

The Mental Health Parity and Addiction Equity Act Evaluation Study:
Impact on Specialty Behavioral Health Utilization and Expenditures
Among “Carve-Out” Enrollees

Susan L. Ettner, Ph.D.^{a,b}
Jessica Harwood, M.S.^a
Amber Thalmayer, Ph.D.^c
Michael K. Ong, M.D., Ph.D.^{a, d}
Haiyong Xu, Ph.D.^a
Michael J. Bresolin, Ph.D.^c
Kenneth B. Wells, M.D., M.P.H.^e
Chi-Hong Tseng, Ph.D.^a
Francisca Azocar, Ph.D.^c

a Division of General Internal Medicine and Health Services Research, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles

b Department of Health Policy and Management, Fielding School of Public Health, University of California, Los Angeles

c Optum®, United Health Group

d VA Greater Los Angeles Healthcare System

e Department of Psychiatry, Neuropsychiatric Institute, University of California Los Angeles

Acknowledgements: We gratefully acknowledge support for this study from the National Institute on Drug Abuse (1R01DA032619-01), data from Optum®, United Health Group, and helpful comments from seminar participants at the Virginia Commonwealth University, the University of Minnesota-Minneapolis, the University of Toronto, and the University of California, Los Angeles. We would like to thank Rosalie Pacula, Ph.D. and Susan Ridgely, J.D. for providing information on state parity mandates for treatment of mental illness and substance use disorders through a RAND subcontract. We would also like to thank Brent Bolstrom and Nghi Ly from Optum®, United Health Group for the extraction, preparation and delivery of data to UCLA. The academic team members analyzed all data independently and retained sole authority over all publication-related decisions throughout the course of the study.

Disclaimer: The views and opinions expressed here are those of the investigators and do not necessarily reflect those of the National Institutes of Health, Optum®, United Health Group or UCLA.

Conflicts of interest: Drs. Azocar and Bresolin are employees of Optum®, United Health Group. As such, they receive salary and stock options as part of their compensation. Dr. Thalmayer was a contractor of Optum®, United Health Group at the time the original manuscript was written. As such, she received salary as part of her compensation. All other authors declare they have no conflicts of interest.

JEL Code: I1

Keywords: Behavioral health, parity, utilization, expenditures, insurance benefits

Author contact information:

Lead/corresponding author: Susan L. Ettner, Ph.D., Division of General Internal Medicine and Health Services Research, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles. Address: 911 Broxton Plaza, Los Angeles, CA 90024. Email: settner@mednet.ucla.edu.

Co-author: Jessica Harwood, M.S., Division of General Internal Medicine and Health Services Research, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles. Address: 911 Broxton Plaza, Los Angeles, CA 90024. Email: JHarwood@mednet.ucla.edu.

Co-author: Amber Thalmayer, Ph.D. Address: 27786 Royal Ave., Eugene, OR 97402. Email: ambergayle@gmail.com.

Co-author: Michael K. Ong, M.D., Ph.D., Division of General Internal Medicine and Health Services Research, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles. Address: 10940 Wilshire Blvd., Ste. 700, Los Angeles, CA 90024. Email: mong@mednet.ucla.edu.

Co-author: Haiyong Xu, Ph.D., Division of General Internal Medicine and Health Services Research, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles. Address: 911 Broxton Plaza, Los Angeles, CA 90024. Email: haiyong.xu@gmail.com.

Co-author: Michael J. Bresolin, Ph.D., Optum®, United Health Group. Address: PO Box 9472 Minneapolis, MN 55440-9472, Email: Michael.Bresolin@optum.com.

Co-author: Kenneth B. Wells, M.D., M.P.H., Department of Psychiatry, Neuropsychiatric Institute, University of California Los Angeles, UCLA Neuropsychiatric Institute. Address: 10920 Wilshire Blvd, Suite 300, Los Angeles, CA 90024-3724. Email: KWells@mednet.ucla.edu.

Co-author: Chi-Hong Tseng, Ph.D., Division of General Internal Medicine and Health Services Research, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles. Address: 911 Broxton Plaza, Los Angeles, CA 90024. Email: CTseng@mednet.ucla.edu.

Senior author: Francisca Azocar, Ph.D., Optum®, United Health Group. Address: 425 Market Street, 14th Floor, San Francisco, CA 94105. Email: Francisca.Azocar@optumhealth.com.

Abstract

Interrupted time series with and without controls was used to evaluate whether the federal Mental Health Parity and Addiction Equity Act (MHPAEA) and its Interim Final Rule increased the probability of specialty behavioral health treatment and levels of utilization and expenditures among patients receiving treatment. Linked insurance claims, eligibility, plan and employer data from 2008-13 were used to estimate segmented regression analyses, allowing for level and slope changes during the transition (2010) and post-MHPAEA (2011-13) periods. The sample included 1,812,541 individuals ages 27-64 (49,968,367 person-months) in 10,010 Optum “carve-out” plans. Two-part regression models with Generalized Estimating Equations were used to estimate expenditures by payer and outpatient, intermediate and inpatient service use. We found little evidence that MHPAEA increased utilization significantly, but somewhat more robust evidence that costs shifted from patients to plans. Thus the primary impact of MHPAEA among carve-out enrollees may have been a reduction in patient financial burden.

Introduction

Historically, insurance coverage in the United States was less generous for behavioral health (BH) disorders than for medical conditions. Starting in the 1970s, states began to address this inequity by passing parity laws, i.e., laws requiring equality of insurance coverage for mental health (MH) and substance use disorders (SUD) compared with medical care. However, state laws varied substantially in their definition of parity in terms of specific benefits (e.g. deductibles, co-insurance, day limits), definitions of mental illness, inclusion of SUD, inclusion of individual and group plans, and exemptions for cost increases and small employers. Moreover, the Employee Retirement Income Security Act of 1974 (ERISA) exempts self-insured firms from state insurance mandates and the proportion of covered workers in such plans has been steadily increasing over time; currently 61% of commercially insured patients are in self-funded plans (Henry, 2008) and hence their benefits are not subject to state parity laws.

Due to the limitations of state benefit mandates, advocates lobbied for federal parity legislation, leading to the passage of the Mental Health Parity Act of 1996 (MHPA). MHPA was effective as of January 1998 and mandated that if insurers covered mental health benefits, they must provide the same annual and lifetime spending limits as they do for medical benefits. However, MHPA did not require parity for SUD and some employers were exempt. Employers also had the option to drop MH coverage altogether. A survey of employers subject to MHPA found that the percent reporting parity in dollar limits grew from 55% in 1996 to 86% in 1999 (Allen, 2000). However, MHPA did not require parity with respect to other cost-sharing features (e.g., copayments) or treatment limitations (e.g., numbers of visits) and most of the newly compliant employers changed plans to be more restrictive in these other ways (Allen, 2000). Thus MHPA improved MH coverage in terms of annual and lifetime financial limits, but may have had unintended consequences in terms of resulting in other forms of limits on benefits, leading to no net gains for consumers.

The limited nature of the MHPA provisions led to a push for stronger state and federal parity laws. By 1998, 14 states had passed stronger parity legislation than MHPA, and in 2001, the Federal Employees Health Benefits Program (FEHBP) was required to offer its 8.7M beneficiaries equal BH coverage in annual and lifetime dollar limits, deductibles, copayments and limits on the number of outpatient visits and inpatient days. Although the FEHBP parity provisions were much more comprehensive than those of MHPA, on average costs increased by only 0.10% over five years (Goldman et al., 2006). The lack of meaningful impact was thought to be the result of increases in direct utilization management in response to the law. These findings underscored the important role played by care management in determining behavioral health care utilization, an issue that was highlighted during the development of the next major piece of federal parity legislation.

On October 3, 2008, the 110th Congress passed the Mental Health Parity and Addiction Equity Act (MHPAEA), which was effective for plans renewing on or after January 1, 2010. MHPAEA prohibited employer groups offering BH coverage (including both MH and SUD) from applying financial requirements (e.g., deductibles and copayments) or treatment limits (e.g., number of visits or days of coverage) that are more restrictive than the “predominant” requirements/limits applying to “substantially all” medical/surgical benefits. It also prohibited separate accumulation of deductibles and out-of-pocket maximums. However, MHPAEA did not specify covered diagnoses and it exempted plans with ≤ 50 employees, disability plans, long-term care plans, government-sponsored plans opting out, hospital or other fixed indemnity insurance, and plans showing that their costs increased by a certain amount as a result of compliance. Importantly, until regulations were issued, there was no formal enforcement of these provisions; employers (and in the case of fully insured plans, insurers) were merely expected to make a “good faith effort” at interpreting and complying with the law.

The MHPAEA Interim Final Rule (IFR) was issued on February 2, 2010 and took effect for most plans on the first day of their plan year beginning or renewing on or after July 1, 2010

(e.g., plans renewing on a calendar year cycle had to comply by January 1, 2011). In addition to “signaling” that formal compliance would now be required and enforced by states, the IFR made a critical extension to the original MHPAEA provisions by clarifying that parity also applied to non-quantitative treatment limits (NQTLs), e.g., pre-authorization, medical necessity review, provider reimbursement rates, etc. The MHPAEA Final Rule (FR) was issued in November 2013, updating and replacing the IFR as each plan renewed on or after July 1, 2014 (for most plans, which renew on the calendar year, the FR became effective on January 1, 2015). The FR retained the IFR’s NQTL provisions and further clarified interactions of MHPAEA with the Affordable Care Act.

