

NBER WORKING PAPER SERIES

THE MEDICAL TREATMENT OF DEPRESSION,
1991-1996: PRODUCTIVE INEFFICIENCY, EXPECTED
OUTCOME VARIATIONS, AND PRICE INDEXES

Ernst R. Berndt
Anupa Bir
Susan H. Busch
Richard G. Frank
Sharon-Lise T. Normand

Working Paper 7816
<http://www.nber.org/papers/w7816>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2000

Funding support from the John D. and Catherine T. MacArthur Foundation, the National Institute for Mental Health, and the National Science Foundation is gratefully acknowledged. We gratefully acknowledge comments from Joseph Newhouse and Will Manning. The views expressed herein are those of the authors, and do not necessarily reflect the views of the research sponsors or the institutions with which they are affiliated.

© 2000 by Ernst R. Berndt, Anupa Bir, Susan H. Busch, Richard G. Frank, and Sharon-Lise T. Normand. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Medical Treatment of Depression, 1991-1996: Productive Inefficiency,
Expected Outcome Variations, and Price Indexes

Ernst R. Berndt, Anupa Bir, Susan H. Busch, Richard G. Frank, and Sharon-Lise T. Normand

NBER Working Paper No. 7816

July 2000

JEL No. I11, O33

ABSTRACT

We examine the price of treating episodes of acute phase major depression over the 1991-1996 time period. We combine data from a large retrospective medical claims data base (MarketScan™, from the MedStat Group) with clinical literature and expert clinical opinion elicited from a two-state Delphi procedure. This enables us to construct a variety of treatment price indexes that include variations over time in the proportion of “off-frontier” production, as well as the corresponding variations in expected treatment outcomes. We also incorporate the fact that the no treatment option (“waiting list”) frequently results in spontaneous remission of depressive symptoms.

We find that in general the incremental cost of successfully treating an episode of acute phase major depression has generally fallen over the 1991-96 time period. Based on hedonic regression equations that account for the effects of changing patient mix, we find price reductions that range from about -1.66% to -2.13% per year. An implication of this is that, since expenditures on depression are thought to be increasing since at least 1991, the source of the spending increases is volume (quantity) increases, and not price increases.

Ernst R. Berndt
Sloan School of Management
MIT, E52-452
50 Memorial Drive
Cambridge, MA 02142
(617) 253-2665
and NBER
eberndt@mit.edu

Anupa Bir
Harvard Medical School
Department of Health Care
Policy
180 Longwood Avenue
Boston, MA 02115
(617) 432-0305
bir@hcp.med.harvard.edu

Susan H. Busch
Harvard Medical School
Department of Health Care
Policy
180 Longwood Avenue
Boston, MA 02115
(617) 432-3430
busch@hcp.med.harvard.edu

Richard G. Frank
Harvard Medical School
Department of Health Care
Policy
180 Longwood Avenue
Boston, MA 02115
(617) 432-0178
and NBER
frank@hcp.med.harvard.edu

Sharon-Lise T. Normand
Harvard School of Public Health
and Harvard Medical School
Department of Biostatistics
180 Longwood Avenue
Boston, MA 02115
(617) 432-3260
sharon@hcp.med.harvard.edu

THE MEDICAL TREATMENT OF DEPRESSION, 1991-1996: PRODUCTIVE INEFFICIENCY,
EXPECTED OUTCOME VARIATIONS, AND PRICE INDEXES

by Ernst R. Berndt, Anupa Bir, Susan H. Busch,
Richard G. Frank and Sharon-Lise T. Normand

I. Introduction

Spending on mental health services continued to grow at rates above changes in general prices in the 1990s.¹ A common perception is that these increases in mental health spending predominantly reflect increasing costs of treatment.² But are increased mental health expenditures due to price increases, quantity (volume) increases, or some combination? This question raises fundamental issues concerning the value of services as well as important measurement issues.

Recently researchers and health care policy analysts have renewed professional interest in constructing price indexes for episodes of treatment for an illness, rather than focusing measurements on changing medical care input prices. By pricing episodes of care, the overall cost impacts of changing treatment modalities can be taken into account, incorporating therapeutic substitution among inputs, technological progress, and the effects of cost containment policies.³ Whether pricing episodes of care rather than medical care input prices results in greater or less measured inflation is unclear. In a series of studies covering between eight and sixteen representative medical conditions treated at the Palo Alto Medical Center from 1954 to 1981, Scitovsky found that an aggregate of costs per episode of treatment across illnesses increased more rapidly than the BLS medical care consumer price index (CPI) from 1951 to 1964, and again from 1964 to 1971. Between 1971 and 1981, however, the difference between the two series was insignificant.⁴ Scitovsky noted that one drawback of her cost of illness calculations, as well as of the BLS' medical care price indexes, is that neither adequately accounts for quality change, and related changes in health outcomes.⁵

Using data on heart attack treatments from 1983 to 1994 that explicitly take into account changing mortality outcomes, Cutler et al. report that treatment price indexes increased substantially less rapidly than an input price index similar to the medical care CPI (2.3% vs. 6.7% per year), and that over this time period the real cost of treatment for heart attacks actually fell 1.1% annually.⁶

In another set of studies, Shapiro et al. constructed price indexes for cataract surgery, 1969-93, and found that over this time interval a CPI-like input price index increased by a factor of nine, while a preferred

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

alternative price index incorporating realized reduced levels of hospital services increased by only a factor of three.⁷ However, neither of the Shapiro et al. indexes incorporated widely-heralded improvements in the quality of medical outcomes from cataract surgery.

Berndt, Busch and Frank (2000) constructed price indexes for the treatment of acute phase depression, 1991-95.⁸ As a first step to control for quality of care and outcomes, they limited the analysis to those episodes of treatment that adhered to depression treatment guidelines published by the Agency for Health Care Policy Research (AHCPR) and the American Psychiatric Association (APA).⁹ Using a strict interpretation of these treatment guidelines, they found that only about 23% of observed depression treatment episodes could be included in the price index calculations, whereas when a less stringent interpretation of guidelines was employed, about 50% of the observed episodes were included. Moreover, over the 1991-95 time period, with the strict interpretation, nominal treatment price indexes fell by about 30%, whereas with the broader interpretation, they fell by about 3%.¹⁰

This use of published treatment guidelines as a proxy measure for quality of care and expected outcomes may be preferable to ignoring outcome variations altogether, but it imposes a strong restriction on the shape of the medical treatment production function. Use of the guideline criterion implies that simply by observing input quantity combinations, one assigns each observed treatment episode either a zero (not compatible, not effective) or a one (guideline compatible, satisfactory outcome and equal across all guideline compatible treatments). Thus the implicit depression treatment production function is a step function, with a single step. Moreover, by focusing only on those episodes meeting guideline treatment standards, 50-75% of delivered care is overlooked. Finally, use of the step function production relationship is very simplistic, and does not make use of a great deal of clinical and medical knowledge.

In this paper we report on results of several significant extensions to previous research on price indexes for the treatment of depression. Although we continue here to study the acute phase outpatient treatment of depression, we extend the time series to include 1996. More importantly, however, we now combine data from a large retrospective medical claims data base (MarketScanTM, from The MedStat Group) with clinical literature and expert clinical opinion elicited from a two-stage modified Delphi procedure. This enables us to construct a variety of treatment price indexes that include variations over time in the proportion of "off-frontier" production,

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

as well as the corresponding variations in expected treatment outcomes.¹¹ We also incorporate the fact that the no treatment option (“watchful waiting”, or “waiting list”) frequently results in spontaneous remission of depressive symptoms. We therefore focus on the *incremental* gains in outcome obtained from treatment over and above waiting list. By integrating knowledge concerning the efficacy and effectiveness of alternative treatments based on micro clinical trial data with the judgments of clinical experts and changes in treatment patterns observed in a large medical claims data base, we construct price indexes that capture the system-wide effects of changing treatment practice.¹²

II. Depression and Its Alternative Treatments

According to the APA, major depression is diagnosed when a very specific set of clinical criteria have been met; details concerning these criteria have been discussed elsewhere.¹³ The 12-month prevalence of major depression in the U.S. has been estimated at about 10%, and is about twice as high for women as men.¹⁴ Although a single episode of the illness is self-explanatory given the diagnostic criteria, recurrent depression is defined by two or more major depressive episodes each separated by at least eight weeks of return to usual functioning.¹⁵ Episodes of depressive illness come and go, last from several weeks to several months, and are often followed by periods of relatively normal mood and behavior. Untreated, the average depressive episode lasts from four to six months, although spontaneous remission is common. Between 20% and 35% of episodes of depression display persistent symptoms, and when these persistent symptoms last for 24 or more months, these cases are referred to as chronic depression. Approximately 50% of all people having a depressive episode can be expected to have a recurrence, usually within two or three years. Once an individual has a second episode, additional recurrence is 70% likely.

