Intra-household Allocation and the Mental Health of Children: Structural Estimation Analysis

by

Morris A. Davis
Federal Reserve Board of Governors
20th and C St., NW
Washington, DC 20551
U.S.A.
mdavis@frb.gov

E. Michael Foster
Georgia State University
School of Policy Studies
Atlanta, GA 30303
U.S.A.
pademf@panther.gsu.edu

PRELIMINARY, July 1999

Abstract:

This paper estimates the structural parameters of a dynamic model where parents with one child periodically decide whether or not their child uses various mental health services. In this model, mental health services improve a child's mental health (which parents care about), however, mental health services may be costly to the parents both in terms of utility and household consumption. Using a panel data set collected as part of the Fort Bragg Mental Health Demonstration, we estimate the model with a maximum likelihood procedure that accounts for unobservable differences in mental health endowments of children, and, population heterogeneity in parental preferences and the effectiveness of mental health services. We estimate that parents experience relatively high disutility when a child uses mental health services, implying parents enroll their children in mental health services only if these services have multi-period effects on their child's mental health. Correspondingly, we find that outpatient and inpatient mental health services have permanent effects on a child's mental health. We conclude that the improvement over time of the mental health of the children in our data is, in a large part, the outcome of forward-looking parents choosing to increase their child's mental health.

The opinions expressed here are those of the authors and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System or its staff. This project began and was nearly completed while Davis was a graduate student at the University of Pennsylvania. Correspondingly, much of the research support for this paper comes from the University of Pennsylvania and Georgia State University. The authors are especially indebted to the University of Pennsylvania faculty and Georgia State University faculty for comments. All errors and omissions are the fault of the authors.

1. Introduction

Economists have devoted considerable attention to the process by which parents allocate resources to children. Researchers have tried to identify the share of total household expenditures dedicated to children (Lazear and Michael (1988)), estimate the influence of parental behaviors on a child's health at birth and beyond (Rosenzweig and Wolpin (1995, 1988)) and determine whether transfers from parents to children are motivated purely by altruism (Altonji, Hayashi, and Kotlikoff (1997)).[1] The specific aspects of parent and child resource allocation examined by economists vary widely, but a common theme of much of this research is that the resources parents give to or invest in a child depends on parental preferences over the child, child characteristics, and a technology that converts investments into increased utility of the child (Sheshinski and Weiss (1982), Behrman, Pollack, and Taubman (1982), Becker and Tomes (1976)).

In this spirit, we formulate and estimate a dynamic model of parental investments in a child's mental health. Even though mental health may be an important component of a child's human capital, economists have not examined the decision process determining parental investments in a child's mental health. Our model encapsulates the basic theme of the economics literature on parental investments in children. The details of this model, however, are specific to the study of parental investments in a child's mental health. In our model, parents with one child receive per-period utility from household consumption and that child's stock of mental health. Parents can obtain different combinations of outpatient, inpatient, and (under certain circumstances) intermediate mental health services in order to temporarily and/or permanently improve their child's mental health. Mental health services, however, may be costly to parents in terms of both household consumption and direct disutility. Given the effectiveness of available mental health services, the monetary costs of these services, the disutility associated with these services, and the degree of substitutability of household consumption and a child's mental health, parents periodically decide whether or not their child uses mental health services until their child turns eighteen, at which point parents make no more decisions for their child. Parents make the decisions that maximize the present value of their expected remaining lifetime utility.

We estimate the structural parameters of this model using data from the Fort Bragg Demonstration, which is one of the largest evaluations ever conducted in the field of children's mental health services (Bickman, et. al. (1995)). Implemented by the U.S. Army at Fort Bragg, North Carolina, the Fort Bragg Demonstration tested the "Continuum of Care" philosophy of providing mental health services to children. As part of the Continuum of Care, the Demonstration offered standard outpatient and inpatient mental health services to children as well as non-residential and residential intermediate mental health services that are generally not available elsewhere. All outpatient, inpatient, and intermediate services were provided at the Demonstration for free.[2]

Data were continuously recorded on the use of services for a sample of children in the Demonstration. The use of mental health services was also recorded for a comparison group of children residing at Fort Campbell, Kentucky, and Fort Stewart, Georgia. The children in this "Comparison" group received services through the military's insurance plan, the Civilian Health and Medical Program for the Uniformed Services (CHAMPUS), and paid copayments on the use of services. In other words, Comparison children had access only to outpatient and inpatient services, and these services were not freely provided. Every six months, parents of Demonstration and Comparison children were asked questions on household income and their child's mental health. All Demonstration and Comparison children were "recruited" into the initial wave of the study from the population of children using mental health services.