Together with its interim and final rules, MHPAEA represented a landmark piece of legislation, as its provisions went well beyond prior federal and state parity laws. In addition to being nationally applicable (with no exemptions for self-insured plans) and explicitly including SUD, the provisions of the law required parity not just in financial requirements and quantitative treatment limits (QTLs), but also in NQTLs. A major reason why previous parity mandates did not lead to higher costs was the cost savings resulting from increased use of managed care techniques following the implementation of parity, such as prior authorization requirements or contracting arrangements with specialty MBHOs with expertise in managing behavioral health utilization and benefits (Barry et al., 2003; Barry and Ridgely, 2008; Frank and McGuire, 1998; Sturm et al., 1998). The IFR, which required parity with regard to NQTLs, reduced insurers’ ability to employ “supply-side” techniques to contain costs. Thus despite evidence that earlier parity legislation had, at most, modest effects on access and utilization of behavioral healthcare (and in turn minimal impact on clinical outcomes and medical care) (California Health Benefits Review Program, 2010), the unique features of MHPAEA and its regulations suggested that it could have had far more dramatic effects than the prior laws.

Our study is an evaluation of the impact of the implementation of MHPAEA (statutory

regulations) and the IFR on managed behavioral health “carve-out” enrollees, conducted in collaboration with researchers from the behavioral health division of Optum®, United Health Group, which is one of the largest managed behavioral health organizations (MBHO) in the country. Optum currently contracts with 2500 facilities and 130,000 providers to serve approximately 2500 customers (including UnitedHealthcare and other commercial medical insurance plans in addition to employer groups), with 60.9 million members distributed across all U.S. states and territories. The analyses presented here use administrative databases from Optum to test the hypotheses that implementation of MHPAEA and its IFR were associated with an increase in penetration rates (i.e, the probability of any use of behavioral benefits) as well as increases in service use, plan expenditures and total expenditures among those receiving treatment. We also examine the association of MHPAEA with out-of-pocket costs, although the direction of this relationship is theoretically indeterminate *a priori* because increases in utilization could offset reductions in the rate of patient cost-sharing.

Literature Review

A comprehensive review of the earlier parity literature (California Health Benefits Review Program, 2010) concluded that among individuals who already had some BH coverage and whose utilization was being managed through a range of techniques, parity was associated with a modest increase in utilization of MH/SA services among certain subpopulations, including persons with SUD. However, consumer out-of-pocket expenditures declined. Effects on outpatient BH visits depended on insurance type, showing an increase in response to parity among HMO patients but a decline among fee-for-service patients (due to increased contracting with managed behavioral healthcare “carve-outs” following parity). Conclusions regarding the effect of parity on inpatient utilization were mixed and depended in part on diagnosis. Some evidence suggested that perceived access to care improved, as did receipt of guideline-concordant care. However, parity was not significantly associated with suicide rates in the only

study of clinical outcomes associated with parity (Klick and Markowitz, 2006).

Two later studies by McConnell and colleagues (McConnell, 2013; McConnell et al., 2012) evaluated the parity law in Oregon, the only state whose law (nominally) included NQTLs. Their evaluation used administrative data from 4 PPOs for commercially insured individuals subject to the parity law, matched with individuals covered by exempt (self-insured) plans from the MarketScan database. MH/SA expenditures did not increase overall in response to the Oregon law, but they did increase modestly among the subsample of individuals with serious mental illness (i.e., among those who were already high utilizers).

Changes in benefit design implemented by insurers and employers in response to MHPAEA have been documented in several publications (Ettner, 2016; Goplerud, 2013; Hodgkin et al., 2003; United States Government Accountability Office, 2015). These analyses provided some evidence that among pooled samples of carve-in and carve-out plans, certain insurance benefit design features did become more generous following MHPAEA implementation (in particular, quantitative treatment limits were removed), suggesting that one might expect a demand response. However, to date little has been published about the effect of MHPAEA on members' behavioral health utilization and expenditures.

The Health Care Cost Institute analyzed inpatient claims from individuals enrolled in employer-sponsored "carve-in" plans, finding increases in MH/SA admissions and inpatient spending between 2007 and 2011 (Herrera et al., 2013). However, the authors acknowledge that the role MHPAEA played in these increases is unclear, as MH/SA inpatient admissions were already increasing among carve-in enrollees just before MHPAEA was implemented. To our knowledge, only two peer-reviewed evaluations of MHPAEA have been published to date and both focused on SUD treatment. Busch et al. (Busch et al., 2014) used administrative data from Aetna "carve-in" plans in states with pre-existing parity laws to compare changes between 2009 and 2010 in the substance abuse treatment patterns of individuals enrolled in fully insured plans (already subject to parity through the state laws) vs. self-insured plans (exempt from state

parity laws). They found no changes in identification, treatment initiation, or treatment engagement for SUD, although spending on SUD treatment did go up by about \$10 per enrollee per year. The authors conclude that MHPAEA did not lead to substantial increases in health plan spending on SUD treatment, but note that it would be critical to study longer-term effects, due to the importance of the NQTL provisions that did not take effect until 2011 for most plans.

McGinty et al. (McGinty et al., 2015) used 2007-2012 insurance claims data from members covered by large self-insured employers from the Truven Health MarketScan Commercial Claims and Encounters Database. Looking within a sample of people who used substance use disorder services, they used an interrupted time series design to determine that MHPAEA was associated with an increased probability of using out-of-network SUD services, with an increased average total spending on out-of-network SUD services, and with an increased average number of out-of-network outpatient SUD visits.

Our study complements the previous literature in several ways. We examine all forms of treatment (inpatient, intermediate and outpatient) and we examine all BH services instead of focusing on SUD treatment only (Busch et al., 2014; McGinty et al., 2015). We also use four years of post-MHPAEA data to account for long-term effects and account for the impact of the IFR provisions (e.g., parity in NQTLs) in addition to the original MHPAEA provisions. Lastly, but perhaps most importantly, we analyze data for “carve-out” patients. Earlier studies either focused exclusively on carve-in patients (Busch et al., 2014) or had an unknown mix of carve-in and carve-out enrollees (McGinty et al., 2015). For two reasons, MHPAEA may have differentially affected carve-out and carve-in plans. First, care management prior to parity tended to be quite different for carve-in versus carve-out plans. Second, the administrative burden associated with parity compliance is quite different for carve-out and carve-in models. To comply with parity, carve-out plans had to first identify all of the medical vendors with whom their customers contracted and then obtain detailed benefit design information from each of

them (a more difficult task when the medical vendor was not affiliated with the BH vendor). They then had to either match the most generous medical benefit across the board or else tailor benefits to those for each plan offered by each medical vendor. In turn, this led to a proliferation of plans and heterogeneity in benefit design in the post-parity period among employer groups choosing to retain the carve-out model for their behavioral health coverage (as we see in our data). These differences suggest that any MHPAEA effects estimated among a subsample of carve-in plans may not generalize to carve-out plans.

Methods

Overview of Study Design We use an individual-level interrupted time series (ITS) study design, with a longitudinal sample of carve-out enrollees enrolled any time from 2008-13. The unit of observation for all models is the person-month, so each individual contributes up to 72 monthly observations (over the six years) to the pooled sample. We use segmented regression analysis with this pooled sample to estimate the change in an outcome's time trend as a function of indicators and spline variables for the post-MHPAEA period (2011-2013) and the transition period (2010) relative to the pre-MHPAEA period (2008-2009), controlling for other explanatory variables. The transition period was chosen to correspond to the early MHPAEA implementation, before the Interim Final Rule became effective. During the transition period, plans were required only to make a "good-faith" effort to comply and there was no rule requiring them to be at parity with regard to NQTLs. The post period (after the IFR took effect) was defined by when plans were legally required to comply with MHPAEA and states were expected to enforce its provisions, including the new IFR provisions regarding NQTL parity.

We use ITS because it is one of the strongest quasi-experimental study designs (Wagner et al., 2002), even in the absence of a comparison group (Fretheim et al., 2013; Lagarde, 2012), and has frequently been used to evaluate important policy changes even when no comparison group is available (Aliu et al., 2014; Du et al., 2012; Hacker et al., 2015;

Kozhimannil et al., 2011). Our main analyses focus on the “treated” group of Optum enrollees in plans newly subject to parity provisions (those in self-insured, large-group plans, also known as “administrative services only” plans). However, sensitivity to possible confounding time trends is explored in a difference-in-differences analysis with a comparison group.

Our models are “intent-to-treat” in the sense that we are trying to estimate the overall impact of parity on expenditures and utilization, rather than how changes in the plan’s benefit design resulting from parity affected these endpoints. For this reason, we use a reduced-form model in which the key predictors have to do with whether MHPAEA and the IFR were in effect at the time of the observation. Although the mechanisms through which MHPAEA and the IFR are hypothesized to affect the endpoints are changes to financial requirements, QTLs and NQTLs, an analysis of the mediating pathways is outside the scope of the current study.

