Effects of Differences in Hospital Bed Scarcity on Type of Use

MARTIN S. FELDSTEIN,* A.B., B.LITT.

Brit. med. J., 1964, 2, 562-565

One of the most striking features of the hospital service is the variation in the number of beds available per thousand population in different regions of the country. This paper investigates the effects of relative scarcity on different uses of hospital beds and various types of patients. The results are not only of interest in themselves but suggest some important policy consequences.

Variation in Hospital Bed Availability

Table I shows the availability and use of non-mental hospital beds in each of the eleven regions outside London. All data about bed use is derived from the recently published Hospital In-Patient Enquiry for 1960 (General Register Office, 1963);

TABLE I.—Availability and Use of Non-mental Hospital Beds

Region	Beds Available* (1)	Beds Used*† (2)	Cases Treated* (3)	Mean Stay (4)
Birmingham ..	5.02	3.73	71.26	19.2
East Anglia ..	5.06	3.66	63.56	21.1
Leeds ..	5.94	3.98	79.37	18.3
Liverpool ..	6.79	5.33	90.66	21.5
Manchester ..	4.89	4.03	78.14	17.0
Newcastle ..	5.15	3.96	85.16	19.5
Oxford ..	5.35	4.46	83.73	17.5
Sheffield ..	4.61	3.24	67.85	20.5
S. Western ..	5.95	4.37	77.81	20.5
Wales ..	5.87	4.56	81.50	20.5
Wessex ..	5.04	4.29	79.46	19.8

* Per 1,000 population, 1960.

† Beds used per 1,000 population =

Cases treated per 1,000 population $\times$ mean case stay (days)

366

bed availability data were supplied by the Ministry of Health, and cover beds included in the H.I.P.E., plus private and staff beds. Column 1 indicates that the Sheffield region has the smallest number of hospital beds per thousand population, only 4.6, while the Liverpool region has almost 50% more—some 6.8 beds per thousand population. Correspondingly, Column 2 shows that only 3.24 beds are used per thousand population in Sheffield while 5.33 beds are used in the Liverpool region. Total bed use is very much determined by total supply; the correlation between the two sets of numbers is 0.84.

The regional variation in hospital bed availability reflects history rather than deliberate policy. Half the hospitals now in use are more than 70 years old and one-fifth have been standing for a century; the number of beds has changed very little since before 1939, and only now, with the new 15-year hospital plan, will the Government do any substantial building. The current pattern of bed availability thus reflects a previous distribution of population and a past type of medical care.

What is the effect of these differences in hospital bed availability? How does relative scarcity influence the ways in which the hospital service allocates its limited bed supply

among competing uses? Does a region in which beds are relatively more scarce reduce the number of cases treated or the average duration of stay? Are older patients more affected by the scarcity of beds than younger ones? Are treatments for tonsils and adenoids more affected than those for heart disease or hernias? We have investigated the effects of bed scarcity on a large number of different types of patients; some of the results conform to expectations, but others are quite surprising and may suggest an important problem in the operation of the health service. A selection of these results is presented in this paper.

Shorter Stays or Fewer Cases

The basic choice in allocating hospital beds is between the number of cases treated and the average stay per case. Before this study was conducted it was anticipated that bed scarcity would influence average duration of stay more than it influenced the number of cases treated. Where beds are scarce it should be possible to reduce the average stay per case, especially since Britain has rather long stays in comparison with the United States and several Continental countries. But the number of cases requiring treatment should be less variable; the same number of hernias are likely to need repair, appendices to need removal, heart-attack victims to need attention. Put another way, a potential patient would no doubt be happier if he thought that the relative scarcity of beds in his area affected the length of time that he might stay in hospital rather than the chance of getting a hospital bed when he needed one.

It is therefore surprising, and somewhat disconcerting, to find that bed scarcity has a 50% greater effect on the number of cases treated than on the average stay per case. Column 3 of Table I shows that in the Sheffield region (where beds are scarcest) there were some 68 cases per thousand population (approximately one person in fifteen) during the year. The Liverpool region, where beds are least scarce, hospitalized some 91 persons per thousand population, or approximately one person in eleven. Living near Sheffield decreases the chance of being admitted to hospital by more than 25% in comparison with residing near Liverpool. Column 4 shows that the average stay per case is less influenced by bed scarcity; the range is narrower (from 17.0 to 21.5 days, or only 23% of the average, in comparison with the 38% spread of the number of cases) and the average stays are less closely correlated with bed availability.

