Capacity in Thai Public Hospitals and the Production of Care for Poor and Nonpoor Patients

Vivian Valdmanis, Lilani Kumanarayake, and Jongkol Lertiendumrong

Objective. To assess the capacity of Thai public hospitals to proportionately expand services to both the poor and the nonpoor. This is accomplished by measuring the production of services provided to poor, relative to nonpoor, patients and the plant capacity of individual public hospitals to care for the patient load.
Study Setting. Thai public hospitals operating in 1999, following the economic crisis when public hospitals were required to treat all patients irrespective of ability to pay.
Study Design and Data Collection. Input and output data for 68 hospitals were collected using databases and questionnaire surveys. A distinction was made between inpatient and outpatient services to both poor and nonpoor patients and the data were assessed statistically.
Data Analysis. Congestion and capacity indices to measure poor/nonpoor service trade-offs and capacity utilization were estimated. The analysis was undertaken by data envelopment analysis (DEA), a nonparametric linear programming approach used to derive efficiency and productivity estimates.
Principal Findings. Increases in the amount of services provided to poor patients did not reduce the amount of services to nonpoor patients. Overall, hospitals are producing services relatively close to their capacity given fixed inputs. Possible increases in capacity utilization amounted to 5 percent of capacity.
Conclusions. Results suggest that some increased public hospital care can be accomplished by reallocation of resources to less highly utilized hospitals, given the budgetary constraints. However, further expansion and increase in access to health services will require plant investments. The study illustrates how DEA methodologies can be used in planning health services in data constrained settings.

Key Words. Equity, hospital efficiency, plant capacity, DEA

Pressure to provide health care for the poor in Thailand has increased substantially since 1997. At that time, the Asian economic crisis led both to a rise in the overall number of Thais living in poverty, and to a reduction in government spending on hospital services. The crisis has affected both private and public hospitals. The impact on private providers has included increased debt, reduced demand, and increased costs of production (Lertiendumrong 2003).

Public hospitals face limited availability of resources (both from governmental budgetary constraints and lower consumption power of households), increased demand due to a shift from private to public providers, and also increased costs of production (Lertiendumrong 2003; Wuttipong 1998; Taearruk 1998). In particular, by 1999, the poorest 40 percent of Thais used public outpatient services more than half of the time, and public inpatient services approximately 95 percent of the time. In contrast, the richest 10 percent of Thais used public outpatient services only 26 percent of the time, and public inpatient services only 54 percent of the time (National Statistical Office 1999). One critical issue for Thai public hospitals is the need to balance the social obligation of providing care for the poor with the need to maintain financial viability, primarily driven by revenue generated by reimbursements from nonpoor patients. Given the Thai financial crisis, funding for public hospitals may be precarious.

Resources used in producing public hospital services can be classified into three groups: capital, labor, and operating costs. In Thailand, decisions regarding the amount of capital and labor employed at public hospitals are made centrally at the Ministry of Public Health (MoPH). Most capital investment expenditure and salary costs are allocated from the central government to hospitals via the budgeting system. The revenue earned by the hospital from nonbudgetary income depends on patients' ability to pay and nonbudgetary income is often earmarked to cover hospital operating costs. Faced with economic difficulties, the Thai government, through the MoPH, launched a package of policies known as "Good Health at Low Cost" (Wibulpolprasert, Tangcharoensathien, and Lertiendumrong 1998). These policies aimed to improve the efficiency of Thailand's health system, focusing on public health care providers. In response to the economic crisis and the growing number of people without access to basic health care, the MoPH required public hospitals to continue providing medical treatment to the poor, irrespective of their ability to pay for these services. Policies promoting equity and efficiency, and reforms to health care financing have continued with the current government (elected in January 2001). In April 2001, the government adopted a universal

Address correspondence to Vivian Valdmanis, Ph.D., Associate Professor of Health Policy, University of the Sciences in Philadelphia, 600 S. 43rd St., Philadelphia, PA 19104. Lilani Kumanarayake, Ph.D., is a Lecturer, Health Policy Unit, London School of Hygiene and Tropical Medicine, London. Jongkol Lertiendumrong, M.D., Ph.D., is a Researcher, International Health Policy Program, Ministry of Public Health, Thailand.

coverage policy, known as the 30 Baht policy, with countrywide implementation in April 2002.

