

Title: Putting providers at-risk: How strong are incentives for upcoding and treatment changes?

Draft: October 1, 2009

Authors: Marisa E. Domino, Ph.D., Edward C. Norton, Ph.D., Jangho Yoon, Ph.D., Gary Cuddeback, Ph.D., Joseph Morrissey, Ph.D.

Abstract

Pre-paid, or capitated payments to health care providers change incentives for treatment, and are widely used in a variety of settings. Certain forms of capitated payments, such as case-rate payments, create incentives to change both diagnosis and treatment patterns that may differ from both fee-for-service and pure capitated models. The purpose of this study is to examine the changes in severity determination and service use associated with changes in case-rate payments used to pay for publicly-funded mental health care. We model provider-assigned severity categories as a function of category-specific capitated payments using conditional logit regressions and counts of service days per month using hurdle models. We find that severity determination is only weakly associated with price but that level of use shows a greater degree of association; these results vary between the early or “transition” period immediately after implementation to the subsequent period, 2-4 years after implementation.

Acknowledgements: This research was supported by grant funding from the John D. and Catherine T. MacArthur Foundation’s Mental Health Policy Research Network and from the National Institute of Mental Health (MH63883). Marisa Domino was also supported by NIMH grant # K01-MH065639. For assistance in assembling the administrative data and input on its interpretation we are indebted to the King County, Washington Mental Health, Chemical Abuse and Dependency Services Division.

1. Background and Motivation

There is a clear, often-studied agency relationship between patients and their health care providers, and the strength of that relationship is predictably altered by the types of contracts implemented to pay for health services. As Tom McGuire writes in his Handbook of Health Economics manuscript (2000), “Payments to physicians are laced with incentives, and these incentives, to hospitalize, to treat, to take time to diagnose carefully... direct resources in health, thus determining costs and outcomes.” Pre-paid, or capitated payments to health care providers change incentives for treatment, and are widely used in a variety of health care settings. Pre-payment involves the calculation of an amount that will provide appropriate incentives for provider behavior. The Medicare program, for example, is a large federal system that is increasing the implementation of capitated provider payments; to date Medicare pre-pays for inpatient medical and psychiatric care, home health care, skilled nursing facility care, and now prescription drug use. Capitation rates that are too low can reduce the feasibility of providing services covered under that capitation rate and may encourage shifting to substitute services that may be free goods to the provider (e.g., Domino et al., 2004). Capitation rates that are too high may distort the incentives to participate in the payment scheme by increasing the supply of providers, further driving up price; this has been seen in the market for specialty hospitals in the Medicare program (Berenson, Bazzoli, and Au, 2006).

Certain forms of capitated payments, such as case-rate or tiered payments, create additional incentives to change both diagnosis and treatment patterns that may differ from

pure capitated models. Pure capitated models pay a fixed amount per-person per-month to a provider for all covered individuals, regardless of whether they use services or not. Case-rate payments are used to pre-pay for care for only those individuals engaged in treatment, leaving the sponsoring agency or employer at risk for changes in the number of individuals using services as well as the severity of service-system users. The Medicare DRG system of payment for inpatient care as a function of diagnosis or severity can be thought of as a case-rate system, although one that only lasts as long as the individual is in the hospital. Providers receiving case-rate payments retain only the risk for the number and intensity of services used, not for the number of individuals engaged in treatment. Research has shown a link between the use of case-rate payments for outpatient mental health services and a large (25%) decrease in the use of mental health services over fee-for-service alternatives (Rosenthal 1999 & 2000). In addition, the size of the case-rate payment has been shown to affect the level of service use. However, the level of case-rate payment may also affect the assignment to a severity category, which may have independent effects on services use; to our knowledge this has not been previously studied in an outpatient mental health context.

In this paper, we develop a theoretical model of upcoding and explore the implications from this model empirically for a single county mental health system. Case-rate payments are one of the principle management tools that a managed care organization has in terms of shaping or influencing provider behavior. Little information is available to help managed care organizations gage the level of responsiveness of providers to changes

in case-rate payments. This research will help policy makers and managed behavioral health care organizations enhance this important management tool.

2. Theoretical Model

Physician behavior in selecting the type and number of treatment services for their patients can be modeled according to a utilization-maximization framework. We adapt our model from McGuire (2000), extending it to the case where capitated payments are a function of illness severity, as is the case with Medicare DRGs and other similar systems.