Although much can be learned by studying the individual regional data in this way, it is more convenient to have a simple single-value index of the response of bed use to scarcity. A number of different measures could be adopted—rank correlations or product moment correlations, for example; the index employed in this study has certain advantages over these and other possible choices. The “responsiveness index” value associated with a particular use of hospital beds is defined as

* Research Fellow in Economics, Nuffield College, Oxford.

100 times the percentage change in the particular type of "bed use" divided by the percentage change in bed availability:

$$\text{Responsiveness} = \frac{\% \text{ Change in "Bed Use"}}{\% \text{ Change in Beds Available per 1,000 population}} \times 100.$$

"Bed use" may refer to any of three things: (1) the number of beds used per 1,000 population; (2) the number of cases treated per 1,000 population; and (3) the mean stay per case.

The responsiveness index is an adaptation of the concept that economists refer to as "elasticity," and is calculated by the usual method of "least squares" regression on the logarithms of the original variables (Johnston, 1960); the statistical problems of estimating the responsiveness values will be discussed more thoroughly elsewhere (Feldstein, in preparation).

TABLE II.—*Responsiveness to Bed Scarcity: Number of Cases and Average Stay*

	Beds Used*	Cases Treated*	Mean Stay
All persons	96 (22)	58 (23)	36 (18)
Males	104 (25)	64 (20)	41 (18)
Females	87 (26)	52 (26)	35 (25)

Figures in parentheses are standard errors of responsiveness index values.
* Per 1,000 population, 1960.

An example will help to clarify the way in which the responsiveness index should be interpreted. Table II shows that for "all persons" the number of cases treated per 1,000 population has a responsiveness index value of 58. This can be interpreted as meaning that, on average, if one region has more beds than another the proportional difference in the number of cases treated will be 58% as great as the proportional difference in the number of beds—that is, if the first region had 10% more beds than the second it would be expected to treat 5.8% more cases than the second.

In contrast, the responsiveness index for average stay is only 36; this was referred to before when it was noted that bed scarcity had a 50% greater effect on the number of cases treated than on the average duration of stay. Table II also shows that this relationship between the responsiveness values of cases treated and average stay holds separately for both males and females; the values are 64 and 41 for males, 52 and 35 for females. The treatment of male patients is more responsive to bed scarcity than the treatment of female patients; the number of beds used for males shows a responsiveness of 104 while for females it is only 87.

The results presented in Table II are quite surprising. How can one explain the uncomfortable fact that the number of cases treated seems to vary more in response to relative bed scarcity than does the average duration of stay. Two types of explanation are possible; the first might be called "behavioural" and the second might be called "environmental." A behavioural explanation would accept the finding as correct and would seek to explain it in terms of theories about the behaviour of doctors, patients, and others. An environmental explanation would suggest that the difference between the two responsiveness values was more apparent than real—that is, that it could be explained by other factors not yet taken into account; such an explanation would not only help us to understand why this result has been found but would also indicate that the seemingly inappropriate behaviour of the hospital system was only a spurious finding.

After considering various possible environmental explanations, it was concluded that the result cannot be justified in this way. There is a real difference between the two responsiveness values that must be explained in terms of the behaviour and attitudes of those involved in the use of hospital beds. Before discussing these behavioural theories we summarize five of the environmental arguments that were considered.

Environmental Explanations

First, it may be that the difference between the two responsiveness values is simply fortuitous—that is, that there is no real difference in responsiveness but an apparent difference due only to chance. Although standard statistical tests cannot lead us to any very firm statement about the observed differences in responsiveness, this is partly due to the paucity of observations—only eleven. However, the same relationship between the two values was found for a high proportion of patient and disease categories investigated; it is very unlikely that this would be due only to chance.

Differences in the *types* of cases treated in the various regions would provide a second type of environmental explanation. If, for some medical or environmental reason, regions with relatively more beds had a disproportionately low number of the types of cases that require lengthy stays, such as older patients or those with heart diseases or bronchitis, this would produce a spurious result suggesting that the number of cases treated is less responsive than the average duration of stay. But this explanation will not do. As already pointed out, the greater responsiveness of cases treated is observed not only for "all persons" taken together but for most types of patient and disease when considered alone.

Another type of environmental explanation would be that the results are due to the way in which the regions are defined or to the particular time period that was selected. The official hospital regions do provide good catchment areas, and there is no reason to suspect that this is a source of bias; we are currently preparing related studies using smaller areas as the fundamental units. Similarly, the annual time period is the most natural, since it overcomes any problems of seasonal variation in disease incidence. The specific choice of 1960 is unlikely to have influenced the results; comparisons with preliminary estimates for 1961 support this belief.