Given the existing hospital and budgetary structures in Thailand, two major questions arise with respect to whether the system can increase production of care for the poor. The first is a congestion issue. On the one hand, public hospitals have an obligation to guarantee the provision of health care to the poor. On the other hand, they need to remain financially viable and are dependent on cost recovery to cover operating costs. One way of achieving the latter is by maintaining or increasing treatment levels for nonpoor patients. The poor are a financial burden on public hospitals because revenue generation from the poor is relatively limited—ranging between 40 percent and 55 percent (Srithamrongsawat 2002). However, given set capacity levels, ensuring provision of services to the nonpoor may inhibit hospitals' ability to provide services to the poor, leading to a failure by public hospitals to fulfill the mandate set up by the MoPH, so that they may fall out of favor and face regulatory sanctions. The second, longer-term question is whether, even if it were financially viable to increase services to the poor, public hospitals currently have the capacity to provide increased services to meet the needs of both poor and nonpoor patients.

OBJECTIVES

Two research questions are addressed in this study of public and regional hospitals under the Regional Hospital Division (RHD)[1] of MoPH operating during 1999. First, we empirically investigate whether providing care to one group of patients negatively impacts on access for all groups of patients (determining whether care for both poor and nonpoor patients are economic goods) given the MoPH mandate and set budget for 1999. For this study, the nonpoor are defined as those individuals who have health insurance or can pay for the full cost of their care, whereas poor patients are those that cannot pay for the full cost of their care, for example, because they are uninsured or are covered by a health insurance scheme that does not pay the full amount of the costs incurred. Second, we empirically measure capacity utilization of public hospitals in Thailand looking at:

- Level of hospitals (reflecting variation in technology and complexity of organization); and
- Region of hospitals (reflecting differences in wealth of catchment areas).

Different hospitals have different technical capacities that may affect their ability to care for patients. For example, the regional hospital (type 1) provides the highest available health care technology among the three types of hospitals. Large provincial hospitals (type 2) are more technology-intensive than the smaller provincial hospitals (type 3). We wish to ascertain whether hospitals operating at higher technical levels are better able to make use of basic infrastructure, such as equipment and manpower, than hospitals with lower technology.

In affluent areas, patients have a greater ability to pay for hospital services than patients from poor areas. In Thailand, different regions vary considerably in their level of wealth. For example, the provinces in central and east regions have a higher average income per capita than other regions and the northeast region has the lowest average income per capita (National Statistical Office 1999).

We use extensions of the data envelopment analysis (DEA) techniques that have been utilized in assessing productivity in health care markets (see, for example, Seiford 1996, for a review of the DEA-health care literature) to address these questions. Our study has a two-step process. First, we apply the "congestion" index derived from the DEA to ascertain whether provision of care to the poor and nonpoor are competing objectives. Second, the paper illustrates how DEA techniques can be used to plan service provision, applying the model of plant capacity utilization. Given that some inputs are set by the MoPH, hospitals can only respond to increased demand by working within their existing plant capacity. Expanding capacity seems unlikely given the constraints on government resources following the Thai economic crisis. In the next section we describe the model.

Model

Data Envelope Analysis

Data envelope analysis (DEA) is a linear programming technique that estimates relationships between inputs and multiple outputs for a sample of decision-making units (DMUs) such as hospitals (Farrell 1957; Charnes, Cooper, and Rhodes 1978; Färe, Grosskopf, and Lovell 1985, 1994). By solving a series of linear programming problems, this nonparametric approach constructs a "best practice frontier" that estimates the maximum possible outputs for set quantities of inputs among DMUs. Commonly used to assess efficiency, the production frontier is considered best practice as it relies on the relative performance of hospitals within the sample rather than a predetermined absolute standard of efficiency. All DMUs lie on or within the interior of the frontier. The latter is considered to be inefficient since, given these levels of inputs, *all* outputs could be increased proportionately.

Extensions of the output-based DEA technique are well suited to this sample of Thai hospitals for several reasons. First, hospitals often do not operate as either cost-minimizers or profit-maximizers, therefore, econometrically specified cost or profit functions may lead to biased findings. Second, DEA does not require information on input or output prices; instead natural units or quantities can be employed in determining the production frontier. This is particularly beneficial in analyzing hospitals from low- and middle-income countries, where price data are often unreliable or missing. Third, the resulting DEA efficiency measure has a straightforward interpretation: how much can output be increased, given inputs, for hospitals lying inside the frontier (hospitals lying on the frontier have an efficiency score of 1). Fourth, DEA readily accommodates multiple inputs and outputs and so is relevant for hospitals that produce multiple services. Fifth, the output-based DEA measure can be adapted to test whether reducing the production of one type of patient care (e.g., nonpoor) leads to a reduction in the production of care for other types of patients (e.g., poor). Data envelope analysis approaches can be extended to examine the question of whether there is sufficient capacity among hospitals to increase services to the poor.