Sources of Data Our study is based on four linked administrative databases from Optum for 2008-2013: (i) member eligibility files, (ii) specialty behavioral health claims, routinely collected and archived for all Optum behavioral health beneficiaries; (iii) the “Book of Business” file; and (iv) provider supply data. Member eligibility data include age, gender, relationship to subscriber, state of residence, and eligibility information. The claims provide information on the patient and provider; setting (inpatient vs. outpatient); date(s) of service; diagnosis and procedure codes; amounts billed and reimbursed; deductibles; and copayment/coinsurance amounts. Optum uses fee-for-service reimbursement, so all records have payment amounts. The “Book of Business” file has information on employer group and plan characteristics, such as funding arrangement (self-funded vs. fully insured), employer group size, type of coverage (behavioral health, EAP, work-life, etc.) and type of plan (HMO, PPO, etc.). Information on whether the employer group uses a carve-out or carve-in model (or both) came from other Optum business records. Information on provider supply included the number of Optum providers by year, state and license type; these measures were then divided by the number of

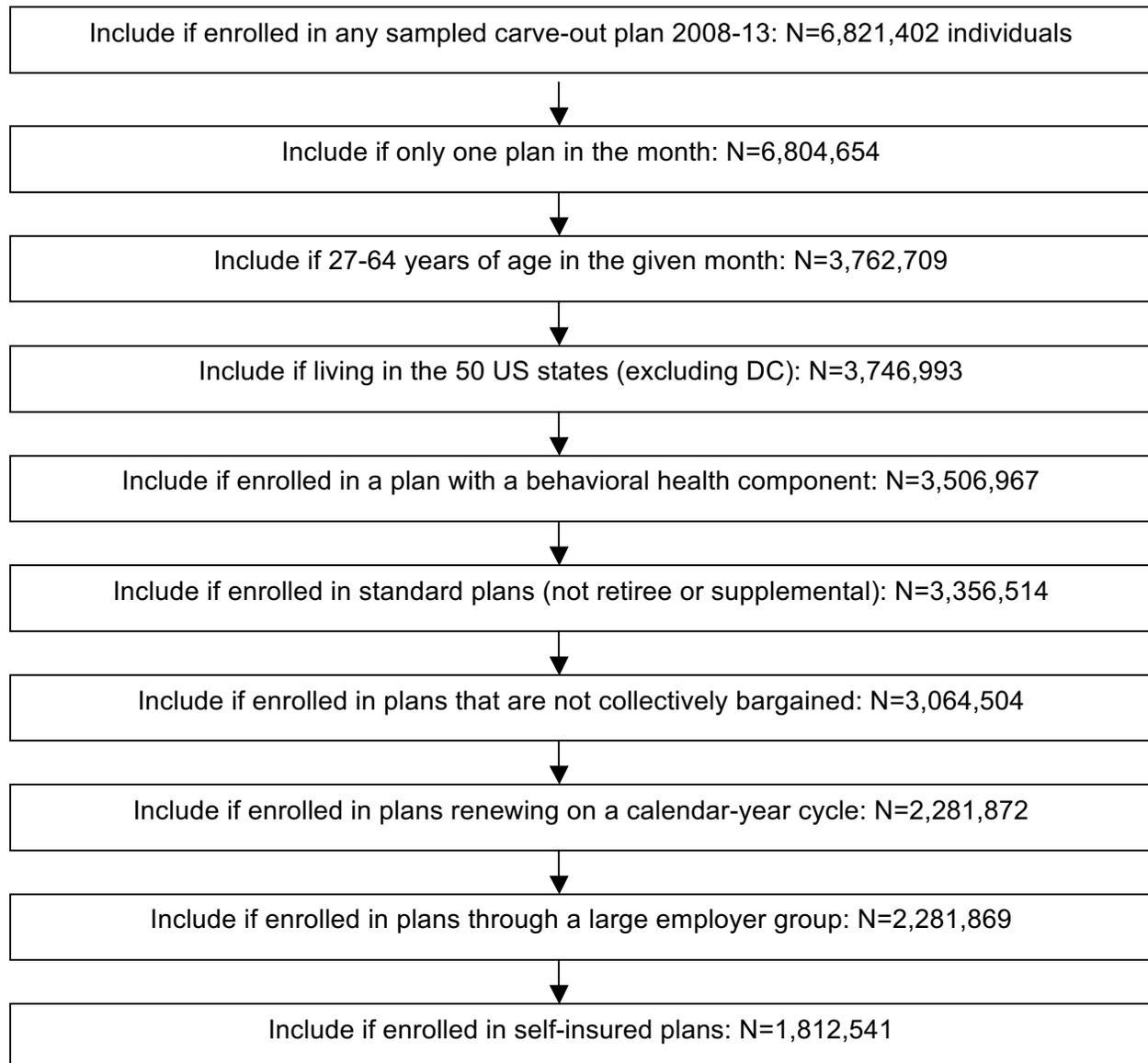
members in each state (in 1,000s) to account for population differences.

Study Cohort Our initial sampling strategy included all employer groups that had Optum carve-out plans at any time between 2008 and 2012. We ultimately obtained 2008-13 administrative data for these employer groups to construct person-month observations, imposing the following inclusion and exclusion criteria (see Figure 1): Include all individuals who were (i) enrolled in an Optum carve-out plan at any time between 2008-2013; (ii) enrolled in no more than one plan in the same month; (iii) aged 27-64 years of age; (iv) living in the 50 U.S. states (excluding DC); (v) were enrolled in plans that included behavioral health coverage; and (vi) were enrolled in plans that were subject to MHPAEA as of January 1, 2010 (thereby excluding retiree and supplemental plans, plans that do not renew on a calendar year cycle, and plans that were collectively bargained or associated with small employers). The main sample for the analyses further limits to individuals enrolled in self-insured plans (accounting for 79% of the sample, or N= 1,812,541 unique individuals in 10,010 plans offered by 63 employer groups, corresponding to 49,968,367 person-month observations).

The two criteria that excluded the most enrollees were limiting to adults aged 27-64 and limiting to plans renewing on a calendar year cycle. Older adults were excluded to ensure that the subject's primary insurance coverage was subject to parity and young adults were excluded to avoid possible overlap with the effects of early Affordable Care Act (ACA) provisions regarding dependent coverage. The limitation to calendar-year plans was imposed because the timing of compliance requirements depended on the plan's "anniversary date." For example, fiscal-year plans renewing July 1st of each year were required to comply with both the original MHPAEA provisions and the IFR on the same date – July 1, 2010 – whereas calendar-year plans had to be compliant with the IFR on January 1, 2011, and may have chosen to be compliant with the original MHPAEA provisions either by January 1, 2010 (when MHPAEA became effective and plans were expected to make a "good faith effort" to comply with its

provisions) or by January 1, 2011 (when the IFR took effect, providing guidance for plans in how to comply and legal compliance was required).

Figure 1. Sample Size Flowchart



Note: Final sample size is N=1,812,541 people, corresponding to 63 employers, 10,010 plans, and 49,968,367 person-months.

Measures For each sampled individual, study outcomes are aggregated across claims incurred within each calendar month and include the following: expenditures, broken down by plan (Optum + “Coordination of Benefit” payments by other insurers), patient out-of-

pocket (including e.g., coinsurance, copayments, deductibles) and total (plan + patient); number of outpatient visits for assessment/diagnostic evaluation, individual psychotherapy, family psychotherapy, and medication management; and number of days of structured (including intensive) outpatient care, day treatment, residential care, and acute inpatient care. Claims that spanned multiple months were prorated.

Expenditure measures were inflation-adjusted to 2013 dollars using the “other medical professionals” component of the Consumer Price Index (CPI) for outpatient expenditures and the inpatient CPI for inpatient expenditures. Next, to adjust expenditures for heterogeneity due to price variation across states, we calculated mean expenditures per service unit across claims within each service type (using individual psychotherapy – accounting for 40% of claims -- for outpatient and acute inpatient services for inpatient), first aggregating across all claims nationally and then aggregating across claims within each state. The state-specific adjustment factor was calculated by dividing the national mean by each state mean. We then multiply expenditures for patients by their state factor to adjust to the national mean. Finally, adjusted outpatient and inpatient dollars were then summed to the total.

Key covariates were a continuous variable for time in months (representing the underlying time trend), spline variables for the transition and post periods and indicators for the transition and post-parity periods (see above definitions). The spline variable for the transition period measured the change in the outcome’s slope for the transition period relative to the outcome’s pre-parity slope. The indicator variable for the transition period measured the discontinuity as of January 2010; in other words, the immediate change in the level of the outcome in the transition period, relative to the level that would be expected based on the pre-parity time trend. The post-parity period indicator and spline variables were similarly defined to measure changes in level and slope that occurred in the post-parity period relative to the pre-parity period. All regressions also controlled for employer group size category; plan type (“more

managed” types such as HMO vs. “less managed” types such as PPO); enrollee sex; enrollee age group; whether the enrollee is the primary insured person (PIP) vs. dependent; provider supply rates in each state and year (by license type); and fixed state effects. We also include fixed effects for each calendar month (e.g., January, February, etc.) to adjust for seasonality.

Statistical Analyses Descriptive statistics were first calculated for all variables used in the analyses. We then estimated the segmented regression analyses. Due to the large sample sizes, linear regression was used to estimate the overall associations of parity period with the dependent variables. To identify whether any changes in unconditional utilization were being driven by changes in penetration rates vs. changes in the level of service use among the treated population, we also estimated logistic regressions of the probability of a positive outcome and gamma regressions of the level of the outcome, based on the conditional subsample of observations with a positive outcome. Based on these regressions, we report the mean risk differences (the predicted probability evaluated using a given configuration of covariate values minus the predicted probability evaluated using a different configuration of covariate values) and the mean conditional predictive margins (differences in conditional expectations). P-values were calculated using first-order Taylor series expansions when computationally feasible and otherwise regression p-values are reported.

Within-person correlation of the residuals may occur in our models because each individual may contribute more than one person-month observation to the pooled sample. Clustering within plans and employers may exist as well. We consider these clustering effects to be nuisance parameters rather than hierarchical variation of particular interest, and cluster at the highest level (Cameron and Miller, 2015). Thus, we use Generalized Estimating Equations with independent covariance structure and robust variance estimation to adjust for clustering at the employer level (Liang and Zeger, 1993). All hypothesis tests are two-tailed and use a cutoff of .05 for Type 1 error. Due to the large sample sizes and multiple outcomes, we look for

broader patterns of significant findings when interpreting our results.

In addition to showing the detailed estimates, we provide a high-level summary of the changes in our expenditure and utilization outcomes associated with MHPAEA, incorporating the effects of both level and slope changes. For each outcome, the penetration rate and conditional and unconditional means are predicted as of the midpoint of our post-parity study period (July 2012) under two scenarios: (1) assuming parity never happened, and (2) assuming parity is in effect. This calculation helps to illustrate the overall magnitude of the changes for one sample month.