If the availability of beds was itself changing in response to demands in the region or to other related factors, the estimated responsiveness values would be biased. But, as indicated before, the actual pattern of bed scarcity is largely a reflection of history.

Finally, there is the possibility that other factors might influence the use of hospital beds—for example, the availability of general-practitioner or local authority services—and that these factors are correlated with bed scarcity. Although it is unlikely that there are any such factors which substantially influence the *overall* division between cases treated and average stay, alternative services, and differences in area, specific morbidity may be important when individual diseases are considered. Although the small number of regions make it impossible to investigate these other factors, they will be analysed as part of the above-mentioned studies of smaller areas.

Behavioural Explanations

If environmental explanations of the types discussed are not satisfactory, what behavioural theories might help us to understand the relative responsiveness values? Four interrelated ideas seem helpful.

First, there is the hospital physician's feeling of greater responsibility for the patient in his own personal care than for the people who are waiting for admission. Medical training, medical ethics, and the law of negligence all emphasize this asymmetry of responsibility. "The best possible care for my patient" forms the basis of the clinician's decision-making. This is, of course, an attitude to be admired in the doctor and one that, as potential patients, we all hope exists. But it does offer an explanation of why doctors allow the number of patients treated to vary more in response to differences in bed scarcity than they allow the care given to the patients already admitted to hospital. It also impinges on another problem that

lies outside the scope of this paper: the extent to which the medical *administrator*, on every level from the Ministry of Health to the local hospital, may be influenced by standards more appropriate for the clinician dealing with an individual case than for the administrator concerned with the health of the public.

A second and closely related explanation is that the doctor and ward sister are "playing safe" by not decreasing patients' durations of stay in response to relative bed scarcity. No one wants to take the risk of making a decision that can later be pointed to as having endangered the health of the patient. Although British durations of stay are comparatively long, decisions are safest if they reflect prevailing practice. This cautiousness is reinforced by the decision-making framework. The decision to discharge a patient from a large hospital often rests on the advice of the ward sister or a junior medical person. They are likely to be less willing to incur the risk of making a "wrong" decision than are the more senior doctors who select cases for admission or the general practitioners whose referral decisions will not come under professional scrutiny.

This desire to do what constitutes "prevailing practice" suggests a third explanation. The prevailing practice with regard to average duration of stay may be easier to *learn* than the prevailing practice in admission; similarly, it may be easier to identify departure from duration of stay norms than from standards for hospital admission.

Finally, the patient's duration of stay may in some cases be determined by the patient himself. To the extent that this is true, average duration of stay would be unaffected by relative bed scarcity.

Each of these four reasons is compatible with the assumption that doctors realize they allow relative bed scarcity to affect admissions more than durations of stay. Such an assumption is almost certainly unwarranted. Explicitly discarding it does not weaken any of the previous theories but does make it easier to understand why it has been possible for the observed relative responsiveness values to exist: doctors do not recognize the choice they are implicitly making. An individual doctor, working in one place, is aware only of the feeling that in his area beds are not as abundant as he would like; he cannot know the local *relative* scarcity of beds nor realize how his standards of admission and discharge compare with those in other places. As a result, he simply may not recognize that, for a variety of unconscious reasons, he is making a choice that he might not find acceptable if it were made explicit.

Before commenting on the significance of this for health service policy, we briefly present some of our findings on the effects of bed scarcity on patients of different ages and diagnostic categories.

Responsiveness by Age and Sex

The most interesting of the problems associated with patient-age is the responsiveness of older patients to bed scarcity. The National Health Service has frequently been criticized for allowing the *hospital* service to provide accommodation and care for aged patients who would be more appropriately looked after by the general practitioner and local authority services. The first three columns of Table III show how older people use the hospital much more than the rest of the population. All three measures of use increase rapidly with age, except for the deviation caused by women of child-bearing age. Whether this is entirely justifiable for medical reasons is outside the scope of this paper; the information is presented only as background to the responsiveness measures given in the last three columns. While these results are neither clear-cut nor easy to summarize, the following pattern seems to emerge: for mean stay, younger patients (those between 35 and 65) tend to show a rather greater responsiveness to bed scarcity than older patients; the same is true for the number of beds used.

Although the number of cases treated is more responsive among older men than younger ones, among women both groups seem equally responsive. In short, although one would probably have expected and wanted to see that bed use, number of cases, and particularly mean stay were more responsive among older patients than among younger ones, the reverse seems to be true. How can we explain this? Again, the most important factor is probably that doctors are not aware that they have implicitly chosen to make younger patients as responsive or more responsive to bed scarcity than older patients. A recognition of this may not only alter doctors' behaviour but also encourage more adequate provision of community facilities for the infirm aged.

