

Medicare Payment Adjustment for Inpatient Psychiatric Facilities: A Review and Potential Refinements

Final Report

First submitted: July 5, 2007

This version: August 22, 2009

Bowen Garrett, Ph.D. (Principal Investigator)

Adela Luque, Ph.D., Korbin Liu, Sc.D., Noga Lewin, M.A., M.S.

Federal Project Officer: Janet Samen, Director, Division of Chronic Care Management

Original Project Officer: Carolyn Rimes, Ph.D.

The Urban Institute

CMS Contract No. 500-00-0025/Task Order #0002, Mod 11

Our colleague and coauthor, Korbin Liu, died in July 2007. We are deeply grateful to Korbin Liu for his friendship, leadership, and insightful research on Medicare, post-acute care, and long term care. We greatly miss the depth and breadth of knowledge that he brought to his work.

Research in this report was supported by the Centers for Medicare and Medicaid Services (CMS) Contract No. 500-00-0025/Task Order #0002, Mod 11 (UI #07108-002-02) "Analyze, Assess and Replicate Data and Analysis for the Psychiatric Hospitals and Exempt Units PPS." The authors are grateful to Emerald Adams, Saad Ahmad, and Paul Masi for excellent research assistance. We thank Douglas Wissoker for many helpful comments and providing internal review on an earlier draft of this report. Any remaining errors are the responsibility of the authors. Inferences and opinions in this report are solely those of the authors and should not be attributed to the Urban Institute, its trustees, its funders, or CMS.

Contents

Executive Summary	i
Findings.....	iii
Conclusion	vii
I. Introduction and Background	1
A. History and Key Features of the Inpatient Psychiatric Facility Prospective Payment System.....	1
B. Considerations in the Design of the IPF PPS.....	3
C. Explaining the Costs of Inpatient Psychiatric Care	4
II. Data and Methods.....	10
A. Data Sources	10
B. Data Exclusions	12
Facility-level Flags.....	12
Stay-level Flags.....	13
C. Summary Statistics for Total, Routine, and Ancillary Costs per Day	15
D. Stay-level Regression Models of IPF Costs per Day.....	16
E. Facility-level Regressions of Average Cost per Day	21
Facility Case-mix Indices	22
Facility-level Model R-squared	24
The CMI Coefficient and CMI Compression	24
Interpreting Compression Results for the Payment and Fully-specified Models	25
F. Consequences of Measuring Routine Cost per Day Only at the Facility Level	26
III. Findings.....	29
A. Variable Frequencies	29
B. Replication of CMS's Regression Model of IPF Cost per Day	33
Overall Findings.....	33
Difference for the DRG 432 (Other Mental Disorders) Adjustment	37
Difference for the Oncology Comorbidity Adjustment	37
C. Sources of Explanatory Power in CMS's Payment Regression.....	39
D. Issues with Adjustment for the Day of Stay	40
E. Evaluation of Alternative Approaches to Predicting Stay-level IPF Costs per Day.....	43
The Distribution of Elixhauser Comorbidities in IPF Patients	43
The Explanatory Power of Alternative Model Specifications	44
F. Results for the Preferred Model.....	50
Summary Evaluation Statistics	50
Detailed Findings for the Preferred Model	51
G. Facility-level Analysis	56
Case-mix Indices by Facility Type	56
Facility-level Regression Results.....	57
IV. Conclusion	62
References.....	64
Appendix I. Details on Dataset Construction	66
Construction of the Dependent Variable (Log of per diem stay cost)	66
Outliers and Data Exclusions.....	67
Facility-level Flags.....	67

Stay-level Flags:.....	69
Appendix II. Details of Models Including Information from Past Stays.....	72
Appendix III. Technical Details on Facility-level Regressions, Case-mix Indices, and CMI Compression	77
The CMI Coefficient and CMI Compression	77
Interpreting Compression Results for the Payment and Fully-specified Models	78

Tables

Table 1. Match of Data Sources for FY 2003	12
Table 2. Facility-level Flags Used in Analysis of FY 2003 Data	13
Table 3. Stay-level Flags Used in Analysis of FY 2003 Data	14
Table 4. Number and Percentage of Stays and Facilities Before and After Data Exclusions (FY 2003)	15
Table 5. Summary Statistics for Total Costs per Day and its Components (FY 2003)	16
Table 6. Distribution of Stays by Explanatory Variable Category (FY 2003)	31
Table 7. Replication of CMS' Final Rule Results with FY 2003 Data.....	35
Table 8. Sources of Explanatory Power in CMS's Payment Regression Using FY 2003 Data ...	40
Table 9. Length of Stay vs. Day of Stay Adjustment	42
Table 10. Percent of Medicare IPF Patients with Elixhauser Comorbidities (FY 2003).....	44
Table 11. Explanatory Power of Alternative Model Specifications	45
Table 12. Summary Evaluation Statistics of the Current CMS IPF PPS Model and the Preferred Model (FY 2003)	51
Table 13. Comparison of Relative Weights in the Current and Preferred Models (FY 2003)	53
Table 14. CMIs by Facility and Ownership Type in the Preferred Model Specification (FY 2003)	57
Table 15. Facility-level Explanatory Power and the Proportionality of Payments with Costs.....	58
Appendix Table I.1. Facility-level Flags Used in Analysis of FY 2003 Data	69
Appendix Table I.2. Stay-level Flags Used in Analysis of FY 2003 Data	71
Appendix Table I.3. Number and Percentage of Stays and Facilities Before and After Data Exclusions (FY 2003)	71
Appendix Table II.1. Estimation Results Using Information from Past Stays	74
Appendix Table III.1. CMI Coefficient Scenarios in Payment and Fully-specified Facility-level Models and their Implications	80

Medicare Payment Adjustment for Inpatient Psychiatric Facilities: A Review and Potential Refinements

Executive Summary

In 2005, the Centers for Medicaid and Medicare Services (CMS) implemented a prospective payment system (PPS) to pay for Medicare services provided by inpatient psychiatric facilities (IPFs). The Balanced Budget Refinement Act of 1999 (BBRA), which mandated the switch to a PPS, also required that the PPS use a per day rather than per stay unit of payment and that it reflect differences in IPF patient service use and costs. Prior to PPS implementation, IPFs had been paid according to reasonable costs per discharge, subject to limits established by the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA). IPFs had been exempted from the PPS established for acute care hospitals by the Social Security Amendments of 1983.

The key components of the IPF PPS include 1) a base rate of \$595.09 (in 2007), 2) a patient classification system that groups patients according to their psychiatric DRG and 17 separate comorbidity categories, 3) payment adjustments for the DRG, comorbidities, age, day of stay, and patients who receive electro-convulsive treatment, and 4) facility-level payment adjustments for area wage levels, cost of living adjustment (COLA) for IPFs located in Alaska and Hawaii, rural location, teaching hospital status, and maintenance of an emergency department. The IPF PPS also includes an outlier policy with a 2 percent outlier pool and a stop-loss policy that guarantees the total payments to facilities under the IPF PPS are no less than 70 percent of their pre-PPS payments under TEFRA.

In shifting from cost-based reimbursement to a PPS, policy makers expect to increase incentives for efficient provision of care. The payment adjustments are meant to account for

clinically appropriate variation in resource use across different types of patients. The facility-level adjustments are meant to reflect cost variation across facilities (net of patient-level adjustments) that are considered to be appropriate or which are designed to offset the costs of desirable activities (e.g., maintaining an emergency department).

An important question for any PPS is whether its patient classification system and corresponding payment adjustments account sufficiently for variation in patient costs. Insufficient adjustments could create strong incentives for providers to stint on care during a stay and avoid admitting patients whose costs would not be fully compensated. Success in this regard depends on the ability of the patient classification system to capture predictable variation in costs at the patient level and matching adjusted payments and costs at the facility level.

This study re-examines the recently developed IPF PPS using an independently-constructed database of IPF stays in fiscal year 2003 (FY 2003). We replicate the regression analyses underlying the IPF PPS, as described in the Federal Register (2003 Preliminary Rule, 2004 Final Rule and corrections published in 2005) and compare our findings with those of the IPF PPS which was based on FY 2002 data.

In the study, we also consider alternative comorbidity classifiers and explore using the history of past inpatient stays to improve cost prediction and suggest new payment adjustors. We conduct the first facility-level analyses of the IPF-PPS in order to understand the relationship between IPF PPS payments and facility-level costs. The facility-level analyses provide information on the way the PPS distributes payments across facilities and the strength of incentives for facilities to select against more or less costly patients.

Findings

Our replication of the regression analyses that form the basis of the patient-specific adjusters using the FY 2003 data produce payment weights that are highly consistent with the payment weights in the current IPF PPS. This result provides an independent validation for the analysis used to construct the weights. One notable exception is for the oncology comorbidity. The current payment weight is 1.07 whereas we estimate a payment weight of 1.28. While only 0.01 percent of IPF patients (54 stays in FY 2003) meet the criteria for the oncology comorbidity, this finding suggests that Medicare payments for these patients are nearly 20 percent underpaid.

The overall R-squared for the stay-level regression producing the payment weights is 32.5 percent. It must be noted, however, that much of this explanatory power arises not from patient level variables used for payment adjustment, but by other patient and facility variables included in the model. If the DRGs, age categories, comorbidity categories, and length of stay categories are dropped from the model, the R-squared falls to 23.1 percent. The difference, 9.4, is the percent of variance explanation that can be uniquely attributed to the patient-level adjustments. Among the patient-level variables used for adjustment, the length of stay variables provide the largest amount of unique variance explanation, followed in order by age categories, DRGs, and comorbidities. The R-squared falls only slightly to 31.9 percent after excluding the current comorbidities.

The current payment system makes day of stay payment adjustments equivalent to the expected percentage differences in costs by length of stay that were estimated from stay-level

payment regression. We find that this is not the most logical or effective use of the estimated differences in costs by length of stay. The current approach results in less variability in payments by day of stay or length of stay than is indicated by the regression analysis, but this concern is easily remedied. All that is needed is to take the adjustment factors that are currently applied to each day of stay and apply them instead to average per day payment for the entire stay according to the length of the stay. This revision would more effectively match payments to expected patient costs.

The IPF PPS comorbidities as currently specified apply to a small percentage of stays. The most common comorbidity used for adjustment is Developmental Disability, making up about 3 percent of cases. This suggests the possibility that using an additional or alternate classification system that captures more common conditions that are also relevant to costs could substantively improve the ability of the PPS to match payments with resource use.

To this end, we examined the use of the Elixhauser comorbidity classification system. Because research using administrative datasets often examines resource use or clinical outcomes, Elixhauser et al. (1998) developed the comorbidity measures to predict hospital charges, length of stay, and in-hospital mortality. These comorbidities have also shown to be predictive of costs in several settings (see, for instance, Elixhauser et al. 1998, Southern 2004, Carney et al. 2006). We also consider a comorbidity classification system developed for IPFs by Research Triangle Institute (RTI).

We find that using the Elixhauser comorbidities rather than the current comorbidities boosts the R-squared from 0.325 to 0.356—a meaningful increase that relies solely on

administrative data already being used within the current system. We find a smaller gain in R-squared, from 0.325 to 0.339, when we use the RTI comorbidities in place of the current ones.

We report the DRG, age category, and length of stay adjustments that would result by using the Elixhauser comorbidities and compare them to the corresponding adjustments using the current comorbidities. We find very little difference, suggesting that shifting to an alternative set of comorbidity adjusters would not require substantial revision of the other weights. We did notice, however, that expected costs increase somewhat less with age after controlling for the Elixhauser comorbidities. This suggests that the Elixhauser comorbidities are picking up the types of conditions that become more frequent with increasing age and lead to increased costs in the IPF.

We attempt to maximize explanatory power by including information from past stays including DRGs and comorbidities from past stays, and type of past stays. We expected that DRGs and comorbidities of past hospitalizations for medical or other conditions or having prior stays in IPFs would be useful in predicting the cost per day of the current stay. We find that including all the past stay variables we constructed increases the R-squared of the CMS base model from 0.325 to 0.345—a modest increase in light of the costs of collecting such data and using it for payment adjustment. This indicates there is little to be gained from incorporating information available in past stays into the patient classification system.

We examine the relationship between PPS payments, measured by a case-mix index (CMI), and costs at the facility level for both the current system and for the payments weights that would result from our preferred model. Under the current system, we find that payments to

facilities are compressed in the sense that facilities with a case-mix that results in 10 percent higher payments (i.e., a CMI of 1.10) are estimated to have 32 percent higher costs. This indicates that facilities with high cost patients are not fully compensated for treating high cost patients and facilities with lower cost patients are overcompensated relative to their case-mix. This situation is referred to as “CMI compression.”

Payments are more proportionate to costs after we control for facility characteristics such as facility type, size, ownership type, and the occupancy rate. In our fully-specified model of facility-level costs, we find that facilities with 10 percent higher payments experience 15.5 percent higher costs. Although the degree of CMI compression is much reduced in the fully-specified models, it indicates that facilities face incentives to avoid high cost patients under the current system.

The degree of CMI compression is lower in our preferred model using the Elixhauser comorbidities. In the fully-specified model, we find that facilities with 10 percent higher payments would experience 13.4 percent higher costs. While CMI compression would remain, we find that incentives to select against more costly payments can be reduced, and payment equity improved, by using a more powerful comorbidity classification system as exemplified by the Elixhauser comorbidities.

The main source of the remaining CMI compression, in our view, is the aggregated nature of the routine cost data. Routine costs are only measured at the facility level and make up 86 percent of total costs per day for a typical stay. The aggregated routine costs lead to log cost regression coefficients that are biased towards zero (and hence payment weights that biased

towards one), which in turn leads to CMIs and payments that vary less than proportionately with expected costs. Methods to counter the expected bias and decompress the relative weights warrant consideration. Such methods could be as simple as a uniform expansion of all payment weights by some factor, the magnitude of which may be decided on the basis of the facility-level analyses we have presented. We believe that research on statistical methods that could reduce the aggregation bias in this context would also be fruitful.

Conclusion

Our findings largely confirm the analyses that form the basis of the patient-level adjustments in the IPF PPS. We noted an exception in the case of the oncology comorbidity and we propose an alternative method for making length of stay adjustments. We found that payments vary less than proportionately with costs at the facility level—a finding that raises concerns about the adequacy of payments to facilities that treat a costlier than average patient mix. We demonstrate how an expansion of the current set of comorbidities using an existing classification system would more closely match payments to patient costs at both the patient- and facility-levels without imposing new administrative requirements, and would thereby improve the overall functioning of the PPS.

I. Introduction and Background

A. History and Key Features of the Inpatient Psychiatric Facility Prospective Payment System

In 2005, the Centers for Medicare and Medicaid Services (CMS) implemented a prospective payment system (PPS) to pay for Medicare services provided by inpatient psychiatric facilities (IPFs). Prior to PPS implementation, IPFs had been paid according to reasonable costs per discharge, subject to limits established by the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA). In 1983, the Congress mandated that a prospective payment system (PPS) be established for hospital service payment under Medicare. Although most inpatient hospital services became subject to the PPS, psychiatric hospitals and psychiatric units in acute care hospitals (among other facilities) were excluded from the PPS because of concerns that the diagnosis-related groups (DRGs) used for payment adjustment were not adequate for psychiatric cases. Instead, Medicare continued to pay reasonable costs subject to limits imposed by TEFRA.

