Did the Balanced Budget Act of 1997 Affect the Quality of Medicare Home Health Services?

Final Report

May 15, 2002

Valerie Cheh
William Black

Submitted to:
Laguna Research Associates
1604 Union Street
San Francisco, CA  94123

Project Officer:
Nelda McCall

Submitted by:
Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ  08543-2393
(609) 799-3535

Project Director:
Valerie Cheh

This report was funded by the Robert Wood Johnson Foundation under project number 037040.
CONTENTS

Chapter | Page
--- | ---
EXECUTIVE SUMMARY | ix

I | INTRODUCTION | 1

II | DATA AND METHODS | 5
A. OVERVIEW OF DESIGN | 5
B. DATA | 7
1. PROQUAL and OASIS | 7
2. Survey Data | 12
C. CONTROL VARIABLES | 18
D. METHODS | 24
1. Logistic Regression Model | 24
2. Issues in Weighting Observations | 26
3. Statistical Hypothesis Testing | 27
4. Sensitivity Tests | 29

III | RESULTS | 31
A. POTENTIAL EFFECTS OF IPS ON PATIENT OUTCOMES | 31
B. EFFECT OF IPS ON FUNCTIONAL STATUS | 34
C. EFFECT OF IPS ON MEDICAL SYMPTOMS AND PATIENT OUTCOMES | 43
D. EFFECT OF IPS ON OVERALL HEALTH STATUS | 46
E. SENSITIVITY TESTS | 50
## CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV CONCLUSIONS</td>
<td>61</td>
</tr>
<tr>
<td>A. LIMITATIONS</td>
<td>62</td>
</tr>
<tr>
<td>B. POLICY RECOMMENDATIONS</td>
<td>64</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>67</td>
</tr>
<tr>
<td>APPENDIX A: DETAILED RESULTS FOR SENSITIVITY ANALYSES</td>
<td>A.1</td>
</tr>
</tbody>
</table>
# TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>II.1</td>
<td>OUTCOME VARIABLES FROM THE OASIS DATA</td>
<td>10</td>
</tr>
<tr>
<td>II.2</td>
<td>CONSTRUCTION OF OASIS AND SURVEY ANALYSIS FILES</td>
<td>13</td>
</tr>
<tr>
<td>II.3</td>
<td>OUTCOME VARIABLES FROM THE SURVEY DATA</td>
<td>17</td>
</tr>
<tr>
<td>II.4</td>
<td>WEIGHTED PRE- AND POST-IPS PERIOD MEANS FOR PATIENT-LEVEL INDEPENDENT VARIABLES USED IN THE ANALYSIS OF OASIS AND SURVEY SAMPLES</td>
<td>19</td>
</tr>
<tr>
<td>II.5</td>
<td>WEIGHTED PRE- AND POST-IPS PERIOD MEANS FROM OASIS FOR PATIENT-LEVEL INDEPENDENT VARIABLES MEASURING HEALTH AND FUNCTIONAL STATUS</td>
<td>23</td>
</tr>
<tr>
<td>III.1</td>
<td>ESTIMATED DIFFERENCES IN BASIC ACTIVITIES OF DAILY LIVING BETWEEN PRE- AND POST-IPS PERIODS, FROM OASIS DATA</td>
<td>35</td>
</tr>
<tr>
<td>III.2</td>
<td>ESTIMATED DIFFERENCES IN INSTRUMENTAL ACTIVITIES OF DAILY LIVING BETWEEN PRE- AND POST-IPS PERIODS, FROM OASIS DATA</td>
<td>37</td>
</tr>
<tr>
<td>III.3</td>
<td>ESTIMATED DIFFERENCES IN FUNCTIONAL STATUS BETWEEN PRE- AND POST-IPS PERIODS, FROM SURVEY DATA</td>
<td>38</td>
</tr>
<tr>
<td>III.4</td>
<td>SENSITIVITY OF SURVEY RESULTS TO ADDITION OF OASIS CONTROL VARIABLES</td>
<td>41</td>
</tr>
<tr>
<td>III.5</td>
<td>ESTIMATED DIFFERENCES IN MEDICAL SYMPTOMS AND OUTCOMES BETWEEN PRE- AND POST-IPS PERIODS, FROM OASIS DATA</td>
<td>44</td>
</tr>
<tr>
<td>III.6</td>
<td>ESTIMATED DIFFERENCES BETWEEN PATIENTS’ SELF-REPORTS OF GENERAL HEALTH DURING THE PRE- AND POST-IPS PERIODS, FROM SURVEY DATA</td>
<td>48</td>
</tr>
<tr>
<td>III.7</td>
<td>SENSITIVITY OF SURVEY RESULTS TO ADDITION OF OASIS CONTROL VARIABLES</td>
<td>49</td>
</tr>
<tr>
<td>III.8</td>
<td>SENSITIVITY OF IPS IMPACTS ON ADL/IADL FUNCTIONAL LIMITATIONS, USING OASIS AND SURVEY DATA</td>
<td>53</td>
</tr>
<tr>
<td>III.9</td>
<td>SENSITIVITY OF IPS IMPACTS ON MEDICAL SYMPTOMS AND OUTCOMES, USING OASIS DATA</td>
<td>55</td>
</tr>
<tr>
<td>III.10</td>
<td>SENSITIVITY OF IPS IMPACTS ON PATIENTS’ SELF-REPORTS OF GENERAL HEALTH</td>
<td>58</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>III.1</td>
<td>CURVES OF HYPOTHESESIZED RELATIONSHIP BETWEEN PATIENT OUTCOME AND AMOUNT OF HOME HEALTH CARE SERVICE</td>
<td>32</td>
</tr>
</tbody>
</table>
EXECUTIVE SUMMARY

The Balanced Budget Amendment of 1997 dramatically altered the Medicare home health program. The most controversial requirement of this legislation was the imposition of financial limits on participating agencies. These limits, known as the Interim Payment System (IPS), led to both substantial decreases in the amount of home health care provided and substantial savings to the Medicare program. The important question is whether these savings resulted in poorer health outcomes for Medicare beneficiaries.

This study examines the effect of IPS using a pre-post design that takes advantage of a special circumstance. At the time that BBA 1997 was implemented, the Centers for Medicare & Medicaid Services was conducting a three-year demonstration project to test the effects of home health prospective payment. Detailed patient data were collected from the demonstration agencies during the pre-IPS period that were unavailable from other agencies. The availability of these data enables us to compare the health outcomes of Medicare home health beneficiaries in agencies operating under the cost reimbursement system with the health outcomes of patients of the same agencies under IPS.\(^1\)

Two main data sources were used for the analysis. The first was an administrative data set collected to monitor patient outcomes: PROQUAL data in the pre-IPS period and OASIS in the post-IPS period. These data are collected at the beginning and end of a home health episode, so change over the course of care can be measured. We also use survey data, which measured patient outcomes at a fixed point in time (120 days) after admission.

We used regression analysis to control for patient characteristics that changed between the pre-IPS and post-IPS period. We conducted numerous sensitivity tests, which included varying the weighting methodology and the sample composition.

FINDINGS

We did not find evidence of serious health consequences associated with IPS. The functional status of post-IPS period patients, despite their having more functional limitations when their episode began, was as likely to improve during their episode of care as was the functional status of pre-IPS period patients. Moreover, the condition of post-IPS period patients was slightly more likely than that of pre-IPS period patients to stabilize (that is, not worsen) (See Table 1). Furthermore, patients were as likely to improve on three measures of instrumental activities of daily living (meal preparation, housekeeping, and medication management) during the post-IPS period as during the pre-IPS period. We did find some evidence from the

---

\(^1\)As we will explain later, some agencies refused to participate in follow-up data collection activities or went out-of-business by the end of the period, so replacement agencies had to be selected.
TABLE 1

SUMMARY OF QUALITY OF CARE MEASURES

<table>
<thead>
<tr>
<th>Type of Measure</th>
<th>Number of Measures</th>
<th>Number of Significant Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>FAVORING PRE-IPS PERIOD</td>
</tr>
<tr>
<td>Improvements in Activities of Daily Living</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization in Activities of Daily Living</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Improvements in Instrumental Activities of Daily</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Living</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stabilization in Instrumental Activities in Daily</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Living</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Status at 120 Days(^a)</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Improvements in Clinical Symptoms</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Stabilization in Clinical Symptoms</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Overall Health and Satisfaction</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

\(^a\)Controlling for full set of baseline characteristics.

survey analysis that IPS may have negatively affected patient’s self-reported functioning (particularly, in ambulating), and this finding is cause for caution. Nevertheless, most of the evidence suggests that functioning was not adversely affected.

Measures of health status and symptoms also suggest that the IPS had little effect on patients’ health outcomes. The status of post-IPS period patients’ surgical wounds was significantly less likely to improve during the episode of care than was that of pre-IPS period patients’ surgical wounds, but the fact that we found no difference in improvement in the status of the most problematic pressure ulcer or number of pressure ulcers between the two periods leads us to believe that post-IPS period patients’ shorter time in home health care, rather than poor treatment of wounds, explains this finding. Furthermore, our measures of six other health symptoms consistently showed that patients who were in care after IPS was implemented were more likely than patients in care before IPS to stabilize during their home health episode (that is, were less likely to worsen).

Three measures of overall health status suggest that IPS could have detrimental effects, but the evidence is weak. Although patients in the post-IPS period were as likely as patients in the pre-IPS period to report good overall health status, they were more likely to report having spent time in bed due to illness, and to feel less satisfied with life. The differences in the baseline patient characteristics that we measured did not account for these differences in patients’ self-
reports. However, because IPS did not negatively affect our measures of health outcomes that are more directly related to home health service use, we are inclined to believe that the finding is the result of the changing nature of Medicare home health patients for which we were unable to control.

LIMITATIONS

This study has a number of limitations. Perhaps the most important limitation is the inherent weakness of a pre-post analysis. In this type of analysis, it is not possible to control for every factor that might have changed over time, and that could have affected health outcomes, so we always must be cautious in drawing conclusions. In our analysis, most of the impacts we estimated suggested that health outcomes improved, and we therefore conclude with confidence that patients were not harmed by IPS. Nevertheless, we hesitate to state that the payment limits actually led to better outcomes, as it is possible that other factors for which we were unable to control may have driven the results. But we do not believe these factors would have had such a large effect as to reverse the results entirely.

The study also is limited by the fact that the list of outcomes we measured was not exhaustive. We may have failed to measure outcomes that are more sensitive to decreases in the number of home health visits than were the ones we did measure. Even so, however, those outcomes probably affect only a small percentage of the Medicare home health population.

A third limitation relates to the small number of agencies that originally volunteered for the Medicare Home Health Prospective Payment (HHPP) Demonstration, and that ultimately agreed to participate in the post-IPS study: these agencies may not be representative of agencies nationwide. It is possible that only agencies that were confident about their quality of care agreed to participate in the study. If this type of self-selection of agencies occurred, then we might have failed to observe adverse outcomes in other, less confident agencies.

Fourth, it is possible that the nurses’ assessments reported in the OASIS data were biased. During the pre-IPS period, agencies had no incentive to complete the OASIS instrument in any particular way. However, if agencies were concerned that their early discharges during the post-IPS period would have resulted in scrutiny of their services, they may have instructed their nurses to ensure that the discharge forms reflected the best possible patient outcomes. To the extent that the nurses were able to “upcode” the status of patients without being untruthful, we might have failed to identify negative consequences of IPS.

Finally, missing data may have affected our results. We know that 30 percent of the cases in the OASIS data set had start-of-care forms that did not match any follow-up forms. Furthermore, we expect that the data failed to include some cases because the cases had no start-of-care form; we do not have an estimate of how often that failure occurred. Although evidence from the HHPP Demonstration suggests the missing data are not correlated with health outcomes, we cannot rule out the possibility that the occurrence of missing observations was in fact correlated in that way. Analysis without the missing data therefore could have biased our results.
POLICY RECOMMENDATIONS

The evidence in this report shows that IPS did not seriously harm patients. Despite the large reductions in home health services use, we found little evidence that patient functioning or health outcomes suffered. Although we do not want to dismiss the possibility that IPS may have had some adverse consequences, it appears that agencies, when given the financial incentive to reduce services, do so safely, without compromising patient care.

The evidence also shows that the type of patient who received care during the post-IPS period differed significantly from the type receiving care during the pre-IPS period. Indeed, the changes we observed suggest that, although patients were functionally more limited at the start of care during the post-IPS period than during the pre-IPS period, they also were less likely to suffer from conditions historically associated with higher home health resource use. This finding suggests that the Centers for Medicare & Medicaid Services (CMS) should monitor access to home health services carefully to ensure that high-use patients are admitted to care. The prospective payment system (PPS) currently in use has mitigated this problem; unlike IPS, PPS does account for differences in the severity of patients’ conditions.

Finally, because CMS uses OASIS data to monitor the quality of care, more resources are necessary to monitor the data reports. Agencies have difficulty implementing the OASIS system, and we found that nearly one-third of the data were missing. Furthermore, emergency service use was greatly underreported. A more thorough system of checks and balances—especially one that is tied to the payment system—may help agencies comply with the data collection while assuring CMS that the lowest-quality cases are not systematically unreported.

---

2 For example, relative to pre-IPS patients, post-IPS period patients in the OASIS sample had fewer incidences of diabetes, cerebrovascular disease, depression and risk factors (Table II.4), all of which are associated with high use of home health resources under cost reimbursement.

3 Note that these data were from the beginning of the OASIS data collection. The problems may have resolved as agencies became familiar with the system.
I. INTRODUCTION

To reverse the rapid growth in Medicare home health care expenditures, Congress legislated a number of changes to the Medicare home health program as part of the Balanced Budget Act of 1997 (BBA 1997). The implementation of the Interim Payment System (IPS) for home health care was the most dramatic change. IPS explicitly limited the amount Medicare would pay per patient, which led to a substantial decrease in the use of home health services by Medicare beneficiaries. The number of Medicare home health users decreased 22 percent during the first year after IPS was implemented, and the number of visits declined 39 percent. These decreases resulted in a drop in Medicare home health payments of 52 percent (McCall et al. 2001).

The question remains, however, whether this decline in service use was detrimental to Medicare beneficiaries. Previous research suggested that reductions in Medicare home health use might worsen outcomes (Shaughnessy et al. 1994). The researchers compared patients in health maintenance organizations (HMOs) with patients who received traditional fee-for-service home health care (that is, home health care with no IPS payment restrictions). They found that HMO patients received approximately 32 percent fewer home health visits during a 60-day period and had substantially lower levels of patient functioning.\(^1\) By contrast, however, more recent research has shown that substantial decreases in home health use had no apparent effect on home health outcomes. Research for the Evaluation of the Medicare Home Health Prospective Payment Demonstration found that prospectively paid agencies reduced the number of Medicare home health visits by 17 percent over a 120-day period (relative to cost-reimbursed agencies),

\(^1\)The differences in utilization did not adjust for differences in patient characteristics.
but without adversely affecting patient outcomes (Chen 2000 and 2001).\(^2\) Thus, research has not determined with certainty whether the decreases in services use associated with IPS will affect the health outcomes of beneficiaries.

It is important to investigate this issue for another reason as well: the estimated decrease in the number of home health visits resulting from IPS appears to be much higher than initially thought. Home health agencies would be expected to respond initially to implementation of payment reform by trimming the least beneficial visits (that is, visits that beneficiaries may want, but that do not necessarily result in better health outcomes). However, given the large reductions in home health service use resulting from IPS, it is possible that agencies felt compelled to do more than simply cut “fat”—and that the additional reductions may have harmed beneficiaries.

This study examines the effect of IPS using a pre-post design that takes advantage of a special circumstance. At the time that BBA 1997 was implemented, the Centers for Medicare & Medicaid Services was conducting a three-year demonstration project to test the effects of home health prospective payment. Detailed patient data were collected from the demonstration agencies during the pre-IPS period that were unavailable from other agencies. The availability of these data enables us to compare the health outcomes of Medicare home health beneficiaries in agencies operating under the cost reimbursement system with the health outcomes of patients of the same agencies under IPS.\(^3\)

We find little evidence that IPS had a detrimental effect on patient health outcomes. Although we identified a few measures of patient health status and functioning that were worse

\(^2\)These utilization differences were estimated with controls for case-mix differences.

