physician-induced demand has played a role in setting Medicare physician payment rates and influenced national policies concerning the physician supply (Reinhardt, 1999).

The role of fee differences in explaining different treatments across patients or areas is unknown. In the case of well-reimbursed patients versus poorly reimbursed patients (either uninsured patients or patients on Medicaid), the effect is clear: many physicians will not even see patients who have Medicaid or no insurance. For patients who are well insured, however, the differences are more difficult to discern. In the Medicare population, for example, fees are similar nationally, yet care differs across patients and regions. We suspect, in this circumstance, that it is not the absolute value of the fee that matters, but rather the relative weight that physicians place upon income versus patient benefits—the β_j term above.

3.3. Physician Specialization and Training

Physician specialty is an example of ε—a factor that may influence treatment separate from patient benefits and fees. It is clear that the specialty of the physician matters for treatment choice. Sommers et al. (2008b) show that when a man with localized prostate cancer sees an oncologist, he is more likely to be treated with chemotherapy, whereas if he sees a surgeon, he is more likely to receive a prostatectomy. Indeed, the physician that a patient sees matters far more for treatment than the patients' preferences (Sirovich et al., 2008). Yet it is also clear that the physician's specialty choice is endogenous. For example, physician wages across specialties vary enormously in the

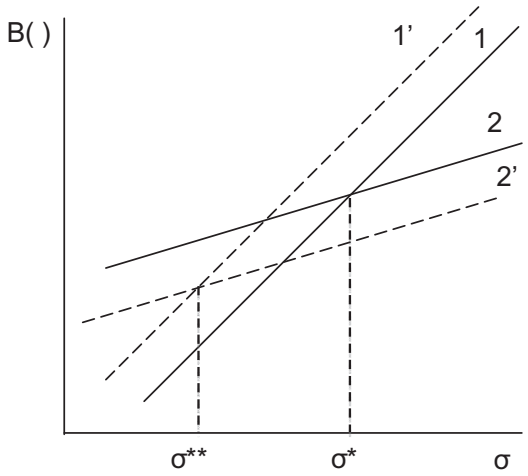

Figure 6.4 Productivity spillovers in the Roy model.

United States, from a mean hourly wage of $132 for neurologic surgery to $50 for primary care (Leigh et al., 2010). A physician-in-training who prefers to provide intensive treatments may specialize in an intensive field, holding constant other factors. Opportunities for specialization in medicine have blossomed in recent years (Cassel and Reuben, 2011).

While it is possible that the effect of physician specialty on utilization reflects physician-induced demand, more benign interpretations are also plausible. For example, it may be that physicians believe what they do is the best course of therapy, because they are trained that their expertise is important and because they experience the success of what they do on successful interventions. They may then translate this into their future recommendations. Alternatively, it may be that specialized physicians who work with uncommon diseases or sicker patients approach each patient as a unique situation and worry more about the marginal benefits specific to that patient, whereas generalists approach patients as part of a population and assign more weight to average benefits (Redelmeier and Tversky, 1990).

Indeed, specialization could matter in another way as well, by changing the true relative benefits of different forms of therapy. Chandra and Staiger (2007) posit increasing returns (positive spillovers) to specialization: in areas that do a lot of one therapy, physicians get better at it and thus render that therapy more optimal over time, even for patients who are less appropriate candidates for that therapy. Figure 6.4 shows an example of this specialization. The two lines denote the benefit from treatment 1 and 2 for patients with different disease severity (σ). Applied to the prostate cancer setting, treatment 1 may denote intensive treatment, which is more appropriate

for younger and sicker men, while treatment 2 may represent active surveillance, more appropriate for older and healthier men. Empirical treatment patterns in the United States largely follow these age and severity dichotomies (Shappley et al., 2009). If treatment were allocated with the goal of maximizing patient outcomes, everyone to the left of σ* would receive treatment 2. If there are benefits to specialization, then the benefit from treatment 1 also depends on the proportion of patients receiving treatment 1, P_1 (analogously, the benefit from treatment 2 depends on 1 minus the fraction of patients receiving treatment 1).

$$\mathbf{U_k}(1) = \alpha_1 P_1 + B_1(\sigma_k) + V(Y_{k1} - P_{k1}) + \theta_{k1} \quad (1')$$

$$\mathbf{U_k}(2) = \alpha_2(1 - P_1) + B_2(\sigma_k) + V(Y_k - P_{k2}) + \theta_{k2} \quad (2')$$

The curves 1′ and 2′ illustrate several implications of this specialization, which produce a new equilibrium, σ**. Patients who are more appropriate for treatment 1 are better off getting it in areas where providers specialize in treatment 1. Patients more appropriate for treatment 2 are worse off in such areas. Any patient who gets treatment 1 in specialized areas would receive a larger benefit relative to getting it in non-specialized areas. Furthermore, the marginal patient receiving treatment 1 in specialized areas will be less appropriate for treatment 1 than the average patient receiving treatment 1.

Chandra and Staiger (2007) argue that this model explains some of the variation in rates and outcomes for heart attacks: areas that do surgery a good deal have better surgical outcomes than areas that are more prone to provide medical management, even given the extent of disease and patient characteristics. The negative and positive spillovers in this model imply that getting better at one type of care necessitates getting worse at the alternative. For example, if physicians have a time budget with which they trade off practicing treatment 1 against practicing treatment 2, and practice matters for outcomes, either choice would produce positive spillovers for the chosen treatment and negative spillovers for the other. In other words, physicians specialize in one skill at the expense of other skills. This simultaneous learning and forgetting of competing treatments may alternatively manifest in a world with selective migration of inputs.

In the clinical literature, a number of studies have uncovered an inverse relationship between procedural volume and adverse outcomes such as mortality, which has been termed the volume–outcome relationship (Birkmeyer et al., 2002; Dudley et al., 2000; van Heek et al., 2005). In this type of learning by doing, the volumes a physician provides of type 1 and type 2 procedures affect patient outcomes (here, N_1 and N_2 are the number of patients receiving each procedure):

$$\mathbf{U_k}(1) = \alpha_1 N_1 + B_1(\sigma_k) + V(Y_k - P_{k1}) + \theta_{k1} \quad (1'')$$

$$\mathbf{U_k}(2) = \alpha_2 N_2 + B_2(\sigma_k) + V(Y_k - P_{k2}) + \theta_{k2} \quad (2'')$$

Chandra and Staiger did not find evidence for this type of specialization in the context of heart attack treatments, but this theory has not been tested rigorously across different types of care, and doing so is a potentially fruitful area of research. Moreover, recent clinical work has suggested that volume may not be as strongly predictive of mortality as previously thought (Finks et al., 2011; Kozower and Stukenborg, 2011). The specialization theory provides a natural explanation for area variation in care, if not individual variation within an area.