The Relationship between Care for the Poor and the Nonpoor: A Congestion Index Approach. Given the MoPH mandate requiring that care be provided for the poor regardless of ability to pay, there may be competing objectives for Thai public hospitals as they also rely on nonpoor patients to generate financial resources. In order to look at the relationship between poor and nonpoor care, DEA techniques can be extended to construct an output "congestion" index. Whereas other outputs may congest the production of care, we focus our study on these patient-based outputs. This index measures whether one type of patient-based output (care for the poor) can be increased without requiring an increase in other types of outputs (care for nonpoor); in other words, whether the ratio of care for the poor to nonpoor may be increased. Technically, congestion is defined as a situation whereby the production of one output hinders the production of another. Congestion could occur when reducing care for the nonpoor (in order to provide care for the poor) would lead to costs to the hospitals in terms of foregone payments, and thereby reduce the capacity to provide care for the poor.

To construct the congestion index, DEA efficiency measures estimated under assumptions of strong and weak disposability of outputs are compared. Mathematical derivation of this index and formal definitions are presented in the Appendix. Strong disposability of outputs means that different outputs may be substituted for one another. *Both* types of output can be increased if the hospital is operating inefficiently (inside the frontier) given inputs. Weak disposability occurs when the reduction of one output leads to the reduction of another output, given constant inputs.

The congestion index (C_I) is defined as the ratio of the efficiency scores derived under the assumption of strong disposability of outputs, E_S, to the efficiency scores derived using the assumption of weak disposability of outputs, E_W. So $C_I = E_S / E_W$. These efficiency scores are estimated assuming variable returns-to-scale technology, as congestion is considered a short-run concept. That is, outputs can be adjusted without any change in fixed inputs. If this index is equal to 1, there is no congestion and both goods can be proportionately increased within the available technology (i.e., given inputs, outputs can be increased if the hospital is not already on the frontier). An index value of less than 1 implies that either care for the poor or care for the nonpoor is congesting production or is not permitting expansion to the frontier, that is, that there are negative marginal products between output types.

Capacity to Provide Increased Services: The Plant Capacity Approach. Once we have looked at the congestion characteristics of these services, our second question is whether care for both the poor and nonpoor can be increased proportionately, given that some inputs are fixed (e.g., those decided by the MoPH). To answer this, we measure the plant capacity of our sample of Thai hospitals, which provides an indication of the current levels of capacity utilization.

The "actual" plant capacity measure is derived by measuring the optimal production of outputs compared to the actual production of outputs (Nelson 1989). This reflects the economic definition of capacity, where the optimal measure of outputs arises at the tangency between the short-run average cost curve and the long-run average cost curve (Nelson 1989; Morrison 1985). However, in this case some inputs are fixed and others are variable, so a measure of short-run performance is required. We adopt Johansen's definition of plant capacity, which is defined as "the maximal amount that can be produced per unit of time with existing plant and equipment without restrictions on the availability of variable production factors" (Johansen 1987). Whereas earlier works have focused on the single-output measure, Segerson and

Squires (1990) and De Borger and Kerstens (2000) have demonstrated that these plant capacity measures can be expanded to multi-output production cases and all outputs can increase proportionately.

To define an output-based efficiency measure (recall that we are only interested in capacity utilization and not technical inefficiency, per se) we first need to factor out any inefficiencies, that is, operating inside the frontier, since we are interested in measuring plant capacity and not inefficient production. Since we assume that all outputs can be increased proportionately, we again use DEA to calculate plant capacity utilization. The plant capacity measure (P_I) is the ratio of the output efficiency where some inputs are variable and others fixed (E_F) to the output efficiency where all inputs are treated as variable (E_V). If $P_I = E_F/E_V = 1$, then the short-run productivity measure equals the long-run measure. If P_I is less than 1, this would indicate unused or underutilized plant capacity, which could be used to treat proportionately more both poor and nonpoor patients. The formal derivation for this index is presented in the Appendix, and uses the methods developed by Färe, Grosskopf, and Kokklenberg (1989), Färe, Grosskopf, and Valdmanis (1989), and Färe, Grosskopf, and Lovell (1994). We arrive at a multi-output measure of plant capacity under restrictions of constant returns to scale. For the plant capacity analysis, we restrict our production function to constant returns to scale because, unlike the analysis determining the disposability of outputs, we are concerned with comparing short-run and long run productivity.