\(^3\)As we will explain later, some agencies refused to participate in follow-up data collection activities or went out-of-business by the end of the period, so replacement agencies had to be selected.
during the post-IPS period relative to the pre-IPS period, the majority of the measures suggested that patient outcomes had remained unchanged or had actually improved. A higher proportion of patients maintained or improved their condition under IPS, which implies that the average patient outcome improved. That is, the condition of home health agencies’ patients was less likely to deteriorate during the post-IPS period than during the pre-IPS period, as it was equally likely to improve.
II. DATA AND METHODS

We estimated the impact of the Interim Payment System (IPS) on measures of patients’ functional status, medical symptoms and outcomes, and self-reports of general health status. To estimate these impacts, we used administrative data from OASIS and data from two waves of a telephone survey to compare health outcomes during the pre-IPS implementation period with those during the post-IPS implementation period. We used regression analysis to control for differences in patient characteristics, and, therefore, to isolate the impact of IPS.

We begin this chapter by describing the study design. We then discuss our data sources and the regression control variables. Finally, we describe our methodology and discuss how we estimated the impacts of IPS.

A. OVERVIEW OF DESIGN

The analysis discussed in this chapter compares the health outcomes of Medicare home health patients before agencies were subject to IPS with outcomes of Medicare home health patients after the agencies were subject to it. The original design planned to use as the analysis sample the 41 agencies in the Medicare Home Health Prospective Payment (HHPP) Demonstration control group that received cost reimbursement during the demonstration, and that had detailed data on their patients collected during the demonstration. Collecting the same data after IPS implementation enabled us to identify the impact of IPS on the health outcomes of the agencies’ patients.

Two key implementation problems weakened the original design. First, nine agencies went out of business toward the end of the demonstration. Furthermore, because of the strict payment
limits that the IPS imposed, agencies were reluctant to participate in the current study.\footnote{Agencies were not paid to participate in data collection efforts during the post-IPS period.} Under cost reimbursement, an agency under the cost limits could pass part of the cost of the data collection on to Medicare. By contrast, the Balanced Budget Act of 1997 (BBA 1997) imposed stricter per-visit cost limits as well as the per-patient reimbursement limit, and many agencies were unwilling to absorb the cost of participating in research. As a result, only 15 of the 41 agencies agreed to participate in the study. To compensate, we selected replacement agencies for the post-IPS period that had the same characteristics (state, urban/rural setting, for-profit/nonprofit status, hospital based/not hospital based, and agency size) as did the agencies that refused to participate. Financial considerations limited our ability to recruit replacement agencies, however, and, in the end, only 27 agencies agreed to participate in the post-IPS data collection efforts, including the 15 agencies from the HHPP Demonstration. Although we conducted our main analysis on all the agencies recruited for the study, we also conducted the analysis on the sample of 15 agencies that were in both the pre-IPS and post-IPS periods to determine whether using the replacement agencies affected the results.

The second implementation problem resulted from IPS restrictions that had the effect of reducing the number of patients admitted to Medicare home health care; evidence suggests that the patients who were admitted during the post-IPS period differed from those admitted during the pre-IPS period (McCall et al. 2001). Indeed, as we demonstrate in this chapter, we identified significant differences in our baseline patient characteristics between the pre-IPS and post-IPS periods. To mitigate this problem, we included as many patient characteristics as possible in our analysis to control for these differences; nevertheless, it is possible that we did not control for all the differences.
B. DATA

We used three sources of data to compare patient outcomes. The first two sources—PROQUAL and OASIS—are administrative data collected to monitor quality of care. The third source is a patient survey, and the survey instrument we used was the same one used in the HHPP Demonstration to collect patients’ self-reported health status measures.

1. PROQUAL and OASIS

The Centers for Medicare & Medicaid Services (CMS) requires home health agencies to collect and report OASIS data to monitor patient care. OASIS was designed and tested by the University of Colorado’s Center for Health Policy Research; the PROQUAL data instrument is a modified version of the OASIS data set. Most of the PROQUAL variables and completion instructions are the same as those in OASIS (Shaughnessey et al. 1995); however, we had to make some modifications to make the two data sets comparable. To simplify the discussion in this report, we refer to the combined PROQUAL and OASIS data as the OASIS data.

A record in our OASIS file consists of a pair of assessments of a patient’s health made by agency nurses. The nurses made the first assessment during the initial admission to home health care, recording the information on a start/resumption-of-care instrument. The nurses made the second assessment at whichever of the following events occurred first: (1) discharge, or (2) a reassessment 120 days after admission. During the pre-IPS period, patients were reassessed only once—120 days after admission. By contrast, during the post-IPS period, agencies completed patient reassessments every 60 days. Thus, to make the time periods between start-of-care and reassessment consistent in the two periods (that is, 120 days in length), we selected only assessments completed at 120 days since the start-of-care.

Although the OASIS file is a rich source of patient information, it has two major drawbacks for research purposes: (1) missing data, and (2) collection of data only at the time a patient is
discharged. Although home health agencies are required to provide OASIS data for 100 percent of their Medicare patients, they had difficulty meeting this requirement. The agencies’ data collection processes sometimes missed patients entirely, and sometimes missed one of the series of forms. During the pre-IPS period, the demonstration evaluator asked agencies why they had missed cases, and the agencies reported reasons that were unrelated to care quality (Cheh et al. 2001); nevertheless, one cannot rule out the potential bias that would result if the occurrence of missing data was correlated with care quality. However, the agencies did report that they had particular difficulty tracking and completing the “transfer to hospital” and “resumption of care after hospitalization” forms that are part of the OASIS data collection system. Because the demonstration quality assurance contractor monitored the agencies during the pre-IPS period to help identify and correct errors occurring while the forms were tracked and completed, and because this process did not occur during the post-IPS period, we believe that including data from these instruments would bias our results. Thus, we excluded data for these instruments from our analysis.

The second drawback is that agencies collect OASIS data when a patient is discharged, rather than at a particular point in time after admission. The average duration of a home health episode was substantially longer in the pre-IPS period (53 days) than in the post-IPS period (29 days), so patients in the post-IPS period generally received assessments earlier in their illness. To the extent that a patient recovers over time, this tendency toward earlier assessment in the post-IPS period could bias our comparison. To address this potential bias, we also analyzed survey data that were collected at the same point in time in both periods (that is, 120 days after admission).
a. **Outcomes Measured Using OASIS Data**

The OASIS outcome variables fall into two broad categories: (1) health measures, and (2) emergency services use. We had planned to analyze both types of variables, but we identified significant gaps in the emergency service use data that rendered such an examination useless. Table II.1 provides a summary list of the outcome variables that we analyzed.

We studied 17 health measures in the OASIS data, including basic activities of daily living (ADLs; for example, bathing), instrumental activities of daily living (IADLs; for example, light meal preparation), and medical symptoms (for example, pain interfering with activity). The nurses scored the items at each assessment, on an ordinal severity scale. For example, the item, “How often does pain interfere with the patient’s activity/movement,” has a four-level response: (1) none of the time, (2) some of the time, (3) most of the time, and (4) all of the time. Binary variables of “improvement” or “stabilization” in each item were calculated by comparing the scores from the admission and follow-up assessments. Improvement in a measure has a value of one if a patient’s score on the scale for that item improved on the second assessment, and zero otherwise. By definition, patients who started the home health episode at the best level of a measure were excluded from the sample for improvement in that measure, because they would not have been able to improve further. Stabilization in a measure has a value of one if a patient’s score on the scale for that measure did not worsen on followup (that is, the score remained the same or improved). Patients who started their episode at the worst level of a measure were excluded from the sample for stabilization in that measure, as their score could not worsen. Thus, the number of patients “eligible” for improvement or stabilization outcomes varied for each measure.
TABLE II.1
OUTCOME VARIABLES FROM THE OASIS DATA

<table>
<thead>
<tr>
<th>Improvement or Stabilization in:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
</tr>
<tr>
<td>Status of most problematic pressure ulcer</td>
</tr>
<tr>
<td>Status of most problematic surgical wound</td>
</tr>
<tr>
<td>Dyspnea</td>
</tr>
<tr>
<td>Urinary incontinence or catheter present</td>
</tr>
<tr>
<td>Confusion</td>
</tr>
<tr>
<td>Grooming</td>
</tr>
<tr>
<td>Toileting</td>
</tr>
<tr>
<td>Ambulating</td>
</tr>
<tr>
<td>Housekeeping</td>
</tr>
</tbody>
</table>
The variables in the second category—reported emergency health services use—were available on the basis of reports by the patient, family, or other providers, which agency nurses recorded on the OASIS follow-up/discharge instrument. The nurses recorded all emergency visits to hospital emergency rooms, physicians’ offices, outpatient clinics, and freestanding urgency care centers since the time of the last completed instrument.

We were unable to analyze the emergency services use data, however, due to significant underreporting. Data collected from OASIS during the HHPP Demonstration indicated that approximately 10 percent of the home health patients went to the emergency room during their episode of care; by contrast, claims data showed that 34 percent had an emergency room visit during the first 120 days after admission (Chen 2000; and Schore 2000). Although a difference in the time periods measured partly may explain this difference (that is, the shorter post-IPS period episodes could account for our observation of less emergency services use during that period), we believe the difference is too large to be due solely to this reason.

If the underreporting of emergency services use was the same during the pre- and post-IPS periods, then we could have used the data. However, we believe that the underreporting was greater in the post-IPS period data. During the pre-IPS period, the University of Colorado monitored the data that the agencies submitted. When it identified an inpatient admission from Medicare claims data that should have triggered the completion of a transfer-to-hospital form, the agency was asked to complete the form based on the nurses’ notes from the appropriate visit. Data from the HHPP Demonstration suggest that, during the first 120 days of an episode, 21 percent of the patients had an emergency room encounter leading to a hospital admission (Schore 2000). Thus, asking the agencies to complete transfer-to-hospital forms that they missed might have produced substantially higher reports of emergency services use. The lack of such
monitoring during the post-IPS period introduces potential bias, so we therefore chose not to analyze these measures.

b. OASIS Sample

The main analysis sample for the OASIS analysis came from 38 agencies in the pre-IPS HHPP Demonstration and the 27 agencies that agreed to be part of the post-IPS study. The University of Colorado’s Center for Health Policy Research collected and provided the pre-IPS data, but the analysis files were created for the HHPP Demonstration. During construction of the analysis files, episodes were excluded for various reasons (Chen 2000); we applied the same exclusion criteria to the post-IPS data received from CMS. The number of observations dropped from the file and the reason for which they were dropped are summarized in Table II.2. At the end of the exclusion process, we had 28,019 pre-IPS period observations that included episodes starting between May 1996 and August 1998, and 7,134 post-IPS period observations with episodes starting between March 2000 and August 2000.

2. Survey Data

A telephone survey was the final source of data for this study. The survey measured patients’ self-reported functioning approximately four months after admission to home health care. As noted, patients in the post-IPS period generally were discharged much earlier than were patients in the pre-IPS period. If a patient’s functioning improves over time regardless of the care provided, then the different episode lengths could bias our results. Measuring functioning at

---

2 The pre-IPS period OASIS sample contains 38 of the 41 agencies because the Center for Health Policy Research dropped 3 agencies during data cleaning.

3 The post-IPS period ended on December 31, 2000, but we included only those episodes that began on or before August 31, 2000, so that the 120-day episode could be completed before the end of the study.
### TABLE II.2

CONSTRUCTION OF OASIS AND SURVEY ANALYSIS FILES

<table>
<thead>
<tr>
<th></th>
<th>Pre-IPS</th>
<th>Post-IPS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OASIS Analysis File</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main File</td>
<td>53,738</td>
<td>22,139</td>
</tr>
<tr>
<td>Exclusions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No standard admission and discharge(^a)</td>
<td>8,173</td>
<td>9,680</td>
</tr>
<tr>
<td>Did not match to claims</td>
<td>8,412</td>
<td>3,165</td>
</tr>
<tr>
<td>Observation in agency with &lt;9 observations per period</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Episode length of ≤1 day</td>
<td>457</td>
<td>59</td>
</tr>
<tr>
<td>Follow-up visit &gt;135 days after admission(^b)</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>Had previous demonstration admission</td>
<td>5,886</td>
<td>1,165</td>
</tr>
<tr>
<td>In HMO during episode</td>
<td>1,777</td>
<td>705</td>
</tr>
<tr>
<td>Had Medicare as second payer during episode (for pre-IPS period) or at start of episode (for post-IPS period)(^c)</td>
<td>1,012</td>
<td>216</td>
</tr>
<tr>
<td><strong>Final OASIS Analysis File</strong></td>
<td>28,019</td>
<td>7,134</td>
</tr>
</tbody>
</table>

| **Survey Analysis File** |         |          |
| Main File                | 1,224   | 1,408    |
| Exclusions\(^d\)         |         |          |
| Did not match to claims  | 0       | 44       |
| Observation in agency with <9 observations per period\(^e\) | 8       | 11       |
| Had Medicare as second payer during episode (for pre-IPS period) or at start of episode (for post-IPS period)\(^f\) | 29      | 29       |
| **Final Survey Analysis File** | 1,187   | 1,324    |

**SOURCE:** OASIS and survey data.

\(^a\)Episodes that started with a return to home health care after hospitalization, or that ended with a discharge from home health care because of hospitalization. These patients were assessed at different points in their course of illness than were patients undergoing assessment because they had their first admission to home health care, were routinely discharged from home health care, or had reached the end of the 120-day episode.
b These observations were excluded because their follow-up data had been collected more than two weeks after the latest date by which they should have been collected. (Follow-up data should have been collected at discharge or 120 days after admission, whichever came first.)

c For the post-IPS period, only claims data at the start of care were available, so we dropped post-IPS period patients for whom Medicare was a second payer at the beginning of the episode. (That is, we were unable to drop post-IPS period patients for whom Medicare became a second payer at some point during the episode, as we did for the pre-IPS period.)

d We were unable to exclude survey observations for patients who were in an HMO during their episode, as we did in the OASIS sample. This variable was available only from OASIS, and we did not match it to the survey because of the match rate was poor.

e One pre-IPS period agency and two post-IPS period agencies had fewer than nine observations. See the text for additional discussion of the agency sample.

HMO = health maintenance organization.
a consistent point after discharge eliminates this potential bias. Roper ASW conducted the telephone survey on a sample of patients admitted to the home health agencies roughly four months after home health admission. The instrument used was the same one used in the pre-IPS period. The four-month period was chosen to correspond to the period in the previous survey. The pre-IPS period data were collected based on admission to the participating agencies from January to August 1997; the post-IPS period data were collected from March 2000 to June 2000.

Although the post-IPS survey attempted to replicate the pre-IPS survey as closely as possible, it differed in a few ways. First, we wanted to obtain a similar sample size, but fewer agencies participated in the post-IPS period survey; thus, the initial target sample size from each agency had to be set higher. Second, the pre-IPS period sample was drawn primarily from a sample frame constructed from Medicare claims. During the post-IPS period, however, claims were unavailable in time to use for a sample frame; thus, we relied on agency lists of new admissions.\(^4\) Third, due to greater concern about patient confidentiality in the post-IPS period, some agencies insisted on giving their patients the opportunity to decline to participate in the post-IPS survey prior to releasing contact information to Roper ASW. Furthermore, one agency required that the patients actually agree to participate prior to releasing contact information.

These changes, combined with increased agency fiscal austerity (which reduced staff availability to help track patients during the post-IPS period), contributed to substantially different survey response rates. The survey response rate was 65 percent during the post-IPS period (RoperASW 2001), in contrast to 90 percent during the pre-IPS period (Zambrowski and

---

\(^4\)In the demonstration, all agencies submitted claims to the demonstration’s fiscal intermediary, which was compensated to provide claims for the evaluation in a timely way. During the post-IPS period, the agencies reported to different fiscal intermediaries, none of which could be paid for this study and thus could not be asked to provide the data.
Cheh 2000). It is unlikely that this difference biased our results, as it occurred for reasons seemingly unrelated to patient outcomes. However, we cannot be certain of this, especially given that patients had an opportunity to self-select out of the sample.

a. Outcomes Measured Using Survey Data

We used the data on the patients’ self-reported health and functional status, collected during the telephone surveys, to construct binary variables. For example, we compared the likelihood of a patient’s reporting that he or she was able to get out of bed or out of a chair independently. We excluded cases from the analysis of a given outcome that had answers of “don’t know” or “refused.”

