PRICE INDEXES FOR ACUTE PHASE
TREATMENT OF DEPRESSION

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

Price Indexes for Acute Phase Treatment of Depression
Ernst R. Berndt, Susan H. Busch, and Richard G. Frank
NBER Working Paper No. 6799
November 1998
JEL No. I1, C8

ABSTRACT

Although broad trends in medical spending in the U.S. over the last decade have received widespread attention from policymakers, very little attention has focused on the components of those changes. For many other industries, economists typically divide nominal expenditures by an official government price index to decompose these expenditures into price and quantity components.

In this paper we construct a new price index for the treatment of one illness, major depression. Making use of results from the published clinical literature and from official treatment guideline standards, we identify therapeutically similar treatment bundles. These bundles can then be linked and weighted to construct price indexes for specific forms of major depression. In doing so, we construct CPI and PPI-like medical price indexes that deal with prices of treatment episodes rather than prices of discrete inputs, that are based on transaction rather than list prices, that take quality changes and expected outcomes into account, and that employ current, time-varying expenditure weights in the aggregation computations.

We find that regardless of which index number procedure is employed, over the 1991-95 time period the treatment price index for the acute phase of major depression has hardly changed, remaining at 1.00 or falling slightly to around 0.97. This index grows considerably less rapidly than the various official PPIs -- thus the "real", inflation-adjusted price index for the treatment of the acute phase of major depression has fallen over the 1991-95 time period. A hedonic approach to price index measurement yields broadly similar results.

These results imply that given a budget for treatment of depression, more could be accomplished in 1995 than in 1991. Our results suggest that at least in the case of acute phase major depression, aggregate spending increases are due to a larger number of effective treatments being provided.

Ernst R. Berndt
Massachusetts Institute of Technology
Alfred P. Sloan School of Management
50 Memorial Drive, E52-452
Cambridge, MA 02142
and NBER
eberndt@mit.edu

Susan Busch
Harvard Medical School
Department of Health Care Policy
180 Longwood Avenue
Boston, MA 02115

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

PRICE INDEXES FOR ACUTE PHASE TREATMENT OF DEPRESSION

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

TABLE OF CONTENTS

Corresponding Author:
Ernst R. Berndt
MIT Sloan School of Management
50 Memorial Drive, E52-452
Cambridge, MA 02142
(617) 253-2665 (tel)
EBERNDT@MIT.EDU

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

> "It is not well to sneer at political economy in its relation to the insane poor. Whether we think it right or not, the question of cost has determined and will continue to determine their fate for weal or woe."
> (Asylum Superintendent George Cook, 1866)

I. INTRODUCTION

Much has been written in the last decade on broad trends in medical care spending in the US. Although the most recent evidence is somewhat ambiguous, the apparent slowdown in the rate of increase in aggregate health expenditures over the last five years has been welcomed by many governments, employers, patients and insurers. Relatively little attention, however, has focused on components of the change in expenditures. Is the trend change in medical care expenditures due to changes in price, quantity, quality or a combination of all three? What has been the role of organizational change, such as the growth of managed care, relative to technological/pharmaceutical innovation, such as new medications, on expenditure trends?

A standard procedure for economists analyzing changing expenditures over time is to employ official government price indexes, divide nominal expenditures by such a price index, and thereby decompose expenditures into price and quantity components. Further analysis might then focus on factors affecting real quantity growth, such as productivity gains.

In fact, the BLS' medical Consumer Price Index (CPI) is often used by analysts interested in undertaking such expenditure decompositions. It is well known that such a practice is frequently inappropriate and misleading, for the medical CPI deals only with consumers' direct out of pocket payments, and does not include payments from employers to insurers.

A deeper issue, however, involves the conceptual foundations underlying the medical CPI and Producer Price Indexes (PPIs). Although some revisions are gradually being implemented, currently the BLS price indexes are based on the repricing over time of a fixed

bundle of inputs. For decades, health economists have argued that a more appropriate price index is one based on the entire episodic treatment of selected illnesses and conditions, incorporating technological and institutional innovations that change the mix of inputs used to treat the condition, and including any effects on changed medical outcomes.

Developing a price index in health care can be viewed in terms of the unique characteristics of the health services in question. Consider creating a price index for a good that displays the following features:

♦ Consumers pay only a portion of the gross price due to insurance. The insurance is associated with moral hazard resulting in "too much" use of the good.

♦ There has been dramatic technical change and great advances in the benefits offered by the good. Yet, the full range of benefits are difficult to measure.

♦ There is great variety in the forms that the good takes and consumers value those forms differently.

♦ The supply side of the market for the good has experienced fundamental changes in its organization.

♦ The good is viewed as so important in some cases that the government compels people to use it.

The treatment of major depression, one of the most prevalent and disabling mental disorders, exhibits all of the features listed above. In this paper we report on the first three years of a research program aimed at measuring prices and output for the treatment of this important disease. The approach taken in this program of research builds on several recent efforts to construct price indexes for medical care. That recent research includes work by Cutler, McClellan, Newhouse and Remler [1996] that contrasts input price indexes for the treatment of heart attacks that rose by 6.7% per year over 1983-94, with an outcomes adjusted index that takes into account a conservative valuation for the extension of life expectancy attributable to new heart attack treatments and increases by only 2.3% per year, implying a net bias of 4.4% annually. Focusing on a price index for cataract surgery, Shapiro and Wilcox [1996] find that a CPI-like

fixed weight input-based index increases by a factor of about nine between 1969 and 1993, whereas their preferred alternative index incorporating realized reductions in hospital lengths of stay, but ignoring any improvements in the quality of medical outcomes, increases by only a factor of three, implying an annual differential of 4.6%.

Because depression is a chronic recurring illness and mortality is not an end point, we must rely on indirect methods of incorporating information on clinical effectiveness and outcomes. In our research we have made use of results from the published clinical literature, and from official treatment guideline standards, to identify therapeutically similar treatment bundles that can then be linked and weighted to construct price indexes for treatment of specific forms of major depression. Preliminary results from this research have been published in Frank, Berndt and Busch [1998] and Frank, Busch and Berndt [1998].

In this paper we consolidate our findings, and extend our analysis considerably. Specifically, we refine our definition of eligible treatment episodes, develop new imputations for missing data, expand the set of treatment bundles from five to seven, and estimate hedonic price regressions.

We begin the paper with an overview of current BLS procedures for constructing medical care price indexes, and then provide a background on the nature of and alternative treatments for acute phase major depression. We outline features of the retrospective medical claims data base from MEDSTAT, and then discuss our implementation of treatment episode definition and identification. We then report on quantities and prices of the treatment bundles from 1991 to 1995, we construct and interpret CPI-like and PPI-like aggregate price indexes, and we comment on an initial analysis involving hedonic price procedures. We end the paper with a discussion section, followed by concluding remarks.

II. CURRENT PROCEDURES FOR MEASURING MEDICAL CARE PRICE INDEXES

In the US context, it is useful to distinguish "supply" and "demand" prices for medical care. By supply prices we mean the total payments received by health care providers for a

particular medical treatment, consisting of the sum of payments from health insurance plans (both private and public) plus that from patients' direct out-of-pocket payments (OOPPs). At this level of generality, our supply price concept relates to components of the Producer Price Index (PPI) published by the Bureau of Labor Statistics (BLS).

The demand price notion we employ here is the consumers' direct OOPPs, consisting of co-payments and deductible cash payments made by the patient/consumer to providers for a particular medical treatment. Thus, in our price index changes in supply price are indicative of changes in the cost of treatment, while changes in the demand price can be traced both to changes in the cost of treatment and changes in benefit design. Our demand price concept relates to the medical care component of the BLS' Consumer Price Index (CPI). As we discuss below, the presence of insurance clouds the interpretation of the demand price index for depression as a CPI. We also discuss important differences concerning the treatment of indirect consumer payments via, for example employees' contributions to employment-related health insurance.

To associate the notion of a price index for treatment episodes of care with official price indexes published by the BLS, we now describe index number procedures currently employed by the BLS, beginning with its PPI.[1]

II.A. BLS Procedures for Medical Care Related PPIs

The PPI "...measures average changes in selling prices received by domestic producers for their output", separately by industry.[2] The PPI takes as its definition of an industry that based on the Standard Industrial Classification (SIC) code, using 4-digit SIC industries and their aggregates.[3] Since its inception in 1902, the PPI has focused rather heavily on the goods-producing sectors of the US economy. In 1986, in recognition of the growing importance of services, the BLS gradually began to broaden the PPIs scope of coverage in the services sectors, including medical care.

Within each industry, the BLS calculates aggregate PPIs using a modified fixed weight Laspeyres price index formula over individual price quotes, where fixed weights are based on value of shipments data. The primary price quote at the "cell index" level of disaggregation is

"...the net revenue accruing to a specified producing establishment from a specified kind of buyer for a specified product shipped under specified transactions terms on a specified day of the month".[4] Although the BLS seeks transactions rather than list prices for its price quotes, the agency is well aware that compliance by firms is easier with list than with actual transactions prices. Participation by firms in the PPI is voluntary, with a "productive" compliance rate being about 63% in 1992.[5]

The BLS currently draws a sample of items for each industry on average about every seven years, and then reprices this fixed set of items monthly until an entirely new sample is drawn. In 1995 the BLS announced that for certain industries, particularly technologically dynamic ones such as pharmaceuticals and electronics in which seven year time lags could yield a sample of products and services much older and quite unrepresentative of market transactions, samples would be supplemented at one or two-year intervals.

For quite some time, the BLS has published PPIs for certain medical-related manufacturing industries, such as pharmaceuticals and diagnostic equipment. It is only recently, however, that the BLS has begun publishing PPIs for medical service industries such as hospital services and physician services. The BLS initiated a PPI for an aggregate of health services in December 1994, for offices and clinics of doctors of medicine in December 1993, and for hospitals in aggregate and by type in December 1992.

A central measurement issue in the construction of medical price indexes involves the specification and implementation of a concept of industry output. For obvious reasons, the SIC structure is not well-equipped to provide guidance to the BLS on what is the appropriate real output concept in medical care industries, nor on how this output quantity and output price can best be measured.

An alternative source for guidance on implementation of a medical care output concept is provided by the Health Care Financing Administration (HCFA). In 1983 HCFA implemented a prospective payment schedule for inpatient hospital care whereby hospitals received a fixed payment for each Medicare patient admission, regardless of the amount or

duration of services actually provided the patient. These Medicare prospective payment schedules distinguish treatments by 24 major diagnostic categories, which are broken down further into 495 groupings of medical diagnoses and surgical procedures, known as diagnostic related groups (DRGs).[6]

DRGs provide one possible output concept, but currently DRGs only measure output for inpatient hospital care; many outpatient commodities (e.g., prescription pharmaceuticals) and services are not included in the DRG system. Classification schemes used by health care payers include version four of Current Procedural Terminology (CPT4) codes, a list containing thousands of procedures for which physicians and hospitals can bill; these CPT4 codes can be envisaged as inputs into the treatment of an illness or condition.[7] A systematic structure of diagnostic codes for illnesses and conditions is version nine of the International Classification of Diseases (ICD-9).[8] Relationships among ICD-9, CPT4 and DRG codes are multifaceted; a single DRG encompasses treatment of somewhat arbitrary aggregations of distinct ICD-9 diagnoses, alternative combinations of CPT4 codes can be used in the treatment of a particular ICD-9 diagnosis, and a given CPT4 procedure can be used in the treatment of various ICD-9 diagnoses. The DRG system makes use of both ICD-9 and CPT4 coding systems. Other diagnostic-related systems used in setting risk-adjusted capitation rates include the Ambulatory Care Group algorithm[9] and the Hierarchical Coexisting Conditions model.[10]

The BLS' PPI program has initiated procedures for construction of medical service PPIs at two rather aggregate levels, physician services and hospital services. In turn these encompass a variety of more detailed industries, such as physician services from psychiatrists, general/family practitioners, internists, and hospital services from general medical and surgical hospitals, psychiatric hospitals, and specialty non-psychiatric hospitals.

With respect to physician services, based on a sampling universe of all physician practices in the US, the BLS constructs a sample frame of physician practice units. From this sample unit, the BLS randomly chooses a bill that measures the net price paid to the practice for the entire set of services and procedures provided during an office visit, distinguished by payor

type. The physician's output from this visit is represented by the content of the patient's bill, including CPT codes associated with that visit, and is related to a particular ICD-9 diagnosis. Given this sample bill, the BLS contacts the physician practice unit each month and asks it to reprice what the current net transactions prices would be for that particular fixed bundle/payor of services. Since the organization of physician practices has undergone considerable upheaval and consolidation in the last few years, however, sample attrition in the BLS' physician services PPI has been considerable, and maintaining a sufficiently large response rate for repricing has proven to be difficult.[11]

The hospital services PPI measures net prices paid to hospitals for the entire bundle of services received during a hospital stay associated with a particular ICD-9 diagnosis, given the payor type. Hospital output is represented by the content of a patient's bill, including all room, medical supplies, drugs and ancillary services provided the patient during a single hospital stay.