The acute phase treatment of major depression typically aims at the stabilization of most acute symptoms, usually over a 12-16 week period. Here we focus on the acute phase, for which there have been rapid advances in both psychotherapy and pharmacotherapy treatments.¹⁶ New psychotherapy treatments include interpersonal therapy, behavior therapy, and cognitive behavior therapy. In pharmacotherapy, the monoamine oxidase inhibitors and the cyclic antidepressants of the 1950's to 1970's (including the widely used tricyclics, TCAs) have now been largely replaced by selective serotonin reuptake inhibitors (SSRIs) such as Prozac, Zoloft and Paxil. Based on clinical trials, the efficacy of the SSRIs has been shown to be about the same as that of the

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

older TCAs for mild to moderate depression. However, since the SSRIs are associated with lower risk of overdose and are generally more easily tolerated by patients, in the naturalistic setting in the U.S., SSRIs typically dominate the TCAs as first line therapy.¹⁷ The advantages of the SSRIs come at a significantly higher pecuniary cost than most TCAs. In practice, combination treatments involving psychotherapy plus antidepressant medications, in varying proportions, constitute a very common treatment strategy.

As in previous studies of treatment for depression, we identify outpatient treatment episodes for depression using retrospective claims data from the publicly available MarketScanTM database. These data consists of 1991-96 medical claims from four large self-insured employers that offered more than 25 health plans to an average of 426,000 employees and their dependents. We identify all ambulatory claims associated with either single or recurrent episodes of depression, as defined by the ICD-9 primary diagnostic codes 296.2x or 296.3x.¹⁸ We construct episodes of care that include all treatments given in the sixteen weeks following initiation of care for depression, with distinct episodes defined in terms of eight or more week gaps in care (including gaps in prescription drug supply).

We classify each episode of care along two dimensions: type of treatment, and type of patient. Treatment types are identified using current procedural (CPT-4) codes, drug claim and billing information contained in the medical claims, and are classified according to type and frequency (pharmacotherapy with TCAs, SSRIs, etc., psychotherapy, office visits, medical management visits, and various combinations of these).¹⁹ Patients are characterized by medically relevant demographic (age, gender) and illness characteristics (history of substance abuse, presence of medical comorbidity, etc.), and whether the ambulatory practice setting where treatment was delivered was a specialty mental health or general medical setting. An example of a treatment/patient cell is women between ages 18 to 45, with no medical comorbid conditions and no recent substance abuse problems, treated with an SSRI for at least 60 days plus three or more psychotherapy visits in a specialty mental health setting over a 16-week period.

To eliminate the most severe cases in the data, and because information on care received in hospitals is incomplete, we exclude all depressive episodes involving inpatient (hospital) claims. Since the claims data end in December 1996, we eliminate both right-censored (new episodes initiated after August 31, 1996) and left-censored (treatments underway in the first eight weeks of 1991) episodes. This leaves us with 13,098 episodes of

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

care. Based on these criteria, we identify about 200 patient-treatment cells. To reduce the burden of the expert panel's deliberations, we eliminate most of the treatment/patient cells having fewer than 30 patients over the six years. This results in 120 treatment/patient cells covering 9,054 episodes (69% of ambulatory treatment episodes). Although 867 patient visits involved an ICD-9 depression diagnosis, their medical claims contained no information on any mental health treatment (n = 816) or there were no expenditures (n = 51). When these "no treatment" episodes are deleted, the number of remaining episodes is 8,187.

Direct medical spending for each treatment episode was calculated using actual transactions data. All payments made by the insurer to the provider, and any cost-sharing assigned to the patient (e.g., patient out of pocket copayment for psychotherapy and prescription drugs) were summed into a nominal dollar total for each treatment episode.

III. Outcomes Assessment: The Expert Panel

To obtain expected outcome assessments for each of the 120 treatment/patient cells, we combined research data on clinical efficacy and effectiveness with expert clinical opinion, using a two-stage modified Delphi method.²⁰ Here we summarize the outcome assessment process; additional details of the methods employed are given in Normand et al. [1999].

Predictive distributions of outcomes for hypothetical cohorts of 100 patients receiving a range of treatments were obtained from a panel of ten experts; this resulted in ten distinct probability distributions for each of the 120 treatment/patient cells.²¹ Expected treatment outcomes were characterized in terms of changes in the Hamilton Depression Rating Scale (HDRS) score, a commonly employed clinical measure quantifying symptoms of depression 16 weeks after entry into treatment for hypothetical cohorts of 100 patients. Standard cutpoints were used to describe the distribution of treatment outcomes: HDRS < 8 representing no depression (full remission); $8 \leq \text{HDRS} \leq 12$, mild depression; $13 \leq \text{HDRS} \leq 17$, moderate depression; and $\text{HDRS} \geq 18$, major depression (no change). Panelists were instructed to assume that at the time patients initiated treatment, their HDRS score was 22 (high likelihood of a major depressive disorder diagnosis).

The elicitation process consisted of three distinct stages. Previous literature has called such a process a modified Delphi technique.²² Similar elicitation processes have been used extensively in the construction of practice guidelines by health services researchers.²³

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

In the first stage of the Delphi process, based on a 1975-1998 MEDLINE literature review, a psychiatrist provided each panelist a summary of published results of treatment efficacy and effectiveness for each of the 120 treatment/patient cells, thereby generating an evidentiary platform for elicitation. This literature review also examined outcome evidence from the "no active treatment" or "waiting list" treatment option. Since placebo treatment in clinical trials typically involves a substantial volume of brief counseling sessions and thereby frequently results in large placebo effects, outcomes evidence on the "waiting list" vs. placebo treatment was reported when available.

In the second stage, the results of the literature review were used in a postal elicitation of outcome distributions (no, mild, moderate or major depression after 16 weeks of treatment) for each of the 120 treatment/patient cells. Experts were directed:

"Using your best clinical judgment, consider an average group of 100 patients presenting with major depressive disorder with an entry HDRS score of 22 in 1998 to an average specialty mental health or primary care physician. After 16 weeks of the indicated treatment, how many patients will most likely fall into each of the following categories: depression-free, mildly depressed, moderately depressed, or no change. When completing the survey, please keep in mind the impact of compliance and the average circumstances of routine practice on symptom reductions."

This postal elicitation yielded 480 parameters from each expert (four parameters, only three of which were independent due to the 100% adding-up identity, for each of the 120 treatment/patient cells). Cells having a "substantial" disagreement among experts were identified (coefficient of variation > 0.5). At the end of this second stage, discordant judgments occurred in 47 of the 120 (39%) of the cells. Disagreement tended to occur disproportionately in those treatment cells for which the number of published studies was small.

In the third stage, re-elicitation occurred, with the goal of creating an opportunity for experts to modify their subjective probability assessments but not forcing group consensus. A face-to-face panel meeting was convened in January 1999, and each expert was given the groups' rating distribution for each discordant cell (as well as his/her own probability assessment, but with each expert blinded to what his/her colleagues reported

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

individually). After literature review and group discussions, each panelist revised opinions for discordant cells confidentially. No substantial disagreements (coefficient of variation > 0.5) remained after this re-elicitation.

The revised probability assessments for the discordant strata were then combined with the initial assessments for the non-discordant data. The panelists' distributions were summarized and combined using arithmetic means, medians and other summary distribution measures across the ten expert raters.

IV. Results: Costs and Expected Outcomes

Because the elicitation process involved the distribution of outcomes for each patient/treatment cell, one can examine different cutpoint outcomes. Full remission is defined as an HDRS score of < 8, while "partial remission" ($\text{HDRS} \leq 12$) represents a roughly 50% improvement by treatment week 16 after an initial HDRS score of 22.