The Balanced Budget Refinement Act (BBRA) of 1999 mandated that the Secretary of Health and Human Services (HHS) develop a *per diem* (not per case) PPS for inpatient hospital services provided by psychiatric hospitals and units (or collectively, inpatient psychiatric facilities [IPFs]) and that its patient classification system reflect differences in patient use and costs among such facilities.¹ After the passage of the BBRA the Centers for Medicare and

¹ It also mandated that the Secretary of Health and Human Services maintain budget neutrality, permitted the Secretary to require psychiatric hospitals and psychiatric units to submit information necessary for the development of the PPS, and required the Secretary to submit a report to the Congress describing the development of the PPS.

Medicaid Services (CMS) engaged in research to develop regulations and payment rates for a per diem PPS, and facilities began receiving PPS payments in January 2005.²

The key components of the IPF PPS include 1) a base rate of \$595.09 (in 2007), 2) a patient classification system that groups patients according to their psychiatric DRG of admission and 17 separate comorbidity codes, 3) payment adjustments for the DRG of admission, comorbidities, age, day of stay, and patients receiving electro-convulsive treatment (ECT), and 4) facility-level payment adjustments for area wage levels, cost of living adjustment (COLA) for IPFs located in Alaska and Hawaii, rural location, teaching hospital status, and maintenance of an emergency department. The IPF PPS also includes an outlier policy with a 2 percent outlier adjustment and a stop-loss policy that guarantees the total payments to facilities under the IPF PPS are no less than 70 percent of their pre-PPS payments under TEFRA.

This study re-examines the recently developed IPF PPS using an independently-constructed database of IPF stays in fiscal year 2003 (FY 2003).³ We replicate the regression analyses underlying the IPF PPS, as described in the Federal Register (2003 Preliminary Rule, 2004 Final Rule and corrections published in 2005) and compare our findings with those of the IPF PPS which was based on FY 2002 data.

In the study, we also consider alternative comorbidity classifiers and explore using the history of past inpatient stays to improve cost prediction and suggest new payment adjusters. We

² For a description of this research, see Federal Register, Vol. 68, No. 229, Friday, Nov 28, 2003. 42 CFR Parts 412, 413 and 424, Medicare Program Prospective Payment System for Inpatient Psychiatric Facilities, Proposed Rule; Federal Register, Vol. 69, No. 219, Monday, Nov 15, 2004. 42 CFR Parts 412 and 413, Final Rule; and Federal Register, Vol. 70, No. 62, Friday, April 1, 2005. 42 CFR Parts 412 and 413, Final Rule; Correction.

³ We supplemented our FY 2003 cost report data with FY 2002 cost report data if the FY 2003 cost report data were missing.

conduct the first facility-level analyses of the IPF-PPS in order to understand the relationship between IPF PPS payments and facility-level costs. The facility-level analyses provide information on the way the PPS distributes payments across facilities and the strength of incentives for facilities to select more or less costly patients.

B. Considerations in the Design of the IPF PPS

In shifting from cost-based reimbursement to a PPS, policy makers expect to increase incentives for efficient provision of care. The payment adjustments are meant to account for clinically appropriate variation in resource use across different types of patients. The facility-level adjustments are meant to reflect cost variation across facilities (net of patient-level adjustments) that are considered to be appropriate or which are designed to offset the costs of desirable activities (e.g., maintaining an emergency department).

An important question for any PPS is whether the patient classification system and corresponding payment adjustments account sufficiently for variation in patient costs. An effective patient classification system should account for a reasonably high proportion of the predictable variation in a provider's patient care costs due to clinically meaningful differences in patient characteristics requiring more resource and care (Liu et al. 2006). The patient classification system should reflect these differences so that facilities with a more severe case-mix are reimbursed as fully as those with a less severe case-mix. In an effective system, financial losses on high-cost patients should be sufficiently offset by gains from low-cost patients (MedPAC 1999).

To the extent that a classification system does not sufficiently account for variation in patient costs, provider incentives to select patients according to risk may increase. Predicting variation in costs that is due to clinically inappropriate variation (e.g., provider inefficiencies or regional practice patterns not related to best practices) is not desirable. In other words, higher predictability of cost variation is generally considered better, as long as it is the “right kind” of variation. The unit of payment currently used for IPFs under Medicare is a patient day. Thus, a classification system should explain sufficient appropriate variation in costs per day at the person level.

C. Explaining the Costs of Inpatient Psychiatric Care

In the last several years, a variety of studies have examined factors that explain variation in the costs of inpatient psychiatric care. Cotterill and Thomas (2004) studied the merit of a per diem vs. a per case payment system. They estimated facility-level regressions of costs on payments under per case and per diem models and found that under the per case scenario, per case payments tend to show little variation around their mean while per diem costs show much more dispersion around their mean.

While Cotterill and Thomas (2004) suggests that a payment system based on per diem costs is better suited to psychiatric care in IPFs, uncovering and developing a patient classification system that encompasses and accurately reflects resource use and cost variation among psychiatric patients has proven difficult.

The sources of this difficulty have been mainly attributed to two factors (Lave 2003). First, the diagnosis of psychiatric disorders is complex, and treatments and resource use by patients do not always align well with diagnoses (see, for instance, Ashcraft 1989). Instead, the cost of treating the patient will depend in part on the patient's symptoms and reason for hospitalization (information not available in administrative data). Second, it is difficult to measure the cost of treating an individual psychiatric patient since many services that psychiatric patients receive (i.e., routine costs) are not billed for separately (see, for instance, Cromwell et al. 2003). Furthermore, routine costs, such as nursing, represent a large proportion (85 percent) of total psychiatric costs. In other words, under Medicare cost reporting, all patients within the same facility are assigned the same routine per day cost even though staff time devoted to individual patients may vary widely.

An added challenge is that while it is difficult to explain cost variation in psychiatric care using patient characteristics, it is not advisable to use facility characteristics over which the facility has control to develop an inpatient psychiatric PPS (see, for instance, Federal Register 2004). Payment adjustors based on this type of facility characteristic would create perverse incentives that are likely to affect the psychiatric care received by the patient. For example, if facilities were to receive payment adjustments based on their occupancy rate, they would have an incentive to admit fewer patients since higher payment adjustments would apply to lower-occupancy facilities.

Given that one of the main shortcomings is the limited ability of diagnoses to accurately capture patient resource use and per diem cost variation, research efforts have mainly focused on

developing alternative comorbidity categorizations and collecting measures of patient conditions that may better proxy the daily resource use and per diem costs of psychiatric patients (e.g., Cromwell et al. 2005 and 2006 with RTI International, Drozd et al. 2006, and Fries et al. 2006 with the University of Michigan).⁴

Cromwell et al. (2006), using existing administrative data, focuses on developing alternative comorbidity conditions and severity groups and interacts some of the comorbidities to identify patients with specific combinations of conditions. Their approach is guided by the idea that a set of adjusters specific to only a small set of very high-cost patients may fail to identify other above-average-cost patients, and that adjusters that do not account for multiple comorbidities may over or underpay for patients with multiple comorbidities.

Fries et al. (2006) developed a survey instrument (the Case Mix Assessment Tool or CMAT) to collect patient information that could be used to refine a patient classification system for payment purposes. CMAT proposes to collect items such as legal status, depressive symptoms, psychotic symptoms, danger to self and others, activities of daily living, instrumental activities of daily living limitations, and items such as dry mouth, tearfulness, nausea, or suicide plan.

Items derived from the CMAT have not yet been used in a regression model to explain IPF per diem cost variation as measured by administrative data and CMAT data were collected for only a limited number of facilities. In our approach, we seek to capture some of the items that

⁴ CMS contracted with the University of Michigan (Fries et al.) starting in 2000 to design and develop a survey instrument that would aid in the improvement of an IPF PPS. A final report was submitted in December 2006. CMS also contracted with RTI International (Cromwell et al.) starting in 2000 to identify patient characteristics and modes of practice believed to account for variation in per diem cost.

CMAT proposes to collect. Some of the items that CMAT would collect might be proxied, using administrative data, by 1) the claims history of the patient and 2) by improving the quality of the diagnosis-code data already collected from IPFs. In regards to 1), CMAT proposes, for instance, to collect data on the number of psychiatric admissions in the lifetime of the patient (question 8), and also on when the patient received his/her last ECT (question 32). Linking the past inpatient stays of the patient for the last 2 years, as we do, allows us to proxy and use such variables as predictors in cost per day regression models, in much the same way the corresponding CMAT questions would.

In regards to 2), CMAT would collect information on patient conditions such as housing and legal status. However, we can proxy some of these items using information that is already captured in diagnosis codes (e.g., lack of housing). Furthermore, results from the initial research conducted by RTI International indicated that some variables expected to affect cost—such as cognitive impairment, risk of falls or measures of assessment of function (Global Assessment of Function)—either “produced inconsistent results or were found to have a minor effect” (Cromwell et al. 2003). Nevertheless, items related to patient functioning or limitations in patient activities of daily living (which would be collected by CMAT) were found to be important drivers of costs in the RTI 2003 study.

This would suggest there may be some room to avoid collecting some information (e.g., housing status) that is already captured by diagnosis codes (or could be with improved coding). Giving IPFs the right incentives and guidance to more accurately code relevant diagnoses might be a way of reducing collection redundancy. There are certain patient conditions, however, not

captured by diagnosis codes that have proven to be significant cost drivers (e.g., limitations in patient daily functioning). Efforts to collect this type of information in a new patient instrument may be well warranted, but need to be tested in their ability to explain per diem costs for a broad and representative sample of IPFs.

As have these earlier studies, we also address the question of how to improve the prediction of resource use/patient costs in IPFs. Our approach focuses on maximizing the use of available administrative data, including data from past inpatient stays. Specifically:

- We test the ability of an existing comorbidity grouping system based on diagnosis codes, the Elixhauser comorbidity groups (which we refer to as ECGs), to capture the type of patient conditions that are correlated with per diem IPF costs. The ECGs were designed to be used with administrative data and with the goal of predicting hospital charges, length of stay, and in-hospital mortality. We compare the Elixhauser comorbidity groups to the current IPF comorbidity groups, as well as those developed by Cromwell et al. (2006).
- We link to current stay data, DRG, diagnosis, and stay-type information from the patient's prior 2-year inpatient history in the Medicare Provider Analysis and Review (MedPAR) files. With this information we seek to capture some aspects of patient condition(s) that other studies propose to collect in a survey from IPFs. Such information could be made available at the time of an IPF stay with an enhanced patient record. These data allow us to calculate, for example, the number of times the patient has had the current stay psychiatric DRG in the past 2 years and how many different DRGs the patient has had during those 2 years. This information provides us with a proxy for chronic conditions that may be correlated with current treatment and resource use.
- We also explore the explanatory power of some other variables, such as blindness, hearing loss, dementia, and homelessness. Furthermore, we examine the value of including some interactive terms in the model, such as the interaction of patient-age categories with certain comorbidities or DRGs.

Unlike prior studies, we address the need for the PPS to be able to account for variation in costs across facilities attributable to differences in facilities' patient case-mix. To deal with this issue, we estimate facility-level cost models using data on nearly all IPFs in FY 2003. We

summarize variation in case-mix-adjusted payments using a case-mix index. We then examine how facility costs vary with case-mix index and other facility characteristics. With these analyses, we are able to examine the proportionality between facility-level payments and IPF costs.

II. Data and Methods

This section describes our data sources and methodological approaches. More detailed information can be found in Appendix I. First, we describe the data sources and creation of the 2003 IPF analysis file and briefly describe the construction of the dependent variable in the current IPF-PPS model. Second, we address the selection of units of observation, and describe our selection criteria. Third, we present the stay-level regression models used to compute payment weights and compare alternative approaches to classifying patients. Finally, we present the facility-level regression models and motivate their use in assessing the relationship between facility payments and cost that is generated by the PPS.

A. Data Sources

We obtained our analytical dataset by linking two primary data sources: 1) the Fiscal Year 2003 (FY 2003) MedPAR file, which includes patient characteristics such as patients' diagnoses, DRGs, age, gender, and charges for each IPF stay; and 2) the 2003 Hospital Cost Report data from the Health Cost Report Information System, which includes facility characteristics such as department-level costs, the number of beds, and the number of residents. Following CMS' approach, we supplemented our FY 2003 cost report data with FY 2002 data if the FY 2003 cost report data were missing. We merged data from the facility-level cost reports to MedPAR records using the Medicare provider ID.

The dependent variable in our analysis is per diem stay cost, which is the sum of per diem routine and ancillary costs at the stay level. We calculate the variable using a standard procedure, also used by CMS, to estimate Medicare costs per discharge for inpatient care (Newhouse et al.

1989). Also using CMS' approach, we exclude emergency department costs from the dependent variable (Federal Register 2004). We use IPF cost report data to obtain facility-level routine costs and ancillary service cost-to-charge ratios (CCRs). We convert stay-level ancillary service charges (from MedPAR claims data) into estimated ancillary costs by multiplying the ancillary CCR of the facility by the ancillary charges. Because charges for routine services (e.g., nursing and accommodations) do not vary across patients in the same facility, facility per diem routine costs are assigned to individual patients in the facility.

Table 1 below shows the total number of stays and facilities in the MedPAR data and the number and percentage matched to the cost data. Only facilities with non-missing per diem routine costs, total ancillary charges, total ancillary cost, Medicare inpatient days, and total inpatient days were selected from the facilities' cost reports, thus not all psychiatric facilities in MedPAR find a match in the cost report data.

Table 1. Match of Data Sources for FY 2003

	MedPAR data	Matched to cost data
# of Stays	510,024	495,484 (97% of MedPAR)
# of Facilities	1,844	1,756

B. Data Exclusions

Our next step was to ensure the quality of our analytical data to the extent possible. With that goal, we identified missing, inconsistent, or extreme values in the variables used to estimate the dependent variable and in the explanatory or cost-adjustment variables used in our regressions.

We first flagged facilities with missing or unreasonably high or low values in key cost or charge variables for exclusion. Analogously, we also flagged stay-level observations with missing or unreasonably high or low values in key variables used in the estimation of *stay-level* per diem total cost. Appendix I describes in more detail the methods used in identifying observations with extreme values. Here, we briefly summarize:

Facility-level Flags

1. Eight facilities were missing the number of beds⁵
2. One facility had total costs (ancillary + routine) per day with an extreme value of \$20,000

⁵ Only facilities with non-missing per diem routine costs, total ancillary charges, total ancillary cost, Medicare inpatient days and total inpatient days were selected from the cost reports.

3. Fourteen psychiatric hospitals were flagged because their total ancillary CCRs were outside the range of 0.05 to 30
4. Two psychiatric units were flagged because their total ancillary CCRs were outside the range 0.05 to 10

Table 2 below shows the number and percentage of facilities flagged (out of the initial number) by each of the criteria above. It also displays the corresponding number and percentage of stays affected by the facility-level exclusions. Only a very small number of facilities and corresponding stays were excluded as a result of these rules.

Table 2. Facility-level Flags Used in Analysis of FY 2003 Data

Flag Names	# Facilities, (% Facilities)	# Stays, (% Stays)
Missing number of beds	8 (0.46)	1,089 (0.22)
Unreliable cost data	1 (0.06)	11 (0.002)
Unreliable cost-to-charge ratios for psychiatric hospitals	14 (0.80)	699 (0.14)
Unreliable cost-to-charge ratios for psychiatric units	2 (0.11)	329 (0.07)
Any facility-level flag	24 (1.37)	2,117 (0.42)
Beginning N (number of observations)	1,756	495,484
Final N after applying facility-level flags	1,732	493,367

Stay-level Flags

1. Stays with zero Medicare covered days
2. Stays with logged per diem cost more than 3 standard deviations from the respective means, with means and standard deviations computed separately for hospitals and psychiatric units
3. In our regressions, we also exclude 3,312 stays that do not have a primary psychiatric DRG

Table 3 below shows the number and percentage of stays affected by the stay-level flags.