The sampling universe for the hospital services PPI is taken from the American Hospital Association, stratified by hospital size, public vs. private ownership, and type of medical specialty. Once a hospital is identified as a sample unit, the BLS chooses a fixed subset of DRGs, and each hospital is then asked on a monthly basis to report on net transactions prices of a single representative patient bill (typically the last patient bill on file for that DRG). Since the identical treatment bundle is not always observed in subsequent months, BLS reporters construct subsequent fictitious DRG bundle prices by repricing the identical inputs.[12]

It is worth noting that with both the physicians' services and hospital services PPIs, BLS use of fixed itemized components for obtaining price quotes does not capture major input substitution of treatment for a condition, such as the changing mix of psychotherapy and psychotherapeutic drugs used for the treatment of acute phase depression. Such a zero-substitutability definition of physician and hospital output leads to the existence of a substitution bias, but that bias is of course not confined to medical PPIs, for it pervades the entire fixed weight Laspeyres price index computational procedures.

Finally, for our purposes it is also useful to note that the BLS publishes a PPI for

prescription pharmaceuticals. Although pricing the output of prescription pharmaceuticals presents some particularly interesting issues involving treatment of generic drugs, and of quality improvements in new products, those issues are beyond the scope of this paper.[13]

II.B BLS Procedures for Medical Care Related CPIs

According to the BLS, the CPI is "...a measure of the average change in the prices paid by urban consumers for a fixed market basket of goods and services."[14] Based on data from its Consumer Expenditure Surveys (CEX), the BLS identifies and defines a fixed 'market basket' of goods, employing a classification system known as the item structure, which is updated approximately every ten years. In January 1998 the BLS introduced its most recent major item structure and fixed weight revisions, based on the 1993-95 CEX; from January 1987 until January 1998, fixed weights had been based on the 1982-84 CEX.

In contrast to the BLS' recently initiated medical PPI program, the medical CPI has been published regularly since 1935.[15] Currently the BLS publishes a monthly aggregate medical care CPI (MCPI), sub-indexes for medical care commodities and medical care services, as well as for prescription drugs and medical supplies, nonprescription drugs and medical supplies, physicians' services, hospital and related services, and health insurance. Each of these medical-related CPIs is based on consumers' out-of-pocket expenditures (OOPPs), and thereby excludes all medical payments by governments, as well as employers' contributions to employee health insurance; only medical OOPPs plus that portion of third party insurance paid for out-of-pocket by employees is included within the scope of the MCPI. Thus, although national health spending in 1996 constituted 13.6% of GDP, the total weight given medical care items in the CPI in 1997 is only 5.4%.[16]

Recently the medical CPI has introduced a number of changes, some of which are similar to those implemented several years earlier in the PPI. Until at least 1990, in most cases the MCPI priced specific input items at list prices, e.g., "chargemaster" fees for x-rays, laboratory tests, and physicians' office visits, rather than at the average actual transaction price for treatment of, say, a child's forearm fracture to a managed care organization obtaining a hospital discount.[17]

Since 1993 the BLS has attempted to obtain hospital transaction rather than list prices, but through 1996, efforts "...yielded slow progress to date."[18] Beginning with the January 1998 major revisions, more aggressive efforts have been made to obtain actual hospital prices by payor, but information is not yet available on the composition and nature of hospital quotes now being obtained. While DRG hospital quotes have been considered for use in the MCPI, the BLS is instead contemplating pricing "package" treatments, consisting of "highly standardized and tightly defined components and risk factors" for conditions such as appendectomies, tonsillectomies and cataract surgery.[19] Details on how such treatment packages would be defined have not been released. Finally, while revisions involving hospital service components of the MCPI have been announced, the BLS has not yet published comparable information on revisions to the physician service components of the MCPI.[20]

II.C BLS Procedures for Pricing Medical Treatments: Comments

For some time now, health economists and government statisticians have pointed to directions toward which the pricing of medical care services should move. Here we briefly summarize several of the directions suggested by this literature. First, as early as 1962 Anne Scitovsky proposed

> "...an index which would show changes, not in the costs of such
> items of medical care such as drugs, physicians' visits, and hospital
> rooms, but in the average costs of the complete treatment of
> individual illnesses such as, for example, pneumonia, appendicitis,
> or measles."[21]

In Scitovsky [1967], this treatment episode approach was implemented on an illustrative basis for five medical conditions. Discussions of shortcomings and biases inherent in the BLS' MCPI approach, and of the preference for the treatment episode approach to medical price measurement, have appeared steadily since 1967; see, for example, the "Measuring Changes in the Price of Medical Care" chapter in various editions of the well-known health economics textbook by Paul Feldstein [1979, 1983, 1988], and in the Baxter Foundation Prize Address by Joseph Newhouse [1989].

Second, experts agree, the measurement of changes in the price of medical services

should incorporate major quality changes, such as those involving adjustments for improvements in health care outcomes. In 1992, for example, Paul Armknecht noted that

> "A new dimension needs to be included in the pricing of medical services that includes outcomes, so that if cancer treatment results in improved survival rates, this is reflected in the index."[22]

The Boskin Commission final report went further, asserting that "...we strongly endorse a move in the CPI away from the pricing of health care inputs to an attempt to price medical care outcomes."[23] Although incorporation of quality and outcomes changes into medical price indexes presents some significant conceptual challenges[24] and implementation difficulties, it is clear that failing to address these aspects is likely to result in price measures that are unreliable and inaccurate.

Third, use of fixed weights over extended periods such as 10-12 years in constructing sub- and aggregate price indexes is particularly inappropriate in the medical care sector, where both institutional and technological changes are rapidly occurring. Note that the frequency with which fixed weights are updated is distinct from the issue of which index number formula (e.g., Laspeyres, Paasche, or Tornquist approximation to the Divisia) is preferable. For the rapidly changing medical care sector, decennial updates of fixed weights with old weights being fifteen years out of date before the new revision occurs (i.e., use of 1982-84 CEX weights through December 1997) results in price indexes whose accuracy and reliability can legitimately be called into question.

In the research reported here, we extend previous research aimed at constructing CPI- and PPI-like medical price indexes that deal with prices of treatment episodes rather than prices of discrete inputs, that are based on transaction rather than list prices, that take quality changes and expected outcomes into account, and that employ more current expenditure weights in the aggregation computations. Before describing the components of this research, however, we digress to discuss the illness whose treatment price we measure, namely, acute phase major depression.

III. THE NATURE AND PREVALENCE OF MAJOR DEPRESSION

Depression is commonly characterized by melancholy, diminished interest in pleasure in all or most activities, sleep disorders, and feelings of worthlessness. In order for a patient's condition to be considered as an episode of major depression, a clinician must determine that a very specific set of clinical criteria have been met. According to the Diagnostic and Statistical Manual of the American Psychiatric Association (Fourth Edition, DSM IV), major depression is diagnosed when the following is observed:

The presence of one of the first two symptoms, as well as at least five of nine total symptoms. The symptoms must be present most of the day almost every day, for at least two weeks. The symptoms include:

1) depressed mood most of the day nearly every day;
2) markedly diminished interest or pleasure in almost all activities most of the day;
3) significant weight loss/gain;
4) insomnia/hypersomnia;
5) psychomotor agitation/retardation;
6) feelings of worthlessness (guilt);
7) fatigue;
8) impaired concentration (indecisiveness); and
9) recurrent thoughts of death or suicide.[25]

Two dimensions of depression involve its persistence. A single episode of the illness is self-explanatory given the above diagnostic criteria. A recurrent depression is defined by two or more major depressive episodes each separated by at least eight weeks of return to usual functioning.[26] Episodes of illness come and go, last from several weeks to several months, and are followed by periods of relatively normal mood and behavior. Untreated, the average depressive episode lasts from four to six months. Although the vast majority of individuals who experience an episode of major depression return to their original level of functioning, between 20 and 35% experience persistent symptoms; when these persistent symptoms last for 24 months or longer, these cases are referred to as chronic depression.[27] 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.[28] The

lifetime average for depressive episodes is five to seven, but as many as forty episodes have been reported.[29]

A number of studies have shown that depression has similar or greater functional impairments than those attributed to other episodic and chronic medical illnesses.[30] Episodes of depression can be classified according to severity: mild, moderate or severe. Mild depression typically involves the minimum number of symptoms required to meet clinical criteria and minor functional impairment. Moderately severe episodes are characterized by an excess in symptoms above the minimum to meet clinical criteria and by greater degrees of functional impairment. Severe major depression involves a number of excess symptoms above the minimum and significant degrees of functional impairment including the ability to work or conduct usual activities.

Epidemiological research indicates that in the early 1990s, 10.3% of the U.S. population met the criteria for major depression at some time during a 12 month period.[31] Depression is often accompanied by other forms of ill health, such as anxiety, eating disorders, substance abuse or other medical conditions.[32] Although the reasons are still not well understood, women -- particularly women under age 25 -- are much more likely to suffer from depression than are men; the relative lifetime female/male prevalence rate is about 1.7. Rates of recurrence and chronicity for major depression appear to be no different for women and men.[33]

IV. ALTERNATIVE TREATMENTS FOR ACUTE PHASE MAJOR DEPRESSION

The acute phase of major depression is typically defined as the stabilization of most acute symptoms. In practice, standards of care typically identify the acute phase as occurring over a 12-week period. In this research we allow protocol levels of acute phase treatment to occur over a 6-month period to recognize that actual practice departs from the controlled environment of efficacy research. In clinical practice, continuation therapy frequently follows on acute phase treatment in hopes of preventing relapse. If a recurrent episode of major depression occurs, maintenance treatments are often initiated to prevent further recurrences [Kupfer, 1991].

Given the information available in claims data, distinguishing between acute and continuation phase treatment is difficult. We employ acute phase standards of care to establish expected outcomes, but in our empirical implementation we undoubtedly will mix acute and other phases of care. Clinical research on the continuation phase of treatment is less developed, and definitive protocols have not been as widely adopted in many clinical settings.

A. Recent Developments in Psychotherapy and Antidepressant Medication

Treatments for acute phase major depression have advanced considerably during the last two decades. In the area of psychotherapy, a variety of new techniques has expanded treatment options well beyond traditional psychodynamic or psychoanalytic approaches. Interpersonal therapy (IPT), behavior therapy (BT), family therapy, and cognitive behavior therapy (CBT) are each relatively new. Evidence from controlled clinical trials suggests that when applied as the single mode of treatment for less severe forms of acute phase major depression, each of these therapies reduces depressive symptoms. Moreover, relative to antidepressant medication as the sole treatment, each has generally been shown to perform at comparable levels of efficacy and to have similar outcomes.[34]

Extraordinary advances have been achieved in the last two decades in the area of antidepressant medication. Although very recent developments expand the therapy set even further, over the 1991-95 period examined here three general classes of antidepressant medications were employed. These are: (i) cyclic antidepressants that include the widely used tricyclic antidepressants (TCAs) and a number of lesser known drugs such as trazodone; (ii) selective serotonin reuptake inhibitors (SSRIs) which include brand name drugs such as Prozac, Zoloft, Paxil and Luvox;[35] and (iii) monoamine oxidase (MAO) inhibitors which, due to considerable side effects and dangerous interactions, are generally used only for patients resistant to other forms of treatment.

Newer SSRIs offer some distinct advantages over older TCAs, although in randomized controlled trials, clinical efficacy rates are similar. SSRIs are associated with lower risk of overdose, and reduced levels and numbers of side-effects. Important side effects

frequently associated with TCAs include drowsiness, dry mouth, impaired ability to concentrate, seizures and weight gain.[36] The side effects most prominently associated with use of SSRIs relate to sexual dysfunction (particularly for males) and anxiety. The advantages of SSRIs come at a significantly higher pecuniary cost than most TCAs.

Psychotherapeutic interventions have frequently been combined with antidepressant medication as a combination strategy for treating major depression. The specific interventions that have been most intensively studied are the use of TCAs in combination with either IPT, BT or a general unspecified form of short-term psychotherapy. To date, no clinical studies have been reported in the literature which systematically assess the combination of the newer SSRIs and psychotherapy.[37] It is generally presumed, however, that such combinations will be at least as efficacious as the combination of TCAs and psychotherapy.

Finally, electroconvulsive therapy (EPT) is an effective treatment, but ECT is typically limited to rather special circumstances when the patient's depression is severe and is complicated by a number of other psychiatric symptoms including psychosis, catatonic stupor, or high risk of suicide.