The weighted average of the median outcome result from the expert panel elicitation process for each treatment (averaged over patient type), as well as 1991-1996 average spending from the MedStat data, are reported in Table 1.²⁴ Also reported in Table 1 are the incremental costs per depression free case and per case of 50% improvement. The relevant increment is the likelihood of being depression free (or 50% improved) relative to the "waiting list" outcome. A number of findings are particularly noteworthy.

First, as seen in the column "Median Probability of No Depression" (i.e., full remission), according to the expert panel, the probability of a full remission after 16 weeks of treatment ranges from 0.15 to 0.35, with the lowest 0.15 probability associated with no treatment ("waiting list"), while the 0.34 - 0.35 highest probabilities are combinations of 60 days of SSRIs plus ≥ 4 psychotherapy (PT) visits, and additional treatments involving anxiolytics, trazodone and TCAs with greater psychotherapy visits. Interestingly, 10-24 PT visits with no drug therapy has almost the same probability of full remission as 60 days of SSRIs plus ≥ 4 PT visits (0.34 vs. 0.35). The results for 50% improvement are consistent with this finding (0.63 and 0.64). Thus to a first order approximation, these two alternative treatment modalities are essentially on the same medical production function isoquant. Also, the marginal expected benefit when ≤ 30 days' SSRI is augmented by 1-3 PT visits vs. no PT visits is almost nil (0.28 vs. 0.29), for the difference in the probability of a full remission is zero. For the probability of 50% improvement the increase in expected outcome from adding visits is also very small, from 0.45 to 0.46. This suggests the possibility that the addition of 1-3 PT visits to ≤ 30 days' SSRI represents a productive

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

inefficiency.

Second, the probability of a partial remission ($\geq 50\%$ improvement in HDRS) is typically about twice as large as the probability of a full remission after 16 weeks of treatment. As seen in Table 1, the "no mental health treatment" probability of attaining a partial remission is 0.34, the same as that for one psychotherapy visit, and only slightly smaller than 1-2 office visits (0.35 - 0.36) that involved no psychotherapy.

Third, about 40% of the treatment episodes involve medical treatments with little or no incremental effectiveness over "waiting list" (probability of full remission 0.15 - 0.17). Though not shown in Table 1, we find that these treatment episodes accounted for about 20% of all spending on treatments.

Fourth, as seen in the "Average Cost per Episode" column of Table 1, if one looks only at the mean spending per depressive episode and ignores outcome variations entirely, the spending per case varies from zero (no treatment) and \$53 (one office visit, but no PT) to a high of \$1,167 (60 days' SSRI, trazadone and PT). Moreover, the cost per episode when treatment consists of 10-24 PT visits and no medication is virtually identical to the cost per case with 60 days' SSRIs plus ≥ 4 PT visits (\$1055 vs. \$1050).

Fifth, if, however, one defines the output of treatment for depression to be the expected probability of full remission over and above that obtained from doing nothing, then one can compute the average spending per incremental remission. Results from such incremental calculations, done separately for full and 50% remissions, are given in the last two columns of Table 1. Average spending per incremental full remission is infinite for those treatments no better than the "waiting list" or "no active treatment" outcomes (e.g., 1-2 office visits, antipsychotics alone), are above \$10,000 for 1 PT visit, SSRIs and trazadone ≤ 30 days along with anxiolytics and PT, and for lithium alone. The lowest costs per incremental full remission are associated with (i) SSRI > 30 days, with an anxiolytic but no PT, \$1,853; (ii) SSRI > 30 days and no PT, \$2,367; (iii) 1 brief office visit \$2,618; (iv) SSRI > 30 days, anxiolytic, 1-3 PT visits, \$2,765; and (v) SSRI < 30 days and no PT, \$3,447. As seen in the bottom row of Table 1, averaged over all cases over the 1991-96 time period, the cost per episode is \$473, the cost per incremental full remission is \$5,909, and the cost per incremental partial remission is \$3,377.

V. Results: Changes Over Time in Episode Expenditure and Price Indexes

The results of the expert panels' outcome assessments can be combined with changes over time in the shares of observed treatment modalities (and their costs) to create various expenditure and price indexes. We

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

now consider a number of such measures.

In constructing various expenditure and price indexes, we make alternative assumptions concerning how to treat the approximately 32% of identified episodes of care (3,910 of 12,097) for which the expert panel did not make an expected remission probability rating. The alternatives we consider are: (i) use only these 8,187 episodes with panel outcome ratings, and exclude the no mental health treatment episodes; (ii) assign a treatment-specific worst case outcome from the panel's outcome rating to all observed episodes in any unrated treatment/patient cells; and (iii) assign the treatment specific median outcome to all observed episodes in any unrated treatment/patient cell.

The simplest calculations involve tracking over time the average spending per full or partial expected remission, given the changing observed prices and quantities of the treatment/patient cells. Results of such calculations are given in the top panel of Table 2. There it is seen that from 1991 to 1996, the average spending per expected full remission, and per partial remission, remains unchanged at 100 when only the rated cases are included. When unrated cases are assigned the worst outcome, the average expenditure indexes fall from 100 in 1991 to 84 (full remission) and 86 (partial remission) in 1996. One reason underlying this decline in average spending per remission when unrated cases are assigned a worst case outcome is that the proportion of cases rated increases from 62% in 1991 to 73% in 1996, and thus the proportion assigned the worst outcome declines from 38% in 1991 to 27% in 1996. When the unrated cases are assigned a median outcome, the average expenditure per full expected remission stays unchanged in 1996 from 1991, but for partial remissions it falls from 100 in 1991 to 90 in 1996.

The expenditure index per incremental expected full or partial remission is both more relevant and theoretically attractive than average expenditure indexes that fail to distinguish average from incremental (marginal) treatment outcomes.²⁵ The incremental remission is defined as the difference between the expected outcome and the outcome that would be expected if the patient received no treatment; incremental measures are given in the bottom panels of Table 2. Based on only the panel's rated outcome episodes, the spending per incremental expected full remission falls from 100 in 1991 to 83 in 1996, an average annual growth rate (AAGR) of about -3.6%. When unrated episodes are added and assigned either the worst or median outcome, by 1996 the index falls to 78 and to 88, respectively, with AAGRs of -4.8% and -2.5%. For incremental partial remissions, use

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

of rated only, rated plus unrated assigned worst outcome, and rated plus unrated assigned median outcome, yield 1991-96 AAGRs of -2.1%, -4.6% and -2.1%, respectively.

One interesting feature of these results is that in almost all cases, expenditure indexes per incremental expected remission fall more rapidly than average expenditure per remission (only in the unrated assigned median outcome case, with partial remissions, does this not occur, and there the indexes coincide in 1996).

While these relatively simple expenditure indexes are of interest, they are not quite the same as conventional price indexes. We have computed a Laspeyres price index with 1991 fixed quantity weights, a Paasche fixed quantity index with 1996 fixed weights, and a Fisher-Ideal price index (the geometric mean of the Laspeyres and Paasche).²⁶ As seen in Table 3, without any reference to outcome, the treatment price index falls from 100 in 1991 to 95.0 (Laspeyres), 99.4 (Paasche), and 97.2 (Fisher-Ideal) in 1996. When the elementary unit is defined as the price per expected incremental full remission, the Laspeyres, Paasche and Fisher-Ideal price indexes show a small rise to 103.4, 104.0 and 103.7, respectively.

In 1992 and 1993, the Laspeyres index for incremental full remissions shows a decline in nominal prices with a rise occurring between 1994-1996. The Paasche index shows price increases in 1993-1995. Table 3 also reveals slightly increasing prices for the expected incremental partial remissions. The Laspeyres index moves from 100 in 1991 to 103.9, in 1996, while the Paasche moves from 100 to 104.7 and the Fisher-Ideal from 100.0 to 104.3 over the same time period. Hence the trends in price indexes are moderately sensitive to how one treats expected outcomes. While outcomes-adjusted price indexes increase slightly in Table 3, the corresponding average expenditure indexes in Table 2 fall, and even more so when incremental remissions are employed. Note that during the 1991-96 time period, increased levels of management have been exercised over the mental health benefit. One implication of this is that the patient population being treated may have been changing (along with the mix of treatment bundles), thereby affecting both expected outcome and cost. These changes are not incorporated into the fixed weight price indexes in Table 3.