We assume patients in the mental health system are one of two types: high severity, H, and low severity, L. The type is unobservable to the sponsor or principal, but observable to the physician agent. The physician reports the patient type to the payor as ϕ_H or ϕ_L and receives capitation payment R_H or R_L , with $R_H > R_L$. That is, providers can misclassify patients. The tradeoff for classifying an individual as a higher severity is that while reimbursement is higher, the level of services provided must be higher too, due to administrative oversight by the payor. The physician's goal is to maximize:

(1) $U(NB, \Pi)$ over the number of patients classified as ϕ_H or ϕ_L and choice of x_i

where the arguments NB and Π represent patient net benefits and provider net income from treating the patient, respectively, with $U_{NB} > 0$, $U_{\Pi} > 0$, $U_{\Pi\Pi} < 0$, $U_{NB\,NB} < 0$, $U_{NB\,\Pi} < 0$.

We assume that patients and physicians are price takers and that patients do not switch doctors.

Net benefits are specified as the improvement in health from the initial endowment due to service provision, subtracting out the costs (out of pocket plus time cost), p , of service receipt. The benefit production function varies with patient type. That is, net benefits are specified according to:

$$(2) \quad NB_i = B^i(x_j) - px_j$$

where i indexes the actual patient type and j indexes the reported patient type. Patients who are reported to be of their true type have $i=j$. Health is assumed to vary weakly with the intensity of services, x_i , with $B_x \geq 0$, indicating that more services may improve health and $B^L < B^H$ at any value of x . That is, we assume that the health benefit production function for both patient types is weakly concave, that B^L is always higher than B^H , for all $x > 0$, and that the two curves converge with greater x (but by definition never cross).

For a given set of patients of type H or L, the provider decides which fraction to label as type ϕ_H or ϕ_L . For patients of type L, the incentive to classify them as type ϕ_H brings with it higher reimbursement, but also the obligation to provide services beyond what maximizes patients' net benefits due to the concavity of the health production function and the inclusion of patient costs in the net benefits equation. For patients of type H, there is no incentive to classify as type ϕ_L since it results in a lower level of reimbursement and an under-provision of services, yielding an unambiguously lower level of physician utility, (1).

Physician-level net benefits are therefore a weighted sum of the three types of patients coding options: N_L is the fraction of patients of type L who are coded as type ϕ_L ; N_H is the fraction of patients of type H who are coded as type ϕ_H ; and δ_L is the fraction of patients of type L who are coded as type ϕ_H .

$$(3) NB = \sum_{i=L,H} N_i (B^i(x_i) - px_i) + \delta_L (B^L(x_H) - px_H)$$

Physicians also receive utility from profits. Providers are paid on a case-rate basis, with different capitation rates based on reported patient severity. Profits are as usual the difference between the capitated rate received and the total cost of treatment provision.

$$(4) \Pi = N_L [R_L - cx_L] + (N_H + \delta_L) [R_H - cx_H]$$

where R_i is the capitation rate, with $R_H > R_L$, c is the full marginal cost of service provision and the size of the caseload has been normalized to one.

This model leads to the following first-order conditions:

$$FOC1: B^L(x_L) - B^H(x_H) - p(x_L - x_H) = \frac{U_\pi}{U_{NB}} (R_L - cx_L - R_H - cx_H)$$

which indicates that if Hs and Ls are equally profitable, then services will be provided at the optimal level, where the difference in benefits equals the difference in marginal costs.

$$FOC2: B^L(x_L) - p = \frac{U_\pi c}{U_{NB}}$$

Since the right-hand side of equation FOC2 is positive, this equation says that services will be under-provided to accurately-coded Ls, since the marginal benefit of services is still greater than their opportunity costs.

$$FOC3: U_{NB}(B^L(x_H) - B^H(x_H)) = -U_\pi(R_H - cx_H)$$

which says that the more profitable the payments for high severity patients, the greater the difference in benefits between providing high levels of service for high severity and mis-coded patients.

$$FOC4: \delta_L^* = \frac{(1 - N_L)[U_\pi c - U_{NB}(B^{H'}(x_H) - p)]}{U_{NB}(B^{L'}(x_H) - B^{H'}(x_H))}$$

FOC4 has a number of implications. First, $\delta_L^* > 0$ when $(B^{H'}(x_H) - p) < 0$, which indicates that too many services will be provided to high severity patients, and when the marginal benefit of services to Ls is lower than the marginal benefit of services to Hs, which is true by design. Second, δ_L^* increases with greater marginal utility of profits, greater marginal cost of service provision, and the closer the two benefit production curves are to each other. The sign of the derivative of δ_L^* with respect to the payment level, R_i , is ambiguous.

