STRUCTURAL ADJUSTMENT, USER FEES, UTILIZATION AND QUALITY OF HEALTH CARE SERVICES

“...User fees, charges imposed for using a health clinic or attending school, have led to increased illness, suffering and death when people cannot pay for health services, and decreased school enrollments when poor families can no longer afford to send their children to school. In a tragic example in Zambia quoted by UNICEF, a researcher observed a 14 year boy with acute malaria turned away from a health clinic for want of a 33 cent registration fee. According to the report, “within 2 hours, the boy was brought back dead”. Soren Ambrose, 2001

User fees and cost recovery schemes have been perhaps the most pervasive topic for debate about the impact of structural adjustment on health. The World Bank has promoted such schemes to achieve the dual goals of generating financial resources for underfunded programs and to encourage better quality and more efficiency. Cost-recovery schemes, the World Bank asserted, can succeed in achieving these goals, only if governments take complementary measures to 1) ensure that services are accessible and of reasonable quality, 2) to redirect the freed revenues to underfunded health programs that provide public benefits and into increasing the number and quality of facilities to serve the poor and 3) to provide protection to the poor who cannot afford new or higher charges [Akin, 1987].
However, throughout the last two decades cost-recovery and user fees schemes have faced increasing criticism that they excluded vulnerable social groups from vital health services, protection systems of the poor have been ineffective, revenue collection has fallen short of target levels and the efficiency gains predicted have failed to materialize. Between appreciation and criticism, the literature provides a contradictory accounts on the effectiveness and implications of these schemes.

**USER FEES AND UTILIZATION**

Careful analysis of the literature reveals that four arguments are used by the proponents of the World Bank’s reforms to justify cost recovery. First, the price elasticity of demand for health care is low, especially if the introduction of user fees was associated with improved quality. Hence, they contend, gradual and proper implementation of cost recovery along with improved quality of service can avoid problems of underutilization and economic inaccessibility of health care. Second, people, including poor people, are already paying a considerable share of the costs of health services in developing countries, which amounts in some countries to 50 percent of the total health spending. Furthermore, free-of-charge services in developing countries are never free and entail hidden and unofficial costs, paid basically by the poor. Third, in developing countries, including the poorest of them, people are willing to pay for better-quality services. Fourth, the failure of protection schemes of the poor occur, in the majority of cases, to half-hearted reforms which do not set up clear and transparent accountability systems.

Opponents of the cost recovery respond to these arguments by stressing that the World Bank’s claim of price elasticity is not true and were refuted by a huge body of literature, which revealed that demand for health care is highly price-elastic. Furthermore, the willingness-to-pay argument was described to be “reductionist” since it tells little about the ability to pay. The argument has been advanced that cost-recovery has resulted in a significant reduction of health care utilization by the poor and that such a reduction is not temporary.

The following sections include a review of literature available about the issue of the relationship between introducing cost-recovery schemes and accessibility to and levels of health care utilization in developing countries.

In one of the earliest studies in this field was undertaken by Heller [1982] in Malaysia. The study presented a theoretical and econometric model of the demand for medical care, and empirically tested whether household demand for outpatient and inpatient care was sensitive to its cost in time and financial resources. The study revealed that the demand for care was highly inelastic to cash price. Cash price did not appear to be a factor differentiating users from non-users of modern medical care, whether outpatient or inpatient. Likewise, the total quantity of outpatient services consumed was not significantly influenced by the cash price. It was
indicated that the estimated elasticities of demand ranged from \(-0.15\) for public outpatient clinics to \(-0.04\) for total outpatient demand.

Nevertheless, consumers were suggested to be clearly responsive in their choice among medical alternatives to the relative cash prices of private and public outpatient clinics. The cross elasticity of demand for public care due to changes in the private outpatient prices was estimated to be \(+0.15\). It was suggested that because the mean and the variance of public sector prices were quite low relative to private sector prices, variation in public sector prices had a smaller impact on the demand for private outpatient clinics. Among those households that had used both kinds of clinics, a higher perceived ratio in the price of public relative to private care appeared to reduce the probability of a public clinic visit. Relative price effects were also suggested to be relevant in differentiating those households that had used a traditional practitioner in normal circumstances or in the event of a severe illness. The cross elasticities of demand for traditional care due to a change in modern outpatient prices were \(0.05\) and \(0.17\) respectively. The demand for inpatient care was indicated to be inelastic to cash prices, particularly where the rate schedule was deliberately subsidized for lower income groups.

Similarly, a study by Akin et al. [1986] in Philippines showed little significant effect of cash prices on the volume and the choice of adults’ demand for health care. This finding was suggested to have several meanings. First, it might reflect real inelasticity for cash prices, which was indicated in the study by Heller [1982] in Malaysia. It was found that private clinics and hospitals charged over 28 times as much as government clinics and public hospitals and that the cost of a visit to traditional healers was 6 times higher than a visit to public clinics. In this context, the insensitivity to price suggested, the author asserted, that public services might have a fair degree of latitude for manipulating their charges to help offset operating costs. Third, since the model used in the study did not control for the severity of illness or the quality of service, it was suggested that the apparent insensitivity to price might be at least partially caused by the willingness to pay more to get higher quality care. Alternatively, it was indicated that private practitioners and traditional healers might informally cut their prices below those reported by using sliding fee scales. Distance was shown also to have not a significant influence over demand for health care. The authors suggested, therefore, that further investments in reducing distance to and the prices of rural health clinics to zero was not required and that there was a room for introducing user fees to public health facilities.

The assertion of these two studies that the demand for health care is inelastic for cash prices led a principal health economist of the WHO to estimate what should consumers in poor countries pay for publicly-provided health services. In the second half of the 1980s, there had been an accumulating evidence from a body of research that reported very small and sometimes positive price effects, most of
which were statistically insignificant [Akin et al, 1986; Schwartz et al, 1988; Birdsall and Chuhan, 1986; Heller, 1982; Ellis, 1987; Shepard, 1986; Bitran, 1986].

These studies provided the theoretical underpinnings of the cost-recovery schemes suggested by the World Bank in its agenda for reforming health systems in developing countries in 1987 [Akin, 1987]. They were extensively used to indicate a high willingness and ability to pay for health services which is not strongly influenced by price level within the ranges studies, or dependent on household income, and thus to support the introduction of user fees on both efficiency and quality criteria. The argument was advanced that equity would be better served by the expansion of services and the improvement of their quality than by maintenance of low prices.

However, these studies were heavily criticizes for drawing conclusions that were not supported by an empirical evidence, even from their results. For example, Non of the variables price, income or quality was found to influence utilization in either the Heller [1982] or Akin et al. [1986] study. And quality was inadequately considered in both studies. Furthermore, the Akin et al. [1986] study was criticized to draw observations from a very limited income range and the rather ambiguous evidence regarding the importance of income level in the Heller [1982] study had been largely ignored.

Meanwhile, there had been a number of studies providing conflicting conclusions and asserting that prices do affect demand for health care. [Mwabu, 1986; Gertler et al, 1987; Alderman and Gertler, 1988]. Furthermore, studies on the demand for medical care in industrial countries uniformly concluded that prices were important determinants of utilization of medical care. The most important and comprehensive of these studies is the Rand Corporation's National Health Insurance Study (HIS), which was a five-year controlled, randomized trial experiment conducted in five sites in the United States, involving over 20,000 individuals [Manning et al, 1987]. The HIS study provides overwhelming evidence that prices are statistically significant determinants of utilization of health care. Price elasticities are found to be on the order of -0.2. Moreover, the HIS results are on the low end of the estimates of price elasticity from the non-experimental literature, which finds statistically significant price elasticities ranging from -0.2 to as high as -2.1 [Rosset and Huang, 1973; Davis and Russel, 1972; Phelps and Newhouse, 1974; Goldman and Grossman, 1978; Colle and Grossman, 1978; and Newhouse and Phelps, 1974 and 1976].

Since the price elasticity is known to increase with the income elasticity, the demand for medical care should be more income elastic in the poorer, developing countries than in the richer, industrial countries. Engel curve estimates for medical care in Birdsall and Chuhan [1986] and Musgrove [1983] report income elasticities close to unity, whereas income elasticities between 0.2 and 0.3 are typically found for
The World Bank explained this contradiction asserting that the models of demand for medical care in developing countries were misspecified. The studies, the World Bank indicated, typically modeled demand as a discrete choice, with the price effect specified to be independent of income. This assumption was described to be restrictive since it is expected that the wealthy would be less sensitive to price differences among providers than the poor. Another explanation, according to the World Bank, would be that many of the studies of demand for medical care in developing countries used data sets of dubious quality. Finally, it was pointed that these studies specified time prices as non-monetary nuisance parameters in the utility function, implying that their coefficients reflected the marginal disutility of traveling. This was indicated to be in conflict with several studies that pointed out that time prices should enter via the budget constraint [Becker, 1965; Dor et al., 1987; Gertler and van der Gaag, 1988].