Sensitivity Analyses We conducted numerous sensitivity analyses, including a comparison of the results from alternative regression specifications; a comparison of unconditional predictive margins based on recombining two-part model estimates instead of using (one-part) linear regressions; estimation of the models using a continuously enrolled subsample; estimation of the models using a subsample of patients with schizophrenia and/or bipolar disorder; and estimation of difference-in-differences models using a comparison group of individuals enrolled in Optum plans less likely to have been affected by MHPAEA.

Alternative regression specifications: We re-estimated the models three ways: first, excluding plan type (in case plan type itself was affected by parity, hence serving as a mediator for its effects); second, excluding the provider supply measures for the same reason; and third, controlling for a full set of indicators for behavioral health diagnoses in the conditional regressions. (We had excluded the diagnostic indicators from the main models due to endogeneity, since diagnoses can only be incurred if services are used and more service use leads to more claims diagnoses.) All three of the alternative regression specifications yielded estimates that were only trivially different from the main specification, so we do not report those estimates here.

Use of two-part models to calculate unconditional predictive margins: Associations of

parity period with utilization and expenditures among the full sample were virtually identical when calculated by recombining two-part model estimates vs. using one-part linear regressions, so for computational ease, the latter was used.

Continuously enrolled subsample: In Appendix Table A1, we present results from a sensitivity analysis using the sample of individuals continuously enrolled for all months 2008-2013. Imposing a continuous enrollment criterion involves a tradeoff between internal and external validity. Looking at changes within the same individuals over time has the advantage of holding constant unmeasured time-invariant patient characteristics that might otherwise confound the analysis if patient populations change over time, yet findings based on the full sample (including those who were discontinuously enrolled) are more likely to generalize. We compared population characteristics at the person-month level for the full sample used in the main analyses vs. the continuously enrolled subsample used in the sensitivity analyses (5% of unique enrollees). With a couple of exceptions, the two populations have similar demographics. In the pre period, continuously enrolled individuals are more likely to fall in the middle age ranges of 35-54 compared to individuals in the main analysis (72% v. 61%). The continuously enrolled subsample also does not see as dramatic a decline in the percentage of people enrolled in more heavily managed plans as the main analysis sample does.

Subsample with schizophrenia and/or bipolar disorder: In Appendix Table A2, we present results from a sensitivity analysis using a subsample of patients who had a diagnosis of schizophrenia or bipolar disorder at any point between 2008-2013. These conditions are the most severe, chronic, and costly among our population, and MHPAEA's effects may be stronger among these enrollees.

Difference-in-differences (DID) models: In evaluating natural experiments such as MHPAEA, the critique of before-and-after comparisons is that the impact of the policy may be confounded by secular time trends. Although interrupted time series methods are considered to

be one of the strongest study designs for analyzing observational data, it is preferable to use it in conjunction with a comparison group that is unaffected by the policy being evaluated (in this case, MHPAEA and the IFR) in order to net out any differences in secular time trends not accounted for by the ITS design. We considered three potential comparison groups for our study. The first was enrollees in large-group retiree or supplemental plans, which were exempt from parity. The second was enrollees in plans offered by small employer groups (≤ 50 employees), which were also not subject to parity compliance during our study period. The third, which is the one we ultimately use for our sensitivity analyses, was enrollees in fully insured (FI) plans from states with strong pre-existing parity laws.

None of these potential comparison groups was ideal. In addition to concerns about small sample sizes, we did not find it plausible that the secular time trends were the same either for the retiree and supplemental plans (relative to plans offering primary behavioral health coverage) or for the small employer groups (since trends in utilization and expenditures among groups with 50 or fewer employees are highly unlikely to generalize to the very large employers in our Optum databases, where groups of fewer than 5,000 employees were already very small in relative terms).

A priori, the most promising comparison was with enrollees in FI plans from states with strong parity laws, so we used this group for our sensitivity analyses. Nonetheless, it is difficult to make the argument that these plans were entirely unaffected by MHPAEA. Heterogeneity in the populations and benefit design features that were included in state parity laws and the details of how they were included makes it difficult to draw a clear distinction between states that had “strong” parity laws and states that did not. Furthermore, even states that appeared to have strong parity laws may not have enforced them. Perhaps most importantly, the NQTL provisions that were such a critical part of MHPAEA’s regulations were virtually unique to federal parity; for example, even Oregon, which had included a similar provision in its own parity

law that took effect not long before MHPAEA, anecdotally never enforced this requirement prior to implementation of the federal parity law. If the comparison group was itself affected by MHPAEA and the IFR, then DID estimates, which rely on an assumption that any changes over time observed among a comparison group reflect a true secular trend rather than intervention effects, would be subject to conservative biases. As a result of these concerns, while we provide formal DID estimates as a sensitivity analysis, these findings should be viewed with caution. The comparison group is enrollees in fully-insured plans in states that had “strong parity” by 2009 (AL, CT, GA, IN, KY, ME, MN, MO, NM, OR, WA, and WI). Enrollees in fully-insured plans from states with no or weak pre-existing parity laws were excluded from analysis.

Results

Descriptive statistics Table 1a describes the person-month sample at two points in time, one pre-parity month (January 2009) and one post-parity month (July 2012). Due to the decline in the use of the carve-out model after parity implementation, more people were enrolled in the pre-parity time period than in the post-parity time period. (The greater average number of plans per employer group in the post-parity period is due to the need to create separate plans corresponding to each medical vendor’s benefits; prior to parity, an employer would typically offer the same BH benefits to all patients in the carve-out plan, but post-parity, the benefits had to be tailored to the medical coverage to achieve compliance.) Enrollees in the post-parity time period tended to be more concentrated in the oldest age group (55- to 64-year-olds) and were less likely to be enrolled in a “more managed” plan than those in the pre-parity time period, but for the most part, differences across the two time periods were modest. Table 1b provides unadjusted descriptive data on service use and expenditures by parity period in order to put our regression estimates into context. Our outcomes (which are at the monthly level) are rare events; for example, the percent of enrollees with any behavioral health specialty expenditures in a given month is about 3%, and the percent of enrollees with any residential care is 0.01%.

Table 1a. Demographics for the pre-parity and post-parity periods.

	Pre (January 2009)		Post (July 2012)	
	n	%	n	%
Number of employers	49		28	
Number of plans*	1,146		2,881	
Number of people	909,393		494,069	
Age group				
27-34 years	182,846	20.1	82,979	16.8
35-44	273,064	30.0	138,672	28.1
45-54	284,658	31.3	145,210	29.4
55-64	168,825	18.6	127,208	25.7
Male (vs. female)	433,196	47.6	243,038	49.2
Primary insured person (vs. dependent)	598,657	65.8	314,844	63.7
Census Division				
Northeast: New England	72,945	8.0	29,290	5.9
Northeast: Middle Atlantic	70,166	7.7	56,162	11.4
Midwest: East North Central	121,204	13.3	101,906	20.6
Midwest: West North Central	76,287	8.4	19,243	3.9
South: South Atlantic	175,675	19.3	87,514	17.7
South: East South Central	25,931	2.9	28,890	5.8
South: West South Central	109,463	12.0	38,513	7.8
West: Mountain	82,385	9.1	25,955	5.3
West: Pacific	175,337	19.3	106,596	21.6
Employer group size				
>40K enrolled employees	401,592	44.2	172,001	34.8
>10K & ≤ 40K	392,538	43.2	270,016	54.7
5,000-10,000	83,943	9.2	25,651	5.2
<5,000	31,320	3.4	26,401	5.3
Plan type is more managed (e.g., HMO) vs. less managed (e.g., PPO)	384,047	42.2	90,360	18.3
<i>Among people with any service use:</i>				
Any adjustment disorder	6,684	25.6	3,824	25.8
Any post-traumatic stress disorder	899	3.4	696	4.7
Any generalized anxiety	4,185	16.0	2,945	19.9
Any obsessive-compulsive disorder	450	1.7	285	1.9
Any panic disorder	1,055	4.0	655	4.4
Any phobia	209	0.8	118	0.8
Any cognitive disorder	74	0.3	47	0.3
Any bipolar disorder	2,167	8.3	1,274	8.6
Any depressive disorder	11,582	44.3	6,519	44.0
Any personality disorder	213	0.8	129	0.9
Any psychotic disorder	217	0.8	166	1.1
Any alcohol use disorder	631	2.4	414	2.8

Any drug use disorder	347	1.3	285	1.9
Any other psychiatric disorder	1,341	5.1	816	5.5
	Mean	SD	Mean	SD
Number of providers per 1000 members, by state & year				
MD	0.9	0.8	0.9	0.7
MSW	3.3	4.6	3.6	4.3
PHD	1.2	1.2	1.2	1.0
RN	0.2	0.4	0.2	0.4
Non-Independent Licensed	0.5	3.9	0.1	1.7

Table 1b. Service use and expenditures for the pre-parity and post-parity periods.

	Pre (January 2009)		Post (July 2012)	
	n	%	n	%
Spending/service use (% with any)				
Total Expenditures	26,133	2.87	14,824	3.00
Plan Expenditures	24,677	2.71	13,878	2.81
Patient Out-Of-Pocket Expenditures	22,491	2.47	12,066	2.44
Outpatient Assessment/Diagnostic Evaluation	2,905	0.32	1,369	0.28
Outpatient Medication Management	8,527	0.94	5,102	1.03
Outpatient Individual Psychotherapy	15,486	1.70	8,838	1.79
Outpatient Family Psychotherapy	1,262	0.14	777	0.16
Structured Outpatient Care	237	0.03	210	0.04
Day Treatment Care	117	0.01	67	0.01
Residential Care	54	0.01	36	0.01
Inpatient Care	237	0.03	141	0.03
Level of spending/service use, among those with any	Mean	SD	Mean	SD
Total Expenditures (\$)	302.4	1,093.0	303.1	1,085.0
Plan Expenditures (\$)	236.0	976.0	262.1	1,056.0
Patient Out-Of-Pocket Expenditures (\$)	92.4	275.0	70.9	147.0
Outpatient Assessment/Diagnostic Evaluation (visits)	1.2	0.8	1.2	0.8
Outpatient Medication Management (visits)	1.3	1.1	1.3	0.8
Outpatient Individual Psychotherapy (visits)	2.3	1.6	2.3	1.6
Outpatient Family Psychotherapy (visits)	1.9	1.3	1.9	1.3
Structured Outpatient (days)	6.3	4.7	5.9	4.9
Day Treatment (days)	7.0	5.4	6.9	4.6
Residential (days)	9.3	7.2	9.9	7.6
Inpatient (days)	5.3	4.7	5.3	4.2

Overall changes associated with MHPAEA As shown in Table 2, which presents the estimated changes associated with MHPAEA and the IFR for all of the outcomes, changes in PMPM utilization and expenditures were mixed. Relative to the pre-parity period, assessment/diagnostic evaluation visits showed an immediate increase in level in the transition

period. However, family psychotherapy visits and patient out-of-pocket expenditures had an immediate decrease in level. Outpatient medication management and individual psychotherapy visits had opposing effects of immediate decreases in level followed by gradual increases throughout the transition period (increases in slope).

