NBER WORKING PAPER SERIES

RULES VS. DISCRETION:
TREATMENT OF MENTAL ILLNESS IN U.S. ADOLESCENTS

Emily Cuddy
Janet Currie

Working Paper 27890
http://www.nber.org/papers/w27890

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2020

The authors thank Armin Falk, John List, James Heckman and seminar participants at the University of Chicago, the University of Pennsylvania, and the University of Lyon, as well as the BCBS Alliance for Health network for helpful comments. Financial support from the NOMIS Foundation and Princeton's Center for Health and Wellbeing is gratefully acknowledged. We are grateful to Pankti Pathak of BCBS and Allen Campbell at IQVIA for assistance accessing data. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Rules vs. Discretion: Treatment of Mental Illness in U.S. Adolescents
Emily Cuddy and Janet Currie
NBER Working Paper No. 27890
October 2020
JEL No. I1,I11

ABSTRACT

Mental health disorders are a leading cause of disability worldwide. Many mental health disorders start in adolescence and appropriate treatment at the outset may improve trajectories. We use a large national data base of insurance claims to examine the impact of initial mental health treatment on the outcomes of adolescent children over the next two years. We find that receiving follow up mental health treatment in the first three months after an initial mental health claim increases the total cost of care over the next 24 months. These higher costs are entirely accounted for by children who receive treatment that is not consistent with practice guidelines. Our estimates imply that, within 24 months, children who initially received a red-flag drug have 205% higher costs than those of the average treated child and are 131% more likely to have used an emergency room or experienced a hospitalization. These results show that large numbers of U.S. children are receiving mental health care that falls outside of accepted guidelines and poses risks to their health. In doing so, they provide support for the guidelines themselves, and demonstrate that analyses of large-scale claims data can provide a useful complement to clinical research studies in identifying best practices.

Emily Cuddy
Department of Economics
Julis Romo Rabinowitz Building
Princeton University
Princeton, NJ 08540
ecuddy@princeton.edu

Janet Currie
Department of Economics
Center for Health and Wellbeing
185A Julis Romo Rabinowitz Building
Princeton University
Princeton, NJ 08544
and NBER
jcurrie@princeton.edu

I. Introduction

Beginning with influential work by James Heckman (See Heckman (2000), Heckman, Stixrud, Urzua, 2006) economists have increasingly focused on non-cognitive skills as a determinant of children's future outcomes. This attention on non-cognitive skills suggests that closely related mental health measures deserve further attention. Borghans et al. (2008) point out that many youths face psychological problems that form a real constraint on their educational and vocational choices.

Many mental health disorders first manifest in adolescence. Mental health disorders are often more debilitating and harmful for a child's future than common physical health problems, increasing health care costs and the likelihood of being disabled while decreasing educational attainment and employment prospects (Currie et al., 2010, Smith and Smith, 2010). Clearly, initial treatment offers an important opportunity to intervene. Yet there is little information available on how U.S. adolescents are typically treated, or how variation in treatment impacts their outcomes.

There are however, widely agreed upon guidelines about what is appropriate for the initial treatment of children with common mental health problems. First, it is important that follow-up treatment begins promptly once a mental health problem is identified. Second, in most cases, one would expect to see a child receive therapy either with or without drug treatment (Birmaher et al., 2007; Connolly et al., 2007, McClellan et al., 2013). If medication is needed, finding a drug that works is often a matter of trial and error, but there are clear guidelines suggesting specific medications as a starting point and providing "red flags" suggesting that other medications be used cautiously if at all.

The use of guidelines in medicine is controversial, especially in psychiatry. Meehl (1954), Grove et al. (2000) and Kahneman and Klein (2009) argue that in general an algorithm could do as least as well as a psychiatrist in the treatment of mental illness. In the case of the many people who are treated for mental illness by family physicians with little specialized knowledge, the argument in favor of guidelines becomes even stronger. On the other hand, Frank and Zeckhauser (2007) argue that guidelines could result in care that is not sufficiently individualized. Again, this argument may be especially relevant in psychiatry given that optimal treatment for mental illness is known to involve trial and error.

This study focuses on the care that adolescent patients receive when they are first treated for mental illness, and how this care affects their health care trajectory over the next two years. We study a large national sample of 97,306 children who were covered by private health insurance for at least one year between 2012 and 2018, had their first mental illness claim between the ages of 10 and 17, and can be followed for at least two years after their initial mental health claim. We examine the effect of initial treatment on total health care costs, costs for mental health care, and the use of hospitals and emergency rooms over the next two years.

Clearly, the initial treatment received is not randomly assigned and is likely to be correlated with both observed and unobserved patient characteristics such as age, gender, and parental preferences. Supply-side variables offer a potentially exogenous source of variation in treatment. For example, shortages of child psychiatrists are an important factor limiting treatment in many areas (Thomas and Holzer, 2006; McBain et al., 2019; Findling and Stepanova, 2008). However, Cuddy and Currie (2020) show that several supply side factors of this kind collectively explain little of the overall variation in treatment probabilities across areas. Hence, they may prove to be weak potential instruments.

In this study, we use measures of practice style in the local health care market and their interactions with patient characteristics as instruments. This procedure generates a large set of potential instruments, raising concerns about weak instruments. That said, given this setting of "big data" and many potential weak instruments, we turn to the post-Lasso procedure proposed by Bellini et al. (2012) to guide our choice of instruments and assist in causal inference.

We find that receiving follow-up mental health treatment in the first three months after an initial claim increases total health care costs over the next 24 months. These higher costs are entirely accounted for by children who receive treatment that is not consistent with practice guidelines, and especially by those who receive "red-flag" drugs such as benzodiazepines, tricyclic anti-depressants, or other drugs which are not approved for use in children of their age. Our estimates imply that, within 24 months, children who initially received a red-flag drug have 205% higher costs than those of the average treated child and are 131% more likely to have used an emergency room or experienced a hospitalization. These results suggest that large numbers of U.S. children are receiving mental health care that falls outside of accepted guidelines and poses risks to their health. In doing so, they provide support for the guidelines themselves and demonstrate that analyses of large-scale claims data can serve as a useful complement to clinical research studies in identifying best practices.

The rest of this paper proceeds as follows. Section II provides background about treatment guidelines for children and about supply-side variations in the supply of physicians and in practice style. Section III describes our data. Section IV provides an overview of methods, and Section V presents our results. We conclude with a discussion and conclusions in Section VI.

II. Background

A. Guidelines for Initial Treatment of Children

As noted above, therapy is a first-line treatment for children. If warranted, medication is, of course, specific to the mental illness being treated. For depression, guidelines suggest a trial of Selective Serotonin Reuptake Inhibitors (SSRIs) (American Academy of Child and Adolescent Psychiatry (Birmaher et al. 2007). The American Psychiatric Association is even more specific and recommends Fluoxetine (generic for Prozac) as the first line treatment for depression in children since it has been the most studied (McQuaid et al., 2019).

These drugs are recommended because they have fewer side effects than older tricyclic antidepressants (TCAs), which can cause heart problems and are more likely to be fatal in overdose. Side effects of psychiatric medications appear to be common in children, though it is difficult to obtain reliable data. Hilt et al. (2014) conducted a large mail survey of parents of pediatric patients and found that 84% reported experiencing side effects.

SSRIs are also the first-line treatment for anxiety disorders, the most common class of disorders in our sample. While benzodiazepines are frequently prescribed for anxiety in adults, they are potentially addictive and the American Academy of Adolescent Psychiatrist's guidelines note that "benzodiazepines have not shown efficacy in controlled trials in childhood anxiety disorders…Clinicians should use benzodiazepines cautiously because of the possibility of developing dependence" (Connolly et al., 2007).

In addition to these guidelines, one can look at whether a drug has been approved for use in children by the U.S. Food and Drug Administration (FDA). The FDA has approved a wide variety of psychiatric drugs for children, depending on the indication. For example, while only two SSRIs are FDA-approved for depression (Fluoxetine and Escitalopram), several others are approved for the treatment of obsessive-compulsive disorder (OCD) and anxiety. The FDA has approved one TCA, Clomipramine, for OCD, and six atypical anti-psychotics.

Based on these guidelines, we examine the following measures of treatment: 1) Did the child receive follow-up treatment in the three months following the initial mental health claim? 2) If the child received treatment in that window, did they receive drugs? Treated children who did not receive drugs received therapy only. 3) If the child received drugs in the three-month window, did they receive TCAs, benzodiazepines, or drugs which are not FDA approved for children in their age group? Collectively, we refer to prescriptions that violate these guidelines as "red flag" drugs.[1] There may be some rare circumstances where their use is warranted, but given the available guidelines, one would expect most children with depression or anxiety to be prescribed an SSRI, and that children with other conditions would start with a trial of an FDA approved drug.

In a study of depression treatment for adults, Currie and MacLeod (2020) find that doctors who violate practice guidelines for adult patients tend to have patients with poorer outcomes. However, the available guidelines for adults are looser than for children and typically focus on transitions from one drug to another. There are several reasons for this. First, given the high costs of switching molecules (in terms of side effects and having to taper off and on), people tend to stick to the same drug they have been taking for long periods of time. Second, adults often have significant medical histories that can be used by their physicians to personalize treatment.

---

[1] We also considered poly-pharmacy as a candidate "red-flag" prescribing practice but found that it often consisted of a child receiving an SSRI with an adjutant atypical anti-psychotic in our sample. This variation seems to be an accepted practice, the theory being that the anti-psychotic helps the SSRI to work better. That said, Hilt et al. (2014) report that the number of children with side effects increased with multiple drug taking and that large numbers of children taking SSRIs with an atypical anti-psychotic experience side effects.

It would make little sense to recommend a drug that a patient had already tried and found ineffective, for instance.

Focusing on a child's initial treatment offers a setting in which more proscriptive guidelines could be useful since there will be little history available to guide the physician's treatment decision. And, given the persistence in treatment choices over time, the choice of initial treatment is likely to have long-term effects.

B. Supply-Side Variation in Access to Mental Health Treatment

"Small area variations" in medical care, the idea that observably similar patients receive quite different treatment in different places, have been extensively documented (c.f. Fisher et al., 2003a,b). Recently, several studies have argued that much of this variation is coming from the supply side of the market for medical care and not from differences in patient demand for medical services (Currie and MacLeod, 2017; Currie et al. 2016; Cutler et al., 2019).

For example, Finkelstein, Gentzkow and Williams (2015) show that half of the variation in treatment is associated with the location rather than the patient in a sample of elderly Medicare patients who change health service areas (HSAs). Moreover, because HSAs are quite large, 50% represents a lower bound on the amount of variation coming from the supply side—as Currie and MacLeod (2017) show, there is typically a great deal of variation in treatments offered within areas and even among doctors working in the same hospital or practice location.

As noted above, many observers blame shortages of child psychiatrists for variations in treatment of mental illness in children. According to the American Academy of Child and Adolescent Psychiatry, there are 8,300 practicing child and adolescent psychiatrists and approximately 15 million children in the U.S. who need their specialized knowledge.[2] There are also shortages of therapists trained to treat children. The National Institute of Health's Health Resources and Services Administration (2015) estimates that there is a shortage of more than 10,000 full-time equivalent mental health counselors, including school psychologists and counselors. Moreover, the available workforce tends to be concentrated in larger urban areas.

In the absence of sufficient mental health professionals, primary care physicians (PCPs) often end up prescribing psychiatric drugs. However, many report that they do not feel comfortable in this role—the majority feel that PCPs ideally should refer children with mental health problems to specialists for treatment but cite long waiting periods to see mental health professionals as a significant barrier to treatment (Heneghan, 2008; Fremont et al., 2008). These considerations suggest that some variation in treatment could be a function of what type of professionals are available to treat the child. However, Currie and Cuddy (2020) find that relatively little of the variation in treatment across areas can be explained by counts of different types of local providers.

There is also a large literature documenting variation in individual doctor's prescribing styles. Frank and Zeckhauser (2007) cite survey data showing that the most prescribed medication for a specific condition is responsible for about 60% of a doctor's prescriptions for that condition, and that different doctors have different favorite drugs. Patient demographics have little explanatory power. Berndt et al. (2015) use data on prescriptions of anti-psychotics. They show that most doctors have a favorite drug and that on average 66% of their prescriptions are for this drug. Again, different doctors have different favorites. Doctors also differ in terms of their diagnostic skill such that different doctors might diagnose the same symptoms

[2] See https://www.aacap.org/AACAP/Resources_for_Primary_Care/Workforce_Issues.aspx.

differently, leading to different treatments (Chan et al. 2019; Currie and MacLeod, 2017; Currie et al., 2016).

These differences in diagnostic skill and practice style could be due to differences in ability or training. They mean that treatments could vary across small areas just because the doctors working in those areas make systematically different decisions for similar patients. For example, if most of the doctors in an area prescribe benzodiazepines for adolescents, then a given adolescent seeking care in that area will be more likely to receive a benzodiazepine. Taken altogether, the literature suggests that small area variations in practice style are a possible instrument for individual treatment choices.

III. Data

We use administrative insurance claims data from the Blue Cross Blue Shield Alliance for Health Research (BCBS), a collaborative effort involving most of the regional BCBS plans.[3] These data have many strengths. The most obvious is the large sample size and detailed information about treatments. Previous large-scale analyses of mental health treatment for adolescents rely on parent/caregiver reports which may be subject to recall bias or bias introduced by survey non-response. Moreover, the questions asked about treatment of mental illness in national surveys are very general (i.e. whether the child has ever been treated) and do not include information about the setting or type of treatment. This lack of specificity makes it

[3] This limited data set is made available through a secure data portal and is drawn from Blue Cross Blue Shield (BCBS) Axis®, the largest source of commercial insurance claims data in the U.S. Accessing insurance claims data often requires extended negotiations with individual insurance carriers, or with government entities. Further information about the BCBS Health of American Initiative, including information about their Axis® data base and contact information is available at: https://www.bcbs.com/the-health-of-america/about.

impossible to say whether any treatment received is broadly consistent with evidence-based treatment guidelines.

While the lack of insurance coverage may be an important driver of overall differences in treatment, our focus on children with health insurance allows us to consider other determinants of variation, such as differences in practice style.[4] Although some mental health providers may be out of network, families can still file claims to BCBS and can usually be at least partially reimbursed. Hence, we would expect to see these cases appear in the claims data.

We select children who are observed before age 11 (typically from age 7 or 8) who had their first insurance claim for mental illness between the ages of 10 and 17. The mean age of these children at the time of the first mental health claim is 12, with most first claims at ages 10 to 14. We focus on first episodes because there are relatively clear guidelines about how these children should be treated.

The BCBS data include 4,356,831children who have a master member ID (which means that they can be followed over time); have claims dates consistent with their coverage period; and meet the age criteria for our study: They must be observed before the age of 11 and for at least one year between the ages of 10 and 18. In addition, they have drug coverage that has never been "carved out" over the time period that we observe them (N=2,223,930). If drug coverage was carved out of the BCBS plan, then they could have claims for psychiatric drugs under a separate drug plan, and we would not observe these claims. Children must also have valid geographic information, and consistent demographic information (age and sex) over the period that we observe them (N=2,201,566).