The study derived a discrete model of the demand for medical care from a theoretical model that implied a natural interrelation between price and income. The basic logic behind this model was that, in the context of a discrete choice model, if health was a normal good, then the price elasticity of the demand for health care must decline as income rises. This implied that the model in previous discrete choice studies that restricted the price effect to be independent of income were misspecified.

The models were estimated with data from the 1985 Cote d'Ivoire's Living Standards Survey and the 1985-86 Peruvian Living Standards Survey. These identical multipurpose household surveys were designed to measure socioeconomic factors relevant to the standard of living. Researchers collected detailed information on individuals' illnesses and utilization of medical care over the four weeks immediately preceding the interview, in addition to many socioeconomic variables relevant to the demand for medical care, such as income, family structure, and education. Surveys also collected community-level information in rural areas. For each village, researchers collected information on travel time to the nearest available medical facility of every type and on village-level agricultural wage rates for males and females.

The results showed that the price elasticity of demand fell with income. Adults' and children's demand for both clinic and hospital care was indicated to be more elastic at lower income levels than at the highest income levels. For example, the adults' and children's demand for clinic and hospital care in the bottom three quarters of the income distribution was in the price-elastic region and these in the top income quartile was well into the inelastic region. Furthermore, it revealed that children's demand for both clinic and hospital care was more price elastic than was adults'
demand and that the difference was smaller in the lower income groups but was substantial in the highest income group.

Figure 145. Arc Price Elasticities of demand for health care for Adults and children in Rural Cote d'Ivoire

The study estimated also travel time elasticity and results were very similar to that of the estimates of price elasticity. The estimates of elasticity showed individuals in the bottom three fourths of the income distribution to be much more sensitive to the opportunity cost of time than richer individuals (those in the top quarter). Moreover, children's utilization of medical care was more sensitive to time than adults' utilization and demand became slightly more time elastic as income rose over the bottom three income quartiles. The travel time elasticities are small relative to the price elasticities, which is unlike the Ivorian results.
The study concluded that user fees is regressive in the sense that they reduce utilization of medical care by the poor substantially more than by the rich. Furthermore, user fees reduce the utilization of medical care by children more than they will reduce utilization by adults. User fees can, however, generate substantial revenues without adverse effects on utilization in relatively better-off communities. It implies also that the opportunity cost of time is a bigger barrier to health care for poorer individuals than it is for richer individuals. Poorer individuals can less afford to lose productive time than can the rich. Therefore, increasing the supply of health care facilities in poor areas, the World Bank suggested, is a sine qua non for improving access. In other words, if improving the poor’s access to medical care is a primary goal of social policy, providing the care free of charge is simply not enough.

Figure 146. Income elasticity of demand for health care for adults and children in Peru

Elasticity of demand for health care was an issue for analysis in a study from Mexico, which examined determinants of household health expenditure, placing a special emphasis on the impact of household income on health expenditure, elasticity of health care expenditures with respect to income for different income group and according to health insurance status. The study revealed that monetary health expenditures were sensitive to changes in household income levels, and the elasticity of monetary health expenditures with respect to income was largest for low-income uninsured groups. It was suggested that in times of economic crisis and recession, these households reduce cash expenditures on health care by proportionally more than higher-income and insured households.

The study indicated also the importance of the health benefit status of the household in determining cash expenditures on health care, which is in turn determined by labor market participation in the formal or informal sector of employment. Monetary health expenditures were also found to be sensitive to changes in in-kind income. Nevertheless, the size of the effects of these changes
was indicated to be much smaller than the estimated impacts of monetary household income.

With respect to residence, rural households were more likely to make cash health expenditures than urban households, which was suggested to reflect the lower availability in rural areas of health clinics at which free services might be obtained, compared to urban areas. One major conclusion was drawn from the results that with economic crisis, even if efforts were paid to make health care available to large sectors of the population, the health care expenditures of the low-income uninsured groups adjust by more than those in the upper-income and insured groups.

A huge body of research of country experience with user fees and cost-recovery supported the model estimated by the World Bank. These studies reported significant decline in utilization rates of health care facilities, shifting to care seeking patterns towards traditional healers and low-cost alternatives of health care.

Yoder examined in Swaziland the impact of increased fees on overall patient use of health services, on types of services (curative vs. preventive) affected, and in health services utilization by higher paying and lower paying groups. He reported a decline in attendance of 32.4% percent in governmental facilities and an increase in mission facilities by 10 percent. Combined, the government and mission facilities experienced a decrease in attendance of 17.4 percent. This decline even increased over three months to reach 28 percent.

The study indicated that mission facilities and traditional healers are being substituted for some mission facilities. More importantly, the greatest decline in utilization was reported in visits for childhood illnesses, BCG and DPT immunization, each showing an average decline of -16, -19, and -24 percent, respectively. Musculoskeletal diseases such as arthritis and rheumatism showed the lowest decline (-1.2%) in average attendance. It was suggested that the decline in utilization was dominated by people who, prior to the fee change, chose the least costly source of health care (the nearest governmental facility) because they had the lowest incomes and consequently are the ones who could least afford a fee increase. The assertion that “people are willing and able to pay for health services, the study concluded, did not appear to be the case in Swaziland.

In the later 1980s, two other facility-based longitudinal studies in Ghana [Waddington, 1989] and Zaire [De Béthune, 1989] reported reduced utilization after the introduction of user fees schemes and that this decline was substantial and worrisome. They indicated also that the revenues generated from the new fees were not entirely reinvested in the facility to improve quality. Using these studies, some policy analysts [Kanji, 1991] draw a connection (or logical leap) and concluded that health facilities which introduced user fees would experience a drop in overall utilization and that the poor would be hit hardest.
In Kenya, two studies reported a substantial decline in health care utilization after the introduction of user fees in 1989. This decline prompted government to suspend the fees for approximately 20 months. Comparison of the levels of utilization before and during and after the introduction of user fees revealed interesting findings.

The first study [Mwabu, 1995] was carried out in two distinct districts: Kiasumu in Western Kenya and Embu in Eastern Kenya. The study indicated that the introduction of fees in the public health sector increased the average cost of medical ailments in all categories of government facilities, except in dispensaries, where fees were not levied. The number of outpatient visits in government health centers, where fees were introduced, fell by 52 percent. Furthermore, the number of patients using government dispensaries, where fees did not rise, increased by 5.9 percent. After suspension of fees, the number of visits to government health centers increased by 41 percent and to dispensaries by some 4 percent.

Demand in the private sector declined by 32 percent after suspension of cost sharing. This drop in private care utilization was attributed to shifting of users to governmental clinics. It was indicated also that the user fees forced some 20 – 26 percent of the patients out of the modern health care system. This was suggested to be the percentage of patients who did not seem to have visited any of the facility categories during cost recovery.

The study gave also a clue on where patient got the money to pay for medical care costs. It revealed the following sources for individuals interviewed in the community and in the health facilities respectively: salary or wage income of a family member (41.2%; 35.3%), loans from relative and friends (21.1%; 34.2%), proceeds from livestock sales (13.2%; 3.5%), and other sources (4.4%; 1.2%).

The second study [Mbugua, 1995] collected data through a household longitudinal survey in Kibwezi, a poor rural region in Kenya. Prior to the introduction of fees, attendance at the hospital and the health centers was higher than at the dispensaries. Attendance at all facilities was indicated to fell sharply after fees were introduced. Attendance at the hospital and health centers, which charged fees, stayed depressed until the registration fees were lifted. Dispensaries, which continued to provide free services, soon rose to above pre-fee levels. Likewise, utilization of inpatient hospital services was shown to substantially decline. For example, there were a 12 percent fewer general hospital admission and 24 percent fewer general inpatient days. The average length of stay fell from six to five days. Finally, maternity admissions were 12 percent lesser than pre-fee levels. With respect to utilization of preventive services, which continued to provide free services, there was no significant change. For example, the utilization of child health services fell by 4 percent, the use of family planning services rose by 17 percent and the attendance for antenatal care increased by 6 percent. After fees were removed, there was a 52 percent increase in
the number of outpatient visits made to government facilities. Simultaneously, utilization of mission and private clinics fell by 36 percent and the use of government dispensaries remained the same.