DATA

The sample consists of 68 public general hospitals operating under RHD—representing 7.1 percent of hospitals and 27.4 percent[2] of all beds in Thailand. These hospitals were selected for five reasons. First, they consume a high proportion of the overall MoPH budget (25 percent in the 1999 budget year). Second, they include tertiary-level hospitals and are the main referral centers in rural areas, so their performance will affect services provided to people in rural areas. Third, data for this group of hospitals are relatively easily available, with an established information system in place. Fourth, these hospitals have a high potential to improve performance for both financial and manpower resources compared with other smaller private facilities. Fifth, these hospitals have been mandated by the MoPH to provide services for the poor.

To achieve the aims of study, various datasets from these hospitals were collected for the year 1999. Some datasets were available from the MoPH database, while others were not. Available datasets were retrieved from the ministry's database and sent back to be verified by appropriate hospital staff. Other data were collected directly from hospitals using a questionnaire survey.[3] From a total of 92 hospitals, complete data were available for 68 hospitals. However, no systematic differences could be noted between responding and nonresponding hospitals.

Data on public hospitals in the MoPH database are reported in a budget year that runs from October 1 to September 30. However, annual data reported from other sources are based on calendar years (January to December). As monthly data were not available for all datasets, it was assumed that there was not much discrepancy in the data between the two periods.

In total, seven inputs and four outputs were included in the study. The inputs were the number of beds, doctors, nurses, and other staff, and allowance expenditures, drug expenditures and other operating expenditures. The first four inputs are classified as *fixed variables* as the levels of capital and labor are set by the MoPH. Due to civil servant regulations, doctors, nurses, and other staff cannot easily move from their place of employment. The last three inputs are variable and can be altered according to patient volume.

The four outputs are number of outpatient visits for poor patients, number of outpatient visits for nonpoor patients, total inpatient cases adjusted with average diagnostic related group (DRG) weighting for poor patients, and total inpatient cases adjusted with average DRG weighting for nonpoor patients. Since the relative weight for each patient was not available, the numbers for inpatient admissions were adjusted with the hospital's average[4] DRG weights instead. Unfortunately, no case-mix adjustments for outpatient visits were available for the study period.

RESULTS

Table 1 presents the descriptive statistics for the input and output variables utilized in 1999.

We found that more outpatient care was provided to the nonpoor (an average of 33 percent more) and there was virtually no difference in the number of adjusted inpatient cases between the poor and nonpoor. The National Socio-Economic survey for 1999 reported that the poor made greater use of public hospitals than the nonpoor (see opening section of this article). However, these findings suggest that the relative utilization of overall services by the poor, particularly outpatient care, is less than the utilization by the

Table 1: Descriptive Statistics of Input and Output Variables, 1999

Year	*Variable*	*Observations*	*Mean*	*Std. Dev.*	*Minimum*	*Maximum*
1999	*Bed*	68	408	194	160	1,082
	Doctor	68	36	28	9	155
	Nurse	68	335	147	148	885
	Other Staff	68	505	256	181	1,390
	Allowance	68	19,500,000	13,400,000	5,091,437	66,700,000
	Drug expense	68	57,400,000	43,400,000	10,900,000	207,000,000
	Other operating expense	68	44,600,000	30,300,000	2,484,917	160,000,000
	OP visit for nonpoor	68	114,807	72,261	19,284	368,975
	OP visit for poor	68	86,587	34,741	26,543	180,863
	IP weight for nonpoor	68	10,578	7,521	1,316	37,687
	IP weight for poor	68	10,565	6,204	2,371	29,736

Note: IP weight is in-patient case adjusted with average DRG relative weight.

nonpoor. Reasons for this difference are not testable because DRG weights for outpatient visits were not available.

In Table 2, the median output congestion ratios for the sample are presented. For both poor and nonpoor services, a comparison is made between the strong and weak disposability specifications. The congestion index equals one for both types of services, indicating that producing less care for the poor does not lead to producing care for the nonpoor, and vice versa. Therefore, both types of treatment can be proportionately increased if the hospitals' production plans lie within the production possibilities frontier.

Since we determined that both types of care can be proportionately increased if a hospital has excess capacity, we explored the second issue, namely, the extent to which capacity is currently utilized and if plant capacity varies by type of hospital or by the region in which the hospital is located. The descriptive statistics of the capacity utilization measures by types and regions, and p-values from the Kruskal-Wallis test for equality of capacity utilization in 1999 by types and regions of hospitals, are presented in Table 3.

In 1999, type 1 hospitals had the highest average capacity utilization. On average, the possible increase in services provided was 2.6 percent, given fixed factors. This means that these types of hospitals operated at nearly full plant capacity. Type 2 hospitals had the lowest capacity utilization among the three types of hospitals, with 5.4 percent unused capacity. However, we did not reject the null hypothesis for sampling variation, so the mean values for each type of hospital are statistically equal (Kruskal-Wallis test $p \sim 0.360$).