[4] Most U.S. children have health insurance either as dependents on their parent's employer-provided health insurance or through various publicly provided plans. However, one critique of public health insurance is that people often face difficulties accessing specialist care.

In our initial sample of 2,201,566 BCBS covered children, there are 202,066 with at least one claim related to mental illness, for an overall mental illness rate of 9.18%. This rate is comparable to estimates from the 2016 National Survey of Child Health, in which parents reported that 6.1% of 12 to 17 year old children had been diagnosed with depression, and 10.5% had been diagnosed with anxiety (Ghandour et al., 2019).

Restricting the sample further to those with non-missing provider information, a first mental health claim between the ages of 10 and 18, and a follow-up period of at least 24 months, yields a sample of 97,306. We would have liked to follow children for a longer period of time, ideally for many years, but longer time periods drastically reduced the sample size. In particular, we lose 39 percent of our current sample with an additional 12 months of follow-up (36 months total), and we lose nearly all (97 percent) with an additional 48 months (72 months total).[6]

Figure 1 is a map of the U.S. illustrating where the BCBS sample is located and the prevalence of child mental illness. The extent of the shaded area confirms BCBS's broad national coverage, although there are some areas of the country with less penetration. As shown in Appendix Table A1, the average child with BCBS coverage lives in an area which tends to be relatively younger, less diverse, and of higher socioeconomic status than the average American. The fraction of children with mental illness varies from zero to 5% in the first vingtile to 12 to 22% in the top vingtile. As we shall see, this is a relatively narrow band compared to variations in treatment.

[6] The maximum follow-up length in our sample is 77 months. This length is a function of both the date of diagnosis and the length of the current data extract, i.e., children who have a first claim later on in our data mechanically have a shorter potential follow-up period. Of the 97306 children in our final sample, we retain 59169 (60.8%) with a 36-month follow-up, 33240 (34.2%) with a 48-month follow-up, 15830 (16.3%) with a 60-month follow-up, and 3833 (3.9%) with a 72-month follow-up.

A. Children's Characteristics, Treatment, and Outcomes

In order to identify claims related to mental illness, we include claims with a diagnosis code F10-F69, F93, or F98 in the ICD10 (or equivalent codes in the ICD9); a procedure code indicating a mental health service, such as therapy; or the prescription of a psychiatric drug.[7] The categories F10-F19, F20-F29 ... to F60-69 cover substance use, psychotic disorders, mood disorders, anxiety, behavioral syndromes (including eating disorders), and personality and behavioral disorders. F93 and F98 are somewhat vague "disorders of childhood." Children with neurodevelopmental conditions are identified using ICD10 codes F80-89, F90-92, F94-F97 and F99, which include developmental disorders, ADHD, conduct disorders, Autism spectrum disorder, and Tic disorders.[8] Seventy-five percent of these children have Attention Deficit Hyperactivity Disorder (ADHD). We treat neurodevelopmental conditions as an important control but not as emerging mental illnesses per se. These neurodevelopmental conditions are generally present very early in life, much before the mental illnesses that are our focus here.

The outcomes we examine are: the total costs of medical care, the costs for mental health treatment, whether the child ever visited the emergency room or was hospitalized, and the number of days that they spent overnight in the hospital. An advantage of using the claims data is that we see the actual costs of care; we use the contracted reimbursable amount, which is the total combined cost to the insurer and patient. We look at both total costs and mental health treatment because mental illness can cause physical health problems like injuries as well as

[7] In our sample, 65.7% initially receive a mental health diagnosis, while 20.7% receive drug treatment without a diagnosis on the claim, and 13.6% receive a mental health procedure without a diagnosis on the claim. By the time children in our mental illness sample reach the end of our available sample in Dec. 2018, 79.4% of them have received a mental health diagnosis.
[8] We are excluding F70-79, intellectual disability, from consideration here.

requiring expenditures for mental health care itself. Visits to the hospital or the emergency room are an important outcome because people having a mental health crisis are universally advised to go to the nearest emergency room (ER), and once there, they may have to be hospitalized. Hence, these visits represent children who are not being successfully stabilized by mental health treatment or who may be having serious negative reactions to a medication. However, there is a large difference in the implications of an ER visit that ends in discharge, and one that ends in hospitalization both in terms of financial cost and in terms of disruption to the patient and family's life. Hence, we also look at the number of nights the patient is hospitalized as an indicator of the severity of any crisis. We examine all these measures at a three-month, 12-month, and 24-month time horizon.

Table 1 provides an overview of our sample. The first two columns in the first panel compare the children with mental health conditions to the full sample of BCBS children meeting our other sample inclusion criteria. Relative to this sample, the children who have a first claim for a mental health issue during our sample period are more than twice as likely to be hospitalized, more likely to visit the ER, and have average monthly costs over the sample period that are double those of the average BCBS child. They are about three times more likely to have an underlying neurodevelopmental condition, which is usually ADHD, confirming that these children are at higher risk of future mental health conditions such as depression and anxiety.

Turning to the type of first claim, about 1.2% have a first mental illness claim that resulted from a hospitalization, 3.3% have a first claim stemming from an ER visit, and 42.1% have a first mental health claim stemming from a formal evaluation. This latter number is interesting because it shows that less than half of the children are receiving a formal mental health evaluations even when they go on to be diagnosed and treated. While it is possible for an evaluation to show that a child does not have a mental illness, the majority of these evaluated children (83.3%) go on to be treated for mental illness by the end of our observation window.

Column 3 shows the further effect of limiting our sample to children who can be followed for two years. This sample is a little younger and may be somewhat sicker given its higher overall utilization of medical care. It is possible that parents who have a sick child are less likely to exit the sample by changing jobs (Bansak and Raphael, 2008).

Columns 4 through 8 show how the children's characteristics are related to the type of treatment that they receive. Looking first at the number of observations in each category, 73% receive some form of follow-up treatment in the next three months after an initial claim. The most common type of treatment received was therapy alone, accounting for 59.6% of treatment. Another 40.4% receive drugs only. Only a small share of treated children, 7.6% receive both drugs and therapy which is surprising in view of guidelines suggesting that most children should receive therapy. And of children receiving drugs, just under half, 45.7% receive a red-flag drug treatment.

Column 4 shows that 27% of children receive no treatment in the three months following an initial claim. For comparison, in the 2016 National Survey of Child Health, parents reported that 79.0% of 12 to 17 year old children with depression diagnoses and 63.7% of children with anxiety diagnoses had been treated (Ghandour et al., 2019). The treated children are actually very similar to our overall sample of children so that there are no obvious differences that might explain why some children are treated and others are not.

Turning to the types of treatment, girls, and children whose first claim resulted from a hospitalization or ER visit, are over-represented in the small group treated with both therapy and drugs. Children who first received an evaluation are over-represented in the group receiving therapy alone, and under-represented in the group who receive any drugs, including red-flag drugs. On average, children who receive drugs have higher costs than those who receive therapy alone, as well as more visits to the ER and hospitalizations. While this may be counter-intuitive since therapy may involve more visits, therapy is often delivered by providers such as social workers who are less expensive per visit than psychiatrists. The highest average costs are incurred by children who get red-flag drugs in the first three months after their initial claim.

Given the guidelines discussed above, we would have expected most children in our sample to be initially prescribed an SSRI, and one that was specifically FDA approved for use in children. Of the children receiving drug treatment in our sample, 45.3% receive an antidepressant (32.5% get SSRIs), 55.9% get anti-anxiety medications, 2.3% get a mood stabilizer, and 4.2% get anti-psychotic medications (recall that 1.2% of sample children have a diagnosis of a psychotic disorder). Of the children receiving anti-psychotics, 40.4% take them with an SSRI, which is a common combination treatment for depression. Of children with a red-flag prescription, 50.7% get benzodiazepines, 23.2% get TCAs, and 58.3% get a drug that is not FDA approved for a child of their age. Nearly 15% of children who receive drug treatment face prescribing that raises more than one red flag.

The second panel of Table 1 breaks the sample down by diagnosis code. In what follows, children may have more than one diagnosis. By far the largest category is anxiety, which affects over half of the children in our sample. These children are disproportionately more likely to get therapy and less likely to get drugs or red-flag drugs. Mood disorders such as depression are the second most important category, affecting 13.6% of sample children. These children are over represented in the therapy plus drugs category, though that is quite a small slice of the treatment pie. The other diagnostic categories are quite small.

Not every child has a diagnosis code in the data—in our sample, almost a third are initially missing diagnosis codes even though they are being treated with therapy or psychiatric drugs. In some cases this may be because the diagnosis code will not affect the provider's reimbursement. For example, if PCPs prescribe anti-depressants during a routine visit, then they may only bill for the routine visit and not include any diagnostic codes on the insurance reimbursement claim. Another possibility is that the doctor is worried about stigmatizing the child with a mental health diagnosis if the condition might prove to be short-lived, or if the doctor is uncertain about the child's diagnosis. We include these children in our main models but also present robustness checks excluding them from the analysis.

B. *Supply-Side Measures*

Following Kessler and McClellan (2000) and Currie and MacLeod (2017), we define the market facing people who live in a particular zip code using information about where children in that zip code actually go to receive mental health care each year. For all children who live in a given Census Zip Code Tabulation Area (ZCTA), we examine up to the 10 most common ZCTAs visited by BCBS children in order to receive mental health treatment over the entire sample period. For example, Figure 2 illustrates the definition of the market area for the ZCTA that includes Princeton, New Jersey.

This procedure has several advantages relative to defining a small area based on an arbitrary geographical definition such as a county or HSA. First, only providers who are actually available to treat BCBS children at some point over the sample period are included. Second, the measure scales naturally. In a rural setting, where people drive long distances to get to a grocery store, for example, it may not be unusual to drive a long distance to see a psychiatrist. Third, providers can serve clients from more than one ZCTA and do not have to be arbitrarily assigned to one market or another. Fourth, the market definition is specific to psychiatric treatment. On average, in ZCTAs that have at least 20 BCBS children, there are 8.6% BCBS children with a first mental health spell per ZCTA.

Because rates for small cells may be noisy, we also turn to an additional source of possible supply-side measures: National data on prescriptions of anti-depressants and anti-anxiety medications from retail pharmacies. These data come from IQVIA's LRx data base.[9] An advantage of using the IQVIA data to supplement information from BCBS is that we can see prescriptions from all of the providers treating children in a particular area, even those who do not treat BCBS children.

In order to identify the types of mental health professionals who treat BCBS children, we first tag any providers who provided mental health treatment to at least one child in the BCBS claims data. We merged provider records from the claims data with data from the National Plan and Provider Enumeration System using the provider's National Provider ID. We then use the NPPES taxonomy codes to recover the provider type.[10] The supply measure is then calculated by dividing the total number of providers within each specialty in the ZCTA market and year by the number of BCBS children present in the ZCTA regardless of mental health status.

[9] IQVIA (formerly known as IMSQuintiles) is a public company specializing in pharmaceutical market intelligence The IQVIA data is available for purchase to qualified researchers. For further information, contact Allen.Campbell@iqvia.com.

[10] Psychiatrists include NPPES codes 2084P0800X and 2084P0804X. Primary care physicians providing mental health services are largely pediatricians, but may also include doctors in family medicine, general practice, adolescent medicine, or developmental/behavioral pediatrics (NPPES codes 208000000X, 2080A0000X, 2080P0006X, 207Q00000X, 208D00000X). Therapists include psychologists, social workers, and mental health counselors (NPPES codes 1041C0700X, 101YM0800X, 101YP2500X, 103TC0700X, 103T00000X, 106H00000X, 101Y00000X, 104100000X, 103TC2200X).

By examining only psychiatrists and therapists who actually treat any BCBS children, we ensure that we are focusing on the relevant group of mental health professionals for the children in our sample. For example, we rule out psychiatrists who only treat adults, or who are not actively practicing. It would not be possible to focus on this more relevant group of clinicians using other sources such as the National Plan and Provider Enumeration System (NPPES).

The area-level measures that we use as possible instrumental variables can be thought of as measures of practice style and include the following six measures:

(1) Of the BCBS providers providing mental health treatment to children, what share are psychiatrists?

(2) Of the BCBS providers providing mental health treatment to children, what share are primary care physicians?

(3) For all of the available primary care physicians treating BCBS children, what share of each physician's child caseload receives a psychiatric drug (averaged over all the physicians[11])?

(4) For all of the available prescribers of psychiatric drugs to BCBS children, what is the share of "red-flag" drugs in each prescriber's portfolio of psychiatric prescriptions (averaged over all physicians[12])?

(5) Of all of the IQVIA prescribers of first psychiatric drug prescriptions for children in the child's ZCTA, what fraction were psychiatrists?

[11] When calculating this average, we weight shares by the physician's caseload, i.e., the number of patients that the physician sees. We also experimented with weighting by the number of prescriptions written by the physician, but results were very similar.
[12] Here, we use a physician's psychiatric caseload to weight shares.

(6) For all of the IQVIA prescribers of first psychiatric drug prescriptions for children in the child's ZCTA, what was the share of "red-flag" drugs in each prescriber's portfolio of anti-depressant and anti-anxiety drugs (averaged over all physicians)?

Measures (1) and (2) capture the composition of the supply of providers serving BCBS children. Note that the total number of BCBS providers providing mental health treatment includes therapists. Measure (5) is also likely to be affected by the supply of psychiatrists available to treat children. It includes all prescribers, not only those seeing BCBS children, but captures only prescribing behavior. Measure (3) looks at whether PCPs who serve BCBS children with mental illness seem to specialize in mental health treatment. Even if there are few PCPs available to serve children with mental health needs, if each of them has a large share of such children in their practice, then they might develop more expertise than a PCP who treats few such patients. Measures (4) and (6) focus on whether the BCBS physicians or the larger set of prescribers follow the prescribing guidelines, or whether they use prescriptions that raise a red flag in terms of violating guidelines.

While Table 1 provided a child-level overview of the data, Table 2 provides an area-level perspective. The first panel explores variability in treatment across areas. If we focus on the interquartile range, which is less likely to be affected by small cell sizes, we can see that the fraction of children who receive follow up treatment within 3 months varies from 56% to 100%. There is an even larger range in the fraction of children who receive therapy alone (29% - 100%). One can also see that in most areas, no children receive drugs and therapy together. It is only when one gets up to the 75th percentile of the distribution that one starts to see any children with combination therapy. The fraction of children treated with drugs alone varies from zero at the 25th percentile to 57% at the 75th percentile. The inter-quartile range in the fraction of children receiving any red-flag drug treatment varies from 0 to 1. Most of this variation is coming from different probabilities of receiving benzodiazepines and drugs which are not FDA approved for children of a given age.