However, a criticism of such a conclusion was advanced in the early 1990s that while these studies reported the response of health facility users to changes in price, the issue of quality of service had been neglected. In this regard, the critique went this way: it was should not be surprising that when user fees were introduced, or increased, without improvements in quality, utilization dropped. People were asked to pay more, yet received the same services; thus they were receiving less for their money. The impact of user fees on demand for health care was attributed to the wrong ways through which cost recovery schemes were introduced and managed.

For example, a study of cost recovery in Papua New Guinea [Thomason, 1994] revealed several of these problems, which were observed at all levels. First, there was neither a standard exemption policy nor a consistent approach to the charging. Second, several centers had imposed other charges, such as charges for transporting patients, including referral cases, in the health center vehicles, which is contrary to the official policy. The study reported that people were being charged large amounts by the mortuary to get relative's dead bodies from the mortuary. It was suggested that some “enterprising” health staff might have been taking advantages of the implementation of fee collection to supplement their incomes.

With respect to management of fees, the study indicated serious inconsistencies. At the lowest service level, the aid post, there was no accountability and the supervising health center, in some cases, did not even know the level of the aid post charges. Records of aid posts takings and even level of fees were not available at any of the supervising health centers. In the aid posts themselves, there were no records maintained of the fee collection, nor reporting of this to the supervising health center. Focus group discussions revealed that the usual method for collecting fees was that the clerks ask for payment from patients in the ward and if the patient cannot pay, say they will return on the following day to collect the money. Health workers reported that it was common for patients to run away before the second visit of the clerk. Non of the money collected was retained by the facility and inability to buy even small things for the operation of the hospital such as mops and brooms was reported.

Needless to say that such a way of introducing and managing fees for services is undermining to the access and the quality of health care. With such pattern of management, which is not uncommon in developing countries, it is not surprising when utilization decreases by the poor, since there is no exemption policy, and by the rich, because of the poor quality of service. The importance of management of user fees to improve quality is further supported by experience reported from facility based longitudinal studies in Benin, Sierra Leone and Guinea, which showed
that when fees were accompanied with a notable improvement in quality of care, overall utilization did not drop, but increased [Knippenberg, 1990].

Whether quality of care is a determinant of health care utilization under user fees programs was a matter of empirical experimental research in Cameroon. The study used three treatment health center areas, to which the fee/quality policy was introduced and two control health center for a total of 5 study health center areas located in the same geographic areas. The results of logistic regression analysis of health care utilization in the treatment and control areas showed that the probability of using the center decreases with higher income in the treatment area and increases with higher income in the control area. It was revealed that the significant group variable indicated that the probability of using the health center was higher in the treatment areas than the control; and the significant interaction term indicated the effect of income among the two groups was different. It was stressed that the group variable was highly significant in the lowest quintile and highly insignificant in the highest quintile. This was suggested to indicate that the effect of treatment (i.e. user fees plus quality policy) was most felt among the lower income groups. Furthermore, the probability of the population using the health center was not shown to significantly increase after introducing fees and quality improvements. The results revealed also that the probability of poor people in the treatment areas of using the health center increased proportionally more than the rest of the population.

A second study from the Niger [Diop, 1995] demonstrated that the combination of cost recovery and quality improvements increased access to quality health care for rural populations in general, and the rural poor in particular. Furthermore, it provided an evidence for the assertion that for access to quality health care for rural populations to be sustained, cost recovery should be accompanied not only by quality improvement measures, but also by cost containment measures.

Using a quasi-experimental design in three districts, the study compared between three methods of financing health care, each implemented in one district. The status quo of free care was maintained in the control district of Illéla and two alternative methods of health financing were introduced in the other two districts. The first method, a fee-per-episode of illness was instituted in the district of Say. The second method was introduced in the district of Boboye and included a type of social financing in the form of a local, annual tax paid by the district taxpayers and a small fee-per-episode to be paid by users in order to moderate the use of public health facilities.

Results from the pilot testing of this study was published in 1995 and showed that in the district of Say, where the fee-for-service method was implemented, the number of initial visits declined slightly but the total quantity of care demanded, measured by the total number of consultations (including follow-up visits),
increased significantly. At the other extreme, the number of initial visits increased by nearly 40% in the tax + fee district of Boboye. Moreover, the total quantity of care demanded increased by 70%. The greater increase in the quantity of care demanded in the tax + fee district compared to the fee-for-service district was due largely to the fact that utilization in the tax + fee district increased among villages where health facilities were located, while it declined mainly for those villages in the fee-for-service district.

The magnitude and direction of changes varied among specific socio-economic groups in the three districts. In the control district of Illéla, where no reforms were introduced, utilization of public health facilities deteriorated among target groups, as well as for the general population, particularly among villages with no health facility; however, there was little change for children. In the fee-for-service district of Say, small changes in utilization among specific target groups were not statistically significant and utilization among the poorest 25% remained at relatively low levels. In the tax + fee district of Boboye, significant improvement in the utilization of public health facilities was observed among children and women, and residents of villages without a public health facility. Among the poorest 25%, the rate of utilization of public health facilities doubled. These positive changes in the tax + fee district of Boboye are statistically significant.

In general, quality improvements introduced with the implementation of the pilot tests – in particular, the improved availability of drugs -stimulated demand for health care at the public health services in the two test districts, by comparison with the low levels of use prevailing prior to the start of the tests and with the declining trend which continued in the control district. In other words, the positive effects of the quality improvements cancelled out the negative effects of the introduction of user fees. The role of quality is confirmed by patient exit interviews undertaken in the three districts which revealed that the availability of drugs was the main reason why utilization of public health facilities increased in the test districts.

Quality improvements introduced with the pilot tests led directly to two important shifts in use of services. Before the introduction of cost recovery in the test districts, home care and the purchase of drugs from the informal market were very prevalent health care-seeking practices. Econometric analyses of baseline and final survey data confirm that the net effects of quality improvements, through higher availability of drugs and the introduction of payments at public health facilities, were translated into shifts in the demand for health care from the informal sector to the formal sector in the two test districts of Say and Boboye. In contrast, the use of informal providers increased in the control district of Illéla during the test year [Ellis and Chawla, 1994].

Results of the pilot tests also demonstrated a positive shift in the use of preventive services in one of the test sites. Logistic regression analyses of baseline and final
prenatal care data confirm that no statistically significant changes in prenatal care behaviour were observed in the control district of Illela and the fee-for-service district of Say. In contrast, utilization of prenatal care services increased significantly in the tax+fee district of Boboye during the test Year. Since fees were not charged for prenatal care either before or during the cost recovery test in Boboye, this increased utilization is likely to have been stimulated as a by-product of increased use of curative services in health facilities in Boboye [Diop Yazbeck and Bitran R, 1995].

The final results of the study was published 5 years later in 2000 and confirmed the results of the pilot testing [Chawla and Ellis, 2000].

The probability of a patient visiting a formal provider increased in Boboye following the policy changes, even though these changes meant a greater financial outlay for the residents of Boboye in terms of a moderate fee-for-service and an indirect tax specifically earmarked for financing health care. This result could have been brought about by one or both of the following reasons. First, the policy changes also included significant quality enhancements which, in the case of Boboye, began three years before the changes in health care financing were implemented. The quality-improvement initiatives included increased drug availability as well as improved management, supervision and training. The probability of visiting formal providers is likely to have been positively affected by these quality enhancements. Secondly, after incurring an ex-ante fixed cost in terms of the indirect tax, the marginal cost of 50 FCFA per visit to a formal provider is low, and demand is likely to increase irrespective of any other changes that might have taken place. In absence of complete data on the quality variables and on out-of-pocket expenditures before the policy changes, we are unable to unequivocally attribute the higher utilization rates to either the quality effect or to the moral hazard effect; both of these could potentially have brought about the observed increases in visits to formal providers. In any case, the observed changes in rates of visits were not due to changes in reported illness, which actually declined in Boboye.