Furthermore, independent of specialization, the physician's institution or location of training may matter—another example of ε. Anecdotal evidence in medicine suggests that where providers train heavily influences their treatment styles. The motto for medical training is "see one, do one, teach one." Training programs, including residency and fellowship, place the budding physician under the tutelage of experienced clinicians in a learning process resembling an apprenticeship. A trainee may often encounter the opinion that for the same procedure or the same protocol, one hospital does it one way while another does it another way. Among cardiologists given patient vignettes, whether their colleagues would have ordered a cardiac catheterization in the same situation predicts whether respondents ordered a catheterization (Lucas et al., 2010). Institutional effects owing to training programs undoubtedly fall into the poorly defined black box of within-firm decision-making determinants, which we further note below (section 4.5). There is some evidence on the differential outcomes of patients treated by physicians from highly ranked medical institutions compared to less highly ranked institutions (Doyle et al., 2010). However, evidence from Cesarean sections in Florida shows that while physicians do learn from other physicians, residency programs explain less than 4 percent of the variation in rates of operations (Epstein and Nicholson, 2009). Physicians did not seem to update their prior beliefs, even newly trained physicians, producing a within-area variation that approximately doubled between-area variation.

3.4. Defensive Medicine

Fear of being sued may affect the care that is provided (Lucas et al., 2010). The financial costs of malpractice associated with providing a particular service could be captured in F_i above. But being sued is rarely about the financial cost. Physicians are insured for malpractice expenses, and such expenses are independent of any specific treatment that is provided—though they depend on the specialty of the practice. Rather, the major costs of being sued are the time and energy involved in defending a lawsuit and the psychological costs of having one's professional judgment questioned. These show up in the X_i terms.

Responses to malpractice concerns may involve providing too much care, or too little care. The former is termed defensive medicine; the latter is stinting, or negative defensive medicine. An example of defensive medicine is doing imaging for a routine

injury just to be sure that there is nothing seriously wrong. An example of stinting is obstetricians avoiding high-risk women, for fear that the baby will be impaired and they will be blamed. Negative defensive medicine would arise in a malpractice environment where adverse events such as injuries receive compensation even though malpractice had not occurred.

The combined effect of defensive medicine and stinting has been estimated in several studies (Brennan et al., 2004; Localio et al., 1991; Mello et al., 2010). These estimates are not without difficulty: measuring malpractice pressure in an area is difficult, and finding an exogenous measure of that is harder still. Even so, studies have surmounted this problem using area variation in malpractice premiums or other measures of malpractice pressure such as the size of indemnity payments. The results show a surprisingly small net contribution of malpractice concerns to what physicians do. Mello et al. (2010) estimate that medical malpractice and efforts to manage its risks cost the national health care system more than $55 billion a year, about 2.4 percent of annual health care spending. The study summed various components of the medical liability system, including payments made to malpractice plaintiffs; defensive medicine costs; administrative costs, such as lawyer fees; and the costs of lost clinician work time. Defensive medicine costs were the largest segment of total malpractice spending and amounted to approximately $45 billion a year annually. In other work, Baicker et al. (2007) note that malpractice pressure increases the use of imaging procedures but exerts a small overall effect on total spending, perhaps because of the presence of negative defensive medicine.

Almost all of the research quantifying the scope of defensive medicine relies on cross-sectional or panel variation in the relationship between medical spending and malpractice pressure. But this design may understate the amount of defensive medicine if physicians in one state are influenced by the malpractice experience of their friends and colleagues in other states. We have little guidance on how physicians assess the threat of litigation and understanding this channel would be a fruitful area for future work.

4. SITUATIONAL FACTORS

This section is about the error term determining treatment choice. Despite knowing the contributions of income, geography, health status, reimbursement, and other factors in treatment variations, there remains substantial variation in treatment choices. What is the source of this variation? We suspect it may derive from influences specific to the clinical situation in which the treatment determination is made. Such influences would do less to explain regional variation as they are largely orthogonal to geography. In particular, we suspect that physicians and patients follow a number of

psychological heuristics, or rules of thumb, which affect clinical decision making in similar ways across regions.

A growing literature in behavioral economics offers a starting point for unpacking these behavioral influences. Frank and Zeckhauser (2007) theorize that physicians deploy heuristics-based ("ready-to-wear") treatments because they minimize a number of behavioral costs relative to patient-specific ("custom-made") treatments. These costs include the cost of patient–doctor communication regarding treatment choices, cognition costs of coming up with custom-made treatments, coordination costs between providers in the increasingly specialized medical profession, and system-level capability or resource costs. Broadly using this framework as a guide, we survey a number of specific behavioral influences below.

4.1. Availability Heuristic

Determining the optimal treatment path among a list of possibilities may require effort and time, which are not costless to the physician. As a result, physicians resort to the *availability heuristic* to make decisions (Groopman, 2007). The availability heuristic refers to the idea that people predict the frequency of an event by the ease with which it comes to mind. Airplane crashes are salient; stomach cancer may be less salient. Thus, people may believe that airplane crashes are a more common cause of death than stomach cancer, even though death from stomach cancer is actually five times more common (Lichtenstein et al., 1978). Since the occurrence of salient events is stochastic, the availability heuristic naturally introduces randomness into treatment choice.

In diagnosis, the availability heuristic suggests that a physician who has just seen a patient with influenza may be more likely to make the diagnosis of influenza for the next patient who walks through the door with a cough, even if this latter patient has a rare lung disease. The reverse may be true if there was a recent case of lung cancer.

Choudhry et al. (2006) found some empirical evidence for this. They showed that physicians who treated a patient with warfarin (a blood thinning drug with the risk of bleeding) and saw that patient experience an adverse bleeding event were 21 percent less likely to prescribe warfarin to other patients for which warfarin is indicated, even 90 days after the adverse bleeding event. The availability heuristic gives the physician a decision-making shortcut that allows the physician to recall and reuse information from the most recent salient cases.

An important note about the availability heuristic is that it is applicable in the absence of easily obtainable information on true risks. Were patients and physicians presented systematic odds about the likelihood of every diagnosis and the appropriate diagnostic tests given those likelihoods, the availability heuristic would be much less operative.

4.2. Framing, Choice, and Risk

Framing and choice architecture affect decision making (Thaler and Sunstein, 2008; Botti and Iyengar, 2006; Ubel, 2009). For example, if a physician faces a growing number of treatment choices, all of which are plausible for a patient, the task of choosing one becomes increasingly difficult. In the classic example, supermarket shoppers presented with six different flavors of jelly were able to choose among them, but shoppers presented with 30 brands found the choice too complex and were one-tenth as likely to make a purchase (Iyengar and Lepper, 2000).