An alternative specification of this model would be the introduction of noise into the classification problem, such that physicians observe the type as $i \pm \xi$, where ξ is a random error term, and attempt to classify patients as accurately as possible. Either approach

yields the same result in this case, but the second interpretation escapes the thorny issue of physician deceptive behavior. Either interpretation can be used to motivate the phenomenon of upcoding or “DRG creep” (e.g., Silverman and Skinner, 2004).

The purpose of this study is to examine the association between changes in case-rate payments used to pay for publicly funded mental health care and severity determination and service use in each severity category. We carefully avoid making causal links, since it is possible that an unaccounted for third factor causes both changes in tiering rates and severity determination, although we have uncovered no such credible candidate. As tiering payments increase for a given severity level, more individuals may be so classified, a phenomenon noted in the Medicare DRG literature as “creep” or up-coding (Hsia et al., 1992). In addition, it is not clear what effect an increase in the tier payment will have on the number of services provided as more patients are classified at higher severity levels.

3. Empirical Analysis

We separately examine the association between changes in tiered payment rates and (i) the tier assignment and (ii) the number of service days provided to patients in each tier. We use service days, or days on which one or more services were provided to an individual, rather than total costs or an intensity of services provided for reasons explained below.

Although the tier or severity categories¹ were numbered (1A, 1B, 2A, 2B, 3A, 3B) implying a natural ordering, the ranking of the tiers in terms of the minimum severity level required changed over time (described further below) indicating that an ordered model may not be appropriate. We therefore run the tier selection model as an unordered conditional logit, expressed as a function of the daily payment rates, quarterly time dummies, and baseline demographic factors, including age defined at the beginning of each month, gender, race/ethnicity (White, African-American, Asian, Native American, Latino/a, or other race/ethnicity). Standard errors were adjusted for clustering based on repeated observations on individuals. Because ordered logit coefficients are difficult to interpret, marginal effects were calculated at the variable means for the first quarter. For the severity assignment model, we calculate own-price as well as cross-price effects for three arbitrarily chosen tiers in each time period.

Count data models were run on the number of service days used in each month. Because of the known instability of the zero-inflated negative binomial model, which did fit these data better than alternative models, we use a two-part hurdle model. In this model, an indicator variable for any service use was specified as a logit, whereas the level of service use (minus one to reintroduce zeros into the model) was specified as a negative binomial. Both parts of the count data models controlled for the same covariates indicated above (age, gender, race/ethnicity). The model was run both with a combined price term and price-severity level interactions.

¹ Tier and severity category are used here interchangeably.

Additional analyses were conducted on individuals identified as having severe mental illness (SMI). This indication was drawn from diagnoses in all data sources available to the larger project, including state psychiatric hospitals, general hospitals, and jails. SMI individuals may have different patterns of service use and providers may have additional constraints in classifying them into severity categories and providing different levels of service use.

4. Data

This study takes advantage of a series of changes in the daily case-rate payments that occurred over an almost 4-year period in King County, Washington after the implementation of a case-rate payment system. Under this payment system during the study period, a private managed care organization monitored the use of the case-rate system, providing oversight on the intensity of use, but payment rates were set by the government. The sample from this study comes from the King County Mental Health, Chemical Abuse and Dependency Services Division (KCMHCADS) in Washington State. The year 2000 population in King County (1.74 million) was concentrated in one principal city, Seattle, but about two thirds of the population was spread across a number of smaller cities and unincorporated areas.

On April 1, 1995, the county implemented a tiered payment system to participating outpatient mental health providers, switching from the previous fee-for-service system. Not all services were paid on a capitated basis; for example, crisis-related and residential services were carved out of the capitated system and reimbursed under separate contracts.

For each individual participating in the case-rate payment program, providers assigned a tier; tiers had to be preauthorized for a fixed period of time. Initially, providers could choose among six tiers for their capitated patients; in September 1996, the tiering system was restructured such that the two middle categories were collapsed into a single tier. We ran separate models for these two time periods because of these differences. Each tier had an associated daily payment rate, which depended on consumer age (child less than 21 although some children age 18-20 received adult rates, adult, or adult age 60 or greater) and other factors (ethnicity, deafness, medically compromised/homebound, sexual minority status). Other adjustments were made to payment rates for providers delivering specialty care, such as to those providing culturally-specific care. Authorizations for approximately 10% of tiered clients received these special payment rates. In addition, actual payment rates sometimes were reduced from published rates for individual benefits by recoupments when service hours fell below identified minimums.