Table 3. Stay-level Flags Used in Analysis of FY 2003 Data

Flag Names	After removing Facility-level Outliers
	# Stays (% Stays)
Zero Medicare-covered days	20,973 (4.2)
Extreme values of per diem stay costs*	3,227 (0.7)
All stay-level flags	24,200 (4.9)
Non-psychiatric primary DRG	3,312 (0.7)
Initial N (number of observations)	493,367
Final N after applying stay-level flags	465,893

* The number and % for this variable is computed after removing those stays with zero Medicare covered days since costs cannot be computed for these observations.

Table 4 shows the number of stays and facilities (N) before and after all flagged facilities and stays have been excluded, as well as the percentage of stays and facilities dropped. In applying the stay-level flags of Table 3, the number of dropped facilities increased from 24 (Table 2) to 35 (Table 4).

Table 4. Number and Percentage of Stays and Facilities before and after Data Exclusions (FY 2003)

	Stays	Facilities
Beginning N	495,484	1,756
Final N	465,893	1,721
Final N - Beginning N	29,591	35
Percent dropped	5.97%	1.99%

C. Summary Statistics for Total, Routine, and Ancillary Costs per Day

The average total cost per day for Medicare stays in IPFs is \$712 (Table 5). Costs per day in units are higher than in psychiatric hospitals (\$766 versus \$528). Routine costs are the largest component, making up 86 percent of total costs. Ancillary costs per day average \$110, and are higher in psychiatric units. While ancillary costs vary across individuals within facilities, routine costs vary only across facilities. This contributes to the low coefficient of variation for routine costs as compared to ancillary costs.

Table 5. Summary Statistics for Total Costs per Day and Its Components (FY 2003)

	All Facilities	Hospitals	Units
Total costs per day			
Mean	\$712	\$528	\$766
Standard deviation	\$238	\$180	\$226
CV	0.33	0.34	0.29
Routine costs			
Mean	\$602	\$452	\$646
Standard deviation	\$197	\$145	\$189
CV	0.33	0.32	0.29
Ancillary costs			
Mean	\$110	\$76	\$120
Standard deviation	\$104	\$84	\$107
CV	0.95	1.10	0.89
N	465,893	105,997	359,896

Note: The coefficient of variation (CV) is the standard deviation divided by the mean.

D. Stay-level Regression Models of IPF Costs per Day

Payment adjustors from the IPF PPS are derived from stay-level regressions in which the dependent variable is the natural log of the cost per day of a stay and the explanatory variables include categories of patient characteristics used for payment adjustment, facility-level characteristics used for payment adjustment, and additional control variables. The current IPF-PPS regression model used by CMS, described in the Federal Register in the 2003 Preliminary Rule and 2004 Final Rule, takes the form:

$$(1) \quad \log(\text{Stay per diem cost})_i = \delta_0 + \delta_1 * \text{DRG}_i + \delta_2 * \text{Comorbidities}_i + \delta_3 * \text{Age}_i + \delta_4 * \text{LOS}_i + \delta_5 * \text{Rural}_i + \delta_6 * \log(\text{Teaching})_i + \delta_7 * X_{1i} + \varepsilon_{1i}.$$

The subscript (i) denotes characteristics that vary by stay. The deltas (δ) are regression coefficients or vectors of regression coefficients. Payment adjustors, say for the DRGs, are equal

to the exponentiated regression coefficients (i.e., $\exp[\delta_k]$). The explanatory variables have the following definitions:

DRG = psychiatric Diagnostic Related Group indicators

Comorbidities = indicators for each comorbidity category currently established for the IPF PPS by CMS

Age = indicator variables for ranges of patient age during the stay

LOS = indicator variables for the length of the IPF stay in number of days

Rural = indicator for whether the stay is in a facility located in a rural area

Teaching = a measure of the level of teaching activity within a facility computed as one plus the ratio of the number of interns and residents assigned to the facility divided by the IPF's average daily census

X_1 = a set of additional control variables (described below)

The DRG and comorbidity categories are intended to capture differences in costs that arise from patients having different diagnoses and combinations of conditions. The age categories are also intended to capture unspecified differences in patients' clinical conditions or treatment patterns that vary with patient age and affect cost. The length of stay categories are included to account for higher resource use in the early days of an IPF stay. In longer stays, the higher costs in the early days get spread across more days leading to falling patient costs by length of stay.

The two facility-level characteristics, rural location and teaching status, are included to reflect higher expected costs in rural facilities and in those with a higher level of teaching.

Because variation by these facility characteristics has been viewed as appropriate or unavoidable (in the case of rural/urban variation) or as resulting from a desirable activity (teaching activity), the IPF PPS includes payment adjustments for these facility characteristics.

Some variables are included in the regression but are not used as payment adjusters themselves. These are identified by the vector X_{1i} in equation (1) above and include the occupancy rate of the facility (measured by the ratio of actual days to available days), a variable indicating whether the facility reported zero ancillary charges, and a 1/0 variable indicating whether the patient received ECT. Each of these variables is correlated with stay-level costs for reasons that do not warrant payment adjustments. For instance, higher costs may reflect inefficiencies in facilities with a low occupancy rate. Also, government facilities often do not report their ancillary costs even though these costs are included in their total cost. The IPF PPS currently provides payment adjustment for ECT, but does so outside the regression framework. Each of these variables is also potentially correlated with the payment-adjustment variables. To the extent that the controls and payment-adjustment variables are correlated, omitting the variables in X_1 from the regression would allow their effect to be loaded onto the coefficients of the payment adjustment variables and would result in somewhat different payment adjustments.⁶

The other model specifications we examine take the same form as the current IPF PPS model in equation (1), but use different sets of variables with the goal of improving its

⁶ Some of the covariance between the payment adjustment variables and control variables may lead to higher costs in ways that would be appropriate for payment adjustment (e.g., if government facilities treat patients who differ from patients treated in other facilities in ways that are not fully captured by the case-mix variables). The current approach strips out all of the influence of such covariance.

explanatory power. The specifications we explore can be summarized by equations (2) and (3) below:

$$(2) \quad \log(\textit{Stay per diem cost})_i = \delta_0 + \delta_1 * \text{DRG}_i + \delta_2 * (\text{Elixhauser Comorbidities})_i + \delta_3 * \text{Age}_i + \delta_4 * \text{LOS}_i + \delta_5 * \text{Rural}_i + \delta_6 * \log(\text{Teach})_i + \delta_7 * X_{1i} + \delta_8 * X_{2i} + \varepsilon_{2i}$$

$$(3) \quad \log(\textit{Stay per diem cost})_i = \delta_0 + \delta_1 * \text{DRG}_i + \delta_2 * (\text{Elixhauser Comorbidities})_i + \delta_3 * \text{Age}_i + \delta_4 * \text{LOS}_i + \delta_5 * \text{Rural}_i + \delta_6 * \log(\text{Teach})_i + \delta_7 * X_{1i} + \delta_8 * X_{2i} + \delta_9 * \text{Past} + \varepsilon_{3i}$$

Equations (2) and (3) differ from the current IPF-PPS model (equation 1) in a variety of ways:

We replace the comorbidity grouping currently used in the IPF-PPS model by a different grouping of conditions—the Elixhauser Comorbidities. Elixhauser et al. (1998) developed the comorbidity measures to predict hospital charges, length of stay, and in-hospital mortality for studies of resource use or clinical outcomes using administrative datasets. These comorbidities have been shown to be predictive of resource use and mortality, and to outperform the widely-used Charlson comorbidity categorization in a variety of settings. For instance, using data from the Veterans Administration inpatient and outpatient databases, Dominick et al. (2005) find that the Elixhauser comorbidities perform better than Charlson’s in predicting use of health services in patients with osteoarthritis. Using California hospital discharge data from 1994-97, Stukenborg (2001) shows that Elixhauser comorbidities improve upon Charlson’s in predicting the risk of in-hospital death. Motivated by the goal of providing appropriate preventive and primary care to patients with schizophrenia, Carney et al. (2006) use a modified version of Elixhauser comorbidities to determine the most common medical conditions in this population.

Southern (2004) finds that Elixhauser comorbidities outperformed the Charlson/Deyo comorbidities in predicting mortality among myocardial patients using Canadian administrative data.

Both specifications also include X_2 , which represents other factors reflecting patient conditions that may be correlated with a patient's daily resource use, and thus, per diem cost variation. These factors include the number of comorbidities a patient has during the current stay, whether the patient suffers from dementia, is blind or suffers from hearing loss, or whether the patient is homeless (a socio-economic condition).

In equation (3), we include an additional set of variables from the past stay history of the patient in MedPAR. This information includes DRGs, diagnosis and surgical procedure codes, and the types of inpatient stays (e.g., long-term psychiatric hospital stay) the patient had during the 2 years prior to the current stay. This information may serve as a proxy for assessing the chronic nature and/or severity of the current medical and psychiatric conditions of the patient, and may also capture some of the diversity of the patient's medical history. For instance, the data allow us to calculate and incorporate into our analysis the number of times the patient has been admitted for the current stay psychiatric DRG, as well as how many different DRGs the patient has been admitted for in the past 2 years. We also examine other combinations of the sets of variables described above, as well as specifications using the comorbidity classification that RTI use to explain IPFs costs.

In our model estimation we adjust the standard errors estimates for within-facility correlations in error terms, i.e., facility-level clustering. We believe the within-facility

correlations across patients are substantial, especially considering that routine costs are only measured at the facility level. This is consistent with the much smaller standard error estimates we obtain when we do not correct for facility-level clustering.

Although the estimation results we present in the Findings section are based on the full sample of stays as documented in the data section, we also conducted out-of-sample validation as a way to evaluate our models. To this end, we divided our data into two smaller samples: a “test” sample and a “validation” sample. The test sample consisted of a random sample of 50 percent of the facilities (and their corresponding stays), while the validation sample contained the remaining facilities (and corresponding stays). We first estimated our models using the test sample. For out-of-sample validation, the results from the models’ estimation were then applied to the validation sample. Out-of-sample validation mimics the way classification systems are used in practice, whereby payment rates developed from older data are used for determining current payments. Large differences between test sample and validation sample results may indicate substantial over-fitting of models to the data in the test sample. The results we obtained from our out-of-sample validation did not indicate over-fitting. We decided therefore to present estimates in this report that use the full-sample.

E. Facility-level Regressions of Average Cost per Day

In addition to explaining patient stay-level costs, effective classification systems should also be able to account for variation in costs across facilities that are attributable to variation in average patient case-mix characteristics.⁷ Facilities may have high average costs due to a more

⁷ The material in this section has been adapted from Liu et al. (2006).

severe or complex patient mix—perhaps because they are well equipped to treat such patients. Facilities with a more severe case-mix would be penalized if the patient classification system does not reflect these differences such that financial losses on high-cost patients sufficiently offset gains from low-cost patients (MedPAC 1999). Differences in IPF payment levels due to case-mix differences should therefore both explain *and be proportional to* facility differences in average costs per patient day.

We estimate facility-level models of average patient cost per day to address these issues, examining how facility costs vary with case-mix-adjusted payments and other facility characteristics. We compute the average cost per patient day for each IPF as the average of stay-level cost per day weighted by Medicare covered patient days in each facility.

Facility Case-mix Indices

Facility case-mix indices (CMIs) summarize the costliness of a facility's patient case-mix as captured by a particular patient classification system. CMIs are constructed by 1) computing the overall payment adjustment for each stay based on patient-level adjustment factors and dividing by the average to create an overall relative payment weight, and 2) averaging the relative payment weights of the stays within each facility, weighting by the number of covered days for each stay.⁸ At the facility level, CMIs are re-centered so that the means of the CMIs across all facilities are equal to 1.0. By construction, case-mix-adjusted payments per day to IPFs are proportional to the CMI.

⁸ The non-adjustment control variables do not enter these calculations.

We estimate facility-level models by regressing the log of facility average costs per day (unadjusted for area wages) for each facility (indexed by j) on a constant, the log of the CMI, and a set of control variables, which varies according to the model specification. We estimate and report two specifications for each classification/payment approach we consider.

Payment model:

$$\log(\text{Facility cost per day}_j) = \gamma_0 + \gamma_1 * \log(\text{CMI}_j) + \gamma_2 * \log(\text{WIF}_j) + \gamma_3 * \text{Rural}_j + \gamma_4 * \text{Teaching status}_j + \varepsilon_{1j}$$

Fully-specified model:

$$\log(\text{Facility cost per day}_j) = \gamma_0 + \gamma_1 * \log(\text{CMI}_j) + \gamma_2 * \log(\text{WIF}_j) + \gamma_3 * \text{Rural}_j + \gamma_4 * \text{Teaching status}_j + \gamma_5 * \text{Other facility controls}_j + \varepsilon_{2j}$$

The payment model contains the log of the CMI, the log of the area wage index factor (WIF), a rural facility indicator, and the teaching status of the facility as control variables. The later three variables are currently accepted as valid facility characteristics for which IPF payments should be adjusted. By including them in the model, we can isolate the relationship between the CMI and costs that is net of any correlation between the CMI and these facility variables that receive separate adjustment.

The “fully-specified” model adds several facility characteristics as control variables including: number of beds (0 to 9 beds, 10 to 19 beds, 20 to 49 beds, 50 to 99 beds, with 100+ beds as the reference category); being a hospital (vs. a unit); ownership type (voluntary/nonprofit, for profit-other, and government, with for profit-chain as the reference category); the occupancy rate; the percentage of patients receiving ECT; and whether the facility

does not bill ancillary charges. These additional variables are expected to be correlated with facility costs and may also be correlated with the CMIs and unobserved aspects of patient case-mix; however, they are characteristics that are not considered appropriate to use in adjusting IPF costs *per se*.⁹

Facility-level Model R-squared

The R-squared of the payment model indicates the share of variation in (log) average IPF costs per day that can be attributed to facility-level payment adjustors and to facility case-mix as measured by the particular classification system. This measure can be compared across different classification systems. It is also informative to examine how the R-squared value changes as additional explanatory variables are added. The change in R-squared from the payment model to the fully-specified model summarizes the unique contribution of the other facility characteristics that are considered inappropriate for payment adjustment in explaining IPF costs.

The CMI Coefficient and CMI Compression

The CMI coefficient (i.e., the regression coefficient on the log (CMI) variable in a facility-level model) measures the percent change in expected facility costs associated with a percent change in the CMI (and, thereby, a facility's average relative payment rate per patient day due to case-mix). The ideal situation is for the CMI coefficient to be 1.0, indicating, for

⁹ Even though the occupancy rate, the percentage of patients receiving ECT, and whether the facility does not bill ancillary charges were used in the stay-level regressions and are therefore uncorrelated with the stay-level relative weights, it remains possible that the CMI is correlated with the three variables at the facility level because of differences in weighting at the facility level and using the log of the CMI in the facility-level regressions rather than the CMI, justifying their inclusion in the fully-specified models. Indeed, we find a correlation of -0.4 between the facility-level CMI and the variable indicating whether the facility does not bill ancillary charges.

example, that a 10 percent higher CMI is associated with 10 percent higher costs. When the CMI is proportional to expected costs, facilities are paid in proportion to the higher cost associated with a higher CMI.

The situation that arises with a CMI coefficient that is greater than 1.0 is called “CMI compression.” If, for example, the estimated CMI coefficient is 1.25, a facility with a CMI and average payments that are 10 percent higher than average would tend to have costs that are 12.5 percent higher than average. Costs tend to rise (and fall) more than payments, so that high case-mix facilities tend to be underpaid and low case-mix facilities tend to be overpaid. The term compression is used because the distribution of case-mix-adjusted payments is narrower, or more compressed, than the distribution of expected costs associated with facility case-mix. A classification system that exhibits CMI compression would tend to penalize facilities that have a more severe case-mix and would create incentives to avoid high cost patients (Cotterill 1986).