The probability of visiting a formal provider did not change significantly in Say, where a fee of 200 FCFA was introduced. Unlike Boboye, quality changes in Say were introduced more or less simultaneously with the financing changes. In contrast, the probability of treatment by a formal provider when ill decreased significantly in the control region of Illela, where there were no cost-recovery or quality-improvement initiatives. The introduction of cost recovery and improvements in quality of care in Boboye and Say changed the probabilities of seeking treatment by informal providers in the desirable direction. Individuals reporting an illness in the two experimental regions were more likely to visit formal providers and less likely to be treated only at home or by healers and other informal providers. In contrast, in the control region of Illela, individuals were more likely to report treatment at home or by informal providers after the policy change.
As far as reported illness is concerned, both the univariate analysis and logit model indicate that rates of reported illness increased substantially in Say, decreased in Boboye and were virtually unchanged in Illela. Because of the short time elapsed since changes in cost recovery and quality improvements, together with the fact that rates of visits were increasing or unchanged in the two experimental regions, it seems implausible that these changes can be attributed to the policy changes. They do suggest the need for further monitoring of this important trend, however.

Overall, the results give a reasonably favourable impression of the policy changes. In neither case is there evidence of serious reduction in access or increase in cost. Particularly notable is the fact that in Say, with moderate cost sharing, the observed decline in rates of visits is statistically insignificant.

The observed increase in the probability of formal visits in Boboye is also striking. Both contrast with the control region of Illela, where visit rates fell substantially even when there was no change in price. An important caveat in using the above results to guide wide-spread health sector reforms is that while many of the observed changes lend positive support to the quality-enhancement and cost-recovery initiatives in Niger, the absolute magnitudes are rather low. Moreover, in the absence of any data on the informal provision sectors during this period of changes, it is difficult to examine any interactions between formal and informal care. This remains an important area for future research.

A study from the Nigerian state of Ogun [Vogel, 1994] empirically examined whether, after effectively controlling for quality of care, the introduction of user fees would affect utilization of care, especially of the poor, and revenue generation. Demand for outpatient health care was estimated by a multinomial probit model using data collected by households and health facilities surveys. To effectively control for the quality of care, three distinct variables were introduced in the model, which were statistically significant: (a) expenditure per person in population served; (b) percentage of times drugs are available; and (c) interviewers evaluation of the physical condition of the facility. The results showed that with quality of care effectively controlled, price of the health care from each provider is a statistically important determinant of the provider chosen. When quality was not controlled for, the magnitude of the price effect was relatively small. In addition, the income level was not significantly associated with the price of care. This finding was suggested to imply that the price response was small for the poor and the non-poor.

It was further suggested that if the quality of care in public facilities was improved to the private sector's levels, prices of the outpatients visits could be increased to the private sector's levels without affecting the utilization level. Under these circumstances, it was asserted that, non-users and about half of the private care users would be drawn to the public facilities.
Quality and Utilization

Quality of health care provision has gained an increasing attention under structural adjustment reforms of health systems in developing countries. Early reform attempts placed a greater emphasis on the operationalization of user fees and cost recovery schemes than improving quality. A recurring lesson from these initial experiences in cost recovery is that without visible and fairly immediate improvements in the quality of care provided, patients will not support the implementation of user fees [Vogel, 1988]. In many cases, user fees are collected at the local health facility and sent directly to a central authority (Ministry of Health, Treasury), creating serious obstacles to recycling the revenues to improve quality at the collection facility. A recent study on health care financing in Ogun State, Nigeria, reconfirmed earlier studies that quality is a major determinant of patient choice of health care providers, strengthening the recommendation that quality improvements should accompany increases in user fees [Denton et al., 1991]. However, the cost implications of these quality improvements are unknown. Quality improvements may also have distributional consequences. Some population groups may be more willing to pay for selected quality improvements than others.

The following sections of this chapter examine the assumption that improving quality of health care will ameliorate the reported negative impact of cost-recovery on utilization of health care in developing countries. This issue is crucial to verify whether the problem of underutilization of governmental health services widely seen in many developing countries is a price or quality effect. In this regard, the empirical research is critically reviewed to ascertain several assumptions imbedded in the debate about this issue. These include providers' (government in the case of most developing countries) behavior in improving quality and the consumers' utilization patterns and their willingness and ability to pay for quality health services.

HEALTH CARE DEMAND STUDIES

With the introduction of cost-recovery, a considerable body of research in health economics has attempted to estimate the importance of financial factors—such as time price, cash price, travel price, and income—in influencing patient choice of provider type (e.g., public, private, mission, traditional). While focusing on financial determinants of choice, health care demand studies have also examined other factors, such as the quality of services, seasonality and patient and household sociodemographic characteristics. Although, the multivariate approach used in econometric health care demand studies resolves many of the weaknesses faced by facility-based studies, it is limited in the extent to which it can examine specific attributes of quality of health care.
Extensive reviews of health care demand studies for developing countries already exist [Bitran, 1988; Creese, 1991]. This review extends the previous literature by taking a closer look at how the particular issue of quality of health care is addressed. The discussion builds on the framework and updates the findings presented in Wouters [1991, 1992].

Many health care demand studies recognize that health care services are valued for their contribution to health rather than for something in and of themselves. Health care services, in combination with several intermediate (proximate) determinants of health such as environmental sanitation and nutrition, have a direct influence on health. With this approach, the demand for health care, otherwise known as the "derived demand for health care," depends on the "health" production technology [Wouters, 1992]. [Wouters [1992] provides a critical analysis of the empirical attempts to estimate the "derived demand" for health.)

The individual's consumption choice is to maximize his or her utility subject to the health production technology and the budget constraint. This decision rule implies that individuals want value for their money. What do patients value in a health care provider? Is it technical competency, a pleasant and friendly atmosphere, accessibility, or other attributes? Creese [1991] makes the general observation that econometric health care demand studies are still in their relative infancy in how they model the importance of user perceptions of quality of care as an influence on service use.

Two general categories of demand studies exist: (a) those that estimate parametric functions of unobservable quality, and (b) those that estimate coefficients for observable proxies of quality. The first category can only imply that quality plays a role in utilization. The specific role of quality cannot be distinguished from other unobservable variables. The second category can identify whether specific dimensions of quality influence utilization patterns.

Selected health care demand studies are reviewed below in terms of what is known about the effects of quality on health care use. Studies that do not have data to proxy quality are briefly discussed first [Bitran, 1989a, 1989b; Gertler et al., 1988; Mwabu, 1984]. Studies that include observable measures of quality are examined next [Akin et al. 1981, 1986a, 1986b; Denton et al., 1991; Heller, 1982]. More recent studies are beginning to pay particular attention to how to deal with multidimensional aspects of quality [Ellis and Mwabu, 1991; Lavy and Germain, 1993; Mwabu, et al., 1993].

**UNOBSERVABLE QUALITY**

For those studies where data on quality were not available, various methods were used to capture its effect. Mwabu [1984] and Mwabu and Mwagi [1986] used provider-specific dummy variables to capture the effects of unobservable provider
attributes as perceived by the patient. Some of these unobservable attributes might relate to quality, but not necessarily. The clinic-specific constants were constructed by interacting patients' preferences with each of the classes of health providers. Mwabu found that these dummies greatly increased the explanatory power of the model. Patients appeared to place a lower value on traditional type of treatment than they placed on the treatment given by any of the other types of facilities. Mission clinics were the most preferred class of providers. Mwabu stated that since mission clinics did not have the shortage of drugs faced by other types of clinics, this empirical result was probably capturing drug availability. Using the preference dummy variables, Mwabu ranked facilities on the basis of patient preference as follows: mission clinics, government hospitals, private clinics, government clinics, pharmacies, and traditional healers. He claimed that this ranking is consistent with a ranking that one would obtain based on facility capacity to treat a range of illnesses and reliability of drug supply. Through simulations using the results of these health care demand estimations, he found that user fees could generate welfare gains if they were used to improve the quality of medical services [Mwabu and Mwangi, 1986].

In the Gertler et al. [1988] and Bitran [1989a, 1989b] studies, quality was also unobservable. Quality was included as a parametric function of its observable determinants, which include patient socio-demographic characteristics. Provider-specific coefficients created provider-specific intercepts, which allowed for a certain core dimension of quality to vary by alternative. In addition, slope coefficients reflected how the provider's quality varied by each socio-demographic characteristic. In other words, when individual and observable provider characteristics were the same, the only factors that accounted for differences in predicted demand for the types of providers were the coefficients associated with individual characteristics that vary by provider and the different provider-specific constant terms. These coefficients were said to capture people's perceptions of the quality of care among facilities. In Gertler's [1988] work in the Ivory Coast, the intercept turned out not to be significant. In Zaire, Bitran [1989a] interpreted the provider-specific constant term to show that health centers were viewed as being of higher quality, followed by private dispensaries and then pharmacies. In the Dominican Republic, the quality of private health services was perceived to be higher. In both studies, the remaining coefficients suggested that quality perceptions vary between individuals of different socio-demographic groups.