TABLE III.—Use and Responsiveness by Age and Sex

Age and Sex	Use			Responsiveness		
	Beds Used* (1)	Cases* (2)	Mean Stay (3)	Beds Used* (4)	Cases* (5)	Mean Stay (6)
Males						
35—	2.3	46.9	17.8	79 (59)	54 (35)	25 (38)
45—	3.5	63.1	20.2	106 (27)	66 (23)	40 (22)
55—	6.1	97.5	23.0	152 (28)	61 (15)	91 (27)
65—	9.5	122.8	28.2	77 (56)	89 (16)	—12 (58)
75—	18.1	162.4	40.7	67 (57)	73 (31)	—6 (45)
85+	30.1	203.0	54.2	182 (111)	138 (69)	44 (110)
Females						
35—	3.6	88.2	14.9	117 (52)	63 (20)	54 (52)
45—	3.4	69.5	17.7	146 (40)	65 (24)	80 (44)
55—	4.2	68.5	22.4	71 (35)	68 (25)	3 (21)
65—	8.2	88.5	33.8	137 (30)	70 (24)	66 (28)
75—	17.9	122.2	53.4	58 (49)	62 (35)	—4 (62)
85+	35.9	139.3	82.5	—42 (92)	55 (32)	—97 (81)
All persons	4.1	77.3	19.2	96 (22)	58 (23)	36 (18)

* Per 1,000 population, 1960. Standard errors indicated in parentheses.

Responsiveness by Diagnostic Category

Table IV shows the responsiveness values for patients in several diagnostic categories. The results for males and females are usually quite similar; all results refer to males, except abdominal hernia and varicose veins.

To study this problem, it is useful to note an additional property of the responsiveness index values. For every type of patient the responsiveness value for the number of beds used per 1,000 population is almost exactly the sum of the admission-rate responsiveness value and the mean-stay responsiveness value. For example, Table IV shows that for all persons the

TABLE IV.—Responsiveness by Diagnostic Category

Diagnostic Category	Beds Used* (1)	Cases Treated* (2)	Mean Stay (3)
Acute appendicitis	15 (36)	—16 (33)	31 (17)
Acute upper respiratory infections ..	257 (100)	153 (52)	104 (74)
Peptic ulcer	85 (52)	29 (40)	56 (51)
Abdominal hernia (female)	139 (44)	52 (22)	87 (44)
Haemorrhoids	114 (62)	70 (48)	44 (24)
Tonsils and adenoids	55 (46)	23 (38)	33 (38)
Arteriosclerotic heart disease	222 (70)	114 (51)	108 (99)
Malignant neoplasms	58 (30)	68 (29)	—10 (20)
Varicose veins (female)	140 (70)	78 (41)	62 (67)
Males	104 (25)	64 (20)	41 (18)
Females	87 (26)	52 (26)	35 (25)
All persons	96 (22)	58 (23)	36 (18)

* Per 1,000 population, 1960. Standard errors indicated in parentheses.

beds-used responsiveness of 96 can be divided into a cases-treated responsiveness of 58 and a mean-stay responsiveness of 36. The importance of this property is that it allows us to see not only how the number of beds used for each particular disease responds to bed scarcity but also how much of this responsiveness is concentrated in each of the two components of bed use.

The responsiveness values for acute appendicitis and acute upper respiratory infections are quite satisfactory. Appendicectomies are very unresponsive to bed scarcity, while the less serious respiratory infections are highly responsive; in addition, the former's responsiveness is properly concentrated on the

mean stay while the latter's is almost completely in the number of cases treated. The values for peptic ulcer also seem appropriate, showing a low overall responsiveness which is concentrated on the mean stay.

The results for the other disease groups are not as satisfactory. Cases of abdominal hernia and of haemorrhoids show greater than average overall responsiveness.

It would seem desirable that admissions for tonsillectomy and adenoidectomy should be highly responsive; in regions where beds are relatively more scarce tonsillectomy and adenoidectomy cases should occupy a smaller proportion of available beds. The decision to operate for tonsillitis has often been cited as an example of medical fashion that is not founded on medical knowledge. Nevertheless, one in twenty hospitalizations in Britain are for tonsillectomy and adenoidectomy. And, more important for our current discussion, the number of these procedures and the length of stay show very low responsiveness to bed scarcity.

In contrast, arteriosclerotic heart disease, including coronary, usually presents very serious medical cases. Nearly a third of these patients die in hospital. Despite this, the number of cases admitted is extremely responsive to bed scarcity.