Two key changes are examined in the empirical models. First, the daily payment rate for each tier changed over time. Including the initial payment schedule, there were six different price periods which occurred during our study period, from April 1, 1995 – December 31, 1998. These price changes, deflated by a GDP index and expressed in 1998 dollars, are plotted in Figure 1. The daily case rate payment is the key variable of interest in the analysis. The second key change that occurred was in the medical necessity determination that served as a guideline to aid providers in tier assignments and to determine thresholds for authorization of tier assignment. Necessity was assessed using

several measures including the Global Assessment of Functioning (GAF) scale and the Children's Global Assessment Scale (CGAS). These two measures are easily implemented but often-criticized scales used to determine the level of functioning of an individual, and take on numeric values ranging from 0 to 100, with higher scores indicating greater functioning (Soderberg, Tungstrom, and Armelius 2005; Mugnaini and Fabiano, 2006). A maximum GAF score, which reflects the difference in the health production functions in the theoretical model, was allowed for each tier and these maxima changed over time. For example, Tier 1A, the Brief Intervention Tier, started with a maximum GAF of 69, which was eventually raised to 80 (on 1/1/97). Similarly, Tier 2A, the Brief Intensive Tier, started with a maximum GAF of 30, and was increased to 60 (on 1/1/96). The (negative) correlation between the tiered payment rate and the maximum GAF scored was high but not perfect, at -0.87 for the first time period and -0.77 for the second period. Two of the six tiers had no changes to their maximum allowable GAF score over time. No changes in the minimum GAF scores occurred in the second policy period; therefore the GAF scores are only used in the first period models.

Many other changes occurred in the tiered payment system over time. Here, we discuss a few of the key differences, but we are not able to capture the full picture of the changing treatment climate. One notable factor which changed over time but is not incorporated here is the length of authorization of each tier. At the outset of the tier program, tiers were authorized for a known period, ranging from 91 to 364 days depending on the tier. On September 1, 1996, 17 months after implementation, the authorization length was standardized across tiers to a constant 365 days. Because of the lack of variation in this

measure, we are unable to incorporate it into our analyses. In as much as other omitted changes are correlated with the factors included here, we are at risk of inappropriately attributing causation to the included factors.

Although providers were no longer paid on a fee-for-service basis, they were mandated to report the number and type of services provided to each service recipient. Incentives were still strong to report services under the capitated regime since administrative oversight and sanctions existed. In our data, we plotted average service use over time and did not find a visible break in monthly service use trends from before to after the tiering system was implemented.

We collapsed service use down to a daily measure of any outpatient service use, regardless of type and added service days up to obtain a measure of the number of service days per month for each individual. We do this for several reasons. First, the capitated system may have distorted the incentives to accurately report the type of services, rather than the provision of services. For example, fee-for-service systems using the Current Procedure Terminology (CPT) coding to report service units distinguish outpatient psychotherapy according to the length of the session. These types of distinctions may be less accurately reported in the non-billing service system, although the incentive to report a service may be preserved. All types of mental health services were recorded in the administrative data, and these included intake assessments, telephone contacts, advocacy and linkage, case management and case termination-related activities. Second, we wanted to aggregate across service types to develop a measure of use. The standard way to do

this is to use dollar amounts as intensity weights, but due to the changes in the service-recording system (changed from FFS to capitated services), there was not a simple method that was independent of the actual capitated amounts. Therefore, we used service days per month. Service days will undercount services if more than one service was provided in a single day. It will also weight equally services of different intensities. We felt the net advantages of this approach were superior to the alternatives.

A random sample of individuals using the King County Mental Health system was collected as part of a larger project (Morrissey et al., 2006; Domino et al., 2004). Individuals were sampled according to a stratified sample based on their use of KCMHCADS services, the King County Jail, and their enrollment in WA State Medicaid, with different service system use resulting in different sampling weights. Use was defined over the period from July 1, 1993 to December 31, 1998. For the present analysis, we used only the sample of individuals for whom tier-based payments were made, from April 1, 1995 (the date the tiered system was implemented) until December 31, 1998. We retain the original sampling weights, as these return us to the population of individuals who used KCMHCADS services during the full study period. Individuals using only non-tiered services were excluded in the present analysis. Tiered payment was made to providers regardless of Medicaid status; therefore we do not distinguish individuals according to their Medicaid enrollment, although the majority of individuals in our sample (71.24%) were enrolled in Medicaid at some point during the sample period.

