The Journal of Mental Health Policy and Economics
J Ment Health Policy Econ **23**, 81-91 (2020)

Putting Providers At-Risk through Capitation or Shared Savings: How Strong are Incentives for Upcoding and Treatment Changes?

Marisa Elena Domino,[1] Edward C. Norton,[2] Jangho Yoon,[3] Gary S. Cuddeback,[4] Joseph P. Morrissey[1]

[1]*Ph.D., Department of Health Policy and Management and Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*
[2]*Ph.D., Department of Health Management and Policy, Department of Economics, Population Studies Center, University of Michigan, Ann Arbor, MI, USA, and National Bureau of Economic Research, Cambridge, MA, USA*
[3]*Ph.D., Department of Preventive Medicine and Biostatistics, F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA*
[4]*Ph.D., School of Social Work, Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*

Abstract

Background: Alternative payment models, including Accountable Care Organizations and fully capitated models, change incentives for treatment over fee-for-service models and are widely used in a variety of settings. The level of payment may affect the assignment to a payment category, but to date the upcoding literature has been motivated largely incorporating financial penalties for upcoding rather than by a theoretical model that incorporates the downstream effects of upcoding on service provision requirements.
Aims of the Study: In this paper, we contribute to the literature on upcoding by developing a new theoretical model that is applicable to capitated, case-rate and shared savings payment systems. This model incorporates the downstream effects of upcoding on service provision requirements rather than just the avoidance of penalties. This difference is important especially for shared-savings models with quality benchmarks.
Methods: We test implications of our theoretical model on changes in severity determination and service use associated with changes in case-rate payments in a publicly-funded mental health care system. We model provider-assigned severity categories as a function of risk-adjusted capitated payments using conditional logit regressions and counts of service days per month using negative binomial models.
Results: We find that severity determination is only weakly associated with the payment rate, with relatively small upcoding effects, but that level of use shows a greater degree of association.
Discussion: These results are consistent with our theoretical predictions where the marginal utility of savings or profit is small, as would be expected from public sector agencies. Upcoding did seem to occur, but at very small levels and may have been mitigated after the county and 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 significant in the first period, and potentially at a clinically important level. Limitations include data from a single county/multiple provider system and potential unmeasured confounding during the post-implementation period.
Implications for Health Care Provision and Use: Providers in our data were not at risk for inpatient services but decreases in use of outpatient services associated with rate decreases may lead to further increases in inpatient use and therefore expenditures over time.
Implications for Health Policies: Health program directors and policy makers need to be acutely aware of the interplay between provider payments and patient care and eventual health and mental health outcomes.
Implications for Further Research: Further research could examine the implications of the theoretical model of upcoding in other payment systems, estimate the power of the tiered-risk systems, and examine their influence on clinical outcomes.

Received 23 October 2019; accepted 17 June 2020

Background and Motivation

Pre-paid, or capitated payments in health insurance programs change incentives for treatment and are increasingly used in a variety of health care settings. Capitation sets a payment for health services use during a specified time period, which may be risk adjusted directly or based on historical benchmarks. The Medicare program, for example, has implemented risk-based provider payments through a large number of mechanisms, including Medicare Advantage, Medicare Part D, bundled payments, case-rate hospital and skilled nursing facility payments, and other types of alternative payment models.[1,2] Capitation rates that are too low can discourage the provision of high-cost, high-intensity services and may

* **Correspondence to:** Marisa Elena Domino, Ph.D., Department of Health Policy and Management and Cecil G. Sheps Center for Health Services Research, The University of North Carolina at Chapel Hill, 135 Dauer Dr., Chapel Hill, NC 27599-7411, USA.
Tel.: +1-919-966 3891
E-mail: domino@unc.edu
Source of Funding: 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 and MH065639).

encourage shifting to substitute services outside the capitation payment that may be free goods to the provider.[3,4] Capitation rates that are too high result in over payments and may distort the incentives to participate in the payment scheme by increasing the supply of providers, further driving up price; this has been observed in the market for specialty hospitals in the Medicare program.[5] Capitation is generally at the group level, while individual providers are still predominantly paid by fee-for-service.[6]

Certain forms of capitated payments, such as diagnosis-based case-rate or tiered payments, create incentives to change both diagnosis and treatment patterns that may differ from population-based capitated models. Population-based models pay a fixed amount per-person per-month for all covered individuals, regardless of whether services are used or not. Case-rate payments are used to pre-pay for care for only those covered individuals engaged in treatment, leaving the sponsoring agency or employer at risk for changes in both the number of individuals using services as well as the intensity of service-system users' needs. In contrast, provider groups 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.