The second panel of Table 2 shows that there is also a great deal of variation in the availability of different types of mental health providers across areas. The number of psychiatrists treating BCBS children per 1,000 BCBS children ranges from 4.91 at the 25th percentile to 11.97 at the 75th percentile. The number of therapists is higher ranging from 19.60 to 37.71 over that range, and the number of PCPs providing mental health care is larger still, ranging from 26.6 at the 25th percentile to 60.07 at the 75th percentile. One issue with these conventional measures of supply side characteristics is that they do not exhibit very much variation within market areas over time—most of the variation is between markets. In practice, this means, for example, that many rural areas are chronically short of child psychiatrists.

In what follows, we do not include these measures in our instrument list. We think that it is important to include market-area level fixed effects in our regression models to control for time invariant or slowly moving factors that may affect both treatment and our outcome measures. But in models with area fixed effects, the lack of within-area variation is problematic. A robustness check discussed further below demonstrates that when these variables are included in the list of potential instruments, they are never selected by the Lasso procedure discussed below.

The third panel of Table 2 focuses on measures of physician practice style, which have greater promise as instruments because several of them show considerable variation within markets over time in addition to variation between markets. These are the six instruments listed above. It is striking that the share of mental health providers who are psychiatrists is low, only reaching 18% at the 90th percentile of the area-level distribution. The share who are PCPs is also quite low, reaching 24% at the 90th percentile. These figures indicate that in most places, it is therapists (psychologists, social workers, and licensed clinical therapists) who make up the bulk of the mental-health workforce. These two variables show the least within-area variation of the six, yet they and their interactions are sometimes chosen in the Lasso procedure, so we continue to include them in our list.

Since PCPs are an important source of mental health care, we also look at the share of each PCP's patients who receive a mental health drug, which has an interquartile range between 5 and 9%. Hence, while many PCPs are doing some prescribing, each one is doing relatively little, suggesting that they may not develop a lot of expertise in providing mental health treatments.

Since we are especially concerned with the question of whether treatment follows guidelines, we look at the share of each mental health provider's prescriptions that are for red-flag drugs. This share is remarkably high, and shows less variation across areas than some of the other measures, ranging from 47% at the 25th percentile to 54% at the 75th percentile. However, 52% of the variation in this measure is within areas, making it a good candidate instrument in our area fixed effects models.

As discussed above, we compute two supplementary measures of practice style using the IQVIA prescriptions data. A nice feature of these data is that it is possible to identify the first prescription that a child received. Table 2 shows that the share of new prescriptions by psychiatrists is remarkably small, consistent with the small share of mental health providers who are psychiatrists. The interquartile range is from 5 to 9%. The IQVIA data on the share of new prescriptions that are for red-flag drugs is consistent with the BCBS data in that it is also quite high: At the 25th percentile, 32% of new prescriptions are for red-flag drugs, rising to 40% at the 75th percentile and 44% at the 90th percentile. Half of the variation in this measure is within area.

IV. Methods

Our goal is to investigate the effect of initial choices about the treatment of adolescents with emerging mental illness on their medical trajectory over the next two years. We are particularly interested in whether treatment that follows widely accepted guidelines is better than treatment that does not.

The baseline linear probability/OLS models take the form:

(1) $Outcome_{ymzi} = \beta_0 + \beta_1 * Treatment_{ymzi} + \beta_2 * Female_{ymzi} + \beta_3 * Age_{ymzi} + \beta_4 *(1st\ claim\ is\ hospitalization)_{ymzi} + \beta_5 *(1^{st}\ claim\ is\ an\ ER\ visit)_{ymzi} + \beta_6 *(1st\ claim\ is\ an\ evaluation)_{ymzi} + \beta_7 *(Any\ ER\ or\ hospital\ visit\ last\ 6\ months)_{ymzi} + \beta_8 *(Diagnosis\ codes)_{ymzi} + \beta_9 *(Neurodevelopmental)_{ymzi} + \mu_m + \gamma_y + \zeta_z + \varepsilon_{ymzi}$.

Treatment is one of the four outcome variables discussed above. The first six variables are all zero-one indicators. *Age* is entered linearly, because estimation using single year of age indicators showed that treatment modalities do change in quite a linear way with age. *Diagnoses* are captured with eight categories corresponding to F10-F19, F20-F29, … F60-69, F93/F99, and a "no diagnosis" category.[13] Measures of whether the first claim was for an evaluation, an ER visit, or a hospitalization are included as indicators of the severity of the child's condition. Whether the child was hospitalized for any reason over the past six months can be viewed as a measure of physical health status (since we are dealing with first claims for mental health, they would not have been hospitalized for a mental health condition in the past six months). *Neurodevelopmental* indicates that the child had an existing neurodevelopmental condition prior

[13] The diagnosis brackets are not mutually exclusive. Some children receive multiple diagnoses, which span the main ICD-10 categories.

to this episode such as ADHD or autism spectrum disorder. The indicators $\mu_m + \gamma_y + \zeta_z$ are month, year, and ZCTA fixed effects. Standard errors are clustered at the level of the ZCTA to allow for correlations between children facing similar supply-side measures.

The month and year fixed effects capture possible seasonal effects (such as those induced by the school calendar) and time trends in outcomes and treatment propensities. The ZCTA fixed effects control for relatively fixed differences across small areas such as urbanicity, racial composition, and median income. They will also capture the mean numbers of mental health professionals per capita over the sample period as well as area-level variables such as the average distance to a hospital or emergency room.

Establishing causality in such a model is difficult because even conditional on variables like diagnosis, age, gender, and the measures of initial severity that we can see in the claims data, there are likely to be omitted unobserved variables that affect treatment choices. One possibility is that children who are sicker in an unobserved way will be more likely to be treated, and will be treated more aggressively (i.e. with drugs, and perhaps with red-flag drugs). Alternatively, it is possible that high SES parents demand prompter and more aggressive treatment for their children, even if their children are on average less sick. It is even possible that some parents demand certain drugs for their children, though that seems unlikely to be a widespread reason for treatment choice in these cases of initial mental health treatment where parents may not have a lot of experience with the condition and available treatments.

While we will estimate linear probability or ordinary least squares models as a baseline, causal inference will require identifying instruments that affect treatment decisions but are uncorrelated with these omitted variables. Area-level measures of practice style offer a possible solution to this problem. If most prescribers of psychiatric drugs in the market area prescribe benzodiazepines to children, then a child should be more likely to receive them than if the child was living in an area where no one prescribed them.

That said, by controlling for ZCTA fixed effects, we control for cross-sectional differences in supply-side variables that might be correlated with omitted characteristics of local areas that impact both treatment and supply. This means that only variations in the practice style measures within small areas over time are left to identify the effects of the supply-side instruments on treatment choices.

The simplest version of a first stage model that we could estimate with our six potential instruments is:

$$(2)\ Treatment_{ymzi} = \beta_0 + \beta_1{*}Female_{ymzi} + \beta_2{*}Age_{ymzi} + \beta_3{*}(1st\ claim\ is\ hospitalization)_{ymzi} + \beta_4{*}(1^{st}\ claim\ is\ an\ ER\ visit)_{ymzi} + \beta_5{*}(1st\ claim\ is\ an\ evaluation)_{ymzi} + \beta_6{*}(Any\ ER\ or\ hospital\ visit\ last\ 6\ months)_{ymzi} + \beta_7{*}(Diagnosis\ codes)_{ymzi} + \beta_8{*}(Neurodevelopmental)_{ymzi} + \mu_m + \gamma_y + \zeta_z + \varepsilon_{ymzi},$$

which includes the same controls as equation (1) but with the addition of the six candidate instruments discussed above and their squared terms.

However, these instruments collectively explain little of the variation in treatment. One potential solution to this problem is to interact the instrumental variables with measures of individual demographics, severity, and diagnosis. Substantively, this procedure is justified if for example, in areas with many prescribers of benzodiazepines, children who first present in a hospital or ER are more likely to receive them.

Interacting the instruments and the controls leads to a very large set of potential instruments, raising concerns about bias from weak instruments in a two-stage least squares (TSLS) setting. Hence, we use the "post-Lasso" TSLS estimator discussed in Bellini et al.

(2012, 2014) and implemented in different contexts by Bellini et al. (2012) and Sands and Gilchrist (2016). This method involves using Lasso to select the instruments to be included in the first stage and then re-estimating OLS on the first stage with the Lasso-selected instruments (hence, "post"-Lasso).

Given our confidence in the need for patient controls such as age and gender in our models in both the first and second stage, we require Lasso to select the individual controls that are to be included in the second stage.[14] The exogenous variables that we require Lasso to select in the first stage include: *Female, Age, 1st claim is hospitalization, 1st claim is an ER visit, 1st claim is an evaluation, Any ER or hospital visit last 6 months, Diagnosis codes, Neurodevelopmental,* μ_m, γ_y, and ζ_z.

As potential instruments, we include a second degree polynomial of the original vector of six instruments as well as interactions of the six instruments variables with *Age, 1st claim is hospitalization, 1st claim is an ER visit, 1st claim is an evaluation, Any ER or hospital visit last 6 months, Diagnosis codes,* and *Neurodevelopmental.* Since including all of the possible diagnosis codes results in a very large instrument set and many of the codes apply to few children, we group diagnoses into four large groups for the purposes of the interactions: Anxiety or depression, adjustment disorders (F43.2, adjustment disorders, the largest single diagnostic category in our data), no diagnosis, and other. In the end, these interactions result in an instrument vector with 87 potential instruments.[15] We use the efficient F-statistic discussed by

[14] Alternatively, we could have partialled-out each control beforehand.

[15] The full potential instrument vector includes a second-degree polynomial of the six instruments (N=27) and interactions with age (N=6), four severity measures (N=24), and five diagnosis indicators (N=30).

Montiel Olea and Plueger (2013) to check that the instruments selected by Lasso pass a test for weak instruments.

Once Lasso selects the first-stage instruments, we estimate the effect of treatment on a child's health outcomes in the second stage. This estimation is similar to the baseline model in equation (1) replacing the child's actual treatment status with predicted treatment from the first-stage regression. Our main results are based on the set of instruments chosen using a refined data-driven penalty, as described in Bellini et al. (2012). In our application, this procedure selects between two and eight instruments.[16] We also present a variety of robustness checks using different subsets of potential instrument vector (e.g. the "top 3" instruments selected by Lasso in the first stage).

V. Estimation Results

A. First stage estimates

We begin with an analysis of the factors that influence the treatment received in the three months following an initial claim for mental illness. An initial set of first stage estimates including only the six original instruments and their squared terms is shown in Table 3. They show that the probability of any treatment, drug treatment, and "red-flag" drug treatments all increase with age, and are higher for girls than for boys. Children whose first claim resulted from a hospitalization are less likely to be treated in the next three months but more likely to be treated with drugs and more likely to receive red-flag drug treatments. Children whose first mental health claim resulted from an ER visit are also less likely to receive any follow-up treatment in the next three

[16] See Appendix Table A2 for the full list of instruments chosen by Lasso in each model discussed in the paper.

months. This might reflect a lack of access to care, which may be one reason they were first seen in an ER (or hospital) setting. Children who were hospitalized at all in the past six months for other reasons are more likely to be treated, more likely to receive drug treatment, and more likely to receive red-flag drug treatment than other children.

Turning to the practice style instruments, we can see that some are individually statistically significant in a first stage regression model. Column (1) shows that the share of BCBS mental health providers who are psychiatrists is negatively associated with getting follow up treatment within three months at low levels and positively associated with treatment probabilities at high levels. Column (2) indicates that conditional on receiving any treatment within three months of the initial claim, the BCBS share of mental health providers who are primary care physicians is positively associated with drug treatment at low levels, turning negative at higher levels. The share of new prescriptions in the IQVIA data that are written by psychiatrists is inversely related to the probability that treatment involves drugs at low levels, but becomes positive at higher levels.

Columns (3) and (4) both look at the probability of receiving red-flag drug treatments, where column (3) is conditional on any treatment, and column (4) is conditional on any drug treatment. None of the supply-side instruments are individually statistically significant in column (3), while in column (4) the BCBS mean share of provider prescriptions that are for red-flag drugs has a positive effect on red-flag treatment at high levels, and the IQVIA share of new prescriptions by psychiatrists has a positive effect at low levels, and a negative effect at high levels.

The key issue, however, is not whether any of these variables are individually statistically significant, but whether they jointly explain much of the treatment variables. This question is addressed in Table 4, which summarizes the R-squared and the Montiel Olea and Plueger effective F statistics (F_{EFF}). The first column shows the R-squared for the OLS regression without any instruments as a baseline. Column (2) repeats the R-squareds from Table 3, the first stage regression with the six supply-side instruments and their squared terms. For the most part, the R-squareds are identical to the baseline column (1) models, showing that the instruments collectively explain little of the variation in the treatments. Not surprisingly then, the F_{EFF} for these models range from 2 to 10, indicating that they do not pass a weak instruments test.

It is reasonable to expect the effects of the practice style variables to vary considerably with demographics, severity, and diagnosis. Column (3) shows that a first stage model that included all of the interactions discussed above does fit better, and the value of F_{EFF} also shows that such a model would pass the weak instruments test in the models for treatment and for drugs conditional on treatment. But the full interaction first stage would not pass a weak instruments test in the models of red-flag drug treatments, which are of particular interest.

The last two columns show first-stage models where Lasso has selected a more limited set of controls. In column (4), the Lasso is constrained to select at most three instruments. These models do not fit as well as the "full interaction" model would, but they do appreciably improve the fit relative to no instruments or to the main 12-item vector of instruments. Column (5) uses the instrument set chosen by Lasso using the refined data-driven penalty discussed in Bellini et al. (2012). These models fit almost as well as the "full interaction" model, but are much more parsimonious. The F_{EFF} statistics in columns (4) and (5) indicate that all of the models pass a weak instruments test. Hence, we will focus most of our attention on the post-Lasso estimates using the refined data-driven penalty instrument sets in what follows.

B. Two Stage Least Squares Estimates of the Effects of Treatments on Outcomes

Table 5 focuses on costs, both overall total costs and costs for mental health care specifically. We examine effects at 3 months, 12 months, and 24 months in order to see both the short-run and medium-term effect of treatment choices. The first row shows OLS estimates. Not surprisingly, the initial effect of treatment on costs is positive, increasing overall costs 44%, where much of this increase is accounted for by the cost of mental health treatment itself. Over the next 12 to 24 months, overall costs increase by a slightly smaller amount, suggesting some cost savings in the longer term, even though costs for mental health care continue to rise. However, these OLS results may be biased by selection into treatment.

The post-Lasso TSLS estimates suggest that the OLS estimate of the effect of treatment on total costs is biased upwards, especially in the short run. However, by 24 months, the effects on total costs are remarkably similar in the OLS and TSLS models. The effects of treatment on mental health costs are consistently larger in the TSLS models than in OLS. Overall, the results indicate that receiving mental health treatment in the first three months after an initial claim is associated with higher costs for mental health treatment, leading to higher overall health care costs in both the short and medium run. That is, the higher initial costs are not offset by lower costs down the road.

