Dear editor and reviewers:

Thank you for the opportunity to revise the manuscript. Changes made in response to reviewer comments are described below, in the order that the comments were made.

Reviewer: Pikhart

Discretionary revisions.

1. All independent variables were left in the model to produce adjusted results. This is because the large sample size provides sufficient power to test the hypotheses even when some of the relevant categories had small percentages of subjects.

Minor essential revisions.

1. Table 3 changed to Table 2.

Major compulsory revisions

2. The background section has been expanded to more fully address the use of a single item for measuring self-rated mental health. The following text has been added. "Epidemiological studies of mental health problems pose special measurement problems, because of the need to cost-effectively collect a broad set of measures that are brief yet valid. Measures of physical and mental health used in epidemiological studies have evolved in recent years. Long instruments are regarded as important for studies of patients in clinical settings, but impractical for community surveys. A single time has become the norm for measuring overall health in population studies. Measurement of mental health in population studies also has evolved from complex diagnostic instruments toward shorter scales. For example, a study of older people in Spain used a single item to measure health but 20 items to measure depression [9]. A recent study of medical inpatients used an eight-item symptom checklist to detect anxiety and depression and a seven-item index to mean hypochondriasis [10]. In contrast, the American Journal of Public Health published the results of a national survey that used a single item to measure mental health. The single item was dichotomized (positive versus negative mental health [11]. The Behavioral Risk Factor Surveillance Survey (BRFSS), which is required of every US state by the Centers of Disease Control, uses a single-item to measure (i.e., the number of recent days when mental health was poor). Single-item measures of mental health are valid because, rather than seeking to assign a clinical diagnoses such as depression, they simply reflect the respondents' perceptions of his or her own mental health. Perceived or self-rated mental health is inherently valid because the respondent is the best judge of his or her own perceptions."

3. The description of the insurance variables has been clarified.

4. Table 1 has been changed as suggested.

5. The discussion section has been expanded. The new text reads as follows. "The medical care
system has been chastised for its insensitivity to underlying mental health problems among patients who present with physical symptoms[12]. The problem is regarded as of sufficient magnitude that medical costs are thought to be higher because of failure to treat mental health problems. The logical conclusion is that treatment of mental health problems would lead to a reduction or 'offset' in the cost of medical care.[13] However, policy analysts dispute the existence an offset effect, calling it a myth.[14] Regardless of whether medical costs would actually decline if mental health problems were to be adequately addressed, few could argue that high quality primary care would recognize and treat these issues. Nevertheless, quality improvement efforts directed at improving the quality of services delivered to depressed or anxious patients in primary care settings have had mixed results [15]. Therefore, it is not surprising that this study reveals that poor mental health is associated with higher utilization of medical care.”

6. A paragraph mentioning the limitations of the study has been added.
7. The statements about anti-depression programs have been sponsored.

Reviewer: Alpass

Major compulsory revisions
1. Statements supporting the hypotheses have been added.
2. The term 'mental status' has been replaced with 'cognitive impairment'. We did not screen out subjects due to poor mental health.
3. About 25 percent of respondents had mental health that was poor all of the time, most of the time, a good bit of the time or some of the time. About 22 percent were blue a little bit of the time. This variable was dichotomized in order preserve consistency with the approach generally used to analyze self-rated overall health. The split was made between 'none of the time' and 'some of the time' because of the natural break in the data.
4. Satisfaction usually is dichotomized because of a bias toward favorable responses which skews distributions to the right.
5. An additional paragraph expanding the results section as been added.