The Medicare DRG system of payment for inpatient care as a function of diagnosis or severity is a case-rate system, although one that only lasts as long as the individual is in the hospital. Outpatient procedure-based payments, common in the fee-for-service system in the U.S., can also be considered a form of case-rate payments, and have been shown to similarly suffer from upcoding.[7] Bundled payment models, such as the Bundled Payments for Care Improvement (BPCI) Initiative[8] or the Oncology Care Model,[9] are another form of capitation with a limited scope and time period. 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.[10,11] In addition, the size of the case-rate payment can affect the level of service use. However, the level of case-rate payment may also affect the assignment to a severity category, a phenomenon noted in the Medicare DRG literature as ''creep'' or upcoding.[12-14]

Notable economic analyses of upcoding have found considerable evidence of its occurrence after exogenous price changes. Dafny examined how changes in prices from a policy reform in hospital DRG payments in 1988 that eliminated DRG categorization based on age affected the use of diagnostic categories[15] and found considerable evidence of large increases in the use of DRGs with the largest price changes concentrated among for-profit hospitals. Sacarny[16] similarly looked at hospitals' response to incentives to report more information that could potentially affect revenues from heart failure admissions. Brunt[7] examined upcoding in Medicare-funded general office visits using the CPT coding system and similarly found evidence that changes in relative Medicare payments increase the odds of upcoding. In addition to directly increasing health sector payments, upcoding on diagnostic codes may have significant downstream effects if those codes form the basis of risk-adjustment payments in the future.[17] Geruso and McGuire[18] further unpack the ''power'' of a payment system in terms of whether greater utilization yields higher payments through new diagnoses that may trigger a different risk-adjusted payment. However, to date the upcoding literature has been motivated largely incorporating financial penalties for upcoding rather than by a theoretical model that incorporates the downstream effects of upcoding on service provision requirements. This difference is important especially for shared-savings models with quality benchmarks.

In this paper, we contribute to the literature on upcoding by developing a new theoretical model that is applicable to capitated, case-rate and shared-savings payment systems. We also explore the implications from this model empirically in a case study of a single county mental health system that implemented a case-rate payment system for outpatient behavioral health services, examining the level of responsiveness of mental health providers to changes in case-rate payments. This research will help policy makers and managed health care organizations further refine rate-setting in the current generation of at-risk alternative payment models, and estimates provider behavior under an alternative payment model in behavioral health.

Method

Theoretical Model

We motivate a new model of provider behavior in selecting the type and number of treatment services for their patients in the context of at-risk payment models using a utility-maximization framework. That is, unlike other models of upcoding (Brunt[7] and Bowblis and Brunt[19]) that rely on financial penalties to explain provider upcoding, we begin with a utility maximizing provider who has imperfect information on patient severity of illness to better understand why upcoding may persist. We adapt our approach from McGuire,[20] which models provider profit as a mix of a single-capitated payment or fee-for-service payments, including corner solutions. Our model extends this profit equation to allow capitated payments as a function of mental illness severity, as would be the case with risk adjustment. This framework is relevant for many types of alternative payment models, including full and partial capitation models, including the current ACO model, which use payment systems built on shared savings calculated as the difference between actual expenditures and historical averages, adjusted for health sector inflation. Historical benchmarks are based on practice case-mix and thus vary across different practices by prior caseload severity.

For simplicity, we assume patients are one of two types: high severity, H, and low severity, L. As is the case in markets with asymmetric information, the type is unobservable to the payor sponsor or principal, but observable to the provider agent. The provider 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, or upcode patients because of unobserved heterogeneity. The tradeoff for classifying an individual as a higher severity is that while reimbursement is higher,* the level or intensity of services, x_H, provided must be higher too, due to administrative monitoring by the payor. The provider's goal is to maximize

$$U(NB, \Pi) \tag{1}$$

over the number of patients classified as ϕ_H or ϕ_L and choice of x_i, the number or intensity of services provided to a patient classified as i, where the arguments NB and Π represent patient net benefits and provider net income from treating the patient, respectively, with $U_{NB} > 0$, $U_\Pi > 0$, $U_{\Pi\Pi} < 0$, $U_{NBNB} < 0$, $U_{NB\Pi} < 0$. For simplicity, we assume that patients and providers are price takers and that patients do not switch providers.

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

$$NB_i = B^i(x_j) - px_j \tag{2}$$

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, indicating a non-decreasing slope of the health production function for persons of either type, $B^i_x \geq 0$. This implies that more services may improve health at a decreasing rate with $B^L_x \leq B^H_x\ \forall x$. 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 become arbitrarily close 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 $\phi_L$since it results in a lower level of reimbursement and an under provision of services, yielding an unambiguously lower level of provider utility.[1] Our model contrasts with Brunt[7] and Bowblis and Brunt[19] in that here the mechanism that prevents corner solutions is the inefficiency of having to provide more services to an upcoded patient, whereas in Brunt[7] and Bowblis and Brunt,[19] provider utility is a function of profit only, and thus the probability of a financial penalty reduces the likelihood of upcoding.

Provider-level net benefits are therefore a weighted sum of the three types of patients coding options: N_L refers to the number of patients of type L who are coded as type ϕ_L; N_H is the number of patients of type H who are coded as type ϕ_H; and δ_L are the number of patients of type L who are upcoded as type ϕ_H. Each are normalized by the provider's total caseload such that $N_L + N_H + \delta_L = 1$.

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

If providers only maximized net health benefits, then services would be provided at a level where the health benefit production curve is tangent to the patient's price of service receipt, p. We refer to these points as x^E_i. Because the slope of B^H is assumed to be steeper than the slope of B^L, this implies that more services would be provided to Hs than Ls. This also implies that patients with lower price of service receipt, p, would be provided a greater level of services.

However, in this model, providers also receive utility from profits. Providers are paid on a capitated basis, with different capitation rates based on reported patient severity. Note that our model assumes a case-rate payment approach, which is the payment model in our empirical example. A pure capitation approach, where a lower capitation rate is paid for both service users and service non-users alike, leaving the providers at risk for the probability of any service use or changes in the external margin, will share the same incentive structure examined here if the size of the service-using population is exogenously determined or the size of non-users is trivial. The model could accommodate an endogenous fraction of service non-users with an additional choice variable, N_0, such that $N_L + N_H + \delta_L + N_0 = 1$.