Evidence shows that both health care professionals and patients are susceptible to the framing of choice in clinical decisions (Akl et al., 2011; Sommers and Zeckhauser, 2008a; Redelmeier and Tversky, 1990; McNeil et al., 1982). In a systematic review of 35 studies, both providers and patients understood natural frequencies better than probabilities in the presentation of risk (Akl et al., 2011). For example, a 50 percent risk reduction was perceived to be substantially larger than an absolute risk reduction from 2 to 1 percent. Statistics presented as number needed to treat were least persuasive, such as 100 people treated to prevent one case of the disease. Perhaps because of this, there is widespread overly aggressive screening and treatment in prostate cancer (Drazer et al., 2011; Schröder et al., 2009).

Indeed, classic risk aversion may drive decision making for both patients and physicians. Moreover, the magnitude of risk aversion may depend on the clinical scenario, implying that the classic Arrow−Pratt coefficient of risk aversion may not be exogenous at the individual level, but may be a function of the situation. In the simplest example, a clinician's experience dealing with a particular situation may affect risk aversion, as younger physicians often order more procedures than more experienced physicians (Woo et al., 1985). We can also think about how risk aversion would influence the theoretical model set forth in sections 2 and 3. On the physician side, risk aversion can take (at least) two forms. First, physicians may avoid prescribing a treatment that is high risk for the fear of harming the patient (uncertainty in $B_i(\sigma_k)$). Second, physicians may avoid such a treatment for the fear of malpractice litigation (uncertainty in $V(F_i)$). Importantly, such a treatment can be the "do nothing" option, especially in a world in which physicians and patients are loss averse. Depending on the clinical scenario, these two forces may combine to push the physician towards prescribing the state-of-the-art laparoscopic surgery, for example, which lowers blood loss and length of stay. Or these two forces may antagonize each other in cases where low risk for patient benefit $B(\ )$ is high risk for the physician's income $V(\ )$, such as watchful waiting in the case of prostate cancer. On the demand side, patients may similarly be risk averse in the level of uncertainty about $B(\ )$.

In our simple model, increased utilization of certain services may arise when both parties express risk aversion or when either party does. Suppose physicians are fully

benevolent ($\beta = 1$), whereby patient preferences dictate the decisions. Suppose also that the patient is risk averse. In this situation, if patients push for the latest and greatest technologies, that is enough to lead to utilization and spending on those treatment options.

4.3. Status Quo and Confirmation Bias

Status quo bias refers to the fact that people tend not to change their behavior unless the incentive to change is strong. In the treatment decision setting, aversion to change can lead to status quo bias. For physicians, changing treatment choices, even the mere process of discussing alternative options with patients, can be very costly (Groopman, 2006). Human beings are also naturally loss averse, so the risk and uncertainty imposed by changing to a new treatment may similarly bias physicians towards familiar paths. For example, physicians may have a favorite drug within a class of equivalent substitutes. The source of this preference may be orthogonal to anything we have considered thus far. Yet the consequence of it is a bias towards this drug.

Patients are also prone to status quo bias, and patients' inertia to change is prevalent outside of the treatment decision. For example, more people stay in their insurance plan from year to year than would be expected to do so, given the value of different policies and the price of those policies. In fact, people often stay in insurance plans dominated in every state of the world, but it is the plan they chose last year (Kahneman et al., 1991).

This status quo bias may be further driven by a self-reinforcing confirmation bias in which prior successes from one treatment induce the provider to continue using that treatment. Confirmation bias describes the idea that people interpret new information in light of their pre-existing biases (Myers and Lamm, 1976; Halpern, 1987). Specialization may be a source of this confirmation bias (Sommers et al., 2008b). A specialist that has experienced success with a particular treatment—for example, radical prostatectomy for patients with localized prostate cancer—will judge that treatment to be effective in a wider range of patients, likely even wider than objective literature would indicate. Whether this judgment is substantiated by any positive spillovers in the learning-by-doing sense, however, is an open question.

4.4. Channel Factors

Channel factors may also play an important role in determining treatment choice. A channel factor is something that affects perceived benefits and costs towards pursuing a particular choice. In a classic example, college seniors given a positive message about tetanus inoculation changed their beliefs about the value of a tetanus inoculation, but few got inoculated. Only when students were also given a map of the campus with the infirmary circled and asked to choose a particular time for inoculation (with no

consequence for missing it) did they get inoculated (Bertrand et al., 2004). Channel factors, including a map, a predetermined time, or physical proximity, seem to trump information alone.

There are also unique channel factors in a variety of clinical scenarios. Drug adherence, for example, is a particularly intriguing area. Only about 50 percent of patients adhere to their prescribed drug regimens. Economists often look to the role of cost sharing in determining such variation (Shrank et al., 2010). But even when drugs come with no cost sharing, adherence is not close to 100 percent. One example of a relevant channel factor here is route of medication intake. Consider the drugs Vancomycin and Linezolid, both treatments for methicillin-resistant *Staphylococcus aureus* (an antibiotic-resistant bacterial infection often picked up during hospitalizations). Linezolid is available as an oral medication, facilitating adherence, while Vancomycin is not. While both are equally effective in survival benefit and readmission rates, Linezolid had a lower rate of therapy discontinuation (Caffrey et al., 2010).

4.5. Resource/Capability Factors

To this point, we have modeled the treatment decision as being made by the patient, the individual doctor, or a combination of the two. That may not be right, however. Most providers work in organizations, and the organization itself, or possibly the insurer, may influence the care that is provided. Suppose that the provider is instead a hospital firm that produces medical care. How does the firm internally decide what to produce (or allow) for each type of patient (consumer)? Inside the firm are doctors, each of whom may have his or her own intrinsic views about utilization, and decision-making structure (Snail and Robinson, 1998; Harris, 1977).

The internal decision-making processes of firms have remained largely a black box to economists (Hart and Holmstrom, 2010; Williamson, 2010; Gibbons, 2005; Holmstrom and Roberts, 1998; Gibbons, 1998). Inside the firm, employment contracts, agency contracts, and relational contracts may all affect the production process and treatment choice of the provider. The difficulty of specifying these contracts precisely makes the analysis of treatment choices by health care providers especially difficult—especially given the unique agency and moral hazard complexities in health care. However, there are identifiable components of the provider organization (either the hospital firm or the physician firm) that may influence treatment choice.