For the post-parity period, declines in slope were seen for assessment/diagnostic evaluation visits (-0.00002 [p=0.03]), medication management visits (-0.00005 [p=0.05]) and family psychotherapy visits (-0.00003 [p=0.00]). However, the post-parity level of structured outpatient days showed a level increase (0.0006 [p=0.03]). Several significant changes were seen for PMPM expenditures, including a decline in slope for total expenditures (-\$0.06 [p=.03]); an immediate increase in level for plan expenditures (\$1.78 [p=.01]); and a decline in both level and slope for patient out-of-pocket expenditures (-\$0.73 [p=.03] and -\$0.03 [p=.03] respectively).

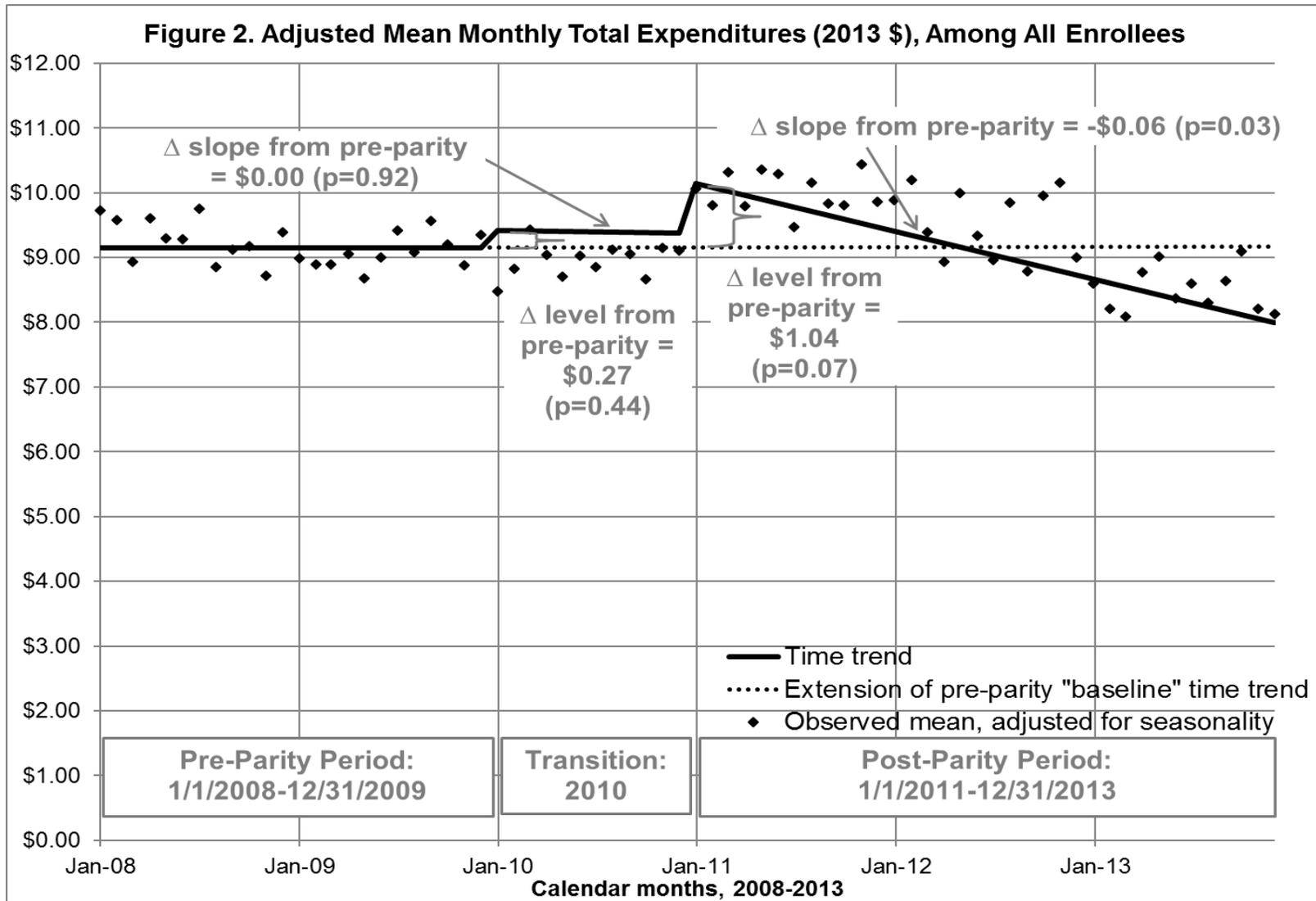
The unconditional means for total, plan, and patient per-member-per-month (PMPM) expenditures associated with MHPAEA and the IFR are displayed graphically in Figures 2-4. Each figure presents the outcome's time trend as predicted for the pre-, transition, and post-parity periods. The figures extend the pre-parity "baseline" time trend forward (shown via dotted line), to represent what would be expected in the absence of parity. The means of the raw data, adjusted only for calendar month, are also shown. The figures show that, relative to the baseline pre-parity time trend, the only significant change in the time trends of the expenditure outcomes in the transition period was a level decrease of -\$0.54 (p=.05) for patient out-of-pocket expenditures (Figure 4). However, in the post-parity period, PMPM patient expenditures (Figure 4) had an immediate decrease of \$0.73 (p=0.03) and additional decreases of -\$0.03 per month (p=0.03). Conversely, post-parity plan expenditures (Figure 3) showed an immediate increase in level of \$1.78 (p=0.01). Total expenditures (Figure 2) also had a more negative slope than would have been expected, had the pre-parity time trend continued (-\$0.06 [p=0.03]).

Table 2. Interrupted time series segmented regression analysis: estimates of the change in unconditional means of expenditures and utilization associated with federal parity.

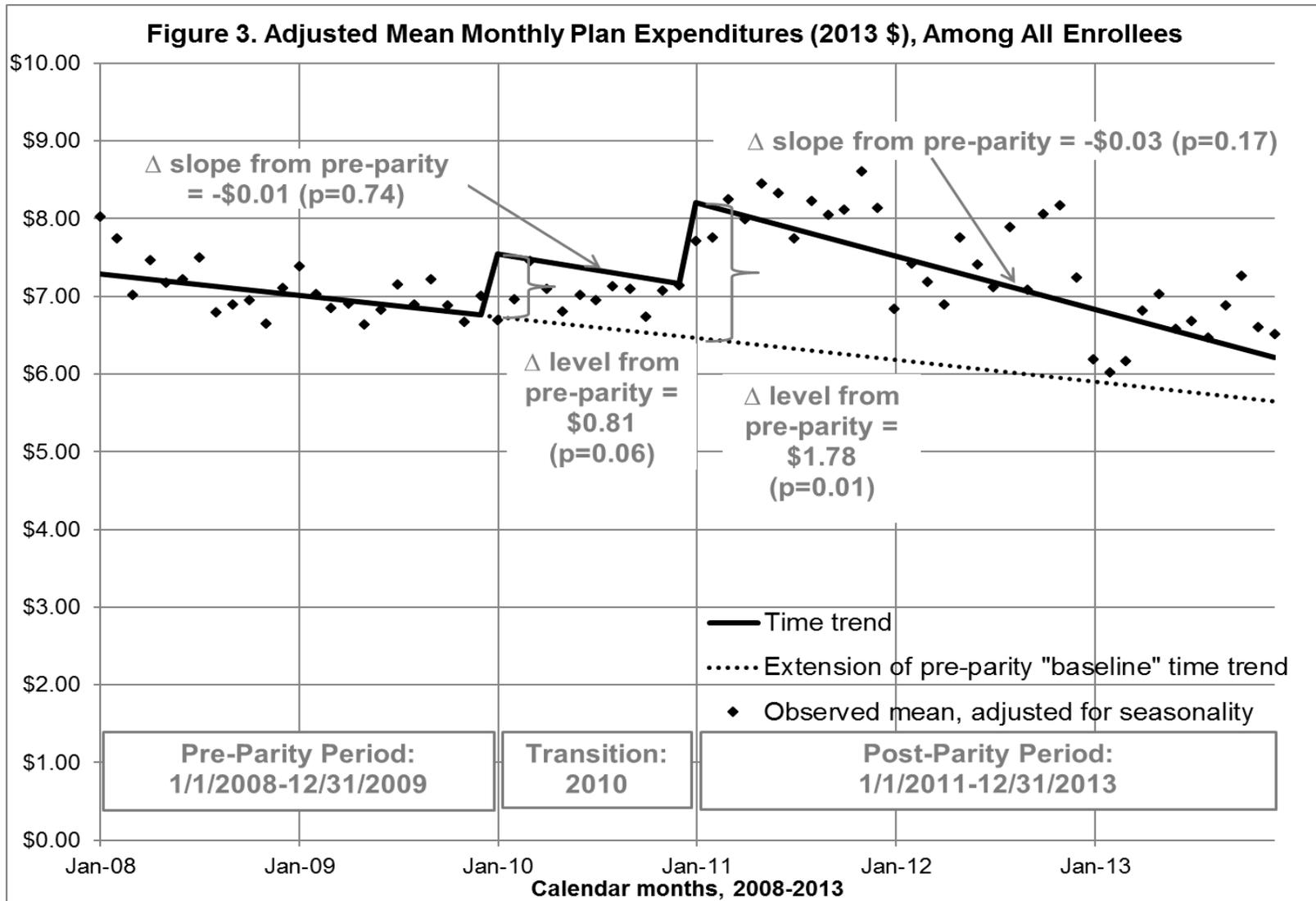
Outcome	Transition Period vs. Pre-Parity Period				Post Period vs. Pre-Parity Period			
	Δ Level ¹	P-value	Δ Slope ²	P-value	Δ Level ¹	P-value	Δ Slope ²	P-value
Expenditures								
Total	\$0.27	0.44	\$0.00	0.92	\$1.04	0.07	-\$0.06	0.03
Plan	\$0.81	0.06	-\$0.01	0.74	\$1.78	0.01	-\$0.03	0.17
Patient Out-Of-Pocket	-\$0.54	0.05	\$0.01	0.75	-\$0.73	0.03	-\$0.03	0.03
Outpatient Visits								
Assessment/Diagnostic Evaluation	0.00029	0.05	-0.00001	0.30	0.00010	0.67	-0.00002	0.03
Medication Management	-0.00081	0.04	0.00007	0.01	0.00006	0.92	-0.00005	0.05
Individual Psychotherapy	-0.00290	0.05	0.00023	0.02	-0.00174	0.61	-0.00022	0.10
Family Psychotherapy	-0.00028	0.05	0.00001	0.48	-0.00002	0.95	-0.00003	0.00
Days of Intermediate Care								
Structured Outpatient	0.00026	0.21	0.00002	0.31	0.00060	0.03	-0.00001	0.35
Day Treatment	0.00024	0.12	0.00000	0.83	0.00030	0.11	0.00000	0.96
Residential	0.00036	0.11	-0.00002	0.14	0.00026	0.16	-0.00001	0.36
Days of Inpatient Care	0.00006	0.58	0.00001	0.54	0.00035	0.08	0.00000	0.40