Profits are defined as the difference between the capitated rate received and the total cost of treatment provision.

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

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.

The provider's task is to choose N_L, N_H, δ_L, x_L and x_H in order to maximize utility. Letting $N_H = 1 - N_L - \delta_L$ and taking the derivative with respect to the remaining four choice variables N_L, x_L, δ_L, and x_H, respectively, leads to the following first-order conditions (equations 5-8):

$$B^L_x(x_L) - p = \frac{U_\pi c}{U_{NB}} \tag{5}$$

Since the right-hand side of equation 5 is positive, this equation indicates that services will be under-provided to accurately-coded Ls, since the marginal benefit of services is still greater than their opportunity costs. This is the classic result in capitated health systems.

$$\begin{aligned} &B^L(x_L) - B^H(x_H) - p(x_L - x_H) = \\ &= -\frac{U_\pi}{U_{NB}}(R_L - cx_L - R_H - cx_H) \end{aligned} \tag{6}$$

which indicates that if Hs are more profitable than Ls, then services will be underprovided more for Ls than Hs, with respect to the net-benefit maximizing point.

Equation 6 also leads to the testable implication that forms the basis of the empirical work that follows. As the capitated payment for high severity patients, R_H, increases, x_H will increase towards the net-benefit maximizing point. Similarly,

* If $R_H = R_L$, there is no incentive to upcode.

as the payment for low-severity patients increases, the level of service provision for *Ls* will also increase. Therefore, the model implies that an increase in capitated payments will lead to greater service provision; but this effect can be diluted if U_Π is small. This also has implications for ACOs in that if shared savings are a fraction of savings from full capitation, then service provision incentives may again be mitigated.

$$U_{NB}\big(B^L(x_H) - px_H\big) = -U_\pi(R_H - cx_H) \qquad (7)$$

Equation 7 gives the expected result that misclassified or upcoded patients will be over-provided services, such that the benefits are less than patient costs.

$$\delta_L^* = \frac{(1 - N_L)\big[U_\pi c - U_{NB}\big(B_x^H(x_H) - p\big)\big]}{U_{NB}\big(B_x^L(x_H) - B_x^H(x_H)\big)} \qquad (8)$$

Equation 8 has a number of implications. First, all providers will upcode some patients, that is $\delta_L^* > 0$, whenever $\big(B_x^H(x_H) - p\big) > \frac{U_\pi c}{U_{NB}} > 0$. This implies that the level of services provided to high severity patients will be less than would be optimal, x_H^E. Second, δ_L^* increases with greater marginal utility of profits, the greater marginal cost of service provision, and the closer the two benefit production curves are to each other. The derivative of δ_L^* with respect to the capitation rate is ambiguous and depends on the relative magnitude of the second derivative of utility with respect to profits. That is, a higher capitation rate will not always yield a greater rate of upcoding but can vary depending on the relative weight of profits and net benefits in the provider's utility function.

Model Extensions

An alternative specification of this model would be the introduction of noise into the classification problem, such that providers 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 intentionally deceptive behavior. Either interpretation can be used to motivate the phenomenon of upcoding or ''DRG creep''.[21]

We also consider the case of a partial capitation system, where only a fraction, a, of the difference between revenues and costs from equation[4] are retained by the provider, such as with risk corridors or payments adjusted post-hoc for quality of service provision, such as in ACO models:

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

where $a \in (0, 1]$. The analogous first order conditions indicate that the size of under- and over-provision of services to correctly and upcoded patients decreases under this payment system. Finally, if a is not exogenously determined, but is a function of service provision, $a(x_i)$, as would be the case for quality adjusted payments based on preventative health services, such as cancer screening, we again find that this modification of equation (4) decreases the incentives for under-provision of services, depending on the strength (magnitude) of $a(x_i)$.

Data Analytic Procedures

We now turn towards testing several implications of this model. Following other studies in this area, our identification strategy uses exogenous price changes.[15] First, we examine the effect of changes in the case-rate payments on the assignment of mental illness severity level. The theoretical model results were ambiguous in this regard. Unlike other upcoding studies,[5] we use individual patient-level, rather than aggregated provider-level data, which allows us to control for sociodemographic characteristics of individuals. Second, we examine the association between changes in case-rate payments and service use in each severity category. The theoretical model predicts that an increase in payments will increase the number of services provided within a severity level because of the incentive to upcode. We use service days, or days on which one or more services were provided to an individual, rather than total costs or intensity of services provided for reasons explained below. We carefully avoid making causal links, since it is possible that an unaccounted for third factor causes both changes in case-rate, or tiering rates and severity determination, although we have uncovered no such credible candidate.