The second panel of Table 5 shows the effect of receiving drug treatment, conditional on any treatment. Children who did not receive drug treatment received therapy alone. The OLS estimates suggest that mental health costs are lower for people who receive drug treatment, but that overall health care costs are higher. However, the post-Lasso TSLS estimates indicate that both types of costs are significantly higher when children initially receive drug treatment. Effects on total costs are somewhat front-loaded with large initial increases in costs that gradually taper off. It is possible that some of the initially high costs reflect negative drug reactions. By two years out, children who are receiving drug treatment have total costs 94% higher than those who received therapy alone. This finding supports guidelines which suggest that it is useful to start with therapy.

The third and fourth panels address the question of the type of drug therapy received, and whether it is broadly consistent with practice guidelines or whether it raises a red flag. The OLS estimates suggest that red-flag drug treatment reduces mental health costs, while increasing overall costs. However, the post-Lasso estimates reverse the sign of the estimates for mental health care costs. They suggest that red-flag prescribing raises both mental health costs and total costs much more than other forms of treatment.

Panel 4 focuses on children who received drug treatment, and looks at the effect of receiving red-flag drugs in this sub-population. Here the post-Lasso estimates confirm that there are higher total costs by 12 months, persisting to 24 months, when the total costs are 238% higher than the costs for children who received drug treatment that was consistent with guidelines (this is just the exponentiated coefficient from Panel D). Initially mental health costs are lower, but by 24 months there is no statistically significant effect. These estimates suggest then that it may be initially cheaper to treat a child with red-flag drugs, but that it increases non-mental health medical care costs.

Table 6 examines effects on facility use. Since ER visits and hospitalizations are very expensive, they could be a significant driver of higher health care costs. These visits are also important outcomes in their own right since ER visits and hospitalizations represent traumatic events for children and their families and suggest that the initial treatment pursued was not effective and/or led to significant side effects. OLS estimates indicate that treatment is associated with higher facility use, measured as an ER visit or hospitalization. Treatment is also associated with 0.56 more nights in the hospital by 24 months out, compared to no treatment within three months.

The post-Lasso TSLS estimates suggest that these effects are biased towards zero in OLS. Among children who are treated, the probability of an ER visit or hospitalization rises in the first three months by 62 percent. The effect falls to a 49.2% higher rate of facility use over the next 24 months. We also see a significant increase in the number of overnight stays in the hospital to 2.58 by 24 months out. These results suggest that the higher costs for treated children observed in Table 5 are in part accounted for by greater facility use. Possibly, once they are in treatment children are more likely to be taken to hospital when there is any deterioration in their condition.

Another possibility is that treated children are more likely to go to the hospital because drug treatment involves a higher risk of side effects. The second panel of Table 6 explores this possibility by examining the effect of drug treatment conditional on any treatment. The TSLS estimates show higher probabilities of facility use in the first three months of treatment. The instruments further suggest an increase at 12 months, though not at 24 months. These results are consistent with the idea that particularly in the short-term after treatment is first initiated, drug therapy is associated with a higher number of visits to the hospital than therapy alone. However, the estimates also suggest that by 24 months there is a reduction in the number of nights in the hospital of about one and a half days. Hence, drug therapy may involve a tradeoff: A higher short-run risk of ER visits due to side effects, but a longer-term benefit as a tolerable drug regime is found by trial and error.

The last two panels of Table 6 turn to the question of whether we can detect a beneficial effect of following drug guidelines when prescribing for children. The third panel shows that compared to children who receive other treatments, children who received red-flag drug treatments in the first three months after an initial claim are much more likely to have a visit to the hospital: The probability rises by 20 percentage points on a baseline of 10.1%. The estimated effects on overnight stays in the hospital are too imprecisely estimated to be informative but suggest that much of the increase in facility use may be coming from visits to the ER which do not result in overnight stays.

The last panel of Table 6 offers a sharper comparison by focusing on children who received drugs in the first three months after their initial claim, and comparing children who received red-flag drugs to those who received other drugs. The estimates on any facility use show a 10.9 percentage point higher probability after three months on a baseline of 14.9%, for an increase of 73%. This effect falls to a 48.4% increase by 24 months. Again, the estimated effects on overnight stays in the hospital are too imprecisely estimated to be informative. Overall, these estimates suggest that receiving a prescription that violates practice guidelines is associated with a higher probability of ER visits, which is one of the drivers of higher costs in these children.

To summarize, both the OLS and the post-Lasso TSLS estimates indicate that total costs are significantly higher for children who receive follow-up treatment in the next three months after their initial claim. Conditional on treatment, those who receive drug treatments have even higher total health care costs, and conditional on receiving drug treatment, those who get red-flag treatments have the highest costs. The post-Lasso estimates are consistently higher than the OLS estimates, suggesting that the marginal child who receives drug treatment or red-flag treatments due to local variations in practice style incurs higher overall health care costs.

The estimated effects of treatment types on mental health costs are more complicated. We see that treatment does increase mental health care costs as one would expect, in both OLS and the post-Lasso estimates. Conditional on treatment, drug treatment has a negative effect in OLS, as does red-flag drug treatment conditional on drug treatment. However, the former effect is reversed in the post-Lasso models, suggesting that drug treatment raises mental health care costs overall. Red-flag treatments are estimated to increase costs further conditional on being treated.

Figure 3 traces out the implications for total costs of each choice in the treatment decision tree holding all else equal. The numbers are based on the estimates in Table 5, and we focus here on the total costs at 24 months. Figure 3 has a number of interesting implications. First, although treated children have higher average costs than untreated children, this is primarily due to higher costs among the group getting drug treatment: Children who receive only therapy have lower costs over the next 24 months ($3415) than children who did not receive follow up treatment within three months ($3537). Second, drug treatment leads to higher costs mainly because of the high cost associated with red flag drug treatments: Children who had drug treatments that followed guidelines have average costs of $4052 over the next 24 months compared to $9653 for children receiving "red flag" drug therapies.

Some of the higher costs associated with any treatment, drug treatment, and red-flag treatments may be due to higher rates of facility use (defined as ER visits or hospitalizations). Figure 4 traces out the implications of each step in the decision tree for cumulative use of ER rooms and hospitalizations over 24 months. Children with red-flag drug treatments have the highest facility use, but this does not entirely explain the higher facility use among treated children: Even children with therapy alone have higher facility use than children who did not immediately receive follow-up treatment. It is possible that people in treatment are more likely to be advised to use the ER as a way to stay safe.

C. Robustness

In the discussion above we have emphasized conclusions based on a post-Lasso TSLS procedure with a set of instruments chosen using the refined data-driven penalty discussed in Bellini (2012). However, it is useful to think about how robust these conclusions are to changes in the instrument set. In this section, we briefly discuss estimates using alternative instrument sets. The first uses the "top three" instruments chosen by the Lasso procedure. These estimates are shown in Table 7 (for costs) and Table 8 (for facility use).

On the whole, these results are remarkably consistent with those discussed above. The results for the effects of any treatment in the three months following an initial claim are almost identical. Estimates for the effects of drug treatment conditional on treatment are smaller than those discussed above, but show the same qualitative patterns. The estimates for the effects of red-flag drugs conditional on treatment are very similar for costs, but differ in sign for the effects on facility use after 24 months. However, the estimated effects of red-flag drugs conditional on any drugs are remarkably consistent in the two specifications. On the whole then, the estimates are quite robust to using this alternative instrument set. We prefer the main results presented previously because in all cases, the optimally chosen instrument set has more explanatory power in the first stage than the top three instruments.

Another issue raised above has to do with children in our sample who are missing official diagnoses. While we believe that most of the children with missing diagnoses do have a mental health condition and so arguably should be included in the main sample, we have also reestimated our models excluding these children. These results are shown in Tables 9 and 10, and Appendix Figures A1 and A2 show calculations analogous to Figures 3 and 4 about the implications of the estimates for total costs and facility use in different groups of children.

For the most part, the qualitative patterns are similar though the magnitudes vary. We find, for example, that by 24 months children who are treated have 51.6% higher total health care costs in this sample, compared to 33.1% higher in the full sample. The estimated effects of treatment on the probability of having any facility visits (ER or hospitalization) and on overnight stays in the hospital is very similar in the two samples.

Compared to other treated children, those who received drug treatment have total costs 86.1% higher after 24 months in this sample, compared to 93.7% higher in the original sample. The estimated effects on facility use are very similar to those discussed above.

Turning to red-flag drug treatments, in the original sample we find that total costs are 205.1% higher for children receiving these treatments than for all treated children after 24 months. The corresponding number is 269.5% higher in the restricted sample.[18] We also see increases in facility use in both samples at three months, and in the restricted sample, we also see an increase at 12 months.

There may also be questions about the exclusion of psychiatrists per 1,000 BCBS children, therapists per 1,000 BCBS children, and PCPs providing mental health services per 1,000 BCBS from our main instrument list. As discussed above, we also tried estimating our models including these variables as potential instruments and found that they were not selected

---

[18] In the main sample, we can take the red-flag drug treated child's costs divided by the average treated child's costs (9653/4708). In the diagnosis only sample, the equivalent ratio is (13270/4924).

in the first stage Lasso procedure so that their inclusion in the potential instrument set had no effect on our estimates.[19]

One caveat to our results is that differences in parental attitudes, preferences, or financial considerations (co-pays) could explain at least some of the variation in the probability that any follow-up treatment is obtained in the three months after the initial mental health claim, or in the probability that a child is prescribed psychiatric drugs. In areas where parents believe strongly in the efficacy of drug therapy, we may see PCPs prescribing drugs to a larger share of the children in their practices, for example. By a similar logic, we might expect then to see a corresponding decrease (increase) in the share of mental health treatment being administered by psychiatrists (PCPs). We have estimated models excluding these instruments and their interactions from the potential instrument list provided to Lasso, and we find that our results, presented in Appendix Tables A3 and A4, are robust to this change. We think it less plausible that parental demand could be the main driver of variation in the types of drugs prescribed conditional on any prescription. In our context, that would imply that many parents are demanding that their children with newly emerging mental health conditions be initially treated with non-FDA approved drugs, benzodiazepines, or tricyclic antidepressants.

Overall then, our results are robust to variations in the instrument set and to excluding children with missing diagnostic data.

VI. Discussion and Conclusions

We use health insurance claims data to explore the effects of variation in mental health treatment received in the first three months after an initial claim for mental illness. We focus on

[19] Column 4 of Appendix Table A2 shows which instruments were chosen in that specification.

initial claims for several reasons. First, this is a setting where treatment guidelines may be especially useful since clinicians will typically have little history to inform their decision making. Second, there is persistence in mental health treatment choices over time, suggesting that the initial choice is especially important. Third, there is a lot of treatment that violates treatment guidelines in this setting, even though the guidelines are not especially stringent.

Most of the children in our sample would have received therapy, or an SSRI in combination with therapy, had they been treated according to evidence-based guidelines. While all of the treatments we single out as red-flag treatments may possibly be appropriate in some cases, guidelines suggest that their use should be rare. And yet, nearly half of children (45.2%) treated with drugs receive benzodiazepines, tricyclic anti-depressants, or drugs that are not FDA-approved in the first three months after their initial health claim.

Since treatment is likely to be affected by many unobserved factors that we cannot observe in claims data, we pursue an instrumental variables strategy. Our instruments are based on the idea that conditional on area fixed effects, children living in different places face doctors with different practice styles, and that this is an exogenous source of variation in the treatment that they will receive. Give a large number of possible instruments, and concerns about bias due to weak instruments, we use the post-Lasso method suggested by Bellini et al. (2012) in which Lasso is used to select instruments from the large pool of potential variables.

We find that receiving mental health treatment in the first three months after an initial claim is associated with higher costs for mental health treatment, and that higher initial costs are not offset by lower costs down the road. The costs are higher among those receiving drug therapy, and within this group, the highest costs are among children who receive what we dub red-flag drug treatments. In fact, our estimates imply that higher costs among treated children can be entirely explained by the costs associated with children getting red-flag drug treatments. Children who receive therapy alone in the first three months after an initial mental health claim have lower costs after 24 months than those who do not. Our estimates imply that, by 24 months out, children who initially received a red-flag drug have 205% higher than the mean for all treated children and are 131% more likely to have used an emergency room or experienced a hospitalization. These costs are front-loaded and taper off over time, suggesting that red-flag drug treatments may raise costs by leading to significant side effects.

These results break new ground and contribute to several strands of literature. First, they focus on an important determinant of children's non-cognitive skills: Their mental health. We show that although the initial emergence of mental health conditions offers an potentially important place to intervene in children's life trajectories, large numbers of U.S. children are receiving mental health care that falls outside of accepted guidelines and poses risks to their health. These risks are demonstrated by higher subsequent overall health care costs among these children.

Second, the results provide support for the guidelines themselves, contributing to the on-going controversy about "rules vs. discretion" in medicine. Third, this research suggests that analyses of large scale claims data can provide a useful complement to clinical research studies in identifying best practices. Such analyses are now more feasible than ever before given advances in machine learning and econometrics.

References

Bansak, Cynthia and Stephen Raphael. 2008. “The State Children’s Health Insurance Program and Job Mobility: Identifying Job Lock among Working Parents in Near-Poor Households,” *Industrial and Labor Relations Review*, *61*(4): 564–579.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen. 2012. “Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain.,” *Econometrica*, *80*(6): 2369–2429.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. “High-dimensional methods and inference on structural and treatment effects.” *Journal of Economic Perspectives*, 28: 29–50.

Berndt Ernest, Robert Gibbons, Anton Kolotilin, Anna Taub. 2015. “The heterogeneity of concentrated prescribing behavior: Theory and evidence from antipsychotics,” *Journal of Health Economics* 40:26–39.

Birmaher Boris, David A. Brent, AACAP Work Group on Quality Issues. 2007. “Practice Parameter for the Assessment and Treatment of Children and Adolescents With Depressive Disorders,” *J . Am. Acad. Child Adolesc. Psychiatry* 46(11).

Borghans, Lex, Angela Duckworth, James J. Heckman and Bas Ter Weel. “The Economics and Psychology of Personality Traits,” *J. Human Resources Fall 2008 vol. 43 no. 4 972-1059*

Connolly Sucheta, Gail Bernstein, AACAP Work Group on Quality Issues. 2007. “Practice Parameter for the Assessment and Treatment of Children and Adolescents With Anxiety Disorders.” *J . Am. Acad. Child Adolesc. Psychiatry* 46(2): 267-283.

Cuddy, Emily and Janet Currie. “Treatment of Mental Illness in the U.S. Adolescents Varies Widely Within and Across Areas,” May 2020.

Currie, Janet, William B. MacLeod, and Jessica Van Parys. 2016. “Physician Practice Style and Patient Health Outcomes: The Case of Heart Attacks,” *Journal of Health Economics*, 47: 64–80.

Currie, Janet, William B. MacLeod. 2017. “Diagnosis and Unnecessary Procedure Use: Evidence from C-Section,” *Journal of Labor Economics*, 35: 1–42.

Currie Janet, William B. MacLeod . 2020. “Understanding Doctor Decision Making: The Case of Depression Treatment,” *Econometrica*, 88 #3, May 2020, 847-878.

