PATIENT WELFARE AND PATIENT COMPLIANCE: AN EMPIRICAL FRAMEWORK FOR MEASURING THE BENEFITS FROM PHARMACEUTICAL INNOVATION

Paul Ellickson
Scott Stern
Manuel Trajtenberg

Working Paper 6890
http://www.nber.org/papers/w6890

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1999

This paper was prepared for the CRIW Medical Care Output and Productivity Conference, June, 1998. We are grateful to Ashoke Bhattacharjya, Judy Hellerstein, Alison Keith, Ariel Pakes, and Tom Hubbard for useful conversations. This research was supported by Pfizer and the MIT Program on the Pharmaceutical Industry. The views expressed here are those of the author and do not reflect those of the National Bureau of Economic Research.

© 1999 by Paul Ellickson, Scott Stern, and Manuel Trajtenberg. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.
Patient Welfare and Patient Compliance: 
An Empirical Framework for Measuring the Benefits from Pharmaceutical Innovation
Paul Ellickson, Scott Stern, and Manuel Trajtenberg
NBER Working Paper No. 6890
January 1999
JEL No. C51, C81, I11, O33

ABSTRACT

The main goal of this paper is to develop an empirical framework for evaluating the patient welfare benefits arising from pharmaceutical innovation. Extending previous studies of the welfare benefits from innovation (Trajtenberg, 1990; Hausman, 1996), this paper unpacks the separate choices made by physicians and patients in pharmaceutical decisionmaking and develops an estimable econometric model which reflects these choices. Our proposed estimator for patient welfare depends on (a) whether patients comply with the prescriptions they receive from physicians and (b) the motives of physicians in their prescription behavior. By focusing on compliance behavior, the proposed welfare measure reflects a specific economic choice made by patients. We review evidence that the rate of noncompliance ranges up to 70%, suggesting an important gulf between physician prescription behavior and realized patient welfare. Since physicians act as imperfect but interested agents for their patients, the welfare analysis based on compliance must account for the nonrandom selection of patients into drugs by their physicians. The key contribution of this paper resides in integrating the choices made by both physicians and patients into a unified theoretical framework and suggesting how the parameters of such a model can be estimated from data.

Paul Ellickson
Department of Economics
MIT
Cambridge, MA 02142
paule@mit.edu

Scott Stern
MIT Sloan School
Cambridge, MA 02142
and NBER
sstern@mit.edu

Manuel Trajtenberg
Department of Economics
Tel Aviv University
Tel Aviv 69978 ISRAEL
and NBER
manuel@post.tau.ac.il
I. Introduction

The pharmaceutical industry's innovative output consists primarily of a small number of new drugs, each of which is required to receive FDA approval. While substantial social value is often attributed to pharmaceutical innovation, there have been only a small number of actual evaluations of the welfare gains stemming from the introduction and diffusion of new drugs (Lichtenberg, 1996). In the absence of a measurement framework to assess the patient benefits arising from new product introduction, regulation of the pharmaceutical industry and other institutions of the health care system turns on an incomplete vision of the relevant costs and benefits of different public policy choices.

The main goal of this paper is to develop an empirical framework for evaluating the patient welfare benefits arising from pharmaceutical innovation. Extending previous studies of the welfare benefits from innovation (Trajtenberg, 1990; Hausman, 1996), this paper unpacks the separate choices made by physicians and patients in pharmaceutical decisionmaking and develops an estimable econometric model which reflects these choices. Our proposed estimator for patient welfare depends on (a) whether patients comply with the prescriptions they receive from physicians and (b) the motives of physicians in their prescription behavior. By focusing on compliance behavior, the proposed welfare measure reflects a specific economic choice made by patients. Moreover, since physicians act as imperfect agents for their patients, physician prescription behavior reflects both the consequences of agency as well as an evaluation of which drug yields the highest benefits for a given patient. The key contribution of this paper resides in integrating the choices made by both patients and physicians into a unified theoretical framework and suggesting how the parameters of such a model can be estimated from data.

Relying on recent advances in the study of differentiated product markets (Berry, 1994; Trajtenberg, 1990), we develop a discrete choice model that lends itself naturally to the evaluation of welfare gains from pharmaceutical innovation. The model highlights two important aspects of pharmaceutical markets. First, pharmaceutical therapies are discrete in nature; for most diseases, one drug regimen is given to each patient to the exclusion of substitutes. Second, patients are
heterogeneous in ways which may be either observed or unobserved by the investigator. Examples include the severity of their illness, their price-sensitivity, or their sensitivity to the side effects associated with specific drugs. As a first step, we propose a “baseline” model of pharmaceutical choice which abstracts away from the institutional details of pharmaceutical decision making. In this model, fully informed patients are assumed to hold authority over their pharmaceutical choices and bear full financial responsibility for their decisions. Extending past characterizations, we present a computationally straightforward method to calculate patient welfare under this baseline model.

We then turn to the heart of the paper – the development of an estimable model of pharmaceutical choice which accounts for the most salient institutions of pharmaceutical decision making. The key problem lies in the fact that the prescription decision is vested with the physician rather than with the patient. While physicians are more informed than their patients about the relative benefits of different therapies and have some incentives to attain an “optimal matching” between patient and drug, a wedge may yet exist between the interests and preferences of the patient and the actual behavior of the physician. Specifically, physicians may be less sensitive than patients to the effective prices of drugs and may underinvest in gathering the types of information about patients and/or drugs which yield the best fit (Stern and Trajtenberg, 1998). As a result, the exclusive use of physician prescription patterns to infer patient welfare is at best problematic. Such a calculation would of course capture patient welfare to the degree that physicians act as a “filter” for patient preferences, but prescription-level analysis cannot reflect the economic choices made by patients directly.