To account for the effect of changing patient mixes on computed price indexes, we employ hedonic-like equations. We have specified and estimated a number of such models, based on data from the 8,187 rated treatment episodes.²⁷ We delineate eight patient categories, depending on whether medical comorbidity is

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

present, whether male, if female whether over age 50, and whether there is comorbid substance abuse. The eight cases are, respectively (Y for yes, N for no) 1 – NYNY; 2 – NNNY; 3 – YYNN; 4 – YNYN; 5 – YNNN; 6 – NYNN; 7 – NNYN; and 8 – NNNN. We estimate three alternative hedonic equations for the price per expected full remission, all of which are reported in Table 4. The specifications differ across the three models. In Equation (1) of Table 4, the left hand variable is the natural log of spending per episode of treatment. The right hand variables include the probability of a full remission associated with the patient's treatment, dummy variables for 7 of 8 patient categories, and year dummies (1991 is excluded). Equation (2) has a left hand variable that measures the natural log of the price divided by the incremental probability of full remission minus a small constant (0.001, to ensure feasibility of taking logarithms). The right hand side variables are the same as in Equation (1), only now the probability of full remission is excluded as a regressor, since incremental remission probability is incorporated into the dependent variable. Equation (3) is the same as Equation (1) except that the probability of full remission is replaced by the log odds of the probability of expected full remission. Hence the difference between Equations (1) and (3) is only the way in which expected outcomes appears as a regressor – linear, or log odds ratio. We estimate these equations by ordinary least squares; standard errors are heteroskedasticity-robust.

As seen in Table 4, variations in patient categories have significant and substantial effects on treatment costs, with categories 2 (comorbid substance abuse) and 4 (comorbid medical conditions) being particularly large and positive. A second important result is that, ceteris paribus, treatments having greater probability of full remission have higher costs. In Equation (1), the remission probability enters in linear form, whereas in Equation (3) it is in the form of a log odds ratio; in both cases, the coefficient is positive and highly significant, with their robust t-statistics being very similar.

In terms of time effects as captured by yearly dummy variables, the results reported in Table 4 display considerable stability across models, with all three equations having estimated price reductions. The magnitude of the reductions from 1991 to 1996 ranges from 8.3% (in price per incremental remission) in Equation 2 to 12.8% in price per remission in Equation 1, with AAGRs from -1.66% to -2.13%. The 1991 to 1996 differences are significantly different from zero at the ($p < 0.05$) level for Equation 1 and 3 and at $p < 0.10$ for Equation 2. For most years, prices are estimated to be lower than in 1991. Yet only in the later years are the differences from 1991 statistically significant. Moreover, the estimates are not very precise in several cases

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

($p < 0.10$). The differences between the hedonic price indexes in Table 4 and the alternative price indexes from Table 3 appear, therefore, to result from a changing and increasingly complex mix of patients, and changes in treatment bundles, over the six year period; we discuss this further below.

Table 5 presents corresponding hedonic regression results for expected partial remissions. The specifications mirror those described for Table 4. Overall, the results are qualitatively quite similar across specifications. The two equations which estimate the price per expected partial remission arrive at nearly identical estimates of the 1991-1996 price change, with both models having an estimated significant reduction of 12.2% over the six years. The price per partial remission in Equation 2 is estimated to have fallen by about 4.8% over the six years, but this estimate is not statistically significant.

It is important to note that the hedonic regressions in Tables 4 and 5 suggest a rather different set of price movements than the simple price indexes reported in Table 3. The main reason for the difference is that the variables describing the composition of cases are included in the regressions, but are not properly controlled for in the traditional price index calculations. For example, the mix of patients has been changing over time. To illustrate this point, consider depressed patients with medical comorbidities. In 1991, about 13% of episodes involved comorbid medical diagnoses, but by 1995 nearly 23% of episodes involved such conditions, and in 1996 this share fell to 19%. Comorbidities are known to complicate treatment and make it more costly. Thus, when the hedonic model controls for the changing patient composition of the treated population, as well as changes in expected outcomes, the price indexes show declines in the nominal prices of expected full and partial remissions.

It is also interesting to note that when price is measured as spending per incremental expected full or partial remission as in Equation (2) in Tables 4 and 5, a great deal of extra variation is introduced into the data (recall from Table 1 that some prices per incremental remission approach infinity), the yearly time coefficients are all less precisely estimated, and the goodness of fit declines dramatically. This is an area that merits further investigation.

Together these results suggest that the price of treating cases of major depression, conditional on expected outcome, generally fell during the 1991-1996 period. To place these results in context, we make two sets of observations. First, for thirty years a basic assumption in health care policy has been that the prices of medical treatments are increasing at rates far in excess of general inflation. Mental health treatments have been

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

viewed as particularly inflationary (Phelps, 1999). Our results reported here refute that conventional wisdom. The second observation is that these results differ notably from standard price indexes that focus on discrete elements of treatment. Table 6 reports on a set of the U.S. Bureau of Labor Statistics Producer Price Indexes (PPIs) that are related to components of mental health care. In all cases, the PPIs rise, and rise at higher rates than any of the estimates presented above – be they expenditure indexes, traditional price indexes, or indexes drawn from the estimated hedonic regressions. Thus, focusing on outcomes of care related to specific episodes of treatment yields different strikingly different price trends than does price measurement based on components of care.²⁸

VI. Discussion

The findings reported here suggest that the use of expected outcomes data, based on an expert panel elicitation procedure, is a feasible and reasonable approach to computing price indexes for medical treatments that take into account variations in expected outcomes, and in patient populations.

The usefulness and credibility of this procedure, however, hinge in large part on the underlying reliability of the effectiveness or outcomes measures. One check for face validity involves examining the relative impacts on remission probabilities of variations in patient comorbidities and demographics generated in this elicitation process, to make sure they are generally consistent with medical literature and practice. Results from such a validity check are reassuring. We find, for example, that the mean probability of full remission (across experts) is lower for patients with a medical comorbidity compared to patients without a comorbidity (0.22 vs. 0.25), and is lower for patients treated in primary care settings compared to those in specialty care (0.22 vs 0.24).

Since our data involves patients over the 1991-96 time frame, a second validity check is available to us. Specifically, we track each episode of treatment for 18 months following the 16 week treatment, to assess whether that patient had a recurrence of diagnosed depression in the form of a new episode. We then perform a simple logistic regression of the probability of the recurrence of depression, where the panel's mean estimated probability of full remission from the 16-week treatment modality initially received by that individual is a regressor. The estimated coefficient on the probability of full remission regressor is negative and significant ($p < 0.001$), indicating that increasing the probability of an expected full remission is associated with a decreased risk of having a repeat episode of care. This result is robust to inclusion of patient and treatment characteristic dummy

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

variables, and provides additional support for the reliability and validity of the outcome probabilities obtained through the elicitation process.

The results we have obtained suggest that incorporating off-frontier treatment variations over time does not materially change findings from previous research on treatment of depression: the cost of treating an episode of depression has fallen from 1991 to 1996. To see this in a different way, in Figure 1 we plot predicted treatment spending on the vertical axis, and probability of full remission on the horizontal axis, based on parameter estimates from Equation (3) in Table 4. We plot predicted spending for 1991, and for 1996, when all other regressors are evaluated at the overall 1991 through 1996 sample means. As can be seen in Figure 1, while predicted treatment costs increase at an increasing rate as the expected probability of a full remission rises (a movement along both curves), there has been a noticeable downward shift in the treatment cost function from 1991 to 1996. An implication is that, since expenditures on depression are thought to be increasing since at least 1991, the source of the spending increases is volume (quantity) increases, and not price increases. An implication of this is that given a budget for treatment of depression, more could be accomplished, in terms of producing health, in 1996 than in 1991. While this result is at odds with many long-held notions about mental health care, it is consistent with the findings of the Surgeon General's Report on Mental Health.²⁹ That report noted advances in both the science and delivery of mental health treatments in the U.S.