Notes: Estimates are from linear regression. Sample is person-months from 2008-2013 (N=49,968,367). Bold denotes significance at $p \leq .05$. Regressions controlled for a linear monthly time trend, indicators and splines for both the transition and post periods, sex, age group, whether the enrollee was the primary insured person, employer group size category, plan type, state fixed effects, provider supply measures, and seasonality.

1. Discontinuity (change in level) at the beginning of the given period, as measured by an indicator for the given period.
2. Spline (change in slope) for the given period.

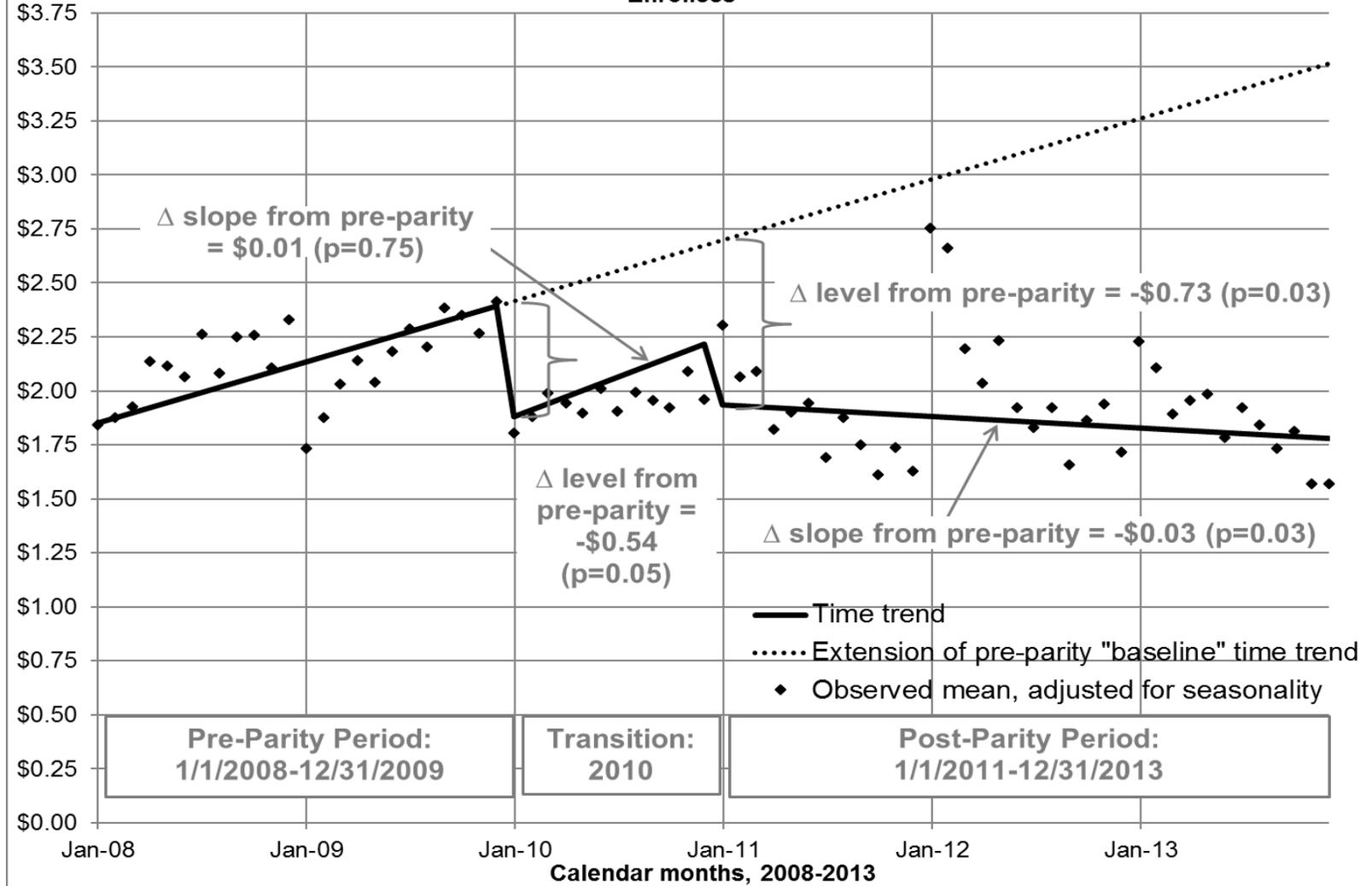


Notes: Sample is person-months from 2008-13 (N=49,968,367). Estimates from linear regression. Interrupted time series segmented regression analysis controlled for a linear monthly time trend, indicators and splines (measuring respective changes in level and slope) for both the transition and post periods, sex, age group, whether the enrollee was the primary insured person, employer group size category, plan type, state fixed effects, provider supply measures, and seasonality.



Notes: Sample is person-months from 2008-2013 (N=49,968,367). Estimates from linear regression. Interrupted time series segmented regression analysis controlled for a linear monthly time trend, indicators and splines (measuring respective changes in level and slope) for both the transition and post periods, sex, age group, whether the enrollee was the primary insured person, employer group size category, plan type, state fixed effects, provider supply measures, and seasonality.

Figure 4. Adjusted Mean Monthly Patient Out-of-Pocket Expenditures (2013 \$), Among All Enrollees



Notes: Sample is person-months from 2008-2013 (N=49,968,367). Estimates from linear regression. Interrupted time series segmented regression analysis controlled for a linear monthly time trend, indicators and splines (measuring respective changes in level and slope) for both the transition and post periods, sex, age group, whether the enrollee was the primary insured person, employer group size category, plan type, state fixed effects, provider supply measures, and seasonality.

Table 3. Interrupted time series segmented regression analysis: estimates of the change in penetration rates and conditional means of expenditures and utilization associated with federal parity, comparing the post period to the pre-parity period.

Outcome	Change in Penetration Rate ¹				Change in Conditional Mean ²			
	Δ Level ³	P-value	Δ Slope ⁴	P-value	Δ Level ³	P-value	Δ Slope ⁴	P-value
Expenditures								
Total	0.00028	0.86	-0.00014	0.01	\$31.37	0.07	-\$0.59	0.42
Plan	0.00024	0.89	-0.00010	0.16	\$58.03	0.00	-\$0.17	0.82
Patient Out-Of-Pocket	-0.00323	0.04	-0.00023	0.00	-\$21.58	0.03	-\$0.33	0.38
Outpatient Visits								
Assessment/Diagnostic Evaluation	0.00008	0.70	-0.00002	0.02	0.005	0.86	-0.001	0.42
Medication Management	0.00017	0.68	-0.00004	0.03	-0.008	0.80	0.000	0.70
Individual Psychotherapy	-0.00045	0.71	-0.00008	0.06	-0.030	0.53	-0.002	0.49
Family Psychotherapy	0.00006	0.73	-0.00002	0.00	-0.049	0.45	-0.001	0.47
Days of Intermediate Care								
Structured Outpatient	0.00012	0.03	0.00000	0.19	-0.275	0.48	0.013	0.28
Day Treatment	0.00002	0.30	0.00000	0.70	0.570	0.28	0.016	0.43
Residential	0.00002	0.29	0.00000	0.49	1.443	0.07	0.003	0.91
Days of Inpatient Care	0.00007	0.01	0.00000	0.63	0.008	0.98	-0.003	0.79

Notes: Sample is person-months from 2008-2013 (full sample N=49,968,367). Bold denotes significance at $p \leq .05$. Regressions controlled for a linear monthly time trend, indicators and splines for both the transition and post periods, sex, age group, whether the enrollee was the primary insured person, employer group size category, plan type, state fixed effects, provider supply measures, and seasonality.