An ordered logit model would typically be used to examine severity assignment as a function of the payment rate for each category.[7,19] In our application, the ranking of the tiers in terms of the minimum severity level set by clinical criteria changed over time (described further below) indicating that an ordered model would not be appropriate. We therefore run the tier selection model as an unordered mixed or McFadden's conditional logit model,[22] 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, and race/ethnicity (White, African-American, Asian, Native American, Latino/a, or other race/ethnicity). That is, we estimate $p_{ij} = \frac{e^{X_i\beta_j + Z_j\beta}}{1 + \sum_{k=2}^{J} e^{X_i\beta_k + Z_k\beta}}$, the probability individual i is classified in category j, as a function of person-level characteristics, X_i, and category-specific characteristics, Z_j, such as payment rates. Standard errors were adjusted for clustering based on repeated observations on individuals.* Because conditional logit coefficients do not provide the direction or magnitude of the estimated effects, we report average marginal effects of payment rates overall and for each severity category or tier and delta-method standard errors. That is, we use this model to test whether changes in payment levels affect the severity assignment as the sign of this derivative is ambiguous from equation 8.

We also examine the response of service days to daily payment rates using negative binomial count models on the

* Clustering on providers is not observable in our data but assumed to be accommodated by clustering on patients.

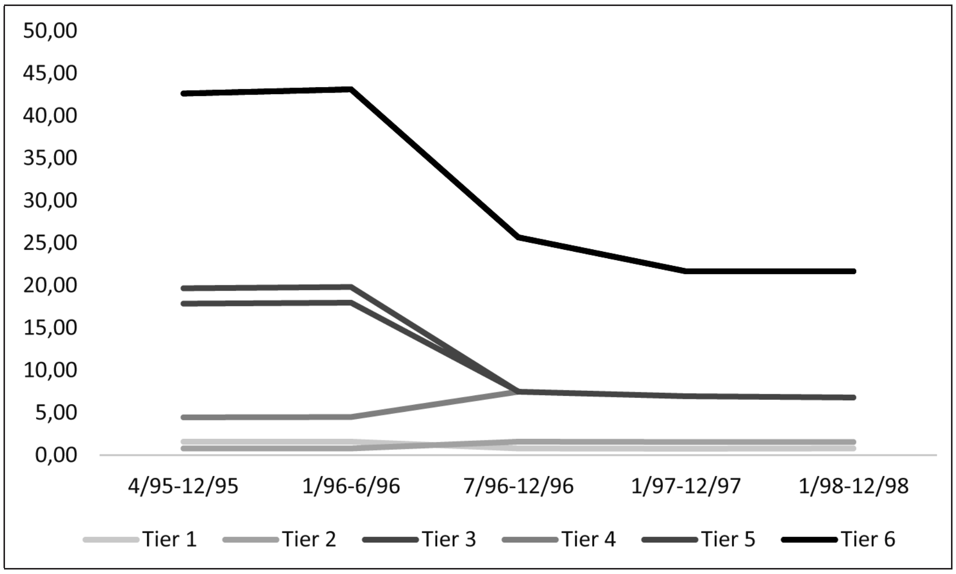

Figure 1. Average Daily Prices by Tier for Each Time Period.

number of service days used in each month. We ran these models stratified by the two time periods when six and five tiers were available, respectively. The marginal effect of the payment rate is again the key result of interest in testing the second hypothesis from equation 6. The models controlled for age, gender, race/ethnicity and time trends. The model was run both across categories with a price term and price-severity level interactions, and conditional on each severity category, to estimate within-tier price effects controlling for selection into each tier. We also examine the association between payment rates and use at the external margin by simulating the effect of price changes on the probability of not observing any services used in a month.

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 (Seattle), Washington after the implementation of a behavioral health case-rate capitation payment system. Payment rates were set by the state government, but a private managed care organization provided oversight on the intensity of use. The sample for this study comes from the King County Mental Health, Chemical Abuse and Dependency Services Division (KCMHCADS) in Washington State. On April 1, 1995, King County implemented a tiered-payment system affecting participating outpatient mental health provider groups, switching from the previous fee-for-service system. Under this system, risk is borne by local provider organizations for Medicaid-enrolled users of mental health services.[3] Typically, these were large community mental health agencies with a variety of programs and services, including homeless services, case management, assertive community treatment teams, integrated dual disorder treatment, medication management and general outpatient services. Providers delivered services to a mix of individuals across tiers (i.e., ranging from those with less serious disorders and less intensive needs to those with the most profoundly disabling disorders and most intensive service needs). Crisis-related and residential services were carved out of the capitated system and reimbursed under separate contracts.

Participating providers assigned a severity ‘‘tier’’ to consumers in the system; tiers were preauthorized for a fixed period of time. Initially providers could choose among six tiers for their capitated patients. In July 1996, the tiering system was restructured such that the two middle categories were collapsed into a single tier. Each tier had an associated daily payment rate, which depended on patient age (child less than 21 generally received a separate rate, 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. Approximately 10% of tiered patients received these special payment rates. *A posteriori* payment rates, not observable in our data, were sometimes reduced post hoc from published rates 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 five different price periods that 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 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; each takes on numeric values ranging from 0 to 100, with higher scores indicating greater functioning.[23,24] A maximum GAF score, which reflects the difference in health 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 the time observed in this study and no changes in the minimum GAF scores occurred after July 1, 1996. We do not observe actual GAF scores for study participants, however, so we are unable to examine how actual coding varied over time.

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 complexity of the changing treatment climate. 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.

After tiered 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. As a check on data quality, we plotted average service use over time and did not find a visible break in monthly service use trends from before to after the implementation of the tiering system.