In the analysis reported below, we focus on outpatient treatments for major depression; these constitute the vast majority of treatment episodes (75 to 80%). We do this for several reasons. First, inpatient claims data typically do not contain information on the drugs prescribed for treatment; thus characterizations of inpatient care are inherently incomplete.[38] Second, because of other incomplete information regarding illness severity and comorbid conditions, it is difficult to use administrative claims data to characterize fully an illness diagnosis, and therefore to make judgments about the appropriate use of hospital services for treating major depression. Third, few clinical trials specifically address inpatient treatment for major depression, making it difficult to assign outcomes to treatments. Finally, because there was considerable evidence of overuse of hospital services in the aggregate during the late 1980s and early 1990s, the inclusion of hospital services in our 1991 base year could make interpretation of price changes troublesome.[39] This strategy of limiting severity cases by excluding individuals hospitalized for

treatment of depression is likely to be only partly successful. During the 1990's there has been a substantial reduction in inpatient psychiatric admissions and in the length of stay for hospitalized cases. The implication of this is that the population of people treated only on an outpatient basis may be getting sicker over time.

In our analysis, we therefore focus on the use of various antidepressant medications alone, several forms of psychotherapy alone, and several drug-psychotherapy combination treatments. Because of very small sample sizes, we do not incorporate treatment involving ECT or the MAO inhibitors.

B. Results From Comparative Efficacy Studies

We now provide a brief summary of research results comparing the efficacy of alternative treatments for depression of varying severity. We divide depression severity into two classes: severe and less severe (hereafter, mild). We have reviewed approximately 30 major clinical trials and meta-analyses from the clinical literature on comparative efficacy of acute phase treatments.[40] This literature points to several key conclusions.

First, psychotherapies of all kinds have been shown to result in superior outcomes compared to no treatment. When compared amongst themselves, the different forms of psychotherapy appear to have no significant differences in outcomes.[41]

Second, for less severe forms of depression, psychotherapies alone, TCAs with medical management, and SSRIs with medical management appear to produce comparable outcomes. Each of these therapies produced significantly better outcomes than placebo treatments. Versions of these results have been reported in numerous large treatment trials and by meta-analyses of smaller clinical trials. Combination treatments with these as components also generate equivalent levels of efficacy for less severe forms of depression.

Third, for more severe forms of depression, the bulk of the evidence suggests that TCAs alone, SSRIs alone, and combinations of drugs and psychotherapy have comparable levels of efficacy, and each results in superior outcomes compared to psychotherapy alone. Recently some evidence has emerged showing some extra improvement from the combination treatments

16

relative to medication alone.[42] We believe it is premature to conclude that combination treatments offer significantly higher levels of efficacy than do antidepressant medications alone (or with medical management, as is typically the case).

Based on these observations from the literature, we view all the major treatment technologies as offering comparable expected outcomes for the average care of less severe acute phase depression. For severe depression, we view TCAs and SSRIs alone as comparable to each other and to combinations of TCAs and SSRIs with psychotherapy.

V. IDENTIFYING COMPARABLE TREATMENT BUNDLES FROM CLAIMS DATA

The results from our review of the clinical trials literature enable us to develop a set of treatment "bundles" that groups therapies into what we envisage as therapeutically similar groups for treatment of a specific form of major depression. This identification of treatment bundles that result in similar expected health outcomes is a crucial step in the construction of medical treatment price indexes that take expected outcomes into account. The implicit assumption we adopt here is that obtaining therapeutically similar outcomes from alternative treatments provides a useful approximation to achieving similar expected utility levels.[43]

A. The MEDSTAT Data

To identify empirically comparable treatment bundles for acute phase major depression, we employ a dataset consisting of retrospective insurance claims from four large self-insured employers that offered 25 health plans to 428,168 employees and their dependents for the years 1991 through 1995. The data were obtained from MEDSTAT Inc., and contain information on prescription drug claims, inpatient hospital treatment, outpatient visits, ICD-9 diagnoses, CPT4 procedures, and the demographic characteristics of all covered individuals. The health benefits offered to enrollees in this data base are quite generous relative to the general market for private health insurance in the US.

During the five years observed, important changes occurred in the terms of insurance coverage for mental health care. While the vast majority of plans represent so-called managed

17

indemnity plans (90%-94%), the management of mental health care benefits changed for a substantial number of enrollees between 1991 and 1995. Beginning on January 1, 1994, about eight percent of enrollees had their mental health coverage "carved out" to a specialty mental health managed care company.[44] In January 1995 an additional 35% of enrollees had their mental health benefits carved out.

It is reasonable to expect that these carve-out arrangements affected both the input prices and the quantities of specific delivered services, such as visits.[45] Changes in the extent of carve-out arrangements might also have affected the general clinical strategies used in treating major depression. Recent analyses by Wells et al. [1996] and by Berndt, Frank and McGuire [1997] show clear differences in treatment patterns between carve-out managed care plans and treatment of mental health care financed by general health insurers, with carve-out arrangements being associated with a higher likelihood of using prescription drug treatments.

Each outpatient and prescription drug claim can accommodate two ICD-9 diagnostic codes. In identifying cases of major depression, we used ICD-9 codes 296.2 (major depressive disorder - single episode) and 296.3 (major depressive disorder - recurrent episode) to define patients diagnosed with major depression. We do this for three reasons. First, the clinical trials literature has for the most part employed these definitions in their entry criteria. Second, chart reviews have indicated that the specificity of these two diagnoses is high, i.e., the proportion of true positives and false negatives is high, while the proportion of false positives is very low. Third, clinicians could employ a more ambiguous diagnosis such as "depression not elsewhere classified" or "neurotic depression". That clinicians designated their diagnosis as either 296.2 or 296.3 indicates a conscious act of volition.[46]

Using the diagnostic information and dates contained in the claims, we construct episodes of treatment. Since we do not directly observe symptoms in retrospective claims data, we cannot make our claims-based definition of an episode of treatment correspond directly to an episode of the illness.[47] When 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 follow American Psychiatric Association [1993] guidelines in defining an episode of depression as new if a diagnosis is preceded by a period of at least eight weeks of not meeting clinical criteria for depression.[48] Thus we use an eight week period without treatment to define new treatment episodes. In preliminary analyses, we experimented with alternative definitions, such as those involving six or twelve week intervals; results were essentially unaffected. For treatments lasting longer than six months we count quantities of visits and drugs meeting guideline standards for acute care occurring with the first six months of care (e.g., 15 psychotherapy visits even if year totals were 25 visits). To ensure that we consider the full set of claims associated with the acute phase of treatment, we exclude episodes beginning in the last six months of 1995, the last year in our sample.

When these criteria were applied to the MEDSTAT dataset, we defined 18,920 episodes of acute phase outpatient care over the 1991-95 time period. Because we cannot fully observe the treatment received for censored cases, we confine our attention to the 15,750 uncensored episodes in which we observe at least eight weeks without treatment before the beginning of an episode; as noted above, 1548 episodes beginning in the last six months of 1995 were also excluded. In order to limit the sample to less severe forms of depression, we eliminated individuals with episodes involving inpatient hospital treatment for a mental illness any time during the five years. This reduced the number of episodes to 10,067.

Using information on procedures (e.g., type of visit, whether drug prescribed) as given by the CPT4 codes, we can describe the composition of treatment that occurred within a treatment episode. Drug treatment is based on the national drug codes (NDC) reported on the claim. The NDC classification of antidepressant medications revealed use of seven TCAs, three SSRIs, two other serotonin-related drugs[49], three MAO inhibitors, four anxiolytics and four heterocyclics for treatment of depression. In terms of composition, 54% of the drug claims involved SSRIs, 19% TCAs, 19% anxiolytics and 7% heterocyclics.[50]

In previous research, we have reported that in the MEDSTAT claims data base, the share of treatment accounted for by treatment involving psychotherapy claims appeared to

increase significantly in 1995, after having fallen steadily from 1991 to 1994.[51] Subsequent research has indicated to us that while procedure codes are missing for many outpatient claims -- a common problem with claims data, the extent of missing claims in the MEDSTAT dataset is particularly large in 1991 through 1994. Thus, the apparent sudden increase in psychotherapy claims in 1995 could simply reflect considerable missing claims from 1991 through 1994. Hereafter we call the data underlying this previous research our "old" dataset.

To identify missing psychotherapy claims not explicitly delineated by CPT4 codes, for our "new" dataset we have developed an algorithm in which a psychotherapy procedure code is assigned to claims with missing CPT4 codes if two visits are within 14 days of each other, if they have the same charge, and if a previously psychotherapy-identified claim in the episode has the same charge.[52] Use of this algorithm contributed to an increase in the number of identified episodes.

Our primary analysis of the claims data considers only "pure" or "strict guideline" treatments. That is, initially we will only consider episodes of care that adhere strictly with treatment definitions as tested in the clinical trials literature. In this way we can directly link the "price" of an episode of a well-defined treatment to the price of other therapeutically similar treatments. However, we will also report results based on findings derived from episode treatment definitions that relax this stringent guideline restriction, in which we include treatment slightly below guideline criteria and those involving some continuation phase treatment.

In our strict guideline analyses we identify seven major classes of treatments that have been proven effective in the treatment of depression: (i) psychotherapy alone, 6-15 visits -- "PT alone"; (ii) short term TCA treatment alone or with medical management -- "TCA alone"; (iii) short term SSRI treatment alone or with medical management, 31-180 days -- "SSRI alone"; (iv) short term TCA treatment (31-180 days) with some psychotherapy -- "TCA+PT"; (v) short term SSRI treatment (31-180 days) with some psychotherapy -- "SSRI+PT"; (vi) short-term combined TCA/SSRI treatment, 31-180 days -- "TCA/SSRI"; and (vii) short term combined TCA/SSRI treatment (31-180 days) with some psychotherapy -- "TCA/SSRI+PT". With the exception of

psychotherapy alone, episodes with anxiolytics also being prescribed are combined into the appropriate class defined above. Use of seven major classes of treatments in this new dataset represents an expansion from the five used in our previously published research. Another significant change in the new data was to include episodes that involved long-term treatments or those extending beyond 20 visits or six months. For those episodes we included the components of care consistent with the acute phase bundles specified above and excluded spending on the remainder of the "episode." Using this expanded definition of an episode, we are more likely to include continuation treatment in the analysis. This may be especially true in the later years as new recommendations about continuation treatment diffused into practice. Thus, the new dataset may be "pricing" acute plus early continuation phase care in the later years. This would tend to bias an index upward. Evidence for this is reflected by an upward trend in the average duration of SSRI treatment during the five years (from 87 days in 1991 to 126 in 1995). This change contributed to a substantial increase in the number of cases used in the analyses.

The new dataset also improved the definition of censoring. In the "old dataset" to limit our analysis to episodes for which we had information on the entire episode, all cases which included treatments in the last 8 weeks of 1995 were eliminated. This means that more episodes were eliminated form 1995 than any other year. More severe episodes in need of longer term treatment may therefore have been disproportionately dropped from previous analyses. This may have biased the price index downward. We have now corrected that "late 95" censoring problem limiting our analysis to the first 6 months of treatment. As mentioned previously, to ensure we consider all treatments in the first 6 months of the episode, we do not consider episodes begun in the last half of 1995.

Finally, the new dataset allows drug switching if they are consistent with standards of care in clinical trials. Again, the change would tend to increase the severity and complexity of the treated population. Drug switching has been increasing over time as new pharmaceutical agents have become available (SSRIs and SNRIs).

It is informative to compare episodes defined in our new dataset, with that from the

old. As seen in Table 1, the most obvious change is that the number and proportion of episodes receiving guideline-compatible treatments increase very substantially, from 2348 to 5039, or, from 23% to 50%. There are 2980 guideline compatible episodes in the new dataset that were not in the old, and 289 in the old that do not appear in the new dataset.

There is also some evidence suggesting that in the new dataset, a greater number of episodes are associated with other mental health comorbidities. While new-old differences are relatively small in terms of substance abuse comorbidities in guideline compatible episodes (2.5% vs. 2.8%), 17.9% of acute phase depression episodes in the new dataset simultaneously involve treatment of schizophrenia or bipolar depression, while in the old dataset this proportion was a smaller 15.8%. For the various panic and anxiety comorbidities, the new-old proportions are 3.3% vs. 2.9. The literature, as well as our own discussions with clinician, suggest (1) that major depression and bipolar illness are often mistaken for one another, and (2) treatment of depression with comorbid psychoses and substance abuse is both common and difficult.

We interpret differences between the new and old datasets as therefore moving us away from the more pristine world of clinical trials to a more realistic and more complicated treatment environment in which mental illness comorbidities are more prevalent.

It is worth noting that differences between the new and old datasets in terms of other, non-mental health comorbidities are rather minor. As seen in the bottom panel of Table 1, for example, the proportion having no such comorbidities is 48.2% in the new dataset, and 48.3% in the old; relative to the old dataset, the proportion having one other medical comorbidity is slightly higher (37.9% vs. 36.2%), but is slightly lower for two or more comorbidities (13.9% vs. 15.6%).