Treatment choices may also be influenced by ancillary relationships between the provider and institutions or entities outside of the provider–patient relationship. Pharmaceutical manufacturers and medical device companies spend billions of dollars each year marketing products directly to physicians (see Scott Morton and Kyle, 2011, this Handbook). There are overt conflicts of interests which mechanically influence treatment, where physicians may receive kickbacks for providing a certain drug,

although there are federal statutes against such relationships. There are informal relationships where physicians receive fees and vacations in exchange for speaking engagements and informal product endorsements. Then there are psychological influences of reciprocity, where even free samples and free lunch offered by industry sales representatives with no strings attached have been shown to change behavior through intrinsic reciprocity (Fehr and Gächter, 2000).

In 2007, Campbell and colleagues published a national survey of 459 medical school department chairs at 140 institutions, which found that 67 percent of departments and 60 percent of department chairs have relationships with industry. In this study, non-clinical departments were more likely to enter into relationships with industry compared to clinical departments, which the authors attribute to the likely greater degree of licensing and product development activities associated with non-clinical departments. Clinical departments, on the other hand, were far more likely to receive discretionary funding to purchase equipment and support research seminars, graduate medical education, and continuing medical education (Campbell et al., 2007a). A study of patients found that 90 percent of participants in cancer-research trials expressed little worry about financial ties between institutions or researchers and drug companies (Hampson et al., 2006).

Systems-level health care resources matter. Not all resources are available in all settings. Resource-poor settings, for example, rarely have the ability to purchase a state-of-the-art 64-slice CT scanner, whereas an urban academic medical center often can. Consumers in some places may demand more new technology than people in other places, thereby creating a more robust market for technology in places with high demand. Direct-to-consumer advertising from drug or device manufacturers can also induce such a market and catalyze diffusion (Law et al., 2008).

A literature in sociology and management shows that technological innovations diffuse at different rates in different places (Rogers, 1983). The dynamics of diffusion are described by Thomas Schelling among others, who wrote about the tipping point theory (Schelling, 2006). In later work, scholars have expanded upon Schelling's intuition to model and empirically investigate diffusion of innovations (Young, 1998). Differences in adoption are both large and persistent. In a well-documented example, the state-level adoption rate of beta-blockers in the 1980s—a technology to reduce recurrent heart attacks in people who have had a first—was highly correlated with the state-level adoption rate of hybrid corn in the 1950s (Skinner and Staiger, 2005).

Diffusion also has long-run implications for productivity in health care. Skinner and Staiger (2009) find that even small differences in the likelihood of adoption can lead to large differences in productivity across hospitals. Specifically, they find that in the case of heart attack treatments, the speed of diffusion of low-cost technologies such as beta-blockers, aspirin, and primary reperfusion explain the bulk of long-run variations in productivity, overwhelming the impact of traditional factor inputs.

Relative to high-diffusion hospitals, survival rates in low-diffusion hospitals can lag by a decade. Adoption, or uptake, of medical technology can vary for many reasons (Chandra and Skinner, 2011). The structure of provider organizations, the culture of physicians, market power, physician specialization, and the payment system can all affect uptake. Differential adoption is particularly important given the large role that medical technology plays in treatment changes over time (Chernew and Newhouse, 2011, this Handbook).

5. CONCLUSION

The geographic variations literature demonstrates that despite nationally standardized training, physicians in different geographic locations can make markedly different treatment decisions for patients with similar clinical profiles. This was the first evidence to suggest that providers of medical care may not, as a profession, systematically choose the clinically optimal treatment path for patients. Where early literature in health economics mostly focused on the possibility of physician-induced demand—which concerned quantity more so than choice—we attempt to decompose potential explanations for variations in treatment choice.

In general, the literature points to the importance of supply-side incentives over demand-side factors in driving treatment choice. Our views are largely consistent with this paradigm. Yet traditional demand factors such as preferences and patient characteristics are undoubtedly still relevant, and behavioral influences are only beginning to be understood. In the end, one of the ultimate goals of understanding how treatment choices are made is to inform policies that move physician and patient choices towards the social optimum, even as our framework for thinking about treatment choice still rests largely on models of local (i.e. individual) optimum.

The challenges are numerous, owing to the peculiarities of medicine. In this world where the very definition of optimality rests on numerous endogenous determinants, where the supplier who judges relative merits of treatments is the same supplier who directly benefits (or loses) from the decision, and where the appropriateness of patients for any given treatment is, at best, open to interpretation and, at worst, completely unknown, understanding the determinants of treatment choice will be an ever-evolving task.

REFERENCES

Abelson, R. (2006). Heart procedure is off the charts in an Ohio City. *New York Times*, August 18.

Akl, E. A., Oxman, A. D., Herrin, J., Vist, G. E., Terrenato, I., Sperati, F., et al. (2011). Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database of Systematic Reviews*, *3*, CD006776.

American College of Cardiology Foundation, Appropriate Use Criteria Task Force, et al. (2011). Appropriate use criteria for echocardiography. *Journal of the American College of Cardiology, 57*(9), 1126–1166.

American Joint Committee on Cancer (2010). *AJCC cancer staging manual* (7th ed.). New York: Springer, Inc.

Angus, D. C., Barnato, A. E., Linde-Swirble, W. T., et al. (2004). Use of intensive care at the end of life in the United States: An epidemiologic study. *Critical Care Medicine*, *32*, 638–643.

Anthony, D. L., Brooke Herndon, M., Gallagher, P. M., Barnato, A. E., Bynum, J. P. W., Gottlieb, D. J., et al. (2009). How much do patients' preferences contribute to resource use? *Health Affairs*, *28*(3), 864–873.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review, 53*, 941–973.

Atkins, D., Best, D., Briss, P. A., Eccles, M., Falck-Ytter, Y., Flottorp, S., et al. (2004). Grading quality of evidence and strength of recommendations. *BMJ*, *328*(7454), 1490.

Back, A. L., Young, J. P., McCown, E., et al. (2009). Abandonment at the end of life from patient, caregiver, nurse, and physician perspectives: Loss of continuity and lack of closure. *Archives of Internal Medicine*, *169*(5), 474–479.

Baicker, K. & Chandra, A. (2004). Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Affairs (Millwood)*, Suppl. Web Exclusives: W4-184-97.

Baicker, K., Chandra, A., Skinner, J. S., & Wennberg, J. E. (2004). Who you are and where you live: How race and geography affect the treatment of Medicare beneficiaries. *Health Affairs (Millwood)*, Suppl. Variation: VAR33–44.

Baicker, K., Fisher, E. S., & Chandra, A. (2007). Malpractice liability costs and the practice of medicine in the Medicare program. *Health Affairs*, *26*(3, May–June), 841–852.