We report the CMI coefficient for each of the two models and discuss results in the next section. In addition to the CMI, we report the standard error of the CMI coefficient, so that confidence intervals can be computed, and the p-value for the t-test of the null hypothesis that the CMI coefficient is equal to 1.0.

Interpreting Compression Results for the Payment and Fully-specified Models

Findings of CMI compression or decompression have somewhat different meanings and implications in the payment and fully specified models. The CMI coefficients in the payment models quantify the relationship between facility case-mix and costs across facilities, holding

area wages, urban status, and teaching status fixed (or as if an average cost adjustment has already been made for those three factors). The CMI coefficient tells us whether case-mix adjusted payments will flow to facilities in proportion to their expected costs.

The CMI coefficients in the fully-specified models measure the extent to which facilities of a given type would face incentives to risk-select patients on the basis of the characteristics included in the classification system. Facilities may be making money overall. However, if costs are related proportionately to the CMI in the fully-specified models, facilities of a given type will not be making or losing more or less money depending on the case-mix characteristics of their patients. We provide more detail on the interpretation of the CMI coefficient results and potential sources of CMI compression in Appendix III.

If the fully-specified model is a reasonably accurate overall model of the drivers of IPF costs, substantial compression in the fully-specified models would cast doubt on the effectiveness of a classification system that is not otherwise modified with payment cushions, such as outlier payments. They would also suggest by how much the payment weights would need to be expanded to bring facility payments in line with costs.

F. Consequences of Measuring Routine Cost per Day Only at the Facility Level

Routine costs are only measured at the facility level, while ancillary costs are measured at the stay-level. The consequences of using total cost data in stay-level cost regressions that is fully or partially aggregated across facilities is well-known and is relevant to other payment systems (Lave 1985). It is particularly important for the IPF PPS, however, because routine costs

make up 86 percent of total costs per day for the typical stay. The aggregated routine costs lead to log total cost regression coefficients that are biased towards zero (and hence payment weights that are biased towards one). Because the dependent variable does not show the full variation in resource use across individual patients, the payment weights do not show the full variation in expected costs across individual patients. This in turn shows up as CMI compression in the facility-level analyses. Even if the best possible predictors were used in the stay-level analyses, all else equal, more aggregation in the stay-level cost measures will lead to more compressed payment weights and CMIs. The facility-level analysis we present below sheds some light on the consequences of this data limitation for the IPF PPS.

Because payment weight compression is likely to be substantial in the IPF PPS, approaches to countering it should be considered. The outlier and stop-loss policies can temper the effects of payment weight compression to some degree, by limiting facilities' risk of significant underpayment. These policies, however, do not address more common cases where costs systematically exceed payment but not by enough to qualify for outlier or stop-loss adjustments. Nor do they address the overpayment that occurs for relatively inexpensive cases when payment weights are compressed.

Thorpe, Cretin, and Keeler (1988) suggested an approach for countering the effects of CMI compression by “decompressing” or expanding the relative weights. In this approach, relative weights greater than 1.0 could be scaled up by some factor, while relative weights less than 1.0 could be scaled down by the same factor. If relative payment weights are compressed, as the evidence we present suggests they are, this method would shift payments towards cases

(facilities) with higher expected costs and away from cases (facilities) with lower expected costs, as might be deemed appropriate.

III. Findings

We present our findings in six sub-sections. First, we provide the distributions of the explanatory variables in the current IPF PPS model. Second, we replicate the results obtained by CMS, and discuss any discrepancies encountered as well as their sources. Third, we examine the explanatory power of the different sets of patient and facility-level characteristics in the current CMS PPS model. Fourth we address the issues involved in the current day of stay adjustment and how it might be improved. Fifth, we evaluate alternative approaches to predicting stay-level IPF costs per day, including the use of other comorbidity classifications and information from past inpatient stays. Finally, in our facility-level analysis, we examine how accurately the current distributes payments across facilities as compared to our “preferred model”.

A. Variable Frequencies

We start by providing the distributions of each set of explanatory variables used in the current IPF PPS model. As seen in Table 6, Psychosis (DRG430) is by far the most common DRG, with patients in almost 73 percent of the stays in our data admitted for this mental condition. A distant second is DRG429 (Organic Disturbances), applying to approximately 9 percent of IPF patient stays. Regarding the comorbidity categories currently used by the IPF PPS, none seem to be very common among IPF patients. The two most common comorbidities are Developmental Disabilities and Infectious Diseases, with each applying to only 3 percent of patient stays. Many of the other comorbidities are quite rare. Patients under the age of 50 make up 42.4 percent of our sample, while patients over the age of 80 constitute 15.9 percent of the

sample. The distribution of length of stay is somewhat skewed towards shorter stays, with 7 days being the modal length of stay.

Table 6. Distribution of Stays by Explanatory Variable Category (FY 2003)

Category	Percent
Psychiatric DRGs	
DRG424 – Procedure w/ principal dgnos of mental illness	0.25
DRG425 – Acute adjustment reaction	1.02
DRG426 – Depressive neurosis	3.62
DRG427 – Neurosis, except depressive	1.16
DRG428 – Disorders of personality	0.65
DRG429 – Organic disturbances	8.80
DRG430 – Psychosis	72.76
DRG431 – Childhood disorders	0.33
DRG432 – Other mental disorders	0.10
DRG433 – Alcohol/Drug use, LAMA	0.27
DRG521 – Alcohol/Drug, w CC	1.61
DRG522 – Alcohol/Drug, w/out CC	0.22
DRG523 – Alcohol/Drug use, w/out rehab	3.04
DRG12 – Degenerative nervous system disorders	6.04
DRG23 – Non-traumatic stupor and coma	0.13
Current IPF PPS comorbidities	
Developmental disabilities	2.92
Infectious diseases	2.70
Drug and/or alcohol induced mental disorders	1.59
Renal failure, chronic	1.24
Uncontrolled type I diabetes mellitus	0.69
Eating and conduct disorders	0.47
Poisoning	0.46
Chronic obstructive pulmonary disease	0.44
Renal failure, acute	0.33
Artificial openings - digestive and urinary	0.32
Musculoskeletal and connective tissue diseases	0.31
Sever protein malnutrition	0.09
Tracheotomy	0.04
Coagulation factor deficit	0.03
Gangrene	0.03
Cardiac conditions	0.02
Oncology treatment	0.01
No comorbidity	89.09
(table continues)	

Table 6 (continued). Distribution of Stays by Explanatory Variable Category (FY 2003)

Age of Patient	
Age < 45	31.30
Age 45-49	11.10
Age 50-54	8.38
Age 55-59	6.03
Age 60-64	4.57
Age 65-69	7.09
Age 70-74	7.31
Age 75-79	8.35
Age 80+	15.87
Length of Stay	
Length of stay=1	3.79
Length of stay=2	4.62
Length of stay=3	6.10
Length of stay=4	6.54
Length of stay=5	6.66
Length of stay=6	7.13
Length of stay=7	7.58
Length of stay=8	6.19
Length of stay=9	5.14
Length of stay=10	4.72
Length of stay=11	4.26
Length of stay=12	3.83
Length of stay=13	3.94
Length of stay=14	4.40
Length of stay=15	3.00
Length of stay=16	2.28
Length of stay=17	1.99
Length of stay=18	1.75
Length of stay=19	1.52
Length of stay=20	1.54
Length of stay=21	1.53
Length of stay=22+	11.50
N	465,893

B. Replication of CMS's Regression Model of IPF Cost per Day

In Table 7, we compare our estimates of the current IPF PPS regression model using FY 2003 data to those obtained by CMS using FY 2002 data as documented in its 2004 Final Rule. The table is broken into several blocks, with each of them containing a different set of payment adjustors (e.g., DRGs, comorbidities). The first column lists the regressors used in CMS' IPF-PPS model. The second and third columns present the payment adjustors obtained by CMS and us respectively, and the fourth column shows the difference between the two.

Overall Findings

Overall, we find the results are quite similar. With a couple of exceptions, the regression coefficients we obtain (and thus payment adjustors) are of very similar magnitude and are statistically indistinguishable from those of CMS. The R-squared we obtain is 0.325, which is essentially identical to the 0.33 reported by CMS.

Regarding the statistical significance of the regression coefficients, we find some discrepancies between CMS' results and ours. Some of the payment adjustments reported in Table 7 are based on coefficients that, according to our estimates, are not statistically significant. In the 2004 Final Rule, however, it is reported that all payment adjustors are based on statistically significant regression coefficients. The discrepancy in part stems from the fact that we sometimes obtain different coefficient estimates. In addition, it appears that we use different methods to estimate the standard errors. Our standard error estimates are adjusted for within-facility correlations in error terms, i.e., facility-level clustering, whereas it appears that those

used in the 2004 Final Rule are not. When we do not use facility-level clustering, all of our coefficients appear to be statistically significant, except that of DRG 433.

Table 7. Replication of CMS' Final Rule Results with FY 2003 Data

Dependent Variable:	CMS Payment Adjustment	Replication Payment Adjustment	Difference Replication – CMS
Log per diem Cost of a Stay			
N	471,536	465,893	
R-squared	0.33	0.325	
Psychiatric DRGs			
DRG 424 – Procedure w principal dgns of mental illness	1.22	1.22**	0.00
DRG425 – Acute adjustment reaction	1.05	1.05**	0.00
DRG426 – Depressive neurosis	0.99	0.98*	-0.01
DRG427 – Neurosis, except depressive	1.02	1.01	-0.01
DRG428 – Disorders of personality	1.02	1.03**	0.01
DRG429 – Organic disturbances	1.03	1.04**	0.01
DRG430 – Psychosis (omitted category)	1.00	1.00	0.00
DRG431 – Childhood disorders	0.99	0.98	-0.01
DRG432 – Other mental disorders	0.92	0.77*	-0.15
DRG433 – Alcohol/Drug use, LAMA	0.97	0.99	0.02
DRG521 – Alcohol/Drug, w CC	1.02	1.00	-0.02
DRG522 – Alcohol/Drug, w/o CC	0.98	0.96	-0.02
DRG523 – Alcohol/Drug use, w/o rehab	0.88	0.85	-0.03
DRG12 – Degenerative nervous system disorders	1.05	1.05**	0.00
DRG23 – Non-traumatic stupor and coma	1.07	1.05**	-0.02
Current IPF PPS comorbidities			
Developmental disabilities	1.04	1.04**	0.00
Coagulation factor deficit	1.13	1.11**	-0.02
Tracheotomy	1.06	1.07**	0.01
Eating and conduct disorders	1.12	1.05**	-0.07
Infectious diseases	1.07	1.07**	0.00
Renal failure, acute	1.11	1.10**	-0.01
Renal failure, chronic	1.11	1.15**	0.04
Oncology treatment	1.07	1.28**	0.21*
Uncontrolled type I diabetes mellitus	1.05	1.04**	-0.01
Sever protein malnutrition	1.13	1.10**	-0.03
Drug and/or alcohol induced mental disorders	1.03	1.03**	0.00
Cardiac conditions	1.11	1.08**	-0.03
Gangrene	1.10	1.05**	-0.05
Chronic obstructive pulmonary disease	1.12	1.12**	0.00
Artificial openings - digestive and urinary	1.08	1.06**	-0.02
Musculoskeletal and connective tissue diseases	1.09	1.10**	0.01
Poisoning	1.11	1.12**	0.01

(table continues)

Table 7 (continued). Replication of CMS' Final Rule Results with FY 2003 Data

Age of Patient			
Age < 45 (omitted category)	1.00	1.00	0.00
Age 45-49	1.01	1.01**	0.00
Age 50-54	1.02	1.03**	0.01
Age 55-59	1.04	1.05**	0.01
Age 60-64	1.07	1.08**	0.01
Age 65-69	1.10	1.12**	0.02
Age 70-74	1.13	1.15**	0.02
Age 75-79	1.15	1.18**	0.03
Age 80+	1.17	1.20**	0.03
Length of Stay			
Length of stay=1	1.19	1.22**	0.03
Length of stay=2	1.12	1.14**	0.02
Length of stay=3	1.08	1.09**	0.01
Length of stay=4	1.05	1.06**	0.01
Length of stay=5	1.04	1.04**	0.00
Length of stay=6	1.02	1.03**	0.01
Length of stay=7	1.01	1.02**	0.01
Length of stay=8	1.01	1.01**	0.00
Length of stay=9	1.00	1.00	0.00
Length of stay=10 (omitted category)	1.00	1.00	0.00
Length of stay=11	0.99	0.99*	0.00
Length of stay=12	0.99	0.99*	0.00
Length of stay=13	0.99	0.99**	0.00
Length of stay=14	0.99	0.99**	0.00
Length of stay=15	0.98	0.98**	0.00
Length of stay=16	0.97	0.98**	0.01
Length of stay=17	0.97	0.97**	0.00
Length of stay=18	0.96	0.96**	0.00
Length of stay=19	0.95	0.96**	0.01
Length of stay=20	0.95	0.95**	0.00
Length of stay=21	0.95	0.94**	-0.01
Length of stay=22+	0.92	0.93**	0.01
Facility-level Adjustors			
Rural location	1.17	1.14**	-0.03
Teaching status	0.52	0.49**	-0.02

Note: Following CMS' specification, our estimation also includes control variables (i.e., occupancy rate, an indicator of whether the facility reports zero ancillary charges, and an indicator for whether the patient has received ECT treatment during the stay) that are not payment adjustors.

** Statistically significant at the 5% level

* Statistically significant at the 10% level

Only two of the payment adjustors we obtain differ substantially from those obtained by CMS—the ones corresponding to DRG 432 (other mental disorders) and the oncology comorbidity. We discuss these below.

Difference for the DRG 432 (Other Mental Disorders) Adjustment

We estimate a payment adjustment of 0.77 for DRG 432 whereas CMS uses an adjustment of 0.92. In order to find the source of discrepancy in the payment adjustment of DRG 432, we replicated the way CMS assigns a patient to a psychiatric DRG in those cases where the primary DRG of the patient is non-psychiatric. In these cases, CMS uses the second diagnosis of the patient to assign the patient to a psychiatric DRG. The payment adjustment we obtained for DRG 432 after this modification (0.78) was virtually the same as the one we previously obtained (0.77). Therefore, it would seem that the discrepancy is not explained by modifying the way patients are assigned to a particular DRG in the manner just described. Furthermore, in their replication of CMS' IPF PPS model, RTI (2006) obtains a payment adjustment for DRG 432 that is almost identical to ours (0.777).

Given these findings, CMS may want to consider reanalyzing the DRG 432 payment adjustment and monitor possible changes in the composition of this DRG. We note, however, that the percentage of patients with DRG 432 is quite small (only 0.10 percent in our sample).

Difference for the Oncology Comorbidity Adjustment

The discrepancy observed in the oncology comorbidity category apparently arises from defining this category by both diagnosis *and* procedure codes according to the modifications

introduced in the 2004 Final Rule. In CMS' 2003 Proposed Rule, the oncology comorbidity was defined only by diagnosis codes.¹⁰ In the 2004 Final Rule, CMS modified the definition of oncology to include surgical codes along with diagnosis codes. We identify approximately 7,900 cases if only diagnosis codes are used versus 54 cases if we use both diagnosis and surgical codes.