Currie Janet, Mark Stabile, Phongsack Manivong, Les Roos. 2010. “Child Health and Young Adult Outcomes,” *Journal of Human Resources* 45(3):517-548.

Cutler, David, Jonathan Skinner, Ariel D. Stern, and David Wennberg. 2019. “Physician Beliefs and Patient Preferences: A New Look at Regional Variation In Health Care Spending,” *American Economic Journal: Economic Policy*, 11: 192–221.

Findling Robert L., Ekaterina Stepanova. 2018. “The Workforce Shortage of Child and Adolescent Psychiatrists: Is It Time for a Different Approach?” *Journal of American Academy of Child and Adolescent Psychiatry*, 57(5):300-301.

Finkelstein Amy, Matthew Gentzkow, Heidi Williams. 2015. “Sources of geographic variation in health care: Evidence from patient migration,” *The Quarterly Journal of Economics* 131(4):1681–1726.

Fisher Elliot S., David Wennberg, Thérèse A. *Stukel*, Daniel J. Gottlieb, F. L. Lucas, and Etoile Pinder. 2003a. “The implications of regional variations in medicare spending,” *Annals of Internal Medicine* 138 (4): 273–298.

Fisher Elliot S., David Wennberg, Thérèse A. *Stukel*, Daniel J. Gottlieb, F. L. Lucas, and Etoile Pinder. 2003b. “The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care,” *Ann Intern Med* 138(4):288–298.

Frank Richard G, Richard J. Zeckhauser. 2007. “Custom-made versus ready-to-wear treatments: Behavioral propensities in physicians’ choices,” *Journal of Health Economics* 26(6):1101 – 1127.

Fremont Wanda P., Robert Nastasi, Nancy Newman, Nancy Roizin. 2008. “Comfort Level of Pediatricians and Family Medicine Physicians Diagnosing and Treating Child and Adolescent Psychiatric Disorders,” *Int J Psychiatry Med.* 38(2):153-68.

Ghandour Reem M, Laura J. Sherman, Catherine J. Vladutiu, Mir M. Ali, Sean E. Lynch, Rebecca H. Bitsko, Stephen J. Blumberg. 2019. “Prevalence and treatment of depression, anxiety, and conduct problems in U.S. children,” *The Journal of Pediatrics*, 206:256–267.

Health Resources and Services Administration/National Center for Health Workforce Analysis; Substance Abuse and Mental Health Services Administration/Office of Policy, Planning, and Innovation. 2015. National Projections of Supply and Demand for Behavioral Health Practitioners: 2013-2025. Rockville, Maryland.

Heckman, James, J. 2000. “Policies to Foster Human Capital,” *Research in Economics*, 54 (1): 3-56.

Heckman, James J., Jora Stixrud, and Sergio Urzua. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics* 2006 24:3, 411-482.

Heneghan Amy, Andrew S. Garner, Amy Storfer-Isser, Karl Kortepeter, Ruth E.K, Stein, Sarah M. Horwitz. 2008. “Pediatricians' Role in Providing Mental Health Care for Children and Adolescents: Do Pediatricians and Child and Adolescent Psychiatrists Agree?,” *Journal of Developmental and Behavioral Pediatrics*, 29(4):262-269.

Hilt, Robert J, Monica Chaudhari, Janice F Bell, Christine Wolf, Kent Koprowicz, and Bryan H King. 2014. *Journal of Child and Adolescent Psychopharmacology*. March: 83-89.

Kessler Daniel P, Mark B. McClellan. 2000. "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics* 115(2):577-615.

McBain, Ryan K. Aaron Kofner, Bradley D. Stein, Jonathan H. Cantor, William B. Vogt, Hao Yu. 2019. "Growth and Distribution of Child Psychiatrists in the United States: 2007-2016," *Pediatrics*, 144 (6).

McClellan J, S Stock, AACAP Work Group on Quality Issues. 2013. "Practice Parameter for the Assessment and Treatment of Children and Adolescents With Schizophrenia," *J . Am. Acad. Child Adolesc. Psychiatry* 52(9):976-990.

McQuaid J, Members of the Guideline Development Panel for the Treatment of Depressive Disorders. 2019. "Clinical Practice Guideline for the Treatment of Depression Across Three Age Cohorts," *American Psychological Association Guideline Development Panel for the Treatment of Depressive Disorde*rs.

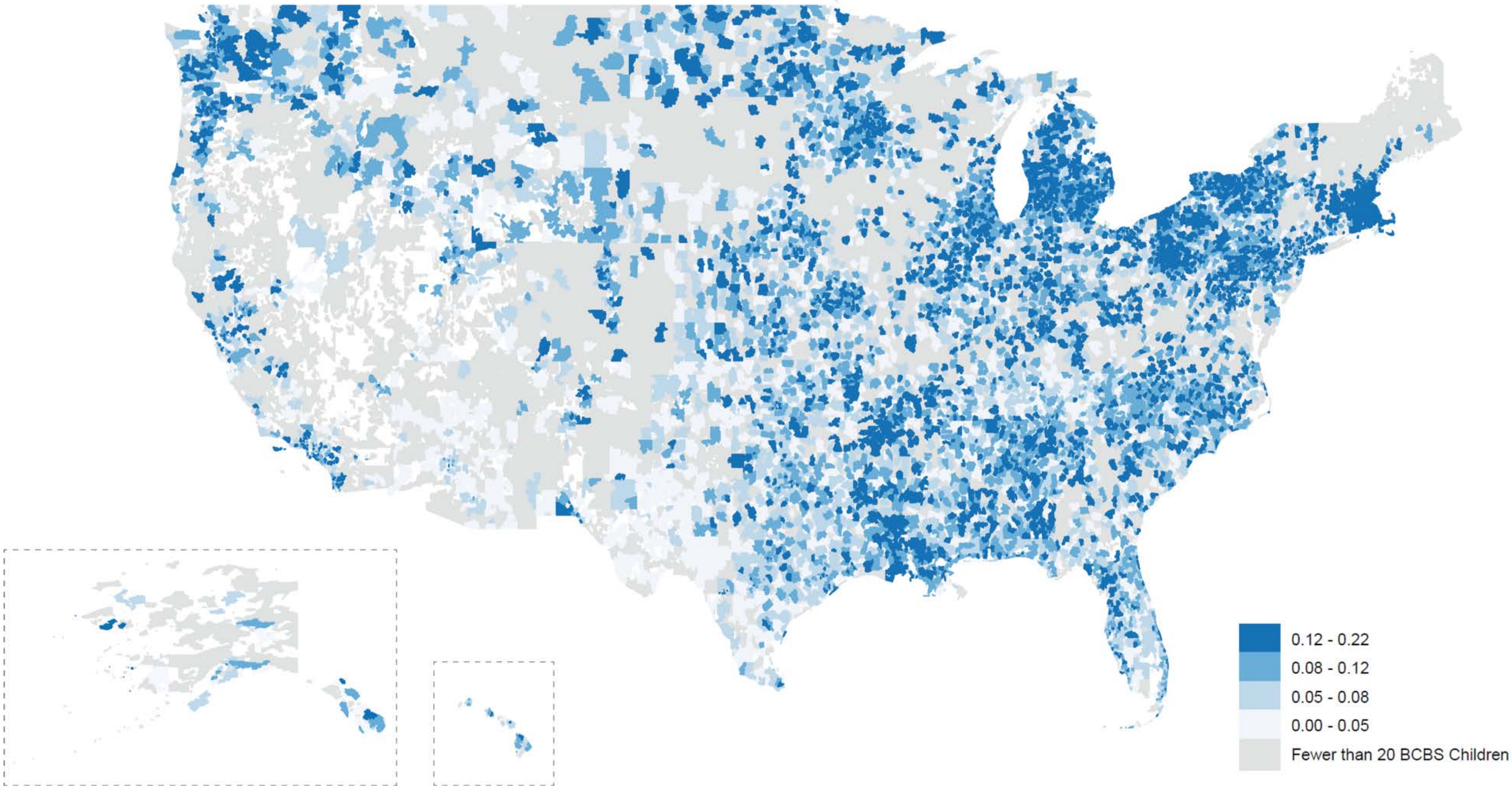
Montiel Olea, Jose L. and Carolin E. Pflueger. 2013. "A Robust Test for Weak Instruments," *Journal of Business and Economic Statistics* 31, 358-369.

Sands Elizabeth G, Duncan S. Gilchrist. 2016. "Something to Talk About: Social Spillovers in Movie Consumption," *Journal of Political Economy* 124 (5) :1339-1382.

Smith James P, Gillian Smith. 2010. "Long-term economic costs of psychological problems during childhood," *Social Science and Medicine*, 71(1): 110-115.

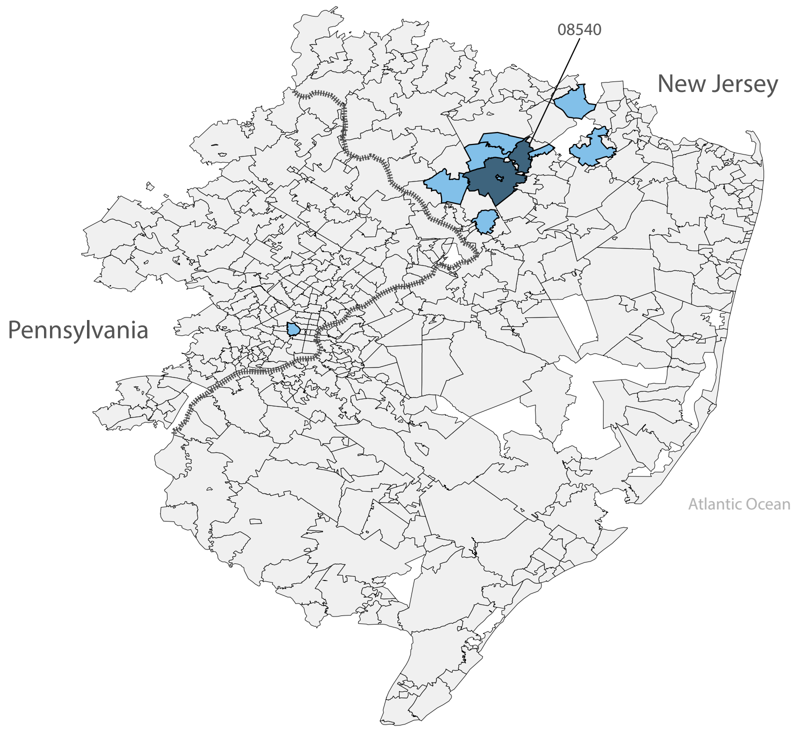
Thomas Christopher R, Cordula E. Holzer. 2006 "The Continuing Shortage of Child and Adolescent Psychiatrists," *Journal of the American Academy of Child & Adolescent Psychiatry* 45(9):1023-1031.

Figure 1
Prevalence of Childhood Mental Illness in BCBS Coverage Areas



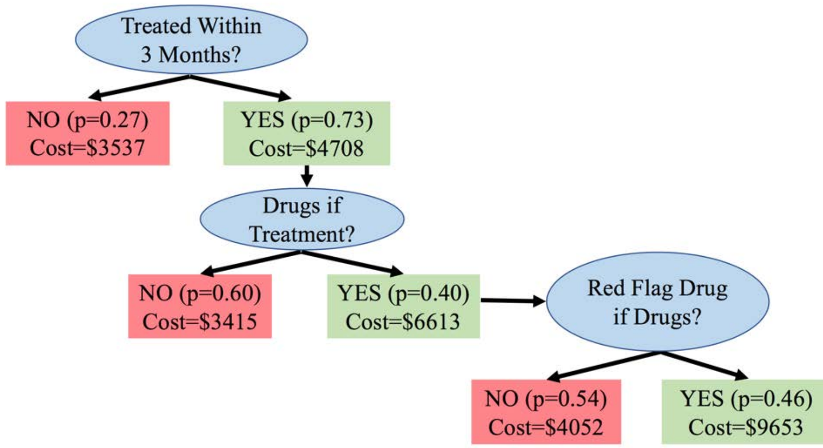
Notes: This figure shows the prevalence of mental illness among children (10-17) with BCBS coverage between 2012-2018. Each polygon represents a Census zip code tabulation area (ZCTA), where the blue-shaded polygons are ZCTAs with at least twenty children with BCBS coverage.

Figure 2
Illustrating the Definition of the Market Area for A Specific Zip Code, Princeton NJ, 08540



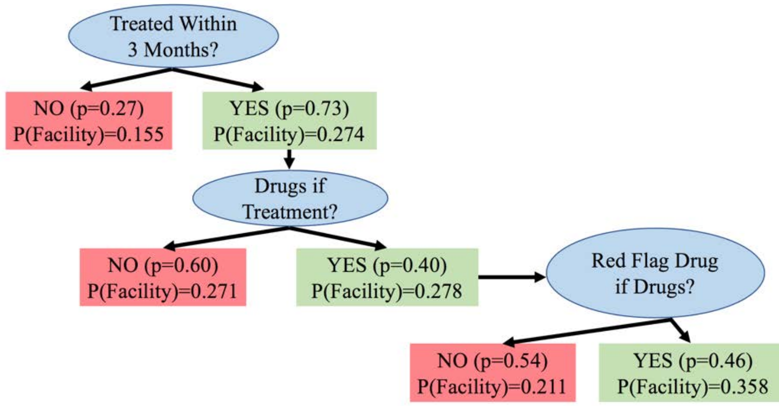
Notes: BCBS children residing in Princeton NJ saw providers for mental health services in their own zip code (08540) and also in the zip codes that are coded in lighter blue. Hence, these zip codes together are considered the relevant market area for children in Princeton NJ. Markets for all other zip codes are defined analogously, so that providers in each zip code may form part of the market for children residing in multiple zip codes. We include up to 10 ZCTAs in each market ranked in terms of the number of children who travel there (as in this example); in markets where children traveled to fewer than 10 ZCTAs, we use only those ZCTAs.

Figure 3
Implications for Total Cost at 24 Months



Notes: This figure traces out the implications of each initial treatment choice on total costs at 24 months. The share of children in each branch appears in parentheses. Cost estimates are based on post-Lasso estimates in column 3 of Table 5.

Figure 4
Implications for Facility Use at 24 Months



Notes: This figure traces out the implications of each initial treatment choice on facility use—ER visit or hospitalization—at 24 months. The share of children in each branch appears in parentheses. Cost estimates are based on post-Lasso estimates in column 3 of Table 6.