In order to address the problems stemming from the wedge between physician and patient preferences in the context of drug choice, we redirect our analysis towards the choices that patients do make – whether or not to comply with prescriptions. After receiving a prescription from a physician for a specific drug, patients choose whether or not to fill the prescription (purchase

---

1 This phenomenon may have significantly changed in recent years: the rise of managed care may have increased the incentives of physicians to respond to the true prices (e.g., through a capitation system) while managed care may have reduced the effective price-sensitivity of many patients (by offering more generous pharmaceutical insurance than many fee-for-service plans).
compliance), whether or not to maintain the regimen once purchased (use compliance), and whether or not to maintain the prescription over the life of refills and follow-up (sustained compliance). Accordingly, variation in compliance rates across drugs reflects different valuations of the incremental utility afforded by drugs as compared to a common baseline, namely no pharmaceutical therapy (the “outside good”). Of course, welfare analysis based on compliance must account for the fact that, since physicians act as informed and interested agents for their patients, patients who choose whether to comply with a prescription for a given drug have been selected into that choice by their physician.

Clearly, if nearly all patients comply (or if there is little variation in compliance across drugs), the fact that patients face choices still would leave little room for the actual measurement of patient welfare. We therefore present a review of evidence from the clinical medical literature suggesting that compliance both within and across drug therapies is an important empirical phenomena, providing a basis for empirical work in this area. While medical researchers focus on different issues and frame them in different terms than economists might (for example, most assume that noncompliance is “irrational” on the part of the patient), the evidence is compelling. First, noncompliance rates are astonishingly high, reaching up to 70%. Second, there is substantial variation in the compliance rate, depending on the type of drug and disease being treated.

With this evidence as motivation, we introduce our alternative model and corresponding welfare function. Two important issues are addressed: first, patients vary in the degree and nature of their insurance and in their unobserved costs of complying with their physician’s approved therapy. Second, since compliance is conditional on the prescription behavior of physicians, we account for the selectivity of patients into drugs, induced by physicians responding to patient characteristics which are not observed by the econometrician. To address this selection problem, we specify a general model of physician behavior and identify the distribution of idiosyncratic utility conditional on prescription. This conditional distribution is a simple function of the parameters of the physician behavior model and can be calculated analytically for special classes of distributions. Thus, our model provides a way of controlling for the fact that physicians match patients to drugs, for any set of assumptions about how that matching process unfolds. With such a control in place, we are able to
propose a consistent estimator for patient welfare for any set of pharmaceutical products available in the market. This estimator is a function of the vector of compliance rates for individual drugs in a therapeutic category, as well as of the characteristics of those drugs and of the underlying patient population. Our estimator can be used to perform counterfactuals such as the welfare loss associated with a year of delay in regulatory approval, or the incremental returns from “one-a-day” pills which increase compliance versus a completely new form of treatment for a particular pathology.

Our analysis suggests that there are high returns to understanding how patients respond to choices in the health care system, even if most of the system involves the delegation of authority to their agents (primarily physicians). Currently, there are few systematic data gathering efforts by the government (or by private data gathering sources) aimed at collecting this type of information, limiting both our understanding of the benefits from pharmaceutical innovation as well as the welfare impact of physician authority.

II. A “Baseline” Model of Pharmaceutical Choice and Welfare

We begin by restating the commonly held assumption that consumer welfare can be measured by the “revealed preferences” of consumers through their observed choices. In the market for a single, homogenous good, only those consumers who value the good above its price purchase the good.\(^2\) Under these conditions, the consumer welfare (surplus) in this market is measured by the area between the demand curve and price; further, the incremental welfare from product innovation requires a comparison of the difference in the area under the demand curve before and after the innovation has been introduced into the market. In this sense, the welfare benefits from technological change result from the diffusion of new technologies rather than their mere invention (Griliches, 1958).

While many studies have attempted to gauge the producer returns to innovation, both in the pharmaceutical industry as well as elsewhere (Hall, 1995), the difficulties associated with estimating

\(^2\) Or, alternatively, each consumer purchases the good until her marginal utility equals the price of the good, making her indifferent between an additional unit of the good and its expense (the price).
demand have limited the calculation of the consumer welfare implications of innovation. These difficulties arise in part because product innovation occurs mostly in product differentiated markets, and hence new goods do not simply augment the prevalent market demand curve but provide an imperfect substitute for older goods. When consumers substitute the new (innovative) good for the old, the consumer welfare benefits from the innovation are composed of two parts: the first consisting of the incremental value placed on the new product by those who substituted (a direct effect); the second capturing the impact of the introduction of the new product on the prices and utility earned by consumers of the old product (which will yield a set of indirect effects). Thus, the calculation of consumer welfare arising from the introduction of a new differentiated product requires an estimate of the degree to which the new product replaces older goods, an estimate of the incremental value gained by those who switch, and finally, an evaluation of the competitive impact of innovation on the market prices of existing products.

Consider the introduction of Zantac by Glaxo in 1982. Zantac provided a differentiated substitute for Tagamet, which had been introduced about five years earlier. While Zantac was believed to be superior to Tagamet along some therapeutic dimensions, Tagamet remained the preferred product for a portion of the market (though Zantac eventually achieved a majority market share). In order to calculate the incremental welfare arising from the introduction of Zantac, we need to estimate a demand system which allows us to calculate consumer surplus both when Zantac was in and out of the market. With such a demand system, it is possible to calculate the impact of Zantac as it diffused into the market. After all, the bulk of consumer benefits from Zantac did not arise upon Zantac’s introduction (at which time it achieved only a small market share), but only as patients substituted over time out of Tagamet (or no therapy) into the newer drug.
While estimating the welfare impacts associated with these market dynamics is challenging, several methods which have been developed in recent years allow for accurate measurement (Bresnahan, 1986; Trajtenberg, 1990; Hausman, 1996), most notably the discrete choice framework which forms the basis of our current approach. The basic notion in these models is that competing products in a given market can be thought of as consisting of different vectors of characteristics (or performance dimensions), selling for different prices. Consumers derive utility from these characteristics (disutility for price), and choose their preferred product by comparing the various options available in the market in terms of the overall utility that different products provide. The econometric estimation of demand models of this sort yields the parameter estimates needed to compute the welfare gains from innovation: the marginal utility of the attributes of the products, the degree of substitutability between new and old products, and other parameters pertinent to the diffusion process of new products. We can exploit our estimate of the value that consumers place on attributes to compute the incremental surplus associated with the introduction of new products incorporating superior characteristics.

