

Berndt, Ernst R.; Busch, Susan H.; Frank, Richard G.

**Article**

## Interpreting changes in mental health expenditures: Minding our Ps and Qs

NBER Reporter Online

**Provided in Cooperation with:**

National Bureau of Economic Research (NBER), Cambridge, Mass.

*Suggested Citation:* Berndt, Ernst R.; Busch, Susan H.; Frank, Richard G. (2000) : Interpreting changes in mental health expenditures: Minding our Ps and Qs, NBER Reporter Online, National Bureau of Economic Research (NBER), Cambridge, MA, Iss. Winter 2000/01, pp. 10-13

This Version is available at:

<https://hdl.handle.net/10419/67011>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

---

## Research Summaries

---

# Interpreting Changes in Mental Health Expenditures: Minding Our Ps and Qs

Ernst R. Berndt, Susan H. Busch, and Richard G. Frank\*

Interpreting growth trends in mental health spending can be puzzling. When mental health spending grows more rapidly than other health expenditures, is it because of provider price inflation? When mental health spending grows less rapidly than other health expenditures, is it because the mental health needs of our population increasingly are not being met?

A first step toward interpreting changes in mental health spending properly involves decomposing these expenditures into Ps and Qs: prices and quantities. But measuring the quantity of real output in the health care sector, particularly in mental health services, is complicated for a number of reasons. For example, questions about the effectiveness of treatments and the welfare losses from moral hazard in insurance have long created concern that the value of spending on mental health may be low relative to spending on other health services. Mental disorders frequently are chronic and recurring conditions, and mortality is

not typically an appropriate measure of treatment outcome. Defining outcomes from various mental health treatments often relies on more subjective and difficult-to-measure constructs. Therefore, creating price indexes that account for the changing quality and effectiveness of mental health treatments poses significant measurement issues.

### Initial Methodology and Data

For some time, health economists have suggested that an appropriate treatment price index would be one based on defined episodes of treatment of selected illnesses and conditions. It would incorporate technological and institutional innovations that change the mix of inputs to treat the condition and would include any effects on changed medical outcomes. Anne A. Scitovsky was the first to implement this type of approach in 1967. She examined changes in the costs of treating episodes for six specific medical conditions at the Palo Alto Medical Research Foundation between 1951 and 1965.<sup>1</sup> In the health-related producer price indexes (PPIs) constructed and published by the Bureau of Labor Statistics (BLS), by contrast, variations in treatment outcomes are not taken into account, nor are major treatment substitutions, for example between pharmacotherapy and psychotherapy.

Over the last four years, we have undertaken a research program at the NBER that builds on the treatment episode tradition begun by Scitovsky and extends it to the most prevalent and costly of the mental

disorders, major depression. We report here on how this research has progressed and how findings have evolved as we developed more refined measures of treatment episodes and outcomes.

In the past two decades, new treatment technologies have been introduced, indicating the potential for changes in outcomes in treatment for depression. Treatment input patterns have shifted within treatment classes (for example, from older to more recently developed pharmacotherapies, particularly the selective serotonin reuptake inhibitors, or SSRIs) and between treatment classes (for example, to less intensive psychotherapy and more intensive pharmacotherapy). Fundamental organizational changes, such as the growth of managed care and specialty mental health and pharmacy "carve-outs," also may have affected prices of and treatment choices for depression.

In our work, we use quantities and prices of outpatient treatment for depression that are based on retrospective medical claims data from MedStat's publicly available MarketScan™ database. These data consist of 1991–6 enrollment records and medical claims from four large self-insured employers offering more than 25 health plans to their 400,000 plus employees and their dependents. These data include inpatient, outpatient, and pharmaceutical claims. The health benefits offered to enrollees in this database are quite generous relative to the general market for private health insurance in the United States.