We derived a number of measures including tier assignment and service use. We collapsed a comprehensive list of mental health services provided by KCMH, including outpatient visits, intake assessments, telephone contacts, advocacy and linkage, case management and case termination-related activities, down to a daily measure of any outpatient service use, regardless of service type. Daily services were summed to obtain a measure of the number of service days per month for each individual. This measure of service days was preferred to other available options for a number of reasons. First, the capitated system may have distorted the incentives to accurately report the type of services using procedure codes, rather than the provision of services per se (e.g., specific HCPCS/CPT codes). Second, counting service days freed us from using dollar amounts as intensity weights, since fee-for-service payments were no longer available in the data after the capitation system was imposed. Service days have some disadvantages, including undercounting services if more than one service was provided per day and equally weighting services of different intensities.

A random sample of individuals using the King County Mental Health system was collected as part of a larger project.[3] Individuals were sampled according to a stratified sample based on their use of KCMHCADS services, the King County Jail, and 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 implementation date of the tiered system) until December 31, 1998, yielding an unweighted sample size of 13,557 individuals contributing a total of 283,322 monthly observations. We retained the original sampling weights, as these return us to the population of individuals who used KCMH services during the full study period. Individuals who were not assigned to a tier and used only non-tiered services were excluded in the present analysis. We additionally removed months in which any days in jail were observed from the estimation sample for the count models. 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 study period.

We calculated daily tiered payment rates at the person-month level, based on the age category (<18, 18-59, and ≥ 60), date of tier authorization, and the tier assigned. We are unable to make adjustments for the specialty rate differentials described above because our data do not contain information 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

Additional analyses were conducted on individuals identified as having a severe mental illness (SMI), defined for this paper using diagnoses of schizophrenia, bipolar disorder or major depression. This indication was drawn from diagnoses in all data sources available to the larger project, including state psychiatric hospitals, general hospitals, and jails. Individuals with SMI 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.

Because the largest price change occurred in July 1996 (**Figure 1**), we run models of severity selection with shorter windows of observation around that time period to increase identification of price effects: one analysis uses a 6-month window (Jan – Dec 1996) and one analysis uses a 12 month window (July 1995 – June 1997). In January of 1998, the county implemented a system of recoupments called a risk corridor. In part, this risk corridor may resemble shared savings models that adjust provider payments based on the

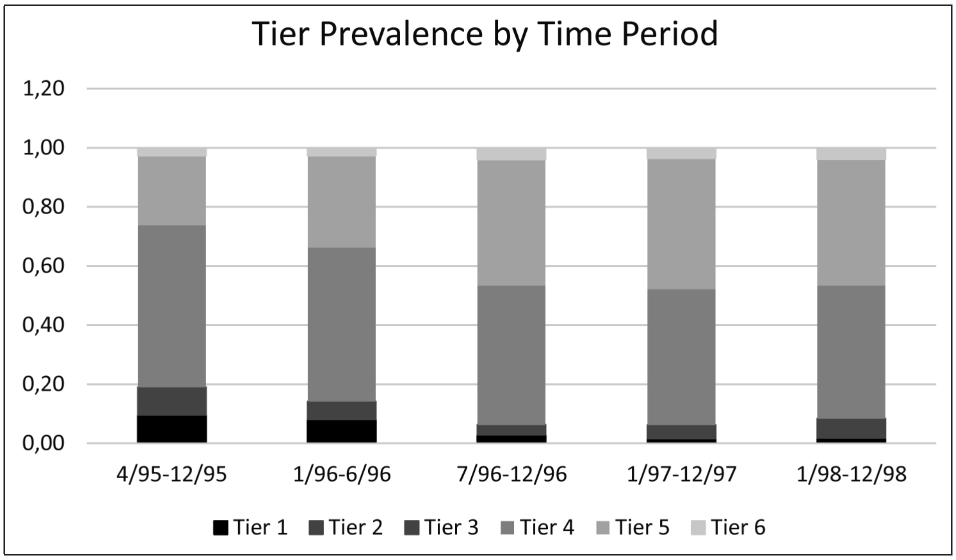

Figure 2. Changes in Tier Prevalence by Key Time Periods.

Table 1. Sample Description*, Weighted with Sampling Weights and Months of Use.

Variable	Mean / %	Standard deviation
Age at services use	36.4	11.3
Female gender	57.2%	
White (referent category)	68.9%	
African American	13.5%	
Other Race	17.6%	
Severely Mentally Ill	67.8%	
Number of months of tier participation (per person)	20.9	14.1
Daily Tiered payments	7.70	5.81

* NT=283,322; N=13,557.

attainment of various quality indicators. In King County, 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 conducted a sensitivity analysis by excluding data from 1998.

Results

Variable means are presented in **Table 1**. The mean age of the sample was 36 years old, ranging from 18-64. Just over half were female (57%) and the majority were white (69%), with 14% African-American, and 18% classified as Other races. Over two-thirds (68%) of the sample were classified as SMI. The average length of participation in the tiering system was just under two years (21 months), with a range of 1 to 45 months. The distribution of assigned categories by time period is displayed in **Figure 2**. The two largest tiers account for the majority of observations throughout the study period, although the use of tiers changed over time. Finally, the unadjusted mean encounter days for each tier are displayed in **Figure 3**. The trends follow the price schedule from **Figure 1**, with utilization generally declining as prices declined until July 1996, then began to increase.