**OBSERVABLE QUALITY**

The study of Ogun State, Nigeria, used facility operational costs per capita, observed physical condition of the facility, percentage of year drugs are available, number of functioning X-ray machines and labs, number of support personnel and nurses per capita, and doctors per capita to proxy quality of care [Denton et al., 1991]. Of all of these proxies, three were significant: facility operational costs per
capita, observed physical condition of facility, and percentage of year drugs are available at the facility. The results of the multinomial probit estimation were then used to simulate probabilities of choosing public care, private care, or self-treatment. When public and private operational expenditures per person were simulated to double, patients moved, to a limited degree, from self-care into public care, with no change in private care usage. If only the public-sector operational expenditures doubled, patients would shift from the private to the public sector. Improvements in drug availability elicited large responses in use. Full availability of drugs in both the public and private sectors induced patients to move in substantial proportions from self-care and the private sector to public care. Improvements in building conditions also increased public-sector usage but not by much. Interestingly, if the quality of the public sector improved to nearly match the private sector at the same time that all prices doubled, the probability of self-care and private care dropped while the probability of using public care increased.

Akin and his colleagues [1981, 1986a, 1986b] explored price and quality trade-offs to some extent in three studies of the Philippines. In the study of demand for primary health care services (outpatient, prenatal, obstetric, well-baby, and immunization), quality was defined by whether or not a physician attended the patient [Akin et al., 1986b]. For a traditional practitioner, the quality variable was a dummy variable for whether he or she treated any of the five common illnesses cited in the survey. The authors found, counter to their expectations, that for adult outpatients, the quality variables were statistically insignificant. They did observe a strong movement of patients toward private physicians for more serious illnesses, which they interpreted as a quality phenomenon. They concluded that patients in the Bicol Region considered more the orientation of the facility (modern versus traditional) and type of facility (private versus public) than the type of attending practitioner.

In the two other studies of the Philippines, Akin and colleagues [1981, 1986a] were unable to control for quality of care, but they interpreted their results with quality of care issues in mind. In the analysis of child health services, they inferred from the variable for mother’s education that modern public, traditional services and self-treatment were considered to be of lower quality than modern private-sector treatment by the more educated people. In the study of adult outpatient services, they included some measure of satisfaction with health services, but they did not define it nor did they interpret it as a measure for quality. It turned out not to be significant. Akin and colleagues recognized how the omission of a proxy for quality could confound the price variable; the apparent insensitivity to price might be at least partially caused by the willingness of individuals to pay more for higher quality care.
In an earlier work, Heller [1982] proxied quality by creating variables that identified the type of provider the patient expected to see based on previous visits. These included the probability of seeing a physician during an outpatient visit, the probability of seeing a physician during a private outpatient visit, and the probability of seeing a physician during a government outpatient visit. Heller found that the likelihood of being seen by a physician had only a statistically insignificant effect on total usage of outpatient care. However, households did respond significantly to a higher probability of care from a physician rather than a paramedic in their choice between public and private clinics.

In a more recent study, Mwabu, Ainsworth, and Nyamete [1993] paid particular attention to the difficulties of examining the effects of quality improvements on utilization patterns using nonexperimental cross-sectional data. The typical expectation was that an improvement in some aspects of quality would increase demand, either by attracting new users or by increasing the number of visits by existing users. However, the coefficients on the quality variables could be biased by three types of endogeneity problems. First, although individuals might have demanded higher quality services, utilization could appear to decrease with improved quality if quality improvements were effective in treating illnesses, thus reducing overall population morbidity in the long run. Second, variables measuring observed quality (especially if measured as stocks of supplies) were actually capturing both supply and demand behavior. If demand was high such that stocks were quickly depleted, the coefficients would imply that individuals preferred low-quality facilities. Third, quality measures were not exogenous if they were collected from respondents who actually chose the service in question. Self-selection was a problem.

Their study of a Kenyan rural health district looked at treatment-related facility measures of quality including the availability of a variety of drugs and diagnostic equipment. Because of the small sample size and difficulties with multicollinearity, only a limited number of quality variables could be used in the estimations. The results indicated that their measures of quality improvements had significant negative and positive effects on demand. Demand was lower for facilities that had no aspirin, but higher for facilities with antimalarial drug shortages. The interpretation was that the high demand for antimalarials actually led to shortages (excess demand). They also found that the probability of a visit to a public facility was most sensitive and positively related to the availability of a broad number of drugs.

The quality attributes were interacted with gender variables to investigate gender-specific quality preferences that affect utilization. The absence of antimalarials was associated with higher demand, especially if the patient was male. Men might have had a higher demand for antimalarial drugs, but facilities could not keep sufficient supplies in stock. Including such interaction terms was an important first step in
understanding the distributional impacts of quality improvements. Further research is needed to look at differences among income groups and other vulnerable populations.

The authors acknowledged the difficulty of disentangling the demand effects of quality improvements from offsetting supply constraints and health effects over time. They suggested that experimental design, in which inputs are exogenously varied by researchers, might be the most promising way to isolate the demand effects of specific quality improvements. Examples of such field studies currently under way include the USAID HFS cost recovery experiment in Niger and the World Bank Indonesian Resource Mobilization Study [World Bank 1993a].

In another study of Kenya (with data from another district), Ellis and Mwabu [1991] examined relationships between willingness to pay and quality using 13 measures of physical attributes of the facility that were available from the Kenyan Ministry of Health. These 13 measures were collapsed into an overall score of quality using principal components. The score or index explained 33.1 percent of the total variability in these measures. This approach to quality measures showed that the facility-level quality index strongly influenced the choice of provider and that consumers were willing to pay significantly more for higher quality services. Simulations were also used to generate nonwelfare-neutral measures of willingness to pay for quality improvements. For a typical individual (estimated using variable means), willingness to pay was about 19.1 KShs per visit (average treatment cost for full sample of facilities is 29.7 KShs) if government dispensaries were upgraded to have the same quality index as government health centers. This adds further support to the importance of quality in influencing health care utilization patterns; however, the aggregate measure of quality does not allow one to determine preferences for specific dimensions of quality.

Lavy and Quigley [1993] attempted to measure specific quality improvements that "can pay for themselves" with increased user fees. The Ghanaian study measured directly the quality of the various options available to the consumer: infrastructure (beds, vehicles, laboratory, operating room, etc.), personnel (number of doctors, nurses, medical assistants, etc.), and availability of health services and drugs (number of hours open per week, type of services provided, stocks of 16 types of drugs). As in other studies problems with multicollinearity forced the authors to group quality variables into five measures of quality: availability of essential drugs (mean availability of ampicillin, chloroquine, paracetemol), number of medical staff (nurses and doctors), provision of basic adult and child health services (measured by the availability of a functioning laboratory, ability to vaccinate children, and the ability to provide prenatal, postnatal and child growth), the availability of electricity, and the availability of running water.
All of the quality factors had significant and large positive effects, both as individual variables and jointly, although the authors noted cautiously that the quality variables might be capturing both supply and demand effects. The coefficient on the quality variable reflected the net effect. When quality variables were omitted from the equation, the coefficients on the price variables were smaller. Elasticities of probability of choice in response to quality improvements suggested that preferences for improvement in drug availability were the largest. Improving infrastructure resulted in the least impact on probability of use. These elasticity measures are difficult to interpret since incremental changes in quality variables are hard to quantify in meaningful terms. Simulations using the estimated multinomial logit coefficients showed that when quality was improved in both the private and the public sectors, the relative changes in both sectors were about the same.

A unique aspect of the Ghana study was the attempt to specifically estimate the willingness to pay using measures of compensating variation: welfare-neutral estimates of willingness to pay (the amount of income individuals are willing to give up to remain at the original level of satisfaction). Compensating variation is a function of the utilities before and after the change and the marginal utility of income. Most previous studies focused on estimating price elasticity rather than on generating actual estimates of willingness to pay. Willingness to pay was measured as the incremental amount of income necessary to keep the individual at least as well off (constant utility) after the improvement as before. Simulations were based on estimated probabilities of use for a typical individual (variable means). Also, willingness-to-pay figures for typical individuals in three different income groups (sample mean, plus one standard deviation, minus one standard deviation) were calculated. Willingness to pay was expressed as a percentage of income, rather than in absolute terms. The results indicated that the willingness to pay (holding utility constant) increased with income, no matter what the scenario. However, the income effect was more important for the quality variables (infrastructure, drug availability, services, and personnel) than for distance or price. The willingness to pay, expressed as a percentage of income, increased as income increased. The authors concluded that the richest were more concerned with quality of health services than its costs (price or distance). Willingness to pay was highest for travel, then drug availability, and lowest for building infrastructure.