Finally, in this research we have confined our attention to the supply side, where prices are envisaged as total receipts obtained by providers for treatment episodes (remissions) delivered by them to patients. Evaluating marginal prices from the point of view of the consumer (patient) is considerably more complicated. Consumers' marginal copayments for treatments may be affected by nonlinear schedules involving the number of treatments, and by other forms of risk-sharing.³⁰ Patients' valuations of their time involved in treatment, and the relative time path of their response to alternative treatments, affect the marginal cost-benefit calculations as viewed by the patient.³¹ Also, patients suffering from mental illness may not be able to reveal their preferences for care according to the tenets of economic theory. The symptomatic benefits of treatment may well also spill over to ability to function, and to be productive while at work.³² Even more difficult issues emerge when evaluating the benefits of an entirely new treatment modality via reservation prices, and tracking out the impacts on future consumption and enjoyment.³³

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

Table 1: Spending and Expected Outcomes of Ambulatory Treatments for Depression

Treatment	No. of Strata	No. (%) of Episodes	Median Probability of:		Average Cost per Episode	Average Incremental Cost Per:	
			No. Depression	50% Improvement		No Depression	50% Improvement
1 Brief Office Visit	11	901 (10)	0.17	0.36	\$53	\$2618	\$2618
2 Brief Office Visits	1	42 (0)	0.15	0.35	95	∞	9405
≥ 3 Brief Office Visits	1	34 (0)	0.20	0.44	235	4700	2350
1 Psychotherapy Visit	9	978 (11)	0.16	0.34	136	13573	∞
2-3 Psychotherapy Visits	8	816 (9)	0.19	0.41	278	6930	3960
4-9 Psychotherapy Visits	9	1292 (14)	0.25	0.55	558	5577	2656
10-24 Psychotherapy Visits	9	637 (7)	0.34	0.63	1055	5553	3638
TCA ≤ 30 Days AND							
1-3 Psychotherapy Visits	1	22 (0)	0.20	0.45	280	5585	2539
> 3 Psychotherapy Visits	2	49 (1)	0.24	0.52	856	9507	4754
TCA >30 Days AND							
No Psychotherapy	2	30 (0)	0.20	0.45	280	5585	2539
1-3 Psychotherapy Visits	1	25 (0)	0.30	0.57	637	4243	2767
> 3 Psychotherapy Visits	2	63 (1)	0.35	0.63	864	4320	2980
SSRI ≤ 30 Days AND							
No Psychotherapy	2	65 (1)	0.20	0.45	173	3447	1915
1-3 Psychotherapy Visits	2	87 (1)	0.20	0.46	304	6070	2530
≥ 4 Psychotherapy Visits	3	147 (2)	0.25	0.50	877	8765	5478
> 3 Psychotherapy Visits, Anx.	2	18 (0)	0.32	0.60	837	4924	3220
Trazodone ≤ 30 Days, Anx., PT	1	19 (0)	0.20	0.50	756	15116	4724
SSRI >30 Days AND							
No Psychotherapy	11	552 (6)	0.28	0.60	308	2367	1184
1-3 Psychotherapy Visits	7	473 (5)	0.29	0.59	504	3593	2013
≥ 4 Psychotherapy Visits	9	801 (9)	0.35	0.64	1050	5249	3499
No Psychotherapy, ANX.	1	20 (0)	0.32	0.62	315	1853	1125
1- 3 Psychotherapy Visits, Anx	1	36 (0)	0.35	0.66	553	2765	1728
Trazodone ≤ 30 Days, Anx., PT	1	35 (0)	0.30	0.60	875	5827	3362
Trazodone > 30 Days, Anx., PT	2	61 (1)	0.35	0.66	1167	5833	3646
Heterocyclics > 30 Days, Anx., PT	1	39 (0)	0.28	0.58	557	4284	2321
Lithium Alone	1	20 (0)	0.20	0.45	538	10753	4888
Lithium + Anti-Depressants	2	67 (0)	0.26	0.55	820	7454	3905
Anti-psychotics Alone	1	15 (0)	0.15	0.35	476	∞	47647
Unspecified MH Treatment	8	894 (10)	0.16	0.37	757	75690	25230
No MH Treatment	9	816 (9)	0.15*	0.34*	-	-	-
Total	120	9054	0.23	0.48	473	5909	3377

*Median probability of being depression-free and 50%-improved were 0.18 and 0.38 respectively but were set to the minimum values (0.15 and 0.34) for calculating incremental spending.

Table 2: Alternative Expenditure Indexes for Treatment of Depression, 1991-1996 (1991 = 100)

Outcomes of Episodes Included	1991	1992	1993	1994	1995	1996
<i>Average Expenditure per Full Remission</i>						
Rated Only (n=8, 187)	100	103	108	106	103	100
Rated + Unrated Worst Case (n=12, 097)	100	99	97	95	90	84
Rated + Unrated Median Case (n=12, 097)	100	102	109	108	105	100
Percent Rated	62%	63%	68%	69%	71%	73%
Percent Meeting AHCPR Guideline Standards ^a	42%	42%	53%	54%	51%	49%
<i>Average Expenditure per Partial Remission</i>						
Rated Only (n=8, 187)	100	102	107	107	103	100
Rated + Unrated Worst Case (n=12, 097)	100	100	100	98	93	86
Rated + Unrated Median Case (n=12, 097)	100	101	104	103	95	90
<i>Average Expenditure per Incremental Full Remission</i>						
Rated Only (n=8, 187)	100	93	90	85	89	83
Rated + Unrated Worst Case (n=12, 097)	100	98	88	83	85	78
Rated + Unrated Median Case (n=12, 097)	100	100	96	92	95	88
<i>Average Expenditure per Incremental Partial Remission</i>						
Rated Only (n=8, 187)	100	102	98	92	93	90
Rated + Unrated Worst Case (n=12, 097)	100	100	91	86	86	79
Rated + Unrated Median Case (n=12, 097)	100	100	102	103	94	90

^aGuideline standards interpreted broadly. See Frank, Berndt and Busch, (1999).

Table 3: Alternative Traditional Price Indexes for Treatment of Depression, 1991-1996 (1991=100)

Price Index Formula	1991	1992	1993	1994	1995	1996
<i>Laspeyres</i>						
Price per Episode (no outcome adjustment)	100.0	97.5	93.7	92.2	95.4	95.0
Price per Incremental Full Remission	100.0	92.0	99.5	101.3	108.3	103.4
Price per Incremental Partial Remission	100.0	93.3	92.5	90.6	93.2	103.9
<i>Paasche</i>						
Price per Episode (no outcome adjustment)	100.0	98.3	97.7	97.1	97.5	99.4
Price per Incremental Full Remission	100.0	92.1	100.1	101.8	107.6	104.0
Price per Incremental Partial Remission	100.0	93.3	93.1	91.2	92.7	104.7
<i>Fisher-Ideal</i>						
Price per Episode (no outcome adjustment)	100.0	97.9	95.7	94.6	96.4	97.2
Price per Incremental Full Remission	100.0	92.0	99.8	101.6	107.9	103.7
Price per Incremental Partial Remission	100.0	93.3	92.8	90.9	93.0	104.3

Table 4: Hedonic Regression Results for Price Index of Expected Full Remission
(heteroskedasticity-robust t-statistics in parentheses)

Variable	Estimates		
	(1) ln (Price)	(2) ln (Price per Incremental Expected Full Remission)	(3) ln (Price)
Probability of Full Remission	***9.283 (77.24)	- -	- -
Pt Category 2	***1.551 (19.43)	***1.128 (13.69)	***1.509 (18.93)
Pt Category 3	***0.226 (7.33)	***0.927 (18.91)	***0.246 (8.01)
Pt Category 4	***0.243 (6.75)	***1.187 (20.33)	***0.268 (7.48)
Pt Category 5	***0.156 (6.14)	***0.696 (17.21)	***0.169 (6.70)
Pt Category 6	***-0.183 (6.71)	***-0.520 (16.24)	***-0.191 (7.00)
Pt Category 7	***-0.120 (2.16)	-0.069 (0.89)	** -0.131 (2.39)
92	*-0.080 (1.91)	0.008 (0.15)	*-0.074 (1.78)
93	-0.049 (1.21)	-0.063 (1.18)	-0.037 (0.93)
94	*-0.067 (1.69)	*-0.096 (1.83)	-0.056 (1.43)
95	***-0.128 (3.41)	-0.050 (1.01)	***-0.120 (3.24)
96	***-0.137 (3.50)	*-0.087 (1.67)	***-0.126 (3.25)
$\ln\left(\frac{P_o}{1-P_o}\right)$, p_0 = full remission probability	- -	- -	***1.703 (75.91)
Constant	***3.619 (77.64)	***8.836 (201.36)	***7.870 (192.83)
Observations	8187	8187	8187
R ²	0.40	0.16	0.41