We merged daily payment rates with data at the person-month level, based on the age category (<18, 18-59, and ≥60) on the date of tier authorization and the tier assigned. We do not make adjustments for the rate differential provided for other factors described above or to specialty providers noted earlier because our data do not contain information sufficient to identify those patients and providers. However, there is an almost perfect correlation (0.999 – 1.0) between the special rates and the standard rates over time, indicating that our use of the standard rates should not bias the results.

Sensitivity analysis

Actual payment rates received may vary from those announced initially during the last year of data (1998). In January of 1998, the county implemented a system of recoupments called a risk corridor. In any given month, if total agency payment exceeded an agency-specific threshold, payments were reduced by the percent the agency exceeded its threshold. Separate risk corridors were specified for Medicaid and non-Medicaid individuals. These risk corridors may have an effect on the severity determination and service use rate. We conduct a sensitivity analysis by excluding data from 1998.

5. Results

Variable means are presented in Table 1. The mean age of the sample was 39 years old, ranging from 18-64. Just over half were female (54%) and the majority were white (69%), with 13% African-American, 10% classified as Other race, and 6% Asian. Native Americans were just over 1% of the sample. Over three-fourths (77%) of the sample were classified as severely mentally ill. The average length of participation in the tiering

system was just under two years (20 months), with a range of 1 to 45 months. In the first period, the fourth category was most widely used, with 53% of person-months. In the second period, tier assignments were clustered between two of the top tiers, accounting for a combined 90% of tier-months.

Table 2 presents the association between the level of the daily payment rate and the probability of being assigned each severity category. Coefficients from the conditional logit models are insignificant in Period 1 and significant and positive in Period 2. The marginal effects of changes in daily tier payment rates are generally quite small, indicating that changes in tiered payment amounts had little effect on the actual probability of tier assignment. The size of the effects in Period 2 indicates that a \$1 increase in the monthly tier price is associated with an increase in the probability of tier assignment of less than one percentage point. Results from the models on the truncated time period and on severely mentally ill individuals were quite similar to those from the full model.

Results from the logit models (Table 3) indicate that payment rates increased the probability of receiving one or more visits only in Period 1, and the marginal effect in that period was small. In the model with the tier-specific price interaction, the largest positive effect of the capitation rate was in Tier 5, with a \$1 increase in the monthly rate associated with a 2% point increase in the probability of receiving some service. In contrast, the monthly tier rate had a negative effect on service receipt in two of the six tiers. The price effects became insignificant in Period 2 overall and across all 5 tiers.

Results from the count model of the number of days tell a somewhat different story. In Period 1, increases in the payment rates were associated with increases in services use overall and for three of the six tiers, with a significant decrease in services use for tier 4, and no change for the other tiers. The magnitude of the marginal effects imply that a \$10 decrease in daily payment rates, a similar magnitude to that actually experienced in some tiers during the study period, is associated with 0.54 fewer service days per month overall or 6.5 fewer service days per year, which may be clinically important. Only in tier 4 are price increases associated with fewer visits. Price effects generally become insignificant in Period 2. Results are again similar for the truncated and SMI models.

6. Discussion

Daily tier payments were seen to be weakly associated with changes in tier assignment and the level of services use in the King County mental health service system. That is to say, upcoding did seem to occur, but at very small levels and only after providers had some experience with the new system.

The association between the payment levels and the number of service days in a month, however, was only significant only in the first period, but potentially at a clinically important level. This may be because a number of factors, including the pathway specified by the theoretical model of a pathway through greater provider utility to patient health through increased service provision. We cannot rule out other explanations, such as greater oversight (or perceived oversight) by the mental health agency when rates

changed. This result may not be especially surprising during Period 2, as no substantial changes occurred in the maximum allowed GAF scores within categories, nor in authorization lengths, and price changes during this period were minimal over time (Figure 1).

The finding that treatment patterns changed in light of payment changes is certainly not surprising to economists, but does indicate that rate setting has important implications. Providers in our data were not at risk for inpatient services, but decreases in use of outpatient services associated with rate changes may lead to further increases in inpatient use and therefore expenditures over time. The magnitude of the reductions indicates that a \$10 reduction in monthly tiering price would be associated with 420 fewer mentally ill persons receiving any services per month in a population of 10,000 SMI and among those who did receive services, an average of 538 fewer visits would be provided.