**HEDONIC PRICES AND QUALITY**

In contrast to health care demand studies, hedonic prices look more closely at willingness to pay (WTP) for specific attributes of a product or a service, some of which may pertain to quality. The basic premise underlying hedonic pricing studies is that goods and services are valued for their utility-bearing attributes [Rosen, 1974]. Hedonic pricing studies look at the markets for goods that can be completely described as bundles of several objectively measured characteristics. Hedonic prices are the implicit prices of product attributes. They are revealed to individuals
through observation of actual market prices and the specific amounts of service attributes associated with those prices. The goal of hedonic studies is to estimate demand and supply functions for product attributes. For example, market researchers have used hedonic pricing methods to identify market segments for their products; in particular, they are interested in the WTP for various attributes of a specific good and whether the WTP varies among types of consumers as described by socio-demographic and economic characteristics. This is similar to the question posed by health care financing policymakers who want to know the WTP for different types of improvement in the quality of health care services; however, the application of hedonic pricing techniques to health care services is almost nonexistent.

Hedonic pricing studies fall in the group of observed/indirect methods for estimating consumer valuations of goods and services [Mitchell and Carson, 1989]. They rely on data from actual market situations; however, for the most part, they can only infer preferences for the product attribute under consideration from consumer choices about related product characteristics. Hedonic pricing studies have often been used to corroborate the results of contingent valuation studies that fall within the group of hypothetical/direct methods for estimating consumer valuation of goods and services [Brookshire et al., 1982]. Because neither type of study is an observed and direct assessment of consumer valuations, they cannot be used as a means of internal validation [Mitchell and Carson, 1989].

This section gives a brief review of hedonic pricing techniques and discusses their potential application in estimating WTP for quality improvements to support cost-recovery initiatives. First, Rosen's theoretical model is summarized. Then, key assumptions and estimation techniques are discussed in terms of their applicability to the health care service market in developing countries.

CONTINGENT VALUATION SURVEYS – WILLINGNESS TO PAY FOR QUALITY

Contingent valuation (CV) studies attempt to directly elicit revealed preferences or willingness to pay (WTP) from respondents by describing a hypothetical market that is as close to actual market conditions as possible. CV studies involve \textit{direct observation of WTP in a hypothetical market}. Contingent valuation studies are most frequently compared with hedonic pricing studies and travel cost studies that involve \textit{indirect observation of WTP in actual markets}. In particular, these latter studies observe consumer price behavior under actual market conditions for services or commodities which are \textit{similar but not identical} to the new or changed service or commodity under consideration. WTP for new services or commodities is \textit{inferred} from observations of pricing behaviors for existing related products. Since neither approach generates actual WTP for the new service or commodity
under actual market conditions, they cannot be used as an ultimate test of internal (criterion) validity. Another related method is contingent ranking in which attitudes (rather than direct WTP) about the new service or commodity are determined. This involves indirect observation of WTP in a hypothetical market. This method is an even weaker test for the internal validity of CV studies.

There are three major types of the contingent valuation empirical studies looking at health-related issues. The first group looks at valuation of health outcomes in terms of reduction in the risk of mortality and morbidity. The second group looks at the valuation of specific health services described by their structural and process attributes. The third major type of contingent valuation studies elicits WTP information for incremental changes in a specific process or structural attribute for a specific health service. Developing-country studies fall in the third group and focus on WTP for structural attributes. Some of this empirical literature is reviewed below. This review is not exhaustive, but instead attempts to outline major trends, issues, and themes.

Contingent valuation analysis has frequently been used in the application of cost-benefit analysis to health services as an alternative to the human capital approach for valuing benefits in monetary terms. The human capital approach is based on the cost of illness method that equates the value of additional life and losses due to morbidity with forgone earnings discounted to their present value. This method is limited in that it omits the costs of pain and suffering, undervalues the illness of those who are not working or retired, and is subject to valuations emanating from market imperfections and unemployment. Moreover, the human capital approach is concerned with maximizing net output of society, while cost-benefit analysis should be directed toward the broader objective of welfare improvement based on individuals' preferences. Some researchers suggest that WTP measures of health benefits are more appropriate than human capital measures since WTP incorporates those quantities and qualities that are difficult to measure, such as the individual's preference for risk aversion and valuation of pain and suffering. However, WTP measures have other limitations, such as being driven by existing patterns of income distribution. One could simulate the effect of different income distributions by assigning weights to different income groups, but this is likely to be extremely value-laden and controversial. Discussion of how to measure benefits is quite complex and beyond the scope of this report. The intention here is to provide an introduction for the application of contingent valuation methods.

**WILLINGNESS TO PAY FOR HEALTH OUTCOMES**

Contingent valuation studies that look at willingness to pay for health outcomes for health services have measured outcome in terms of reducing the risks of mortality or morbidity associated with specific treatments [Acton, 1973; Thompson, 1986].
Similar outcome measures have been used to assess the willingness to pay for reduced health risks associated with the use of various consumer goods such as cleaning products and automobiles [Magat et al., 1988; Muller and Reutzel, 1984].

Acton's [1973] study looked at how much people would pay (in the 1970s) for the availability of mobile coronary care units. The average response was a willingness to pay $56 for a 0.002 chance that one's life would be saved in the next year, implying a value of life saved of $28,000. Reviews of this study suggest that well over half of the respondents gave irrational responses [Fischer, 1979].

In the Thompson [1986] study, 247 patients with rheumatoid arthritis enrolled in a randomized control trial were asked about their willingness to pay money and to incur mortal risk to cure their disease. Subjects were asked to think about all the ways their arthritis affected their lives and their families. They were also told to assume that a complete cure for arthritis existed and that they would have to pay the full costs of receiving it if they chose it. Respondents were asked what percentage of their household income they would be willing to pay for this treatment. Thirty-one other independent variables were considered in the analysis covering various sociodemographic characteristics, medical history, other measures of health status, and other economic costs. The rate of plausible responses was substantially higher than previous studies. Thompson suggested that substantial improvements in questionnaire design (explanatory introductions, question repetition for confused respondents, opportunities to revise earlier answers) and performance of the interviewers accounted for improved response rates. The results indicated that both the likelihood of the response and the likelihood of plausible responses increased with the education of the respondent. Income did not appear related to these WTP measures; however, personal investment measures (proxying accumulated wealth) were. Many of the alternative measures of health status were also positively related to WTP.

In a study by Muller and Reutzel [1984], 87 students were asked about their willingness to pay for reduction in fatality risk due to car crash protection. Internal validity of the CV study was checked by estimating theoretical proposed relationships, such as the relationship of WTP to respondents' disposable income, ownership of a car, accident experience, exposure to potential accidents, risk aversion and mathematical ability. The results showed enormous variation in value-of-life estimates, suggesting that the students were not able to evaluate small risks well. In addition, none of the variables posited to be related to the WTP were significant. The authors conclusions about the usefulness of such WTP questions were quite mixed.

Another study elicited consumer valuations of morbidity risk reductions associated with safer cleaning products, in terms of reducing the probabilities of suffering several types of accidents [Magat et al., 1988]. This study differs from many others
in that it estimates consumer values for a private good. WTP was elicited using the
take-it-or-leave-it approach. The contingent valuation approach was contrasted
with another type of direct, hypothetical revealed preference method known as
paired comparisons. This approach is widely used by marketing research because it
simulates actual choices in the marketplace. Respondents are given a pair of
products where the description of the current product matches the current market
situation and the description of the new product includes a new risk level and a new
price. Respondents rate, on a scale of one to nine, which of the two products they
prefer. Depending on the responses, the first comparison is followed up with
questions about the preferences for other paired comparisons. Their results show
that the contingent valuation method yields valuation estimates different from those
in the paired comparison.

To summarize, these studies suggest that it can be quite difficult for respondents to
evaluate certain types of hypothetical information, such as incremental changes in
the risk of mortality and morbidity. Results are ambiguous. Comparing results
with similar hypothetical direct observation methods cannot resolve concerns for
internal validity. Other applications of CV methods have also found that accurately
conveying risk levels and risk reductions is quite challenging [Mitchell and Carson,
1989].