It seems that the current CMS payment adjustor for oncology does not reflect these modifications, but is still based only on diagnosis codes. When we construct the oncology variable following the 2004 CMS' guidelines, the payment weight we obtain is higher (1.28 vs. the current 1.07) and statistically significant as indicated in Table 7 above. While only 0.01 percent of IPF patients (54 stays in FY 2003) meet the criteria for the oncology comorbidity, this finding suggests that Medicare payments for these patients are nearly 20 percent underpaid.¹¹

While the relative weight estimate of 1.28 is significantly different from zero at the 95 percent confidence level, the standard error (0.041) is relatively large. Given the small number of stays with the oncology comorbidity (as well as some of the other rare conditions), we recommend that any future revision of the relative weights consider pooling multiple years of data to obtain more reliable relative weight estimates for rare comorbidities. Providers may respond to the incentives the IPF PPS creates to code diagnoses more completely. In such a case, more reliable estimates may be possible without pooling multiple years of data.

¹⁰ November 28, 2003 Federal Register, 42 CFR parts 412, 413, and 424.

¹¹ Cromwell et al. 2006 also show discrepancies for this comorbidity, although they obtain a slightly lower relative weight (1.246) than we do.

C. Sources of Explanatory Power in CMS's Payment Regression

Table 8 below shows the explanatory power of the different sets of patient and facility-level characteristics. Starting with the full model used in the replication analysis, we exclude different sets of variables to estimate their unique contribution to explanatory power. All the regressions include the occupancy rate, an indicator of whether the facility reports zero ancillary charges, and an indicator for whether the patient has received ECT treatment during the stay as control variables, as in the replication analysis table above.

The overall R-squared for the stay-level regression producing the payment weights is 32.5 percent. It must be noted, however, that much of this explanatory power arises not from patient level variables used for payment adjustment, but from other patient and facility variables in the model, including the control variables that are not used for payment adjustment. If the DRGs, age categories, comorbidity categories, and length of stay categories are dropped from the model, the R-squared only falls to 23.1 percent. The difference, 9.4, is the percent of variance explanation that can be uniquely attributed to the patient-level adjustments. Among the patient-level variables used for adjustment, the length of stay variables provide the largest amount of unique variance explanation, followed in order by age categories, DRGs, and comorbidities. The R-squared falls only slightly to 31.9 percent after excluding the current comorbidities, indicating that they only add 0.6 percentage points to the current R-squared.

Table 8. Sources of Explanatory Power in CMS’s Payment Regression Using FY 2003 Data

Model	R-squared
Full CMS model	0.325
Excluding DRGs	0.315
Excluding age	0.295
Excluding CMS IPF-PPS comorbidities	0.319
Excluding length of stay categories	0.291
Excluding all patient-level payment adjustors	0.231
Excluding facility-level payment adjustors	0.294
Excluding all patient and facility-level payment adjustors	0.196

Note: N = 465,893. Regressions include the same control variables (occupancy rate, an indicator of whether the facility reports zero ancillary charges and an indicator for whether the patient has received ECT treatment during the stay) as the replication analysis table above.

D. Issues with Adjustment for the Day of Stay

The length of stay variables add significantly to the prediction of per diem costs—more so than other sets of patient characteristics as we found above. To reflect the declining marginal cost of each patient day during the course of the stay, the current payment system makes day of stay payment adjustments equivalent to the expected percentage differences in costs by length of stay that were estimated from stay-level payment regression. The present situation is demonstrated in the first three columns of Table 9.

We find that applying the length of stay adjustments to the day of stay is not the most logical or most effective use of the estimated differences in costs by length of stay. The coefficients on length of stay in the cost regressions refer not to the marginal cost of each additional day, but rather to the average cost per day of stays with a particular length of stay.¹² A falling pattern of average cost per day with length of stay results from declining marginal costs

¹² Cromwell et al. (2003) also address this issue (p. 248).

for each day, but the decline in marginal costs by day of stay should be steeper than the decline in average per day costs by length of stay.

Column 5 of Table 9 shows the average per day payment over the entire stay, by length of stay, using the current day of stay adjustment (e.g., the average payment per day for someone staying three days is the average of the first three dollar amounts in column three). Were we to apply an overall adjustment factor to the average cost per day, by length of stay, we would use the implied LOS adjustment factors in column 6. These range from 1.19 to 1.01, which is less variability than the CMS adjustment factors in column 2 (which range from 1.19 to 0.92). The adjustment factors in column 2, however, reflect the actual variation in average per day costs by length of stay. In dollars, payments currently vary by length of stay as calculated in column 5, although we argue that they should vary as shown in column 7, which applies the current adjustment factors by length of stay.

Table 9. Length of Stay vs. Day of Stay Adjustment

(1) Day of stay	(2) CMS adjustment factor	(3) CMS daily payment by day (base=\$530)	(4) Length of stay	(5)	(6)	(7)
				Average payment per day over entire stay (CMS method)	Implied LOS adjustment factor (using base of \$530)	Average payment per day over entire stay, applying LOS adjustment to entire stay
1	1.19	\$631	1	\$631	1.19	\$631
2	1.12	\$594	2	\$612	1.16	\$594
3	1.08	\$572	3	\$599	1.13	\$572
4	1.05	\$557	4	\$588	1.11	\$557
5	1.04	\$551	5	\$581	1.10	\$551
6	1.02	\$541	6	\$574	1.08	\$541
7	1.01	\$535	7	\$569	1.07	\$535
8	1.01	\$535	8	\$564	1.07	\$535
9	1.00	\$530	9	\$561	1.06	\$530
10	1.00	\$530	10	\$558	1.05	\$530
11	0.99	\$525	11	\$555	1.05	\$525
12	0.99	\$525	12	\$552	1.04	\$525
13	0.99	\$525	13	\$550	1.04	\$525
14	0.99	\$525	14	\$548	1.03	\$525
15	0.98	\$519	15	\$546	1.03	\$519
16	0.97	\$514	16	\$544	1.03	\$514
17	0.97	\$514	17	\$542	1.02	\$514
18	0.96	\$509	18	\$541	1.02	\$509
19	0.95	\$504	19	\$539	1.02	\$504
20	0.95	\$504	20	\$537	1.01	\$504
21	0.95	\$504	21	\$535	1.01	\$504
22	0.92	\$488	22	\$533	1.01	\$488

Although the magnitude of the difference is modest, the current approach results in less variability in payments by day of stay or length of stay than is indicated by the regression analysis. Fortunately, this concern is easily remedied. All that is needed is to take the adjustment factors that are currently applied to each day of stay and apply them instead to average per day payment for the entire stay according to the length of the stay. This revision would more effectively match payments to expected patient costs.¹³

¹³ Such a revision would involve recomputing the base rate in order to be budget neutral.

E. Evaluation of Alternative Approaches to Predicting Stay-level IPF Costs per Day

As an alternative to the current set of comorbidities used in the IPF PPS, we examine the predictive ability of the Elixhauser comorbidities in predicting stay-level IPF costs per day. In this section we also examine the set of comorbidities proposed by RTI, and explore the predictive value of information from past inpatient stays.

The Distribution of Elixhauser Comorbidities in IPF Patients

We report the percentage of patients who meet criteria for each Elixhauser comorbidity in Table 10. Elixhauser comorbidities are a lot more common among IPF patients than the ones employed by the current IPF PPS. Compare, for instance, the most common comorbidity under the two alternative categorizations. While only 3 percent of the patients have the most common comorbidity (developmental disability) in the current IPF PPS model, a considerably higher 31 percent experienced hypertension—the most frequent Elixhauser comorbidity in our sample. Furthermore, the following top five Elixhauser conditions (chronic pulmonary disease, diabetes without chronic complications, drug abuse, other neurological disorders and alcohol abuse) occur in more than ten percent of IPF cases. We note that most of the Elixhauser comorbidities are physical health conditions. Such conditions have the potential to affect resource use in the IPF, as we demonstrate below.

Table 10. Percent of Medicare IPF Patients with Elixhauser Comorbidities (FY 2003)

Elixhauser Comorbidities	Percent
Hypertension	31.25
Chronic pulmonary disease	14.17
Diabetes w/out chronic complications	12.66
Drug abuse	11.01
Other neurological disorders	10.10
Alcohol abuse	10.06
Hypothyroidism	9.51
Psychoses	6.09
Depression	5.88
Fluid and electrolyte disorders	5.23
Deficiency anemias	5.04
Congestive heart failure	4.72
Obesity	4.70
Liver disease	1.51
Weight loss	1.46
Peripheral vascular disease	1.44
Valvular disease	1.35
Paralysis	1.30
Renal failure	1.29
Diabetes w/ chronic complications	1.21
Rheumatoid arthritis/collagen vascular disease	1.00
Solid tumor w/out metastasis	0.57
Coagulopathy	0.52
Acquired immune deficiency syndrome	0.47
Metastatic cancer	0.21
Lymphoma	0.15
Pulmonary circulation disease	0.14
Chronic blood loss anemia	0.11
Peptic ulcer disease excluding bleeding	0.04
N	465,893

The Explanatory Power of Alternative Model Specifications

We now present our findings on the explanatory power of the different model specifications and compare them with that obtained in the base model. Table 11 shows the R-squared for each of the main alternative model specifications we tested. It is divided into two

main columns. The left half of the table shows model specifications that use only current stay variables. The right half shows models that incorporate current and past stay information.

Table 11. Explanatory Power of Alternative Model Specifications

Current Stay Variable Models			Current and Past Stay Variable Models		
Model #	Model Description	R ²	Model #	Model Description	R ²
1a	CMS's current model	0.325	1b	1a + all past variables	0.345
			1c	1a + past DRGs	0.329
			1d	1a + past comorbidities	0.326
			1e	1a + past type of stay	0.341
2a	CMS(2)* + Elixhauser comorbidities	0.356	2b	2a + past Elixhauser comorbidities	0.358
			2c	2a + all past variables**	0.371
3	CMS(2)* + RTI comorbidities	0.339			

* CMS(2) does not include the comorbidities in the current CMS PPS model.

** This model specification uses past Elixhauser comorbidities.

Model 1a corresponds to our replication of the current IPF-PPS specification. Model 2a employs the Elixhauser comorbidities instead of the ones currently used by the IPF-PPS model (IPF-PPS comorbidities from this point forward), and model 3 uses a comorbidity classification produced for IPFs by RTI.¹⁴

¹⁴ To estimate model 3 we re-created RTI comorbidities following the guidelines provided in RTI (2006).

To assess how much explanatory power Elixhauser comorbidities add to the existing CMS IPF-PPS model, we compare models 1a and 2a. We find that using ECGs increase the R-squared by approximately 3.1 percentage points. The R-squared we obtained when using the RTI comorbidities (model 3) is 0.339, a boost of 1.7 percentage points over model 1a.¹⁵ In terms of additional explanatory power, we find that the Elixhauser comorbidities provide substantial improvement.¹⁶

We now turn to the addition of information from prior patient stays including DRGs, diagnoses, and the types of stay the patient had in the 2 years prior to the current stay (see Appendix II for detailed estimation results). The R-squared after adding all of these variables to the current IPF-PPS model (in model 1b) is 0.345—an increase of 2 percentage points over the CMS model. This boost in explanatory power can be mostly attributed to the addition of the past type of stay categorical variables (model 1e). In fact, DRGs and comorbidities from prior stays add very little to the explanatory power of the current CMS model (models 1c and 1d).

We also estimated a model (not shown in table) that included all comorbidity categorizations under consideration (i.e., CMS', Elixhauser and RTI's). The current stay variable version of this model produced an R-squared of 0.361, a very small 0.005 improvement over the R-squared of model 2a, leading to the conclusion that the Elixhauser comorbidities alone are

¹⁵ For the assessment to be complete, we should also mention RTI's (2006) own estimation results of models 1a and 3 using FY 2004 data. They obtain an R-squared of 0.318 for model 1a and of 0.329 for model 3—an increase of 1.1 percentage points.

¹⁶ We find an identical improvement in adjusted R-squared. Adjusted R-squared adds a penalty for additional explanatory variables. Finding the similar improvement in adjusted R-squared using the Elixhauser suggests is it the content of their information, and not just the additional number of categories, that leads to their increased explanatory power.

sufficient to capture most of the information contained within the two other sets of comorbidities.¹⁷

We examined several other specifications using past DRGs and comorbidities (not reported in the table) that included 1) the number of times the patient had the current stay DRG/comorbidity in the past; 2) whether the patient had other DRGs/comorbidities in the past; 3) how many other DRGs/comorbidities; and 4) the number of times he/she had them (see Appendix II for details). Our results indicated that adding information on whether the patient had other DRGs/comorbidities in the past (other than the current stay DRG) does not seem to increase explanatory power. Likewise, the other permutations we explored increased the R-squared, but only slightly.

Including information about the type of stays in the last 2 years (model 1e) boosted the R-squared by approximately 1.6 percentage points over the CMS model. This variable may capture the nature of the patient's condition in a way that is relevant in explaining some of the resource requirements the patient has in the current stay.

The addition of past stay variables to the Elixhauser comorbidity specification increases the explanatory power of model 2a by 1.5 percentage points (see model 2c). However, the R-squared of Model 2b shows that the inclusion of past Elixhauser comorbidities by themselves only slightly increases explanatory power (0.2 percentage points over model 2a). See Appendix II for detailed results on the estimation of model 2c.

¹⁷ The R-squared of the past stay variable version of this model was 0.376, also only a small 0.005 improvement over the R-squared of model 2c.

Given these results using past stay variables, our overall conclusion is that the additional benefit of including past stay variables, which require more effort to collect and construct, is not large enough to warrant using such information in the payment system. A possible exception is past type of stay information. Such information may be relatively straightforward to collect (e.g., on a new patient instrument such as CMAT) and has the most predictive ability of the sets of past stay variables we examined.

We also explored the explanatory power of variables that proxy for patient conditions that may be correlated with particularly high resource use or per diem costs, but that are not necessarily captured by comorbidities or DRGs. From the diagnosis codes available in administrative data, we created two indicator variables to identify patients who lack housing or suffer from blindness or hearing loss. Our results indicate that these variables do not add much explanatory power to the IPF-PPS model (the R-squared is 0.326, a very negligible 0.1 percentage point increase over our base CMS model). However, we need to exercise caution in interpreting these results since the reporting quality of diagnosis codes pertaining to these items in administrative claims data may be low and it may be that such cases are under-reported.¹⁸

Finally, we included indicator variables that tallied the number of comorbidities a patient had and their interaction with other regressors such as the patient's age or DRG. To test these variables, we estimated regressions using model 2a (Elixhauser comorbidity model) as our base specification. Our results indicate that while the number of (Elixhauser) comorbidity indicators increases the R-squared to 0.363—an increase of 0.7 percentage points over model 2a—the

¹⁸ This seems to be particularly true for patients suffering from blindness and hearing loss. The number of cases with these conditions was too low to yield reliable estimates.

interaction terms do not add explanatory power. The regression coefficients of the comorbidity-count indicator—and thus their implied payment adjustment factors—are all positive and significant. The addition of these variables tends to slightly decrease the payment weights of the individual comorbidities. This suggests that these variables are picking up some effects that are currently being attributed to individual comorbidities. Though crude, a count of the number of comorbidities may be a simple but worthwhile addition to the set of payment adjusters. However, concerns about up-coding and other coding problems may reduce interest in including such a variable. Also, some may also object to a simple comorbidity count on the grounds that it lacks clinical specificity.

For the remainder of the report, we adopt model 2a as our “preferred model”. In considering which model specifications are the most suitable for future implementation, two important criteria are the increase in explanatory power that certain variables provide, and the ease with which the model can be implemented within the current administrative framework. We decided that model 2a best meets both criteria, as the Elixhauser comorbidities substantially boost explanatory power and simply involve classifying data that are already being collected in a different way. We present our regression results and discuss the full model specification for our preferred model in the next sub-section.