To implement a price index for

---

\* Berndt is Director of the NBER's Program on Productivity and Technological Change and the Louis B. Seley Professor of Applied Economics at MIT's Sloan School of Management. Busch is an assistant professor of health care policy at Yale University Medical School. Frank is a Research Associate in NBER's Programs on Health Care and Health Economics and the Margaret T. Morris Professor of Health Economics at Harvard University Medical School.

treatment episodes, we combine individual claims using patient identifiers, diagnostic information, and dates of services rendered. For depression, a chronic disease, defining an acute episode requires extensive knowledge of the disorder, its course, and the administration of treatments in practice. At numerous times, therefore, we benefited from consultations with clinicians about these issues.

To match our medical claims with clinical data, we identify all ambulatory claims associated with either single or recurrent episodes of major depression, as defined by the International Classification of Diseases. When the claims data indicate that psychotherapeutic drugs were prescribed, we consider the number of days of treatment provided by the prescription as the time period over which an individual received care. We define an episode of depression as new if the diagnosis is preceded by a period of at least eight weeks without treatment. We eliminate episodes if the entire episode is not observed or if we do not observe eight weeks both before and after an acute phase episode. As a control for severity and because of a lack of information on the details of treatment, we exclude patients with psychiatric hospitalizations.

Using information on procedures (for example, a 20- or 50-minute psychotherapy visit, or whether a drug was prescribed) available in the medical claims data, we describe the composition of treatment that occurred within a treatment episode. Prescription drug treatment is based on the national drug codes (NDCs) reported on the claim. The NDC classification reveals the use of seven older-generation tricyclic antidepressants, three SSRIs, two other serotonin-related drugs, and various other drugs used to treat depression, including monoamine oxidase (MAO) inhibitors, anxiolytics, and

heterocyclics. We calculate direct medical spending for each treatment episode using actual transaction data. All insurer payments made to the provider and any cost-sharing assigned to the patient (for example, patient out-of-pocket copayment for prescription drugs) are summed to a nominal dollar total for each treatment episode. Thus, the treatment price indexes we construct are analogous to the PPI (supply side) rather than to the consumer price index or CPI (demand side). This process yields 10,368 identified episodes of depression between 1991 and 1995 in the claims data.

In our initial research, we used results from published treatment guidelines and our review of the clinical trial literature<sup>2</sup> to develop a set of "treatment bundles" grouping therapies into what we interpreted as similar groups for treatment of acute phase major depression. Our five treatment bundles vary in mix and length of psychotherapeutic drug treatment and/or number of psychotherapy visits, but they have similar *ex ante* expected outcomes. All bundles are confined to at most six months of treatment (the "acute" phase). The assumption in this methodology is that obtaining therapeutically similar outcomes from alternative bundles provides a useful approximation to achieving similar expected utility levels. However, an additional implicit assumption is that the production function for treatment of depression has a step-function form. For example, an individual receiving six psychotherapy sessions (barely meeting treatment guidelines) is treated as receiving "effective" treatment, although an individual receiving four or five visits (slightly less than treatment guidelines) is viewed as receiving "ineffective" treatment. The proportion of identified episodes receiving "effective" care was only 50.1 percent in our data.

When we limit our treatment

episodes that meet guideline criteria in this way and aggregate over treatment bundles using 1991 fixed quantity weights (analogous to the BLS's use of a Laspeyres price index), we obtain a treatment price index of 100 in 1991 and 68.4 in 1995, observing a negative average annual growth rate (AAGR) of 9.1 percent. Similar time patterns result when we use alternative index number aggregation formulas. Over this same time period, the official PPI (not based on episodes of treatment) for antidepressant drugs grew at an AAGR of 3.8 percent, while the PPI for physicians' services increased at one percent per year.

Next, we extended this research by reconstructing episodes to identify missing psychotherapy procedure codes, by adding two additional treatment bundles, and by incorporating episodes that involved longer treatment (but only including the first six months of treatment for such individuals).<sup>3</sup> With this expanded set of episodes and bundles, the Laspeyres-type treatment price index was essentially flat between 1991 and 1995, falling from 100 to 97.6, or -0.6 percent per year. This still represents considerably less growth than the official PPIs shows.