Table 2: Descriptive of Output Congestion Index for Poor and Nonpoor Services, 1999

Specification	*Mean*	*Median*	*95th Percentile*	*Std. Dev.*	*Minimum*	*Maximum*
Nonpoor	1.002	1	1.02	0.012	1	1.086
Poor	1.021	1	1.10	0.059	1	1.360

Note: $n = 68$. Weak and strong disposability specifications were compared for all outputs specification to analyze output congestion. Using a Kruskal-Wallis test for statistical significance, we fail to reject the null hypothesis that the rankings are equal at the $p < 0.05$ level.

Table 3: Plant Capacity Utilization

Descriptive Statistics of Capacity Utilization, 1999

Variable	*Type*	*Observations*	*Mean*	*Std. Dev.*	*Minimum*	*Maximum*
1999	1	16	**0.974**	0.047	0.840	1
	2	37	0.946	0.072	0.760	1
	3	15	0.951	0.082	0.680	1

Descriptive Statistics of Plant Capacity, 1996–1999, by Type of Hospital

Year	*Region*	*Observations*	*Mean*	*Std. Dev.*	*Minimum*	*Maximum*
1999	Central	10	0.965	0.054	0.850	1
	East	6	0.907	0.099	0.760	1
	North	15	0.961	0.054	0.840	1
	Northeast	13	**0.989**	0.033	0.880	1
	West	10	0.939	0.070	0.790	1
	South	14	0.935	0.091	0.680	1

Descriptive Statistics of Plant Capacity, 1996–1999

Variable	*Observations*	*Mean*	*Std. Dev.*	*Minimum*	*Maximum*
1999	68	0.954	0.069	0.680	1

P-values from Kruskal-Wallis Test for Equality of Population of Capacity Utilization, 1999, by Types and Regions of Hospitals

Type	0.360
Region	0.065

Looking by region, we find that hospitals in the northeast region have the highest average capacity utilization (98.9 percent). This indicates that, on average, hospitals in this region already produce services near the optimal

Table 4: Plant Capacity Utilization by Sample and Selected Individual Hospitals

Mean	*Standard Deviation*	*Minimum*	*Maximum*
0.954	0.069	0.689	1.00
Individual Hospital	*Hospital Type*	*Statistics Region*	1999
87	3	s	**0.68***
34	2	w	**0.79***
20	1	e	**0.84***
64	2	n	**0.86***
91	2	s	**0.87***
22	2	e	**0.76***
13	2	c	**0.85***
25	2	w	**0.86***

level with given numbers of beds and personnel. On the other hand, hospitals in the east had the lowest average capacity utilization (90.7 percent). Given our findings and sampling error, we did not reject the hypothesis that capacity utilization of hospitals in different regions had the same mean at the 0.05 significance level.

We also examined the eight hospitals with the lowest level of plant capacity utilization during 1999. Table 4 shows there is no systematic relationship between plant capacity utilization and the region in which the hospital operates. Type 2 hospitals, however, did dominate the majority of the hospitals operating with excess capacity.

KEY FINDINGS AND CONCLUSIONS

This paper used DEA-based models to empirically investigate the performance of Thai public hospitals operating in 1999. The paper tested to see whether regulations regarding supply of services to the poor would lead to an overall reduction in the provision of services to the nonpoor. This paper is the first analysis to use DEA methodology to consider competing objectives for hospitals. By relaxing the strong disposability of output assumption in our model, we were able to test whether reducing care for the nonpoor would lead to a decrease in care for the poor. This is a relevant policy question since if providing care to one group of patients harms access for all groups of patients, then the equity condition is not met (i.e., Pareto Optimality), a central aim of the MoPH reforms. Our results suggest that social welfare is optimized since all types of patients are treated equally by the hospital. The hospital-based statistics did not disclose a dramatic difference in utilization between the poor and nonpoor. However, since the proportion of poor that seek care in public hospitals is greater than for the nonpoor, the question naturally arises as to whether access to hospital care in general among the poor is lacking. Assessing hospital care alone may not be sufficient to underpin statements regarding equity.

Results of the congestion index found that the marginal product of poor and nonpoor services are nonnegative and that the financial incentives related to increased cost recovery from nonpoor services did not affect the extension of services to the poor. This indicates that different patient types are considered as equals in a productive sense. In general, however, there may have been a decrease in the number of nonpoor in Thailand due to the economic crisis resulting in a lack of ability to pay for private health services. Individuals covered by the low-income scheme grew from 17.9 million in 1997 (before the economic crisis) to 21.3 million in 1999. If the demand for public hospital care continues to grow, it is important that hospitals both become more productively efficient and have enough capacity to serve the population in need.