F. Results for the Preferred Model

Summary Evaluation Statistics

In Table 12 we compare overall evaluation statistics on two alternative models: the CMS IPF PPS model and our preferred model (models 1a and 2a, respectively, in Table 11). The evaluation statistics we present are the R-squared, “sensitivity,” and the standard error of the relative weights.¹⁹ Sensitivity measures how well the classification system correctly predicts high cost cases to be high cost. We define high cost cases for this purpose as cases in the highest 10 percent of costs. Sensitivity is calculated as the percent of cases in the highest 10 percent of actual costs that are also in the highest 10 percent of predicted costs. As such, the sensitivity measures tell us the probability that the most expensive stays will be paid at the highest payment rates under a given classification system. All else equal, a more sensitive classification system will be less likely to create incentives for providers to avoid the most costly patients.

A patient classification system yields a set of predicted costs that depend on patient characteristics. Relative weights (also called relative values) measure the costliness of a patient type relative to the average patient. Relative weights are constructed by dividing the predicted cost for a particular type of patient by the average predicted cost. The relative weights have a mean of 1.0 in the sample from which they are derived. The standard deviation of the relative weights provides an indication of the extent of variability of the payment rates that will result from a particular payment system.

¹⁹ Our discussion of evaluation statistics in this section has been adapted from Liu et al. (2006).

Table 12. Summary Evaluation Statistics of the Current CMS IPF PPS Model and the Preferred Model (FY 2003)

	R-squared	% Predicted >=90th percentile given actual >=90th percentile	Standard Deviation of Relative Weights
CMS IPF PPS Model	0.325	0.233	0.104
Preferred Model	0.356	0.237	0.122

As shown in Table 12, the R-squared of our preferred model (35.6 percent) is 3.1 percentage points higher than the one yielded by our replication of the current IPF-PPS model using FY 2003 data. It is also 2.6 percentage points higher than the one obtained by CMS with 2002 data. The sensitivity indicator shows that 23.7 percent of the highest cost cases are correctly predicted by the preferred model, compared with 23.3 percent for the CMS IPF PPS model. This indicates that both models perform similarly in predicting the highest cost cases—although the preferred model’s prediction capability is slightly higher. This finding indicates that the main benefit of the Elixhauser comorbidities over the current comorbidities is to provide more accurate payment for patients whose costs lie below the top decile (and possibly within the top decile) and does little to improve the identification of cost outliers. The standard deviation of the relative weights highlights the range of values the two models cover. CMS’ model has a standard deviation of 0.104, compared with 0.122 for the preferred model, indicating that the latter model would have a greater spread in payment values.

Detailed Findings for the Preferred Model

Table 13 shows the detailed estimation results for the preferred model. Each of the Elixhauser comorbidities is statistically significant and requires IPF payment compensation. We

display the different Elixhauser comorbidities according to their t-statistics, which indicate their relative explanatory power.²⁰ Renal failure and diabetes without chronic complications top the list. The two comorbidities with the highest relative weights are metastatic cancer and renal failure, two resource-use intensive medical conditions.

Substituting the Elixhauser comorbidities for the current ones has little effect on the remaining payment adjustors, suggesting that shifting to an alternative set of comorbidities would not require substantial revision of the other weights. We note, however, that expected costs per day increase somewhat less with age after controlling for the Elixhauser comorbidities. This suggests that the Elixhauser comorbidities are picking up the types of conditions that become more frequent with increasing age and lead to increased costs in the IPF. All else equal, it should be preferable to adjust payments on the basis of explicit clinical conditions rather than indirectly through a demographic characteristic like age.

²⁰ We note that common conditions with lower relative weights may provide more explanatory power than uncommon conditions with high relative weights.

Table 13. Comparison of Relative Weights in the Current and Preferred Models (FY 2003)

	CMS IPF PPS Payment Adjustment	Preferred Model Payment Adjustment	Difference Preferred - CMS
N	465,893	465,893	–
R-squared	0.325	0.356	–
Psychiatric DRGs			
DRG424 – Procedure w/ principal dgns of mental illness	1.22**	1.20**	-0.02
DRG425 – Acute adjustment reaction	1.05**	1.04**	-0.01
DRG426 – Depressive neurosis	0.98*	0.99	0.01
DRG427 – Neurosis, except depressive	1.01	1.01	0.00
DRG428 – Disorders of personality	1.03**	1.03**	0.00
DRG429 – Organic disturbances	1.04**	1.03**	-0.01
DRG430 – Psychosis (omitted category)	1.00	1.00	0.00
DRG431 – Childhood disorders	0.98	0.99	0.01
DRG432 – Other mental disorders	0.77*	0.80*	0.03
DRG433 – Alcohol/Drug use, LAMA	0.99	1.00	0.01
DRG521 – Alcohol/Drug, w CC	1.00	0.99	-0.01
DRG522 – Alcohol/Drug, w/out CC	0.96	0.99	0.03
DRG523 – Alcohol/Drug use, w/out rehab	0.85	0.88**	0.03
DRG12 – Degenerative nervous system disorders	1.05**	1.06**	0.01
DRG23 – Non-traumatic stupor and coma	1.05**	1.06**	0.01

(table continues)

Table 13 (cont'd). Comparison of Relative Weights in the Current and Preferred Models (FY 2003)

	CMS IPF PPS Payment Adjustment	Preferred Model Payment Adjustment	Difference Preferred - CMS
CMS' IPF PPS Comorbidities			
Developmental disabilities	1.04**	—	—
Coagulation factor deficit	1.11**	—	—
Tracheotomy	1.07**	—	—
Eating and conduct disorders	1.05**	—	—
Infectious diseases	1.07**	—	—
Renal failure, acute	1.10**	—	—
Renal failure, chronic	1.15**	—	—
Oncology treatment	1.28**	—	—
Uncontrolled type I diabetes mellitus	1.04**	—	—
Sever protein malnutrition	1.10**	—	—
Drug and/or alcohol induced mental disorders	1.03**	—	—
Cardiac conditions	1.08**	—	—
Gangrene	1.05**	—	—
Chronic obstructive pulmonary disease	1.12**	—	—
Artificial openings - digestive and urinary	1.06**	—	—
Musculoskeletal & connective tissue diseases	1.10**	—	—
Poisoning	1.12**	—	—
Elixhauser Comorbidities			
Renal failure	—	1.09**	—
Diabetes w/o chronic complications	—	1.05**	—
Fluid and electrolyte disorders	—	1.07**	—
Chronic pulmonary disease	—	1.07**	—
Other neurological disorders	—	1.07**	—
Paralysis	—	1.07**	—
Hypothyroidism	—	1.05**	—
Rheumatoid arthritis/collagen vascular disease	—	1.06**	—
Congestive heart failure	—	1.04**	—
Deficiency anemias	—	1.06**	—
Solid tumor w/out metastasis	—	1.07**	—
Hypertension	—	1.05**	—
Valvular disease	—	1.05**	—
Coagulopathy	—	1.07**	—
Metastatic cancer	—	1.10**	—
Diabetes w/ chronic complications	—	1.08**	—
Obesity	—	1.05**	—
Liver disease	—	1.06**	—
Depression	—	1.04**	—
Lymphoma	—	1.08**	—
Pulmonary circulation disease	—	1.08**	—
Weight loss	—	1.07**	—
Chronic blood loss anemia	—	1.07**	—
Peripheral vascular disease	—	1.04**	—
Alcohol abuse	—	1.02**	—
Drug abuse	—	1.02**	—
Acquired immune deficiency syndrome	—	1.07**	—
Psychoses	—	1.02**	—

(table continues)

Table 13 (cont'd). Comparison of Relative Weights in the Current and Preferred Models (FY 2003)

	CMS IPF PPS Payment Adjustment	Preferred Model Payment Adjustment	Difference Preferred - CMS
Age of Patient			
Age <45 (omitted category)	1.00	1.00	0.00
Age 45-49	1.01**	1.01**	0.00
Age 50-54	1.03**	1.02**	-0.01
Age 55-59	1.05**	1.03**	-0.02
Age 60-64	1.08**	1.05**	-0.03
Age 65-69	1.12**	1.08**	-0.04
Age 70-74	1.15**	1.11**	-0.04
Age 75-79	1.18**	1.12**	-0.06
Age 80+	1.20**	1.14**	-0.06
Length of Stay			
Length of stay=1	1.22**	1.23**	0.01
Length of stay=2	1.14**	1.14**	0.00
Length of stay=3	1.09**	1.09**	0.00
Length of stay=4	1.06**	1.07**	0.01
Length of stay=5	1.04**	1.04**	0.00
Length of stay=6	1.03**	1.03**	0.00
Length of stay=7	1.02**	1.02**	0.00
Length of stay=8	1.01**	1.01**	0.00
Length of stay=9 (omitted category)	1.00	1.00	0.00
Length of stay=10	1.00	1.00	0.00
Length of stay=11	0.99*	0.99	0.00
Length of stay=12	0.99*	0.99**	0.00
Length of stay=13	0.99**	0.99**	0.00
Length of stay=14	0.99**	0.99**	0.00
Length of stay=15	0.98**	0.98**	0.00
Length of stay=16	0.98**	0.97**	-0.01
Length of stay=17	0.97**	0.97**	0.00
Length of stay=18	0.96**	0.96**	0.00
Length of stay=19	0.96**	0.95**	-0.01
Length of stay=20	0.95**	0.95**	0.00
Length of stay=21	0.94**	0.94**	0.00
Length of stay=22+	0.93**	0.93**	0.00
Rural location	1.14**	1.13**	-0.01
Teaching status	0.49**	0.49**	0.00

Note: Following CMS' specification, the regressions also include control variables (i.e., occupancy rate, an indicator of whether the facility reports zero ancillary charges, and an indicator for whether the patient has received ECT treatment during the stay) that are not payment adjusters.

** Statistically significant at the 5% level

* Statistically significant at the 10% level

G. Facility-level Analysis

Case-mix Indices by Facility Type

Facility case-mix indices (CMIs) summarize the expected costliness of a facility's patient case-mix *as captured by a particular patient classification system* relative to an average facility. Facilities with CMIs above (below) 1 are more (less) costly than average. Due to the PPS design, relative payments to facilities per patient day are proportional to the CMI. Table 14 below shows the mean and standard deviation of CMIs by facility ownership type.

We first notice that psychiatric units treat a costlier (per diem) patient case-mix than psychiatric hospitals. CMI means are all above 1 for psychiatric units whereas the means for psychiatric hospitals are all below 1. While government hospitals generally have the highest *per case* costs, they also tend to have the longest stays, thus making their *per diem* costs low relative to other types of facilities.²¹ We also notice that for-profit units have higher CMIs on average than other types of psychiatric units. To explore this issue we examined the length of stay, patient age, number of comorbidities, and DRG mix across types of facilities. Our findings indicate that for-profit units tend to have an older patient mix with a higher number of comorbidities than other IPFs. Their lengths of stay also tend to be slightly lower than those of most other IPFs (except for non-profit units).

²¹ This finding is consistent with CMS' findings (Federal Register 2004, p.66934): "The government-operated psychiatric hospitals have relatively low per diem costs, relatively long lengths of stay, and relatively high per case costs... Psychiatric units have the highest per diem and per case costs, followed by non-profit hospitals, and last by for-profit hospitals."

Table 14. CMIs by Facility and Ownership Type in the Preferred Model Specification (FY 2003)

	CMI Mean	CMI Standard Deviation
Government hospital	0.87	0.034
For-profit hospital	0.92	0.052
Non-profit hospital	0.95	0.060
Government unit	1.01	0.076
For-profit unit	1.05	0.064
Non-profit unit	1.02	0.057

Note: N = 1,721 facilities.

Facility-level Regression Results

We estimate facility-level regressions to examine the relationship between average facility costs per day and case-mix-adjusted payments as summarized by the CMI. We compare payment models and fully-specified models (described in the Data and Methods section) under three scenarios. The first scenario corresponds to the current IPF PPS payment adjustment method in which the length of stay effects are applied (inappropriately in our view) to the day of stay. In this scenario, the CMI is based on relative weights estimated for model 1a above, except that we reduce the variation in the length of stay adjustments to reflect the way they are applied to day of stay.²²

²² We reduce the variation in length of stay effects by multiplying each estimated length of stay adjustment factor from the regression by the corresponding ratio of payment amounts in columns (5) and (7) in Table 8, where column (5) corresponds to the effective length of stay adjustment that occurs when the length of stay coefficients are applied to day of stay.

The second scenario uses the relative weights estimated for model 1a (without any rescaling) to compute the CMIs, which implies making adjustments to length of stay instead of day of stay as we recommended above, but otherwise reflects the current payment adjustment method. The last scenario we consider uses the relative weights of our preferred model—model 2a—to compute the CMIs. The only difference between this third scenario and the second is that it replaces the current comorbidities with the Elixhauser comorbidities. We report the results in Table 15.

Table 15. Facility-level Explanatory Power and the Proportionality of Payments with Costs

	(1)	(2)	(3)
	CMS' IPF PPS Model w/ LOS adjustments as currently applied by CMS	CMS' IPF PPS Model w/ avg adjustment by LOS	Preferred Model w/ avg adjustment by LOS
Payment model			
R-squared	0.373	0.393	0.400
CMI coefficient	3.31	3.18	2.52
SE on CMI coefficient	0.112	0.103	0.082
p-value to test coeff.=1	<0.001	<0.001	<0.001
Fully-specified model			
R-squared	0.573	0.574	0.575
CMI coefficient	1.58	1.55	1.34
SE on CMI coefficient	0.125	0.123	0.106
p-value to test coeff.=1	<0.001	<0.001	0.0015

Note: N = 1,721 facilities.

Starting with the findings for the current payment adjustment method in column 1, we find that the CMI and other payment variables explain a large share (37.3 percent) of the variation in facility-level costs per day. The change in R-squared from the payment model to the fully-specified model summarizes the marginal contribution of the other facility characteristics

(e.g., facility type, size, ownership type, and the occupancy rate—considered inappropriate for payment adjustment) in explaining IPF costs. The R-squared increases to 57.3 percent in the fully-specified model in column 1, which indicates a strong independent influence relationship to costs of facility characteristics, net of case-mix and facility-level payment adjustment factors.

The CMI coefficient of 3.31 in the payment model provides evidence of substantial CMI compression under the current system. Payments to facilities are compressed in the sense that under this payment system, facilities with a case-mix that results in 10 percent higher payments (i.e., a CMI of 1.10) are found, on average, to have 33.1 percent higher costs. This indicates that facilities with high cost patients are not fully compensated for treating high cost patients and facilities with lower cost patients are overcompensated relative to their case-mix.

Payments are more proportionate to costs after we control for facility characteristics. In our fully-specified model of facility-level costs, we estimate a CMI coefficient of 1.58, which suggests that facilities with 10 percent higher payments experience 15.8 percent higher costs. Although the degree of CMI compression is much reduced in the fully-specified models, it indicates that facilities still face incentives to avoid high cost patients under the current system and that even within facilities of a given type, facilities with a more expensive case-mix are not fully compensated for treating more expensive patients.

Comparing the CMI coefficient findings in columns 1 and 2 shows the impact on the extent of CMI compression of using the stay-level cost model estimates to make length of stay payment adjustments rather than day of stay adjustments. We find that making the suggested

change would move things in the right direction by bringing the CMI coefficients closer to 1, but the overall impact would be modest.