## The Next Phase

One major problem with all of our initial research, in addition to the restrictive step-function production assumption, is that by confining our analyses to those treatment episodes that meet guideline criteria, we ignore about 50 percent of delivered care. The share of episodes treated with guideline care in this claims database only increased from 35 percent to 55 percent between 1991 and 1995. Therefore, we wanted to relax the step-function production assumption and to make use of a great deal of clinical and medical information that is now known, as well as to incorporate treatments that reflect the

real-world environment but do not meet guideline standards. So, in the next phase of our research, we incorporated two major changes. First, we classified a broader set of episodes, including those that did not meet guideline criteria, according to two dimensions: type of patient and type of treatment. In that way, we identified about 200 patient treatment cells. When we eliminated treatment cells having fewer than 30 patients between 1991 and 1996, we were left with 120 patient treatment cells.

Next, we convened an expert panel of ten clinicians and researchers and elicited from them the outcomes they would expect for each of the 120 patient treatment cells. More specifically, we asked the expert panel members: of 100 patients meeting specific criteria for depression at initial visits, what number would fully respond to treatment after 16 weeks of treatment, what number would evidence a significant but partial response, and what number would not evidence any medically significant response? We also asked the panel to assess what number would remit or respond without any treatment (we called this the "waiting list"). Using a modified Delphi procedure, our expert panel assessment process converged in two steps. This process allowed us to infer outcome information for a wider range of treatment types and quantities than was available in our initial research, and it allowed us to integrate knowledge concerning the efficacy and effectiveness of real-world treatments with the MedStat retrospective claims data.

The results from this second phase of our research are reported in two recent papers.<sup>4</sup> Without making any adjustment for variations in expected outcomes, the Laspeyres-type treatment episode price index fell from 100 in 1991 to 95 in 1996, an AAGR of minus one percent. Since some individuals improve without receiving

any treatment, outcomes are best incorporated as expected mental health improvements over and above no treatment (that is, as price per incremental full remission or price per incremental partial remission). From 1991–6, the Laspeyres-type price index per incremental partial remission increased from 100 to 103.9 (an AAGR of 0.8 percent), while the index per incremental full remission increased slightly less, from 100 to 103.4 (AAGR of 0.7 percent). Indexes based on other weighting formulas revealed similar trends. Hence, from 1991–6, the total treatment cost of attaining an expected incremental partial or full remission from depression (including the costs of those treatments that were not likely to have been effective) increased by less than 1 percent per year.

Over this same period of time, however, increased levels of management were exercised over mental health benefits. This implies that the patient population may have been changing along with the mix of treatment bundles, thereby affecting both expected outcome and cost. Because our expected outcomes are assigned based on both treatment and patient type, changes in the mix of patients will affect the price per incremental remission. For example, the expert panel rated patients with comorbid substance abuse to have lower expected outcomes than patients without comorbid substance abuse receiving the same treatment. These changes in the patient mix over time in our MedStat claims database are not incorporated in the price index calculations described above.

To account for the effect of changing patient mix on computed price indexes, we delineate eight patient categories (whether medical comorbidity is present, whether male, if female whether over age 50, and whether there is comorbid substance abuse.) Then we estimate hedonic price equations for the price per

expected full remission. The dependent variable is the natural log of spending for each of the 8,187 treatment episodes; the regressors are the probability of a full remission associated with the patient's treatment and type, dummy variables for seven of the eight patient categories, and annual dummy variables. As expected, variations in patient categories have significant and substantial effects on treatment costs, and the coefficient on remission probability is positive and highly significant. The resulting price index falls from 100 in 1991 to 87.2 in 1996, implying an AAGR of -2.7 percent. The differences between this hedonic and the previous set of price indexes reflect the changing and increasingly complex mix of patients, along with changes in treatment bundles, over the six-year period.

## Conclusion

In summary, our analysis suggests that between 1991 and 1996, based on our preferred price index and adjusting both for expected outcomes from changing treatments and from varying patient mixes, the cost of treatment for depression has declined 2.7 percent per year. This contrasts with a price increase of 2.6 percent per year when we use BLS-like methods with these same data. The source of increased expenditures on treatment for depression since 1991 is the increased quantity of treatments and remissions, not increases in their prices. Since roughly half of spending on mental health involves treatment for depression, our results imply that much of the recent increase in spending on mental health care has been driven by increased productivity and expanded quantities of care, a result that is contrary to much conventional wisdom. Therefore, decomposing expenditures into their P and Q components is critical in interpreting expenditure variations.