Trajtenberg (1989, 1990) applies this framework to the case of Computed Tomography (CT) Scanners, one of the most remarkable medical innovations of the last few decades. Even though CT Scanners are complex systems from a technological viewpoint, there are just a few attributes that characterize their performance (primarily scan speed and resolution). In the decade following the introduction of the first scanner by EMI in 1973, a tremendous amount of entry and innovation took place in the CT scanner market, leading to dramatic improvements in those attributes (as well as the introduction of new features). For example, scan time dropped from 5 minutes to 1-2 seconds, and spatial resolution improved to less than one millimeter. Each year, buyers of CT Scanners faced much improved choice sets; the question is how valuable those improvements were. Using detailed data on the prices, attributes, and sales of each model in the market each year, Trajtenberg (1989) estimated a discrete choice model of demand for these systems. The estimated demand parameters were used to compute the (substantial) welfare benefits stemming from the innovations introduced year after year. These calculated gains were then used to compute the social rates of return to investments in R&D and examine the pattern of those gains over time.
In view of the peculiarities of pharmaceutical markets, this methodology needs to be extended and modified in order to apply it successfully to the study of innovation in pharmaceuticals. In particular, while Trajtenberg abstracted away from the institutional details of hospital decisionmaking, our approach tackles these issues head-on. In particular, our preferred framework (developed in Section IV) assesses patient welfare from the analysis of compliance behavior, rather than simply relying on observed prescriptions, since compliance is a choice made by patients, while prescription is a choice made by physicians acting as agents for patients. However, in order to fix ideas we first abstract away from the institutional context and agency problems, and introduce a baseline model of pharmaceutical decision making predicated on the assumption of “optimal” prescribing and purchasing behavior by informed patients with authority over their treatment choices.

Our point of departure is a simple discrete choice model, as in Berry (1994). Each patient maximizes the utility derived from pharmaceutical purchasing by choosing among \( J_t + 1 \) alternatives \((J_t \text{ marketed products in year } t \text{ and the option of no purchase (} j=0))\), as follows,

\[
\text{MAX}_{j \in \{0, \ldots, J_t\}} \quad V_{ij} = X_j' \beta_i + \alpha_i \text{PRICE}_j + \xi_j + \epsilon_{ij} = \delta_j + \mu_{ij} \tag{1}
\]

where \( V_{ij} \) is the value of drug \( j \) to patient \( i \), \( \beta_i \) the marginal valuations of the observed characteristics of drug \( j \), \( \xi_j \) the disutility associated with price, \( \epsilon_{ij} \) an idiosyncratic patient-drug specific effect. Berry (1994) suggests rewriting such a value function in terms of the mean utility accruing to a representative consumer, \( \delta_j \), and the deviation from that mean valuation for an individual consumer, \( \mu_{ij} \), where the joint distribution of idiosyncratic utility, \( F(\mu;\sigma) \), is parameterized according to \( \sigma \).

The choice problem in (1) determines the probability that a patient of a given type chooses the \( j^\text{th} \) drug, i.e., \( Pr( V_{ij} > V_{ik} \forall j \neq k ) = Pr( \delta_j + \mu_{ij} > \delta_k + \mu_{ik} \forall j \neq k ) \). Moreover, by considering a population of such patients, it is possible to estimate the demand for each drug (i.e., its market share, \( s_j \)) as a function of its own price and characteristics (conditional on the prices and characteristics of alternatives). Product-level demand depends both on the average utility level, \( \delta_j \), and the degree of substitutability with other products (i.e., whether \( \mu_j \) is correlated with other...
elements of $\mu$). The empirical characterization of the demand system therefore requires estimates of the elements of $\delta$ and $\sigma$. As discussed in Berry (1994), estimating the estimation of distributional parameters may require the repeated evaluation of a $(J_t + 1)$-dimensional integral, a computationally intensive task in many circumstances. Prior research has overcome this challenge by drawing upon distribution functions from the Generalized Extreme Value (GEV) class, allowing for the analytical computation of the market share function (and, as will be seen below, of the welfare function). Our proposed baseline model of pharmaceutical demand follows this methodology, and hence we develop the calculation of welfare under the assumption that the distribution of idiosyncratic utility follows a GEV distribution. In this context, it is useful to recall that the logit function for market share, 

$$s_j = \frac{e^{\delta_j}}{\sum_{j \in J} e^{\delta_j}},$$

results from the imposition of an independent Type I Extreme Value distribution, the simplest distribution function in this class (McFadden, 1978). While our framework is flexible enough to accommodate more general distributional assumptions (including semiparametric models), our focus on GEV and variants allows us to sharpen the issues associated with welfare calculation in the context of a computationally feasible model.