Cases with malignant neoplasms also have a high hospital fatality rate; more than a fifth of all admissions die in hospital. This is about 50% higher than in the United States, reflecting in part the greater tendency in Britain to keep terminal cancer cases in hospital after there is no longer any hope of helping them. As the table shows, this length of stay is unaffected by relative bed scarcity.

Finally, bed use for varicose veins is highly responsive with the greater proportion of this due to the responsiveness of the number of cases treated.

It is difficult to understand why such strikingly unsatisfactory results should have been found. It may be possible that some of the results can be explained by differences among the regions in the actual incidence of the diseases. Since the hospital statistics on which these studies have been based provide the only measures of morbidity, the influence of area-specific incidence rates could be studied only for those diseases in which mortality is a good indicator of morbidity. Doing so for heart disease and cancer does not suggest that the existing variability would explain our results.

We shall not try to offer any behavioural explanations of the individual responsiveness values. To do so properly we would have to develop a complex sociology of medical admission and treatment decisions based on the way in which doctors diagnose and treat each type of disease. But again, it is likely that any explanation should begin by recognizing that the doctors themselves are not aware of the allocation patterns that they have established.

Summary and Conclusions

The availability of non-mental hospital beds per head differs significantly among the regions. The number of cases treated is 50% more responsive to relative bed scarcity than the average duration of stay; this is true for all types of patients taken together, for males and females separately, and for most diagnostic categories that have been studied. This surprising and uncomfortable finding cannot be explained away in terms of underlying environmental factors. The observed pattern of behaviour can be understood if we recognize that the doctors themselves may not be aware of the implicit choices they have made, if we consider the institutional framework within which

admission and discharge decisions are made, and if we realize the pressures which unconsciously induce medical people to choose as they do.

The care of older patients is no more responsive to bed scarcity than the rest of the community. Responsiveness values for specific diagnostic categories indicate that when beds are scarce the wrong type of work may be neglected.

Neither the statistical results nor the behavioural explanations can be considered a proved description of the way in which differences in relative bed scarcity affect admission rates and durations of stay. Nevertheless, the evidence does strongly suggest that a previously unrecognized problem has been uncovered. This finding has implications for current hospital practice, for the changes proposed in the *Hospital Plan* (1962), and for future research in this field.

Hospital care to-day may be improved if individual doctors and regional medical committees compare the current use of hospital beds in their own areas with the pattern of the rest of the nation. In particular, those regions in which beds are relatively scarce might consider having shorter stays in order to treat more cases and might question whether the wrong types of cases are now being kept out of hospital in their area.

The *Hospital Plan* calls for a major reduction in acute beds during the next decade. Some of this will be facilitated by improved medical techniques which reduce duration of stay. But the reduced number of beds will also mean a change in the composition of work done in each region. It is important that everyone concerned be aware of which patients are being affected by the reduction of beds. This is especially true for regions such as Liverpool, where, per capita, the acute-bed supply is to be reduced by more than 35%.

Finally, these results indicate the need for more detailed studies of the effects of bed scarcity. Regional differences in the numbers of cases treated in each disease category are likely to imply differences in the standards used by the general practitioner or the consultant to decide whether a person should receive hospital care. A study of the differences in these standards would provide a basis for evaluating the medical effect of the greater responsiveness of admission rates. At the same time, such a study could perhaps identify the type of person who does not get hospital care for a particular disease when beds are more scarce. In doing these things, it would indicate more specifically the nature of the implicit allocation choice that doctors are currently making. This could both improve the current operation of the health service and help in planning future development. (Feldstein, 1963.)

I am grateful to several people for their contributions to this study: Mr. A. Barr, Professor W. M. Gorman, Dr. M. Heasman, and several colleagues at Nuffield College for discussions about the research; Miss Dorothy Milne and Mrs. Anne Godfrey for assistance with the calculations; and the Oxford University Computing Laboratory for the use of facilities. This paper presents part of the results of a broader study on the operation and economics of the National Health Service. The research is assisted by grants from the Ministry of Health, the United States Public Health Service, and the Health Information Foundation.

REFERENCES

- A Hospital Plan for England and Wales* (1962). Cmd. 1604. H.M.S.O., London.
- Feldstein, M. S. (1963). *Lancet*, 1, 491.
- (in preparation). "Hospital Bed Scarcity: An Analysis of the Effects of Inter-regional Differences."
- General Register Office (1963). *Report on Hospital In-patient Enquiry for the Year 1960, Part II*. H.M.S.O., London.
- Johnston, J. (1960). *Econometric Methods*. McGraw-Hill, New York.