We estimate the structural parameters of our model using a simulated maximum likelihood procedure similar to that of Keane and Wolpin (1994). This procedure embeds the solution of the parents' dynamic programming problem directly into the likelihood function. The likelihood procedure accounts for measurement error in household income and occasional missing data in later waves of the sample. The likelihood procedure also accounts for unobserved heterogeneity among children and households: specifically, the mean to which a child's mental health reverts in the absence of treatment and the mean around which log household income reverts vary unobservably by household. Finally, the likelihood calculations allow preferences and technologies to unobservably vary between households. We posit that there are two types of households in the world ("Type one" and "Type two" households), and that households of different types vary in preferences for mental health and the temporary and permanent efficacy of mental health therapy.

We estimate our model with four waves (covering a span of two years) of data on 201 single-child households. This panel data set is large enough to estimate the structural parameters of a model where parents routinely make decisions involving their child's use of mental health services. The nature of the data, however, introduces complexity to our model. As noted, children at the Demonstration have access to a broader array of mental health services than Comparison children. Our model, therefore, specifies that parents of Demonstration children face a different choice set than parents of Comparison children. Furthermore, the nature of the sample selection induced by the data collection procedure limits our ability to generalize our results. Specifically, children were recruited into the Demonstration and Comparison groups if they had used services near the time of the start of the study. In the population service use is rare; we estimate that the propensity to use services is correlated with household type. As a result, we cannot use these data to identify the distribution of types in the population.

One of our findings is that the average disutility parents experience when their child uses mental health services is large relative to the range of utility parents experience when their child does not use services. Therefore, parents rarely purchase mental health services if these services only have one-period effects on their child's mental health. In the first wave of our data, most parents purchase mental health services for their child, which is behavior that corresponds with another of our findings: outpatient and inpatient mental health services have beneficial and permanent effects on a child's mental health. When these results are taken together they reveal that parents' choices over their child's use of mental health services appears to be the outcome of forward-looking behavior.

We determine that the use of services is a major factor underlying the improvement over time of the mental health of the children in this study. Evidence from previous research regarding the effectiveness of mental health services for children is mixed.[3] Most of the studies in this literature that attempt to randomly assign children to mental health therapies find that treatment is not effective in improving a child's mental health, although the sample size in these studies is generally very small (Smyrnios and Kirkby (1993), Weisz, et. al. (1995)). Many studies using data collected from community or clinical settings also find that services are ineffective at improving a child's mental health. These studies, however, are subject to a host of econometric problems. Summarizing, even though these studies generally compare treatment and comparison groups across a wide range of characteristics (Weisz, et. al. (1995)), they ignore the fact that treatment decisions are correlated with unobservable characteristics of the child and unobserved variations in preferences and treatment effectiveness in the population (Foster (1998)). As a result, these studies confound the marginal impact of mental health services with unobservable characteristics of children and households that are determinants of service use.

We also find that the use of mental health services is both price and income inelastic, a result that is driven by the relatively high average disutility parents experience from their child's use of mental health services. This finding contradicts research from the Rand Health Insurance Experiment (HIE), which suggests that children's use of medical services is sensitive to the financial terms under which care is provided.[4] Although in any wave of our data, service use by Demonstration children (who receive freely provided services) is higher than use of Comparison children, we estimate that these differences do not stem from a high price elasticity of parents, but from between site differences in the baseline distribution of household types.

Finally, we assess the impact of the Demonstration itself on mental health outcomes. For the most part, prior research has assessed the impact of the Demonstration by comparing outcomes of Comparison children with those of Demonstration children (Bickman et. al. (1995)). However, such comparisons may be misleading if the distribution of types differs among households in the Demonstration and Comparison. Therefore, to assess the impact of the Demonstration, we simulate outcomes of Demonstration children if Demonstration parents had the same choice set and paid the same costs as Comparison children. These simulations suggest that neither the intermediate services offered by the Demonstration nor the fact that outpatient and inpatient services were free altered the mental health of the enrolled children. This result coincides with most assessments of the Demonstration.

We make several restrictive assumptions throughout. We do not observe household assets, so we assume parents neither borrow nor save out of household income. Even though we believe this is one of the first papers to accurately estimate the effects of mental health services for children, we assume all parents are knowledgeable about these effects when making their decisions. We have no time-use data, so we do not model the time-allocation problem parents must solve when determining their child's use of mental health services. Finally, we limit ourselves to modeling households with one child for two reasons: the multiple-child household model is considerably more complicated than our current model, and, in our data we do not observe use of mental health services of siblings.