This research suggests that while

constructing episode-based, outcomes-adjusted price indexes is a complex and cumbersome task, it is critically important for informed policy discussions. Although it may not be sensible or practical for the BLS to produce such an index on a monthly basis, it is important that policy analysts use episode-based, outcomes-adjusted price indexes when evaluating sources of expenditure variation in the National Health Accounts.

<sup>1</sup> A. A. Scitovsky, "Changes in the Costs of Treatment of Selected Illnesses, 1951-65," *American Economic Review*, 57(5) (December 1967), pp. 1345-57.

<sup>2</sup> R. G. Frank, S. H. Busch, and E. R.

Berndt, "Measuring Prices and Quantities of Treatment for Depression," *American Economic Review*, 88(2) (May 1998), pp. 106-11; R. G. Frank, E. R. Berndt, and S. H. Busch, "Price Indexes for the Treatment of Depression," NBER Working Paper No. 6417, February 1998, and in *Measuring the Prices of Medical Treatments*, J. E. Triplett, ed. Washington, D.C.: The Brookings Institution, pp. 72-102.

<sup>3</sup> E. R. Berndt, S. H. Busch, and R. G. Frank, "Treatment Price Indexes for Acute Phase Major Depression," NBER Working Paper No. 6799, November 1998, forthcoming in *Medical Care Output and Productivity*, D. M. Cutler and E. R. Berndt, eds. Chicago: University of Chicago

Press; S. H. Busch, E. R. Berndt, and R. G. Frank, "Creating Price Indexes for Measuring Productivity in Mental Health Care," forthcoming in *Frontiers in Health Policy Research*, vol. 3, A. M. Garber, ed. Cambridge, MA: MIT Press.

<sup>4</sup> S. H. Busch, E. R. Berndt, and R. G. Frank, "Creating Price Indexes for Measuring Productivity in Mental Health Care," forthcoming in *Frontiers in Health Policy Research*, vol. 3, A. M. Garber, ed. Cambridge, MA: MIT Press; E. R. Berndt, A. Bir, S. H. Busch, R. G. Frank, and S. T. Normand, "The Medical Treatment of Depression, 1991-6: Productive Inefficiency, Expected Outcome Variations, and Price Indexes," NBER Working Paper No. 7816, July 2000.

## International Trade and National Factor Markets

Donald R. Davis\*

### Trade, Wages, and Unemployment

Of primary concern in international trade policy is the impact of trade on national factor markets. Such concerns have sharpened as the extent of international integration has risen both among OECD countries and between these countries and poorer non-OECD nations. How does trade affect unemployment, we wonder, or wages? Does it matter that countries have very different labor market institutions?

For the analyst, these questions raise both theoretical and empirical issues. What theoretical framework is appropriate for thinking about these issues? Is there a model of international exchange and national factor markets with enough empirical support to give us confidence in the comparative static assessments that we make? These questions have motivated my work in recent years.

One strand of this work explicitly focuses on the analytics of trade, wages, and unemployment when countries have different labor market institutions. This work is motivated by the contrast between the experience of the Anglo-American economies and that of many Continental European economies. In the Anglo-American economies after 1980 there was a large decline in the

relative wages of unskilled workers, yet relatively low unemployment. In contrast in many of the European economies at that time, relative wages of the unskilled did not decline, but unemployment reached very high levels. The contrast in experience has been noted by many other researchers, and differences in the flexibility of factor market institutions often were cited as an important contributing factor.

One key question that had not been addressed, though, is how the costs borne by each of these country-types were affected by the fact that these countries trade in a unified global goods market. I address that question in a series of papers. The first paper considers a benchmark case between a stylized America and Europe, where the countries are

\*Davis is a Research Associate in the NBER's Program on International Trade and Investment and a Professor of Economics at Columbia University. His "Profile" appears later in this issue.