The results from the analysis of plant capacity show that hospitals are generally operating at relatively high capacity (90–95 percent), given levels of fixed inputs. Type 1 hospitals have the highest average capacity utilization and type 2 hospitals have the lowest average capacity utilization. However, there is no significant ($p<0.05$) difference between capacity utilization across the three types of hospitals. This is striking considering the very different services provided by the different hospitals, in particular, type 1 hospitals are more capital/technology intensive. What may be inferred from these findings is that all hospitals could be expanded to some degree to care for the patients admitted, irrespective of type or intensity of services. This may bode well for smaller institutions that can care for a variety of patients' needs but without the expensive technology that is needed at higher-level hospitals. Similarly, hospitals in the northeast region have the highest-capacity utilization, and those of east region the lowest. However, there is no significant difference of capacity utilization across the different regions at $p<0.05$.

Regional variation suggests that hospitals in the east region could provide more services within the current fixed inputs and unconstrained operating expenses. However, the sample size of hospitals in the east is rather small. The high plant capacity in the northeast region is coupled with the fact that this is poorer than the central and east regions. This finding may lead to policy debates over reallocation of resources (from rich to poor) that would require an infusion of public monies since the poor northeast region may not be able to raise funds from paying patients.

We find that, in general, the public hospitals in our sample are producing outpatient and inpatient services close to their capacity given the MoPH-mandated fixed inputs. Furthermore, with the plant capacity currently in operation, proportionately more poor patients could be treated, and given that no weak disposability between the patient-based outputs existed, care for both the poor and nonpoor could increase.

However, the number of poor patients may continue to grow, especially if more poor people gain access to hospital services, which is likely with the implementation of the 30 Baht scheme. Furthermore, if the MoPH continues to be reliant on the public sector to provide services to a potentially growing number of poor, further expansion or a reallocation of resources to hospitals operating at full capacity from hospitals with less capacity will be necessary. This may make the most sense in a middle-income country with limited budgets for public services.

This paper extends existing DEA methods and considers the question of competing hospital objectives with respect to the poor and nonpoor. It also measures capacity in the public health system. The use of DEA in the Thai setting shows that these methodologies are useful in developing country settings where data can be limited, making it difficult to estimate marginal costs accurately. The specific analysis of plant capacity is a good tool for planning and our analysis illustrates how these methods can be used to identify hospitals where potential capacity utilization improvements can be made.

ACKNOWLEDGMENTS

The authors would like to thank Thailand Research Fund and Health System Institute for funding, J. Lertiendumrung and Toby Mendel for editing, and two anonymous referees and the editors of this journal for helpful comments.

APPENDIX
MODELING HOSPITAL BEHAVIOR UNDER THE ASSUMPTIONS OF STRONG VERSUS WEAK DISPOSABILITY

Under the Farrell (1957) framework that was updated by Färe, Grosskopf, and Lovell (1994), production under the assumptions of variables returns to scale (V) and strong disposability of outputs (S) can be modeled as:

$$P(x,\ y|\ V,\ S) = \{y : y \leq z{\cdot}M,\ z \cdot K \leq x,\ z \in -, \sum_{j=1}^{N} z_j = 1\},$$

where y is the individual amount of output produced by a firm, M is the total amount of output produced by the other firms in the sample, x is the amount of input used by the firm, and K is the total amount of inputs employed by the other firms in the sample. The z denotes the intensity parameters that permit the convex combinations of inputs and outputs. The summations of the z parameters equaling one permit variable returns to scale. In order to permit weak disposability of outputs (W), we add the μ parameter, which further constrains the movement along the production possibilities frontier. In other words, the tradeoffs among outputs are no longer allowed.

$$P(x,\ y|\ V,\ W) = \{y : y \leq \mu \cdot z{\cdot}M,\ z \cdot K \leq x, 0 \leq \mu \leq 1,\ z \in -, \sum_{j=1}^{N} z_j = 1\},$$

The difference between the $P(x,\ y|V, S)$ and $P\ (x, y|V, W)$ allows us to gauge "desirable" output loss due to production of "undesirable" outputs; called output congestion (Färe et al. 1989). The resulting measures can be determined by solving the following two linear programming problems, where output is maximized given input levels.[5]

$$\begin{aligned} P(x|\ V,\ S) &= \max \theta \\ s.t. \theta \cdot y &\leq z \cdot M \\ z \cdot K &\leq x \\ z &\in \mathscr{R}^{+} \\ \sum_{j=1}^{N} z_j &= 1, \end{aligned} \tag{1A}$$

$$\begin{aligned} P(x|\ V,\ W) &= \max \theta \\ s.t. \theta \cdot y &\leq \mu \cdot z \cdot M \\ z \cdot K &\leq x \\ 0 \leq \mu &\leq 1 \\ z &\in \mathscr{R}^{+} \\ \sum_{j=1}^{N} z_j &= 1. \end{aligned} \tag{2A}$$

Once again θ is the maximum radial expansion of outputs but the added constraint μ permits in the weak disposability case, that is, the backward bend in the production possibility curve (and so nonconvexity).