For items on ADLs and IADLs, we asked patients who reported that someone usually assisted them with an activity whether they could have performed the activity in the absence of help. If we had measured only whether a patient actually performed an activity, the comparison between the pre- and post-IPS periods might have been distorted—pre-IPS period patients were more likely than post-IPS period patients to have received services at four months since their care began (due to the longer average episode length in the post-IPS period), and patients receiving assistance might have had fewer opportunities to perform activities independently, thus appearing to be relatively more impaired. Table II.3 lists the survey outcome variables.

b. Survey Sample

The main analysis sample consists of 40 agencies and 1,187 observations from the pre-IPS period and 25 agencies and 1,324 observations from the post-IPS period. The pre- and post-IPS period samples originally consisted of 41 and 27 agencies, respectively, but we dropped 1 agency from the pre-IPS period and 2 from the post-IPS period because they had fewer than 9 observations per period. Table II.2 describes how we excluded certain observations from the original data file.
**TABLE II.3**

OUTCOME VARIABLES FROM THE SURVEY DATA

<table>
<thead>
<tr>
<th>Reported Health Good or Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any Days in Bed Within Past Two Weeks</td>
</tr>
<tr>
<td>Satisfied with Life</td>
</tr>
<tr>
<td>Did Bathe Independently</td>
</tr>
<tr>
<td>Could Bathe Independently</td>
</tr>
<tr>
<td>Did Eat Independently</td>
</tr>
<tr>
<td>Could Eat Independently</td>
</tr>
<tr>
<td>Did Transfer Independently</td>
</tr>
<tr>
<td>Could Transfer Independently</td>
</tr>
<tr>
<td>Did Ambulate Independently</td>
</tr>
<tr>
<td>Could Ambulate Independently</td>
</tr>
</tbody>
</table>
C. CONTROL VARIABLES

The weakness of a pre-post design is that the health outcomes may have changed between the pre- and post-IPS periods for reasons other than IPS. As noted, evidence suggests that the types of patients admitted to Medicare home health care changed after IPS was implemented (McCall et al. 2001). If these time trends are correlated with health outcomes, we may erroneously attribute differences to IPS that actually were due to changes in other factors over time. It is therefore important to use control variables to account for patient characteristics that could have changed over time.

We drew our patient-level control variables from three data sources: (1) Medicare administrative data, (2) OASIS start-of-care instruments, and (3) the telephone survey. For the analysis of OASIS outcomes, we used control variables constructed from Medicare claims and OASIS; for the analysis of survey outcomes, we used control variables constructed from Medicare claims and the survey. Ideally, we would have used OASIS as a source of control variables for the survey as well, but we lost 29 percent of the sample due to poor match rates. Thus, we used the OASIS control variables only for supplemental survey analyses.

We identified a number of significant differences in the control variables that are consistent with the hypothesis that home health agencies were admitting patients with different characteristics during the pre- and post-IPS periods. Although not all the differences are consistent with the notion that the post-IPS patients were “sicker,” some measures are consistent with the notion that agencies in the post-IPS period were admitting patients who were historically less costly. A comparison of the means of the control variables (Table II.4) and additional indicators of health and functional status (Table II.5) suggest that patients in the post-IPS period had more of some functional limitations than did patients in the pre-IPS period,
### TABLE II.4

**WEIGHTED PRE- AND POST-IPS PERIOD MEANS FOR PATIENT-LEVEL INDEPENDENT VARIABLES USED IN THE ANALYSIS OF OASIS AND SURVEY SAMPLES**

<table>
<thead>
<tr>
<th>Demographic Measures</th>
<th>OASIS</th>
<th>Four-Month Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-IPS</td>
<td>Post-IPS</td>
</tr>
<tr>
<td>Original Reason for Medicare: Reached Age 65 Years (EDB)</td>
<td>82.76</td>
<td>83.12</td>
</tr>
<tr>
<td>Age (Years; EDB)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65 Years</td>
<td>8.91</td>
<td>8.25</td>
</tr>
<tr>
<td>65 to 74 Years</td>
<td>29.11</td>
<td>26.65**</td>
</tr>
<tr>
<td>75 to 84 Years</td>
<td>40.32</td>
<td>39.51</td>
</tr>
<tr>
<td>≥85 Years</td>
<td>21.66</td>
<td>25.59***</td>
</tr>
<tr>
<td>White (EDB)</td>
<td>81.42</td>
<td>88.75***</td>
</tr>
<tr>
<td>Female (EDB)</td>
<td>63.99</td>
<td>65.89**</td>
</tr>
<tr>
<td>Was Married at Time of Home Health Admission (Survey)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Had Medicaid Buy-In for Part A and B Medicare (EDB)</td>
<td>22.86</td>
<td>22.14</td>
</tr>
<tr>
<td>Enrolled in Medicare for Fewer than Six Months Before Home Health Admission (EDB)</td>
<td>1.43</td>
<td>3.38***</td>
</tr>
<tr>
<td>Enrolled in HMO at Some Time During Six Months Before Home Health Admission (EDB)</td>
<td>0.99</td>
<td>1.47*</td>
</tr>
<tr>
<td>Was Survey Respondent (Survey)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>OASIS</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>Pre-IPS</td>
<td>Post-IPS</td>
</tr>
<tr>
<td><strong>Medical Conditions, Symptoms, and</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Needs at Home Health Admission</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer (EDB)</td>
<td>5.82</td>
<td>4.86</td>
</tr>
<tr>
<td>Diabetes (EDB)</td>
<td>6.65</td>
<td>4.13***</td>
</tr>
<tr>
<td>Cerebrovascular Disease (EDB)</td>
<td>6.58</td>
<td>5.12**</td>
</tr>
<tr>
<td>Skin Ulcers (EDB)</td>
<td>2.17</td>
<td>2.18</td>
</tr>
<tr>
<td>Had Risk Factors*a (OASIS)</td>
<td>23.74</td>
<td>19.47***</td>
</tr>
<tr>
<td>Had Depressed Feelings (OASIS)</td>
<td>8.11</td>
<td>4.11***</td>
</tr>
<tr>
<td>Demonstrated Disruptive Behaviors</td>
<td>23.85</td>
<td>20.69</td>
</tr>
<tr>
<td>(OASIS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prognosis at Home Health</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Admission (OASIS)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prognosis Was Good/Fair</td>
<td>84.77</td>
<td>93.29***</td>
</tr>
<tr>
<td>Life Expectancy Less than Six Months</td>
<td>6.62</td>
<td>1.97***</td>
</tr>
<tr>
<td>Rehabilitative Prognosis Was Good</td>
<td>67.66</td>
<td>79.74***</td>
</tr>
<tr>
<td><strong>Availability of Informal Care at</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Home Health Admission</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Had Live-In Informal Help (OASIS)</td>
<td>46.04</td>
<td>48.47</td>
</tr>
<tr>
<td>Had Paid Help or Was in Assisted-</td>
<td>15.01</td>
<td>25.04***</td>
</tr>
<tr>
<td>Living Residence (OASIS, Survey)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Had Nonpaid Help from Friends and</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Family in Month Before Home Health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admission (Survey)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure</td>
<td>OASIS</td>
<td>Four-Month Survey</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>--------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td>Pre-IPS</td>
<td>Post-IPS</td>
</tr>
<tr>
<td>Had Nonpaid Friend of Family Living at Home During Month Before Home</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Health Admission (Survey)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was in a Nursing Home During Month Before Home Health Admission</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(Survey)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was in Adult Day Care in Month Before Home Health Admission</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(Survey)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received Home-Delivered Meals During Month Before Home Health Admission</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(Survey)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measures of Patient’s Prior Service Use Before Home Health Admission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was in Hospital During Two Weeks Before Home Health Admission (Claims)</td>
<td>46.00</td>
<td>47.27</td>
</tr>
<tr>
<td>Length of Inpatient Stay Ending During Two Weeks Before Home Health</td>
<td>4.10</td>
<td>3.42</td>
</tr>
<tr>
<td>Admission (Days; Claims)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whether in SNF During Two Weeks Before Home Health Admission (Claims)</td>
<td>16.79</td>
<td>17.52</td>
</tr>
<tr>
<td>Whether Admitted to SNF During Six Months Before Home Health Admission</td>
<td>21.77</td>
<td>21.87</td>
</tr>
<tr>
<td>(Claims)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalizations During Six Months Before Home Health Admission</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(Number; Claims)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whether Hospitalized During Six Months Before Home Health Admission</td>
<td>67.58</td>
<td>63.57*</td>
</tr>
<tr>
<td>(Claims)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE II.4 (continued)

<table>
<thead>
<tr>
<th></th>
<th>OASIS</th>
<th>Four-Month Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-IPS</td>
<td>Post-IPS</td>
</tr>
<tr>
<td><strong>Agency Characteristics</strong> a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was For-Profit</td>
<td>53.26</td>
<td>33.51**</td>
</tr>
<tr>
<td>Had &lt;30,000 Visits</td>
<td>13.79</td>
<td>22.60</td>
</tr>
<tr>
<td>Was Hospital Based</td>
<td>16.88</td>
<td>16.30</td>
</tr>
<tr>
<td><strong>Area Characteristics</strong> a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>87.33</td>
<td>95.17</td>
</tr>
<tr>
<td>State</td>
<td></td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>24.31</td>
<td>32.59</td>
</tr>
<tr>
<td>Florida</td>
<td>8.44</td>
<td>8.15</td>
</tr>
<tr>
<td>Illinois</td>
<td>22.51</td>
<td>20.37</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>8.44</td>
<td>12.22</td>
</tr>
<tr>
<td>Texas</td>
<td>36.30</td>
<td>26.67</td>
</tr>
<tr>
<td>Patients (Number)</td>
<td>28,019</td>
<td>7,134</td>
</tr>
<tr>
<td>Agencies (Number)</td>
<td>38</td>
<td>27</td>
</tr>
</tbody>
</table>

**SOURCE:** OASIS and survey data, Medicare enrollment database, and Medicare claims.

*a Agency and area characteristics were the baseline values from the time the agency was first observed. They remained unchanged throughout the study.

*Significantly different from zero at the .10 level, two-tailed test
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

EDB = Medicare enrollment database; SNF = skilled nursing facility.
## TABLE II.5
WEIGHTED PRE- AND POST-IPS PERIOD MEANS FROM OASIS FOR PATIENT-LEVEL INDEPENDENT VARIABLES MEASURING HEALTH AND FUNCTIONAL STATUS

<table>
<thead>
<tr>
<th></th>
<th>Pre-IPS</th>
<th>Post-IPS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activities of Daily Living</strong>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grooming</td>
<td>54.3</td>
<td>46.7***</td>
</tr>
<tr>
<td>Bathing</td>
<td>77.9</td>
<td>84.7***</td>
</tr>
<tr>
<td>Toileting</td>
<td>33.2</td>
<td>32.5</td>
</tr>
<tr>
<td>Transferring</td>
<td>54.6</td>
<td>58.3</td>
</tr>
<tr>
<td>Ambulating</td>
<td>76.5</td>
<td>80.9**</td>
</tr>
<tr>
<td><strong>Instrumental Activities of Daily Living</strong>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparing Light Meals</td>
<td>63.0</td>
<td>70.0***</td>
</tr>
<tr>
<td>Housekeeping</td>
<td>91.7</td>
<td>97.0***</td>
</tr>
<tr>
<td>Managing Medications</td>
<td>54.9</td>
<td>52.1</td>
</tr>
<tr>
<td><strong>Medical Conditions</strong>b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>66.1</td>
<td>63.6</td>
</tr>
<tr>
<td>Most Problematic Pressure Ulcer</td>
<td>7.3</td>
<td>4.4***</td>
</tr>
<tr>
<td>Number of Pressure Ulcers</td>
<td>7.0</td>
<td>4.4***</td>
</tr>
<tr>
<td>Status of Most Problematic Surgical Wound</td>
<td>18.2</td>
<td>25.2***</td>
</tr>
<tr>
<td>Dyspnea</td>
<td>66.4</td>
<td>57.6***</td>
</tr>
<tr>
<td>Urinary Incontinence or Catheter Present</td>
<td>25.3</td>
<td>28.1*</td>
</tr>
<tr>
<td>Treated for a Urinary Tract Infection During Past 14 Days</td>
<td>11.2</td>
<td>9.4*</td>
</tr>
<tr>
<td>Confusion</td>
<td>41.0</td>
<td>40.3</td>
</tr>
<tr>
<td>Behavior Problem Frequency</td>
<td>14.5</td>
<td>9.8*</td>
</tr>
</tbody>
</table>

**Source:** OASIS data.

*a* Shows weighted means of binary indicators equal to one if a patient had any limitation and equal to zero otherwise.

*b* Shows weighted means of binary indicators equal to one if a patient had any symptom of a given condition and equal to zero otherwise.

*Significantly different from zero at the .10 level, two-tailed test.

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.
and that they were more likely to have had paid help.\textsuperscript{6} Despite these limitations, post-IPS period patients also had better prognoses, were less likely to have had risk factors, were less likely to be depressed, and were less likely to have had a hospitalization during the previous six months. Three of these measures were historically correlated with higher home health costs (Brown et al. 1991; and Phillips et al. 1992). Furthermore, means of the OASIS sample indicate that post-IPS period patients were less likely to suffer from diabetes or cerebrovascular disease—diseases that have also been demonstrated to be associated with historically higher home health costs.

D. METHODS

We used logit models to improve statistical precision, and to control for changes in factors that could affect outcomes and that changed between the pre- and post-IPS periods. Estimation of outcome variables in the OASIS and survey analyses will demonstrate how IPS affected patients’ satisfaction, health status, functional limitations, medical conditions, and emergency services use. The independent variable of primary interest indicates the effect of an observation occurring in the pre- or post-IPS period and will be interpreted as the impact of IPS. The remaining independent variables measure patient-level baseline characteristics at the time of home health admission in each period.

1. Logistic Regression Model

All the dependent variables studied in this report are binary. We used logistic regression models to estimate the impact of IPS. The logit model for estimating IPS impacts is:

\[
p(Y = 1) = \frac{e^{\beta x + POST\delta}}{1 + e^{\beta x + POST\delta}},
\]

\textsuperscript{6}We did not control for the functional limitations presented in Table II.5 because these variables were used to generate the OASIS outcomes.
where:

\[ p(Y_i = 1) = \text{the probability that the binary outcome variable } Y_i \text{ for the } i\text{th observation equals one} \]

\[ x_i = \text{the data vector of control variables for the } i\text{th observation, with elements } x_i = [1, x_{i1}, x_{i2}, \ldots, x_{ip}] \]

\[ \beta = \text{the vector of coefficients for the control variables, with elements } \beta = [\beta_0, \beta_1, \beta_2, \ldots, \beta_p], \beta_0 \text{ is the intercept term} \]

\[ POST = \text{the indicator variable for the analysis period and has the value of one if the observation belongs to the post-IPS period and zero if it is in the pre-IPS period} \]

\[ \delta = \text{the coefficient on the post-IPS period indicator } POST. \]

The coefficient \( \delta \) on the variable for analysis time period (\( POST \)) does not directly measure the estimated impact of IPS. It is the log of the estimated odds-ratio of experiencing the outcome for the post-IPS period relative to the pre-IPS period. That is,

\[ \delta = \ln \left( \frac{P_{Post}}{(1 - P_{Post})} \right) \times \left( \frac{P_{Pre}}{(1 - P_{Pre})} \right) \],

where \( P_{Post} \) is the predicted probability that an individual in the post-IPS period has \( Y = 1 \), and \( P_{Pre} \) is the analogous estimate for the same individual if he or she were in the pre-IPS period. We used the p-value for the coefficient to test the hypothesis that the IPS impact is significantly different from zero. To estimate the IPS impact on the probability that \( Y_i = 1 \), we used the coefficients estimated from the model to generate two predicted probabilities for each observation: one assuming that the observation belongs to the post-IPS period (\( POST = 1 \)), and the other assuming that it belongs to the pre-IPS period (\( POST = 0 \)). The impact estimate is the mean difference between these estimated probabilities, across the sample.
Throughout the tables of results, we present as a point of reference the regression-adjusted pre-IPS period mean of each outcome variable alongside the estimated impact. The regression-adjusted pre-IPS period mean provides a reasonable estimate of the mean value for the outcome variable that might be expected to occur in the absence of IPS implementation. We used this mean to assess the relative magnitude and importance of the estimated impact.