While the maximization problem in (1) abstracts away from agency and learning problems, it does highlight important elements of pharmaceutical choice. First, the model makes it clear that the benefits from new pharmaceutical products arise from substitution out of old therapies (or no therapy) into the new drug; second, the model highlights the centrality of patient heterogeneity. By specifying that the parameters which govern the value placed on each drug be patient-specific, the model accommodates heterogeneity along several dimensions, including differential sensitivity to price and to therapeutic characteristics of the drug (such as side effects, dosing regimens, or bioavailability). This patient-drug interaction captures the idea that pharmaceutical choice involves “matching” each patient with the drug which is most appropriate for their specific condition (Melmon et al, 1993; Stern and Trajtenberg, 1998).
As suggested above, the parameters of the baseline model can be estimated from the relationship between observed market shares and the prices and characteristics of different drugs available in the market. With the aid of these parameters, one can then characterize the incremental benefits associated with expansions in the choice set or changes in the characteristics of particular choices (such as changes in prices or dosing regimens). Concretely, as shown in Trajtenberg (1990), \( W_t \) can be computed just as the summation of the consumer surplus associated with each product, conditional on the prices and characteristics of available substitutes:

\[
W_t = \sum_{j=0}^{J} \sum_{p_j} Z(q_j|q_k = p_k, \forall k < j, q_k = \infty, \forall k > j) dq_j
\] (2)

It is useful to note that, even in this general formulation, the incremental welfare from a new good depends upon the level and steepness of the slope of the demand curve for it. To the extent that the new product is a close substitute for old products (and thus faces a flat demand curve), the welfare gains from its introduction will be less dramatic than if existing products are poor substitutes for the new good.

Calculating each element of (2) requires integrating the J-dimensional integral which determines market share, \( s_j = \sum_{k=0}^{J} \sum_{p_j} Z(q_j|q_k = p_k, \forall k \neq j, \forall k > j) dq_j \). However, as mentioned above, if \( \mu \) is drawn from the Generalized Extreme Value distribution, then this computational complexity is substantially eased, and it is feasible to calculate the market share function analytically (McFadden, 1978; Bresnahan, Stern, and Trajtenberg, 1997). Here we extend this prior result and show that the welfare function is also an analytically defined function of the GEV distribution function and depends only on estimating the parameters of the discrete choice model. This can be seen most clearly in the case where price-sensitivity is constant:

**Proposition 1** Under the maximization model in (1) and \( \alpha_i = \alpha \), if \( G_i: \mathbb{R}^{i+1} \rightarrow \mathbb{R}^1 \) is a non-negative, homogenous of degree one function satisfying certain restrictions,\(^3\) then:

\[
F(\mu_{i,0}, \ldots, \mu_{i,J}) = \exp -G_i(e^{-\mu_{i,0}}, \ldots, e^{-\mu_{i,J}})
\]

is the cumulative distribution function of a multivariate GEV distribution and,

\(^3\) The limit of \( G(\bullet) \) as any argument goes to \( \infty \) must be equal to \( \infty \), mixed partials of \( G(\bullet) \) alternate in sign, and first derivative with respect to each argument is nonnegative (McFadden, 1978).
\[ W_t = \frac{\ln(G_t(\delta_0, \ldots, \delta_{J_t}))}{-\alpha} \]  

(3)

is the per-capita expected utility (or average level of consumer welfare) from participating in the market.

**Proof:** By Roy’s identity, we know that \( \frac{\partial W}{\partial p_j} = -s_j \). From Theorem 1 of McFadden (1978),

\[ s_j = \frac{\partial G(\delta)}{\partial \delta_j}. \]

Assume that \( W_t \neq \frac{\ln(G_t(\delta_0, \ldots, \delta_{J_t}))}{-\alpha} \). Then, \( \frac{\partial (\ln(G_t(\delta_0, \ldots, \delta_{J_t}))}{\partial p_j} \neq -s_j \). But,

\[ \frac{\partial (\ln(G_t(\delta_0, \ldots, \delta_{J_t}))}{\partial p_j} = \frac{\partial G(\delta)}{G(\delta)} \frac{\partial G(\delta)}{\partial \delta_j} \]

establishing a contradiction.

The functional form for the choice probabilities in a GEV model makes calculating patient welfare particularly straightforward. As additional products are introduced into the market (or the features of existing products are enhanced), the value of \( G_t \) increases, and so does the welfare function \( W_t \). Proposition 1 also suggests a useful way to conceptualize the measurement and meaning of patient welfare: it is the monetary amount a patient would “pay” to be faced with the choice set \( J_t \) prior to observing the realization of idiosyncratic utility (\( \mu_i \)).

Extending Proposition 1 to accommodate heterogeneity in \( \alpha \) is immediate. When patients differ in their price sensitivity, total patient welfare requires the calculation of just a single-dimensional integral over the distribution of price sensitivities, as follows:

\[ W_t = \int_n(G_t(\delta_0(\alpha_i), \ldots, \delta_{J_t}(\alpha_i)) \frac{1}{-\alpha_i} dF(\alpha_i; \sigma_a) \]  

(4)

Under this framework, calculating the incremental welfare benefits from innovation is straightforward, involving just the difference \( W = W_{t+1} - W_t \), which captures the gains in consumer welfare as the product set changes between the two time periods.
III. Economic Implications of Physician Authority and Patient Compliance

Two related features of pharmaceutical decisionmaking suggest that the baseline model presented above may lead to a biased and potentially misleading assessment of the welfare gains arising from pharmaceutical innovation. First, there are strong reasons to believe that physician authority over the prescription decision may lead to systematic biases in prescribing patterns. Since information asymmetry is at the heart of any “expert” relationship, physicians may have the opportunity to take advantage of their informational advantage in their prescription behavior. For example, several recent studies point to the presence of strong habit effects, whereby some physicians tend to prescribe in the same way across patients, even though the heterogeneity of patients’ conditions may call for matching different drugs to different patients (Hellerstein, 1998; Stern and Trajtenberg, 1998; Coscelli, 1998). To the extent that individual patients find it difficult to monitor such behavior, physicians may earn an information rent through underinvestment in “matching” individual patients to drugs. As developed in related work (Stern and Trajtenberg, 1998), this would manifest itself in a high degree of concentration in the physician’s prescribing portfolio and a tendency to prescribe drugs which are most appropriate for an “average” patient.5

Second, to the extent that patients choose not to comply with prescribed therapies, a gulf may arise between physician prescribing patterns and realized patient welfare. Although patients may have relatively little control over the medication prescribed, they are free to ignore their physician’s recommended regimen. In fact, since compliance rates reflect patients’ valuations of particular therapies, we can take advantage of this observed behavior to infer welfare. By approaching welfare in such a way, our analysis builds on a growing literature aimed at acknowledging the information

---

4 For example, in the simplest case of logit probabilities, \( w_i = \frac{\ln\left( \sum_j e^{\theta_{ij}} \right)}{\theta} \).