Several important caveats are in order. First, data are from a single county agency which used an innovative approach to paying for mental health services during a complex administrative period and may not generalize to other systems using a similar payment scheme. Other changes that were not captured in the present analyses may have occurred during this period. In particular, since we do not have an independent measure of patient severity in our data we are unable to control for the severity level of consumers using county-funded mental health services, nor do we control for the myriad institutional changes that were occurring over this period. For example, Medicaid qualifying categories were changing during this period, and these changes may have changed the

severity mix of individuals over time. To the extent that changes in those factors were correlated with the time-varying variables analyzed here, we are at risk of inappropriately attributing those factors to changes in the tiered payment system. We urge caution in inferring causation to these results.

Methods of paying health care providers for their services have important implications on the delivery of services to patients. Health care is an unusually complex environment and an ideal payment method which is free from the uncertainty surrounding the delivery of health care services in terms of their known effects on patient health has not yet emerged. Payment systems which provide differential payments based on patient characteristics such as time-variant severity represent an important improvement in mitigating some of the perverse incentives inherent in non-standardized fee-for-service payment (e.g., over-treatment) or pure capitation methods (e.g., under treatment), but still require a level of oversight and sophistication in determining payment levels that give appropriate incentives for optimal care. Results from this study indicate that health program directors and policy makers need to be acutely aware of the interplay between provider payments and patient care and eventual health outcomes.

References

- Berenson RA, Bazzoli GJ, Au M (2006), “Do Specialty Hospitals Promote Price Competition?” Center for Studying Health System Change Issue Brief #103, Available: <http://www.hschange.com/CONTENT/816/#ib1>, Accessed August 15, 2006.
- Domino ME, Norton EC, Morrissey JP and Thakur N (2004), “Cost Shifting To Jails After A Change To Managed Mental Health Care,” *Health Services Research* 39(5):1379-401.
- Greene WH (2000) *Econometric Analysis*, 4th Edition. New Jersey: Prentice Hall.
- Hsia DC, Ahern CA, Ritchie BP, Moscoe LM, Krushat WM (1992) “Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988,” *JAMA* 268(7):896-9.
- McGuire TC (2000). “Physician Agency” in *The Handbook of Health Economics*, Cuyler AJ and Newhouse JP, Eds. Amsterdam: Elsevier.
- Morrissey JP, Steadman HJ, Dalton KM, Cuellar A, Stiles P, Cuddeback GS (2006) Medicaid enrollment and mental health service use following release of jail detainees with severe mental illness. *Psychiatric Services* 57(6):809-15.
- Mugnaini D and Fabiano GA (2006). The evaluation of impairment in the attention deficit/hyperactivity disorder: the contribution of the impairment rating scale. *Minerva Pediatr.* 58(2):159-66.
- Rosenthal MB (1999). Risk sharing in managed behavioral health care. *Health Affairs* 18(5):204-13.

- Rosenthal MB (2000) Risk sharing and the supply of mental health services. *Journal of Health Economics* 19(6):1047-1065.
- Silverman E and Skinner J (2004). Medicare Upcoding and Hospital Ownership. *Journal of Health Economics* 23(2):369-89.
- Soderberg P, Tungstrom S, Armelius BA (2005). Reliability of global assessment of functioning ratings made by clinical psychiatric staff. *Psychiatric Services*. 56(4):434-8.
- Vuong QH (1989) "Likelihood Ratio Tests for Model Selection and Non-Nested Hypothesis," *Econometrica*, 57:307-334.