1. Estimates are from logistic regression marginal effects post-estimation; p-values are from logistic regression.
2. Estimates and p-values are from gamma regression marginal effects post-estimation.
3. Discontinuity (change in level) at the beginning of the post period, as measured by an indicator variable for the post period.
4. Spline (change in slope) for the post period.

Changes in penetration rates and conditional means associated with MHPAEA Table

3 presents the results from the separate parts of the two-part model, focusing on the changes from pre- to post-parity (transition period estimates not shown). The penetration rate (the percent of people with any expenditures or use) generally decreased (in either level or slope or both) in the post-parity period for total and patient expenditures and outpatient visits and increased for intermediate and inpatient care. More specifically, the probability of having any (total) expenditures decreased in slope in the post-parity period, relative to the pre-parity period (-0.00014 [p=0.01]). The likelihood of having any out-of-pocket expenditures decreased both in level (-0.00323 [p=0.04]) and slope (-0.00023 [p=0.00]), relative to the pre-parity period. The probabilities of using any assessment/diagnostic evaluation, medication management and family psychotherapy visits decreased in slope respectively by -0.00002 (p=0.02), -0.00004 (p=0.03) and -0.00002 (p=0.00). In contrast, the probability of using structured outpatient care and inpatient care was higher in the post-parity period than would have been expected based on the pre-parity trend (level changes of 0.00012 [p=0.03] and 0.00007 [p=0.01] respectively).

There were no significant changes in mean utilization among users, either in level or slope changes. However, among the conditional samples of people with plan and patient expenditures, the post-parity period was respectively associated with a level increase of \$58.03 (p=0.00) in per-user-per-month plan expenditures and a level decrease of \$21.58 (p=0.03) in per-user-per-month patient expenditures.

Predicted penetration rates and conditional and unconditional means as of July 2012

Table 4 displays how the above-described changes to level and slope for each outcome's time trend work together to determine parity's effect at one point in the post period. In some cases, this table highlights the relatively modest effect of parity; for example, although the trend in the penetration rate of medication management significantly decreased, the predicted percentages of people using this service during the given month are similar without and with parity: 1.09% vs.

1.04%.

Table 4 also shows how the changes in penetration rate and conditional use work together to affect unconditional use. For example, the rate of patient expenditures decreases, as does the mean amount spent among people with any patient expenditures; because the two effects reinforce each other for this particular outcome, the unconditional mean of PMPM patient expenditures among the full sample in the “parity” scenario is only about half that in the “no parity” scenario (\$1.85 vs. \$3.12). The overall pattern that emerges from the unconditional predictions suggests that the main effect of parity may have been shifts in the incidence of costs from patients to health plans, rather than increases in treatment.

Table 4. Interrupted time series segmented regression analysis: predicted penetration rates, conditional means, and unconditional means as of July 2012. Each prediction is made under two assumptions: once assuming no parity, and again assuming parity is in effect.

Outcome	Predicted Penetration Rate ¹		Predicted Conditional Mean ²		Predicted Unconditional Mean ³	
	No Parity	Parity	No Parity	Parity	No Parity	Parity
Expenditures						
Total	3.26%	3.03%	\$279.03	\$298.56	\$9.16	\$9.03
Plan	2.91%	2.74%	\$210.11	\$261.74	\$6.04	\$7.18
Patient Out-Of-Pocket	3.28%	2.46%	\$105.07	\$74.19	\$3.12	\$1.85
Outpatient Visits						
Assessment/Diagnostic Evaluation	0.32%	0.29%	1.25	1.23	0.004	0.004
Medication Management	1.09%	1.04%	1.29	1.27	0.014	0.013
Individual Psychotherapy	2.00%	1.79%	2.35	2.29	0.047	0.041
Family Psychotherapy	0.19%	0.16%	1.96	1.88	0.004	0.003
Days of Intermediate Care						
Structured Outpatient	0.03%	0.04%	5.78	5.75	0.002	0.002
Day Treatment	0.01%	0.01%	6.61	7.45	0.001	0.001
Residential	0.01%	0.01%	8.18	9.60	0.001	0.001
Days of Inpatient Care						
	0.02%	0.03%	5.07	5.03	0.001	0.001

Notes: Sample is person-months from 2008-2013 (full sample N=49,968,367). Regressions controlled for a linear monthly time trend, indicator and spline for the transition and post periods, sex, age group, whether the enrollee was the primary insured person, employer group size category, plan type, state fixed effects, provider supply measures, and seasonality.

1. Predictions are from logistic regression.
2. Predictions are from gamma regression.

Sensitivity analyses using continuously enrolled subsample Appendix Table A1

displays results from the sample of individuals continuously enrolled in 2008-13. Compared with the findings from the main analyses in Table 2, fewer significant associations were found. Relative to pre-parity, the post-parity period was still significantly associated with declines in slope for total expenditures, patient out-of-pocket expenditures and family psychotherapy visits, as was the case in the main analysis. However, none of the other significant associations with the post-parity period found among the full sample remained significant among the continuously enrolled subsample (perhaps unsurprisingly, given the smaller sample sizes). In the transition period, the estimates were also mostly insignificant and had mixed signs.

Sensitivity analyses using subsample of patients with schizophrenia and/or bipolar disorder Appendix Table A2 provides the results for patients who had a diagnosis of schizophrenia (psychosis) or bipolar disorder at any point during the study period. Among this subgroup, the results that remained statistically significant (such as increases in the use of medication management and individual psychotherapy visits in the transition period and reductions in the level of patient out-of-pocket expenditures in the post-parity period) were larger in magnitude and had smaller p-values than the comparable estimates for the full study cohort. However, a number of previously significant associations lost significance altogether. Thus the results did not provide consistent evidence supporting the hypothesis that patients with more severe illnesses demonstrate a greater response to parity.

Sensitivity analyses using difference-in-differences approach Appendix Table A3 shows the difference-in-differences analyses using a comparison group composed of individuals enrolled in fully insured plans in states that had strong parity laws by 2009. Again, fewer associations were significant than in the main analyses. In the transition period, there was a significant increase in slope for plan expenditures and in level for assessment/diagnostic evaluation visits. In both the transition and post periods, there were significant increases in level

for day treatment. However, the associations of the post-parity period with the expenditure measures were no longer significant.

Discussion

Using 2008-2013 administrative data from Optum, we used an interrupted time series study design with segmented regression to examine the association of MHPAEA and its accompanying regulations with changes in utilization and expenditures among a large sample of behavioral healthcare carve-out members. Associations of parity with penetration rates and levels of service use were mixed in terms of both sign and significance; even when statistical significance was demonstrated, associations tended to be modest in magnitude. Overall, we did not identify any broad, consistent patterns of findings to suggest that parity had a notable impact on treatment. Our finding of modest to no effects on service use is consistent with conclusions of prior parity studies, including the Busch et al. (Busch et al., 2014) study of the impact of MHPAEA on SUD treatment patterns. We did find somewhat more robust evidence suggesting that there may have been a shift in costs from patients to health plans. Thus the primary impact of parity on specialty behavioral healthcare among carve-out enrollees, if any, may have been a reduction in patient financial burden.

Our study is subject to several limitations, the most significant being the inability to identify a strong control group. Secular time trends not accounted for by our interrupted time series analysis could have biased our results in either direction. However, ITS is one of the strongest observational study designs and to the extent that our comparison group is valid, the DID analyses support our conclusion that there was no widespread pattern of utilization increases resulting from parity implementation. Furthermore, although our main ITS models assume a linear secular time trend pre-parity, in sensitivity analyses, we did not find any evidence of significant nonlinear time trends in the expenditure equations.

Another possible limitation is self-selection of employers into retaining behavioral

healthcare coverage. Groups that were the most generous to begin with, and hence least affected by the passage of a strong parity law, could have been more likely to retain coverage. If our sample excluded groups that responded to parity by dropping coverage altogether, then we might be missing adverse effects on behavioral healthcare treatment patterns among patients who lost coverage. However, in data not shown here, we found no evidence that carve-out groups dropped coverage altogether. Instead, it appears that after 2009, some of Optum's carve-out employers chose to "carve in" their behavioral health benefits with their medical vendors (including carving in with Optum's sister company, UnitedHealthcare). Anecdotally, this shift from carve-out to carve-in model was the result of the administrative burden of matching benefits and coordinating combined deductibles with multiple independent medical vendors.

A related limitation is possible threats to internal validity when using the full sample, since the individuals included in the study cohort change over time. However, the characteristics of the carve-out enrollees changed little over the course of the study, and the findings from our sensitivity analysis of the continuously enrolled subsample were consistent with the main results, even though due to small sample sizes, most of the estimates were non-significant. Self-selection into the user subsample might also lead to bias. For example, increases in penetration rates would likely lead to the "marginal users" being less severely ill than existing patients, in turn leading to possible declines in conditional use and expenditures. However, no consistent parity effects were seen on penetration rates and self-selection into service use would not explain why patient expenditures go down while plan expenditures go up.

Study outcomes are based on administrative data, so we cannot look at the impact of MHPAEA on important endpoints such as quality of care, clinical and functional outcomes, or quality of life. Any effects of MHPAEA on these measures should be mediated through changes in treatment patterns, however, so starting by looking at utilization and expenditures is an appropriate first step in evaluating this legislation. Also, MHPAEA effects on utilization may be

overstated because our measures are based on claims. In the pre-parity period, patients who exceeded their visit limits and paid entirely out-of-pocket for additional visits would have their utilization underestimated unless they continued to submit claims in the hope of getting reimbursement. Post-parity, the virtual elimination of treatment limits (Hodgkin et al., 2003) (Ettner, 2016) meant that this additional utilization would show up in our claims. However, if anything, this bias should have led to an overstatement of MHPAEA effects on service use, supporting our conclusion that parity did not have a substantial impact on utilization among carve-out patients.