Table 2 presents the average marginal effects of daily payment rates on the tier assignment from the mixed logit models. We find a positive association between payment rates and tier selection, suggesting an upcoding effect. This effect was relative constant across model specifications. The effects ranged from 0.0026 to 0.0116 across tiers. These effects are very small, however, with a $1 increase in daily payments (approximately a 13% increase from the mean payment) associated with a 1% point increase in tier assignments in each category. The upcoding effect is slightly larger for individuals with SMI and also using a smaller window of identification around the largest price change, as compared with results from the full population and study period.

Results from the negative binomial models (**Table 3**) indicate that increasing payments are associated with greater

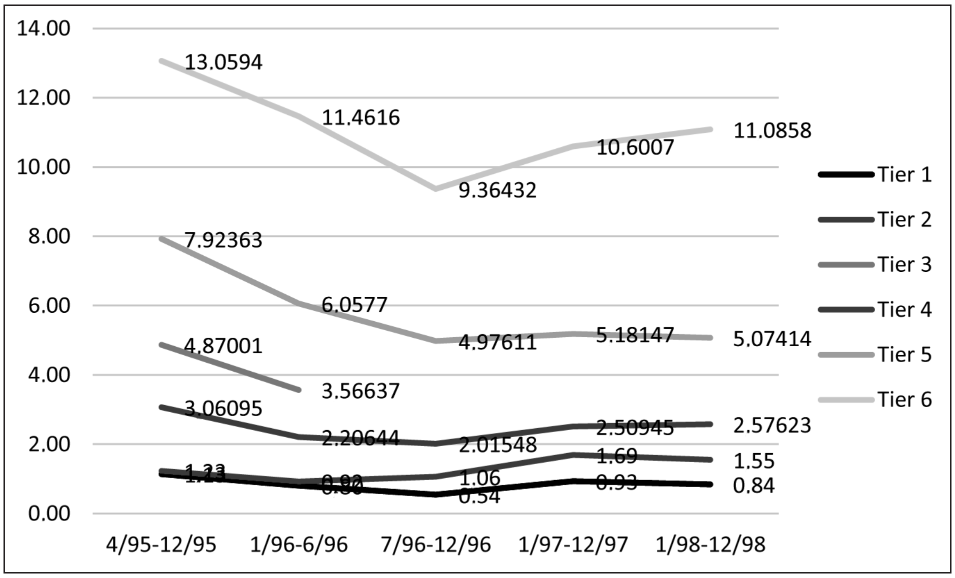

Figure 3. Changes in Average Number of Encounters Per Month by Tier and Time Period.

Table 2. Average Marginal Effects of Daily Tier Payments on Tier Selection from Conditional Logit Models.

		Time-limited models			
	Full sample	Six-month window (1996 only)	One-year window (July 1995 - June 1997)	Excluding 1998	SMI only
Overall effect	0.0078*	0.0084*	0.0076*	0.0078*	0.0091**
	(0.0032)	(0.0035)	(0.0032)	(0.0032)	(0.0030)
Estimated for:					
Tier1	0.0116**	0.012**	0.0109**	0.0111**	0.0141**
	(0.0039)	(0.0038)	(0.0037)	(0.0037)	(0.0036)
Tier2	0.0088**	0.0110**	0.0093**	0.0092**	0.0091**
	(0.0028)	(0.0037)	(0.0031)	(0.0030)	(0.0022)
Tier3	0.0055**	0.00185	0.0029*	0.0049**	0.0081**
	(0.0019)	(0.00095)	(0.0013)	(0.0019)	(0.0021)
Tier4	0.00263**	0.0033**	0.0029**	0.0029**	0.00272**
	(0.00096)	(0.0012)	(0.0010)	(0.0010)	(0.00083)
Tier5	0.0049*	0.0064*	0.0059*	0.0056*	0.0051**
	(0.0020)	(0.0029)	(0.0025)	(0.0024)	(0.0018)
Tier6	0.0090**	0.0078**	0.0071**	0.0077**	0.0108**
	(0.0030)	(0.0020)	(0.0023)	(0.0025)	(0.0033)
Observations	1,520,461	440,775	845,194	1,126,636	1,205,733
Unique individuals in sample	13,557	8,934	10,528	11,522	9,509

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. Delta method standard errors are adjusted for clustering on individuals.
$^{**}p<0.01$; $^{*}p<0.05$.

provision of services overall in both time periods, with a $1 increase associated with 6 more visits per 100 population in Period 1, and 38 more visits per 100 people in Period 2. The overall effect of price is slightly larger in the SMI population in Period 1 but is not significant for persons with SMI in Period 2. Individual price effects within tiers are positive in all tiers in Period 1 and significant in 4 out of 6 tiers, ranging in magnitude from 3 to 51 more outpatient visits in a population of 100 for a $1 increase in price. Price effects among the SMI are similar in Period 1, although the magnitude of the effect in the lowest tier is reduced by 50% and loses significance, while the magnitude of the effect in Tier 3 nearly doubles. Price effects were generally insignificant in Period 2.

Table 3. Average Marginal Effects of Daily Tier Rate on the Number of Service Days in Each Month.