This congestion index is defined as the ratio of the production technology assuming strong and weak disposability of outputs $(P(x|V, S)/P(x|V, W)$, which is the ratio of the solutions from Equations (1A) and (2A). If this ratio equals 1 then there is no weak disposability of outputs and both outputs can be considered economic goods, implying a possible increase in production of poor services does not affect nonpoor services or vice-versa and so movement along the frontier is possible. Any measure less than 1 indicates the percent amount by which total output is reduced due to the imposition of the regulation that poor patients must be treated and the financial reality that a reduction in non-poor patients will lead to a reduction in poor patients and so a movement along the frontier is not possible. In order to gauge whether this is the case, we assess an output based model relaxing the constraint of strong disposability of output. In the first case, we will assess whether care for poor patients is an economic "bad" and in the second case, we will assess whether care for nonpoor patients is an economic "bad." If neither case arises, then reducing one type of output will not reduce production of the other type of output. Assuming that no congestion is present in production, the next step of our analysis is ascertaining if expansion of both goods is possible given constraints (mandated by law) of certain types of inputs. This leads us to describing the second modeling approach used in this paper.

Computationally, it is relatively straight forward to measure the Johansen (1987) definition of plant capacity utilization employing the methods developed by Färe, Grosskopf, and Kokklenberg (1989), Färe, Grosskopf, and Valdmanis (1989), and Färe, Grosskopf, and Lovell (1994). The modeling is presented below.

Since we are assuming that the outputs can all be increased proportionately we again use nonparametric data envelopment analysis (DEA) to calculate plant capacity utilization. This is done by fixing some of the inputs in the model's constraints.

In order to define an output-based efficiency measure (recall that we are only interested in the issue of capacity utilization and not technical inefficiency, per se) we first need to remove any inefficiencies, that is, operating inside the frontier, since we are interested in measuring plant capacity and not inefficient production. This is done by calculating the familiar DEA output-based measure:

$$F_0(x^k, u^k | C, S) = \max_{\theta,z} \theta$$
$$s.t. \theta u_{km} \leq \sum_{k=1}^{K} z_k u_{km}, m = 1, 2, \ldots M$$
$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{kn}, n = 1, 2, \ldots N$$
$$z_k \in \mathscr{R}^+$$

In the second model, we hold certain inputs fixed by adding the constraint where $n \in \hat{a}$ indicate fixed inputs. In this way we have the optimal output that is possible when variable inputs are unrestricted and fixed inputs are restricted, which is consistent with the Johansen definition of plant capacity (Färe, Grosskopf, and Lovell 1994).

$$\hat{F}_o(x_f^k, u^k | C, S) = \max_{\theta z \lambda} \theta$$
$$s.t. \theta u_{km} \leq \sum_{k=1}^{K} z_k u_{km}, m = 1, 2, \ldots M$$
$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{kn}, n = 1, 2, \ldots N$$
$$\sum_{k=1}^{K} z_k x_{kn} = \lambda_{kn} x_{kn}, n \in \hat{a}$$
$$z_k \in \mathscr{R}^+$$
$$\lambda_{kn} \geq 0, n \in \hat{a}$$

Dividing the output-based measure whereby all inputs are allowed to vary by the plant capacity model wherein some variables are held fixed, we arrive at a multi-output measure of plant capacity under restrictions of constant returns to scale. We note here we are restricting our production function to constant returns to scale because unlike the analysis determining the disposability of our two outputs, we are concerned with long-run implications. In other words, we have derived the measure wherein the economic definition of plant capacity is met, that is, that the short-run productivity measure equals the long-run measure of productivity when the plant capacity measure equals one.

Dividing the output-based measure whereby all inputs are allowed to vary by the plant capacity model wherein some variables are held fixed, we arrive at a multi-output measure of plant capacity under restrictions of constant returns to scale. In other words, we have derived the measure wherein the economic definition of plant capacity is met if the results are less than 1, this would indicate unused or underutilized plant capacity, which could permit treating proportionately more of both nonpoor and poor patients.