2. Issues in Weighting Observations

To estimate IPS impacts with patient-level data, one could weight the data in two different ways. The first method, called “agency-equal” weighting, weights the data so that each agency is given equal representation. The main benefit of this weighting scheme is that the analysis will reflect the level of the intervention, as IPS is an agency-level restriction. Furthermore, agency-equal weighting ensures that the experiences of a few large agencies do not dominate the impact estimates. The second method is to analyze the data unweighted, which gives each individual equal representation. The benefit of this approach is that it more accurately portrays the experience of Medicare beneficiaries, assuming the size distribution of the agency sample reflects the overall distribution of agencies. Because each weighting method has merit, we conducted the analysis using both methods. For simplicity, we present results obtained from using the agency-equal weighting as the main impact analysis. The unweighted analysis results are presented in Appendix A.

For each outcome variable to be analyzed, we constructed the agency-equal weight \( w_i \) for the \( i \)th agency as the ratio of the average number of patient-level observations per agency \( n/k \) to the number in the agency \( n_i \):

\[
(2) \quad w_i = \frac{n_i}{n/k} = \frac{2k}{n_i},
\]

where \( n \) is the total number of observations across all agencies, and \( k \) is the number of agencies. Each observation from agency \( i \) was then weighted by \( w_i \). Because \( n, n_i, \) and \( k \) vary with the
outcome, we constructed a separate set of weights for each outcome. Patient observations were thus weighted in inverse proportion to the size of the patients’ agency.

Extremely large weights (from agencies with very few observations) inflate the design effect, decrease the precision of estimated impacts, and, if a heavily weighted agency has an anomalous mean value, may distort the estimates. The weights in the OASIS data ranged from 0.23 to 77.81. For 16 OASIS outcomes, the maximum weight exceeded 50. The OASIS weight distributions were heavily skewed to the left, with only a few extremely large weights in the tail. A typical weight distribution was that for stabilization in pain, for which the 75th percentile was at 2.9, the 90th percentile at 7.9, the 95th percentile at 15.8, and the maximum at 56.7. The survey weights also were skewed to the left, but to a lesser degree. We therefore (1) truncated the weights for both the OASIS and survey data (Cox and Cohen 1985), and (2) deleted agencies that had fewer than nine observations.7 Each observation was assigned the lesser of the actual calculated weight \(w_i\) or a maximum allowable weight \(w_{\text{max}}\). We determined by inspection that the 85th percentile was a natural cut-off point for all weight distributions, so we used the value of \(w_{\text{max}}\) for each outcome.

### 3. Statistical Hypothesis Testing

To test the null hypothesis that IPS had no impact on the outcomes, it is necessary to estimate correctly the standard errors of the coefficient on the post-IPS period indicator. It also requires a decision about the level of significance at which to reject the null hypothesis.

---

7Deleting agencies with fewer than nine observations, a potential source of large weights, did not affect the OASIS sample and resulted in the removal of 19 observations from the survey sample in one pre-IPS period agency and two post-IPS period agencies.
a. Design Effects

We used STATA software to obtain the appropriate standard errors for the impact estimates, as this software is well equipped to analyze clustered data. In our data, the clustering of patients within home health agencies might generate a correlation among patients in the same agency that is due to the agency’s effect on patient outcomes. Standard statistical software, which assumes simple random sampling from an infinite population, generally underestimates variances in clustered data by an amount, called the “design effect,” that increases as the intracluster correlation increases. Underestimation of the variances, in turn, leads to falsely small p-values, so that impacts appear to be more significant than they actually are. STATA is able to correct for this design effect, as well as to account for the increase in the standard errors caused by the use of sample weights.

b. Significance Levels

For each outcome, we used a two-tailed $t$-statistic to test the null hypothesis that there is no difference between the regression-adjusted population means for the pre- and post-IPS periods. The associated p-value is the probability of obtaining the observed estimate under the null hypothesis. We chose a p-value of less than 0.10 to reject the null hypothesis, and, thus, to

---

8Our OASIS sample actually includes the entire population of patients admitted to the agencies in this study. (The time period varies depending on whether agency chose to participate in the study.) However, because we wished to make inferences about how outcomes would be affected for patients admitted at other times and to other agencies, we treated patients in the data as though they were drawn in a two-stage random sample from the pool of all (future) patients in all agencies.

9In general, we were concerned that prospective payment would have adverse effects on the quality of home health care and on patient outcomes. We used two-tailed tests in our analyses to avoid confusion, and to flag estimated pre-/post-IPS period differences of the “wrong” expected sign that were large enough to be statistically significant. For impacts with the “correct” expected sign, a two-tailed test is less likely than a one-tailed test at the same significance level to reject the hypothesis of no demonstration effect (all else equal).
establish that an IPS impact is statistically significant. At this p-value, however, approximately 10 percent of independent tests would show, simply by chance, a statistically significant IPS impact in the absence of a true program effect (known as a Type I error). Therefore, in assessing whether a statistically significant impact, especially one with a p-value between 0.05 and 0.10, should be interpreted as a true program impact, we also considered whether the sign and magnitude of the estimated effect were consistent with those for related outcomes.

4. Sensitivity Tests

We performed sensitivity tests to determine whether our estimates were contingent on our methodology. First, agency attrition between the pre- and post-IPS periods caused the sample of agencies in both periods to vary slightly and could have had an effect on the results. Fifteen of the 41 agencies that participated in the pre-IPS period also participated in the post-IPS period. To minimize the effects of agency attrition, replacement agencies in the post-IPS period were chosen to resemble as closely as possible the agencies that dropped out. However, the change in sample composition, if not sufficiently controlled for by the explanatory variables, could affect the results. Therefore, we reestimated our impacts with an agency sample that was constant in both periods, and the similarity of these results to those of the main analysis will indicate the effect of the differences in sample composition.

Second, it is possible that a portion of the IPS impact occurred during what we have thus far defined as the pre-IPS period. The pre-IPS period of the analysis consists of data from May 1996 through December 1998—months during which the agencies were not subject to IPS restrictions, which was signed into law in August 1997 and took effect the following October as part of BBA 1997. The post-IPS period began after the HHPP Demonstration ended and lasted from March to December 2000. By beginning the post-IPS period in the main analysis after the
demonstration ended, we are assuming that agencies sustained the IPS impact after they first were subject to the restrictions of IPS (that is, when they completed the HHPP Demonstration).

However, it is possible that the agencies’ expectation of the IPS restrictions may have caused them to alter their behavior before the regulations actually took effect. In this case, part of the IPS impact would have occurred during what we have thus far considered as the pre-IPS period. To examine the extent to which this may have occurred, we altered our definition of the pre-IPS period to include only the agencies’ first year in the demonstration, which was completed before the IPS restrictions took effect. We then reestimated the impacts, using the altered definition of the pre-IPS period. If agencies did indeed alter their behavior before they actually were subject to IPS restrictions, then we would expect results from the main analysis to understate the true IPS impact, and our reestimated impacts to be of larger magnitude. If they did not alter their behavior, then we would expect the results from this sensitivity test to be similar to those in the main analysis.

\[10\] Indeed, service utilization in these agencies fell substantially before the agencies were subject to IPS restrictions (Archibald and Cheh 2001).
III. RESULTS

The Interim Payment System (IPS) is a fundamentally new approach to agency payment that gives agencies an incentive to reduce the costs of Medicare services by whatever means feasible. This chapter presents the results of our analysis to determine whether the substantial reduction in the provision of home health care as a result of IPS had a detrimental effect on patients. We begin the chapter by describing the theoretical predictions of the impact of IPS on patient outcomes. We then present our results for the impact of IPS on patients’ functional status, medical symptoms and outcomes, and self-reports of general health status. Finally, we discuss the results from our sensitivity analyses, which determine whether our results were contingent on the methodology we used.

A. POTENTIAL EFFECTS OF IPS ON PATIENT OUTCOMES

Research has not yet determined how substantial reductions in service provision, such as those made by home health agencies after IPS was imposed, affect health outcomes. As discussed in Chapter I, empirical studies have reached mixed conclusions.

Improved, unchanged, and worsened patient outcomes accompanying reductions in the amount of home health care all are consistent with theoretical considerations. Suppose, for example, that an agency responds to IPS payment incentives by providing “less” home health care, such as fewer home visits and earlier discharge. Furthermore, suppose the agency does not otherwise change the “content” or quality of the care. The hypothetical solid curve in Figure III.1 (labeled “typical quality of care”) relates patient outcomes to the amount of home health care and shows possible effects of a reduction of this type. In the first instance, the amount of home care falls, from pre-IPS to post-IPS. Because the curve is flat, we do not observe any decrement in
The curves show hypothesized relationships between a patient outcome and the amount of home health care provided. If quality of care remains the same on the solid “typical quality of care” curve, a decrease in the amount of home health care provided from Pre-IPS to Post-IPS causes no change in patient outcome (vertical movement from a to b). However, if the curve actually has a downward portion (dotted line), the decrease from Pre-IPS to Post-IPS causes the patient outcome to improve (vertical movement from f to b). If the decrease in the amount of home health care provided goes from Pre-IPS* to Post-IPS*, then the patient outcome also worsens. Finally, if quality of care improves under the IPS from the solid “typical quality of care” curve to the dashed “improved quality of care” curve, a decrease in the amount of home health care is associated with an improvement in patient outcome (vertical movement from a to e).
patient function (that is, no vertical movement from point \textit{a} to point \textit{b}). In the second instance, the decrease in the amount of care, from \textit{pre-IPS*} to \textit{post-IPS*}, is the same; in contrast to the first instance, however, the curve has a steep slope, so patient functioning worsens (that is, substantial vertical movement from point \textit{c} to point \textit{d}).

It is conceivable that too much home health care could actually worsen some patient outcomes (Fisher and Welch 1999). For example, too much help from agency staff could promote patient passivity and dependency. This possibility is shown by the dotted curve in Figure III.1, where a fall in the amount of care, from \textit{pre-IPS} to \textit{post-IPS}, actually improves patient outcomes (that is, vertical movement from point \textit{f} to point \textit{b}).

Finally, suppose agencies responded to the need to reduce visits with innovations that simultaneously reduced the amount of care and improved the content and quality of care. For example, agencies might improve patient education and provide specialized nursing care, such as a wound care specialist. The possibility of less albeit better care is shown by the dashed curve in Figure III.1 (labeled “improved quality of care”). For any given amount of home care services, improved quality of care achieves better patient outcomes than does typical care. With a shift to higher-quality care, a fall in the amount of home health care, from \textit{pre-IPS} to \textit{post-IPS}, is accompanied by improved patient outcomes (that is, vertical movement from point \textit{a} to point \textit{e}).

1The shapes of the curves reflect our assumptions about the relationship between the amount of home health care and patient outcomes. When the amount of home health care services is small, the curve initially slopes upward because providing additional home care services improves patients’ outcomes substantially. With larger amounts of home health care services, however, more visits produce dwindling marginal gains, so the curve flattens.

2In the Home Health Prospective Payment (HHPP) Demonstration, treatment agencies reported trying several new strategies to reduce home health care, all of which had the potential to improve the quality of care. The strategies included implementing standardized treatment protocols and care maps, hiring specialists (for example, wound care specialist nurses), using the telephone more frequently, using laptop computers to improve documentation, instructing patients and their families intensively in self-care, and involving community-based providers early in the episode of care (Phillips and Thompson 1997).
This discussion illustrates that theory does not point clearly to how reductions in home health care service will affect patient outcomes. The issue must be addressed empirically.

**B. EFFECT OF IPS ON FUNCTIONAL STATUS**

We did not find significant evidence from our analysis of OASIS data that IPS had detrimental effects on patient functioning. Conclusions about the impact of IPS on functional status based on the survey analysis are less clear, however; in fact, that analysis raises some concerns. For the analysis of OASIS data, none of the measures of activities of daily living (ADLs), which measure change in ADL functioning over the course of an episode of care, indicated that patients were less likely to improve or to stabilize during the post-IPS period (relative to the pre-IPS period). In fact, we found significant evidence suggesting that IPS may have had beneficial effects on the ability of patients to perform ADLs and instrumental activities of daily living (IADLs). Every ADL that we measured demonstrated that the functional status of post-IPS period patients was more likely to stabilize than was the functional status of pre-IPS period patients (Table III.1). In addition, patients’ ability to groom themselves and to use the toilet was more likely to improve during the post-IPS period than during the pre-IPS period. Similarly, post-IPS period patients’ performance of IADLs, such as preparing light meals and managing medications, was more likely to stabilize (Table III.2).

Based on the survey analysis, however, the impact of IPS is unclear. Survey estimates of the impact of IPS are inconsistent with those from the OASIS analysis, suggesting that IPS had detrimental effects on patient functioning. For example, the proportions of patients who did bathe, who did transfer, and who did ambulate were significantly lower during the post-IPS period than during the pre-IPS period, as were the proportions who could bathe, who could eat, and who could ambulate (Table III.3). The reductions in the proportions who did bathe, who did transfer, and who did ambulate could be explained by an increase in the use of non-Medicare
TABLE III.1
ESTIMATED DIFFERENCES IN BASIC ACTIVITIES OF DAILY LIVING BETWEEN PRE- AND POST-IPS PERIODS, FROM OASIS DATA

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N^a</th>
<th>Predicted Pre-IPS Group Mean (Percentage)^b</th>
<th>Estimated Pre-Post Difference^c (p-Value)^d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grooming</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>17,465</td>
<td>59.24</td>
<td>5.97*** (.01)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>32,518</td>
<td>88.07</td>
<td>4.67*** (.00)</td>
</tr>
<tr>
<td>Bathing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>27,301</td>
<td>55.71</td>
<td>2.36 (.20)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>32,781</td>
<td>85.16</td>
<td>3.58*** (.00)</td>
</tr>
<tr>
<td>Toileting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>10,902</td>
<td>63.02</td>
<td>4.03** (.05)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>33,142</td>
<td>91.88</td>
<td>3.33*** (.00)</td>
</tr>
<tr>
<td>Transferring</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>18,935</td>
<td>52.45</td>
<td>2.38 (.36)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,219</td>
<td>89.48</td>
<td>4.20*** (.00)</td>
</tr>
<tr>
<td>Ambulating</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>26,449</td>
<td>33.64</td>
<td>-0.43 (.83)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,368</td>
<td>90.23</td>
<td>3.83*** (.00)</td>
</tr>
</tbody>
</table>

SOURCE: OASIS data.

NOTE: Observations have been weighted to give agencies equal representation in the analysis.

^aPatients who already were at the best level of a measure at the start of an episode of care could not improve in that measure. Patients who already were at the worst level of a measure at the start of care could not stabilize in that measure. Thus, the number of patients “eligible” for improvement or stabilization differed within and across measures.

^bThe regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

^cEstimated differences are regression adjusted through logit models to control for preexisting differences between pre- and post-IPS period patients.
The p-value corresponds to a test of whether the pre/post-IPS difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.
### Table III.2

**Estimated Differences in Instrumental Activities of Daily Living Between Pre- and Post-IPS Periods, From OASIS Data**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Predicted Pre-IPS Group Mean (Percentage)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Estimated Pre-Post Difference&lt;sup&gt;c&lt;/sup&gt; (p-Value)&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Light Meal Preparation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>22,290</td>
<td>52.68</td>
<td>1.09 (.63)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>25,585</td>
<td>85.20</td>
<td>2.96** (.02)</td>
</tr>
<tr>
<td><strong>Housekeeping</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>32,174</td>
<td>43.02</td>
<td>-0.72 (.67)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>20,039</td>
<td>75.75</td>
<td>1.72 (.30)</td>
</tr>
<tr>
<td><strong>Management of Medications</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>17,365</td>
<td>44.69</td>
<td>-0.63 (.76)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>27,343</td>
<td>87.28</td>
<td>2.48*** (.01)</td>
</tr>
</tbody>
</table>

**Source:** OASIS data.

**Note:** Observations have been weighted to give agencies equal representation in the analysis.

<sup>a</sup> Patients who already were at the best level of a measure at the start of an episode of care could not improve in that measure. Patients who already were already at the worst level of a measure at the start of care could not stabilize in that measure. Thus, the number of patients “eligible” for improvement or stabilization differed within and across measures.

<sup>b</sup> The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

<sup>c</sup> Estimated differences are regression adjusted through logit models to control for preexisting differences between pre- and post-IPS period patients.