	Average Marginal Effects on count of service days in month - Full sample		Average Marginal Effects on count of service days in month - SMI only		Average Marginal Effects on count of service days in month - Truncated Sample
Full Sample	Period 1	Period 2	Period 1	Period 2	Period 2
Overall	0.062**	0.38*	0.071**	0.27	0.27
	(0.016)	(0.19)	(0.019)	(0.24)	(0.20)
Estimated for:					
Tier1	0.51**	–1.74	0.31	–3.36	0.41
	(0.14)	(2.82)	(0.17)	(5.44)	(2.99)
Tier2	0.29*	–0.43	0.32*	–2.10	–0.59
	(0.12)	(1.58)	(0.13)	(1.89)	(1.70)
Tier3	0.36*		0.74*		0.07
	(0.14)		(0.30)		(0.17)
Tier4	0.001	0.18	0.032	0.07	0.43
	(0.019)	(0.18)	(0.024)	(0.22)	(0.37)
Tier5	0.0319**	0.63	0.0366**	0.55	0.35
	(0.0090)	(0.38)	(0.0097)	(0.41)	(0.21)
Tier6	0.010	0.41*	0.006	0.45*	
	(0.045)	(0.20)	(0.048)	(0.19)	

Note: All models presented weighted results and control for age, gender, an array of race/ethnicity variables, and time trends. Period 1 models examine the time period from April 1995 - August 1996, which had 6 tiers, and also control for maximum gaf score, which did not change within tiers during Period 2, which ran from September 1996 - Decmeber 1998. Standard errors are adjusted for clustering on individuals.
**p<0.01; *p<0.05.

We also examine the association between payment rates and use at the external margin by simulating the effect of price changes on the probability of not observing any services used in a month in **Table 4**. In Period 1, we find that price increases do increase the external margin in the sense of decreasing the probability of having no services received in a month. The effects are not trivial, with a $1 increase in the payment rate was associated with a 1.4% point increase in the probability of using some services in the month overall, with specific effects within tiers ranging from less than 1% point increase to over 12 % point increase (Tier1). The effects were again similar in the SMI sample. In Period 2, however, most of these price effects have again decreased in magnitude or significance.

Discussion

Daily tier payments were seen to be weakly associated with changes in tier assignment in the King County mental health service system. That is to say, upcoding did seem to occur, but at very low levels. Several prior studies generally found differences between for-profit and not-for-profit hospital providers in their use of upcoding.[25] Mental health providers examined here were all non-profit entities, which may help explain the weak evidence of upcoding. Steinbusch *et al.*[26] also speculate that systems that are less vulnerable to upcoding individuals are locked into a severity category for longer periods of time, as was the case in the second period examined here.

The association between the payment levels and the number of service days in a month, however, was significant in the first period, and potentially at a clinically important level. This may be because a number of factors, including the pathway specified by the theoretical model from greater provider utility to patient health through increased service provision. This finding is somewhat different that the framework established by Geruso and McGuire.[18] Here, additional services do not serve as opportunities to diagnose additional conditions, yielding higher payment rates as do Marketplace plans with concurrent risk adjustment. Here, payments would only increase if additional visits serve as an opportunity to observe changes in functioning, which might lead to subsequent changes (positive or negative) in tier assignments. That is, the power of this payment system is higher than those in Marketplace plans, and possibly close to 1.0. We cannot rule out other explanations, such as greater monitoring (or perceived monitoring) 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 it does indicate that rate setting has important

Table 4. Average Marginal Effects of Daily Tier Rate on the Probability of Having No Service Days in a Month.

	Average Marginal Effects on probably of no services received in month - Full sample		Average Marginal Effects on probably of no services received in month - SMI only		Average Marginal Effects on probably of no services received in month - Truncated Sample
Full Sample	Period 1	Period 2	Period 1	Period 2	Period 2
Overall	–0.0138**	0.007	–0.0126**	0.021	0.011
	(0.0039)	(0.023)	(0.0042)	(0.029)	(0.024)
Estimated for:					
Tier1	–0.120**	0.48	–0.070	0.96	–0.11
	(0.033)	(0.78)	(0.038)	(1.55)	(0.86)
Tier2	–0.083*	0.09	–0.091**	0.40	0.12
	(0.032)	(0.32)	(0.035)	(0.38)	(0.35)
Tier3	–0.0195*	–0.019	–0.034*	–0.006	–0.007
	(0.0073)	(0.018)	(0.013)	(0.020)	(0.019)
Tier4	–0.0001	–0.021	–0.0027	–0.017	–0.015
	(0.0019)	(0.013)	(0.0021)	(0.012)	(0.013)
Tier5	–0.00070**	–0.0025*	–0.00074**	–0.0024*	–0.0023
	(0.00020)	(0.0013)	(0.00020)	(0.0010)	(0.0014)
Tier6	–0.00006		–0.00003		
	(0.00025)		(0.00023)		

Note: All models presented weighted results and control for age, gender, an array of race/ethnicity variables, and time trends. Period 1 models also control for maximum gaf score, which did not change within tiers during period 2. Standard errors are adjusted for clustering on individuals.
**p<0.01; *p<0.05.

implications. Providers in our data were not at risk for inpatient services but decreases in use of outpatient services associated with rate decreases may lead to further increases in inpatient use and therefore expenditures over time. This has especially important implications as partially capitated ACO models are being phased in nationwide.

Several caveats are in order. First, data are from a single county/multiple provider system which used an innovative approach to pay for mental health services during a complex administrative period and may not generalize to other systems using a similar payment scheme. Capitation was at the level of large agencies which may or may not have provided strong incentives to individual clinicians. 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 altered the severity mix of individuals over time, although eligibility for Medicaid was not tied directly to the severity categories used in the mental health system. 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 monitoring and sophistication in determining payment levels that give appropriate incentives for optimal care.[27] 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 and mental health outcomes.