Table 1

COMPARISON OF DATASETS BASED ON NEW AND OLD EPISODE DEFINITIONS

	New Data	Old Data	In New but Not in Old Data	In Old but Not in New Data
Total Episodes	10,067	10,368		
Episodes Guideline Compatible	5,039	2,348	2,980	289
Mental Health Comorbidities*				
Substance Abuse	2.5%	2.8%	2.6%	4.2%
Schizophrenia / Bipolar	17.9%	15.8%	20.1%	23.2%
Anxiety / Panic	3.3%	2.9%	3.9%	6.2%
Sum of Other Medical Comorbidities**				
0	48.2%	48.3%	47.5%	41.5%
1	37.9%	36.2%	39.3%	38.8%
≥ 2	13.9%	15.6%	13.2%	19.7%

* Panic/anxiety diagnoses are ICD-9 300, 300.1-300.4; schizophrenia/bipolar are all 295s, 296.0-296.1, 296.4-296.8; other includes obsessive/compulsive disorders (300.3), bulimia (783.6), depression not otherwise specified (311) and dysthymia (300.4).

** Other medical comorbidities are sense organs, circulatory system, cerebrovascular, digestive system, kidneys, prostate, pregnancy, and nervous system.

B. On Guideline Compatibility

Earlier we noted that of the 10,067 episodes identified in the new dataset from 1991 through 1995, 5039 (50%) were guideline-compatible. It is interesting to note that in our new dataset the proportion of episodes receiving guideline-compatible treatment rose from 35% in 1991 to 42% in 1992, then increased sharply in 1993 to 56%, and remained at 56% in 1994 and 1995.

The interpretation of guideline-incompatible treatments is difficult with claims data. For example, 1308 episodes (33% of the 3965) treated with psychotherapy alone consisted of a single visit. In addition, 1672 or 16% of all episodes received neither psychotherapy nor an antidepressant drug. Such single visits might have taken place for the purpose of "ruling out"

major depression as the relevant condition to be treated in favor of a somatic condition or another mental disorder. In such cases, the visit should not be viewed as "inappropriate treatment" but rather as an appropriate assessment, consistent with the Depression Guideline Panel statement [1993, p. 36]:

> "Effective treatment rests on accurate diagnosis. The practitioner must first determine whether the patietn has a clinical depression or is simply suffering normal sadness or distress...For patients who have very mild cases of major depression or whose diagnosis is unclear and who are not in immediate danger or are not suffering significant functional impairment, the practitioner may want to schedule one or two additional weekly evaluation visits to determine whether symptoms will abate without formal treatment..."

The implication of this is that with retrospective claims data, distinguishing treatment and assessment is quite uncertain. Hence we are somewhat unclear as to whether only 50-60% of care lies on the production frontier, or whether another 10-20% of treatments were dealt with properly but did not require treatment of the type studied here. It is important to note here that our use of guideline standards of care imposes a rather unrealistic shape to the production function for treatment of depression. It takes on a step function form. For example, if one were to receive 6 psychotherapy visits for treatment of depression, our analysis would treat the case as "effective" whereas 4 or 5 visits would be viewed as "ineffective." This is unlikely to be an accurate representation of clinical reality. Nevertheless, there is little systematic analysis upon which to make alternative assumptions. Thus, we use the step function production model as a point of departure.

With these difficulties in appropriate interpretation as a caveat, we now examine in greater detail the extent of guideline compatibility in the various treatment bundles of the new dataset.

Of the 10,067 episodes, 3765 were treated with psychotherapy alone, and 376 with psychotherapy and anxiolytics. Within the claims data, no distinction was made between different types of psychotherapy. Considering the 3044 treatments involving 50 minute psychotherapy

sessions, we find that 1658 received four or fewer visits, and 800 only a single visit. Of the 921 episodes involving 20 minute psychotherapy alone, 508 involved just one brief visit, while 178 episodes were treated with between two and five brief sessions. An additional 20 episodes involved some form of group therapy treatment, with only 25% of those having more than three visits.

While clinical trials data indicate that individuals show partial response to psychotherapy with six weeks of treatment (with weekly sessions) and remission in twelve weeks, published guidelines for the treatment of acute phase depression do not indicate any demonstrated effectiveness for fewer than six visits. The benefits of short psychotherapy visits, in the absence of antidepressant medication, have not been studied and therefore cannot be considered either effective or ineffective treatment. Although clinical trials and published treatment guidelines indicate psychotherapeutic treatment alone is an effective treatment, in our data, of the 3044 episodes given this treatment, only 1386 episodes (46%) can be considered to have completed a psychotherapy regimen that is consistent with guideline treatments.[53]

Turning attention now to treatments involving antidepressant medications, we begin by noting that claims data do not include information on how many days' medication was actually taken; as a proxy, we use the number of days of treatment for which a prescription was filled.

Of the 10,067 episodes considered, 231 were treated with TCA either alone (176) or in combination with anxiolytic medication (55). Of these 231 episodes, 55 (24%) were treated with less than 30 days of medication and 29 with ten or fewer days. Generally, the clinical literature (and the APA guidelines) indicate that while patients may show some improvement to antidepressant medication by the end of the first week, full response to acute phase depression may take four to six weeks.

An additional 741 episodes were treated with SSRIs either alone (604) or in combination with anxiolytic medication (137); 77 (10%) of these episodes were treated with fewer than 30 days of medication. Because individuals have differing reactions to drugs, some individuals are appropriately treated with one class of antidepressants, and then switched to

another class. In our sample, 95 episodes were treated with both TCAs and SSRIs, but 20 of these episodes had less than 30 days of both drugs.

Several episodes were treated with medications other than SSRIs and TCAs. In our sample ten episodes were treated with MAO inhibitors, while 164 were treated with heterocyclics. An additional 124 episodes were treated with anti-anxiety medication alone, a protocol which has not been approved by the FDA for the treatment of major depression. The use of anti-anxiety medication in the treatment of depression remains controversial. The use of alprazolam may be appropriate, if other medication is contraindicated. There is no clinical trial evidence for the efficacious use of other anxiolytic medications.[54]

Finally, in terms of combination treatments, the share of episodes treated with psychotherapy and antidepressant medication grew over time. Over the entire five year time period, 552 treatment episodes involved both some TCA and some psychotherapy, while 2169 included both SSRI and some psychotherapy. A large share (36%) of the episodes treated with combination treatments had three or fewer psychotherapy visits, and an additional 382 episodes were treated with some TCA, some SSRI and some psychotherapy.

C. Comparison With Results From Other Studies

The patterns of care observed in this dataset raise issues related to the likely effectiveness of treatment, given the substantial proportion of episodes involving guideline-incompatible treatments. One potential criticism of the patterns of treatment bundle data presented above is that it is based on retrospective claims data. Claims data are useful in that the retrospective medical treatment of many individuals can be analyzed efficiently and at minimal expense. In addition, such observational data reveal the "real world" practice of medicine, not the pristine clinical trial environment. Claims data are also used for quality assessment by organizations constructing "report cards" on health care organizations. Yet claims data has been fairly criticized for several reasons. The accuracy of diagnoses and recorded data are sometimes questioned, and omissions in records are common. For example, depression has been shown to be under-diagnosed by primary care physicians and over-diagnosed by psychiatric clinicians.[55]

Other studies have found treatment patterns for depression that are generally consistent with the patterns we observe here. The Medical Outcomes Study (MOS) consisted of 635 individuals diagnosed with depression or with current depressive symptoms for whom data was collected by self-administered questionnaires, patient diaries, phone interviews and health examinations. The MOS found that only 23% of depressed patients had used an antidepressant medication in the prior month or used it daily for a month or more in the prior six months. Of those patients using an antidepressant medication, 39% used an inappropriately low dose.[56]

The MOS did not report number of psychotherapy visits. Instead it reported 'counseling', defining it as three or more minutes of counseling during the screening interview. This makes comparison with published standards of care difficult. Although 90% of patients of mental health specialists were counseled, among general medical practitioners where most study participants were treated, only 20% of managed care and 40% of fee-for-service patients were counseled.

Another study of 88 outpatients enrolled in the NIMH Clinical Research Collaborative Program on the Psychobiology of Depression found that prior to entry into the study, only 19% of patients received an adequate dose and duration of antidepressant medication, while 24% received some anti-anxiety medication. Regarding psychotherapy visits, 44% were seen for at least one hour weekly.[57]

Thus, the substantial proportion of patients in the MOS and NIMH studies apparently not obtaining efficacious treatments, while perhaps surprising, is consistent with the treatment patterns found in the claims data we observe here.

For our purposes of constructing price indexes for the treatment of acute phase major depression, we must decide whether to utilize the data suggestive of treatment not consistent with FDA approvals and AHCPR guidelines. Since the interpretation of such treatments is problematic, in our "strict guideline" analysis we confine our attention to episodes of treatment defined as being consistent with AHCPR guidelines; however, in a separate analysis we also include treatment bundles that are "close" to guideline standards. Additional research on

guideline-incompatible care and the shape of the production function for treatment of depression is currently underway.

VI. QUANTITIES AND PRICES OF ALTERNATIVE TREATMENT BUNDLES

We begin our empirical analysis by characterizing quantities and prices of the seven treatment bundles. In Table 2 we report guideline-compatible treatment quantities, by year. Censoring at the beginning and end of our time span implies that in 1991 and in 1995, about 1500 episodes are identified, whereas in 1992-1994 the number is considerably larger at about 2400 per year. The two most common treatment bundles are SSRI+PT and PT alone. Because of varying sample size by year, quantity trends are more easily discerned by examining quantity proportions rather than absolute levels; guideline-compatible treatment bundle quantity proportions are reported in Table 3.

Table 2

GUIDELINE-COMPATIBLE TREATMENT BUNDLE QUANTITIES, BY YEAR

Treatment	1991	1992	1993	1994	1995	Total
PT Alone	214	311	321	311	229	1386
TCA Alone	32	45	43	37	19	176
SSRI Alone	52	113	188	205	110	668
TCA+PT	63	111	129	73	33	409
SSRI+PT	125	292	528	644	400	1989
TCA/SSRI	9	14	23	23	6	75
TCA/SSRI+PT	27	53	121	98	37	336
Sum	522	939	1353	1391	834	5039
Total Episodes	1479	2211	2426	2468	1483	10067
Percent Guideline-Compatible	35.3%	42.5%	55.8%	56.4%	56.4%	50.1%

A number of notable trends appear in Table 3. First, there has been a very substantial decline in PT alone treatments, from 41% in 1991 to 22% in 1994, and then up slightly to nearly 28% in 1995. Second, for the medication only treatments, SSRI alone has grown from 10 to 13%, even as TCA alone declined from six to two percent; the sum of the two medication only

treatments has remained relatively constant at 15-16%. Third, most of the compositional change among treatment bundles has involved the medication-psychotherapy combination treatments. While the TCA+PT combination fell from 12 to 4% between 1991 and 1995, the SSRI+PT treatment share doubled, from 24 to 48% of all treatments. By 1995, the SSRI+PT combination had become the modal treatment bundle.

Table 3

GUIDELINE-COMPATIBLE TREATMENT BUNDLE QUANTITY PROPORTIONS, BY YEAR

Treatment	1991	1992	1993	1994	1995	Total
PT Alone	0.409	0.331	0.237	0.224	0.275	0.275
TCA Alone	0.061	0.048	0.032	0.027	0.023	0.035
SSRI Alone	0.100	0.120	0.139	0.147	0.132	0.133
TCA+PT	0.121	0.118	0.095	0.052	0.040	0.081
SSRI+PT	0.239	0.311	0.390	0.463	0.480	0.395
TCA/SSRI	0.017	0.015	0.017	0.017	0.001	0.015
TCA/SSRI+PT	0.052	0.056	0.089	0.070	0.044	0.067

In Table 4 we report the average supply price for each treatment bundle, by year; recall that this supply price captures the sum of payments from insurers and patients/consumers to providers. As is seen there, annual price movements represent a mix of increases and decreases. Between 1993 and 1994, for example, the price of one bundle increases and that of all other six bundles decreases, while for 1994-95, price movements are reversed -- six up, and one down. Of particular interest is the extent to which compositional changes in quantities of treatment bundles appear to be negatively related to changes in relative prices. When entries in Tables 3 and 4 are considered together, it becomes clear that compositional quantity changes do not follow simple conceptions of downward sloping demand curves where quantity and "price" changes are clearly negatively related. From 1991 to 1992, for example, even as the supply price of SSRI alone more than doubled from \$52 to \$113, the SSRI alone share increased from four to six percent; similarly, the cost of the SSRI+PT combination bundle increased by about 150% from \$125 to

$292 between 1991 and 1992, yet the quantity share grew from 24% to 31%. It is important to note that the quality of treatment bundles may vary (e.g., side-effects). Also, insurance coverage drives a wedge between the supply and demand prices. Both of these may help explain observed patterns and they are discussed below.