<sup>d</sup> The p-value corresponds to a test of whether the pre/post-IPS difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test  
**Significantly different from zero at the .05 level, two-tailed test.  
***Significantly different from zero at the .01 level, two-tailed test.
### TABLE III.3
ESTIMATED DIFFERENCES IN FUNCTIONAL STATUS BETWEEN PRE- AND POST-IPS PERIODS, FROM SURVEY DATA

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Predicted Pre-IPS Group Mean (Percentage)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Estimated Pre-Post Difference&lt;sup&gt;b&lt;/sup&gt;</th>
<th>(p-Value)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activities of Daily Living</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bathing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did bathe</td>
<td>2,285</td>
<td>56.57</td>
<td>-7.01***</td>
<td>(.01)</td>
</tr>
<tr>
<td>Could bathe</td>
<td>2,268</td>
<td>64.26</td>
<td>-5.91**</td>
<td>(.02)</td>
</tr>
<tr>
<td>Eating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did eat</td>
<td>2,254</td>
<td>82.23</td>
<td>-2.99</td>
<td>(.25)</td>
</tr>
<tr>
<td>Could eat</td>
<td>2,246</td>
<td>92.27</td>
<td>-2.82*</td>
<td>(.08)</td>
</tr>
<tr>
<td>Transferring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did transfer</td>
<td>2,292</td>
<td>68.42</td>
<td>-5.84***</td>
<td>(.01)</td>
</tr>
<tr>
<td>Could transfer</td>
<td>2,280</td>
<td>82.35</td>
<td>-1.78</td>
<td>(.39)</td>
</tr>
<tr>
<td>Ambulating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did ambulate</td>
<td>2,226</td>
<td>68.06</td>
<td>-6.10**</td>
<td>(.02)</td>
</tr>
<tr>
<td>Could ambulate</td>
<td>2,216</td>
<td>79.74</td>
<td>-6.03**</td>
<td>(.02)</td>
</tr>
<tr>
<td><strong>Instrumental Activities of Daily Living</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Take Medications Independently</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did take medications independently</td>
<td>2,250</td>
<td>52.62</td>
<td>-0.76</td>
<td>(.73)</td>
</tr>
<tr>
<td>Could take medications independently</td>
<td>2,236</td>
<td>69.89</td>
<td>1.95</td>
<td>(.35)</td>
</tr>
</tbody>
</table>

**Source:** OASIS data.

**Note:** Observations have been weighted to give agencies equal representation in the analysis.
a Patients who were already at the best level of a measure at the start of an episode of care could not improve in that measure. Patients who were already at the worst level of a measure at the start of care could not stabilize in that measure. Thus, the number of patients “eligible” for improvement or stabilization differed within and across measures.

b The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

c Estimated differences are regression adjusted through logit models to control for preexisting differences between pre- and post-IPS period patients.

d The p-value corresponds to a test of whether the pre/post-IPS difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

* Significantly different from zero at the .10 level, two-tailed test
** Significantly different from zero at the .05 level, two-tailed test.
*** Significantly different from zero at the .01 level, two-tailed test.
services and are not necessarily inconsistent with the OASIS results. That is, a patient who received services from another source would be less likely to perform these ADLs independently—an outcome consistent with the OASIS data. However, the fact that a higher proportion of post-IPS period patients than pre-IPS period patients reported an inability to perform some of these tasks independently suggests that patients in the former group either improved less during the post-IPS period than during the pre-IPS period (which is inconsistent with the OASIS data) or began their episode at a lower level of functioning (as we noted in Chapter II).

To examine whether the survey impact estimates in Table III.3 were driven by differences in baseline functional status between the pre- and post-IPS periods, we matched the survey sample to the OASIS data and reestimated the impacts.⁴ The results of this analysis are presented in columns B and C of Table III.4. We first estimated the survey impacts using the observations in which the survey data and the OASIS sample matched (column B).⁵ To observe how the impacts changed, we then reestimated the models with additional controls for baseline health and functional status obtained from the OASIS sample (column C).

Adding the OASIS control variables for baseline health and functional status to the survey analysis substantially altered the results. All five of the significant negative impacts estimated with the matched sample (that is, did bathe, did transfer, did ambulate, could bathe, and could

---

³Note that we originally planned to control for the baseline functional status using OASIS data. However, a substantial proportion of the survey (34 percent) did not match to the OASIS data, so we would have lost one-third of the sample.

⁴See the note to Table III.4 for a list of control variables we added from OASIS.

⁵These impacts are different from those in column A of Table III.4 because the sample changed—867 survey observations did not match to OASIS data, and 34 additional observations were deleted because they were in agencies that had fewer than nine observations.
### TABLE III.4

**SENSITIVITY OF SURVEY RESULTS TO ADDITION OF OASIS CONTROL VARIABLES**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Full Sample (A)</th>
<th>Matched Survey Sample (B)</th>
<th>Matched Survey Sample with Additional Controls (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activities of Daily Living</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bathing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did bathe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,285</td>
<td>1,627</td>
<td>1,627</td>
</tr>
<tr>
<td>Mean^d</td>
<td>56.57</td>
<td>59.22</td>
<td>56.86</td>
</tr>
<tr>
<td>Impact</td>
<td>-7.01***</td>
<td>-8.28***</td>
<td>-4.33</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.00)</td>
<td>(.12)</td>
</tr>
<tr>
<td>Could bathe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,268</td>
<td>1,614</td>
<td>1,614</td>
</tr>
<tr>
<td>Mean^d</td>
<td>64.26</td>
<td>66.38</td>
<td>63.96</td>
</tr>
<tr>
<td>Impact</td>
<td>-5.91**</td>
<td>-6.38**</td>
<td>-2.83</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.02)</td>
<td>(.02)</td>
<td>(.20)</td>
</tr>
<tr>
<td><strong>Eating</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did eat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,254</td>
<td>1,610</td>
<td>1,610</td>
</tr>
<tr>
<td>Mean^d</td>
<td>82.23</td>
<td>81.69</td>
<td>79.13</td>
</tr>
<tr>
<td>Impact</td>
<td>-2.99</td>
<td>-0.70</td>
<td>1.95</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.25)</td>
<td>(.81)</td>
<td>(.40)</td>
</tr>
<tr>
<td>Could eat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,246</td>
<td>1,607</td>
<td>1,607</td>
</tr>
<tr>
<td>Mean^d</td>
<td>92.27</td>
<td>91.50</td>
<td>89.31</td>
</tr>
<tr>
<td>Impact</td>
<td>-2.82*</td>
<td>-1.26</td>
<td>1.02</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.08)</td>
<td>(.48)</td>
<td>(.49)</td>
</tr>
<tr>
<td><strong>Transferring</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did transfer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,292</td>
<td>1,632</td>
<td>1,632</td>
</tr>
<tr>
<td>Mean^d</td>
<td>68.42</td>
<td>69.49</td>
<td>67.79</td>
</tr>
<tr>
<td>Impact</td>
<td>-5.84***</td>
<td>-5.76**</td>
<td>-4.18**</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.03)</td>
<td>(.04)</td>
</tr>
<tr>
<td>Could transfer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,280</td>
<td>1,624</td>
<td>1,624</td>
</tr>
<tr>
<td>Mean^d</td>
<td>82.35</td>
<td>83.57</td>
<td>81.26</td>
</tr>
<tr>
<td>Impact</td>
<td>-1.78</td>
<td>-1.74</td>
<td>0.48</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.39)</td>
<td>(.50)</td>
<td>(.84)</td>
</tr>
</tbody>
</table>
### TABLE III.4 (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Full Sample (A)</th>
<th>Matched Survey Sample (B)</th>
<th>Matched Survey Sample with Additional Controls (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ambulating</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did ambulate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,226</td>
<td>1,590</td>
<td>1,590</td>
</tr>
<tr>
<td>Mean (^d)</td>
<td>68.06</td>
<td>68.80</td>
<td>65.49</td>
</tr>
<tr>
<td>Impact</td>
<td>-6.10**</td>
<td>-5.16*</td>
<td>-1.47</td>
</tr>
<tr>
<td>p-Value</td>
<td>.02</td>
<td>.06</td>
<td>.55</td>
</tr>
<tr>
<td>Could ambulate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,216</td>
<td>1,582</td>
<td>1,582</td>
</tr>
<tr>
<td>Mean (^d)</td>
<td>79.74</td>
<td>82.93</td>
<td>78.78</td>
</tr>
<tr>
<td>Impact</td>
<td>-6.03**</td>
<td>-7.69***</td>
<td>-3.26*</td>
</tr>
<tr>
<td>p-Value</td>
<td>.02</td>
<td>.00</td>
<td>.10</td>
</tr>
<tr>
<td><strong>Take Medications Independently</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did take medications independently</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,250</td>
<td>1,589</td>
<td>1,589</td>
</tr>
<tr>
<td>Mean (^d)</td>
<td>52.62</td>
<td>53.70</td>
<td>52.61</td>
</tr>
<tr>
<td>Impact</td>
<td>-0.76</td>
<td>0.37</td>
<td>1.75</td>
</tr>
<tr>
<td>p-Value</td>
<td>.73</td>
<td>.88</td>
<td>.49</td>
</tr>
<tr>
<td>Could take medications independently</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,236</td>
<td>1,578</td>
<td>1,578</td>
</tr>
<tr>
<td>Mean (^d)</td>
<td>69.89</td>
<td>71.40</td>
<td>70.11</td>
</tr>
<tr>
<td>Impact</td>
<td>1.95</td>
<td>1.95</td>
<td>3.78*</td>
</tr>
<tr>
<td>p-Value</td>
<td>.35</td>
<td>.31</td>
<td>.08</td>
</tr>
</tbody>
</table>

**SOURCE:** Survey data.

**NOTE:** Observations have been weighted to give agencies equal representation in the analysis.

\(^a\)Column A represents the survey results presented in Table III.3

\(^b\)Column B replicates column A with the matched survey/OASIS sample.

\(^c\)Column C estimates survey impacts with the following additional control variables from OASIS: original reason for Medicare (reached age 65), had risk factors, had depressed feelings, demonstrated disruptive behaviors, prognosis was good or fair, life expectancy was less than six months, rehabilitative prognosis was good, and binary indicators of the functional limitations and medical conditions presented in Table II.5.

\(^d\)The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

*Significantly different from zero at the .10 level, two-tailed test
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.
ambulate) decreased in magnitude, but three of the five (did bathe, did transfer, and could ambulate) remained at least marginally significant. We were most concerned about the negative IPS impacts for the measures, “could ambulate” and “could eat”; an increase in the use of non-Medicare services may have driven the negative impacts on “did bathe” and on “did ambulate,” but the “could” measures are more likely to indicate the effect of IPS. The impact on “could eat” was insignificant when estimated using the matched sample, but was positive when we added OASIS controls. This change indicates that the estimated impact on “could eat” may have been driven by uncontrolled for baseline characteristics. However, the fact that the negative impact for “could ambulate” remained marginally significant after controlling for baseline functional status is cause for some concern.

C. EFFECT OF IPS ON MEDICAL SYMPTOMS AND PATIENT OUTCOMES

How the IPS system affected medical symptoms and outcomes is another important issue. The discussion in the previous section indicates that IPS did not negatively affect the physical functioning of patients. Given that agencies reduced their physical therapy services by only a small amount, however, this finding may not be surprising (McCall 2001). By contrast, medical symptoms and outcomes are more likely to be affected by the two services that agencies reduced the most—skilled nursing visits and home health aide visits.

We found little evidence to suggest that IPS had a detrimental impact on wound care. We did find that there was a significant negative impact on the probability of the most problematic surgical wound improving (Table III.5). However, given that the probability of the most problematic pressure ulcer improving was the same during the pre- and post-IPS periods, we are inclined to believe the finding most likely reflects the shorter post-IPS period episode length (an average episode length of 53 days in the pre-IPS period, compared with just over half that—29 days—in the post-IPS period). It is likely that some patients simply did not receive care long
TABLE III.5

ESTIMATED DIFFERENCES IN MEDICAL SYMPTOMS AND OUTCOMES
BETWEEN PRE- AND POST-IPS PERIODS, FROM OASIS DATA

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N^a</th>
<th>Predicted Pre-IPS Group Mean (Percentage)^b</th>
<th>Estimated Pre-Post Difference^c (p-Value^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Problematic Pressure Ulcer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>2,277</td>
<td>76.52</td>
<td>3.16 (.45)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,603</td>
<td>97.41</td>
<td>1.78*** (.00)</td>
</tr>
<tr>
<td>Number of Pressure Ulcers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>2,022</td>
<td>76.63</td>
<td>3.74 (.28)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,835</td>
<td>97.69</td>
<td>1.52*** (.00)</td>
</tr>
<tr>
<td>Status of Most Problematic Surgical Wound</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>7,986</td>
<td>83.29</td>
<td>-6.30*** (.00)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,396</td>
<td>97.50</td>
<td>-0.34 (.30)</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>21,846</td>
<td>54.33</td>
<td>-0.64 (.74)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>32,783</td>
<td>83.95</td>
<td>2.96*** (.00)</td>
</tr>
<tr>
<td>Dyspnea</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>20,917</td>
<td>49.16</td>
<td>7.03*** (.00)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,060</td>
<td>81.00</td>
<td>5.15*** (.00)</td>
</tr>
<tr>
<td>Urinary Incontinence or Catheter Present</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>7,490</td>
<td>54.43</td>
<td>-3.24 (.21)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>33,380</td>
<td>95.46</td>
<td>1.30*** (.00)</td>
</tr>
<tr>
<td>Outcome</td>
<td>N&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Predicted Pre-IPS Group Mean (Percentage)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Estimated Pre-Post Difference&lt;sup&gt;c&lt;/sup&gt; (p-Value&lt;sup&gt;d&lt;/sup&gt;)</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>---------------</td>
<td>-----------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Treated for Urinary Tract Infection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>2,393</td>
<td>79.60</td>
<td>0.96 (.78)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>30,913</td>
<td>98.22</td>
<td>0.37* (.10)</td>
</tr>
<tr>
<td>Confusion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>11,833</td>
<td>48.17</td>
<td>-1.40 (.57)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>31,536</td>
<td>86.29</td>
<td>2.92*** (.01)</td>
</tr>
<tr>
<td>Frequency of Behavioral Problems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>3,906</td>
<td>77.10</td>
<td>0.83 (.78)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>33,420</td>
<td>92.99</td>
<td>2.91*** (.00)</td>
</tr>
</tbody>
</table>

**Source:** OASIS data.

**Note:** Observations have been weighted to give agencies equal representation in the analysis.

<sup>a</sup>Patients who already were at the best level of a measure at the start of an episode of care could not improve in that measure. Patients who already were at the worst level of a measure at the start of care could not stabilize in that measure. Thus, the number of patients “eligible” for improvement or stabilization differed within and across measures.

<sup>b</sup>The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

<sup>c</sup>Estimated differences are regression adjusted through logit models to control for preexisting differences between pre- and post-IPS period patients.

<sup>d</sup>The p-value corresponds to a test of whether the pre/post-IPS difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test  
**Significantly different from zero at the .05 level, two-tailed test.  
***Significantly different from zero at the .01 level, two-tailed test.
enough for some slow-healing wounds to improve substantially. Furthermore, we also found that, during the post-IPS period, the most problematic pressure ulcer was less likely to worsen, and that the number of pressure ulcers was less likely to increase. Both findings suggest that IPS had a positive impact on wound care.

With respect to other measures of health status, the results of our analysis consistently suggest that IPS improved the quality of care. For every measure, patients were significantly more likely to stabilize during the post-IPS period than during the pre-IPS period; for almost all measures, they were at least as likely to improve on a given health outcome during the post-IPS period (Table III.5). Furthermore, patients with symptoms of dyspnea (shortness of breath) were more likely to improve during the post-IPS period than during the pre-IPS period.

Why were patients’ outcomes and symptoms more likely to stabilize in the post-IPS period? One explanation is that the stabilization was the result of home health nurses’ efforts to improve patient self-care, so that patients would be able to leave home health earlier. Indeed, McCall et al. (2002) found that, in the post-BBA period, patients were more likely to report that nurses encouraged their independence. Moreover, agency staff reported in the HHPP Demonstration that improved teaching was a key method used to reduce the length of the episode. If the teaching of self-care helped patients monitor symptoms and medications more effectively, then it is likely that patients also identified and sought treatment for potential problems early, before the problems developed into significant complications.