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 and MH065639). For assistance in assembling the administrative data and input on their

interpretation we are indebted to the King County Mental Health, Chemical Abuse and Dependency Services Division (King County, Washington State).

References

1. McClellan M, McKethan AN, Lewis JL, Roski J and Fisher ES. A National Strategy to Put Accountable Care into Practice. *Health Aff* 2010; **29**(5): 982-990.
2. Center for Medicare and Medicaid Services [CMS] (2017). *Health Care Payment Learning and Action Network*. Available: https://innovation.cms.gov/initiatives/Health-Care-Payment-Learning-and-Action-Network/. Accessed August 14, 2018.
3. Domino ME, Norton EC, Morrissey JP and Thakur N. Cost Shifting to Jails After a Change to Managed Mental Health Care. *Health Serv Res 2004;* **39**(5): 1379-401.
4. Huskamp HA. 2003. Managing Psychotropic Drug Costs: Will Formularies Work? *Health Aff* 2003; **22**(5): 84-96.
5. Berenson RA, Bazzoli GJ, Au M. 2006. Do Specialty Hospitals Promote Price Competition? *Center for Studying Health System Change Issue Brief* #103. PMID: 16465700.
6. Zuvekas SH and JW Cohen. Fee-For-Service, While Much Maligned, Remains The Dominant Payment Method For Physician Visits. *Health Aff* 2016; **35**(3): 411-414.
7. Brunt CS. CPT fee differentials and visit upcoding under Medicare Part B. *Health Econ* 2011; **20**(7): 831-841.
8. CMS (2018a). *Bundled Payments for Care Improvement (BPCI) Initiative: General Information*. Available: https://innovation.cms.gov/initiatives/bundled-payments/. Accessed August 14, 2018.
9. CMS. 2018b. *Oncology Care Model*. Available: https://innovation.cms.gov/initiatives/oncology-care/. Accessed August 14, 2018.
10. Rosenthal MB. Risk sharing in managed behavioral health care. *Health Aff* 1999; **18**(5): 204-213.
11. Rosenthal MB. Risk sharing and the supply of mental health services. *J Health Econ* 2000; **19**(6): 1047-1065.
12. Simborg DW. DRG creep: a new hospital-acquired disease. *N Engl J Med* 1981 June 25; **304**(26): 1602-1604.
13. Hsia DC, Ahern CA, Ritchie BP, Moscoe LM, Krushat WM. Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988. *JAMA* 1992; **268**(7): 896-899.
14. Carter GM, Newhouse JP, Relles DA. How much change in the case mix index is DRG creep? *J Health Econ* 1990; **9**(4): 411-428.
15. Dafny LS. How Do Hospitals Respond to Price Changes? *Am Econ Review* 2005; **95**(5): 1525-1547.
16. Sarcany A. Adoption and Learning Across Hospitals: The Case of a Revenue-Generating Practice. *J Health Econ* 2018; **60**: 142-164.
17. Geruso M, Layton T. 2015. *Upcoding: Evidence from Medicare on Squishy Risk Adjustment*. NBER Working Paper w21222. DOI: 10.3386/w21222.
18. Geruso M, McGuire T. 2016. Tradeoffs in the Design of Health Plan Payment Systems: Fit, Power, and Balance. *J Health Econ* 2016; **47**: 1-19.
19. Bowblis JR and Brunt CS. Medicare Skilled Nursing Facility Reimbursement and Upcoding. *Health Econ* 2014; **23**(7): 821-840.
20. McGuire TC. 2000. Physician Agency. In: *The Handbook of Health Economics*, Cuyler AJ and Newhouse JP (Eds). Amsterdam: Elsevier.
21. Silverman E and Skinner J. Medicare Upcoding and Hospital Ownership. *J Health Econ* 2004; **23**(2): 369-89.
22. McFadden, Daniel. 1973. Conditional Logit Analysis of Qualitative Choice Behavior. In: *Frontiers in Econometrics*, P. Zarembka (ed.), 105-142, Academic Press: New York.
23. Soderberg P, Tungstrom S, Armelius BA. Reliability of global assessment of functioning ratings made by clinical psychiatric staff. *Psychiatr Serv* 2005; **56**(4): 434-438.
24. Mugnaini D and Fabiano GA. The evaluation of impairment in the attention deficit/hyperactivity disorder: the contribution of the impairment rating scale. *Minerva Pediatr* 2006; **58**(2): 159-166.
25. Berta P, Callea G, Martini G, and Vittadinia, G. The Effects of Upcoding, Cream Skimming and Readmissions on the Italian Hospitals Efficiency: A Population-Based Investigation. *Economic Modelling* 2010; **27**(4): 812-821.
26. Steinbusch PJ1, Oostenbrink JB, Zuurbier JJ, Schaepkens FJ. The Risk of Upcoding in Casemix Systems: A Comparative Study. *Health Policy* 2007; **81**(2-3): 289-299.
27. Frakt, AB and Mayes R. Beyond Capitation: How New Payment Experiments Seek to Find the ‘Sweet Spot’ in Amount of Risk Providers and Payers Bear. *Health Aff* 2012; **31**(9): 1951-1958.