Finally, we move to the results we obtain using our preferred model (column 3), which addressed the day of stay issue, but more importantly, replaces the current comorbidities with the Elixhauser comorbidities. Comparing results in columns 2 and 3, we find that the CMI coefficient in the payment model falls from 3.18 to 2.52, and falls from 1.55 to 1.34 in the fully-specified model. In the fully-specified model, this implies that facilities with 10 percent higher payments due to their case-mix experience 13.4 percent (down from 15.8 percent) higher costs. This represents a substantial improvement in the match between payments and costs across facilities, which is driven entirely by a more complete description of the clinical characteristics of IPFs patients through the alternative set of comorbidities.²³

While significant CMI compression would remain even with our preferred model, we find that incentives to select against more costly payments can be reduced, and payment equity improved, by using a more powerful comorbidity classification system as exemplified by the Elixhauser comorbidities.

The main source of the remaining CMI compression is, in our view, the aggregated nature of the routine cost data. Routine costs are only measured at the facility level and make up 86 percent of total costs per day for a typical stay. The aggregated routine costs lead to log cost regression coefficients that are biased towards zero (and hence payment weights that are biased

²³ It is interesting to note that in moving from column 2 to 3, R-squared increases very little, even in the payment model. This suggests that much of the new information contained in the Elixhauser comorbidities is correlated with other facility-level variables included in the models. It is consistent, for example, with higher costs in psychiatric units being driven by a medically more complex case-mix.

towards one), which in turn leads to CMI's and payments that vary less than proportionately with expected costs. Methods to counter the expected bias and decompress the relative weights warrant consideration. Such methods have been proposed by Thorpe, Cretin, and Keeler (1988), who suggested countering the effects of compression by "decompressing" or expanding the relative weights. In this approach, relative weights greater than 1 could be scaled up by some factor, while relative weights less than 1 could be scaled down by the same factor. The magnitude of the scaling factor may be decided on the basis of facility-level analyses like those we have presented. We also believe that research on statistical methods that could reduce the aggregation bias in this context would also be fruitful.

IV. Conclusion

Our findings in replicating the data set construction and regression analysis that form the basis for the patient-level adjustments in the IPF PPS with FY 2003 data are extremely similar to those reported by CMS using 2002 data. We noted exceptions in the cases of the current oncology comorbidity and the “other mental disorders” DRG. We discussed the way payments are currently adjusted according to day of stay, and find that the same adjustment factors should be applied instead to the average payment per day according to length of stay.

Payments vary less than proportionately with costs at the facility level—a finding that raises concerns about the adequacy of payments to facilities that treat a costlier than average patient mix. Adopting our proposal on the length of stay issue would make a slight improvement in this regard.

We also demonstrate how using an “off the shelf” comorbidity classification system, the Elixhauser comorbidity groups, as an alternative to the current set of comorbidities, can substantially improve the prediction of per diem costs at the stay level and also improve the correspondence between adjusted payments and costs in the PPS. Use of the Elixhauser comorbidity groups would distribute payments across facilities substantially more proportionately with costs than the current system. The Elixhauser comorbidity groups can be formed using the same administrative data that is already being collected for the current system, and would therefore be relatively easy to implement. In implementing a revised comorbidity classification for the IPF PPS, CMS could draw from the current comorbidities, those developed

by RTI, and the Elixhauser comorbidity groups examined here. In such a case, CMS should consider engaging a clinical panel to screen out conditions that may be overly subject to gaming and to resolve overlaps while retaining the best aspects of each classification system.

References

- Ashcraft, Marie L.F., Brant E. Fries, David R. Nerenz, Spencer P. Falcon, Sujan V. Srivastava, Caryl Z. Lee, S.E. Berki, Paul Errera. 1989. "A Psychiatric Patient Classification System: an Alternative to Diagnosis-Related Groups." *Medical Care* 27 (5): 543-57.
- Carney, Caroline P., Laura Jones, Robert F. Woolson. 2006. "Medical Comorbidity in Women and Men with Schizophrenia." *Journal of General Internal Medicine* 21 (11): 1133-37.
- Cotterill, Philip G., and Frederic G. Thomas. 2004. "Prospective Payment for Medicare Inpatient Psychiatric Care: Assessing the Alternatives." *Health Care Financing Review* 26 (1): 85-101.
- Cotterill, Philip G. 1986. "Testing a Diagnosis-Related Group Index for Skilled Nursing Facilities." *Health Care Financing Review* 7 (4): 75-85.
- Cromwell, Jerry, Barbara Gage, Edward Drozd, Jan Maier, Erin Richer. 2001. "Report on the Feasibility of a Per diem Prospective Payment System for Psychiatric Facilities Excluded from PPS." CMS Contract No. 500-95-0058 prepared by the Health Economics Research, Inc., Waltham, MA.
- Cromwell, Jerry, Barbara Gage, Edward Drozd, Jan Maier, Deborah Osber, Christian Evensen, Erin Richter, Howard H. Goldman. 2003. "Psychiatric Inpatient Routine Cost Analysis, Final Report." CMS Contract No. 500-95-0058 prepared by RTI International, Waltham, MA.
- Cromwell, Jerry, Edward M. Drozd, Barbara Gage, Jan Maier, Erin Richter, Howard H. Goldman. 2005. "Variation in Patient Routine Costliness in U.S. Psychiatric Facilities." *Journal of Mental Health Policy and Economics* 8(1): 15-28.
- Cromwell, Jerry, Edward M. Drozd, Barbara Gage, Jan Maier, Jan F. Hales. 2006. "Comorbidity Options for the Inpatient Psychiatric Facility Prospective Payment System (IPF-PPS)" CMS Contract No. 500-00-0024 prepared by RTI International, Waltham, MA.
- Dominick, Kelli L., Tara K. Dudley. 2005. "Comparison of three Comorbidity Measures for Predicting Health Service Use in Patients with Osteoarthritis." *Arthritis and Rheumatism* 53(5): 666-72.
- Drozd, Edward M., Jerry Cromwell, Barbara Gage, Jan Maier, Leslie M. Greenwald, Howard H. Goldman. 2006. "Patient Casemix Classification for Medicare Psychiatric Prospective Payment." *American Journal of Psychiatry* 163 (4): 724-32.
- Elixhauser, Anne, Claudia Steiner, D. Robert Harris. 1998. "Comorbidity Measures for Use with Administrative Data." *Medical Care* 36 (1): 8-27.
- Federal Register. 2003. *Medicare Program Prospective Payment System for Inpatient Psychiatric Facilities, Proposed Rule*. Washington, DC: Office of the Federal Register 28 (229), November 28.

- Federal Register. 2004. *Medicare Program Prospective Payment System for Inpatient Psychiatric Facilities, Final Rule*. Washington, DC: Office of the Federal Register 69 (219), Nov. 15.
- Federal Register. 2005. *Medicare Program Prospective Payment System for Inpatient Psychiatric Facilities, Final Rule; Correction*. Washington, DC, Office of the Federal Register 70 (62), April 1.
- Frank, Richard G., and Judith R. Lave. 1986. "Per case Prospective Payment for Psychiatric Inpatients: an Assessment and Alternatives." *Journal of Health Politics, Policy, and Law* 11 (1): 83-96.
- Fries, Brandt E. and Carl Gibson. 2006. "Design, Development, and Implementation of a Prospective Payment System for Inpatient Psychiatric Hospitals and Exempt Units." CMS Contract No. 500-96-0007 prepared by Michigan Public Health Institute, Okemos, MI.
- Lave, Judith R. 1985. "Is Compression Occurring in the DRG Prices?" *Inquiry* 22 (2): 142-147.
- Lave, Judith R. 2003. "Developing a Medicare Prospective Payment System For Inpatient Psychiatric Care." *Health Affairs* 22 (5): 97-109.
- Liu, Korbin, Bowen Garrett, Douglas Wissoker, et al. "Options for Improving Medicare Payment for SNFs." 2006. Final Report to CMS Contract No. 500-00-0025.
Available at: http://www.cms.hhs.gov/SNFPPS/Downloads/snf_report_final_03-01-07-2.pdf
- Medicare Payment Advisory Commission (MedPAC). 1999. *Report to the Congress: A Framework for Considering Medicare Payment Policy Issues*. Washington, DC: MedPAC.
- Newhouse, Joseph P., Shan Cretin, and Christina J. Witsberger. 1989. "Predicting Hospital Accounting Costs." *Health Care Financing Review* 11 (1): 25-33.
- Pettengill, Julian, and James Vertrees. 1982. "Reliability and Validity in Hospital Case-Mix Measurement." *Health Care Financing Review* 4(2):101-28.
- Secretary of Health and Human Services Report to Congress. 2002. *Prospective Payment System for Inpatient Services in Psychiatric Hospitals and Exempt Units*. Washington, DC: U.S. Department of Health and Human Services.
- Stukenborg, George J. 2001. "Comparison of the Performance of two Comorbidity Measures, with and without Information from Prior Hospitalizations." *Medical Care* 39 (7): 727-39.
- Southern, Danielle. 2004. "Comparison of the Elixhauser and Charlson/Deyo Methods of Comorbidity Measurement in Administrative Data." *Medical Care* 42 (4): 355-60.
- Thorpe, Kenneth E., Shan Cretin, and Emmett B. Keeler. 1988. "Are the Diagnosis-Related Group Weights Compressed?" *Health Care Financing Review*. 10 (2): 37-46.

Appendix I. Details on Dataset Construction

Construction of the Dependent Variable (Log of per diem stay cost)

The dependent variable in our analysis is constructed using the same calculation as CMS' IPF-PPS.²⁴ The per diem total cost of a stay has two components: per diem routine costs and per diem ancillary costs. We obtained per diem routine costs from the FY 2003 Hospital Cost Report data from the Health Cost Report Information System. Following CMS' approach, we supplemented our FY 2003 cost report data with FY 2002 data if the FY 2003 cost report data were missing. Per diem ancillary costs were calculated by first deriving the total ancillary cost of a stay and then dividing it by the number of Medicare covered days in the stay. Total ancillary costs were derived by multiplying each facility's departmental ancillary cost-to-charge ratio (obtained from IPFs' cost reports) by the corresponding ancillary charges on the MedPAR stay record.

Following CMS' Final Rule methodology, we also made two adjustments to the estimated per diem costs:

- 1) CMS pays for emergency department (ED) costs of IPFs with qualifying EDs and IPFs that are part of hospitals with qualifying ED through a specific adjustment to the day one variable per diem adjustment factor. Thus, ED costs were excluded from the dependent variable used in the cost regressions. This was done in order to remove the effects of ED costs from other payment adjustment factors with which ED costs may be correlated
- 2) The resulting per diem cost was adjusted for differences in labor cost across geographic areas by dividing per diem stay cost by an adjustment factor:

$$\text{Adjustment Factor} = (\text{labor share} * \text{WI}) + (1 - \text{labor share}),$$

²⁴ Federal Register November 28, 2003, 42 CFR Parts 412, 413 and 424, Medicare Program Prospective Payment System for Inpatient Psychiatric Facilities, Proposed Rule; and Federal Register, Vol. 69, No. 219, Monday, Nov 15, 2004. 42 CFR Parts 412 and 413, Medicare Program Prospective Payment System for Inpatient Psychiatric Facilities, Final Rule.

where WI is the FY 2003 area-wage index for the Metropolitan Statistical Area in which the facility is located and *labor share* is the labor-related share of cost determined by CMS for the IPF-PPS.²⁵

Also as CMS, we take the natural log of the dependent variable. IPFs' costs per day are typically positively skewed with a small number of observations having very high per diem costs. Taking the log of the dependent variables reduces the skew, and thus reduces the undue influence of the small set of patients with very high per diem costs. It also results in multiplicative, rather than additive, payment adjustments.

Outliers and Data Exclusions

Facility-level Flags

We identified facilities whose data we deemed to be too unreliable to be included in our analysis sample. We encountered three types of problems on facility cost reports that led to the creation of flags for specific facilities: (1) missing values, (2) unreliable cost data on the cost report, and (3) unreliable cost-to-charge ratios.

1) Missing values

Only facilities with non-missing values in the variables used in the creation of our dependent variable (i.e., per diem routine costs, total ancillary charges, total ancillary cost, Medicare inpatient days, and total inpatient days) were selected from the cost reports. We then looked for missing values in other variables used in the creation of our explanatory variables. We found 8 facilities with the number of beds missing.

²⁵ The labor share is 0.72247 for 2003.

2) Unreliable cost data

Cost data for some facilities were determined to be unreliable based upon unlikely cost breakdowns or extreme cost levels. For facilities that report any ancillary costs, we searched for facilities that had unacceptable ratios of ancillary to routine costs based upon ratios less than or equal to .005 or ratios greater than 5.0. We also looked for facilities with total ancillary costs that exceeded the sum of the costs of the ancillary cost components by a significant amount. In addition, we searched for facilities with extreme routine costs per day and per diem cost. We found and flagged one facility with an extreme value of \$20,000 in per diem costs (see Appendix Table I.2).

3) Unreliable cost-to-charge ratios

We conducted an analysis by facility type to determine which cost-to-charge ratios (CCRs) were “out of range.” It is important to note that there is no “right” answer. Our goal was to be as inclusive as possible, while identifying IPFs that were likely to contribute clearly erroneous data to our analysis.

We conducted detailed analyses of CCR distributions separately for psychiatric hospitals and units, flagging IPFs with CCRs that appeared to be extreme values. From this analysis, we determined that psychiatric hospitals CCRs outside the range 0.05 – 30.0 (see Appendix Table I.1) were unreasonable, and that psychiatric units with CCRs outside the range of 0.05 – 10 were unreasonable (see Appendix Table I.2)

Appendix Table I.1 below shows the number and percentage of facilities being flagged under each criteria.

Appendix Table I.1. Facility-level Flags Used in Analysis of FY 2003 Data

Flags	# Facilities (% Facilities)	# Stays (% Stays)
Missing number of beds	8 (0.46)	1089 (0.22)
Unreliable cost data	1 (0.06)	11 (0.002)
Unreliable cost-to-charge ratios for psychiatric hospitals	14 (0.80)	699 (0.14)
Unreliable cost-to-charge ratios for psychiatric units	2 (0.11)	329 (0.07)
Any facility-level flag	24 (1.37)	2,117 (0.42)
Beginning N (number of observations)	1,756	495,484
Final N after applying facility-level flags	1,732	493,367

Stay-level Flags:

We excluded stays that met two separate criteria: (1) those with zero Medicare covered days, and (2) those with extreme values of per diem costs.

(1) Zero Medicare covered days

We require the number of Medicare covered days to compute our dependent variable. Thus, we excluded stays with missing or zero Medicare covered days (see Appendix Table I.2 below).

(2) Extreme values in per diem total costs

The method we used follows CMS', which is fairly standard in previous work on hospital costs. We first computed the means and standard deviations of the logged per diem total cost for hospital and psychiatric units separately. Then we proceeded to trim stays where the logged per diem cost was outside plus or minus 3 standard deviations from the respective means for hospitals and psychiatric units. The reason for having separate calculations for hospitals and units is that units tend to have higher per diem costs than hospitals, and thus stays in units would be more likely to be considered extreme values, and be eliminated at a higher rate, had we computed a common mean and standard deviation for all IPFs. In our regressions, we also exclude 3,312 observations that do not have a primary psychiatric DRG (see Appendix Table I.2 below).

Appendix Table I.2. Stay-level Flags Used in Analysis of FY 2003 Data

Flags	After removing Facility-level Outliers
	# Stays (% Stays)
Zero Medicare-covered Days	20,973 (4.2)
Extreme high values of per diem stay costs*	3,225 (0.7)
Extreme low values of per diem stay costs*	2 (0.0)
ALL STAY-LEVEL FLAGS	24,200 (4.9)
Non-psychiatric primary DRG	3,312 (0.7)
Initial N	493,367
Final N after applying stay-level flags	465,893

* The number and percent for this flag is computed after removing those stays with zero Medicare covered days (costs cannot be computed for these observations).