D. EFFECT OF IPS ON OVERALL HEALTH STATUS

IPS is unlikely to have changed patients’ overall health status; however, the evidence on this issue is not conclusive. Patients reported roughly the same self-perceived health status before and after implementation of IPS, but they also reported that they were significantly more likely to have spent time in bed during the two weeks before their survey date, and that they were less
satisfied with life (Table III.6). The latter two results suggest that IPS was having a detrimental effect.

As noted, the possibility that post-IPS patients were sicker than pre-IPS patients at the start of their episodes may explain these findings. To determine whether the inclusion of additional patient characteristics at baseline substantially decreased the estimated impacts, we obtained additional control variables from the OASIS data and reestimated our results on the smaller matched sample. Including the additional variables did not decrease the magnitude of our estimates (Table III.7). The difference in the probability of having any days in bed during the two-week period preceding the survey remained near five percentage points, regardless of whether one included or excluded the patient characteristics. Furthermore, including additional patient characteristics actually increased the negative effect on patients’ satisfaction with life. This result suggests that the impacts were not driven by patient-level differences in the functional status and health measures that we added to the model.

How is it that post-IPS period patients could have done just as well as pre-IPS period patients in individual measures of functioning and health status, yet report inferior levels of overall health status? We have three potential explanations for this seeming inconsistency. First, it is possible that IPS negatively affected aspects of health that we did not measure, and that these aspects led to the overall decline in health status. Second, we know that post-IPS period patients were more functionally limited than were pre-IPS period patients (Table II.5). Although we concluded that these functional limitations did not drive the IPS impacts on overall health status, it is possible that additional patient characteristics for which we did not control affected the results. Third, the finding may have resulted from a statistical anomaly in the data.
### TABLE III.6

**ESTIMATED DIFFERENCES BETWEEN PATIENTS’ SELF-REPORTS OF GENERAL HEALTH DURING THE PRE- AND POST-IPS PERIODS, FROM SURVEY DATA**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N</th>
<th>Predicted Pre-IPS Group Mean (Percentage)</th>
<th>Estimated Pre-Post Difference&lt;sup&gt;b&lt;/sup&gt; (p-Value)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported Health Good or Excellent</td>
<td>2,262</td>
<td>42.84</td>
<td>-2.71 (.33)</td>
</tr>
<tr>
<td>Any Days in Bed Within Past Two Weeks</td>
<td>2,347</td>
<td>27.53</td>
<td>6.50*** (.01)</td>
</tr>
<tr>
<td>Satisfied with Life</td>
<td>2,197</td>
<td>62.54</td>
<td>-6.75*** (.01)</td>
</tr>
</tbody>
</table>

**SOURCE:** Survey data.

**NOTE:** Observations have been weighted to give agencies equal representation in the analysis.

<sup>a</sup>The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

<sup>b</sup>Estimated differences are regression adjusted through logit models to control for preexisting differences between pre- and post-IPS period patients.

<sup>c</sup>The p-value corresponds to a test of whether the pre/post-IPS difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.
### TABLE III.7

SENSITIVITY OF SURVEY RESULTS TO ADDITION OF OASIS CONTROL VARIABLES

<table>
<thead>
<tr>
<th></th>
<th>Full Sample (A)</th>
<th>Matched Survey Sample (B)</th>
<th>Matched Survey Sample with Additional Controls (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reported Health Good or Excellent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,262</td>
<td>1,615</td>
<td>1,615</td>
</tr>
<tr>
<td>Mean</td>
<td>42.84</td>
<td>43.36</td>
<td>45.18</td>
</tr>
<tr>
<td>Impact</td>
<td>-2.71</td>
<td>-1.01</td>
<td>-5.18</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.33)</td>
<td>(.75)</td>
<td>(.14)</td>
</tr>
<tr>
<td><strong>Any Days in Bed Within Past Two Weeks</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,347</td>
<td>1,669</td>
<td>1,669</td>
</tr>
<tr>
<td>Mean</td>
<td>27.53</td>
<td>27.05</td>
<td>26.88</td>
</tr>
<tr>
<td>Impact</td>
<td>6.50***</td>
<td>4.83*</td>
<td>5.43**</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.08)</td>
<td>(.05)</td>
</tr>
<tr>
<td><strong>Satisfied with Life</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,197</td>
<td>1,564</td>
<td>1,564</td>
</tr>
<tr>
<td>Mean</td>
<td>62.54</td>
<td>59.92</td>
<td>61.62</td>
</tr>
<tr>
<td>Impact</td>
<td>-6.75***</td>
<td>-3.98</td>
<td>-7.05**</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.17)</td>
<td>(.02)</td>
</tr>
</tbody>
</table>

**SOURCE:** Survey data.

*Column A replicates the survey results presented in Table III.3.

*Column B replicates column A with the matched survey/OASIS sample.

*Column C estimates survey impacts with the following additional control variables from OASIS: original reason for Medicare (reached age 65), had risk factors, had depressed feelings, demonstrated disruptive behaviors, prognosis was good or fair, life expectancy was less than six months, rehabilitative prognosis was good, and binary indicators of the functional limitations and medical conditions presented in Table II.5.

The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

*Significantly different from zero at the .10 level, two-tailed test.

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.
Because the overall health status measures are the measures that we believe are the least sensitive to the provision of home health services, we are inclined to believe that the results reflect unmeasured baseline characteristics that changed during the analysis period. However, we cannot definitively rule out the possibility that the reduction in home health services led to the declines.

E. SENSITIVITY TESTS

We performed three sensitivity tests to determine whether our results were contingent on our weighting scheme or our sample composition. First, we investigated the robustness of our results to not weighting the observations; this test is equivalent to weighting agencies proportional to their size.\(^6\) Analyses in which observations are unweighted are useful because small agencies that have anomalous results may distort the “agency equal weighted” analysis. Furthermore, although we chose to weight agencies equally, reflecting the fact that IPS was implemented at the agency-level, it is also important to determine the IPS impact on the average Medicare beneficiary, as in the unweighted analysis.

Second, agency attrition between the pre- and post-IPS periods required that we recruit replacement agencies during the post-IPS period. We included agency-level control variables in the main analysis to control for this additional source of variation. However, if this agency-level variation is correlated with our outcomes and is not sufficiently controlled for with our explanatory variables, then it will bias our impact estimates. To minimize agency-level variation, we therefore estimated the impacts using a sample restricted to agencies that

\(^6\)In this case, size would be the number of observations in a given agency.
participated in both periods.\textsuperscript{7} One drawback of the constant agency sample composition is the lower statistical power due to the smaller sample size. It will be important to take this drawback into account when comparing significance levels from the full sample with those obtained using a constant agency sample composition.

Third, it is possible that the impacts estimated with the full sample understate the actual effect of IPS. Such an understatement of effects would occur if pre-IPS period agencies began changing their behavior in anticipation of imminent IPS restrictions. We analyzed this possibility by restricting the pre-IPS period to observations that occurred during the first year of the HHPP Demonstration (that is, agency fiscal year 1996-1997), which ended before IPS took effect in October 1997.\textsuperscript{8} If pre-IPS period agencies did indeed alter their behavior in anticipation of IPS, then we might expect to find impacts of larger magnitude when we restrict the pre-IPS period to year one only.

We present the sensitivity tests on an outcome-by-outcome basis in Tables III.8 through III.10. Presenting the tests in this way enables us to demonstrate how little the sign and significance of impact estimates change across specifications and, therefore, why we believe our results are robust. Detailed results of these sensitivity analyses are shown in Tables A.1 through A.5 (in Appendix A).

\textbf{a. IPS Impacts on Patient Functioning}

IPS impacts on patient functioning generally were robust to changes in specification, further reinforcing the conclusion that IPS did not have detrimental effects on patients’ functional status.

\textsuperscript{7}Estimation with this sample does not eliminate agency-level variation completely, however, because agencies may change over time in unobservable ways.

\textsuperscript{8}This restriction was not possible with the pre-IPS period of the survey, which was fielded entirely during year two of the HHPP Demonstration.
None of the OASIS outcomes (with the exception of the management of medications) yielded significant negative results in any specification (Table III.8). Although the change in specification had an effect on some of the results (specifically, on improvement in bathing, toileting, and management of medications and on stabilization in light meal preparation), these changes do not alter the overarching conclusion based on the OASIS analysis that IPS did not have detrimental effects on patients’ functional status.

Survey impacts were more likely than impacts from OASIS to change across specifications. In the case of did transfer, did ambulate, and could ambulate, for example, significant impacts became insignificant when estimated on a constant sample composition. However, the insignificant impacts for the constant sample composition had the same sign as those of the weighted full sample, suggesting that the lack of significance could partially be due to the smaller sample size (see Table A.3). As noted, it is unclear what portion of the survey impacts are due to changes in patient characteristics that were not controlled for, so it is difficult to draw definitive conclusions based on these impact estimates.

b. IPS Impacts on Medical Symptoms and Outcomes

Changes in specification had little effect on whether we detected an IPS impact on medical symptoms and outcomes. All significant impacts estimated with the weighted full sample, with the exception of improvement in dyspnea, remained significant for each alternate specification (Table III.9). The results for the unweighted full sample yielded a significant negative impact for four cases (improvement in the status of the most problematic surgical wound, pain, urinary

---

9The survey sample size for a constant agency sample composition is approximately 50 percent smaller than that of the full sample.
### TABLE III.8

**SENSITIVITY OF IPS IMPACTS ON ADL/IADL FUNCTIONAL LIMITATIONS, USING OASIS AND SURVEY DATA**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample</th>
<th>Unweighted, Full Sample</th>
<th>Weighted, Constant Sample Composition</th>
<th>Weighted, Pre-Period One Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)^a</td>
<td>(B)^b</td>
<td>(C)^c</td>
<td>(D)^d</td>
</tr>
</tbody>
</table>

#### ADL/IADL Functioning (OASIS)

**Grooming**
- Improvement: + + + +
- Stabilization: + + + +

**Bathing**
- Improvement: 0 0 + 0
- Stabilization: + + + +

**Toileting**
- Improvement: + 0 0 +
- Stabilization: + + + +

**Transferring**
- Improvement: 0 0 0 0
- Stabilization: + + + +

**Ambulating**
- Improvement: 0 0 0 0
- Stabilization: + + + +

**Preparing Light Meals**
- Improvement: 0 0 0 0
- Stabilization: + + + +

**Housekeeping**
- Improvement: 0 0 0 0
- Stabilization: 0 0 0 0

**Managing Medications**
- Improvement: 0 – 0 0
- Stabilization: + + + +

#### ADL/IADL Functioning (Survey)

**Bathing**
- Did bathe: – – – n.a.
- Could bathe: – – – n.a.
TABLE III.8 (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample</th>
<th>Unweighted, Full Sample</th>
<th>Weighted, Constant Sample Composition</th>
<th>Weighted, Pre-Period Contains Year One Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>(B)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>(C)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>(D)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Eating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did eat</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Could eat</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>n.a.</td>
</tr>
<tr>
<td>Transferring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did transfer</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Could transfer</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Ambulating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did ambulate</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Could ambulate</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Take Medications Independently</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did take medications independently</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>Could take medications independently</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

**Source:** OASIS and survey data.

**Note:** See tables in Appendix A for detailed results.

+ = estimated post-IPS period value exceeds estimated pre-IPS period value, p \( \leq .10 \).

− = estimated post-IPS period value falls below estimated pre-period value, p \( \leq .10 \).

0 = estimated post-IPS period value not significantly different from estimated pre-IPS period value, p = .10.

<sup>a</sup>Column A presents impacts on the weighted full sample.

<sup>b</sup>Column B presents impacts for the unweighted full sample.

<sup>c</sup>Column C presents weighted impacts for a constant agency sample composition.

<sup>d</sup>Column D presents weighted impacts with the pre-IPS period consisting of year one observations only.

ADL = activity of daily living; IADL = instrumental activity of daily living; n.a. = not applicable.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample</th>
<th>Unweighted, Full Sample</th>
<th>Weighted, Constant Sample Composition</th>
<th>Weighted, Pre-IPS Period Contains Year One Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Problematic Pressure Ulcer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Number of Pressure Ulcers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Status of Most Problematic Surgical Wound</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Stabilization</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>0</td>
<td>–</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Dyspnea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Urinary Incontinence or Catheter Present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>0</td>
<td>–</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Treated for Urinary Tract Infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Confusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>0</td>
<td>–</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Frequency of Behavioral Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stabilization</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**Source:** OASIS data.

**Note:** See tables in Appendix A for detailed results.
\textit{TABLE III.9 (continued)}

$+$ = estimated post-IPS period value exceeds estimated pre-IPS period value, $p \leq .10$.
$-$ = estimated post-IPS period value falls below estimated pre-period value, $p \leq .10$.
$0$ = estimated post-IPS period value not significantly different from estimated pre-IPS period value, $p = .10$.

\textsuperscript{a}Column A presents impacts on the weighted full sample.

\textsuperscript{b}Column B presents impacts for the unweighted full sample.

\textsuperscript{c}Column C presents weighted impacts for a constant agency sample composition.

\textsuperscript{d}Column D presents weighted impacts with the pre-IPS period consisting of year one observations only.
incontinence or catheter present, and confusion), but these results were significant in the remaining three specifications for only one outcome (improvement in the status of the most problematic surgical wound), suggesting little cause for concern.

c. IPS Impacts on General Health Status

The impact of IPS on self-reports of general health generally was consistent across specifications. For two measures (reported health good or excellent and satisfied with life), the significance level remained relatively constant for each specification (Table III.10). For the third measure (any days in bed within the past two weeks), a significant impact for the weighted full sample became insignificant when we restricted the sample to agencies that participated in both periods. However, the coefficient decreased only slightly when we used the weighted constant agency sample composition (see Table A.5). As was the case for survey impacts on functional limitations, the sample size decreased by approximately 50 percent when we used a constant agency sample composition. Therefore, the lack of significance for whether a patient was in bed during the two weeks before the survey may simply reflect the decrease in statistical power that accompanies the smaller sample size for a constant agency sample composition.

d. Effect of Agencies’ Anticipation of IPS on IPS Impacts

We found some evidence that pre-IPS period agencies may have begun altering their behavior in anticipation of IPS. This evidence was most convincing for the impacts on medical symptoms and outcomes, as all 10 significant impacts using the weighted full sample increased in magnitude at least slightly when the pre-IPS period contained year one observations only (see Table A.4). This finding suggests that the main impacts may understate the actual impact of IPS. The pattern of the impacts on functional limitations, although consistent with those for medical symptoms and outcomes, provides weaker evidence that the impact of IPS was understated.
TABLE III.10
SENSITIVITY OF IPS IMPACTS ON PATIENTS’ SELF-REPORTS OF GENERAL HEALTH

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Unweighted, Full Sample (B)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Weighted, Constant Sample Composition (C)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported Health Good or Excellent</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Any Days in Bed Within Past Two Weeks</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Satisfied with Life</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

**SOURCE:** Survey data.

**NOTE:** See tables in Appendix A for detailed results.

+ = estimated post-IPS period value exceeds estimated pre-IPS period value, p ≤ .10.
− = estimated post-IPS period value falls below estimated pre-period value, p ≤ .10.
0 = estimated post-IPS period value not significantly different from estimated pre-IPS period value, p = .10.

<sup>a</sup>Column A presents impacts on the weighted full sample.

<sup>b</sup>Column B presents impacts for the unweighted full sample.

<sup>c</sup>Column C presents weighted impacts for a constant agency sample composition.
Compared with estimates using the weighted full sample, using the sample restricting the pre-IPS period to only observations in year one caused the magnitude of six of the nine significant impacts on functional status to increase, whereas the magnitude of the remaining three decreased (see Tables A.1 and A.2).
IV. CONCLUSIONS

The Balanced Budget Amendment of 1997 dramatically altered the Medicare home health program. The most controversial requirement of this legislation was the imposition of financial limits on participating agencies. These limits, known as the Interim Payment System (IPS), led to both substantial decreases in the amount of home health care provided and substantial savings to the Medicare program. The important question is whether these savings resulted in poorer health outcomes for Medicare beneficiaries.

We did not find evidence of serious health consequences associated with IPS. The functional status of post-IPS period patients, despite their having more functional limitations when their episode began, was as likely to improve during their episode of care as was the functional status of pre-IPS period patients. Moreover, the condition of post-IPS period patients was slightly more likely than that of pre-IPS period patients to stabilize (that is, not worsen). Furthermore, patients were as likely to improve on three measures of instrumental activities of daily living (meal preparation, housekeeping, and medication management) during the post-IPS period as during the pre-IPS period. We did find some evidence from the survey analysis that IPS may have negatively affected patient's self-reported functioning (particularly, in ambulating), and this finding is cause for caution. Nevertheless, most of the evidence suggests that functioning was not adversely affected.