Table 4

AVERAGE SUPPLY PRICE OF TREATMENT BUNDLE, IN DOLLARS BY YEAR

Treatment	1991	1992	1993	1994	1995	Average[**]
PT Alone	$876	$853	$866	$818	$819	$846
TCA Alone	236	310	318	204	226[*]	267
SSRI Alone	311	387	309	308	344	328
TCA+PT	932	882	745	706	941	820
SSRI+PT	941	983	949	958	972	961
TCA/SSRI	649[*]	361[*]	449	423	328[*]	439
TCA/SSRI+PT	1002	1021	1062	953	977	1010
AVERAGE[**]	805	817	802	788	825	

Notes: [*] denotes a cell with less than 20 observations
          [**] denotes a weighted average

This positive relationship between changes in supply price and quantity share is observed in other years as well. From 1992 to 1993, for example, the supply price of TCA+PT fell by about 18% from $882 to $745, but its quantity share also dropped from 12% to 10%. Between 1993 and 1994, while the cost of PT alone fell about 5% from $866 to $818, the PT quantity share fell from 26% to 23%. Similarly, for both TCA alone and TCA+PT, between 1993 and 1994 both prices and quantity shares fell.

What these quantity and supply price data suggest, therefore, is that factors other than supply price are likely to have induced compositional changes among treatment bundles for acute phase major depression. Prominent among these, we conjecture, is the increased knowledge and experience gained by physicians on the efficacy and effectiveness of the SSRI medications,

particularly in combination with a limited amount of psychotherapy and changing insurance arrangements.

Although compositional changes among treatment bundles may reflect a host of price and non-price influences, it is clear that the dollar expenditure shares among the seven treatment bundles have been greatly affected between 1991 and 1995. These dramatic changes are evident in Table 5. While the expenditure share going to PT alone fell from 45% to 27% between 1991 and 1995, by 1995 the SSRI+PT bundle was responsible for more than half (56%) of all dollar expenditures, up from 28% in 1991. In 1991 the three singleton bundles of PT alone, TCA alone and SSRI alone accounted for 50% of all treatment dollars in 1991, but by 1995 they only captured 33% of expenditures.

Table 5

TOTAL EXPENDITURE SHARES BY TREATMENT BUNDLE, BY YEAR

Treatment	1991	1992	1993	1994	1995	Average
PT Alone	0.446	0.346	0.256	0.232	0.273	0.310
TCA Alone	0.018	0.018	0.013	0.007	0.006^*	0.012
SSRI Alone	0.038	0.057	0.054	0.058	0.055	0.052
TCA+PT	0.140	0.128	0.089	0.047	0.045	0.090
SSRI+PT	0.280	0.374	0.461	0.563	0.565	0.449
TCA/SSRI	0.014^*	0.007^*	0.010	0.009	0.003^*	0.008
TCA/SSRI+PT	0.064	0.071	0.118	0.085	0.053	0.078

Note: * denotes a cell with less than 20 observations

Together, the patterns observed in Tables 2 through 5 suggest to us that rather than PT and the SSRIs being viewed as simple singleton substitutes in the production of treatments for acute phase depression, they are more accurately envisaged as being complementary in the sense that the PT+SSRI combination has now become the modal treatment.

To this point, our discussion of treatment bundle price has focussed on the supply price -- the sum of payments received from insurers and patients/consumers. The demand price

notion reflects the fact that patients receiving treatment for acute phase depression provide direct payments in the form of copayments and deductibles. In Table 6 we report average demand prices for each of the seven treatment bundles, annually. Trends in the demand price differ quite markedly from those of the supply price.

Recall from Table 4 that the average supply price across all bundles fell slightly from $805 in 1991 to $788 in 1994, and then increased to $825 in 1995. In contrast, as seen in Table 6, from 1991 to 1994 the average demand price increased from $100 to $105, and then it increased very sharply to $128 in 1995. This sharp increase of 22% between 1994 and 1995 in the average demand price across all treatment bundles reflects corresponding jumps in the PT alone bundle (26%), SSRI alone (19%), SSRI+PT (17%), and TCA/SSRI+PT (21%).

Two main factors may account for changes in demand prices that differ so starkly from supply price changes. First are the changes in cost sharing provisions. These are documented in the MEDSTAT datset. The MEDSTAT data indicated to us that in 1995, for a number of insurance plans, benefit changes were introduced that increased substantially the copayments and deductibles required of plan enrollees.[58] Second is the tendency to use out of network providers in the context of managed care. Coverage of care provided by non-network providers typically carries higher copayments than does care received from a network clinician. During the time period observed, networks were becoming more restrictive and behavioral health carve-outs were introduced, which may have increased the share of care provided by non-network providers. The impact of these benefit design changes can be seen by computing the portion of the total treatment cost borne directly by the patient, i.e., computing the ratio of the demand price in Table 6 to the corresponding supply price in Table 4; we call this the OOPPs ratio.

Table 6

**AVERAGE DEMAND PRICE OF TREATMENT BUNDLE,
IN DOLLARS BY YEAR**

Treatment	1991	1992	1993	1994	1995	Average**
PT Alone	121	122	122	123	155	128
TCA Alone	39	46	46	48	51	46
SSRI Alone	23	30	24	27	32	27
TCA+PT	128	124	116	113	117	120
SSRI+PT	102	105	115	121	142	120
TCA/SSRI	43*	47*	46	51	21*	45
TCA/SSRI+PT	102	121	131	135	163	132
AVERAGE**	100	101	102	105	128	

Note: * denotes a cell with less than 20 observations
       ** denotes a weighted average

For the SSRI alone bundle, the OOPPs ratio increased only marginally, from 7% in 1991 to 9% in 1995. For the very low total cost TCA alone bundle, the OOPPs ratio increased more markedly, from 16% in 1991 to 23% in 1995; in absolute terms, this additional burden borne by the patient was only $12 (from $39 to $51). For the combination treatments involving PT and the SSRIs, however, the increase was much more substantial, both in terms of OOPPs ratio and absolute dollar burden. In the SSRI+PT bundle, for example, the OOPPs ratio rose from 11% to 15% between 1991 and 1995, with the demand price increasing $40 from $102 to $142; even more dramatically, for the TCA/SSRI+PT bundle the OOPPs ratio rose from 10% to 17%, while the demand price jumped $61 from $102 in 1991 to $163 in 1995. The changes resulted from increased cost sharing for psychotherapy.

One important implication of these findings is that when considering construction of price indexes for the treatment of a medical condition such as depression, it is imperative that demand prices be distinguished from supply prices, for changes in insurance plan design benefits

33

over time can introduce sharp differences into their time trends.

VII. AGGREGATE PRICE INDEXES FOR THE TREATMENT OF DEPRESSION

        With the prices and quantities of the seven treatment bundles as elementary building blocks, we now move on to the construction of aggregate supply and aggregate demand price indexes. A variety of price indexes can be computed, each having implicit assumptions on the extent of *ex ante* substitutability among the seven bundles. Since we have discussed these alternative index number formulae and underlying assumptions elsewhere[59], and since they are well-known in the literature on index numbers[60], we provide only a very brief discussion here.

        Fixed quantity weight indexes such as the Laspeyres and the Paasche reflect an assumption of zero *ex ante* substitutability among the seven treatment bundles; the Laspeyres employs a base period fixed weight, the Paasche the final period fixed weight. The Cobb-Douglas employs a fixed expenditure weight (here, equal to the mean expenditure share for each bundle over time), and assumes the elasticity of substitution among treatment bundles equals unity for all pairs of bundles. When demand curves are downward sloping and bundle prices are all increasing, between any two adjacent time periods the Laspeyres price index is greater than the Paasche.

        In general the chained, or sequentially updated price indexes are viewed as being preferable to the fixed weight indexes. The Tornquist discrete approximation to the Divisia, and the Fisher Ideal, make no *a priori* assumptions on the extent of substitutability among treatment bundles. The Tornquist weights percentage changes in each of the treatment bundles by the arithmetic mean of the expenditure share in the two time periods, whereas the Fisher Ideal is the geometric mean of the corresponding Laspeyres and Paasche indexes. A commonly observed empirical pattern is that the Tornquist and Fisher Ideal price indexes are very close to one another, and both tend to be in between the Laspeyres and Paasche price indexes.

        With this as background, we now report results of various aggregate price index calculations, first in Table 7 as viewed from the supply price (i.e., PPI) vantage, and then in Table 8 from the demand side (i.e., CPI).

The most striking result that immediately emerges from Table 7 is that, regardless of which index number procedure is employed, over the 1991-95 time period the treatment price index for acute phase major depression has hardly changed, remaining at 1.00 or falling slightly to around 0.97. Each of the price indexes reveal an increase from 1991 to 1992, they all fall in 1993, all fall again in 1994 (to between 0.95 and 0.97), and then all increase in 1995. Given some of the data problems in 1995, this last finding should be viewed cautiously. Differences among the various fixed and chained indexes are relatively minor.

Table 7

ALTERNATIVE AGGREGATE PRODUCER ("SUPPLY") PRICE INDEXES FOR THE TREATMENT OF ACUTE PHASE MAJOR DEPRESSION

Price Index	1991	1992	1993	1994	1995
Fixed Weights:					
Laspeyres	1.000	1.003	0.975	0.931	0.976
Paasche	1.000	1.027	0.996	0.974	1.000
Cobb-Douglas	1.000	1.018	0.985	0.953	0.992
Chained Weights:					
Laspeyres	1.000	1.003	0.969	0.938	0.968
Paasche	1.000	1.011	0.976	0.952	0.978
Fisher Ideal	1.000	1.007	0.97	0.945	0.973
Tornquist	1.000	1.007	0.972	0.945	0.972
BLS PPIs[*]					
Aggregate PPI	1.000	1.012	1.025	1.031	1.051
Health Services				1.000	1.024
Antidepressants	1.000	1.076	1.134	1.162	1.204
Psychiatric Hospitals		1.000	1.024	1.060	109.9

[*] BLS indexes normalized to appropriate base year

By comparison, in the bottom panel of Table 7 we report various PPIs published by the BLS. Over the 1991-95 time period, the aggregate PPI for all finished goods increased about 5%, and that for antidepressant prescription drugs increased by about 20%. From 1992 to 1995, the PPI for psychiatric hospital services increased by about 10%, while between 1994 and 1995

the overall health services PPI increased 2.4%. Thus, while our various supply price indexes for the treatment of acute phase depression are either flat or very slightly falling, they all grow considerably less than the various official PPIs, i.e., the "real", inflation adjusted price index for the treatment of acute phase depression has fallen over the 1991-95 time span.

Table 8

ALTERNATIVE AGGREGATE CONSUMER ("DEMAND") PRICE INDEXES FOR THE TREATMENT OF ACUTE PHASE MAJOR DEPRESSION

Price Index	1991	1992	1993	1994	1995
Fixed Weights:					
Laspeyres	1.000	1.028	1.041	1.063	1.266
Paasche	1.000	1.037	1.080	1.118	1.334
Cobb-Douglas	1.000	1.037	1.069	1.110	1.300
Chained Weights:					
Laspeyres	1.000	1.028	1.048	1.081	1.281
Paasche	1.000	1.032	1.065	1.104	1.317
Fisher Ideal	1.000	1.030	1.056	1.092	1.299
Tornquist	1.000	1.030	1.057	1.093	1.298
BLS CPIs[*]					
All Items	1.000	1.030	1.061	1.088	1.119
Medical Care	1.000	1.075	1.146	1.205	1.266
Prescription Drugs	1.000	1.075	1.117	1.155	1.177

[*] BLS indexes normalized to 1991 base year

Index number afficionados might notice, however, that the typically observed inequality relationships among the Laspeyres and Paasche price indexes are not present here. For example, for both fixed and chained weight versions, the Paasche price index is larger than the Laspeyres. The reason this occurs is that, as noted in Section VI above, in a number of cases changes in treatment bundle supply prices and quantity shares are positively rather than negatively correlated. Because the price index calculations do not take into account factors affecting demand other than price, and because consumers do not face supply prices, there is the appearance of a positive relation between quantity and supply price. We discuss other price-

related issues in Section VII below, where we deal with hedonic price indexes. Finally, note that as is usually the case, the Fisher Ideal and Tornquist indexes are always in between the chained Paasche and Laspeyres indexes.

We now turn to a discussion of aggregate price indexes viewed from the demand side, i.e., the patient/consumers' OOPPs expenditures. As seen in Table 8, the demand aggregate price indexes have very different time trends than do the supply aggregate price indexes. Each of the price indexes exhibits an increase between each pair of adjacent years, with the 1994-95 price increase being particularly dramatic. By 1995, the fixed weight price indexes (normalized to 1.000 in 1991) ranged between 1.27 and 1.33, while the chained indexes had a slightly smaller variation, between 1.28 and 1.32. Between 1991 and 1994, the price indexes increased by only about 10%, but in 1995 the increase was much larger, averaging around 20%. As discussed in Section VI, we attribute this sharp increase to benefit design changes in 1995 that increased patient copayments and deductibles.