Barry, M. J., Fowler, F. J., Jr., Mulley, A. G., Jr., Henderson, J. V., Jr., & Wennberg, J. E. (1995). Patient reactions to a program designed to facilitate patient participation in treatment decisions for benign prostatic hyperplasia. *Medical Care*, *33*(8), 771–782.

Bertrand, M., Mullainathan, S., & Shafir, E. (2004). A behavioral-economics view of poverty. *American Economic Review, 94*(2), 419–423.

Berwick, D. M. (1996). Quality of health care. Part 5: Payment by capitation and the quality of care. *New England Journal of Medicine*, *335*(16), 1227–1231.

Bill-Axelson, A., Holmberg, L., Ruutu, M., Häggman, M., Andersson, S. O., Bratell, S., et al. (2005). Radical prostatectomy versus watchful waiting in early prostate cancer. *New England Journal of Medicine*, *352*(19), 1977–1984.

Birkmeyer, J. D., Siewers, A. E., Finlayson, E. V., et al. (2002). Hospital volume and surgical mortality in the United States. *New England Journal of Medicine*, *346*, 1128–1137.

Boden, W. E., O'Rourke, R. A., Teo, K. K., et al. (2007). Optimal medical therapy with or without PCI for stable coronary disease. *New England Journal of Medicine*, *356*(15), 1503–1516.

Bolla, M., Gonzalez, D., Warde, P., Dubois, J. B., Mirimanoff, R. O., Storme, G., et al. (1997). Improved survival in patients with locally advanced prostate cancer treated with radiotherapy and goserelin. *New England Journal of Medicine*, *337*(5), 295–300.

Botti, S. & Iyengar, S. S. (2006). The dark side of choice: When choice impairs social welfare. *Journal of Public Policy and Marketing*, *25*(1), 24–38.

Brennan, T. A., Leape, L. L., Laird, N. M., Hebert, L., Localio, A. R., Lawthers, A. G., et al. (2004). Incidence of adverse events and negligence in hospitalized patients: Results of the Harvard Medical Practice Study I. 1991. *Quality & Safety in Health Care*, *13*(2), 145–151.

Buntin, M. B., Haviland, A. M., McDevitt, R., & Sood, N. (2011). Healthcare spending and preventive care in high-deductible and consumer-directed health plans. *American Journal of Management Care*, *17*(3), 222–230.

Caffrey, A. R., Quilliam, B. J., & LaPlante, K. L. (2010). Comparative effectiveness of linezolid and vancomycin among a national cohort of patients infected with methicillin-resistant Staphylococcus aureus. *Antimicrobial Agents Chemotherapy, 54*(10), 4394–4400.

Campbell, E. G., Regan, S., Gruen, R. L., Ferris, T. G., Rao, S. R., Cleary, P. D., et al. (2007b). Professionalism in medicine: Results of a national survey of physicians. *Annals of Internal Medicine*, *147*(11), 795–802.

Campbell, E. G., Weissman, J. S., Ehringhaus, S., Rao, S. R., Moy, B., Feibelmann, S., et al. (2007a). Institutional academic industry relationships. *JAMA*, *298*(15), 1779–1786.

Cassel, C. K. & Reuben, D. B. (2011). Specialization, subspecialization, and subsubspecialization in internal medicine. *New England Journal of Medicine*, *364*(12), 1169–1173.

Chandra, A. & Skinner, J. S. (2011). Technology growth and expenditure growth in health care. NBER Working Paper 16953.

Chandra, A. & Staiger, D. O. (2007). Productivity spillovers in health care: Evidence from the treatment of heart attacks. *Journal of Political Economy*, *115*(1), 103–140.

Chandra, A., Gruber, J., & McKnight, R. (2010). Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review*, *100*(1), 193–213.

Choudhry, N. K., Anderson, G. M., Laupacis, A., Ross-Degnan, D., Normand, S. L., & Soumerai, S. B. (2006). Impact of adverse events on prescribing warfarin in patients with atrial fibrillation: Matched pair analysis. *BMJ*, *332*(7534), 141–145.

Cooper, R. A. (2004). Weighing the evidence for expanding physician supply. *Annals of Internal Medicine*, *141*(9), 705–714.

Cooperberg, M. R., Broering, J. M., Kantoff, P. W., & Carroll, P. R. (2007). Contemporary trends in low risk prostate cancer: Risk assessment and treatment. *Journal of Urology*, *178*(3 Pt 2), S14–S19.

Coulam, R. F. & Gaumer, G. L. (1991). Medicare's prospective payment system: A critical appraisal. *Health Care Financing Review*, Annual Supplement, 45–77.

Cromwell, J. & Mitchell, J. B. (1986). Physician-induced demand for surgery. *Journal of Health Economics*, *5*, 293–313.

Cutler, D. M. (2004). *Your money or your life: Strong medicine for America's healthcare system*. Oxford University Press.

Dahm, P., Yeung, L. L., Chang, S. S., & Cookson, M. S. (2008). A critical review of clinical practice guidelines for the management of clinically localized prostate cancer. *Journal of Urology*, *180*(2), 451–459.

Dale, W., Bilir, S. P., Hemmerich, J., Basu, A., Elstein, A., & Meltzer, D. (2011). The prevalence, correlates, and impact of logically inconsistent preferences in utility assessments for joint health states in prostate cancer. *Medical Care*, *49*(1), 59–66.

Doyle, J. J., Jr., Ewer, S. M., & Wagner, T. H. (2010). Returns to physician human capital: Evidence from patients randomized to physician teams. *Journal of Health Economics*, *29*(6), 866–882.

Dranove, D. & Wehner, P. (1994). Physician-induced demand for childbirths. *Journal of Health Economics*, *13*, 61–73.

Drazer, M. W., Huo, D., Schonberg, M. A., Razmaria, A., & Eggener, S. E. (2011). Population-based patterns and predictors of prostate-specific antigen screening among older men in the United States. *Journal of Clinical Oncology*, *29*(13), 1736–1743.

Dudley, R. A., Johansen, K. L., Brand, R., Rennie, D. J., & Milstein, A. (2000). Selective referral to high-volume hospitals: Estimating potentially avoidable deaths. *JAMA*, *283*, 1159–1166.

Ellis, R. P. & McGuire, T. G. (1986). Provider behavior under prospective reimbursement. *Journal of Health Economics*, *5*, 129–151.

Ellis, R. P. & McGuire, T. G. (1993). Supply-side and demand-side cost sharing in health care. *Journal of Economic Perspectives*, *7*(4), 135–151.