The ability to generalize our study findings may be limited to Optum patients, although given the enormous size, geographic coverage (all 50 states) and diversity of this patient population, we do not consider this to be a significant limitation. As Optum was the largest MBHO in the nation during our study period, we believe that Optum enrollees are representative of the MBHO population overall. In turn, MBHOs administer behavioral benefits on behalf of two-thirds of insured patients (Fox et al., 2000).

Finally, MHPAEA effects may have evolved over time, especially after 2014, when the most relevant provisions of the ACA (e.g., requirements for the Exchanges and essential health benefits, or EHBs, which include MH and SUD treatment) were rolled out. We exclude young adults (who could have been affected by changes to dependent coverage) from our study cohort and our follow-up period in this analysis ends in 2013, so an examination of the interactions between MHPAEA and the ACA is outside the scope of the current study. However, our study provides a baseline for future work examining how the ACA extends the MHPAEA requirements to additional populations (e.g., individual and small group markets, Medicaid expansion) through the EHBs.

The apparently modest impact of parity on service use has several possible explanations, in addition to any conservative biases due to study limitations. The Optum patient

population may not have been sick enough for the parity law to matter. For most people, insurance benefits were unlikely to have posed a binding constraint on their behavioral healthcare utilization even before parity. For example, Mark et al. (Mark et al., 2012) found that fewer than 10% of members used more than the maximum inpatient day or outpatient visit limits common before parity. Parity effects are likely to be found disproportionately among the sickest enrollees. Our sensitivity analysis of adults with schizophrenia and bipolar disorder – two of the most severe, chronic and expensive mental illnesses – found mixed results to support this conjecture.

As changes in benefit design were hypothesized to mediate the impact of MHPAEA on service use, a lack of association between MHPAEA and service use could be due to the absence of strong effects of MHPAEA on benefit design. Although (as noted in the literature review) there is evidence that MHPAEA did increase certain aspects of plan generosity, carve-out employers tend to be larger than carve-in employers and may have already been offering their employees generous behavioral healthcare benefits, so that few changes might have been required to carve-out plans in order to comply with MHPAEA and the IFR. Alternatively, lack of adequate enforcement could have led to non-compliance, particularly with regard to NQTLs.

Commercially-insured populations tend to be moderate- to high-income, suggesting that modest improvements in cost-sharing would have less impact than among more socioeconomically vulnerable populations. Non-financial constraints may be equally or even more important than cost-sharing in determining service use among this population. For example, factors such as stigma (especially for treatment of substance use disorders), availability of providers and geographic access may play a strong role in determining psychiatric specialty treatment patterns even when cost is not a major concern.

Knowledge of the law also appears to have been extremely limited. An APA survey found that only 4% of Americans knew about MHPAEA and a high proportion described their behavioral health care coverage in a manner suggesting that the coverage was not parity-

compliant (American Psychological Association, 2014). Patients may be reluctant to use services if they believe their financial exposure is high and it takes time and effort for information about coverage improvements to be disseminated, especially given the complexity of benefit design for behavioral healthcare. Preliminary work by our team suggests that a large number of variables would be required to fully describe benefit design, as the generosity of coverage for any particular service depends on the combination of benefit design feature (including numerous financial requirements, QTLs and NQTLs), type of disorder (MH vs. SUD), type of service and setting (with carve-out plan databases allowing for distinctions between many dozens of services), and network status (in-network vs. out-of-network). Given the challenges facing patients trying to figure out what their coverage is and how it has changed as a result of parity, it may simply be too soon to see a large impact on service use, although we did evaluate three years post-IFR.

In summary, even if plans are compliant, enacting a law may not change consumer behavior unless consumers are aware of their behavioral health coverage (and are able to overcome any other perceived constraints). Our findings suggest that other barriers to care should be evaluated (and if identified, addressed) in order for commercially insured patients to take full advantage of their behavioral healthcare benefit, especially as the ACA has now extended the MHPAEA provisions to other populations.

References

- Aliu O, Auger KA, Sun GH, Burke JF, Cooke CR, Chung KC, Hayward RA. The effect of pre-Affordable Care Act (ACA) Medicaid eligibility expansion in New York State on access to specialty surgical care. *Medical care* 2014;52; 790-795.
- Allen K. Mental Health Parity Act: Despite New Federal Standards. Mental Health Benefits Remain Limited. Hearing Before the Committee on Health, Education, Labor, and Pensions, US Senate 2000.
- American Psychological Association. Few Americans Aware of Their Rights for Mental Health Coverage. 2014.
- Barry CL, Gabel JR, Frank RG, Hawkins S, Whitmore HH, Pickreign JD. Design of mental health benefits: still unequal after all these years. *Health affairs (Project Hope)* 2003;22; 127-137.
- Barry CL, Ridgely MS. Mental health and substance abuse insurance parity for federal employees: how did health plans respond? *Journal of policy analysis and management : [the journal of the Association for Public Policy Analysis and Management]* 2008;27; 155-170.
- Busch SH, Epstein AJ, Harhay MO, Fiellin DA, Un H, Leader D, Jr., Barry CL. The effects of federal parity on substance use disorder treatment. *The American journal of managed care* 2014;20; 76-82.
- California Health Benefits Review Program. 2010. Analysis of Assembly Bill 1600: Mental Health Services. A Report to the 2009-2010 California Legislature
- Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 2015;50; 317-372.
- Du DT, Zhou EH, Goldsmith J, Nardinelli C, Hammad TA. Atomoxetine use during a period of FDA actions. *Medical care* 2012;50; 987-992.
- Ettner S. 2016. The Mental Health Parity and Addiction Equity Act (MHPAEA) Evaluation Study: Impact on Financial Requirements and Quantitative Treatment Limits for Substance Abuse Treatment among "Carve-In" Plans. 6th Biennial Conference of the American Society of Health Economists. Ashecon; 2016.
- Fox A, Oss ME, Jardine E, Marsh E. Open Minds Yearbook of Managed Behavioral Health Market Share in the United States, 2000-2001s. Open Minds; 2000.
- Frank RG, McGuire TG. Parity for mental health and substance abuse care under managed care. *The journal of mental health policy and economics* 1998;1; 153-159.
- Fretheim A, Soumerai SB, Zhang F, Oxman AD, Ross-Degnan D. Interrupted time-series analysis yielded an effect estimate concordant with the cluster-randomized controlled trial result. *Journal of clinical epidemiology* 2013;66; 883-887.
- Goldman HH, Frank RG, Burnam MA, Huskamp HA, Ridgely MS, Normand SL, Young AS, Barry CL, Azzone V, Busch AB, Azrin ST, Moran G, Lichtenstein C, Blasinsky M. Behavioral health insurance parity for federal employees. *The New England Journal of Medicine* 2006;354; 1378-1386.
- Goplerud EN. Consistency of Large Employer and Group Health Plan Benefits with Requirements of the Paul Wellstone and Pete Domenici Mental Health Parity and Addition Equity Act of 2008s.; 2013.
- Hacker KA, Penfold RB, Arsenault LN, Zhang F, Soumerai SB, Wissow LS. Effect of Pediatric Behavioral Health Screening and Colocated Services on Ambulatory and Inpatient Utilization. *Psychiatric services (Washington, D.C.)* 2015;66; 1141-1148.
- Henry J. Kaiser Family Foundation and Health Research and Educational Trust. Employee Health Benefits: 2003 Annual Survey 2008.
- Herrera C, Hargraves J, Stanton G. The impact of the Mental Health Parity and Addiction Equity Act on inpatient admissions. Health Care Cost Institute, Washington, DC 2013.

Hodgkin D, Horgan CM, Garnick DW, Merrick EL. Cost sharing for substance abuse and mental health services in managed care plans. *Medical care research and review : MCRR* 2003;60; 101-116.

Klick J, Markowitz S. Are mental health insurance mandates effective? Evidence from suicides. *Health economics* 2006;15; 83-97.

Kozhimannil KB, Adams AS, Soumerai SB, Busch AB, Huskamp HA. New Jersey's efforts to improve postpartum depression care did not change treatment patterns for women on medicaid. *Health affairs (Project Hope)* 2011;30; 293-301.

Lagarde M. How to do (or not to do) ... Assessing the impact of a policy change with routine longitudinal data. *Health policy and planning* 2012;27; 76-83.

Liang KY, Zeger SL. Regression analysis for correlated data. *Annual review of public health* 1993;14; 43-68.

Mark TL, Vandivort-Warren R, Miller K. Mental health spending by private insurance: implications for the mental health parity and addiction equity act. *Psychiatric services (Washington, D.C.)* 2012;63; 313-318.

McConnell KJ. The effect of parity on expenditures for individuals with severe mental illness. *Health services research* 2013;48; 1634-1652.

McConnell KJ, Gast SH, Ridgely MS, Wallace N, Jacuzzi N, Rieckmann T, McFarland BH, McCarty D. Behavioral health insurance parity: does Oregon's experience presage the national experience with the Mental Health Parity and Addiction Equity Act? *The American journal of psychiatry* 2012;169; 31-38.

McGinty EE, Busch SH, Stuart EA, Huskamp HA, Gibson TB, Goldman HH, Barry CL. Federal parity law associated with increased probability of using out-of-network substance use disorder treatment services. *Health affairs (Project Hope)* 2015;34; 1331-1339.

Sturm R, Goldman W, McCulloch J. Mental health and substance abuse parity: a case study of Ohio's state employee program. *The journal of mental health policy and economics* 1998;1; 129-134.

United States Government Accountability Office. Mental Health and Substance Abuse: Employers' Insurance Coverage Maintained or Enhanced Since Parity Act, but Effect of Coverage on Enrollees Varied 2015;GAO-12-63.

Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *Journal of clinical pharmacy and therapeutics* 2002;27; 299-309.