* p-value < 10%
 ** p-value < 5%
 *** p-value < 1%

**Table 5: Hedonic Regression Results for Price Index of Expected Partial Remission
(heteroskedasticity-robust t-statistics in parentheses)**

Variable	Estimates		
	(1) ln (Price)	(2) ln (Price per Incremental Partial Remission)	(3) ln (Price)
Probability of Partial Remission	***6.046 (76.87)	- -	- -
Pt Category 2	***1.581 (19.68)	***1.289 (16.29)	***1.574 (19.60)
Pt Category 3	***0.247 (8.20)	***0.099 (3.16)	***0.258 (8.59)
Pt Category 4	***0.204 (5.98)	0.027 (0.77)	***0.215 (6.33)
Pt Category 5	***0.172 (6.92)	***0.079 (2.98)	***0.181 (7.27)
Pt Category 6	***-0.100 (3.77)	***-1.141 (4.84)	***-0.102 (3.83)
Pt Category 7	** -0.105 (1.99)	***-0.236 (4.00)	** -0.105 (1.98)
92	*-0.072 (1.78)	-0.048 (1.11)	*-0.071 (1.76)
93	-0.044 (1.13)	-0.046 (1.11)	-0.044 (1.13)
94	*-0.064 (1.67)	-0.042 (1.04)	*-0.064 (1.67)
95	***-0.131 (3.62)	-0.061 (1.57)	***-0.130 (3.60)
96	***-0.131 (3.45)	-0.049 (1.21)	***-0.131 (3.46)
$\ln\left(\frac{P_o}{1-P_o}\right)$, p_0 = partial remission prob.	- -	- -	***1.471 (77.51)
Constant	***2.852 (53.27)	***6.947 (201.36)	***5.873 (177.77)
Observations	8187	8187	8187
R ²	0.43	0.02	0.43

* p-value < 10%
 ** p-value < 5 %
 *** p value < 1 %

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

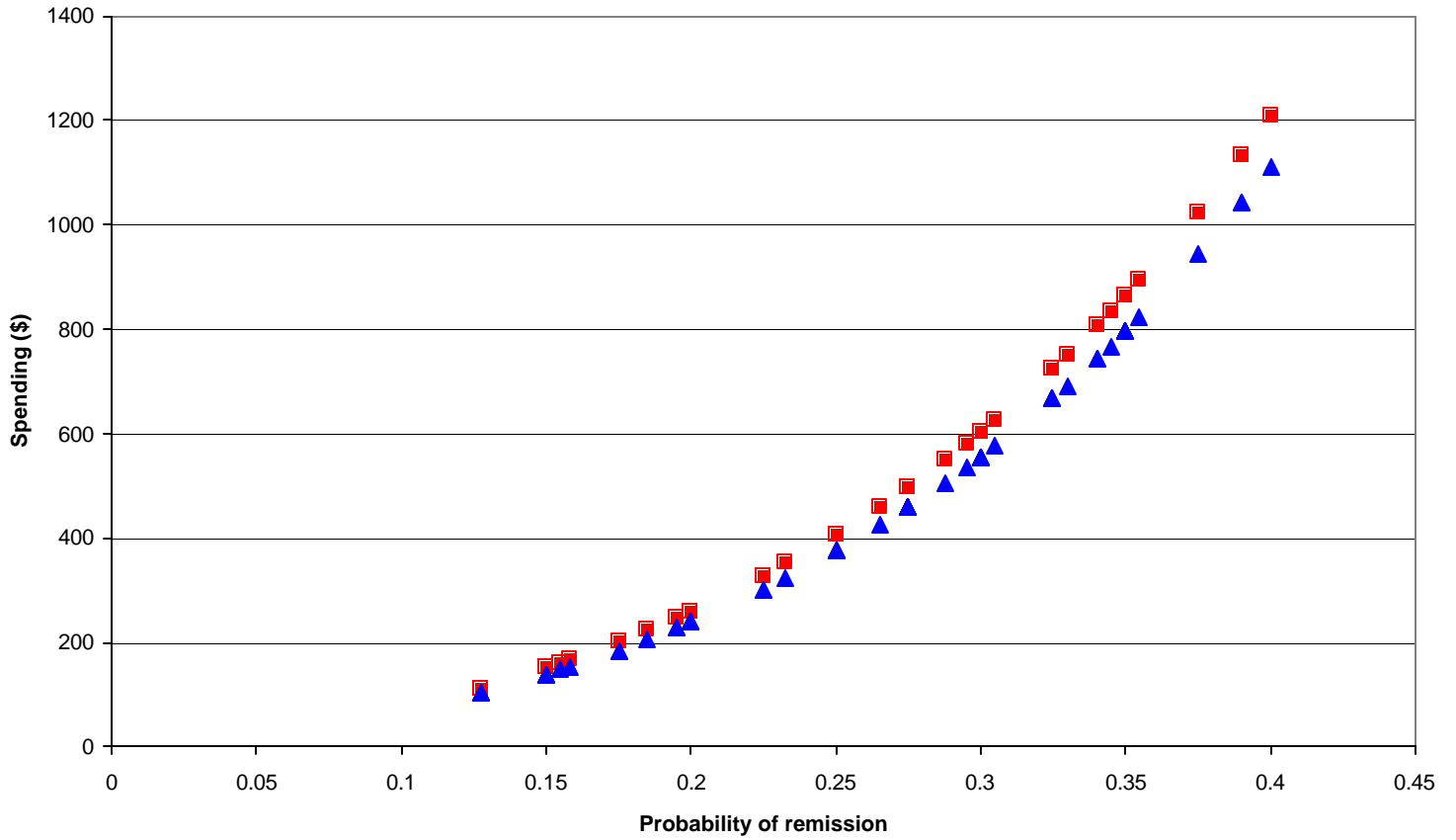
Table 6: Mental Health Related BLS Producer Price Indexes

Producer Category	1992	1993	1994	1995	1996
Office of Doctors	-	100	102.8	105.9	107.8
Psychiatrists Small Group/Solo	-	100	102.9	104.7	106.8
All Outpatient Services	-	100	102.4	106.9	114.2
Psychiatric Hospital Outpatient Services	-	100	103.3	113.0	119.4
Antidepressant Prescription Pharmaceuticals	100	105.3	107.9	118.9	119.0

Source: U.S. Bureau of Labor Statistics, Producer Price Indexes

Figure 1: Projected Spending per Remission, adj. for Patient Category, 1991 and 1996

Based on Equation 3 of Table 4



References

Advisory Commission to Study the Consumer Price Index [1996], "Toward a More Accurate Measure of the Cost of Living: Final Report to the Senate Finance Committee", Washington DC: U.S. Government Printing Office, December.

Agency for Health Care Policy Research [1999], Treatment of Depression - Newer Pharmacotherapies, Bethesda MD: AHCPR Publication No. 99-E014.

American Medical Association [1999], Current Procedural Terminology 1999, Chicago.

American Psychiatric Association [1994], Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Washington DC.

American Psychiatric Association [1993], "Practice Guidelines for Major Depressive Disorder in Adults," American Journal of Psychiatry, 150:4, Supplement, April, 1-26.

Becker, Gary S. [1965], "A Theory of the Allocation of Time," Economic Journal, Vol. 75, September, pp. 493-517.

Berndt, Ernst R., Iain Cockburn and Zvi Griliches [1996], "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs," Brookings Papers on Economic Activity: Microeconomics 1996, 133-188.

Berndt, Ernst R., David Cutler, Richard G. Frank, Zvi Griliches, Joseph P. Newhouse and Jack E. Triplett [2000a], "Medical Care Prices and Output," forthcoming as chapter 29 in Joseph P. Newhouse and Anthony C. Culyer, eds., Handbook of Health Economics, Amsterdam: Elsevier Science.