Epstein, A. J. & Nicholson, S. (2009). The formation and evolution of physician treatment styles: An application to cesarean sections. *Journal of Health Economics*, *28*(6), 1126–1240.

Fagerlin, A. & Schneider, C. E. (2004). Enough: The failure of the living will. *Hastings Center Report*, 34, 30–42.

Fehr, E. & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, *14*(3), 159–181.

Feldman, R. & Sloan, F. (1988). Competition among physicians, revisited. *Journal of Health Politics, Policy and Law*, *13*, 239–261.

Finks, J. F., Osborne, N. H., & Birkmeyer, J. D. (2011). Trends in hospital volume and operative mortality for high-risk surgery. *New England Journal of Medicine*, *364*(22), 2128–2137.

Fisher, E. S., Bell, J., Tomek, I. M., Esty, A. R., & Goodman, D. C. (2010). Trends and regional variation in hip, knee, and shoulder replacement. *The Dartmouth Institute for Health Policy and Clinical Practice*, April 6.

Fisher, E. S., Wennberg, D. E., Stukel, T. A., Gottlieb, D. J., Lucas, F. L., & Pinder, E. L. (2003a). The 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, E. S., Wennberg, D. E., Stukel, T. A., Gottlieb, D. J., Lucas, F. L., & Pinder, E. L. (2003b). The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine*, *138*(4), 288–298.

Frank, R. G. & Zeckhauser, R. J. (2007). Custom-made versus ready-to-wear treatments: Behavioral propensities in physicians' choices. *Journal of Health Economics*, *26*(6), 1101–1127.

Fuchs, V. R. (1978). The supply of surgeons and the demand for operations. *Journal of Human Resources*, *XIII*, 35–56.

Gawande, A. (2009). The cost conundrum: What a Texas town can teach us all about health care. *The New Yorker*, June 1.

Gibbons, R. (1998). Incentives in organizations. *Journal of Economic Perspectives*, *12*(4), 115–132.

Gibbons, R. (2005). Four formal(izable) theories of the firm? *Journal of Economic Behavior & Organization*, *58*, 200–245.

Glover, J. A. (1938). The incidence of tonsillectomy in children. *Proceeding of the Royal Society of Medicine*, *31*, 1219–1236.

Goldman, D. & Philipson, T. J. (2007). Integrated insurance design in the presence of multiple medical technologies. *American Economic Review, 97*(2), 427–432.

Groopman, J. (2007). *How doctors think*. New York: Houghton Mifflin.

Gruber, J. & Owings, M. (1996). Physician financial incentives and Cesarean section delivery. *RAND Journal of Economics*, *27*, 99–123.

Halpern, D. F. (1987). *Critical thinking across the curriculum: A brief edition of thought and knowledge*. Lawrence Erlbaum Associates.

Hampson, L. A., Agrawal, M., Joffe, S., Gross, C. P., Verter, J., & Emanuel, E. J. (2006). Patients' views on financial conflicts of interest in cancer research trials. *New England Journal of Medicine*, *355*(22), 2330–2337.

Harris, J. E. (1977). The internal organization of hospitals: Some economic implications. *Bell Journal of Economics*, *8*(2), 467–482.

Hart, O. & Holmstrom, B. (2010). A theory of firm scope. *Quarterly Journal of Economics*, *CXXV*(2), 483–513.

Hayes, J. H., Ollendorf, D. A., Pearson, S. D., Barry, M. J., Kantoff, P. W., Stewart, S. T., et al. (2010). Active surveillance compared with initial treatment for men with low-risk prostate cancer: A decision analysis. *JAMA*, *304*(21), 2373–2380.

Hodgkin, D. & McGuire, T. G. (1994). Payment levels and hospital response to prospective payment. *Journal of Health Economics*, *13*(1), 1–29.

Holmstrom, B. & Roberts, J. (1998). The boundaries of the firm revisited. *Journal of Economic Perspectives*, *12*(4), 73–94.

Howlader, N., Noone, A. M., Krapcho, M., Neyman, N., Aminou, R., & Waldron, W. (Eds.) (2011). *SEER cancer statistics review, 1975–2008*. Bethesda, MD: National Cancer Institute.

Hsu, J., Price, M., Huang, J., Brand, R., Fung, V., Hui, R., et al. (2006). Unintended consequences of caps on Medicare drug benefits. *New England Journal of Medicine*, *354*(22), 2349–2359.

Huskamp, H. A., Deverka, P. A., Epstein, A. M., Epstein, R. S., McGuigan, K. A., & Frank, R. G. (2003). The effect of incentive-based formularies on prescription-drug utilization and spending. *New England Journal of Medicine*, *349*(23), 2224–2232.

Iyengar, S. S. & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology, 79*(6), 995–1006.

Jacobson, M., Earle, C. C., Price, M., & Newhouse, J. P. (2010). How Medicare's payment cuts for cancer chemotherapy drugs changed patterns of treatment. *Health Affairs (Millwood)*, *29*(7), 1391–1399.

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., & Thun, M. J. (2009). Cancer statistics, 2009. *CA Cancer Journal for Clinicians*, *59*(4), 225–249.

Joyce, G. F., Escarce, J. J., Solomon, M. D., & Goldman, D. P. (2002). Employer drug benefit plans and spending on prescription drugs. *JAMA*, 288(14), 1733–1739.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, *5*(1), 193–206.

Kolata, G. (2010). Results unproven, robotic surgery wins converts. *The New York Times*, February 13.

Kozower, B. D. & Stukenborg, G. J. (2011). The relationship between hospital lung cancer resection volume and patient mortality risk: study design: original observational research (population based cohort study). *Annals of Surgery* [Epub ahead of print].

Law, M. R., Majumdar, S. R., & Soumerai, S. B. (2008). Effect of illicit direct to consumer advertising on use of etanercept, mometasone, and tegaserod in Canada: Controlled longitudinal study. *BMJ*, *337*, a1055.

Lee, D. H. & Vielemeyer, O. (2011). Analysis of overall level of evidence behind Infectious Diseases Society of America practice guidelines. *Archives of Internal Medicine*, *171*(1), 18–22.

Leigh, J. P., Tancredi, D., Jerant, A., & Kravitz, R. L. (2010). Physician wages across specialties: Informing the physician reimbursement debate. *Archives of Internal Medicine*, *170*(19), 1728–1734.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 551–578.

Litwin, M. S., Gore, J. L., Kwan, L., Brandeis, J. M., Lee, S. P., Withers, H. R., et al. (2007). Quality of life after surgery, external beam irradiation, or brachytherapy for early-stage prostate cancer. *Cancer*, *109*(11), 2239–2247.