5 On the other hand, the impact of agency in pharmaceutical decisionmaking should not be overstated. In contrast to other areas of health care which are subject to physician inducement (Gruber and Owings (1996)), physicians receive no direct pecuniary benefit from prescribing one drug over another. Of course, to the extent that the patients can choose their physician, there exists a practice-building incentive to provide high-quality care; however, this practice-building has the effect of ameliorating the agency problem rather than exacerbating it. The presence of induced demand considerations may impact the overall level of pharmaceutical demand, as physicians may order expensive (and revenue-producing) procedures and substitute away from pharmaceutical therapy. In other words, while direct pecuniary-based incentive issues may shape the overall substitution between drugs and other therapies, agency within
value inherent in patient decision-making and the effect of patient choice over health care outcomes (Philipson and Posner, 1993; Meltzer, 1999). In particular, our methodology complements the work of Philipson and Hedges (1998) which argues that the statistical evaluation of clinical trials must account for the active role that subjects play in evaluating treatments. Specifically, a patient’s decision to withdraw from an experiment reflects their evaluation of the effectiveness of the therapy (which the patient knows may be simply a placebo). Those patients who receive the greatest disutility from being placed on the placebo may opt out of the clinical trial, leading to a downward bias in the measured effectiveness of the drug as calculated by a difference between the (ex-post) treated and control groups. Our model extends these prior analyses by focusing on the implication of compliance for the doctor-patient relationship itself, using the observed compliance share to quantify the wedge between physician and patient valuations.

The degree of observed patient noncompliance is truly surprising. Several studies put overall patient noncompliance at around 50% (Sacket, 1979), indicating a sizeable difference between the benefits perceived by physician and patient. The problem of patient noncompliance has garnered sustained interest in the medical literature for the past 25 years. Two studies have estimated the cost of noncompliance as a result of hospital re-admissions and lost productivity at $100 billion annually (National Pharmaceutical Council, 1992; Task Force for Non-Compliance, 1994). Patient noncompliance extends to a variety of chronic conditions, cuts across demographic categories, and covers a wide range of gravity of cases (Dunbar-Jacob et al, 1995). From a clinical perspective, noncompliance involves a variety of costs above and beyond the simple reduced effectiveness of the medication, including reduced ability by physicians to assess drug regimen effectiveness, increased drug resistance, and a higher probability of the onset of a more severe condition. In addition, overestimation of compliance on the part of physicians (given that noncompliance is hard to detect) may lead to inadvertent increases in dosage, decreased incentives to consider alternative therapies.

There are a host of additional issues that arise in the context of compliance, and constitute interesting lines of investigation: how does the principal’s (patient’s) reluctance to truthfully reveal their pharmaceutical consumption impact the prescribing behavior of the agent (the physician)? Does optimism regarding compliance on the part of physician lead to inefficiently high levels of medication? Can this problem be mitigated through the development of compliance enhancing one-a-day medications?
and discontinuance of effective therapies which are simply not implemented by the patient.

A primary concern of the clinical medical literature is simply measuring compliance. Patients choose whether to comply with a prescription in two stages: first, whether to purchase the medication (purchase compliance) and then whether to follow the prescribed regimen (use compliance). Measurement of compliance has been attempted using patient interviews, pill counting, urine and blood tests, and, most recently, electronic and chemical monitoring. Studies using more sophisticated methodologies (such as monitoring) tend to find higher levels of noncompliance (McGavack, Britten, and Weinman, 1996). Use noncompliance has been found to vary significantly across therapeutic categories: 36% in hypertension (Dunbar-Jacob et al, 1991), 40% to 60% in arthritis (Belcon, Haynes, & Tugwell; 1984, Hicks, 1985), 15% to 43% among organ transplant recipients (Didlake et al, 1988; Rovelli et al, 1989), and 18% to 70% in the treatment of depression (Engstrom, 1991; Myers and Branthwaite, 1992).

The medical literature has also established a strong link between noncompliance and adverse medical outcomes. Indeed, it is estimated that more than one third of hospital re-admissions for heart failure result from noncompliance with dietary and medication regimens (Ghali et al, 1988; Vinson et al, 1990) while among patients who sustain myocardial infarction, those with poor compliance records were 250% more likely to die within a year of follow up (Horwitz, 1990). Another study suggests that “actual compliance…might reduce stroke risks by about one half and coronary heart disease by about one fifth within a few years” (Collins, 1990). In insulin dependent diabetes, 39% of single and 31% of multiple admissions have been attributed to poor compliance (Fishbein, 1985), while in tuberculosis and HIV infections, there is an established link between noncompliance and drug resistance (Bloom & Murray, 1992). At the extreme, Rovelli (1989) estimates that the probability of tissue rejection (or death) can be as much as four times higher as the result of noncompliance by patients.

These studies can be usefully framed within a health care production function framework:

---

7 Although this difference is not always clearly spelled out in the literature, studies have found noncompliance to be around 20% in purchase and 50% in use (Bearden et al, 1993).
how does noncompliance impact the production of health? Not surprisingly, decreasing a key “input” reduces overall output. What is missing from this analysis is a discussion of patient welfare. Are the long-term health benefits of compliance outweighed by more immediate concerns? In other words, do patients substitute decreased long term health prospects for an immediate reduction in negative side effects or other inconveniences associated with drug therapies? Is noncompliance a problem of information or a response to the true psychic and other costs associated with maintaining a drug regimen? Addressing these questions requires understanding how patient and drug characteristics impact the compliance decision.