The discrepancy between the PPI and CPI results highlight several important points in interpretation of the data. First is that if the elasticity with respect to the demand price is less than unity (in absolute value), spending must increase. Since cost sharing increased notably and demand estimates suggest an elasticity of less than unity, it is not surprising that the CPI increases over time. Second, and more important, is the welfare interpretation of these changes. It is well known that low levels of demand side cost sharing in medical insurance is associated with welfare losses due to moral hazard [Newhouse, 1993]. Thus, changes in the CPI for depression may reflect reduced moral hazard in addition to higher prices per unit of "effective care." Research by Manning and Marquis [1993] suggests that 50% cost sharing for mental health care was optimal in a fee-for-service context. Because observed rates in our data are quite a bit lower, gains from reduced moral hazard may be significant. Because interpretation of the CPI is unclear, we focus more attention on the PPI.

VII. HEDONIC PRICE INDEXES

An enduring issue in price index measurement is how one should make adjustments for quality change. Quality adjustments could be significant here, for it is reasonable to expect that the characteristics and attributes of the acute phase of treatment for depression has changed over time, and that changes in the underlying patient population might also have occurred.

Although its use to date in the medical marketplace has been limited, the hedonic price approach has been used in many other contexts to adjust price movements for (not fully priced) quality changes.[61] We now briefly outline our implementation of hedonic price measurement for the treatment of acute phase depression.

We expect that treatment costs for depression are affected by patient characteristics, various attributes of antidepressant medications, the type of treatment given the patient, and the year in which the treatment episode began.

Patient characteristics for which we have data include the patient's gender, age, general medical condition (measured by the number of comorbid general medical conditions, as noted on the bottom of Table 1), the industry in which the plan enrollee is employed (transportation, communication, utilities; services; or government), and information on the patient's mental health comorbidities. With respect to these mental health comorbidities, we construct dummy variables for whether the patient is also diagnosed with a substance abuse disorder, a panic/anxiety disorder, schizophrenia or bipolar depression, depression not otherwise specified, and whether the current episode involves a recurrence of major depression.

For attributes of the various antidepressant medications, we employ data on the half-life of the medication, and on the number of side effects that frequently accompany use of the medication.[62]

Regarding treatments, we create dummy variables for each of the seven guideline-compatible treatment bundles described earlier. We also create a dummy variable indicating whether at the start of the episode the individual was enrolled in a behavioral health carve-out. Finally, we create dummy variables indicating the year in which the treatment episode began.

The dependent variables in the hedonic regressions are the natural logarithm of the

supply price and of the demand price. The explanatory variables are the patient, medication, treatment and time variables discussed above. The set of omitted dummy variables, and thus the reference case, is that for a female; where the enrollee is employed in the transportation, communication or utilities industry; has none of the mental health comorbidities noted above; is not being treated for a recurrent episode of acute phase depression; treatment consists of PT alone; and the treatment episode began in 1991. Parameter estimates for the supply price equation are in the left panel of Table 9, while estimates for demand price equation are in the right panel. A number of results are worth mentioning in detail. We begin with the supply price.

First, characteristics of the patient and their illness have a statistically significant impact on treatment costs. While age does not appear to affect the supply price, mental health comorbidities such as anxiety/panic, schizophrenia/bipolar depression, and depression not otherwise specified all have a positive and significant impact. Somewhat surprisingly, comorbid substance abuse has a negative coefficient estimate, although not significant. The sum of the patient's other general medical comorbidities is also negative, but not significant. Finally, treatment costs for patients in which the enrollee is employed in the services or government sectors are considerably lower than if the industry involves transportation, communications and/or utilities.

Second, attributes of the antidepressant medications also affect supply prices of treatment, even when one controls for type of treatment. Specifically, within the various bundles involving prescription drugs, greater side effects are associated with lower supply prices (lower side effects with higher supply prices), and increased half-life of the drug (mitigating the negative impacts of occasional forgetting to take the medication) also raises the supply price. Relative to PT alone, treatment with TCA alone or SSRI alone reduces the supply price significantly, while use of SSRI+PT or TCA/SSRI+PT increases supply price.

Table 9

PARAMETER ESTIMATES FROM HEDONIC PRICE REGRESSIONS
(Absolute value of t-statistic in parentheses, from robust standard error)

Variable	Supply Price ("PPI") Equations		Demand Price ("CPI") Equations	
Constant	6.528	6.587	4.614	4.658
	(112.60)	(98.91)	(51.88)	(47.39)
Male	-0.016	-0.005	-0.090	-0.056
	(0.75)	(0.18)	(2.26)	(1.29)
Age	0.003	0.000	0.011	0.008
	(3.38)	(0.09)	(6.71)	(4.19)
Anxiety/Panic	0.258	0.350	0.126	0.266
	(5.68)	(7.06)	(0.89)	(1.84)
Schizophrenia/	0.235	0.393	0.233	0.458
Bipolar	(9.94)	(14.62)	(5.86)	(10.56)
Other Depression	0.128	0.183	0.070	0.125
	(5.32)	(6.16)	(1.75)	(2.61)
Substance Abuse	-0.136	-0.093	-0.529	-0.478
	(1.38)	(0.87)	(2.71)	(2.36)
Recurrent	0.080	0.124	0.050	0.132
	(4.04)	(5.32)	(1.44)	(3.42)
Sum Medical	-0.025	-0.024	-0.029	-0.031
Comorbidities	(2.04)	(1.58)	(1.29)	(1.24)
Government	-0.121	-0.172	-0.858	-0.936
	(3.07)	(3.74)	(17.21)	(16.10)
Services	-0.184	-0.225	-0.991	-1.053
	(4.62)	(4.92)	(21.59)	(19.53)
Sum Side Effects	-0.084	-0.129	0.025	-0.104
	(3.33)	(10.45)	(0.78)	(6.26)
Half-Life	0.001	0.001	0.000	-0.001
	(4.01)	(4.47)	(1.31)	(3.53)
TCA Alone	-1.170	na	-1.483	na
	(10.15)		(8.77)	
SSRI Alone	-0.892	na	-1.847	na
	(13.50)		(19.18)	
TCA+PT	0.137	na	-0.116	na
	(1.38)		(0.89)	
SSRI+PT	0.195	na	-0.122	na
	(3.21)		(1.41)	
TCA/SSRI	-0.654	na	-1.320	na

	(5.71)		(8.32)	
TCA/SSRI+PT	0.301	na	0.007	na
	(4.15)		(0.07)	
Carve-Out	-0.047	0.025	-0.028	0.092
	(0.90)	(0.43)	(0.44)	(1.24)
Start 1992	0.048	0.072	0.108	0.099
	(1.23)	(1.60)	(1.62)	(1.35)
Start 1993	0.004	0.054	0.114	0.094
	(0.11)	(1.21)	(1.71)	(1.30)
Start 1994	-0.021	0.025	0.145	0.097
	(0.54)	(0.58)	(2.09)	(1.31)
Start 1995	0.026	0.059	0.366	0.296
	(0.59)	(1.14)	(4.92)	(3.53)
R^2	0.328	0.088	0.277	0.083
RMSE	0.689	0.802	1.168	1.315
N	5034	5034	5034	5034

Somewhat surprisingly, we find that whether at the beginning of the treatment episode the patient was enrolled in a mental health carve-out has no significant impact on supply price. To check whether the carve-out impact was diluted because of differential treatment bundles used by the carve-out, we also estimated a supply price equation in which the treatment bundle dummy variables were deleted; results are given in the second column of parameter estimates in Table 9. Although the goodness of fit was reduced considerably when the treatment bundle variables were deleted, the carve-out parameter estimate remained insignificant. This may be due in part to the fact that the comparison conditions involved managed indemnity arrangements which tend to negotiate price discounts and shift care patterns.

Of particular interest are the parameter estimates on the 1992-95 yearly dummy variables. In the context of the hedonic price equation, these parameter estimates indicate how the supply price changed relative to 1991, holding constant the various patient characteristics, medication attributes, and the treatment bundle composition. Each of the four parameter estimates is small and insignificantly different from zero; hedonic price indexes computed as anti-logarithms of these coefficients are 1.049, 1.004, 0.980 and 1.027 for 1992 through 1995, respectively (1991=1.000). These quality-adjusted hedonic price indexes are consistent with

41

traditional aggregate supply price indexes reported in Table 7 above. Adjusting for various patient, medication and treatment changes over time does not appear to affect the central finding that from 1991 to 1995, there was very little change in the supply price for treating acute phase major depression.

We now turn to the demand price hedonic equations, whose parameter estimates appear in the last two columns of Table 9. In general, the pattern of results from the hedonic demand price equations involving quality attributes is not as strong and dramatic as is that of the hedonic supply price equation. For example, while parameter estimates on the patient's mental health comorbidities are of the same sign, in the demand price equation a number are no longer statistically significant; attributes of the antidepressant medication are not significant as well. As with the supply equation, the mental health carve-out coefficient is insignificant.

However, in contrast to the supply price equation, for the demand hedonic price equation the parameter estimates on the 1992-95 year dummy variables are positive, and by 1994 and 1995 they become statistically significant. Relative to 1991 = 1.000, holding constant patient characteristics, medication attributes and treatment composition, the hedonic price indexes for 1992-95 (computed as anti-logs of the estimated parameters) are 1.114, 1.121, 1.155 and 1.443. Compared to the traditional aggregate demand price indexes reported in Table 8, the time trends in these hedonic price indexes are broadly similar. While the traditional aggregate demand price index in Table 8 averages about 1.30 in 1995, the corresponding hedonic quality-adjusted demand price index is 1.44; if one takes the antilog of the 1995 parameter estimate in Table 9 minus one times the robust standard error, one obtains a 1995 price index of about 1.34, not much different from that in Table 8. We conclude, therefore, that traditional and hedonic approaches to price index measurement yield broadly similar results.

IX. DISCUSSION

Our goal in this research project has been to extend previous research aimed at constructing CPI- and PPI-like medical price indexes that deal with prices of treatment episodes

rather than prices of discrete inputs, that are based on transaction rather than list prices, that take quality changes and expected outcomes into account, and that employ more current expenditure weights in the aggregation computations. Although we have made considerable progress in achieving this goal, we believe a number of caveats are in order.

First, the results reported here differ considerably from those reported earlier in Frank, Busch and Berndt [1998] and Frank, Berndt and Busch [1998]; there we reported substantial supply and demand price declines over the 1991-95 time period. Although we are still examining why the differences occurred, we believe that the primary difference is that in the new dataset we bring in many more patients. The evidence suggests that over time the new data has an increasing share of patients with: (1) more complicated conditions, (2) greater severity of illness, and (3) elements of longer term treatments such as continuation phase care. While the older dataset may have corresponded with episodes more closely approximating the pristine world of clinical trials, the larger, new dataset involves patients with more typical complicated illnesses.

It is therefore of interest to examine how robust are our new dataset findings when one relaxes the strict guideline-compatible standards we imposed in defining episodes of care. We have created yet another dataset, where now we relaxed the definitions of the seven bundles. Bundles which consider drugs alone as treatment remain unchanged at 30-180 days of treatment. Definitions of bundles which include psychotherapy have been expanded such that the cost of up to 20 psychotherapy visits is now included in the cost of the episode. Furthermore, for the treatment "psychotherapy alone" we previously allowed for a minimum of 5 visits. This constraint was relaxed to include episodes treated with 3 or 4 visits. The appendix table shows that with the broader definition of standards of care the price indexes are largely unchanged. Specifically, they are rather flat, not falling below 2 percentage points in comparing 1991 and 1995.

It is also useful to comment on implications of our research findings. Using alternative index number procedures, we find that over the 1991-95 time period, the nominal supply price of acute phase treatment for depression is essentially unchanged, while the real

supply price (relative to GDP deflator) has declined. Note that this may well create an upward

bias in estimated price movements for acute phase treatment of depression. Recall that we

showed that the percentage of episodes meeting guideline standards has increased over time. The

price indexes reported above do not take account of this apparent move toward the production

frontier. Improved compliance with standards of care implies greater benefits are being generated

over time, thereby understating price reductions of effective care. The demand price index

revealed a steady 2-3% per year increase from 1991 to 1994, which is then followed by a sharp

20% increase in 1995. Since the demand price encompasses only consumers'/patients' direct

OOPPs expenditures, and since the supply price is the sum of consumer plus insurer payments, it

follows that from 1991 to 1994 there was a steady 2-3% annual *decrease* in insurers'

expenditures, and that in 1995 a very substantial 20% *decrease* took place. While the nominal

supply price of acute phase treatment for depression was apparently unchanged between 1991 and

1995, the compositional burden shifted from third party insurer to patient. The change in

incidence of payment, in large part due to increased patient copayments and deductibles, likely

reflects efforts by insurers to deal with moral hazard via design of mental health benefits and

tightening of the size of provider networks used in managed care.