Localio, A. R., Lawthers, A. G., Brennan, T. A., Laird, N. M., Hebert, L. E., Peterson, L. M., et al. (1991). Relation between malpractice claims and adverse events due to negligence. Results of the Harvard Medical Practice Study III. *New England Journal of Medicine*, *325*(4), 245–251.

Lucas, F. L., Sirovich, B. E., Gallagher, P. M., Siewers, A. E., & Wennberg, D. E. (2010). Variation in cardiologists' propensity to test and treat: Is it associated with regional variation in utilization? *Circulation: Cardiovascular Quality and Outcomes*, *3*(3), 253–260.

Lurie, N., Christianson, J., Finch, M., & Moscovice, I. (1994). The effects of capitation on health and functional status of the Medicaid elderly: A randomized trial. *Annals of Internal Medicine*, *120*, 506–511.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E., Leibowitz, A., & Marquis, M. S. (1987). Health insurance and the demand for medical care: Results from a randomized experiment. *American Economic Review*, *77*(3, June), 251–277.

Marshall, S., McGarry, K. M., & Skinner, J. S. (2010). The risk of out-of-pocket health care expenditure at end of life. NBER Working Paper No. 16170.

McClellan, M. (2011). Reforming payments to healthcare providers: The key to slowing healthcare cost growth while improving quality? *Journal of Economic Perspectives*, *25*(2), 69–92.

McClellan, M. & Skinner, J. (2006). The incidence of Medicare. *Journal of Public Economics*, *90*, 257–276.

McGuire, T. G. (2000). Physician agency. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of health economics*. Amsterdam: Elsevier (Chapter 9).

McGuire, T. G. & Pauly, M. V. (1991). Physician response to fee changes with multiple payers. *Journal of Health Economics*, *10*(4), 385–410.

McNeil, B. J., Pauker, S. G., Sox, H. C., Jr., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine*, *306*(21), 1259–1262.

Mello, M. M., Chandra, A., Gawande, A. A., & Studdert, D. M. (2010). National costs of the medical liability system. *Health Affairs (Millwood)*, *29*(9), 1569–1577.

Miller, R. H. & Luft, H. S. (1994). Managed care plan performance since 1980: A literature analysis. *JAMA*, *271*, 1512–1519.

Myers, D. G. & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*, 602–627.

National Comprehensive Cancer Network (2010). Prostate Cancer: NCCN Guidelines for Patients, version 2010. http://www.nccn.com/images/patient-guidelines/pdf/prostate.pdf. Accessed on June 7, 2011.

Newhouse, J. P. (1992). Medicare care costs: How much welfare loss? *Journal of Economic Perspectives*, *6*(3), 3–21.

Newhouse, J. P. (2006). Reconsidering the moral hazard-risk avoidance tradeoff. *Journal of Health Economics*, *25*(5), 1005–1014.

Newhouse, J. P. & the Insurance Experiment Group (1993). *Free for all? Lessons from the RAND Health Insurance Experiment*. Cambridge, MA: Harvard University Press.

Nguyen, N. X. & Derrick, F. W. (1997). Physician behavioral response to a Medicare price reduction. *Health Services Research*, *32*, 283–298.

O'Hare, A. M., Rodriguez, R. A., Hailpern, S. M., Larson, E. B., & Kurella Tamura, M. (2010). Regional variation in health care intensity and treatment practices for end-stage renal disease in older adults. *JAMA*, *304*(2), 180–186.

Pauly, M. V. (1968). The economics of moral hazard: Comment. *American Economic Review*, *58*(3), 531–537.

Phelps, C. E. (1986). Induced demand: Can we ever know its extent? *Journal of Health Economics*, *5*, 355–365.

Potosky, A. L., Legler, J., Albertsen, P. C., Stanford, J. L., Gilliland, F. D., Hamilton, A. S., et al. (2000). Health outcomes after prostatectomy or radiotherapy for prostate cancer: Results from the Prostate Cancer Outcomes Study. *Journal of the National Cancer Institute*, *92*(19), 1582–1592.

Pritchard, R. S., Fisher, E. S., Teno, J. M., Sharp, S. M., Reding, D. J., Knaus, W. A., et al. (1998). Influence of patient preferences and local health system characteristics on the place of death. SUPPORT Investigators. Study to Understand Prognoses and Preferences for Risks and Outcomes of Treatment. *Journal of the American Geriatric Society*, *46*(10), 1242–1250.

Redelmeier, D. A. & Tversky, A. (1990). Discrepancy between medical decisions for individual patients and for groups. *New England Journal of Medicine*, *322*(16), 1162–1164.

Reinhardt, U. E. (1999). The economist's model of physician behavior. *JAMA*, *281*(5), 462–465.

Rice, T. (1983). The impact of changing Medicare reimbursement rates on physician-induced demand. *Medical Care*, *21*, 803–815.

Rogers, E. M. (1983). *Diffusion of innovations*. New York: Free Press.

Rossiter, L. F. & Wilensky, G. R. (1984). Identification of physician-induced demand. *Journal of Human Resources*, *19*, 231–244.

Sanda, M. G., Dunn, R. L., Michalski, J., Sandler, H. M., Northouse, L., Hembroff, L., et al. (2008). Quality of life and satisfaction with outcome among prostate-cancer survivors. *New England Journal of Medicine*, *358*(12), 1250–1261.

Schelling, T. (2006). *Micromotives and macrobehavior*. W. W. Norton and Co.

Schröder, F. H., Hugosson, J., Roobol, M. J., Tammela, T. L., Ciatto, S., Nelen, V., et al. (2009). Screening and prostate-cancer mortality in a randomized European study. *New England Journal of Medicine*, *360*(13), 1320–1328.

Sepucha, K. & Mulley, A. G., Jr. (2009). A perspective on the patient's role in treatment decisions. *Medical Care Research Review*, *66*(Suppl. 1), 53S–74S.

Shaneyfelt, T. M., Mayo-Smith, M. F., & Rothwangl, J. (1999). Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA*, *281*(20), 1900–1905.

Shappley, W. V., 3rd, Kenfield, S. A., Kasperzyk, J. L., Qiu, W., Stampfer, M. J., Sanda, M. G., et al. (2009). Prospective study of determinants and outcomes of deferred treatment or watchful waiting among men with prostate cancer in a nationwide cohort. *Journal of Clinical Oncology*, *27*(30), 4980–4985.

Shrank, W. H., Choudhry, N. K., Fischer, M. A., Avorn, J., Powell, M., Schneeweiss, S., et al. (2010). The epidemiology of prescriptions abandoned at the pharmacy. *Annals of Internal Medicine*, *153*(10), 633–640.