Indeed, a growing literature focuses on identifying the patient and drug characteristics associated with noncompliance. Perhaps surprisingly, simple demographic characteristics (sex, income, etc…) have not been consistently linked to compliance (Royal Pharmaceutical Society of Great Britain, 1998). On the other hand, regimen features such as complexity, number of medications, and duration have been associated with the compliance rate (Goodall & Hallford, 1991; Col et al, 1990, Parkin et al, 1976). The patient’s evaluation of effectiveness or the severity of side effects are also significant (Conrad, 1985; Basler and Weissbach, 1984). This suggests that patients are responding to perceived costs, both monetary and psychic, when choosing whether or not to comply. The economics of compliance are particularly salient in asymptomatic conditions, where patients are trading off a reduction in immediate and noticeable side effects for an increased risk of future pathology. In the case of insured patients, there is an additional incentive to discount future costs of health care.

Traditionally, health care researchers (particularly noneconomists) have treated noncompliance as the result of irrational or at best misinformed behavior. However, in response to findings that compliance is responding to such factors as the level of side effects, the health care community is reevaluating the rationale for this type of patient behavior. This new approach stresses the importance of factoring the patient’s “beliefs” into the determination of appropriate therapies (Royal Pharmaceutical Society of Great Britain, 1998), emphasizes education for both patients and practitioners, and giving patients greater control over health decisions. While these policy
recommendations seem eminently sensible, it is quite clear that both economists and health care professionals have yet to develop a clear understanding of the causes and consequences of imperfect compliance.

In sum, the evidence clearly indicates that patient compliance is an important empirical phenomena, with far-reaching economic and productivity measurement consequences. Two specific examples may shed further light on such issues. First, most prior studies of health care productivity and health care “production” have abstracted away from the reformulation of drugs (such as one-a-days), assuming that such formulations simply pose an “aggregation” problem. However, to the extent that compliance is increasing in once-a-day formulations, a revealed preference perspective suggests that there may be substantial incremental welfare gains associated with such therapies. Second, failing to account for patient compliance behavior can also lead to biased measures of the welfare gains arising from the introduction of generic brands. In many instances, the choice between the generic and branded versions of drugs resides at least in part with the patient in consultation with the pharmacist (Ellison et al, 1997). No extant study has examined how the availability of a generic formulation impacts the purchase compliance associated with a drug. Such an exercise could provide direct evidence about patient sensitivity to price conditional on prescription. Motivated by these measurement concerns, we now turn to an estimable empirical model of patient welfare, which focuses on the patient compliance decision while fully incorporating physician prescription substitution patterns.

IV. An Empirical Framework for Measuring Patient Welfare Based on Patient Compliance

The two principal insights to be drawn from Section III are that physician prescription patterns may not reflect patients’ preferences, and that patient compliance represents an economic choice which should allow for identification of the incremental benefits of a given drug over the

---

8 This underestimation of welfare is similar to the concerns raised by Hausman (1996), who suggests that even relatively small changes in the product set may have large absolute welfare consequences in the presence of consumers who are sufficiently sensitive to the degree of the match.
alternative of no drug at all. The goal of this section is to incorporate these insights into a estimable model of patient welfare. We start out by expanding the framework of Section II and consider a two-stage sequential decision process. In the first stage, the physician chooses one drug regimen among \( J \) available regimens. In the second stage, the patient chooses whether or not to comply with the prescription (see Figure A).

Two key issues arise in such a model. First, the welfare function needs to be modified to reflect the nature of the choices facing patients. Second, to obtain a consistent estimate of the appropriate welfare function, the model must account for the selection by physicians of patients into particular drugs. To the extent that there exists positive dependence between the physician’s evaluation of idiosyncratic patient-drug utility, and the underlying (true) patient utility, the sample of patients who are prescribed a particular drug will be biased towards patients who have particularly high valuations for that drug. We start by specifying a simple model of physician choice over drugs:

\[
\max_{j \in \{0, \ldots, J\}} V_{ij}^{MD} = X_j^j \beta_{MD} + \alpha_{MD} \text{PRICE}_j + \xi_j + \varepsilon_{ij}^{MD} = \delta_j + \mu_{ij}^{MD}
\]  

(5)

As will be seen below, a tractable version of the physician behavior model is key to ensuring estimability of a welfare formula based on patient compliance but controlling for physician selection. Consequently, we repeat our suggestion from Section II and resort to a tractable distribution drawn from the GEV family, yielding prescription shares for the total population equal to \( s_j = \frac{\partial G(\delta; \sigma)}{\partial \delta_j} \).

Conditional on the physician’s prescription in (5), each patient chooses whether or not to comply with that choice. If patients respond to exactly the same factors which determine the solution to (5), the compliance rate would of course be equal to one. To the extent that the physician chose a particular drug over the outside good in the first stage, then the patient would also choose that prescribed drug over the outside good in the second stage. In order to have a

---

9 Of course, one could expand the compliance model to incorporate dynamic elements such as the hazard rate of noncompliance. For example, the clinical literature distinguishes between complete noncompliance and partial noncompliance or “drug holidays.”
meaningful compliance model, then, the patient’s decision model must include elements observed by the patient but is not accounted for by the physician:

$$\begin{align*}
    \text{MAX} \\
    j \in \text{COMP}, 0 \\
    V_{i,\text{COMP}} &= X_j' \beta + \alpha_i \text{PRICE} + \xi_j + \mu_{i,j} + \eta_{i,\text{COMP}} = \lambda_j + V_{i,j} \\
    V_{i,0} &= 0 + V_{i,0} 
\end{align*}$$

 Patients choose whether or not to comply according to their valuation of observed product characteristics, $X_j$, their disutility for price, idiosyncratic valuation which was both observed and responded to by the physician ($\mu$), and an additional element of idiosyncratic valuation unobserved by the physician ($\eta$). Note that potentially important components of $\eta$ are the opportunity and attention costs associated with compliance.