Table 1
Descriptive Statistics

	(1) All BCBS	(2) All MH BCBS	(3) Sample	(4) Treated	(5) Therapy Only	(6) Therapy and Drugs	(7) Drugs	(8) Red-flag Drugs
A: Child Characteristics								
Number of children	2,201,566	202,066	97,306	70,886	42,236	5,362	28,650	13,102
% of sample	-	-	-	*73%*	*43%*	*6%*	*29%*	*13%*
Female	0.490	0.528	0.505	0.516	0.503	0.584	0.536	0.557
Age 1st appearance in sample	7.581	8.395	8.728	8.756	8.686	8.969	8.860	8.879
Hospitalized, any reason	0.012	0.030	0.036	0.039	0.027	0.100	0.056	0.069
ER, any reason	0.099	0.158	0.176	0.187	0.182	0.298	0.193	0.196
Average monthly costs ($2018)	$157	$302	$346	$358	$291	$514	$456	$596
Neurodevelopmental condition	0.133	0.318	0.348	0.331	0.369	0.464	0.275	0.255
Neuro condition is ADHD	0.100	0.233	0.265	0.255	0.288	0.346	0.206	0.183
Age, 1st mental illness episode	-	12.023	11.492	11.501	11.409	11.866	11.637	11.690
Hospitalized, 1st mental illness	-	0.012	0.010	0.009	0.004	0.055	0.016	0.015
ER, 1st mental illness	-	0.033	0.031	0.027	0.027	0.053	0.026	0.023
1st episode is an evaluation	-	0.421	0.446	0.473	0.722	0.452	0.106	0.078
B: Prevalance of ICD10 Diagnoses in the Sample								
Substance use (F10-F19)		0.013	0.008	0.003	0.003	0.014	0.004	0.004
Non-mood psychotic disorders (F20-F29)		0.013	0.012	0.010	0.007	0.053	0.015	0.012
Mood disorders (F30-F39)		0.161	0.136	0.144	0.141	0.501	0.148	0.104
Anxiety / stress disorders (F40-F49)		0.573	0.547	0.553	0.736	0.828	0.284	0.227
Behavioral syndromes (F50-F59)		0.027	0.025	0.014	0.013	0.046	0.015	0.016
Personality and behavioral disorders (F60-F69)		0.020	0.020	0.014	0.015	0.045	0.013	0.011
Other disorders originating in childhood		0.060	0.053	0.021	0.024	0.045	0.017	0.014
No MH diagnosis		0.267	0.308	0.363	0.165	0.050	0.654	0.724

Table 1. Data is from the BCBS Axis data base of insurance claims for 2012 to 2018. It covers children who have a valid master member ID, pharmacy coverage, valid geographic information, and who were observed both before age 11 and for at least one year between the ages of 10 and 18. Children in column 2 had at least one claim related to mental illness. Column 3 includes all children who can be followed for at least 3 months; column 4 includes all children who can be followed for at least 24 months: this is our population of interest in the analysis. Column 4 includes those who received any follow up treatment in the 3 months following an initial claim. Children in column 5 received only therapy (no drugs) in the three months following the initial claim. Column 6 includes children who received both therapy and drugs in the three months following the initial claim. Column 7 includes children who received only drug treatment (no therapy) in the three months following the initial claim. Column 8 includes children who received benzodiazepines, tricyclic anti-depressants, or a non-FDA approved drug in the 3 months following the first claim. The variables "Hospitalized, any reason," "ER, any reason," and "Average Monthly Costs" are computed taking the average over all of the months that a child appears in the data, and then both facility measures are annualized. Dashes indicate that the category is not applicable.

Table 2
Small-Area Variation in Treatment, Provider Supply, and Practice Style

Percentiles of the Area-Level Distribution:	**10th**	**25th**	**50th**	**75th**	**90th**	**Within-ZCTA Variation**	**Between-ZCTA Variation**
A. Treatment (Source is BCBS)							
Child treated within 3 months	0.20	0.56	0.77	1.00	1.00	37%	63%
Therapy Only (if any treatment)	0.00	0.29	0.58	1.00	1.00	51%	49%
Drugs & Therapy (if any treatment)	0.00	0.00	0.00	0.06	0.25	57%	43%
Drugs Only (if any treatment)	0.00	0.00	0.33	0.57	1.00	51%	49%
Red-flag Drug (if any drug treatment)	0.00	0.00	0.50	1.00	1.00	58%	42%
Benzodiazepines	0.00	0.00	0.00	0.33	1.00	60%	40%
Tricyclic antidepressants	0.00	0.00	0.00	0.00	0.40	61%	39%
Not FDA approved	0.00	0.00	0.00	0.50	1.00	60%	40%
B. Provider Supply (Source is BCBS)							
Psychiatrists per 1,000 BCBS children 10-17	3.49	4.91	7.45	11.97	17.90	6%	94%
Therapists per 1,000 BCBS children 10-17	14.74	19.60	26.70	37.71	51.59	9%	91%
PCPs providing MH treatment per 1,000 BCBS children 10-17	17.80	26.60	39.77	60.07	89.23	9%	91%
C. Practice Style (Source is BCBS or IQVIA)							
BCBS: Share MH providers who are psychiatrists	0.07	0.09	0.11	0.15	0.18	8%	92%
BCBS: Share MH providers who are PCPs	0.12	0.15	0.18	0.22	0.24	11%	89%
BCBS: Share PCP's patients who receive a MH drug	0.03	0.05	0.07	0.09	0.11	35%	65%
BCBS: Share provider prescriptions for red-flag drugs	0.43	0.47	0.51	0.54	0.58	52%	48%
IQVIA: Share new prescriptions by psychiatrists	0.03	0.05	0.07	0.09	0.12	35%	65%
IQVIA: Share new prescriptions for red-flag drugs	0.28	0.32	0.36	0.40	0.44	50%	50%

Notes: This table is calculated by computing small area-level rates and then calculating percentiles of the distributions of those rates. Small areas are defined using information about where children in a particular zip code actually go to receive mental health care. Each row represents a separate distribution of ZCTAs. For example, places at the 90th percentile in terms of psychiatrists per capita could be at the 10th percentile in terms of the fraction of PCPs treating mental health. BCBS indicates that the variable is calculated using our main sample. IQVIA indicates that the data was calculated using that data base. The variables "within" and "between" ZCTA variation represent the decomposition of total variation in each treatment, supply, and style measure.

Table 3
Regressions with 6 Main Interactions and Quadratic Terms

	(1)	(2)	(3)	(4)
	Treated	**Drugs**	**Red-flag Drugs**	**Red-flag Drugs**
Conditional on:	-	*Treatment*		*Drugs*
Child female	0.023**	0.025**	0.029**	0.043**
	(0.002)	(0.002)	(0.003)	(0.006)
Child age (Years)	0.004**	0.018**	0.014**	0.015**
	(0.001)	(0.001)	(0.001)	(0.003)
1st claim is hospitalization	-0.057**	0.116**	0.089**	0.095**
	(0.017)	(0.022)	(0.020)	(0.024)
1st claim is ER visit	-0.084**	-0.013	-0.018*	-0.028
	(0.026)	(0.012)	(0.009)	(0.018)
1st claim is evaluation	0.062**	-0.428**	-0.197**	-0.034**
	(0.008)	(0.008)	(0.004)	(0.011)
Hospitalized last 6 months	0.042**	0.097**	0.194**	0.226**
	(0.014)	(0.013)	(0.016)	(0.023)
ICD10: F10-F19	-0.144**	0.094**	0.036	0.018
	(0.017)	(0.029)	(0.024)	(0.041)
ICD10: F20-F29	-0.005	0.176**	0.083**	0.040
	(0.014)	(0.019)	(0.015)	(0.025)
ICD10: F30-F39	0.253**	0.208**	0.030**	-0.106**
	(0.005)	(0.008)	(0.005)	(0.011)
ICD10: F40-F48	0.336**	0.108**	0.013*	-0.046**
	(0.006)	(0.009)	(0.007)	(0.014)
ICD10: F50-F59	-0.082**	0.111**	0.072**	0.094**
	(0.010)	(0.015)	(0.013)	(0.023)
ICD10: F60-F69	0.015	0.090**	0.026*	0.013
	(0.010)	(0.016)	(0.012)	(0.024)
Other MH diagnosis	-0.126**	0.087**	0.032**	-0.005
	(0.008)	(0.012)	(0.009)	(0.024)
No MH diagnosis	0.503**	0.468**	0.235**	0.081**
	(0.009)	(0.012)	(0.009)	(0.016)
Neurodevelopmental condition	-0.042**	-0.074**	-0.046**	0.003
	(0.004)	(0.004)	(0.003)	(0.007)
BCBS: Share MH providers	-0.386*	0.096	0.071	0.092
who are psychiatrists	(0.181)	(0.178)	(0.150)	(0.310)
share squared	1.054+	-0.642	-0.360	-0.302
	(0.593)	(0.547)	(0.458)	(1.005)
BCBS: Share MH providers	-0.272	0.742**	0.197	-0.140
who are PCPs	(0.265)	(0.249)	(0.209)	(0.444)
share squared	0.468	-1.410*	-0.338	0.417

	(0.672)	(0.680)	(0.576)	(1.148)
BCBS: Share PCP's patients	0.201	0.297	-0.024	-0.345
who receive an MH drug	(0.134)	(0.182)	(0.154)	(0.331)
share squared	-0.435	-0.579	-0.301	-0.092
	(0.522)	(0.636)	(0.517)	(1.063)
BCBS: Share provider prescriptions	0.159	0.159	-0.003	-0.359
for red flag drugs	(0.175)	(0.180)	(0.146)	(0.347)
share squared	-0.108	-0.265	0.145	0.777*
	(0.175)	(0.181)	(0.149)	(0.354)
IQVIA: Share new prescriptions	0.146	-0.289**	-0.005	0.333+
by psychiatrists	(0.135)	(0.093)	(0.088)	(0.195)
share squared	-0.500	0.556+	-0.140	-1.029+
	(0.376)	(0.295)	(0.231)	(0.561)
IQVIA: Share new prescriptions	-0.350+	-0.113	0.113	0.393
for red flag drugs	(0.199)	(0.211)	(0.184)	(0.380)
share squared	0.469+	0.221	-0.042	-0.338
	(0.256)	(0.280)	(0.250)	(0.500)
Constant	0.257**	0.110+	-0.067	0.118
	(0.057)	(0.056)	(0.048)	(0.104)
Number of observations	97306	70886	70886	28650
R-squared	0.170	0.491	0.215	0.062

Notes: This table presents first-stage results, where a child's treatment status is instrumented fo by the six original instruments and their squared terms. Each column is from a single regressioı model, where the sample is conditional on treatment in columns 2 and 3 and on drug treatment in column 4. All models include month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are in parentheses.

+ Significant at the 10 percent level
* Significant at the 5 percent level
** Significant at the 1 percent level

Table 4
Test Statistics for Alternative First-Stage Regressions

	(1)	(2)	(3)	(4)	(5)
	OLS - No Instruments	**Main Instruments + Quadratic**	**Full Interactions**	**Post-Lasso, 3 Instruments**	**Post-Lasso, All Instruments**
Any Follow Up Treatment in 3 Months After 1st Claim					
N	97306	97306	97306	97306	97306
R-squared	0.170	0.170	0.195	0.188	0.190
F_{eff}	-	2	23	383	207
Drug Treatment in 1st 3 Months After 1st Claim *Conditional on Treatment*					
N	70886	70886	70886	70886	70886
R-squared	0.491	0.491	0.515	0.509	0.512
F_{eff}	-	6	27	376	189
Red-flag Drug Treatment in 1st 3 Months After 1st Claim *Conditional on Treatment*					
N	70886	70886	70886	70886	70886
R-squared	0.215	0.215	0.223	0.217	0.220
F_{eff}	-	6	6	75	57
Red-flag Drug Treatment in 1st 3 Months After 1st Claim *Conditional on Drug Treatment*					
N	28650	28650	28650	28650	28650
R-squared	0.058	0.062	0.068	0.064	0.064
F_{eff}	-	10	3	55	42

Notes: This table presents post-estimation results from the first-stage model, where the instrument choice set varies across specifications. F_{eff} refers to the effective F statistic of Montiel Olea and Plueger (2013). The main instruments include the shares of psychiatrists and PCPs providing mental health treatment to children, the average share of patients in PCP: caseloads receiving psychiatric drugs, the share of red-flag drugs in physician's caseloads in BCBS and IQVIA, and the share of psychiatrists among initial psychiatric prescribers. In columns 1-3, the instruments are pre-selected: column 1 uses no instruments (N=0); column 2 uses the main instruments as well as their squared terms (N=12); and column 3 includes all potential instruments, i.e., a second-degree polynomial in the main instruments, as well as interactions with age, severity (indicators for first event in ER, first event in hospital, first event is evaluation, and hospitalization in previous 6 months), and diagnosis (indicators for anxiety/depression, adjustment disorders, no diagnosis, neurodevelopmental, and other MH diagnosis) (N=87). In columns 4 and 5, Lasso selects the instruments. Column 4 is a post-Lasso model with a three-instrument constraint (N=3). Column 5 is a post-Lasso model wher a refined data-driven penalty is used for the Lasso (6≤N≤8). See Appendix Table A2 for a ful list of the instruments selected in each model. All models include patient controls (age, femal severity, and diagnosis) as well as month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are in parentheses.

Table 5
Effects of Treatment on Costs

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Total Costs)			**Log(Total MH Costs)**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
A. Any Follow Up Treatment in 3 Months After 1st Claim (N=97306)						
OLS						
Treatment	0.366**	0.339**	0.280**	0.311**	0.519**	0.555**
SE	(0.012)	(0.011)	(0.011)	(0.016)	(0.017)	(0.018)
R2	0.146	0.150	0.150	0.510	0.267	0.427
Post-Lasso 2SLS						
Treatment	0.253**	0.295**	0.286**	0.617**	0.786**	0.869**
SE	(0.063)	(0.062)	(0.053)	-0.077	(0.088)	(0.081)
Mean Dependent Variable	6.785	7.761	8.387	5.205	5.669	6.010
B. Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=70886)						
OLS						
Treatment	0.183**	0.246**	0.259**	-1.027**	-0.835**	-0.695**
SE	(0.017)	(0.015)	(0.014)	(0.025)	-0.024	(0.024)
R2	0.143	0.149	0.155	0.626	0.557	0.504
Post-Lasso 2SLS						
Treatment	1.128**	0.874**	0.661**	0.969**	1.216**	1.203**
SE	(0.065)	(0.058)	(0.051)	(0.087)	(0.089)	(0.083)
Mean Dependent Variable	6.871	7.844	8.456	5.212	5.728	6.081
C. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=70886)						
OLS						
Treatment	0.343**	0.408**	0.388**	-0.665**	-0.542**	-0.430**
SE	(0.029)	(0.023)	(0.020)	(0.034)	(0.037)	(0.037)
R2	0.148	0.157	0.162	0.610	0.547	0.497
Post-Lasso 2SLS						
Treatment	2.607**	2.199**	1.768**	1.990**	2.623**	2.574**
SE	(0.235)	(0.203)	(0.182)	(0.263)	(0.299)	(0.300)
Mean Dependent Variable	6.871	7.844	8.456	5.212	5.728	6.081
D. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=28650)						
OLS						
Treatment	0.347**	0.384**	0.348**	-0.104**	-0.061*	-0.011
SE	(0.031)	(0.025)	(0.022)	(0.021)	(0.028)	(0.033)
R2	0.185	0.199	0.201	0.687	0.608	0.548
Post-Lasso 2SLS						
Treatment	0.530	0.803**	0.868**	-0.786**	-0.537*	-0.294
SE	(0.329)	(0.265)	(0.242)	(0.248)	(0.274)	(0.321)
Mean Dependent Variable	6.781	7.848	8.495	3.736	4.334	4.783

Notes: This table presents results from OLS and post-Lasso 2SLS models of mental health treatment on log total health costs (columns 1-3) and log mental health costs (columns 4-6). The instruments in the post-Lasso 2SLS models are chosen using a refined data-driven penalty and are displayed in Appendix Table A2. Panel A includes 6 instruments; Panel B includes 7 instruments; Panel C includes 8 instruments; and Panel D includes 4 instruments. Post-estimation results from the first-stage of each model are included in Table 4. See Appendix Table A2 for a full list of the instrument selected in each model. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZI 3 level, are in parentheses.