　　　　Our research findings based on index number procedures are broadly consistent with

those based on simple hedonic price methods, particularly with the supply price. On the demand

side, the hedonic price index increases more rapidly to 1994, and then increases more sharply in

1995 than do the traditional index numbers. Note that what we call "traditional" here, however,

represents a substantial departure from current BLS practice, in that we have used as elementary

building blocks the transaction prices accompanying alternative treatment bundles, and not the

(list) prices of fixed set of inputs. Nevertheless, it is interesting that, as in Berndt, Cockburn and

Griliches [1996], the marginal impact of incorporating hedonic pricing methods is not that

significant, once one defines output quantities carefully. Our hedonic research could have gone

several steps further, allowing for changing parameters over time and integrating the estimated

hedonic price equations with index number procedures, but we leave those nuances for further

44

research.[63]

Finally, it is worth commenting on implications of this research for procedural revisions currently under possible consideration at the BLS. The BLS is responsible for constructing and publishing PPIs and CPIs for a very large number of medical-related services and commodities. Our research has focussed on but one disorder -- major depression -- albeit one whose output measurement presented considerable challenges. We believe that economies of scope and scale are undoubtedly available, and we have experienced a steep learning curve in working with the MEDSTAT data. We also believe that extension of the treatment episode approach to other illnesses and disorders is feasible. Nevertheless, we are struck by how difficult, time-consuming and expensive this research has been for us. If a treatment episode approach is to be implemented efficiently at the BLS (or elsewhere, such as at HCFA), considerable care needs to be exercised in choosing a set of illnesses/disorders for which treatment bundles can be well-defined, for which quality comparisons can readily be made (facilitated by AHCPR or other published professional guideline standards), and for which claims data can readily be employed.

X. CONCLUDING REMARKS

A very brief summary of our research is as follows. We have implemented an approach that employs transaction data from a publicly available retrospective medical claims data base of almost half a million lives, annually from 1991 to 1995. Based on a review of the clinical research literature dealing with the acute phase of treatment for depression, we identify seven alternative service bundles that combine varying types and quantities of prescription drugs, medical management and psychotherapy. We construct data on episodes of treatment, and their cost. Because the treatment bundles are viewed by the medical community as being therapeutically similar in terms of *ex ante* efficacy, our linking of treatment bundles provides an important step toward incorporation of expected medical outcomes. We distinguish a supply price index, similar to the BLS' Producer Price Index, that represents the total receipts received by providers of medical treatment (from the insurer and the patient), from a demand price index

(similar to the Consumer Price Index) that incorporates only the out of pocket payments (via copayments and deductibles) by the patient/consumer. Finally, we employ a variety of traditional aggregate index number formulae, consistent with varying assumptions concerning the *ex ante* substitutability among the seven treatment bundles, as well as hedonic price procedures in computing price indexes.

For the supply side price index, the various price indexes all point to a flat and essentially unchanging price over the 1991-95 time period. Since the BLS' various producer price indexes all indicate a significant increase over the same time period, our results suggest that in real terms, the supply price for treatment of acute phase depression has fallen over time. On the demand side, however, the indexes show a total increase in price of about 10% between 1991 and 1994, and then a very sharp 20% increase. The usual welfare interpretation of the CPI is not possible in this case. The price increase largely reflects increased cost sharing arrangements which are certain to reduce welfare losses stemming from moral hazard. This is especially the case in MEDSTAT data which generally covers plans with substantially lower cost sharing for mental health care than is typical in the economy as a whole (20% vs. 50%).

Our research can be extended in a number of ways. Two issues are particularly important. First, although we have experimented with use of both strict and somewhat relaxed guideline criteria to define episodes of care, in each case we have assumed an "all or nothing" treatment. The implicit production function therefore involves a steep step or cliff in which treatment either is or is not effective. An alternative is to consider some sub-guideline treatment bundles as having a lower probability of being efficacious than are those meeting guideline criteria, and therefore to assign the sub-guideline episodes a lower quantity weight. Together with a group of clinicians and psychiatric researchers, we are currently initiating an effort to more completely define the specific depression treatment production function.

Second, we have adhered rather closely to the notion of defining output in terms of episodes of acute phase treatment. This involves a somewhat arbitrary delineation of an output that often does not correspond well to the reality of clinical practice. Since depression is often a

chronic recurring illness, continuation and maintenance phase care are part of the overall treatment of the disorder. Unfortunately, since the bulk of clinical research has focused on the acute phase of treatment, our links to outcomes were tied to a somewhat imperfect characterization of care. It is therefore important to more completely specify treatment in this context.

Finally, and perhaps most surprisingly, all of our results point to the fact that given a budget for treatment of depression, much more could be accomplished in 1995 than in 1991. That is, real prices of care have fallen. This runs counter to most public and expert perceptions. The implications for interpreting spending changes are enormous. For years the view has been that spending increases on mental health care (of which depression is 50% in private insurance) was due to (a) provision of increasing low benefit services, and (b) higher payments to providers. Our results point to a different story where spending increases are due to a larger number of "effective" treatments being provided. These treatments represent a shift towards new treatment technologies that are provided under a new set of organizational arrangements.

Appendix Table

**AGGREGATE PRODUCER ("SUPPLY") PRICE INDEXES CALCULATED
USING BROADER STANDARD OF CARE DEFINITIONS**

Price Index	1991	1992	1993	1994	1995
Fixed Weights:					
Laspeyres	1.000	1.015	0.984	0.929	0.997
Paasche	1.000	1.029	1.000	0.968	1.010
Cobb-Douglas	1.000				
Chained Weights:					
Laspeyres	1.000	1.015	0.981	0.941	0.985
Paasche	1.000	1.017	0.986	0.952	0.993
Fisher Ideal	1.000	1.016	0.983	0.946	0.989
Tornquist	1.000	1.016	0.983	0.946	0.988

American Medical Association [1990], <u>Current Procedural Terminology (CPT)</u>, 4th edition, Chicago: American Medical Association.

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

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

Armknecht, Paul A. [1996], "Improving the Efficiency of the U.S. CPI in the Future," Washington DC: International Monetary Fund, unpublished manuscript, March.

Armknecht, Paul A. and Daniel H. Ginsburg [1992], "Improvements in Measuring Price Changes in Consumer Services: Past, Present and Future," in Zvi Griliches, editor, assisted by Ernst R. Berndt, Timothy F. Bresnahan and Marilyn E. Manser, <u>Output Measurement in the Service Sectors</u>, Conference on Research in Income and Wealth, Studies in Income and Wealth, Vol. 56, Chicago: University of Chicago Press for the National Bureau of Economic Research, 109-156.

Armknecht, Paul A., Brent R. Moulton and Kenneth J. Stewart [1994], "Improvements to the Food at Home, Shelter and Prescription Drug Indexes in the U.S. Consumer Price Index," US Department of Labor, Bureau of Labor Statistics, CPI Announcement-Version I, October 20.

Beach, S.R.H., E.E. Sandeen and K.D. O'Leary [1990], <u>Depression in Marriage</u>, New York: Guilford Press.

Beck, A.T., A. John Rush, B.F. Shaw and G. Emergy [1979], <u>Cognitive Therapy of Depression</u>, New York: Guilford Press.

Berndt, Ernst R. [1991], <u>The Practice of Econometrics: Classic and Contemporary</u>, Reading, MA: Addison Wesley.

Berndt, Ernst R., David Cutler, Richard G. Frank, Zvi Griliches, Joseph Newhouse and Jack E. Triplett [1998], "A Review of Issues in Measuring Medical Care Prices", paper prepared for the NBER/CRIW Conference, Medical Care Output and Productivity, Bethesda, MD, June 12-13.

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

Berndt, Ernst R., Richard G. Frank and Thomas G. McGuire [1997], "Alternative Insurance

Arrangements and the Treatment of Depression: What Are The Facts?," <u>American Journal of Managed Care</u>, 3:2, February, 243-250.

Berndt, Ernst R. and Zvi Griliches [1993], "Price Indexes for Microcomputers: An Exploratory Study," Chapter 2 in Murray F. Foss, Marilyn E. Manser and Allan H. Young, eds., <u>Price Measurements and their Uses</u>, Studies in Income and Wealth, Vol. 57, Chicago: University of Chicago Press for the National Bureau of Economic Research, 63-93.

Berndt, Ernst R., Zvi Griliches and Neal Rappaport [1995], "Econometric Estimates of Price Indexes for Personal Computers in the 1990's," <u>Journal of Econometrics</u>, Vol. 68, No. 1, July 1995, pp. 243-268.

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," <u>Journal of the American Medical Association</u>, 264:2524-2528.

Busch, Susan H., Richard G. Frank and Ernst R. Berndt [1996], "Effectiveness, Efficacy and Price Indexes for Depression: A Review of the Literature," unpublished manuscript, Boston: Harvard Medical School, Department of Health Care Policy, December.

Cardenas, Elaine M. [1996a], "The CPI for Hospital Services: Concepts and Procedures," <u>Monthly Labor Review</u>, 119:7, July, 34-42.

Cardenas, Elaine M. [1996b], "Revision of the CPI Hospital Services Component," <u>Monthly Labor Review</u>, 119:12, December, 40-48.

Catron, Brian and Bonnie Murphy [1996], "Hospital Price Inflation: What Does the New PPI Tell Us?", <u>Monthly Labor Review</u>, 119:7, July, 124-31.

Croghan, Thomas W., Catherine A. Melfi, Deobrah G. Dobrez and Thomas J. Kniesner [1998], "Effect of Mental Health Specialty Care on Antidepressant Length of Therapy," <u>Medical Care</u>, in press.

Crown, William E., Robert L. Obenchain, Luella Englehart, Tamra Lair, Don P. Buesching and Thomas W. Croghan [1996], "Application of Sample Selection Models to Outcomes Research: The Case of Evaluating Effects of Antidepressant Therapy on Resource Utilization," unpublished manuscript, Cambridge, MA: The MEDSTAT Group, January.

Depression Guideline Panel [1993], <u>Depression in Primary Care</u>, 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., <u>Essays in the Theory and Measurement of Consumer Behavior in Honour of Sir Richard Stone</u>, Cambridge: Cambridge University Press, 163-208.

Diewert, W. Erwin [1976], "Exact and Superlative Index Numbers," <u>Journal of Econometrics</u>, 4:2, 115-145.

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," <u>Archives of General Psychiatry</u>, 46, 971-982.

Ellis, Randall P., Gregory C. Pope, Lisa I. Iezzoni, John Z. Ayanian, David W. Bates, Helen Burstin and Arlene S. Ash [1996], "Diagnosis-Based Risk Adjustment for Medicare Capitation Payments," <u>Health Care Financing Review</u>, 17:3, Spring, 101-128.

Feldstein, Paul J. [1988], <u>Health Care Economics</u>, Third Edition, New York: Wiley.

Feldstein, Paul J. [1983], <u>Health Care Economics</u>, Second Edition, New York: Wiley.

Feldstein, Paul J. [1978], <u>Health Care Economics</u>, First Edition, New York: Wiley.

Fixler, Dennis and Mitch Ginsburg [1997], "Health Care Output and Prices in the Producer Price Index," slides from presentation to the NBER Summer Institute, Franco-American Seminar, Cambridge, MA, July 23.

Ford, Ina Kay and Daniel H. Ginsburg [1997], "Medical Care in the CPI," paper given in presentation to the NBER Summer Institute, Franco-American Seminar, Cambridge, MA, July 23.

Frank, Ellen, David J. Kupfer, J.M. Perel, Cleon Cornes, D.B. Jarrett, A.G. Mallinger, Michael E. Thase, A.B. McEachran, and V.J. Grochocinski [1990], "Three-year Outcomes for Maintenance Therapies in Recurrent Depression, <u>Archives of General Psychiatry</u>, 47, 1093-1099.

Frank, Richard G., Susan H. Busch and Ernst R. Berndt [1998], "Measuring Prices and Quantities of Treatment for Depression," <u>American Economic Review</u>, 88:2, May, 106-111.

Frank, Richard G., Ernst R. Berndt and Susan H. Busch [1998], "Price Indexes for the Treatment of Depression," in Jack E. Triplett, ed., <u>Measuring the Prices of Medical Treatments</u>, Washington DC: The Brookings Institution, forthcoming.

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

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

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

Gilbert, Milton [1962], "Quality Change and Index Numbers: A Reply," <u>Monthly Labor Review</u>, 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," <u>Health Affairs</u>, 17:2, March/April, 40-52.

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

Griliches, Zvi [1988], <u>Technology, Education, and Productivity</u>, New York: Basil Blackwell Inc.

Griliches, Zvi and Iain Cockburn [1994], "Generics and New Goods in Pharmaceutical Price Indexes," <u>American Economic Review</u>, 84:5, December, 1213-1232.

Kanoza, Douglas [1996], "Supplemental Sampling in the PPI Pharmaceuticals Index," <u>Producer Price Indexes Detailed Price Report</u>, January, 8-10.

Katon, Wayne, Michael von Korff, Elizabeth Lin, Terry Bush and Johann Ormel [1992], "Adequacy and Duration of Antidepressant Treatment in Primary Care," <u>Medical Care</u>, 30:67-76.

Keeler, Emmett B., Kenneth B. Wells, Willard G. Manning, J. David Rumpel and Janet M. Hanley [1986], "The Demand for Episodes of Mental Health Services," Santa Monica, CA: RAND Report R-3432-NIMH, October.