NOTES

1. After the structural reform of the MoPH in October 2002, RHD, in charge of 92 regional and provincial hospitals, was merged with the Rural Hospital Division to become the Department of Health Service Support, responsible for 800 district hospitals.
2. In 1999, 72.2% and 60.7% of hospitals and beds, respectively, in Thailand were publicly owned. However, these figures also reflect the fact that a large proportion of beds in private hospitals were not in operation due to the crisis.
3. Two rounds of telephone and mail follow-ups were done before the fieldwork was terminated. The main reasons that hospitals were unable to provide data were changes in staffing, and inadequate records to verify MoPH data or to retrieve the additional required data.
4. In 1999, only average DRG weights for all patients for each hospital were available. These DRG weights were used to adjust numbers of inpatient cases for the poor and nonpoor groups.
5. We used the *OnFront* software package, which is a user-friendly program that easily derives the measures we present in this paper.

REFERENCES

Charnes, A., W. Cooper, and E. Rhodes. 1978. "Measuring the Efficiency of Decision Making Units." *European Journal of Operational Research* 2 (6): 429–44.

DeBorger, B., and K. Kerstens. 2000. "The Malmquist Productivity Index and Plant Capacity Utilization." *Scandinavian Journal of Economics* 102 (2): 303–10.

Färe, R., S. Grosskopf, and E. Kokklenberg. 1989. "Plant Capacity and Technical Change: A Non-Parametric Approach." *International Economic Review* 30 (3): 655–66.

Färe, R., S. Grosskopf, and C. A. K. Lovell. 1985. *The Measurement of Efficiency of Production.* Boston: Kluwer-Nijhoff.

———. 1994. *Production Frontiers.* Cambridge, U.K.: Cambridge University Press.

Färe, R., S. Grosskopf, and C. Pasurka. 1989. "Measuring Efficiency When Some Outputs Are Undesirable: A Non-Parametric Approach." *Review of Economics and Statistics* 71 (1): 90–8.

Färe, R., S. Grosskopf, and V. Valdmanis. 1989. "Capacity, Competition, and Efficiency in Hospitals: A Non-Parametric Approach." *Journal of Productivity Analysis* 1: 123–38.

Farrell, M. 1957. "The Measurement of Productive Efficiency." *Journal of the Royal Statistical Society* 120 (part 3): 253–78.

Johansen, L. 1987. "Production Functions and the Concept of Capacity." In *Collected Works of Leif Johansen* p. 362, edited by F. R. Førsund, Amsterdam: North Holland Press.

Lertiendumrong, J. 2003. "Efficiency of Public Hospitals in the Context of Change: A Case Study of Public General Hospitals in Thailand Pre- and Post-1997 Economic Crisis." Ph.D. Dissertation, London School of Hygiene and Tropical Medicine.

Morrison, C. 1985. "Primal and Dual Capacity Utilization: An Application to Productivity Measurement in the U.S. Automobile Industry." *Journal of Business and Economic Statistics* 3 (4): 312–24.

National Statistical Office. 1999. *Report of the 1999 Household Socio-Economic Survey Whole Kingdom.* Thailand: Office of the Prime Minister.

Nelson, R. 1989. "On the Measurement of Capacity Utilization." *Journal of Industrial Economics* 38 (3): 273–86.

Segerson, K., and D. Squires. 1990. "On the Measurement of Economic Capacity Utilization for Multi-Product Industries." *Journal of Econometrics* 44 (3): 347–61.

Seiford, L. 1996. "Data Envelopment Analysis: The Evolution of the State of the Art (1978–1995)." *Journal of Productivity Analysis* 7: 99–137.

Srithamrongsawat, S. 2002. "The Health Card Scheme: A Subsidized Voluntary Health Insurance Scheme." In *Health Insurance Systems in Thailand*, edited by P. Pramualratana and S. Wibulpolprasert, chap. 6. Desire, Thailand: Health Systems Research Institute.

Taearruk, P. 1998. "Health of Thai under Economic Crisis." In *Social and Cultural Impacts from the Economic Crisis.* Bangkok, Thailand: Social Research Institute, Chulalongkorn University.

Wibulpolprasert, S., V. Tangcharoensathien, and L. Lertiendumrong. 1998. "The Economic Crisis and Responses by the Health Sector in Thailand 1997–1998." *WHO Regional Consultation on Health Implications of the Economic Crisis in the South-East Asia Region.* March, 1998 Bangkok, Thailand.

Wuttipong, P. 1998. "Response Policy to Economic Crisis in Health System." Keynote Address, Ministry of Public Health, Medical Society of Thailand, Bangkok, Thailand.