WILLINGNESS TO PAY FOR A SPECIFIC HEALTH SERVICE

Another application of contingent valuation methods looks at willingness to pay for
specific services described in terms of structural and process attributes rather than
specific health outcomes. For example, a Swedish pilot study attempted to elicit the
willingness to pay for antihypertensive therapy [Johannesson and Jönsson, 1991].
In this case, since all patients were already under treatment, willingness to pay is the
amount of money that, if paid by the consumer, leads to the same level of expected
utility as without treatment [Hicksian equivalent variation]. Two elicitation methods
were used: an open-ended question, in which patients were asked to state their
maximum

willingness to pay for their current treatment for hypertension; and a discrete (yes-
no) question, in which patients were asked to accept or reject a specified increase in
user fees for their current treatment. Additional information collected included
patient subjective estimates of the risk of death from myocardial infarction or stroke
if treatment were not given and of the risk reduction under treatment; and patient
intentions to change life styles (i.e., improved diet, exercise) if user fees increased—a
measure of perceived substitutions between medical and non-medical prevention.
The availability of non-medical therapies could lower WTP. The study found that
discrete valuation questions work better that open-ended CV questions; the non-
response rate for the latter was much higher. The analysis produced estimates of
both the mean and the median WTP for antihypertensive therapy. The authors
interpret the results with considerable caution. They acknowledge that reliability
STRUCTURAL ADJUSTMENT AND HEALTH

and validity of the contingent valuation method are not yet well established. The assessment of validity is more difficult because there is no "true" value for WTP with which to compare.

WILLINGNESS TO PAY FOR STRUCTURAL AND PROCESS IMPROVEMENT

EVIDENCE FROM THE UNITED KINGDOM

In the last two sections, we look at how contingent valuation studies have been used to evaluate specific structural and process improvements in a specific health care service. Developing-country studies using contingent valuation fall in this latter group.

A United Kingdom (U.K.)–based study attempted to measure an individual's willingness to pay for public health services for elderly care provided under two different settings: National Health Service (NHS) nursing homes and hospitals [Donaldson, 1990]. Patients had been randomly assigned to these delivery sites. The study elicited information on what individuals believe the government should be prepared to pay for each type of care since this is the typical scenario in the British Health Service. To minimize strategic behavior (free riding), respondents were given information about the tax consequences of the WTP figures they gave. In looking at the individual responses, most of the patients preferred to stay in the facility type in which they were currently being cared for. However, calculation of aggregate benefits (summation of the net benefits for each group) showed that the group that preferred NHS nursing homes could potentially compensate the group that preferred hospital care and still remain better off, satisfying the Pareto optimal criteria. As reasons for their responses, patients cited factors such as privacy, freedom of choice or independence, "at home" atmosphere, and availability of immediate medical treatment. Differences in actual health outcomes between the types of facilities were not assessed. Again, as in other studies, the author acknowledges that the validity and reliability of the contingent valuation method are uncertain. Several questionable assumptions had to be made to be able to estimate WTP for those respondents who would never be willing to change delivery systems and to aggregate responses (simple summation) for making the Pareto comparison.

In another study of the U.K. National Health Service, Propper [1990] elicits information on the willingness to pay to reduce waiting time for admission by a specified amount. In a public system where user fees are not paid, one can only infer preferences by whether the patient stays on the waiting list for admission or opts out for some alternative private care setting. The study used contingent valuation methods in conjunction with the stated preference approach, in which respondents rate or rank prespecified alternatives characterized by various selected attributes. Waiting time, price, and uncertainty of admission were the focus of this study. The hypothetical context was specified as a choice between immediate treatment at some positive price in a public-sector hospital and
treatment after some positive wait in the same hospital at no monetary cost. Treatment, nursing care, and recovery were specified as identical under the two options. Average willingness to pay for reductions in waiting time was estimated. Results from the standard preference and contingent valuation methods did not differ significantly.

**EVIDENCE FROM DEVELOPING COUNTRIES**

This review found only two studies about developing-country health service issues that used contingent valuation techniques. Of these studies, only the study by Weaver and colleagues [1993] follows CV methods rigorously.

A study by Abel-Smith and Rawal [1992] was part of a larger effort to look at the introduction of user fees for public health services in Tanzania. In surveys of 893 outpatients at three referral hospitals and of 1,820 households with 11,918 persons in a population-based survey, respondents were asked, among other things, their willingness to pay for public services. Interestingly, the study notes that contrary to expectations, respondents were not reluctant to answer questions about willingness to pay. Specifically, respondents were asked to assume that waiting time was reduced to less than one hour and/or that drug supplies had been improved, and then whether they would be willing to pay a specified list of charges. Abel-Smith and Rawal did not explicitly cite contingent valuation techniques, but the survey appears to have used related techniques, such as the payment card elicitation method.

The Tanzanian study found that the proportion of outpatients willing to pay for a one-hour reduction in waiting time varied with the charge proposed. WTP fell as higher prices were suggested. Overall, outpatients were willing to pay more at referral hospitals than at district hospitals for a one-hour reduction in waiting time. In comparison, outpatients were also asked if they would be willing to pay if services were improved in general. The percentage of those unwilling to pay dropped at both district and referral hospitals. The population-based survey repeated the questions concerning reductions in waiting time and added questions about WTP for improved availability of drugs. One of the most interesting results of the study was that respondents in the lowest income groups indicated a fairly high WTP: 45 percent were willing to pay Tsh 200 for government hospitals. This is quite close to the 60 percent figure estimated for those in the highest income group who were willing to pay the same amount. Questions about health centers produced similar findings. Strikingly, the higher income groups had a higher proportion of those who said they were unwilling to pay. In comparing the two types of quality improvements, substantially fewer people said that they were unwilling to pay if drug supplies were always available than if waiting time were reduced to under one hour. Also, individuals were willing to pay higher amounts for improvements in the
reliability of drug supplies. In listing preferences for quality improvements, respondents cited improved availability of drugs most frequently, followed by attitudes of the doctors. Next came quicker service. The researchers interpreted these results cautiously, readily acknowledging that answers to hypothetical questions may not be good predictors of actual behavior.

Using some of the other survey information about individual and household preferences, attitudes, and sociodemographic and economic characteristics, the authors provided some potential explanations for willingness to pay for some of the quality improvements. In particular, the greatest challenge seemed to be to explain the higher willingness to pay for quality improvements among the poor, in light of the already sizable financial payments for transport and other travel-related costs. As noted, rural households appeared willing to pay fairly substantial amounts for quality improvements and relatively fewer rural households were unwilling to pay for these improvements—in spite of the evidence showing that the use of public services was already quite costly if one considered opportunity costs and other nonfacility-related costs, such as transportation, waiting time, etc. For example, a government visit took about three hours, a health center visit about two and a half hours, and a dispensary visit or missionary-operated facility visit about one hour and 40 minutes. Travel costs were also sizable. Although the average price of an admission to a mission hospital was reported to be lower than for a government hospital, total user costs for the government hospital tended to be higher because patients had to supply food, drugs, and other medical supplies. Moreover, the study reports that rural households have difficulty paying for services.

Taking into consideration these various data, the authors made two key recommendations. First, charges could be appropriate if the money could be used to improve the quality of services. Second, policymakers should focus on those quality improvements that might ultimately reduce the financial burden on the poor. These might include improved availability of drugs (and food for inpatients), and reduced travel costs, especially for admissions.

A study of willingness to pay for quality improvements in the Central African Republic by Weaver and colleagues [1993] closely follows contingent valuation methods. Additional documentation of this case study includes Barker [1992] and Kornfield [1992]. The purpose of the study was to determine whether a particular combination of user fees and quality improvements would increase utilization of public facilities. The results of contingent valuation methods were compared with estimations of total actual health expenditures for health services as a function of different quality improvements. Six quality improvements were considered: (a) facility maintenance, (b) supervision of personnel, (c) availability of pharmaceuticals to treat malaria, (d) availability of pharmaceuticals to treat acute respiratory infections, (e) availability of pharmaceuticals to treat intestinal parasites, and (f)
availability of pharmaceuticals to treat diarrhea. The take-it-or-leave-it elicitation method was used. Prices varied around the estimated cost of the quality improvement.