The maximization problem in (6) characterizes the fundamental economic decision faced by individual patients in the context of pharmaceutical treatment. Other features of the health care environment can be incorporated easily into this framework (insurance, demographics, a dynamic specification specifying the hazard rate for noncompliance rather than a single discrete decision, etc.). From the perspective of calculating welfare, however, the most subtle element in (6) concerns the overall distribution of random utility, $\nu$, which is a function in part of the draw observed by both the patient and the physician ($\mu_j$). Note that this overall random term, $\nu$, cannot be mean zero if physicians “skew” their prescription behavior towards patients with particularly high valuations of this drug. In the particular form of selection suggested by (5) and (6), the distribution of $\mu_j$ is the distribution of $\mu_j$ from the physician’s multinomial choice equation, conditional on having prescribed $j$. To see the implications of this selectivity, consider repeated trials of (5), and, for each drug $j$, select out the $\mu_j$ of those trials for which $V_{i,j} > V_{i,k}$ $\forall j \neq k$ (i.e., drug $j$ is chosen in that trial). The distribution of $\mu_j$ in (6) is then simply the distribution of these selected trials:
Proposition 2. Let \( g^*(\mu_j) = f(\mu_j | V_{ij} > V_{ik} \forall k \neq j) \) be the distribution of the selection effect in (6). Then
\[
g^*(\mu_j) = \frac{1}{s_j} f(\mu_j; \sigma_j) \prod_{k \neq j, k \in J} F_{ik}(\delta_j - \delta_k + \mu_j; \sigma)
\]
where \( s_j \) is the overall share of \( j \) in the physician portfolio, and \( F(\mu; \sigma) \) is the assumed distribution of idiosyncratic utility in (5).

**Proof:** By Bayes’ Theorem, \( f(\mu_j | V_{ij} > V_{ik} \forall k \neq j) = \frac{Pr(V_{ij} > V_{ik} \forall k \neq j | \mu_j) f(\mu_j)}{Pr(V_{ij} > V_{ik} \forall k \neq j)} \).

The denominator is simply \( s_j \), and
\[
Pr(V_{ij} > V_{ik} \forall k \neq j | \mu_j) = Pr(\delta_j + \mu_j > \delta_k + \mu_k \forall k \neq j | \mu_j)
\]
which can be rewritten in terms of the product of the distribution functions associated with each \( k \) evaluated at \( \delta_j - \delta_k + \mu_j \).

The distribution of \( g^*(\mu_j) \) is therefore simply a function of the distribution of the maximum realization when \( J \) random variables are drawn from the unconditional distribution \( F(\mu; \sigma) \). It is important to note that Proposition 2 holds for any \( F(\mu; \sigma) \) and so calculating \( g^*(\mu_j) \) only requires the ability to calculate \( F(\mu; \sigma) \) for any particular point in the distribution. When \( F(\mu; \sigma) \) is drawn from the GEV class of distributions, calculating \( g^*(\mu_j) \) is a simple analytical function of observables and of the (estimated) parameters of the model. To see the statistical logic behind Proposition 2 more clearly, consider the case where the draws are independent and \( \delta_j = \delta_k \ \forall j, k \). In this extreme case, Proposition 2 reduces to the distribution of the \( J \)th order statistic of an iid random variable, \( g^*(\mu_j) = J_i f(\mu_j; \sigma_j) (F(\mu_j; \sigma))^j-1 \) (Larsen and Marx, 1986).

Proposition 2 is crucial to our ability to calculate patient welfare because the realized sample selection distribution depends on the choices available for physicians to prescribe. For example, as new drugs enter the market, physicians will tend to substitute the new drug for patients with relatively low valuations of the older drugs. Consequently, the realized average utility for the older drug (conditional on prescription) may be increasing in the offered product set. Note that this increase is not due to any change on the part of the drug itself but on how the portion of the population which is prescribed this drug is changing as new drugs enter the market.
We can now establish a patient welfare measure in the context of pharmaceutical prescription and compliance. Recall that in our earlier discussion, we suggested that the welfare measure can be conceptualized as the maximum monetary amount that a risk-neutral individual would be willing to pay to “access” the prescription decision tree (Figure A) prior to observing their individual draws ($\mu$, $\eta$). In the current model, this is simply the expected welfare from any prescription conditional on having received that prescription, times the probability of receiving that prescription, i.e.,

$$W_t = \sum_{j \in J} w_t^j \cdot \Pr(MD \text{ prescribes } j) \quad (7)$$

where $w_t^j$ is the expected patient welfare from drug $j$ conditional on having been prescribed drug $j$ by the physician.

To illustrate (7), we consider two cases under the assumption that $\eta$ is also distributed according to the extreme value distribution. First, consider the case where physicians allocate patients randomly among drugs and so there is no selectivity in the distribution of individuals facing the compliance decision associated with any one drug. In that case, $w_t^j$ is calculated to be the average welfare over the entire population of potential patients, equal to

$$\frac{\lambda_j}{-\alpha} (w_t^j = \frac{\ln G(\lambda)}{-\alpha} = \frac{\ln(e^{\lambda_j} + 1)}{-\alpha} \approx \frac{\lambda_j}{-\alpha}).$$

To estimate such a model, all that would be required is to regress the log-odds ratio of compliance on observed drug characteristics (including price):

$$\ln \left( \frac{c_s_j}{1 - c_s_j} \right) = \lambda_j \cdot \beta + \alpha_i \cdot \text{PRICE}_j + \xi_j \quad (8)$$

where $c_s_j$ is the compliance share associated with drug $j$. Consistent estimation of (8) can be achieved using OLS (under the assumption that price is exogenous), or, as Berry (1994) suggests, using instrumental variables for price which reflect the marginal cost of $j$ or elements of competition facing $j$ which are unrelated to $\xi_j$. To evaluate $W$ in this case, we would calculate the predicted values for $\lambda_j$ resulting from estimation of (8) and plug these values, along with therapeutic category prescription shares, into (7) for the two periods under consideration.
Calculating welfare in the presence of patient selection into drugs by their physician is somewhat subtler. We must account for selectivity in two distinct ways, first when estimating the parameters determining underlying compliance and second when calculating the average utility received by patients. Consider the following procedure:

1. For any set of products and assumptions about physician behavior, calculate the set of conditional densities, $g^*_{j}(\mu)$. Under (5), this requires estimation of $\delta_{MD}$ and $F(\mu; \sigma)$, which is feasible using data on overall physician prescription behavior.