Keller, Martin B., Gerald L. Klerman, Philip W. Lavori, William Coryell, Jean Endicott and John Taylor [1984], "Long-term Outcome of Episodes of Major Depression," <u>Journal of the American Medical Association</u>, 252:6, August 10, 788-792.

Keller, Martin B., Philip W. Lavori, John Rice, William Coryell and Robert M.A. Hirschfeld [1986], "The Persistent Risk of Chronicity in Recurrent Episodes of Nonbipolar Major Depressive Disorder: A Prospective Follow-up," <u>American Journal of Psychiatry</u>, 143:1, January, 24-28.

Keller, Martin B., Philip W. Lavori, Gerald L. Klerman, Nancy C. Andreasen, Jean Endicott, William Coryell, Jan Fawcett, John P. Rice and M. A. Robert [1986], "Low Levels and Lack of Predictors of Somatotherapy and Psychotherapy Received by Depressed Patients," <u>Archives of General Psychiatry</u>, 43:5, 458-466.

Kelly, Gregory G. [1997], "Improving the PPI Samples for Prescription Pharmaceuticals," <u>Monthly Labor Review</u>, 120:10, October, 10-17.

Kendler, K.S., E.E. Walters, M.C. Neale, Ronald C. Kessler, A.C. Heath and L. J. Eaves [1995], "The Structure of Genetic and Environmental Risk Factors for Six Major Psychiatric Disorders in Women: Phobia, Generalized Anxiety Disorder, Panic Disorder, Bulimia, Major Depression, and Alcoholism," Archives of General Psychiatry 52. (need to complete)

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, 51:1, 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.

Kupfer, David J., Ellen Frank, J.M. Perel, C. Cornes, A.G. Mallinger, Michael E. Thase, A.B. McEachran and V.J. Grochocinski [1992], "Five-year Outcomes for Maintenance Therapies in Recurrent Depression," Archives of General Psychiatry, 49, 769-773.

Kupfer, David J. [1991], "Long-Term Treatment of Depression," Journal of Clinical Psychiatry, 52:5 (supplement), 28-34.

Langford, Elizabeth A. [1957], "Medical Care in the CPI: 1935-1956," Monthly Labor Review, 80:9, September, 1053-1058.

Levit, Katharine R., Helen C. Lazenby, Brandley R. Braden and the National Health Accounts Team [1998], "National Health Spending Trends in 1996," Health Affairs, 17:1, January/February, 35-51.

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.

Marion Merrell Dow [annual], Managed Care Digest: HMO Edition, Kansas City, MO, 1993, 1994 and 1996.

McGuire, Thomas G. [1989], "Financing and Reimbursement for Mental Health Services," in Carl A. Taube, David Mechanic and Ann A. Hohmann, eds., The Future of Mental Health Services Research, Rockville, MD: U.S. Department of Health and Human Services, 87-112.

Mechanic, David [1989], "The Evolution of Mental Health Services and Mental Health Services Research," in Carl A. Taube, David Mechanic and Ann A. Hohmann, eds., The Future of Mental Health Services Research, Rockville, MD: U.S. Department of Health and Human Services, 1-8.

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

Newhouse, Joseph P. and the Insurance Experiment Group [1993], Free for All? Lessons from the RAND Health Insurance Experiment, Cambridge, MA: Harvard University Press.

Papolos, Demitri F. and Janic Papolos [1992], Overcoming Depression, Revised Edition, New York: Harper Collins Publishers.

Prospective Payment Assessment Commission [1995] -- need to get reference

Schulberg, Herbert C., Marjorie Saul, Maureen McClelland, Mary Ganguli, Wallace Cristy and Ellen Frank [1985], "Assessing Depression in Primary Medical and Psychiatric Practices," Archives of General Psychiatry, 42:12, 1164-1170.

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

Trajtenberg, Manuel [1990], Economic Analysis of Product Innovation: The Case of CT Scanners, Cambridge, MA: Harvard University Press.

Triplett, Jack E. [1997a], SIC Code Foundation Paper -- need to get reference

U.S. Department of Health and Human Services [1980], International Classification of Diseases (ICD-9-CM), second edition, Washington, DC: U.S. Government Printing Office (PHS)-80-1260, September.

U.S. Department of Labor, Bureau of Labor Statistics [1992], Handbook of Methods, Bulletin 2414, Washington, DC: U.S. Government Printing Office, April.

U.S. Department of Labor, Bureau of Labor Statistics [1995], "Improvements to CPI Procedures: Prescription Drugs," Washington DC.

U.S. Department of Labor, Bureau of Labor Statistics [not dated], "A Description of the PPI Hospital Services Initiative," Washington DC.

U.S. Senate Finance Committee [1996], Final Report from the Advisory Commission To Study The Consumer Price Index, Washington DC: December 4, updated version.

Weiner, Jonathan P., Allen Dobson, Stephanie L. Maxwell, Kevin Coleman, Barbara Starfield and Gerald Anderson [1996], "Risk-Adjusted Medicare Capitation Rates Using Ambulatory and Inpatient Diagnoses," Health Care Financing Review, 17:3, Spring, 77-100.

Wells, Kenneth B., Wayne Katon, William Rogers and Patti Camp [1994], "Use of Minor Tranquilizers and Antidepressant Medications by Depressed Outpatients: Results from the Medical Outcomes Study," American Journal of Psychiatry, 151:5, 694-700.

Wells, Kenneth B., Anita Stewart, R.D. Hays, William Rogers and K. Spritzer [1995], "Functioning and Well-being Outcomes of Patients with Depression Compared with Chronic General Medical Illness," Archives of General Psychiatry, 52:11-19.

Wells, Kenneth B., Roland Sturm, Cathy D. Sherbourne and Lisa S. Meredith [1996], Caring for Depression: A RAND Study, Cambridge, MA: Harvard University Press.

Wilde, Michelle I. and Paul Benfield [1998], "Fluoxetine: A Pharmacoeconomic Review of its Use in Depression," Pharmacoeconomics, 13:5 (1), May, 543-561.

Wingert, Terence D., John E. Kralewski, Tammie J. Lindquist and David J. Knutson [1995], "Constructing Episodes of Care from Encounter and Claims Data: Some Methodological Issues," Inquiry, Vol. 32, Winter, pp. 430-443.

ENDNOTES

---

[1] A more detailed discussion of the medical CPI and PPI data construction procedures is given in Berndt, Cutler, Frank et al. [1998].

[2] U.S. Department of Labor, Bureau of Labor Statistics [1992], p. 140.

[3] See Triplett [1997a] for economic issues involved in defining and aggregating industries within the SIC system.

[4] U.S. Department of Labor, Bureau of Labor Statistics [1992], p. 141.

[5] Catron-Murphy [1996], Table A-2, p. 31.

[6] For a recent list of DRGs, see Prospective Payment Assessment Commission [1995], Appendix E. Note also that some private insurers and Medicaid programs make use of DRGs for purposes of hospital payment.

[7] For a discussion of CPT4, see American Medical Association [1990].

[8] ICD-9 codes are discussed and listed in U.S. Department of Health and Human Services [1980]. The ICD-9 has recently been updated to version 10.

[9] Weiner et al. [1996].

[10] Ellis et al. [1996].

[11] For further discussion, see Fixler and Ginsburg [1997].

[12] For further details, see Catron-Murphy [1996], Fixler-Ginsburg [1997] and U.S. Department of Labor, Bureau of Labor Statistics [not dated].

[13] For further discussion, see Berndt, Cockburn and Griliches [1996], Griliches-Cockburn [1994], Kanoza [1996] and Kelly [1997].

[14] U.S. Department of Labor, Bureau of Labor Statistics [1992], p. 176.

[15] For historical discussions, see Langford [1957] and Getzen [1992].

[16] Levit, Lazenby, Braden et al. [1998], and Ford-Ginsburg [1997].

[17] See, for example, Armknecht and Ginsburg [1992], and Cardenas [1996a,b].

[18] Cardenas [1996], p. 40.

[19] Cardenas [1996], p. 40.

[20] For a discussion of the pricing of prescription pharmaceuticals in the MCPI, and treatment of generic drugs, see Armknecht, Moulton and Stewart [1994], and U.S. Department of Labor, Bureau of Labor Statistics [1995].

[21] Scitovsky [1964]; also see Scitovsky [1967].

[22] Armknecht [1996], p. 33.

[23] U.S. Senate Finance Committee [1996], p. 60.

[24] See, for example, the discussion between Gilbert [1961,1962] and Griliches [1962].

[25] American Psychiatric Association [1994], p. 161.

[26] American Psychiatric Association [1993].

[27] See Keller, Klerman, Lavori et al. [1984], and Keller, Lavori and Rice et al. [1986].

[28] American Psychiatric Association [1993].

[29] Papolos and Papolos [1992], p. 7.

[30] See, for example, Broadbent, Blazer, George and Tse [1990]; Wells, Stewart, Hays et al. [1989]; Hays, Wells, Sherbourne et al. [1995].

[31] Kessler, McGonagle, Zhao et al. [1994], p. 12.

[32] Kendler, Walters, Neale et al. [1995].

[33] Kessler, McGonagle, Swartz et al. [1993], p. 91.

[34]See Beck et al. [1979], Elkin et al. [1989], Frank et al. [1990], Kupfer et al. [1992] and Beach et al. [1990].

[35]Recent variations slightly distinct from the SSRIs include the brand name drugs Effexor, Serzone and Remeron.

[36]Considerable variation exists in side effects among the TCAs; see Berndt, Cockburn and Griliches [1996], Table 1, pp. 142-3 for details.

[37]Using MEDSTAT retrospective claims data, Croghan et al. [1998] provide evidence suggesting that the combination of SSRIs and psychotherapy is more effective than psychotherapy alone in patients' receiving continuous treatment. For a discussion of related SSRI-psychotherapy research, see Wilde-Benfield [1998].

[38]A significant portion of inpatient episodes have unspecified outpatient follow-up, thereby limiting that avenue for identifying treatments.

[39]For a discussion, see Mechanic [1989] and McGuire [1989].

[40]See Busch, Frank and Berndt [1996].

[41]The AHCPR Depression Guidelines Panel [1993] provides a summary and interpretation of the evidence on this point.

[42]See American Psychiatric Association [1993].

[43]We recognize this is only an approximation. This is particularly the case for depression, where the constellation of side effects across treatment can vary significantly, and can lead to differential patient compliance and patient preferences between the SSRIs and TCAs. See Crown, Obenchain, Englehart et al. [1996] and Wilde-Benfield [1998] for evidence on differential TCA-SSRI compliance among patients.

[44]Specialty carve-out management occurs when a portion of the health risk is separately managed from the rest of health care. See Frank, McGuire and Newhouse [1995] for further discussion.

[45]For discussion, see Goldman, McCulloch and Sturm [1998], and Ma and McGuire [1998].

[46]Thus we exclude other 296 ICD-9 diagnoses, depression not elsewhere classified, as well as some other broad depression-related conditions such as neurotic depression.

[47]For discussion of defining episodes of care, see Keeler et al. [1986] and Wingert et al. [1995].

[48]We count days without treatment only after the number of days of supply in a drug prescription has been exhausted, thereby assuming full compliance with the daily recommended dosage.

[49]These were brand name drugs Effexor and Serzone.

[50]These compositional figures are consistent with IMS aggregate national sales data for antidepressant medication over this time period, as reported by Berndt, Cockburn and Griliches [1996].

[51]See Frank, Berndt and Busch [1998] and Frank, Busch and Berndt [1998].

[52]Additional checks were done to ensure that these newly identified psychotherapy visits were not in fact instead a sequence of medical management visits. Robustness checks revealed that only a small number of claims were reclassified when cost thresholds of about $50 were used to mark medical management from psychotherapy.

[53]For a related discussion, see Katon et al. [1992].

[54]See Wells et al. [1994,1996] for further discussion.

[55]See Schulberg et al. [1985].

[56]Wells et al. [1994].

[57]Keller, Lavori, Klerman et al. [1986].

[58]Data from Marion Merrell Dow's <u>Managed Care Digest</u> [1993-4,1996] indicate that over all HMOs surveyed, average copayment per prescription rose from $6.78 in 1993 to $7.43 in 1995 for brand-name drugs, and from $4.75 to $5.22 for generics. While the MedStat increases appear to be larger, they also include deductibles for psychotherapy visits, not just copayments for drugs.

[59]See Frank, Berndt and Busch [1998] and Frank, Busch and Berndt [1998].

[60]See, for example, W. Erwin Diewert [1976,1981].

[61]For an introductory discussion to the hedonic method, see Griliches [1988], chapters 7 and 8, and Berndt [1991], chapter 4. Applications in the medical context include Trajtenberg [1990] and Berndt, Cockburn and Griliches [1996].

[62]The side effect data are discussed in detail in Berndt, Cockburn and Griliches [1996]. We sum up the 0 (rare) to 4 (common) ratings across the various side effects. Here we also assume that psychotherapy has no side effects.

[63]See Berndt and Griliches [1993] and Berndt, Griliches and Rappaport [1995].