Appendix Table I.3 summarizes the number of stays and facilities before and after all exclusions were applied as well as the percentage of stays and facilities flagged.

Appendix Table I.3. Number and Percentage of Stays and Facilities Before and After Data Exclusions (FY 2003)

	Stays	Facilities
Beginning N	495,484	1,756
Final N	465,893	1,721
Final N - Beginning N	29,591	35
Percent dropped	5.97	1.99

Appendix II. Details of Models Including Information from Past Stays

As discussed briefly in the Data and Methods section, we pulled variables from the inpatient stay history of the patient. The information includes DRGs, diagnosis and surgical procedure codes, and the types of stays (e.g., long-term psychiatric hospital stay) the patient had during the 2 years prior to the current stay.

These data allowed us to calculate and incorporate in our analysis a variety of variables that, in principle, could give us information for assessing the chronic nature and/or severity of the current medical and psychiatric conditions of the patient, and for capturing some of the diversity of the patient's medical history. We create several variables based on past DRG, past comorbidity, and past type of stay information.

From past DRG information, we created several sets of variables:

- (i) The number of times a patient had the current stay DRG in the past two years,
- (ii) Alternatively to (i), we created a 1/0 variable that indicated whether the patient had the current stay DRG in the past two years,
- (iii) A set of dummy variables for each non-current stay psychiatric DRG the patient had in the past two years,
- (iv) A variable that captures the number of times the patient had a given non-current stay psychiatric DRG in the past two years, and
- (v) Alternatively to (iv), a variable indicating whether the patient had a given non-current stay psychiatric DRG in the past two years

From past comorbidity information, we also created alternative sets of variables. One indicated how many times the patient had a given comorbidity in the past two years. The other set of 1/0 indicator variables tell whether the patient had a given current stay comorbidity in the last 2 years. Finally, we created 1/0 variables for each type of stay the patient had in the last 2 years.

We estimated a variety of models to test what combination of the variables above gave us the highest increase in explanatory power. We found that the additional benefit of including these variables, which require more effort computing, does not seem to be large enough to warrant the effort. For instance, employing an indicator variable that tells us whether the patient had a given current stay DRG in the past two years (ii. above), relative to how many times the patient had it (i. above), is simpler to derive and gives us virtually the same explanatory power.

Appendix Table II.1 below displays the estimation results for the past stay variables when we include them in our preferred model specification (model 2c in Findings section). Because the coefficients—and thus, payment adjustments—of current stay variables do not significantly change when we add past stay information, we are not presenting those results here. The inclusion of past stay variables boosts the R-squared of our preferred model by 1.5 percentage points. This increase can be mostly attributed to the inclusion of past type of stay indicators.

Appendix Table II.1. Estimation Results Using Information from Past Stays

	Regression Coefficient	Payment Adjustment
Current Stay DRG in the Past 2 Years		
Patient had current stay DRG424 in last 2 yrs	-0.04	0.96
Patient had current stay DRG425 in last 2 yrs	-0.02	0.98
Patient had current stay DRG426 in last 2 yrs	-0.03	0.97**
Patient had current stay DRG427 in last 2 yrs	0.03	1.03**
Patient had current stay DRG428 in last 2 yrs	-0.01	0.99
Patient had current stay DRG429 in last 2 yrs	-0.04	0.96**
Patient had current stay DRG430 in last 2 yrs	-0.03	0.97**
Patient had current stay DRG431 in last 2 yrs	-0.06	0.94**
Patient had current stay DRG432 in last 2 yrs	-0.05	0.95**
Patient had current stay DRG433 in last 2 yrs	0.02	1.02
Patient had current stay DRG521 in last 2 yrs	-0.02	0.98**
Patient had current stay DRG522 in last 2 yrs	-0.02	0.98
Patient had current stay DRG523 in last 2 yrs	-0.01	0.99
Patient had current stay DRG12 in last 2 yrs	-0.05	0.95**
Patient had current stay DRG23 in last 2 yrs	0.01	1.01
DRGs in the Past 2 Years		
Patient had DRG424 in a previous stay	0.01	1.01
Patient had DRG425 in a previous stay	0.00	1.00
Patient had DRG426 in a previous stay	0.00	1.00
Patient had DRG427 in a previous stay	0.03	1.03**
Patient had DRG428 in a previous stay	0.02	1.02**
Patient had DRG429 in a previous stay	-0.02	0.98**
Patient had DRG430 in a previous stay	-0.03	0.98**
Patient had DRG431 in a previous stay	0.00	1.00
Patient had DRG432 in a previous stay	0.00	1.00
Patient had DRG433 in a previous stay	-0.02	0.98**
Patient had DRG521 in a previous stay	-0.02	0.98**
Patient had DRG522 in a previous stay	-0.04	0.96**
Patient had DRG523 in a previous stay	-0.03	0.97**
Patient had DRG12 in a previous stay	0.00	1.00
Patient had DRG23 in a previous stay	0.01	1.01*
Patient had a non-mental DRG in a previous stay	0.01	1.01*

(table continues)

Appendix Table II.1 (continued). Estimation Results Using Information from Past Stays

	Regression Coefficient	Payment Adjustment
Elixhauser Comorbidities in Past 2 Years		
Aids in last 2 yrs	0.00	1.00
Alcohol abuse in last 2 yrs	-0.01	0.99
Deficiency anemias in last 2 yrs	0.01	1.01
Rheumatoid arthritis/collagen vascular disease in last 2 yrs	-0.01	0.99
Chronic blood loss anemia in last 2 yrs	-0.04	0.96**
Congestive heart failure in last 2 yrs	-0.02	0.98**
Chronic pulmonary disease in last 2 yrs	-0.02	0.98**
Coagulopathy in last 2 yrs	-0.03	0.97**
Diabetes w/out chronic complications in last 2 yrs	-0.03	0.97**
Diabetes w/ chronic complications in last 2 yrs	0.01	1.01
Drug abuse in last 2 yrs	-0.02	0.99**
Hypertension in last 2 yrs	0.00	1.00
Hypothyroidism in last 2 yrs	-0.01	0.99**
Liver disease in last 2 yrs	0.01	1.01
Lymphoma in last 2 yrs	-0.07	0.94**
Fluid and electrolyte disorders in last 2 yrs	0.00	1.00
Metastatic cancer in last 2 yrs	0.01	1.01
Other neurological disorders in last 2 yrs	-0.01	0.99
Obesity in last 2 yrs	-0.01	0.99
Paralysis in last 2 yrs	0.00	1.00
Peripheral vascular disease in last 2 yrs	0.00	1.00
Pulmonary circulation disease in last 2 yrs	0.01	1.01
Renal failure in last 2 yrs	-0.01	0.99**
Solid tumor w/out metastasis in last 2 yrs	-0.02	0.98**
Peptic ulcer disease excluding bleeding in last 2 yrs	-0.02	0.98**
Valvular disease in last 2 yrs	-0.02	0.98**
Weight loss in last 2 yrs	-0.03	0.97**
Psychoses in last 2 yrs	-0.01	0.99**
Depression in last 2 yrs	-0.01	0.99*

(table continues)

Appendix Table II.1 (continued). Estimation Results Using Information from Past Stays

	Regression Coefficient	Payment Adjustment
Past Type of Stay Categories		
Stay in psych unit of short-term hospital in last 2 yrs	0.06	1.06**
Long-term stay in psych hospital in last 2 yrs	-0.06	0.94**
Stay in short-term hospital in last 2 yrs	0.01	1.01
Stay in skilled nursing facility hospital in last 2 yrs	0.02	1.02**
Stay in rehab unit of short-term hospital in last 2 yrs	0.02	1.02**
Stay in long-term hospital in last 2 yrs	0.03	1.03**
Long stay in critical access hospital in last 2 yrs	0.05	1.05**
Long stay in rehab hospital in last 2 yrs	0.04	1.04**
Skilled nursing facility stay in short-term hospital swing bed in last 2 yrs	0.05	1.05**
Skilled nursing facility stay in short-term critical access hospital swing bed in last 2 yrs	0.00	1.00
Other type of long stay in last 2 yrs	0.09	1.10**
N	465,893	
R-squared	0.371	

Note: Regression also includes all current stay control variables in the preferred specification, as reported in Table 13.

** Statistically significant at the 5% level

* Statistically significant at the 10% level

Appendix III. Technical Details on Facility-level Regressions, Case-mix Indices, and CMI Compression

As described in the Data and Methods section, the payment model has as regressors the log of the area wage index factor, a rural facility indicator and the log of the facility's teaching status because these three variables are currently accepted as valid facility characteristics for which IPF payments should be adjusted.²⁶ The fully-specified model adds control variables for number of beds (0 to 9 beds, 10 to 19 beds, 20 to 49 beds, 50 to 99 beds, with 100+ beds as the excluded category), being a hospital-based facility, ownership type (voluntary/nonprofit, for profit-other, and government, with for profit-chain as the excluded category), the occupancy rate, and whether the facility reports zero ancillary charges. These additional variables are expected to correlate with facility costs and may be correlated with the CMIs and unobserved aspects of patient case-mix; however, they are characteristics that are not considered appropriate to use in adjusting IPF costs *per se*.

The CMI Coefficient and CMI Compression

With the log-log model specification, the CMI coefficient (i.e., the regression coefficient on the log (CMI) variable in a facility-level model) measures the percent change in expected facility costs associated with a percent change in the CMI (and, thereby, its average relative payment rate per patient day due to case-mix). The ideal situation is for the CMI coefficient to be 1.0, indicating, for example, that a 10 percent higher CMI is associated with 10 percent higher costs. When the CMI is proportional to expected costs, facilities are paid in proportion to the higher cost associated with a higher CMI.

²⁶ This section has been adapted from Liu et al. (2006).

The situation that arises with a CMI coefficient that is greater than 1.0 is called “CMI compression.” If, for example, the estimated CMI coefficient is 1.25, a facility with a CMI and average payments that are 10 percent higher than average would tend to have costs that are 12.5 percent higher than average. Costs tend to rise (and fall) more than payments, so that high case-mix facilities tend to be underpaid and low case-mix facilities tend to be overpaid. The term compression is used because the distribution of case-mix-adjusted payments is narrower, or more compressed, than the distribution of expected costs associated with facility case-mix. A classification system that exhibits CMI compression would tend to penalize facilities that have a more severe case-mix and create incentives to avoid high cost patients (Cotterill 1986, Pettengill and Vertrees 1982).

A CMI coefficient that is less than 1.0 would indicate what is called “CMI decompression.”²⁷ Decompression is a less common finding in the literature, and does not occur in any of the IPF regressions we estimated. When decompression is present, facilities with a CMI greater than 1.0 will tend to be overpaid and facilities with a CMI less than 1.0 will tend to be underpaid.

Interpreting Compression Results for the Payment and Fully-specified Models

The CMI coefficients in the payment models quantify the relationship between facility case-mix and costs across facilities, holding area wages, rural status, and teaching status fixed (or as if an average cost adjustment has already been made for those factors). The CMI coefficient

²⁷ The term “decompression” could also be used to describe a procedure to adjust for compressed weights, but here we mean the opposite of compression.

tells us whether case-mix adjusted payments will flow to facilities proportionately with their expected costs.

The fully-specified models estimate the relationship between case-mix-adjusted payments and costs within classes of facilities defined by the additional control variables. As such, the CMI coefficients measure the extent to which facilities of a given type would face incentives to risk-select patients on the basis of the characteristics included in the classification system. Facilities of a given type may be making or losing money overall. However, if costs are related proportionately to the CMI in the fully-specified models, they will not be making or losing more or less money depending on the case-mix characteristics of their patients.

If the fully-specified model is in fact a reasonably accurate overall model of the drivers of IPF costs, compression in the fully-specified models would raise concerns about the effectiveness of a classification system that is not otherwise modified with payment cushions, such as outlier payments.

Appendix Table III.1 shows the 9 possible scenarios for compression, proportionality, and decompression in the payment and fully-specified models and their basic interpretation. Compression in the fully-specified model is very likely to be accompanied by compression in the payment model (since CMI coefficients usually fall when additional control variables are added to the model). It would, therefore, be unlikely that we would find a higher CMI coefficient in the fully-specified model than in the payment model. Thus, we focus on the more likely scenarios.

Appendix Table III.1. CMI Coefficient Scenarios in Payment and Fully Specified Facility-level Models and Their Implications

Payment model	Fully-specified model		
	CMI coefficient > 1 Compression	CMI coefficient = 1 Proportionality	CMI coefficient < 1 Decompression
CMI coefficient > 1 Compression	<p>Problematic</p> <p>High cost facilities' costs exceed payments and low cost facilities' payments exceed costs</p> <p>Incentive to select against high cost patients</p>	<p>Ambiguous</p> <p>Depends on reason for compression</p> <p>High cost facilities' costs exceed payments and low cost facilities' payments exceed costs</p> <p>No incentive to select patients based on characteristics in classification system</p>	<p>Problematic</p> <p>Indication of powerful facility effects</p> <p>High cost facilities' costs exceed payments and low cost facilities' payments exceed costs</p> <p>Incentive to select against low cost patients</p>
CMI coefficient = 1 Proportionality	<p>Unlikely</p>	<p>Ideal</p> <p>Facilities' case-mix-adjusted payments equal to expected costs</p> <p>No incentive to select patients based on characteristics in classification systems</p>	<p>Problematic</p> <p>Facilities' case-mix-adjusted payments equal to expected costs, but</p> <p>Incentive to select against low cost patients</p>

Proportionality in both equations is ideal and would suggest both payment equity across equations and minimized incentives to risk-select on the basis of factors included in the patient classification system.

One likely case with an ambiguous interpretation is a finding of proportionality in the fully-specified model but compression in the payment model. Under this scenario, the classification system may or may not be adequate, depending on the source or sources of the compression. A specific example best illustrates this case. Suppose we find a CMI coefficient of 1.2 in the payment regression and 1.0 in the fully-specified regression. This result could only occur if the added facility characteristics in the fully-specified model were related both to costs and to the CMI. Assume the CMIs are related to facility characteristics simply because different types of patients are more appropriately treated in different types of facilities. In this case, interpreting the differences in CMI coefficients requires an understanding of, or assumptions about, why the facility characteristics relate to costs.

Facility characteristics could be related to costs for some combination of 4 main reasons:

- Cost reporting or accounting differences across facility types
- Case-mix differences across facility types not captured by the classification system
- Operational efficiency differences by facility type
- Quality of care differences by facility type

Each reason would have a different implication for the compression in the payment regression, and, although we lack sufficient information to know the *relative* contribution of each with any degree of certainty, we can at least consider their separate implications.

If compression in the payment model were driven by cost accounting differences across facility types, there would be no concern. Facilities would receive case-mix adjusted payments proportionate to their true underlying costs.

If the compression resulted from case-mix differences related to the type of facility, but those differences were not captured by the classification system and thus not reflected in payments, it would be problematic. Costly case-mix facilities would not be paid sufficiently or fairly for their more costly case-mix and low cost facilities would be overpaid.

If the compression was due to inefficiency in certain types of facilities, inefficient facilities would have an incentive to become more efficient, which would further the goals of the Medicare program in the long run. However, financial strain on inefficient providers could potentially disrupt patient care in the short run, should providers close or reduce their level of patient care.

If the compression resulted from differences by facility type in the quality of IPF patient care, although quality-adjusted costs would vary in proportion to payments, high quality facilities would receive no reward for superior quality and low quality facilities would incur no penalty for providing inferior care. Policymakers would need to balance patient care objectives against fiscal

objectives, and there is currently no provision for compensating higher quality IPF care through Medicare payment adjustments.