2. For each drug, solve for $\lambda_j$, the average valuation for that drug over the entire patient population. To do this, invert the compliance share equation accounting for the selection distribution $g^*$ from Proposition 2:

$$\ln\left(\frac{c_j}{1-c_j}\right) = \sum_{\mu_j} \Delta_j + \mu_j \right) g^*_{j}(\mu_j) d\mu_j$$

(9)

Note that (9) involves the functional relationship (one-to-one mapping) between the compliance share for drug $j$ and $\lambda_j$, conditional on the derived distribution, $g^*_{j}(\mu)$.

3. Over the drugs in the sample, regress $\lambda_j$ on observed drug characteristics, prices, and measures of compliance cost, instrumenting for factors which are associated with unobserved patient compliance (such as price).

4. If the distribution of $F(\mu; \sigma)$ is estimated, minimize the GMM objective function of the regression in (c) over $\sigma$. This minimization yields the parameter vector $/\alpha, \beta, \sigma /$ as well as the predicted values for $\lambda$.

5. To calculate welfare for a given product set, one must calculate the expected welfare for patients who receive a prescription for a given drug:

$$w^j_i = \sum_{\mu_j} \Delta_j + \mu_j \right) g^*_{j}(\mu_j) d\mu_j$$

(10)

Note that (10) yields a higher estimate of welfare than the estimate based on no selectivity $\frac{\lambda_j}{-\alpha}$, since (10) accounts for the fact that those patients who actually receive a prescription for drug $j$ tend to have higher valuations for that drug than the average patient in the overall population.

6. Calculate (7) using the vector of conditional welfare estimates from (10) and the therapeutic category prescription shares.

This procedure yields consistent estimates for the average value of all drugs in the market, the
substitutability between those drugs at the physician level, and the parameters of the distribution allowing for correction of sample selection due to physician maximization. All that is required to perform this calculation is data at the individual or the product level on prescription compliance, along with characteristics of the drugs and their prices. For individual data, insurance information can be directly modeled as an interaction with price; in the case of product-level data, all that is required is information about the distribution of insurance by product category. It should be noted that while our model is general enough to accommodate any specific model of physician behavior (exploiting Proposition 2), the welfare measure will depend on the model chosen for physician behavior. In other words, by focusing on compliance data, our model frames a relevant economic decision for the patient; however, a model of substitution between drugs at the physician level is still required in order to understand the diffusion process and its welfare benefits.

V. Concluding Remarks

The basic intuition driving this line of work is easy to state: the welfare of any given economic agent cannot be properly assessed when the choices that determine her actual consumption are made by other agents with different information or incentives. This is clearly the case with pharmaceuticals, where physicians retain authority over the choice of which drug to prescribe. Attempts to evaluate the welfare gains from pharmaceutical innovation cannot simply rely on the shares of each drug in the market. Estimating a discrete choice model based simply on prescription shares would confound the preferences of physician and patient in the resulting welfare measure. This basic conundrum has seriously hampered efforts to conduct systematic studies of welfare in this all-important area.

This paper suggests that the patient’s compliance decision may provide a key ingredient in addressing the wedge between physician choices and true patient welfare. Incorporating the compliance decision into an augmented two-stage discrete choice model may indeed provide a way of uncovering the “true” preferences of patients, and hence the means to compute the “correct” welfare benefits accruing to patients. Furthermore, this approach may be usefully applied to other areas in the economics of health care where similar problems occur, namely examining those specific margins where patients do exercise choice. For example, the behavior of patients who refuse
treatments of various types reflects information about their preferences; evaluating the welfare benefits from invasive health care technologies could focus around such decisionmaking.

However, the implementation of the approach put forward here requires detailed data on prescription compliance at either the individual or the product level. We are currently exploring several options in this regard. Unfortunately, most data sources examine patient or physician behavior individually but not in concert; as a result, there are no public data sources which “track” prescriptions from the physician’s decision through the patient’s compliance. We are therefore investigating proprietary data sources collected by private firms for use in the health care sector, which may allow us to carry out the analysis outlined in this paper. However, given the potential importance of this line of research for the evaluation of public policy, we urge and would like to strongly encourage the systematic gathering and publication of data on prescription compliance by public institutions; such data efforts should significantly foster innovative research in this area.
BIBLIOGRAPHY


Hellerstein, J. (1996), The Importance of the Physician in the Generic versus Trade-Name Decision,


Figure A
Two-Stage Model of Prescription and Compliance Behavior

Physician Prescription Decision:

\[
\underset{j \in \{0,..., J\}}{\text{MAX}} \quad V_{ij}^{\text{MD}} = X_j' \beta_{MD} + \alpha_{MD} \text{PRICE}_j + \xi_j + \varepsilon_{ij}^{\text{MD}} = \delta_j + \mu_{ij}^{\text{MD}}
\]

Patient’s Compliance Choice

\[
\underset{j \in \text{COMP},0}{\text{MAX}} \quad V_{i,\text{COMP}} = X_j \beta + \alpha_i \text{PRICE} + \xi_j + \mu_{i,j} + \eta_{i,\text{COMP}} = \lambda_i + \nu_{i,j}
\]

\[
V_{i,0} = 0 + \nu_{i,0}
\]