Measures of health status and symptoms also suggest that the IPS had little effect on patients’ health outcomes. The status of post-IPS period patients’ surgical wounds was significantly less likely to improve during the episode of care than was that of pre-IPS period patients’ surgical wounds, but the fact that we found no difference in improvement in the status of the most problematic pressure ulcer or number of pressure ulcers between the two periods
leads us to believe that post-IPS period patients’ shorter time in home health care, rather than poor treatment of wounds, explains this finding. Furthermore, our measures of six other health symptoms consistently showed that patients who were in care after IPS was implemented were more likely than patients in care before IPS to stabilize during their home health episode (that is, were less likely to worsen).

Three measures of overall health status suggest that IPS could have detrimental effects, but the evidence is weak. Although patients in the post-IPS period were as likely as patients in the pre-IPS period to report good overall health status, they were more likely to report having spent time in bed due to illness, and to feel less satisfied with life. The differences in the baseline patient characteristics that we measured did not account for these differences in patients’ self-reports. However, because IPS did not negatively affect our measures of health outcomes that are more directly related to home health service use, we are inclined to believe that the finding is the result of the changing nature of Medicare home health patients for which we were unable to control.

A. LIMITATIONS

This study has a number of limitations. Perhaps the most important limitation is the inherent weakness of a pre-post analysis. In this type of analysis, it is not possible to control for every factor that might have changed over time, and that could have affected health outcomes, so we always must be cautious in drawing conclusions. In our analysis, most of the impacts we estimated suggested that health outcomes improved, and we therefore conclude with confidence that patients were not harmed by IPS. Nevertheless, we hesitate to state that the payment limits actually led to better outcomes, as it is possible that other factors for which we were unable to control may have driven the results.
The study also is limited by the fact that the list of outcomes we measured was not exhaustive. We may have failed to measure outcomes that are more sensitive to decreases in the number of home health visits than were the ones we did measure. Even so, however, those outcomes probably affect only a small percentage of the Medicare home health population.

A third limitation relates to the small number of agencies that originally volunteered for the Medicare Home Health Prospective Payment (HHPP) Demonstration, and that ultimately agreed to participate in the post-IPS study: these agencies may not be representative of agencies nationwide. It is possible that only agencies that were confident about their quality of care agreed to participate in the study. If this type of self-selection of agencies occurred, then we might have failed to observe adverse outcomes in other, less confident agencies.

Fourth, it is possible that the nurses’ assessments reported in the OASIS data were biased. During the pre-IPS period, agencies had no incentive to complete the OASIS instrument in any particular way. However, if agencies were concerned that their early discharges during the post-IPS period would have resulted in scrutiny of their services, they may have instructed their nurses to ensure that the discharge forms reflected the best possible patient outcomes. To the extent that the nurses were able to “upcode” the status of patients without being untruthful, we might have failed to identify negative consequences of IPS.

Finally, missing data may have affected our results. We know that 30 percent of the cases in the OASIS data set had start-of-care forms that did not match any follow-up forms. Furthermore, we expect that the data failed to include some cases because the cases had no start-of-care form; we do not have an estimate of how often that failure occurred. Although evidence from the HHPP Demonstration suggests the missing data are not correlated with health outcomes, we cannot rule out the possibility that the occurrence of missing observations was in
fact correlated in that way. Analysis without the missing data therefore could have biased our results.

B. POLICY RECOMMENDATIONS

The evidence in this report shows that IPS did not seriously harm patients. Despite the large reductions in home health services use, we found little evidence that patient functioning or health outcomes suffered. Although we do not want to dismiss the possibility that IPS may have had some adverse consequences, it appears that agencies, when given the financial incentive to reduce services, do so safely, without compromising patient care.

The evidence also shows that the type of patient who received care during the post-IPS period differed significantly from the type receiving care during the pre-IPS period. Indeed, the changes we observed suggest that, although patients were functionally more limited at the start of care during the post-IPS period than during the pre-IPS period, they also were less likely to suffer from conditions that were associated with higher home health resource use under cost reimbursement.1 This finding suggests that the Centers for Medicare & Medicaid Services (CMS) should monitor access to home health services carefully to ensure that high-use patients are admitted to care. The prospective payment system (PPS) currently in use has mitigated this problem; unlike IPS, PPS does account for differences in the severity of patients’ conditions.

Finally, because CMS uses OASIS data to monitor the quality of care, more resources are necessary to monitor the data reports. Agencies have difficulty implementing the OASIS

---

1For example, relative to pre-IPS patients, post-IPS period patients in the OASIS sample had fewer incidences of diabetes and cerebrovascular disease (Table II.4), both of which were historically associated with high use of home health resources.
system, and we found that nearly one-third of the data were missing.\textsuperscript{2} Furthermore, emergency service use was greatly underreported. A more thorough system of checks and balances—especially one that is tied to the payment system—may help agencies comply with the data collection while assuring CMS that the lowest-quality cases are not systematically unreported.

\textsuperscript{2}Note that these data were from the beginning of the OASIS data collection. The problems may have resolved as agencies became familiar with the system.
REFERENCES


APPENDIX A

DETAILED RESULTS FOR SENSITIVITY ANALYSES
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)</th>
<th>Unweighted, Full Sample (B)</th>
<th>Weighted, Constant Sample Composition (C)</th>
<th>Weighted, Pre-IPS Period Contains Year One Only (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)^a</td>
<td>(B)^b</td>
<td>(C)^c</td>
<td>(D)^d</td>
</tr>
<tr>
<td>Grooming</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>17,465</td>
<td>17,465</td>
<td>8,595</td>
<td>6,468</td>
</tr>
<tr>
<td>Mean^e</td>
<td>59.24</td>
<td>59.22</td>
<td>58.43</td>
<td>56.84</td>
</tr>
<tr>
<td>Impact</td>
<td>5.97***</td>
<td>5.48**</td>
<td>5.14*</td>
<td>8.35***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.02)</td>
<td>(.07)</td>
<td>(.01)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>32,518</td>
<td>32,518</td>
<td>15,569</td>
<td>12,161</td>
</tr>
<tr>
<td>Mean^e</td>
<td>88.07</td>
<td>87.91</td>
<td>87.82</td>
<td>86.02</td>
</tr>
<tr>
<td>Impact</td>
<td>4.67***</td>
<td>3.58***</td>
<td>5.07***</td>
<td>6.30***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Bathing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27,301</td>
<td>27,301</td>
<td>13,456</td>
<td>10,598</td>
</tr>
<tr>
<td>Mean^e</td>
<td>55.71</td>
<td>56.08</td>
<td>53.20</td>
<td>54.45</td>
</tr>
<tr>
<td>Impact</td>
<td>2.36</td>
<td>1.22</td>
<td>4.21*</td>
<td>3.50</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.20)</td>
<td>(.47)</td>
<td>(.07)</td>
<td>(.15)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>32,781</td>
<td>32,781</td>
<td>15,763</td>
<td>12,251</td>
</tr>
<tr>
<td>Mean^e</td>
<td>85.16</td>
<td>85.03</td>
<td>84.85</td>
<td>85.14</td>
</tr>
<tr>
<td>Impact</td>
<td>3.58***</td>
<td>3.22***</td>
<td>4.44***</td>
<td>3.33***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Toileting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>10,902</td>
<td>10,902</td>
<td>5,315</td>
<td>4,270</td>
</tr>
<tr>
<td>Mean^e</td>
<td>63.02</td>
<td>62.74</td>
<td>62.39</td>
<td>62.43</td>
</tr>
<tr>
<td>Impact</td>
<td>4.03**</td>
<td>2.31</td>
<td>3.05</td>
<td>4.53*</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.05)</td>
<td>(.29)</td>
<td>(.24)</td>
<td>(.09)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>33,142</td>
<td>33,142</td>
<td>15,896</td>
<td>12,377</td>
</tr>
<tr>
<td>Mean^e</td>
<td>91.88</td>
<td>91.55</td>
<td>91.69</td>
<td>90.65</td>
</tr>
<tr>
<td>Impact</td>
<td>3.33***</td>
<td>3.37***</td>
<td>3.59***</td>
<td>4.43***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
</tbody>
</table>
### TABLE A.1 (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)(^{a})</th>
<th>Unweighted, Full Sample (B)(^{b})</th>
<th>Weighted, Constant Sample Composition (C)(^{c})</th>
<th>Weighted, Pre-IPS Period Contains Year One Only (D)(^{d})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transferring Improvement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>18,935</td>
<td>18,935</td>
<td>9,352</td>
<td>7,383</td>
</tr>
<tr>
<td>Mean</td>
<td>52.45</td>
<td>51.67</td>
<td>49.74</td>
<td>51.52</td>
</tr>
<tr>
<td>Impact</td>
<td>2.38</td>
<td>1.28</td>
<td>4.82</td>
<td>3.15</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.36)</td>
<td>(.64)</td>
<td>(.21)</td>
<td>(.34)</td>
</tr>
<tr>
<td><strong>Stabilization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>34,219</td>
<td>34,219</td>
<td>16,449</td>
<td>12,856</td>
</tr>
<tr>
<td>Mean</td>
<td>89.48</td>
<td>89.12</td>
<td>90.10</td>
<td>88.88</td>
</tr>
<tr>
<td>Impact</td>
<td>4.20***</td>
<td>3.80***</td>
<td>4.47***</td>
<td>4.62***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td><strong>Ambulating Improvement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>26,449</td>
<td>26,449</td>
<td>13,013</td>
<td>10,271</td>
</tr>
<tr>
<td>Mean</td>
<td>33.64</td>
<td>33.73</td>
<td>32.06</td>
<td>33.97</td>
</tr>
<tr>
<td>Impact</td>
<td>-0.43</td>
<td>-0.61</td>
<td>3.55*</td>
<td>-0.39</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.83)</td>
<td>(.69)</td>
<td>(.10)</td>
<td>(.89)</td>
</tr>
<tr>
<td><strong>Stabilization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>34,368</td>
<td>34,368</td>
<td>16,531</td>
<td>12,889</td>
</tr>
<tr>
<td>Mean</td>
<td>90.23</td>
<td>90.06</td>
<td>90.94</td>
<td>90.77</td>
</tr>
<tr>
<td>Impact</td>
<td>3.83***</td>
<td>3.27***</td>
<td>3.44***</td>
<td>3.01***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
</tbody>
</table>

**SOURCE:** OASIS data.

\(^{a}\)Column A presents impacts on the weighted full sample.

\(^{b}\)Column B presents impacts for the unweighted full sample.

\(^{c}\)Column C presents weighted impacts for a constant agency sample composition.
TABLE A.1 (continued)

d Column D presents weighted impacts with the pre-IPS period consisting of year one observations only.

e The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

* Significantly different from zero at the .10 level, two-tailed test
** Significantly different from zero at the .05 level, two-tailed test.
*** Significantly different from zero at the .01 level, two-tailed test.
TABLE A.2

SENSITIVITY OF OASIS IPS IMPACTS ON INSTRUMENTAL ACTIVITIES OF DAILY LIVING TO ALTERNATIVE SPECIFICATIONS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample</th>
<th>Unweighted, Full Sample</th>
<th>Weighted, Constant Sample Composition</th>
<th>Weighted, Pre-IPS Period Contains Year One Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)^a</td>
<td>(B)^b</td>
<td>(C)^c</td>
<td>(D)^d</td>
</tr>
<tr>
<td>Light Meal Preparation Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>22,290</td>
<td>22,290</td>
<td>10,998</td>
<td>8,763</td>
</tr>
<tr>
<td>Mean^c</td>
<td>52.68</td>
<td>52.50</td>
<td>51.66</td>
<td>52.10</td>
</tr>
<tr>
<td>Impact</td>
<td>1.09</td>
<td>-0.30</td>
<td>0.72</td>
<td>1.14</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.63)</td>
<td>(.88)</td>
<td>(.81)</td>
<td>(.71)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>25,585</td>
<td>25,585</td>
<td>12,068</td>
<td>9,223</td>
</tr>
<tr>
<td>Mean^c</td>
<td>85.20</td>
<td>84.08</td>
<td>83.65</td>
<td>84.91</td>
</tr>
<tr>
<td>Impact</td>
<td>2.96**</td>
<td>3.02***</td>
<td>3.98***</td>
<td>2.69</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.02)</td>
<td>(.00)</td>
<td>(.01)</td>
<td>(.14)</td>
</tr>
<tr>
<td>Housekeeping Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>32,174</td>
<td>32,174</td>
<td>15,626</td>
<td>12,288</td>
</tr>
<tr>
<td>Mean^c</td>
<td>43.02</td>
<td>43.47</td>
<td>41.35</td>
<td>42.83</td>
</tr>
<tr>
<td>Impact</td>
<td>-0.72</td>
<td>-1.79</td>
<td>-0.67</td>
<td>-1.24</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.67)</td>
<td>(.20)</td>
<td>(.72)</td>
<td>(.63)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20,039</td>
<td>20,039</td>
<td>9,512</td>
<td>6,964</td>
</tr>
<tr>
<td>Mean^c</td>
<td>75.75</td>
<td>74.70</td>
<td>75.69</td>
<td>74.48</td>
</tr>
<tr>
<td>Impact</td>
<td>1.72</td>
<td>1.85</td>
<td>2.84*</td>
<td>1.52</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.30)</td>
<td>(.20)</td>
<td>(.10)</td>
<td>(.53)</td>
</tr>
<tr>
<td>Management of Medications Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>17,365</td>
<td>17,365</td>
<td>8,363</td>
<td>6,575</td>
</tr>
<tr>
<td>Mean^c</td>
<td>44.69</td>
<td>44.88</td>
<td>42.66</td>
<td>44.94</td>
</tr>
<tr>
<td>Impact</td>
<td>-0.63</td>
<td>-3.58**</td>
<td>-1.01</td>
<td>-2.88</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.76)</td>
<td>(.05)</td>
<td>(.63)</td>
<td>(.37)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27,343</td>
<td>27,343</td>
<td>13,019</td>
<td>10,132</td>
</tr>
<tr>
<td>Mean^c</td>
<td>87.28</td>
<td>86.52</td>
<td>86.40</td>
<td>86.49</td>
</tr>
<tr>
<td>Impact</td>
<td>2.48***</td>
<td>2.73***</td>
<td>3.22***</td>
<td>2.74*</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.06)</td>
</tr>
</tbody>
</table>
**TABLE A.2 (continued)**

| Column A | Presents impacts on the weighted full sample. |
| Column B | Presents impacts for the unweighted full sample. |
| Column C | Presents weighted impacts for a constant agency sample composition. |
| Column D | Presents weighted impacts with the pre-IPS period consisting of year one observations only. |
| Column E | The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS. |

*Significantly different from zero at the .10 level, two-tailed test.  
**Significantly different from zero at the .05 level, two-tailed test.  
***Significantly different from zero at the .01 level, two-tailed test.