Figure 1. Changes in Case Rates, Apr. 95 - Dec. 98

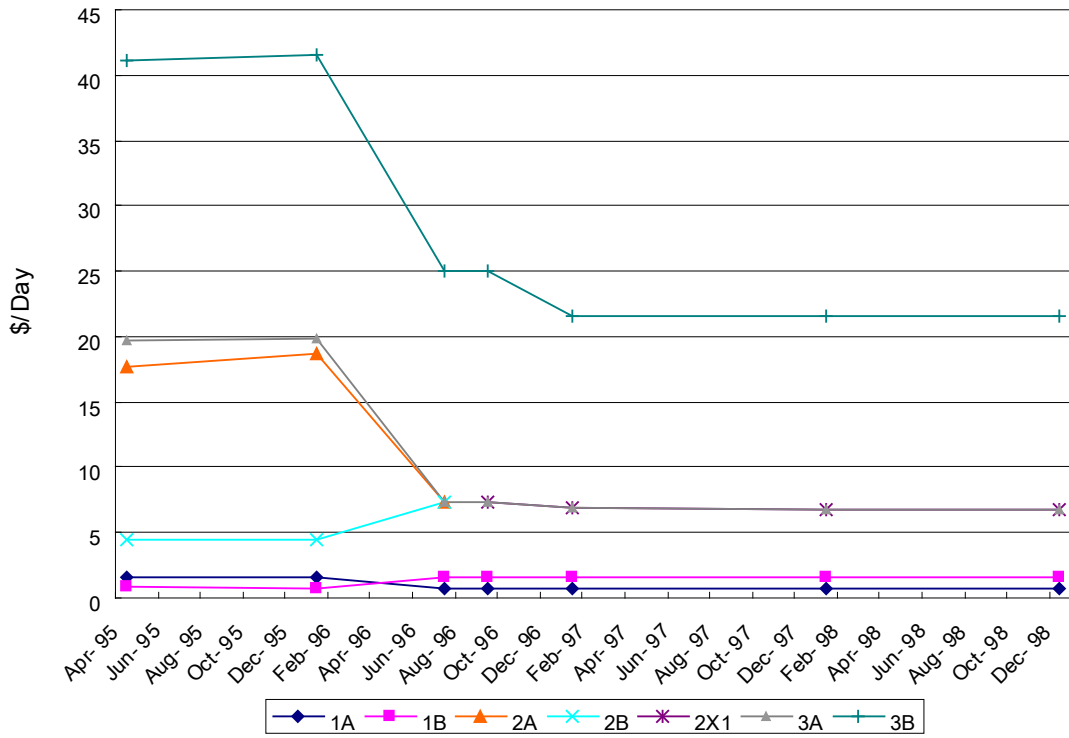


Table 1: Sample Description (n=283,322), weighted with sampling weights and months of use

Variable	Mean	Standard deviation
Age at services use	38.62	0.12
Female gender	0.544	
White (referent category)	0.685	
African American	0.130	
Asian	0.061	
Native American	0.014	
Other Race	0.103	
Severely Mentally Ill	0.772	
Number of months of tier participation (per person)	20.40	14.06
Daily Tiered payments	7.707	0.046
Percent of person-months in each severity category (first period)		
1	9 %	
2	7 %	
3	1 %	
4	53 %	
5	29 %	
6	3 %	
Percent of person-months in each severity category (second period)		
1	2 %	
2	5 %	
3	46 %	
4	44 %	
5	4 %	

Table 2: Tier Selection model estimates of daily tier payments

	Conditional Logit Results	
	Period 1	Period 2
<u>Full Sample (1995-1998)</u>		
Coefficient	0.0009 (0.0038)	0.150* (0.071)
Marginal Effect		
Tier1	0.00002	0.00086
Tier2	0.00002	0.00115
Tier3	2.22e-07	0.00004
Tier4	0.00009	0.00003
Tier5	0.00002	0.00131
Tier6	3.19e-07	
<u>Truncated Sample (1995-1997)</u>		
Coefficient	0.0009 (0.0038)	0.137* (0.066)
Marginal Effect		
Tier1	0.00002	0.00103
Tier2	0.00002	0.00142
Tier3	2.22e-07	0.00005
Tier4	0.00009	0.00003
Tier5	0.00002	0.00125
Tier6	3.19e-07	
<u>SMI Sample (1995-1998)</u>		
Coefficient	-0.0015 (0.0043)	0.0137 (0.079)
Marginal Effect		
Tier1	-0.00002	0.00428
Tier2	-0.00005	0.00488
Tier3	-5.34e-07	0.00039
Tier4	-0.00015	0.00024
Tier5	-0.00004	0.01272
Tier6	-8.17e-07	

Note: All models presented weighted results and control for the maximum GAF score, age, gender, an array of race/ethnicity variables, and quarterly dummy indicators. Standard errors are adjusted for clustering on individuals. **p<0.01; *p<0.05.