The analysis generated the median willingness to pay for each of the six quality improvements. Respondents had the highest median WTP for pharmaceuticals (especially ARI and diarrhea) and the lowest WTP for supervision of personnel. Notably, median WTP estimates were substantially larger than the estimated costs of implementing the quality improvements. Logistic regression of the WTP function for the various quality improvements indicated that WTP was significantly related to various factors in the following directions: positively related to income, negatively related to the price level, positively related to rural residence, sometimes negatively related to health status (especially for diarrheal, ARI, and STD drugs), and negatively related to cleanliness with competing nearest facilities implying that substitution was possible. Geographic differences were also observed.

The study took important steps to check internal validity and reliability. Theoretical validity was supported by obtaining significant and appropriately signed coefficients for key determinants of WTP, some of which have already been mentioned above (i.e., price decreases with higher levels of quality downward sloping demand curve in terms of price and quality). Comparisons of the contingent valuation median WTP results with estimates of median health expenditures at facilities with different quality attributes indicated greater similarities for drugs than for personnel supervision or facility maintenance. These results are somewhat clouded since it is not quite clear what expenditure components are included in total expenditures. Reliability is partially confirmed by significant chi-squared tests and a prediction rate of 30 percent.

**INTERVENTION STUDIES: COST RECOVERY AND QUALITY**

All of the types of studies reviewed in the four previous sections attempted to estimate the willingness to pay for quality improvements through indirect methods. The facility-based studies looked at longitudinal trends in utilization corresponding to the introduction of user fees, without explicit measures of changes in quality of care. Health care demand studies simulate changes in choice of providers due to changes in price and quality, using estimated coefficients derived from actual market behavior. Since they depend on cross-sectional data, the potential demand effects of quality improvements are difficult to estimate because of problems with endogeneity (interactions of supply and demand on observed quality). Hedonic pricing studies involve indirect observation of willingness to pay in actual markets. Contingent valuation studies use direct observation of willingness to pay in hypothetical markets.
In contrast, interventions studies using experimental design methods can directly observe willingness to pay for quality improvements under actual market conditions. In this case, the intervention includes the introduction of user fees accompanied by a package of quality improvements. Experimental design studies provide the most convincing evidence to test hypotheses establishing cause-and-effect relationships. However, they are often not feasible because of the substantial resource requirements; the absence of long-term commitments from governments, donors, and researchers; the inappropriate timing of interventions; and the difficulty in controlling a variety of confounding factors that may contaminate the research design (political, economic, and other events). In addition, health economists typically focus on secondary data analysis using econometric techniques. Application of experimental design methods to study issues in health economics is fairly new. It is not surprising, then, that there is only one study to review in this section the Litvack [1992] study. Discussions of both the experimental design methods and the results are included.

The study took place in the Adamaoua province of Cameroon. As part of a USAID project, drug revolving funds were initiated at each health center in the province. Revenues from drug sales were used to replenish the drug supply and subsidize other aspects of primary health care (PHC) delivery. The study looked at three main hypotheses. First, the desire for an improved drug supply was sufficiently strong that the introduction of health facility user fees would actually result in an increase in facility utilization when fees were accompanied by an improvement in the quality of care. Second, the poorest quintile would be less positively affected by the simultaneous imposition of fees and improved drug supply than the rest of the population. Third, utilization would increase where quality was improved because even with the health facility fees, the local availability of drugs did not represent an increase in total cost and may even have represented a real savings in the total cost per episode of illness when both direct and indirect costs were considered.

The results showed that the probability of using the health center increased significantly for people in the intervention areas compared to those in the control areas. The evidence also indicated that the probability of the poorest quintile seeking care increased at a rate proportionately greater than the rest of the population. Litvack suggested that travel and time costs to get prescriptions filled for the poorer households might have been high enough that they were benefiting from the local availability of drugs. Health expenditures per episode of illness did not change significantly with the intervention suggesting that individuals were using lower cost providers. Litvack proposed that improved quality in the form of improved drug supply was valued because it may actually reduce total costs (by providing efficacious treatments locally and avoiding additional transportation and time costs). For the poor, this may have represented a real savings.
Health care demand studies all strongly suggest that quality of health care is an important factor in choosing a health care provider and that patients are willing to pay for improvements in quality. Most of the studies indicated that the strongest preference was for the availability of drugs, both in amounts and types of essential drugs. Preferences for building infrastructure and type of health providers were weaker, but apparent.

Investigations of demand patterns—in this case, patient choice of type of provider—should include patient and not provider perceptions of quality. However, most studies measured quality using only structural attributes (the most easily collectable quality data) without careful investigation as to whether these were appropriate or complete measures of what patients were looking for. Perceptions of process or outcome measures of quality are almost totally lacking in health care demand studies. Further investigation of patient perceptions of quality is merited.

Little has been done to investigate the distributional aspects of quality improvements. The most recent study of Kenya tested gender-specific effects of quality attributes [Mwabu et al., 1993]. The results implied higher demand for antimalarials by men. The Lavy and Quigley [1993] study looked at welfare-neutral willingness to pay for three different income groups. The results indicated that willingness to pay (as measured by percentage of income) increased with income. The income effect was more important for the quality variables (infrastructure, drug availability, services, and personnel) than for distance or price. The authors concluded that the richest are more concerned with quality of health services than its costs (price or distance).

These findings lend mixed support to an increasingly posited hypothesis that user fees could be a source of welfare transfers if fees are based on ability to pay and are channeled into improvements in service quality and accessibility [Creese, 1991]. Anecdotal evidence suggests that quality improvements, especially improved drug supplies, might be more important to the poor. What remains to be done is to understand how preferences and willingness to pay for various types of quality improvements differ over income groups, gender groups, and other vulnerable groups (i.e., children under five).

The more recent health care demand studies paid particular attention to measurement problems associated with quality variables, especially in the analysis of nonexperimental static cross-sectional studies. The coefficients on the quality variables could be biased for a variety of problems related to endogeneity. First, quality improvements might actually affect health outcomes, reducing population morbidity. If so, demand for health care services would drop in future periods. Second, observed quality reflected both supply and demand conditions. An
estimated negative relationship between probability of use of a provider and the quality attribute of the provider might simply reflect excess demand rather than the counterintuitive behavior that individuals do not prefer better quality. Third, data on quality should be exogenous indicators of quality. Data collected solely from patients who actually use the facility are hampered by selection bias.

Other data problems encountered in these cross-sectional studies were the small sample sizes of facilities and multicollinearity of various indicators of quality. Typically, the solution has been to use aggregate measures of quality or omit more detailed measures of quality. This greatly limits one's ability to look at willingness to pay for specific dimensions of quality, the information most pertinent to health care managers.

Some health care demand studies have attempted to measure elasticities of demand with respect to quality improvements, where demand is measured in terms of probabilities of choosing a given provider. These elasticity measures should be interpreted with caution since it is difficult to quantify incremental changes in quality, especially in a manner that can be compared across quality attributes. Is a 10 percent improvement in drug availability comparable to a 10 percent improvement in building conditions or provider knowledge?

Finally, one of the most recent health care demand studies attempted to go beyond traditional estimates of price elasticities of demand to estimate willingness to pay (at constant utility). The interest here was to generate actual price levels for various quality improvements. However, one should use these estimated price levels with caution. They were based on simulations using the discrete choice empirical results where quality was typically measured in an aggregate form and coefficients were potentially biased due to endogeneity problems. As we will see in a later section, compensating variation can also be measured through contingent valuation surveys. Such surveys can elicit information on more specific aspects of quality, but only in a hypothetical scenario.

The studies of WTP for quality improvements in developing countries suggested that patients intend to pay for quality improvements, especially for pharmaceuticals and that these amounts are quite substantial. Strikingly, rural populations exhibit strong intentions to pay for improved services. Although higher income groups tend to be willing to pay more for quality improvements, the results suggest that WTP can be quite high in lower income groups. Strong preferences for improvements in drug availability at facilities over other types of quality improvements, such as provider attitudes and building conditions, suggest willingness to pay for quality improvements may be closely tied to financial considerations, a hypothesis that should continue to be tested. Also, some quality improvements (e.g., improved accounting systems) are less visible to patients. Another interpretation might be that rural populations have fewer alternatives.
It is difficult to determine the extent to which these results may be artifacts of methodology. For example, rural populations may be less able to deal with the hypothetical nature of contingent valuation studies. Over-pledging might occur if they believe their preferences will affect provision of quality improvements, but that they will not have to pay the full costs (i.e., donor aid programs). As previous literature has shown, lower education levels, less experience with private health markets, and difficulties in dealing with hypothetical situations may all contribute to the estimated differences in median WTP between different groups, such as rural-urban and/or lower-higher populations.