Berndt, Ernst R., David Cutler, Richard G. Frank, Zvi Griliches, Joseph P. Newhouse and Jack E. Triplett [2000b], "Price Indexes for Medical Care Goods and Services: An Overview of Measurement Issues," paper given at NBER-CRIW Conference on Medical Output and Productivity, June 12-13, 1998. Forthcoming in David Cutler and Ernst R. Berndt, eds., Medical Output and Productivity, Chicago: University of Chicago Press for the National Bureau of Economic Research.

Berndt, Ernst R., Stan N. Finkelstein, Paul E. Greenberg, Robert H. Howland, Alison Keith, A. John Rush, James Russell and Martin B. Keller [1998], "Workplace Performance Effects from Chronic Depression and Its Treatment," Journal of Health Economics, Vol. 17, No. 5, Fall, pp. 511-535.

Berndt, Ernst R., Susan H. Busch and Richard G. Frank [2000], "Treatment Price Indexes for Acute Phase Major Depression," paper given at NBER-CRIW Conference on Medical Output and Productivity, June 12-13, 1998. Forthcoming in David Cutler and Ernst R. Berndt, eds., Medical Output and Productivity, Chicago: University of Chicago Press for the National Bureau of Economic Research.

Berndt, Ernst R., Richard G. Frank and Thomas G. McGuire [1997], "Alternative Insurance Arrangements and the Treatment of Depression: What Are The Facts?," American Journal of

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS
Managed Care, 3:2, February, 243-250.

Broadbent, W. Eugene., Dan G. Blazer, Linda K. George and Chiu Kit Tse [1990], "Depression, Disability Days and Days Lost from Work in a Prospective Epidemiologic Survey," Journal of the American Medical Association, 264:2524-2528.

Busch, Susan [2000], "The Effects of New Procedures for Constructing Official Price Indexes: The Case of Treatment for Depression," Cambridge, MA: Harvard University, Essay 2 in unpublished Ph.D. dissertation, Health Policy, April.

Cockburn, Iain M. and Aslam Anis [2000], "Hedonic Analysis of Arthritis Drugs," paper given at NBER-CRIW Conference on Medical Output and Productivity, June 12-13, 1998. Forthcoming in David Cutler and Ernst R. Berndt, eds., Medical Output and Productivity, Chicago: University of Chicago Press for the National Bureau of Economic Research.

Cutler, David M., Mark McClellan and Joseph P. Newhouse [1999], "The Costs and Benefits of Intensive Treatment for Cardiovascular Disease," in Jack E. Triplett, ed., Measuring the Prices of Medical Treatments, Washington DC: The Brookings Institution, pp. 34-71.

Cutler, David M., Mark McClellan, Joseph P. Newhouse and Dahlia Remler [1998], "Are Medical Prices Declining? Evidence from Heart Attack Treatments," Quarterly Journal of Economics, Vol. 113, No. 4, November, pp. 991-1024.

Depression Guideline Panel [1993], Depression in Primary Care, Vol. 2, Rockville, MD: U.S. Department of Health and Human Services, AHCPR publication 93-0551.

Diewert, W. Erwin [1981], "The Economic Theory of Index Numbers: A Survey," in Angus Deaton, ed., Essays in the Theory and Measurement of Consumer Behavior in Honour of Sir Richard Stone, Cambridge: Cambridge University Press, pp. 163-208.

Elkin, Irene, T. Shea, J.T. Watkins, S.D. Imber, S.M. Sotsky, J.F. Collins, D.R. Glass, P.A. Pilkonis, W.R. Leber, J.P. Docherty, S.J. Fiester and M.B. Parloff [1989], "National Institute of Mental Health Treatment of Depression Collaborative Research Program: General Effectiveness of Treatments," Archives of General Psychiatry, 46, 971-982.

Feldstein, Martin S. [1969], "Improving Medical Care Price Statistics," American Statistical Association 1969 Proceedings of the Business and Economic Statistics Section, Washington DC: American Statistical Association, pp. 361-364.

Fisher, Franklin M. and Zvi Griliches [1995], "Aggregate Price Indexes, New Goods, and Generics," Quarterly Journal of Economics, Vol. 110, No. 1, February, pp. 229-244.

Frank, Richard G., Ernst R. Berndt and Susan H. Busch [1999], "Price Indexes for the Treatment of Depression," in Jack E. Triplett, ed., Measuring the Prices of Medical Treatments, Washington, DC: Brookings Institution Press, pp. 72-102.

Frank, Richard G., Susan H. Busch, and Ernst R. Berndt [1998], "Measuring Prices and Quantities of Treatments for Depression," American Economic Review, Vol. 88, No. 2, pp. 106-111.

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

Frank, Richard G., Thomas G. McGuire and Joseph P. Newhouse [1995], "Risk Contracts in Managed Mental Health Care," Health Affairs, 14:3, 50-64.

Frank, Richard G., Thomas G. McGuire, Sharon-Lise Normand and Howard H. Goldman [1999], "The Value of Mental Health Care At The System Level: The Case of Treating Depression," Health Affairs, Vol. 18, No. 5, September/October, pp. 71-88.

Frazier H.S. and F.M. Mosteller [1995], Medicine Worth Paying For: Assessing Medical Innovations, Cambridge, MA: Harvard University Press.

Getzen, Thomas E. [1992], "Medical Care Price Indexes: Theory, Construction and Empirical Analysis of the US Series 1927-1990," in Advances in Health Economics and Health Services Research, JAI Press, 83-128.

Gilbert, Milton [1961], "The Problem of Quality Change and Index Numbers," Monthly Labor Review, 84:9, September, 992-997.

Gilbert, Milton [1962], "Quality Change and Index Numbers: A Reply," Monthly Labor Review, 85:5, May, 544-545.

Goldman, William, Joyce McCulloch and Roland Sturm [1998], "Costs and Use of Mental Health Services Before and After Managed Care," Health Affairs, 17:2, March/April, 40-52.

Griliches, Zvi [1962], "Quality Change and Index Numbers: A Critique," Monthly Labor Review, 85:5, 532-544.

Griliches, Zvi and Iain Cockburn [1994], "Generics and New Goods in Pharmaceutical Price Indexes," American Economic Review, Vol. 84, No. 5, pp. 1213-1232.

Hanson, Kristina W. [1998], "Public Opinion and the Mental Health Parity Debate: Lessons from the Survey Literature," Psychiatric Services, Vol. 49, No. 8, August, pp. 1059-1066.

Hausman, Jerry A. [1997], "Valuation of New Goods Under Perfect and Imperfect Competition," in Timothy F. Bresnahan and Robert J. Gordon, eds., The Economics of New Goods, Chicago: University of Chicago Press for the NBER, pp. 209-237.

Heidenreich, Paul and Mark McClellan [2000], "Trends in Heart Attack Treatments and Outcomes, 1975-1995: A Literature Review and Synthesis," forthcoming in David Cutler and Ernst R. Berndt, eds., Medical Output and Productivity, Chicago: University of Chicago Press for the National Bureau of Economic Research.

Kessler, Ronald C., Katherine A. McGonagle, Shanyang Zhao, Christopher B. Nelson, Michael Hughes, Suzann Eshlerman, Hans-Ulrich Wittchen and Kenneth S. Kendler [1994], "Lifetime and Twelve-Month Prevalence of DSM-III-R Psychiatric Disorders in the United States: Results from the National Comorbidity Survey," Archives of General Psychiatry, Vol. 51, No. 1, January, pp. 8-19.

Kessler, Ronald C., Katherine A. McGonagle, Marvin Swartz, Dan G. Blazer and Christopher B. Nelson [1993], "Sex and Depression in the National Comorbidity Survey I: Lifetime Prevalence, Chronicity and Recurrence," Journal of Affective Disorders, 29:85-96.

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

Leibenstein, Harvey [1976], Beyond Economic Man: A New Foundation for Microeconomics, Cambridge, MA: Harvard University Press.

Ma, Ching-to Albert and Thomas G. McGuire [1998], "Costs and Incentives in a Behavioral Health Carve-Out," Health Affairs, 17:2, March/April, 53-69.

McKusick, David, Tami L. Mark, Edward King, Rick Harwood, Jeffrey A. Buck, Joan Dilonardo, and James S. Genuardi [1998], "Spending for Mental Health and Substance Abuse Treatment, 1996," Health Affairs, Vol. 17, No. 5, September/October, pp. 147-157.