Table 3: Coefficient estimates of the association of service days with Daily Tiered Rates

	Logit Results on any Service Use		Negative Binomial Results on Service Days	
	Period 1	Period 2	Period 1	Period 2
<u>Full Sample (1995-1998)</u>				
Model 1: overall price				
Coefficient	0.0228** (0.0044)	0.015 (0.052)	0.0206** (0.0017)	-0.018 (0.021)
Marginal Effect	0.0042	0.0036	0.0538	-0.0351
Model 2: price*tier interactions				
Tier 1				
Coefficient	2.05** (0.16)	-4.61 (3.22)	0.32 (0.24)	-0.85 (4.02)
Marginal Effect	0.0000	-0.0000	2.4515	-0.0045
Tier2				
Coefficient	-1.26** (0.14)	0.27 (1.72)	-0.16 (0.17)	-2.39 (1.52)
Marginal Effect	-0.0006	0.0387	-0.0350	-0.0000
Tier3				
Coefficient	0.076 (0.072)	0.04 (0.11)	0.319** (0.074)	-0.26* (0.10)
Marginal Effect	0.0143	0.0096	0.0475	-0.3869
Tier4				
Coefficient	-0.453** (0.015)	0.12 (0.12)	-0.109** (0.010)	-0.003 (0.102)
Marginal Effect	-0.1108	0.0232	-0.1687	-0.0067
Tier5				
Coefficient	0.1095** (0.0045)	-0.014 (0.045)	0.0353** (0.0021)	-0.018 (0.021)
Marginal Effect	0.0197	-0.0012	0.1528	-0.0560
Tier6				
Coefficient	0.036 (0.018)		0.0185** (0.0036)	
Marginal Effect	0.0058		0.1395	
<u>Truncated Sample (1995-1997)</u>				
Model 1: overall price				
Coefficient	0.0228** (0.0044)	0.088 (0.051)	0.0206** (0.0017)	-0.002 (0.018)
Marginal Effect	0.0042	0.0202	0.0538	-0.0045
Model 2: price*tier interactions				
Tier1				
Coefficient	2.05**	-6.80	0.32	5.49

	(0.16)	(5.59)	(0.24)	(7.07)
Marginal Effect	0.0000	-0.0000	2.4515	2.0e+18
Tier2				
Coefficient	-1.26**	-0.38	-0.16	-2.91
	(0.14)	(2.11)	(0.17)	(2.10)
Marginal Effect	-0.0006	-0.0086	-0.0350	-0.0000
Tier3				
Coefficient	0.076	0.08	0.319**	-0.09
	(0.072)	(0.13)	(0.074)	(0.13)
Marginal Effect	0.0143	0.0208	0.0475	-0.1526
Tier4				
Coefficient	-0.453**	0.44**	-0.109**	-0.06
	(0.015)	(0.15)	(0.010)	(0.13)
Marginal Effect	-0.1108	0.0533	-0.1687	-0.0634
Tier5				
Coefficient	0.1095**	0.046	0.0353**	-0.006
	(0.0045)	(0.044)	(0.0021)	(0.022)
Marginal Effect	0.0197	0.0017	0.1528	-0.0010
Tier6				
Coefficient	0.036		0.0185**	
	(0.018)		(0.0036)	
Marginal Effect	0.0058		0.1395	

SMI Sample (1995-1998)

Model 1: overall price				
Coefficient	0.0335**	0.048	0.0196**	-0.015
	(0.0045)	(0.061)	(0.0017)	(0.019)
Marginal Effect	0.0053	0.0104	0.0562	-0.0318
Model 2: price*tier interactions				
Tier1				
Coefficient	1.96**	-14.75	0.14	13.87
	(0.19)	(7.87)	(0.26)	(9.94)
Marginal Effect	0.0000	-0.0000	0.3814	1.4e+45
Tier2				
Coefficient	-1.25**	0.09	-0.14	-4.81**
	(0.16)	(2.52)	(0.19)	(1.76)
Marginal Effect	-0.0009	0.0050	-0.0355	-0.0000
Tier3				
Coefficient	0.02	0.12	0.32**	-0.26*
	(0.11)	(0.15)	(0.11)	(0.11)
Marginal Effect	0.0023	0.0290	0.0493	-0.5253
Tier4				
Coefficient	-0.390**	0.19	-0.116**	-0.01
	(0.017)	(0.15)	(0.012)	(0.11)
Marginal Effect	-0.0855	0.0106	-0.1938	-0.0026

Tier5				
Coefficient	0.1013**	-0.001	0.0347**	-0.019
	(0.0050)	(0.050)	(0.0022)	(0.021)
Marginal Effect	0.0173	-0.0000	0.1561	-0.0001
Tier6				
Coefficient	0.030		0.0158	
	(0.016)		(0.0035)	
Marginal Effect	0.0048		0.1279	

Note: All models presented weighted results and control for the maximum GAF score, age, gender, an array of race/ethnicity variables, and quarterly dummy indicators. Standard errors are adjusted for clustering on individuals. **p<0.01; *p<0.05.