+ Significant at the 10 percent level
* Significant at the 5 percent level
** Significant at the 1 percent level

Table 6
Effects of Treatment on Facility Use

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: Visited Facility?			**Number Nights in Hospital**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
A. Any Follow Up Treatment in 3 Months After 1st Claim (N=97306)						
OLS						
Treatment	0.014**	0.022**	0.027**	0.199**	0.337**	0.556**
SE	(0.002)	(0.003)	(0.003)	(0.056)	(0.071)	(0.093)
R2	0.366	0.226	0.178	0.029	0.035	0.025
Post-Lasso 2SLS						
Treatment	0.062**	0.101**	0.119**	1.117+	1.648*	2.582**
SE	(0.015)	(0.020)	(0.023)	(0.571)	(0.682)	(0.972)
Mean Dependent Variable	0.100	0.169	0.242	0.289	0.541	0.915
B. Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=70886)						
OLS						
Treatment	0.069**	0.083**	0.088**	0.362**	0.597**	1.019**
SE	(0.003)	(0.005)	(0.005)	(0.070)	(0.117)	(0.289)
R2	0.328	0.209	0.169	0.028	0.039	0.026
Post-Lasso 2SLS						
Treatment	0.085**	0.038*	0.007	-0.531	-0.403	-1.559+
SE	(0.013)	(0.017)	(0.022)	(0.772)	(0.738)	(0.806)
Mean Dependent Variable	0.101	0.173	0.247	0.333	0.607	1.033
C. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=70886)						
OLS						
Treatment	0.080**	0.085**	0.084**	0.371**	0.610**	0.948**
SE	(0.004)	(0.005)	(0.006)	(0.098)	(0.097)	(0.295)
R2	0.330	0.209	0.169	0.028	0.039	0.026
Post-Lasso 2SLS						
Treatment	0.201**	0.061	0.003	-1.063	-0.854	-1.685
SE	(0.034)	(0.043)	(0.054)	(1.721)	(1.645)	(1.864)
Mean Dependent Variable	0.101	0.173	0.247	0.333	0.607	1.033
D. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=28650)						
OLS						
Treatment	0.062**	0.062**	0.058**	0.273*	0.444**	0.669*
SE	(0.004)	(0.005)	(0.006)	(0.132)	(0.124)	(0.293)
R2	0.265	0.190	0.164	0.040	0.056	0.040
Post-Lasso 2SLS						
Treatment	0.109*	0.121*	0.147*	-1.930	-1.782	-1.603
SE	(0.049)	(0.061)	(0.065)	(2.668)	(3.212)	(4.224)
Mean Dependent Variable	0.149	0.229	0.304	0.665	1.076	1.638

Notes: This table presents results from OLS and post-Lasso 2SLS models of mental health treatment on facility use (columns 1-3) and the number of overnight stays in the hospital (columns 4-6). The instruments in the post-Lasso 2SLS models are chosen using a refined data-driven penalty and are displayed in Appendix Table A2. Panel A includes 6 instruments; Panel B includes 7 instruments; Panel C includes 8 instruments; and Panel D includes 4 instruments. Post-estimation results from the first-stage of each model are included in Table 4. See Appendix Table A2 for a full list of the instruments selected in each model. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are in parentheses.

+ Significant at the 10 percent level

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 7
Treatment Effects on Costs using Three-Instrument Constraint

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Total Costs)			**Log(Total MH Costs)**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
A. Any Follow Up Treatment in 3 Months After 1st Claim (N=97306)						
Treatment	0.141*	0.200**	0.209**	0.599**	0.729**	0.825**
SE	(0.071)	(0.068)	(0.059)	(0.078)	(0.09)	(0.083)
Mean Dependent Variable	6.785	7.761	8.387	5.205	5.669	6.010
B. Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=70886)						
Treatment	0.878**	0.660**	0.505**	0.464**	0.731**	0.761**
SE	(0.063)	(0.062)	(0.057)	(0.078)	(0.087)	(0.082)
Mean Dependent Variable	6.871	7.844	8.456	5.212	5.728	6.081
C. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=70886)						
Treatment	2.871**	2.254**	1.826**	1.224**	2.067**	2.263**
SE	(0.276)	(0.26)	(0.239)	(0.279)	(0.32)	(0.314)
Mean Dependent Variable	6.871	7.844	8.456	5.212	5.728	6.081
D. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=28650)						
Treatment	0.541	0.804**	0.869**	-0.781**	-0.563*	-0.295
SE	(0.332)	(0.268)	(0.243)	(0.25)	(0.274)	(0.316)
Mean Dependent Variable	6.781	7.848	8.495	3.736	4.334	4.783

Notes: This table presents results from post-Lasso 2SLS models of mental health treatment on log tota health costs (columns 1-3) and log mental health costs (columns 4-6), where the Lasso is constrained t select at most three instruments. These instruments are shown in Appendix Table A2. Post-estimation results from the first-stage of each model are included in Table 4. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding to the dat of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are in parenthes

+ Significant at the 10 percent level
* Significant at the 5 percent level
** Significant at the 1 percent level

Table 8
Treatment Effects on Facility Use using Three-Instrument Constraint

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: Visited Facility?			**Number Nights in Hospital**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
Any Follow Up Treatment in 3 Months After 1st Claim (N=97306)						
Treatment	0.060**	0.100**	0.124**	1.101+	1.670*	2.525*
SE	(0.015)	(0.020)	(0.024)	(0.575)	(0.684)	(1.037)
Mean Dependent Variable	0.100	0.169	0.242	0.289	0.541	0.915
Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=70886)						
Treatment	0.039**	-0.011	-0.046*	-0.042	-0.131	-1.635*
SE	(0.013)	(0.018)	(0.023)	(0.327)	(0.389)	(0.759)
Mean Dependent Variable	0.101	0.173	0.247	0.333	0.607	1.033
Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=70886)						
Treatment	0.094*	-0.109+	-0.219**	0.266	-0.022	-3.983+
SE	(0.044)	(0.064)	(0.076)	(0.872)	(1.059)	(2.250)
Mean Dependent Variable	0.101	0.173	0.247	0.333	0.607	1.033
Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=28650)						
Treatment	0.110*	0.121+	0.149*	-2.384	-2.416	-2.614
SE	(0.051)	(0.062)	(0.066)	(3.042)	(3.617)	(4.618)
Mean Dependent Variable	0.149	0.229	0.304	0.665	1.076	1.638

Notes: This table presents results from post-Lasso 2SLS models of mental health treatment on of menta health treatment on facility use (columns 1-3) and the number of overnight stays in the hospital (colum 4-6), where the Lasso is constrained to select at most three instruments. These instruments are shown ir Appendix Table A2. Post-estimation results from the first-stage of each model are included in Table 4. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are in parentheses.

+ Significant at the 10 percent level
* Significant at the 5 percent level
** Significant at the 1 percent level

Table 9
Effects of Treatment on Costs Among Diagnosed

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Total Costs)			**Log(Total MH Costs)**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
A. Any Follow Up Treatment in 3 Months After 1st Claim (N=67336)						
Treatment	0.553**	0.485**	0.416**	1.628**	1.779**	1.816**
SE	(0.078)	(0.075)	(0.065)	(0.092)	(0.103)	(0.096)
Mean Dependent Variable	6.863	7.795	8.405	5.933	6.393	6.709
B. Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=45157)						
Treatment	0.998**	0.802**	0.621**	0.421**	0.703**	0.777**
SE	(0.057)	(0.052)	(0.047)	(0.066)	(0.070)	(0.067)
Mean Dependent Variable	7.035	7.941	8.522	6.288	6.810	7.128
C. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=45157)						
Treatment	3.231**	2.628**	2.039**	1.856**	2.639**	2.797**
SE	(0.207)	(0.178)	(0.153)	(0.214)	(0.233)	(0.226)
Mean Dependent Variable	7.035	7.941	8.522	6.288	6.810	7.128
D. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=9911)						
Treatment	1.322**	1.738**	1.706**	0.858+	0.949+	0.842
SE	(0.425)	(0.460)	(0.462)	(0.472)	(0.562)	(0.579)
Mean Dependent Variable	7.462	8.387	8.950	6.295	7.024	7.456

Notes: This table presents results from post-Lasso 2SLS models of mental health treatment on log total health costs (columns 1-3) and log mental health costs (columns 4-6), where children who had not received an official mental health diagnosis are excluded from the sample. The instruments in each mod are chosen using a refined data-driven penalty and are displayed in Appendix Table A2. Panel A include 5 instruments; Panel B includes 7 instruments; Panel C includes 7 instruments; and Panel D includes 1 instrument. See Appendix Table A2 for a full list of the instruments selected in each model. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-level, are in parentheses.

+ Significant at the 10 percent level
* Significant at the 5 percent level
** Significant at the 1 percent level

Table 10

Effects of Treatment on Facility Use Among Diagnosed

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: Visited Facility?			**Number Nights in Hospital**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
Any Follow Up Treatment in 3 Months After 1st Claim (N=67336)						
Treatment	0.078**	0.115**	0.127**	1.539**	2.119*	2.964*
SE	(0.019)	(0.024)	(0.027)	(0.763)	(0.899)	(1.283)
Mean Dependent Variable	0.104	0.173	0.245	0.351	0.658	1.109
Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=45157)						
Treatment	0.082**	0.043**	0.012	-0.466	-0.404	-1.577*
SE	(0.012)	(0.015)	(0.019)	(0.647)	(0.611)	(0.727)
Mean Dependent Variable	0.104	0.177	0.251	0.426	0.778	1.327
Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=45157)						
Treatment	0.276**	0.161**	0.088	-1.733	-1.199	-4.211+
SE	(0.035)	(0.045)	(0.056)	(2.363)	(2.225)	(2.365)
Mean Dependent Variable	0.104	0.177	0.251	0.426	0.778	1.327
Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=9911)						
Treatment	0.114	0.397**	0.502**	1.799	5.661*	8.382
SE	(0.117)	(0.147)	(0.150)	(1.729)	(2.704)	(8.282)
Mean Dependent Variable	0.216	0.307	0.386	1.512	2.401	3.602

Notes: This table presents results from post-Lasso 2SLS models of mental health treatment on facility use (columns 1-3) and the number of overnight stays in the hospital (columns 4-6). The instruments in each model are chosen using a refined data-driven penalty and are displayed in Appendix Table A2. Panel A includes 5 instruments; Panel B includes 7 instruments; Panel C includes 7 instruments; and Panel D includes 1 instrument. See Appendix Table A2 for a full list of the instruments selected in eacl model. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are in parentheses.

+ Significant at the 10 percent level

* Significant at the 5 percent level

** Significant at the 1 percent level

Appendix Table A1
Area Characteristics of BCBS Children

	(1)	(2)	(3)	(4)
	National Average	**All BCBS**	**All MH BCBS**	**Sample**
Total Population (000s)	31.5	27.1	26.8	26.5
Total Population: Female	51%	51%	51%	51%
Total population: 0-17	23%	23%	23%	23%
Total population: 10-17	10%	11%	11%	11%
White Alone	73%	80%	82%	82%
Black or African American Alone	13%	9%	8%	8%
Other Race	15%	11%	10%	10%
Hispanic or Latino	19%	12%	10%	10%
Family Households	67%	69%	68%	68%
Married (Not Including Separated)	48%	53%	53%	53%
Less than High School	13%	10%	9%	9%
High School Diploma	56%	56%	55%	55%
Bachelor's Degree or Better	31%	34%	36%	36%
Labor Force Participation Rate	63%	64%	65%	65%
Employment Rate	94%	95%	95%	95%
Average Household Income (In $2018)	$84,303	$90,742	$94,584	$95,209
Gini Index	0.44	0.43	0.43	0.43
Owner Occupied Housing Units	64%	71%	71%	72%
Families Below Poverty Level	11%	8%	8%	8%
Adult Poverty Rate	14%	11%	10%	10%
Average Commute to Work (In Min.)	26.7	25.6	25.5	25.6
Uninsured	9%	8%	7%	7%
Public Health Coverage	35%	32%	31%	31%
Private Health Insurance	67%	73%	75%	75%
Children Living with Single Parents	31%	26%	26%	26%
Number of ZCTAs	33120	27906	16487	13848

Note: This table presents estimates from the American Community Survey 2018 5-Year files for ZCTA geographies. Each row represents a weighted average of each ZCTA measure, where the weights correspond to the relevant population in each ZCTA. Column 1 includes all ZCTAs in the country, and the weights correspond to the total reported population in the ZCTA. Column 2 includes all ZCTAs with at least one child aged 10-17 with BCBS coverage, and the weights correspond to the number of BCBS-insured children 10-17. Column 3 includes all ZCTAs with at least one child aged 10-17 with BCBS coverage and a mental health condition, and the weights correspond to the number of BCBS-insured children 10-17 with mental illness. Column 4 includes all ZCTAs with at least one child aged 10-17 with BCBS coverage, a mental health condition, and at least 2 years of continuous coverage after their initial mental health event--this column corresponds to our main analysis sample. The three health insurance variables ("Uninsured," "Public Health Coverage," and "Private Health Coverage") do not add to 100 given overlapping coverage.