Medicode [1998], International Classification of Diseases, Ninth Revision, Clinical Modification, Fifth Edition, Salt Lake City: Medicode.

Meltzer, David [1997], "Accounting for Future Costs in Medical Cost-Effectiveness Analysis," Journal of Health Economics, Vol. 16, pp. 33-64.

Newhouse, Joseph P. [1989], "Measuring Medical Prices and Understanding Their Effects: The Baxter Prize Address," Journal of Health Administration Education, 7:1, 19-26.

Normand, Sharon-Lise T., Richard G. Frank and Thomas G. McGuire [1999], "Elicitation in Observational Studies: Estimating the Value of Acute Treatments for Depression," Boston MA: Harvard Medical School, draft manuscript, December 26.

Phelps, Charles E. and Cathleen Mooney [1993], "Variations in Medical Practice Use: Causes and Consequences," in Richard J. Arnould, Robert F. Rich and Wiliam D. White, eds., Competitive Approaches to Health Care Reform, Washington DC: Urban Institute Press.

Phelps, Charles E. [1992], "Diffusion of Information in Medical Care," Journal of Economic Perspectives, Vol. 6, No. 3, Summer, pp. 23-42.

Reder, Melvin W. [1969], "Some Problems in the Measurement of Productivity in the Medical Care Industry," in Victor R. Fuchs, ed., Production and Productivity in the Service Industries, NBER Studies in Income and Wealth, Vol. 34, New York: Columbia University Press, pp. 95-131.

Scitovsky, Anne A. [1985], "Changes in the Costs of Treatment of Selected Illnesses, 1971-1981," Medical Care, Vol. 23, No. 12, December, pp. 1345-1357.

Scitovsky, Anne A. [1967], "Changes in the Costs of Treatment of Selected Illnesses, 1951-65," American Economic Review, 57:5, December, 1182-1195.

Scitovsky, Anne A. [1964], "An Index of the Cost of Medical Care--A Proposed New Approach," in The Economics of Health and Medical Care, Proceedings of the Conference on the Economics of Health and Medical Care, May 10-12, 1962, The University of Michigan, Ann Arbor, 1964.

Scitovsky, Anne A. and Nelda McCall [1976], "Changes in the Costs of Treatment of Selected Illnesses: 1951-1964-1971," Washington DC: U.S. Department of Health,

EXPECTED OUTCOMES AND PRICE INDEXES FOR DEPRESSION TREATMENTS

Education, and Welfare, National Center for Health Services Research, DHEW Publication No. (HRA) 77-3161.

Shapiro, Irving, Matthew D. Shapiro and David W. Wilcox [2000], "A Price Index for Cataract Surgery," paper given at the NBER-CRIW Conference on Medical Care Output and Productivity, Bethesda MD, June 12-13. Forthcoming in David Cutler and Ernst R. Berndt, eds., Medical Output and Productivity, Chicago: University of Chicago Press for the National Bureau of Economic Research.

Shapiro, Matthew and David Wilcox [1996], "Mismeasurement in the Consumer Price Index: An Evaluation," NBER Macroeconomics Annual, Vol. 11, pp. 93-142.

Triplett, Jack E. [2000], "What's Different About Health? Human Repair and Car Repair in the National Accounts," paper presented at the NBER-CRIW Conference on Medical Output and Productivity, June 12-13, 1998. Forthcoming in David Cutler and Ernst R. Berndt, eds., Medical Output and Productivity, Chicago: University of Chicago Press for the National Bureau of Economic Research.

Triplett, Jack E. and Ernst R. Berndt [1999], "Introduction: New Developments in Measuring Medical Care," in Jack E. Triplett, ed., Measuring the Prices of Medical Treatments, Washington DC: The Brookings Institution, pp. 1-33.

Triplett, Jack E. [1999], "Accounting for Health Care: Integrating Price Index and Cost-Effectiveness Research," in Jack E. Triplett, ed., Measuring the Prices of Medical Treatments, Washington DC: The Brookings Institution, pp. 220-250.

United States Department of Health and Human Services [1999], Mental Health: Report of the Surgeon General, Washington, D.C.: U.S. Government Printing Office.

Footnotes

¹ McKusick et. al. [1998].

² Hanson [1998].

³ See, for example, Berndt, Cutler, Frank, Griliches, Newhouse and Triplett [2000a,b], Triplett [2000] and Triplett-Berndt [1999]. Earlier discussions include Feldstein [1969], Newhouse [1989] and Getzen [1992].

⁴ See Scitovsky [1964,1967,1985] and Scitovsky-McCall [1976].

⁵ This was emphasized by the Boskin Commission report (Advisory Commission to Study the Consumer Price Index [1996]), which recommended that changes in medical outcomes be taken into account in official price index calculations.

⁶ Cutler, McClellan, Newhouse and Remler [1998], and Cutler, McClellan and Newhouse [1999]. A related study is that by Heidenreich and McClellan [2000].

⁷ Shapiro and Wilcox [1996], and Shapiro, Shapiro and Wilcox [2000].

⁸ See also the related papers by Frank, Berndt and Busch [1999], and Frank, Busch and Berndt [1998].

⁹ See AHCPR [1999], APA [1993] and Depression Guideline Panel [1993].

¹⁰ With the latter, there appeared to be an increased share of patients with more complicated conditions, greater severity of illness, and elements of longer term treatment.

¹¹ There is a large literature on the cross-sectional variation in medical treatments for other conditions and illnesses. See, for example, Phelps [1992] and Phelps-Mooney [1993]. On the impacts of cost containment and risk-sharing on choice of treatment, see Berndt, Frank and McGuire [1997], Goldman, McCulloch and Sturm [1998], and Ma and McGuire [1998].

¹² For related discussions, see Frank, McGuire, Normand and Goldman [1999], and Triplett [1999]. One way of viewing this research is as measuring the impacts of reductions over time in what Leibenstein [1976] called X-inefficiency.

¹³ See APA [1994].

¹⁴ Kessler et al. [1993,1994].

¹⁵ APA [1993].

¹⁶ For a discussion, see Elkin et al. [1989].

¹⁷ For a discussion of antidepressant sales histories, as well as side effect and other characteristic profiles of these antidepressants, see Berndt, Cockburn and Griliches [1996].

¹⁸ Medicode [1998].

¹⁹ See American Medical Association [1999].

²⁰ This integration of clinical outcomes with price index research has been advocated by, among others, Triplett [1999].

²¹ The panel consisted of four psychiatrists, four clinical psychologists, and two general internists. Panelist averaged 19.9 years in professional practice (range, 11-42). The mean number of depressive patients treated by panelists in the previous year was 63.7 (range, 4-200). Panel members were David Adler, MD; Christopher Callahan, MD; Ellen Frank, Ph.D., co-chair; Wayne Katon, MD; Janice Krupnick, Ph.D.; Jeanne Miranda, Ph.D.; A. John Rush, MD; H.C. Schulberg, Ph.D., co-chair; Michael Thase, MD; and John Williams, Jr., MD.

²² Frazier and Mosteller [1995].

²³ See Landrum and Normand [1999] and the references therein.

²⁴ This table is reproduced from Normand, Frank and McGuire [1999], and is based on a slightly smaller sample of episodes.

²⁵ Triplett [1999].

²⁶ Use of the Tornqvist approximation to the translog price index resulted in rather volatile shares and quantities. Moreover, with the translog index the use of logarithms and the adding up of shares implies that

the price index per incremental remission is the same as the price per remission. For these reasons, we employ the Fisher rather than the translog.

²⁷For a discussion of index number and hedonic methods to control for quality in the medical care context and elsewhere, see Cockburn-Anis [2000], Gilbert [1961,1962], Griliches [1962], and Reder [1969]. Price indexes with heterogeneous purchasers are considered by Diewert [1981], Fisher-Griliches [1995] and Griliches-Cockburn [1994].

²⁸For further discussion, see Busch [2000].

²⁹See U. S. Department of Health and Human Services [1999].

³⁰See, for example, Ma and McGuire [1998], and Frank, McGuire and Newhouse [1995].

³¹Becker [1965] and Triplett [1999].

³²On this, see for example Berndt, Finkelstein, Greenberg et al. [1998] and Broadbent, Blazer, George and Tse [1990].

³³Hausman [1997], and Meltzer [1997].