Sirovich, B., Gallagher, P. M., Wennberg, D. E., & Fisher, E. S. (2008). Discretionary decision making by primary care physicians and the cost of U.S. health care. *Health Affairs (Millwood)*, *27*(3), 813–823.

Skinner, J. & Staiger, D. (2005). Technology adoption from hybrid corn to beta blockers. NBER Working Paper No. 11251.
Skinner, J. & Staiger, D. (2009). Technology diffusion and productivity growth in health care. NBER Working Paper No. 14865.
Smith, D. M., Brown, S. L., & Ubel, P. A. (2008). Are subjective well-being measures any better than decision utility measures? *Health Economics, Policy & Law*, *3*(Pt 1), 85–91.
Smith, D. M., Sherriff, R. L., Damschroder, L., Loewenstein, G., & Ubel, P. A. (2006). Misremembering colostomies? Former patients give lower utility ratings than do current patients. *Health Psychology*, *25*(6), 688–695.
Smith, M. R. (2011). Effective treatment for early-stage prostate cancer—possible, necessary, or both? *New England Journal of Medicine*, *364*, 1770–1772.
Snail, T. S. & Robinson, J. C. (1998). Organization diversification in the American hospital. *Annual Review of Public Health*, *19*, 417–453.
Sommers, B. D. & Zeckhauser, R. (2008a). Probabilities and preferences: What economics can teach doctors and patients making difficult treatment decisions. *Urologic Oncology*, *26*(6), 669–673.
Sommers, B. D., Beard, C. J., D'Amico, A. V., Kaplan, I., Richie, J. P., & Zeckhauser, R. J. (2008b). Predictors of patient preferences and treatment choices for localized prostate cancer. *Cancer*, *113*(8), 2058–2067.
Song, Y., Skinner, J., Bynum, J., Sutherland, J., Wennberg, J. E., & Fisher, E. S. (2010). Regional variations in diagnostic practices. *New England Journal of Medicine*, *363*(1), 45–53.
Steinhauser, K. E., Christakis, N. A., Clipp, E. C., et al. (2000). Factors considered important at the end of life by patients, family, physicians, and other care providers. *JAMA*, *284*, 2476–2482.
Stewart, S. T., Lenert, L., Bhatnagar, V., & Kaplan, R. M. (2005). Utilities for prostate cancer health states in men aged 60 and older. *Medical Care*, *43*(4), 347–355.
Sutherland, J. M., Fisher, E. S., & Skinner, J. S. (2009). Getting past denial—the high cost of health care in the United States. *New England Journal of Medicine*, *361*(13), 1227–1230.
Teno, J. M., Licks, S., Lynn, J., et al. (1997). Do advance directives provide instructions that direct care? *Journal of the American Geriatric Society*, *45*, 508–512.
Thaler, R. H. & Sunstein, C. R. (2008). *Nudge*. New Haven, CT: Yale University Press.
Thompson, I., Thrasher, J. B., Aus, G., Burnett, A. L., Canby-Hagino, E. D., Cookson, M. S., et al. (2007). Guideline for the management of clinically localized prostate cancer: 2007 update. *Journal of Urology*, *177*(6), 2106.
Tricoci, P., Allen, J. M., Kramer, J. M., Califf, R. M., & Smith, S. C., Jr. (2009). Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA*, *301*(8), 831–841.
Trivedi, A. N., Moloo, H., & Mor, V. (2010). Increased ambulatory care copayments and hospitalizations among the elderly. *New England Journal of Medicine*, *362*(4), 320–328.
Ubel, P. A. (2009). *Free market madness*. Boston, MA: Harvard Business Press.
Ubel, P. A., Loewenstein, G., Schwarz, N., & Smith, D. (2005). Misimagining the unimaginable: The disability paradox and health care decision making. *Health Psychology*, *24*(Suppl. 4), S57–S62.
Van Heek, N. T., Kuhlmann, K. F., Scholten, R. J., de Castro, S. M., Busch, O. R., van Gulik, T. M., et al. (2005). Hospital volume and mortality after pancreatic resection: A systematic review and an evaluation of intervention in the Netherlands. *Annals of Surgery*, *242*(6), 781–788.
Wei, J. T., Dunn, R. L., Sandler, H. M., McLaughlin, P. W., Montie, J. E., & Litwin, M. S., et al. (2002). Comprehensive comparison of health-related quality of life after contemporary therapies for localized prostate cancer. *Journal of Clinical Oncology*, *20*(2), 557–566.
Weintraub, W. S., Spertus, J. A., & Kolm, P., et al. (2008). Effect of PCI on quality of life in patients with stable coronary disease. *New England Journal of Medicine*, *359*(7), 677–687.
Welch, H. G., Sharp, S. M., Gottlieb, D. J., Skinner, J. S., & Wennberg, J. E. (2011). Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. *JAMA*, *305*(11), 1113–1118.
Wennberg, J., & Gittelsohn, A. (1973). Small area variations in health care delivery. *Science*, *182*(117), 1102–1108.
Wennberg, J., & Gittelsohn, A. (1982). Variations in medical care among small areas. *Scientific American*, *246*(4), 120–134.

Wennberg, J. E., Freeman, J. L., & Culp, W. J. (1987). Are hospital services rationed in New Haven or over-utilised in Boston? *Lancet*, *1*(8543), 1185–1189.

Williamson, O. E. (2010). Transaction cost economics: The natural progression. *American Economic Review, 100*, 673–690.

Wilt, T. J., MacDonald, R., Rutks, I., Shamliyan, T. A., Taylor, B. C., & Kane, R. L. (2008). Systematic review: Comparative effectiveness and harms of treatments for clinically localized prostate cancer. *Annals of Internal Medicine*, *148*(6), 435–448.

Woo, B., Woo, B., Cook, E. F., Weisberg, M., & Goldman, L. (1985). Screening procedures in the asymptomatic adult. Comparison of physicians' recommendations, patients' desires, published guidelines, and actual practice. *JAMA*, *254*(11), 1480–1484.

Yip, W. (1998). Physician responses to medical fee reductions: Changes in the volume and intensity of supply of Coronary Artery Bypass Graft (CABG) surgeries in the Medicare and private sectors. *Journal of Health Economics*, *17*, 675–700.

Young, H. P. (1998). *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton, NJ: Princeton University Press.

Zeckhauser, R. J. (1970). Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory, 2*(1), 10–26.

Zuckerman, S., Waidmann, T., Berenson, R., & Hadley, J. (2010). Clarifying sources of geographic differences in Medicare spending. *New England Journal of Medicine*, *363*(1), 54–62.