Appendix Table A2

Instruments Selected by Lasso

(1)	(2)	(3)	(4)	(5)	(6)
Main Analysis	**Alternative Specifications**				
Lasso, Data-driven Penalty	**Lasso, Top 3 Instruments**	**Lasso, Excluding Children Without Diagnoses**	**Lasso, Including Physician Supply**	**Lasso, Excluding PCP Prescribing Instrument**	**Lasso, Excluding All Potential Demand Instruments**
A. Probability of Any Follow Up Treatment in 3 Months After 1st Claim					
first_eval#bad_u18n	F43#bcbs_bad	first_eval#bad_u18n	first_eval#bad_u18n	first_eval#bad_u18n	first_eval#bad_u18n
dax#bad_u18n	F43#s1_by_psych	dax#bad_u18n	dax#bad_u18n	dax#bad_u18n	dax#bad_u18n
F43#share_mh	dax#bad_u18n	F43#share_mh	F43#share_mh	F43#bcbs_bad	F43#bcbs_bad
F43#bcbs_bad		F43#bcbs_bad	F43#bcbs_bad	F43#s1_by_psych	F98#bad_u18n
F43#s1_by_psych		F43#s1_by_psych	F43#s1_by_psych	F98#bad_u18n	
F98#bad_u18n			F98#bad_u18n		
B. Probability of Receiving Drug Treatment in 1st 3 Months After 1st Claim, Conditional on Treatment					
dax#bad_u18n	dax#share_gp	first_eval#share_gp	dax#bad_u18n	dax#bad_u18n	dax#bcbs_bad
dax#share_gp	F43#bcbs_bad	dax#share_mh	dax#share_gp	dax#share_gp	dax#bad_u18n
F43#bcbs_bad	F43#share_gp	dax#bcbs_bad	F43#bcbs_bad	F43#bcbs_bad	F43#bcbs_bad
F43#share_gp		dax#share_gp	F43#share_gp	F43#share_gp	F98#bcbs_bad
F98#bcbs_bad		F43#bcbs_bad	F98#bcbs_bad	F98#bcbs_bad	bcbs_bad#bad_u18n
F98#bad_u18n		F43#share_gp	F98#bad_u18n	F98#bad_u18n	
F98#share_psych		F98#bcbs_bad	F98#share_psych	F98#share_psych	
C. Probability Receiving Red-Flag Drug Treatment in 1st 3 Months After 1st Claim, Conditional on Treatment					
first_eval#bad_u18n	F43#share_mh	dax#bcbs_bad	first_eval#bad_u18n	first_eval#bad_u18n	first_eval#bad_u18n
dax#bad_u18n	dax#bad_u18n	dax#bad_u18n	dax#bad_u18n	dax#bad_u18n	dax#bad_u18n
dax#share_psych	dax#share_psych	dax#share_gp	dax#share_psych	dax#share_psych	F43#bcbs_bad
F43#share_mh		F43#bcbs_bad	F43#share_mh	F43#bcbs_bad	F98#bcbs_bad
F43#bcbs_bad		F43#bad_u18n	F43#bcbs_bad	F43#share_gp	bcbs_bad#bad_u18n
F98#bcbs_bad		F98#bcbs_bad	F98#bcbs_bad	F98#bcbs_bad	
F98#share_psych		F98#bad_u18n	F98#share_psych	F98#share_psych	
no_diagnosis#bcbs_bad			bcbs_bad#bad_u18n	bcbs_bad#bad_u18n	
D. Probability Receiving Red-Flag Drug Treatment in 1st 3 Months After 1st Claim, Conditional on Drug Treatment					
dax#share_mh	dax#share_mh	dax#share_mh	dax#share_mh	dax#share_gp	first_eval#bad_u18n
bcbs_bad#bcbs_bad	bcbs_bad#bcbs_bad		bcbs_bad#bcbs_bad	bcbs_bad#bcbs_bad	dax#bad_u18n
bcbs_bad#bad_u18n	bcbs_bad#bad_u18n		bcbs_bad#bad_u18n	bcbs_bad#bad_u18n	F43#bcbs_bad
no_diagnosis#bcbs_bad					F98#bcbs_bad
					bcbs_bad#bad_u18n
E. Number of Potential Instruments					
87	87	87	93	70	23

Notes: This table presents the instruments selected by Lasso in first-stage models, where the instrument choice set varies across specifications. In columns 1-3, the Lasso is presented with the full 87-element potential instrument vector discussed in the text. In columns 4-6, the Lasso is presented w alternative potential instrument vectors. Column 1 lists the instruments selected in the main analysis, corresponding to Tables 5 and 6. Column 2 lists t instruments selected when the Lasso routine is constrained to chose at most 3 instruments, corresponding to Tables 7 and 8. Column 3 lists the instruments selected when only children with diagnoses are included in the analysis, corresponding to Tables 9 and 10. Column 4 expands the potentia instrument vector to include three additional physician supply measures--psychiatrists, therapists, and PCPs per 1,000 BCBS children. Columns 5 and (remove variables potentially reflecting demand from the original set, corresponding to Appendix Tables A3 and A4. The character "#" denotes an interaction between two variables. "Share_psych" and "share_gp" refer to the shares of psychiatrists and PCPs providing mental health treatment to children; "share_mh" refers to the average share of patients in PCPs caseloads receiving psychiatric drugs; "bcbs_bad" and "bad_u18n" refer to the sha of red-flag drugs in physician's caseloads in BCBS and IQVIA, respectively; and "s1_by_psych" refers to the share of psychiatrists among initial psychiatric prescribers. Indicators for first event in ER ("first_er"), first event in hospital ("first_hosp"), first event is evaluation ("first_eval"), and diagnoses of anxiety/depression ("dax"), adjustment disorders ("F43"), no diagnosis ("no_diagnosis"), and other MH diagnosis ("F98") are also shown.

Appendix Table A3
Effects of Treatment on Costs, Excluding Possible Demand Instruments

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Total Costs)			**Log(Total MH Costs)**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
A. Any Follow Up Treatment in 3 Months After 1st Claim (N=97306)						
Post-Lasso, Excl. PCP						
Treatment	0.253**	0.298**	0.290**	0.607**	0.776**	0.866**
SE	(0.064)	(0.061)	(0.052)	(0.078)	(0.090)	(0.083)
Post-Lasso, Excl. All Demand						
Treatment	0.262**	0.307**	0.289**	0.625**	0.794**	0.879**
SE	(0.065)	(0.063)	(0.054)	(0.077)	(0.087)	(0.079)
Mean Dependent Variable	6.785	7.761	8.387	5.205	5.669	6.010
B. Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=70886)						
Post-Lasso, Excl. PCP						
Treatment	1.128**	0.874**	0.661**	0.969**	1.216**	1.203**
SE	(0.065)	(0.058)	(0.051)	(0.087)	(0.089)	(0.083)
Post-Lasso, Excl. All Demand						
Treatment	1.170**	0.920**	0.696**	0.971**	1.214**	1.200**
SE	(0.073)	(0.063)	(0.056)	(0.089)	(0.091)	(0.084)
Mean Dependent Variable	6.871	7.844	8.456	5.212	5.728	6.081
C. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=70886)						
Post-Lasso, Excl. PCP						
Treatment	3.879**	3.006**	2.321**	2.512**	3.351**	3.365**
SE	(0.284)	(0.261)	(0.228)	(0.318)	(0.357)	(0.350)
Post-Lasso, Excl. All Demand						
Treatment	2.212**	1.917**	1.495**	1.238**	1.846**	1.891**
SE	(0.254)	(0.220)	(0.190)	(0.239)	(0.279)	(0.278)
Mean Dependent Variable	6.871	7.844	8.456	5.212	5.728	6.081
D. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=28650)						
Post-Lasso, Excl. PCP						
Treatment	0.424	0.719*	0.834**	-0.192	-0.158	-0.173
SE	(0.346)	(0.289)	(0.267)	(0.249)	(0.301)	(0.357)
Post-Lasso, Excl. All Demand						
Treatment	0.400	0.723**	0.860**	0.869**	0.728*	0.676+
SE	(0.322)	(0.262)	(0.238)	(0.283)	(0.341)	(0.393)
Mean Dependent Variable	6.781	7.848	8.495	3.736	4.334	4.783

Notes: This table presents results from post-Lasso 2SLS models of mental health treatment on log total health costs (columns 1-3) and log mental health costs (columns 4-6). In the "post-Lasso, Excl. PCP" model, the instruments used in the Lasso estimation exclude any instrument involving the average share of patients in PCPs caseloads receiving psychiatric drugs (N=70). In the "post-Lasso, Excl. All Demand" model, the instruments used in the Lasso estimation further exclude any instrument involving the shares of psychiatrists and PCPs providing mental health treatment to children and the share of psychiatrists among initial psychiatric prescribers (N=23). The instruments in the post-Lasso 2SLS models are chosen using a refined data-driven penalty and are displayed in Appendix Table A2. Panel A includes 5(4) instruments; Panel B includes 7(5) instruments; Panel C includes 8(5) instruments; and Panel D include 3(5) instruments in the post-Lasso, Excl. PCP (post-Lasso, Excl. All Demand) model. See Appendix Table A2 for a full list of the instruments selected in each model. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are in parentheses.
+ Significant at the 10 percent level
* Significant at the 5 percent level
** Significant at the 1 percent level

Appendix Table A4

Effects of Treatment on Facility Use, Excluding Possible Demand Instruments

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: Visited Facility?			**Number Nights in Hospital**		
	3 months	**12 months**	**24 months**	**3 months**	**12 months**	**24 months**
A. Any Follow Up Treatment in 3 Months After 1st Claim (N=97306)						
Post-Lasso, Excl. PCP						
Treatment	0.062**	0.099**	0.116**	1.120+	1.659*	2.598**
SE	(0.015)	(0.020)	(0.023)	(0.572)	(0.685)	(0.983)
Post-Lasso, Excl. All Demand						
Treatment	0.062**	0.102**	0.113**	1.222*	1.834*	2.790**
SE	(0.016)	(0.020)	(0.023)	(0.622)	(0.729)	(1.029)
Mean Dependent Variable	0.100	0.169	0.242	0.289	0.541	0.915
B. Drug Treatment 1st 3 Months After 1st Claim, Cond. on Treatment (N=70886)						
Post-Lasso, Excl. PCP						
Treatment	0.085**	0.038*	0.007	-0.531	-0.403	-1.559+
SE	(0.013)	(0.017)	(0.022)	(0.772)	(0.738)	(0.806)
Post-Lasso, Excl. All Demand						
Treatment	0.080*	0.034+	0.003	-0.482	-0.376	-1.597*
SE	(0.013)	(0.017)	(0.021)	(0.761)	(0.732)	(0.802)
Mean Dependent Variable	0.101	0.173	0.247	0.333	0.607	1.033
C. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Treatment (N=70866)						
Post-Lasso, Excl. PCP						
Treatment	0.204**	0.039	-0.105	1.935**	1.639	-2.016
SE	(0.048)	(0.061)	(0.071)	(0.692)	(1.163)	(3.056)
Post-Lasso, Excl. All Demand						
Treatment	0.170**	0.027	-0.035	-2.375	-2.732	-5.680+
SE	(0.033)	(0.045)	(0.054)	(2.767)	(2.803)	(3.023)
Mean Dependent Variable	0.101	0.173	0.247	0.333	0.607	1.033
D. Red-Flag Drug Treatment in 1st 3 Months, Cond. On Drugs (N=28650)						
Post-Lasso, Excl. PCP						
Treatment	0.022	0.022	0.068	-3.223	-4.173	-2.437
SE	(0.057)	(0.072)	(0.076)	(3.162)	(3.830)	(4.984)
Post-Lasso, Excl. All Demand						
Treatment	0.108+	0.136+	0.195**	3.716	-4.099	-2.744
SE	(0.056)	(0.072)	(0.072)	(3.577)	(4.054)	(5.344)
Mean Dependent Variable	0.149	0.229	0.304	0.665	1.076	1.638

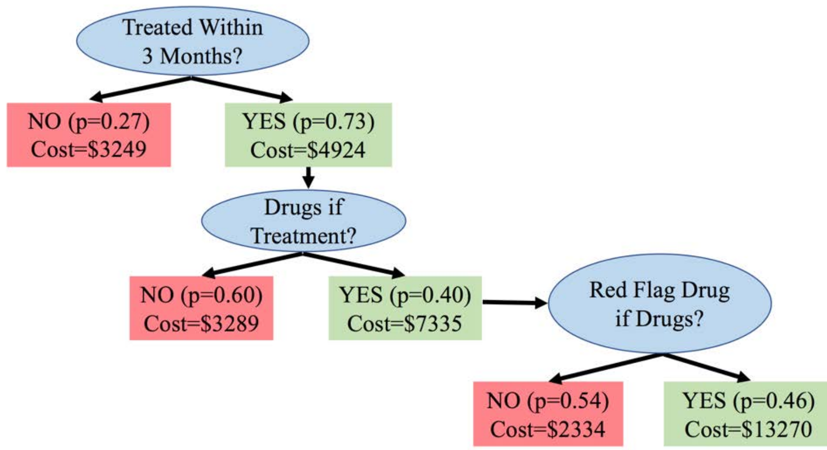
Notes: This table presents results from OLS and post-Lasso 2SLS models of mental health treatment on facility use (columns 1-3) and the number of overnight stays in the hospital (columns 4-6). In the "post-Lasso, Excl. PCP" model, the instruments used in the Lasso estimation exclude any instrument involving the average share of patients in PCPs caseloads receiving psychiatric drugs (N=70). In the "post-Lasso, Excl. All Demand" model, the instruments used in the Lasso estimation further exclude any instrument involving the shares of psychiatrists and PCPs providing mental health treatment to children and the share of psychiatrists among initial psychiatric prescribers (N=23). The instruments in the post-Lasso 2SLS models are chosen using a refined data-driven penalty and are displayed in Appendix Table A2. Panel A includes 5(4) instruments; Panel B includes 7(5) instruments; Panel C includes 8(5) instruments; and Panel D includes 3(5) instruments in the post-Lasso, Excl. PCP (post-Lasso Excl. All Demand) model. See Appendix Table A2 for a full list of the instruments selected in each model. All models include patient controls (age, female, severity, and diagnosis) as well as month and year fixed effects, corresponding to the date of the child's initial mental illness claim. Standard errors, clustered at the ZIP-3 level, are parentheses.

+ Significant at the 10 percent level

* Significant at the 5 percent level

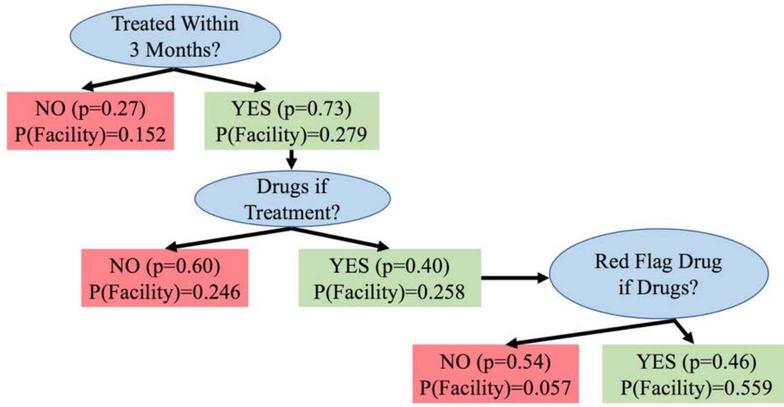
** Significant at the 1 percent level

Appendix Figure A1
Implications for Total Cost at 24 Months Among Diagnosed Children



Notes: This figure traces out the implications of each initial treatment choice on total costs at 24 months among the subset of children who received a mental health diagnosis within 3 months. The share of children in each branch appears in parentheses. Cost estimates are based on post-Lasso estimates in column 3 of Table 9.

Appendix Figure A2
Implications for Facility Use at 24 Months Among Diagnosed Children



Notes: This figure traces out the implications of each initial treatment choice on facility use—ER visit or hospitalization—at 24 months among the subset of children who received a mental health diagnosis within 3 months. The share of children in each branch appears in parentheses. Cost estimates are based on post-Lasso estimates in column 3 of Table 10.