**SOURCE:** OASIS data.
### TABLE A.3

SENSITIVITY OF OASIS IPS IMPACTS ON PATIENT OUTCOMES TO ALTERNATIVE SPECIFICATIONS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)</th>
<th>Unweighted, Full Sample (B)</th>
<th>Weighted, Constant Sample Composition (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)</td>
<td>(B)</td>
<td>(C)</td>
</tr>
<tr>
<td><strong>Activities of Daily Living</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bathing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did bathe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,285</td>
<td>2,285</td>
<td>1,086</td>
</tr>
<tr>
<td>Mean$^d$</td>
<td>56.57</td>
<td>54.67</td>
<td>5.09</td>
</tr>
<tr>
<td>Impact</td>
<td>-7.01***</td>
<td>-3.24*</td>
<td>-6.81*</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.10)</td>
<td>(.08)</td>
</tr>
<tr>
<td>Could bathe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,268</td>
<td>2,268</td>
<td>1,076</td>
</tr>
<tr>
<td>Mean$^d$</td>
<td>64.26</td>
<td>63.61</td>
<td>64.51</td>
</tr>
<tr>
<td>Impact</td>
<td>-5.91**</td>
<td>-4.10***</td>
<td>-6.97**</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.02)</td>
<td>(.01)</td>
<td>(.04)</td>
</tr>
<tr>
<td><strong>Eating</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did eat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,254</td>
<td>2,254</td>
<td>1,065</td>
</tr>
<tr>
<td>Mean$^d$</td>
<td>82.23</td>
<td>80.93</td>
<td>78.25</td>
</tr>
<tr>
<td>Impact</td>
<td>-2.99</td>
<td>-1.40</td>
<td>-0.09</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.25)</td>
<td>(.51)</td>
<td>(.98)</td>
</tr>
<tr>
<td>Could eat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,246</td>
<td>2,246</td>
<td>1,060</td>
</tr>
<tr>
<td>Mean$^d$</td>
<td>92.27</td>
<td>91.45</td>
<td>92.05</td>
</tr>
<tr>
<td>Impact</td>
<td>-2.82*</td>
<td>-1.90</td>
<td>-3.58*</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.08)</td>
<td>(.21)</td>
<td>(.06)</td>
</tr>
<tr>
<td><strong>Transferring</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did transfer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,292</td>
<td>2,292</td>
<td>1,089</td>
</tr>
<tr>
<td>Mean$^d$</td>
<td>68.42</td>
<td>66.96</td>
<td>64.24</td>
</tr>
<tr>
<td>Impact</td>
<td>-5.84***</td>
<td>-4.79***</td>
<td>-2.49</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.01)</td>
<td>(.44)</td>
</tr>
<tr>
<td>Could transfer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,280</td>
<td>2,280</td>
<td>1,082</td>
</tr>
<tr>
<td>Mean$^d$</td>
<td>82.35</td>
<td>82.16</td>
<td>80.90</td>
</tr>
<tr>
<td>Impact</td>
<td>-1.78</td>
<td>-1.59</td>
<td>-0.50</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.39)</td>
<td>(.40)</td>
<td>(.87)</td>
</tr>
<tr>
<td>Outcome</td>
<td>Weighted, Full Sample (A)</td>
<td>Unweighted, Full Sample (B)</td>
<td>Weighted, Constant Sample Composition (C)</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ambulating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did ambulate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,226</td>
<td>2,226</td>
<td>1,052</td>
</tr>
<tr>
<td>Mean</td>
<td>68.06</td>
<td>67.35</td>
<td>65.62</td>
</tr>
<tr>
<td>Impact</td>
<td>-6.10**</td>
<td>-4.64***</td>
<td>-4.21</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.02)</td>
<td>(.01)</td>
<td>(.22)</td>
</tr>
<tr>
<td>Could ambulate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,216</td>
<td>2,216</td>
<td>1,047</td>
</tr>
<tr>
<td>Mean</td>
<td>79.74</td>
<td>80.06</td>
<td>78.66</td>
</tr>
<tr>
<td>Impact</td>
<td>-6.03**</td>
<td>-6.42***</td>
<td>-5.35</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.02)</td>
<td>(.00)</td>
<td>(.19)</td>
</tr>
</tbody>
</table>

**Instrumental Activities of Daily Living**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)</th>
<th>Unweighted, Full Sample (B)</th>
<th>Weighted, Constant Sample Composition (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Take Medications Independently</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did take medications independently</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,250</td>
<td>2,250</td>
<td>1,073</td>
</tr>
<tr>
<td>Mean</td>
<td>52.62</td>
<td>51.65</td>
<td>47.89</td>
</tr>
<tr>
<td>Impact</td>
<td>-0.76</td>
<td>0.13</td>
<td>3.50</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.73)</td>
<td>(.94)</td>
<td>(.15)</td>
</tr>
<tr>
<td>Could take medications independently</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,236</td>
<td>2,236</td>
<td>1,065</td>
</tr>
<tr>
<td>Mean</td>
<td>69.89</td>
<td>69.56</td>
<td>65.02</td>
</tr>
<tr>
<td>Impact</td>
<td>1.95</td>
<td>1.80</td>
<td>7.17***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.35)</td>
<td>(.31)</td>
<td>(.00)</td>
</tr>
</tbody>
</table>

**SOURCE:** OASIS data.

\(^{a}\)Column A presents impacts on the weighted full sample.

\(^{b}\)Column B presents impacts for the unweighted full sample.

\(^{c}\)Column C presents weighted impacts for a constant agency sample composition.

\(^{d}\)Column D presents weighted impacts with the pre-IPS period consisting of year one observations only.

\(^{e}\)The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.
TABLE A.3 (continued)

*Significantly different from zero at the .10 level, two-tailed test
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.
TABLE A.4
SENSITIVITY OF OASIS IPS IMPACTS ON MEDICAL SYMPTOMS AND OUTCOMES TO ALTERNATIVE SPECIFICATIONS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample</th>
<th>Unweighted, Full Sample</th>
<th>Weighted, Constant Sample Composition</th>
<th>Weighted, Pre-IPS Period Contains Year One Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>(B)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>(C)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>(D)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Most Problematic Pressure Ulcer Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,277</td>
<td>2,277</td>
<td>974</td>
<td>725</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;e&lt;/sup&gt;</td>
<td>76.52</td>
<td>76.91</td>
<td>75.34</td>
<td>79.67</td>
</tr>
<tr>
<td>Impact</td>
<td>3.16</td>
<td>4.39</td>
<td>2.52</td>
<td>0.95</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.45)</td>
<td>(.18)</td>
<td>(.68)</td>
<td>(.86)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,603</td>
<td>34,603</td>
<td>16,651</td>
<td>12,897</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;e&lt;/sup&gt;</td>
<td>97.41</td>
<td>97.46</td>
<td>97.49</td>
<td>97.23</td>
</tr>
<tr>
<td>Impact</td>
<td>1.78***</td>
<td>1.62***</td>
<td>1.70***</td>
<td>1.89***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Number of Pressure Ulcers Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,022</td>
<td>2,022</td>
<td>882</td>
<td>631</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;e&lt;/sup&gt;</td>
<td>76.63</td>
<td>77.55</td>
<td>76.06</td>
<td>79.72</td>
</tr>
<tr>
<td>Impact</td>
<td>3.74</td>
<td>3.69</td>
<td>0.51</td>
<td>1.47</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.28)</td>
<td>(.20)</td>
<td>(.92)</td>
<td>(.77)</td>
</tr>
<tr>
<td>Stabilization</td>
<td>34,835</td>
<td>34,835</td>
<td>16,767</td>
<td>13,025</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;e&lt;/sup&gt;</td>
<td>97.69</td>
<td>97.75</td>
<td>97.50</td>
<td>97.60</td>
</tr>
<tr>
<td>Impact</td>
<td>1.52***</td>
<td>1.33***</td>
<td>1.69***</td>
<td>1.56***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.01)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Status of Most Problematic Surgical Wound Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>7,986</td>
<td>7,986</td>
<td>3,954</td>
<td>3,367</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;e&lt;/sup&gt;</td>
<td>83.29</td>
<td>82.71</td>
<td>83.85</td>
<td>84.36</td>
</tr>
<tr>
<td>Impact</td>
<td>-6.30***</td>
<td>-6.28***</td>
<td>-7.26***</td>
<td>-7.53***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.01)</td>
<td>(.01)</td>
</tr>
</tbody>
</table>
TABLE A.4 (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)</th>
<th>Unweighted, Full Sample (B)</th>
<th>Weighted, Constant Sample Composition (C)</th>
<th>Weighted, Pre-IPS Period Contains Year One Only (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status of Most Problematic Surgical Wound (continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>34,396</td>
<td>34,396</td>
<td>16,556</td>
<td>12,815</td>
</tr>
<tr>
<td>Mean</td>
<td>97.50</td>
<td>97.37</td>
<td>97.67</td>
<td>96.95</td>
</tr>
<tr>
<td>Impact</td>
<td>-0.34</td>
<td>-0.28</td>
<td>0.07</td>
<td>0.35</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.30)</td>
<td>(.26)</td>
<td>(.81)</td>
<td>(.46)</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>21,846</td>
<td>21,846</td>
<td>10,475</td>
<td>8,206</td>
</tr>
<tr>
<td>Mean</td>
<td>54.33</td>
<td>55.74</td>
<td>54.09</td>
<td>51.81</td>
</tr>
<tr>
<td>Impact</td>
<td>-0.64</td>
<td>-3.55*</td>
<td>-2.36</td>
<td>1.97</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.74)</td>
<td>(.06)</td>
<td>(.40)</td>
<td>(.46)</td>
</tr>
<tr>
<td>Dyspnea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20,917</td>
<td>20,917</td>
<td>10,112</td>
<td>7,654</td>
</tr>
<tr>
<td>Mean</td>
<td>49.16</td>
<td>51.00</td>
<td>49.41</td>
<td>46.38</td>
</tr>
<tr>
<td>Impact</td>
<td>7.03***</td>
<td>2.31</td>
<td>3.91</td>
<td>9.61***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.30)</td>
<td>(.21)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>34,060</td>
<td>34,060</td>
<td>16,402</td>
<td>12,781</td>
</tr>
<tr>
<td>Mean</td>
<td>81.00</td>
<td>80.88</td>
<td>80.62</td>
<td>79.74</td>
</tr>
<tr>
<td>Impact</td>
<td>5.15***</td>
<td>4.08***</td>
<td>5.53***</td>
<td>5.59***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
</tbody>
</table>
TABLE A.4 (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample</th>
<th>Unweighted, Full Sample</th>
<th>Weighted, Constant Sample Composition</th>
<th>Weighted, Pre-IPS Period Contains Year One Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)(^a)</td>
<td>(B)(^b)</td>
<td>(C)(^c)</td>
<td>(D)(^d)</td>
</tr>
<tr>
<td>Urinary Incontinence or Catheter Present Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>7,490</td>
<td>7,490</td>
<td>3,840</td>
<td>3,043</td>
</tr>
<tr>
<td>Mean(^e)</td>
<td>54.43</td>
<td>55.75</td>
<td>50.55</td>
<td>54.52</td>
</tr>
<tr>
<td>Impact</td>
<td>-3.24</td>
<td>-5.87**</td>
<td>-3.19</td>
<td>-2.83</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.21)</td>
<td>(.02)</td>
<td>(.42)</td>
<td>(.27)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>33,380</td>
<td>33,380</td>
<td>16,077</td>
<td>12,524</td>
</tr>
<tr>
<td>Mean(^e)</td>
<td>95.46</td>
<td>95.44</td>
<td>94.96</td>
<td>94.76</td>
</tr>
<tr>
<td>Impact</td>
<td>1.30***</td>
<td>1.07**</td>
<td>2.01***</td>
<td>2.07***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.02)</td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Treated for a Urinary Tract Infection Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,393</td>
<td>2,393</td>
<td>1,101</td>
<td>959</td>
</tr>
<tr>
<td>Mean(^e)</td>
<td>79.60</td>
<td>80.01</td>
<td>80.00</td>
<td>75.94</td>
</tr>
<tr>
<td>Impact</td>
<td>0.96</td>
<td>0.01</td>
<td>-2.63</td>
<td>4.82</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.78)</td>
<td>(.99)</td>
<td>(.62)</td>
<td>(.16)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>30,913</td>
<td>30,913</td>
<td>14,847</td>
<td>11,615</td>
</tr>
<tr>
<td>Mean(^e)</td>
<td>98.22</td>
<td>98.20</td>
<td>97.86</td>
<td>98.11</td>
</tr>
<tr>
<td>Impact</td>
<td>0.37*</td>
<td>0.37**</td>
<td>0.71*</td>
<td>0.40</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.10)</td>
<td>(.03)</td>
<td>(.07)</td>
<td>(.24)</td>
</tr>
<tr>
<td>Confusion Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>11,833</td>
<td>11,833</td>
<td>5,540</td>
<td>4,488</td>
</tr>
<tr>
<td>Mean(^e)</td>
<td>48.17</td>
<td>48.92</td>
<td>44.08</td>
<td>48.92</td>
</tr>
<tr>
<td>Impact</td>
<td>-1.40</td>
<td>-4.66**</td>
<td>-1.29</td>
<td>-2.07</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.57)</td>
<td>(.05)</td>
<td>(.67)</td>
<td>(.50)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>31,536</td>
<td>31,536</td>
<td>15,071</td>
<td>11,727</td>
</tr>
<tr>
<td>Mean(^e)</td>
<td>86.29</td>
<td>86.07</td>
<td>84.21</td>
<td>85.00</td>
</tr>
<tr>
<td>Impact</td>
<td>2.92***</td>
<td>1.11</td>
<td>4.05***</td>
<td>3.12***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.19)</td>
<td>(.00)</td>
<td>(.01)</td>
</tr>
</tbody>
</table>
TABLE A.4 (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Unweighted, Full Sample (B)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Weighted, Constant Sample Composition (C)&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Weighted, Pre-IPS Period Contains Year One Only (D)&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of Behavior Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>3,906</td>
<td>3,906</td>
<td>1,880</td>
<td>1,546</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;e&lt;/sup&gt;</td>
<td>77.10</td>
<td>75.73</td>
<td>73.65</td>
<td>81.30</td>
</tr>
<tr>
<td>Impact</td>
<td>0.83</td>
<td>-2.88</td>
<td>1.30</td>
<td>-4.12</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.78)</td>
<td>(.26)</td>
<td>(.78)</td>
<td>(.20)</td>
</tr>
<tr>
<td>Stabilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>33,420</td>
<td>33,420</td>
<td>15,966</td>
<td>12,483</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;e&lt;/sup&gt;</td>
<td>92.99</td>
<td>92.56</td>
<td>90.91</td>
<td>92.11</td>
</tr>
<tr>
<td>Impact</td>
<td>2.91***</td>
<td>2.34**</td>
<td>4.73***</td>
<td>3.10***</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.00)</td>
<td>(.05)</td>
<td>(.00)</td>
<td>(.01)</td>
</tr>
</tbody>
</table>

**Source:** OASIS data.

<sup>a</sup>Column A presents impacts on the weighted full sample.

<sup>b</sup>Column B presents impacts for the unweighted full sample.

<sup>c</sup>Column C presents weighted impacts for a constant agency sample composition.

<sup>d</sup>Column D presents weighted impacts with the pre-IPS period consisting of year one observations only.

<sup>e</sup>The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

*Significantly different from zero at the .10 level, two-tailed test

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.
TABLE A.5
SENSITIVITY OF SURVEY IPS IMPACTS ON PATIENTS’ SELF-REPORTS
OF GENERAL HEALTH TO ALTERNATIVE SPECIFICATIONS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Weighted, Full Sample (A)</th>
<th>Unweighted, Full Sample (B)</th>
<th>Weighted, Constant Sample Composition (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported Health Good or Excellent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,262</td>
<td>2,262</td>
<td>1,074</td>
</tr>
<tr>
<td>Mean(^d)</td>
<td>42.84</td>
<td>41.80</td>
<td>41.74</td>
</tr>
<tr>
<td>Impact</td>
<td>-2.71</td>
<td>-1.36</td>
<td>-2.98</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.33)</td>
<td>(.62)</td>
<td>(.43)</td>
</tr>
<tr>
<td>Any Days in Bed Within Past Two Weeks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,347</td>
<td>2,347</td>
<td>1,126</td>
</tr>
<tr>
<td>Mean(^d)</td>
<td>27.53</td>
<td>28.46</td>
<td>28.81</td>
</tr>
<tr>
<td>Impact</td>
<td>6.50***</td>
<td>5.72***</td>
<td>4.36</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.01)</td>
<td>(.31)</td>
</tr>
<tr>
<td>Satisfied with Life</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,197</td>
<td>2,197</td>
<td>1,049</td>
</tr>
<tr>
<td>Mean(^d)</td>
<td>62.54</td>
<td>61.99</td>
<td>63.51</td>
</tr>
<tr>
<td>Impact</td>
<td>-6.75***</td>
<td>-4.85**</td>
<td>-7.91*</td>
</tr>
<tr>
<td>p-Value</td>
<td>(.01)</td>
<td>(.04)</td>
<td>(.06)</td>
</tr>
</tbody>
</table>

**SOURCE**: OASIS data.

\(^a\)Column A presents impacts on the weighted full sample.

\(^b\)Column B presents impacts for the unweighted full sample.

\(^c\)Column C presents weighted impacts for a constant agency sample composition.

\(^d\)Column D presents weighted impacts for a constant agency sample composition.

\(^e\)The regression-adjusted estimate of what the outcome variable mean would have been in the absence of IPS.

*Significantly different from zero at the .10 level, two-tailed test

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.