2. Model

As noted, in our data children are either enrolled in the Demonstration (which took place at Fort Bragg, North Carolina) or they are part of a Comparison group of children from either Fort Stewart, Kentucky, or Fort Campbell, Georgia. At the start of every six month period (periods are denoted by the index *t*), parents of a Comparison child must decide whether or not their child will receive no mental health services, outpatient mental health services, or both outpatient and inpatient mental health services during the current six month period.[5] Parents of a child participating in the Fort Bragg Demonstration in period *t* have all of these choices as well. However, parents at the Demonstration additionally decide whether or not their child receives non-residential services (such as after-school counseling) and/or residential services (such as group home services). We deem last two mental health services as "intermediate" services throughout; as with inpatient mental health services, outpatient services must accompany any intermediate services. Parents make these choices for their child in all periods in which their child is between the ages of 5 and 18 (inclusive).[6]

Let the dummy variable d_t^1 denote parents' period *t* outpatient choice, d_t^2 represent parents' period *t* inpatient choice, d_t^3 represent parents' period *t* intermediate non-residential services choice, and d_t^4 denote parents' period *t* intermediate residential services choice. Finally, let the dummy variable FB denote whether or not parents have a child that is enrolled in the Fort Bragg Demonstration: if the child is enrolled in the Demonstration $FB=1$, otherwise the child is part of the Comparison group and $FB=0$. Using this notation, Table 1 displays the mental health choice set of parents of Comparison children, $D(0)$, and parents of Demonstration children, $D(1)$. Since parents of Comparison children have no access to intermediate services, $D(0)$ consists of the top three choices of Table 1. Parents of children in the Fort Bragg Demonstration have access to all outpatient, inpatient, and intermediate services, so $D(1)$ consists of the bottom nine choices of Table 1. Both $D(0)$ and $D(1)$ reflect the fact that individuals using inpatient or intermediate services also receive outpatient services.

In every period, parents receive utility from their child's stock of mental health, denoted M_t, and from consumption of market goods denominated in thousands of dollars[7], denoted C_t. We define M_t as $100 - CBCL_t$, where $CBCL_t$ is the child's reported "Child Behavior Check-List" score. The CBCL score has a possible range from 0 (best mental health) to 100 (worst mental health).[8] Parents also experience random direct additive marginal utility or disutility from the period *t* choices themselves: they experience additive utility or disutility of b_t^i if $d_t^i = 1$, for $i = 1, \ldots, 4$. We define parents' period *t* utility from household consumption, the child's mental health, and their mental health choices as

(1) $$\log\left[\left(C_t^{\beta} + \alpha M_t^{\beta}\right)^{\frac{1}{\beta}}\right] + \sum_{i=1}^{4} d_t^i b_t^i .$$

In (1), we assume α is greater than zero and β is less than one.[9] Parents are risk averse in the composite[10] of C_t and M_t, $\left(C_t^{\beta} + \alpha M_t^{\beta}\right)^{\frac{1}{\beta}}$. If b_t^i is less than zero for any $i = 1, \ldots, 4$, then b_t^i can be thought of as the "stigma" associated with mental health service *i*.

For all choices $i = 1, \ldots, 4$, the marginal utility of choice *i* (b_t^i) consists of a known deterministic component and a random shock realized at the beginning of the period before any decisions are made, i.e.

(2) $$b_t^{\,i} = \bar{b}^{\,i} + e_t^{\,i}.$$

In (2), $\bar{b}^{\,i}$ is the nonrandom average marginal utility for choice *i* and $e_t^{\,i}$ is a random shock to the marginal utility of choice *i*. All of the marginal utility shocks are jointly drawn before any of the period *t* decisions are made. Although the shocks to marginal utility may be contemporaneously correlated, they are independently drawn over time and independently drawn of all other variables in the model.

After all of the period *t* decisions have been made, the child's CBCL score evolves according to the following known stochastic process,

(3) $$\begin{aligned} CBCL_{t+1} &= \theta_{0,t} + \theta_1 \left(CBCL_t\right) + \sum_{i=1}^{4} d_t^{\,i} \theta_2^i + m_{t+1} \\ \ln\left(\theta_{0,t+1}\right) &= \ln\left(\theta_{0,t}\right) + \sum_{i=1}^{4} d_t^{\,i} \theta_3^i. \end{aligned}$$

In (3), m_{t+1} is a random shock to the child's CBCL score that is realized after all period *t* decisions are made and before period *t+1* begins. This specification for the evolution of a child's CBCL score (and thus the child's mental health) is more general than almost all previous specifications in the empirical mental health literature. For example, some authors (Foster (1998)) implicitly set θ_3^i to zero for all $i=1,\ldots,4$, implying all effects of mental health services are temporary: when $\theta_3^i=0$ for all $i=1,\ldots,4$, the child's CBCL reverts to a fixed mean $\left(\theta_{0,t}\right)\left(1-\theta_1\right)^{-1}$ (as long as $\theta_1<1$) regardless of any mental health therapy the child undertakes or the number of episodes of treatment the child has experienced. Other authors first difference their mental health measure, and regress the first difference on a treatment dummy and other covariates (Target and Fonagy (1994)). These regressions implicitly specify all effects of services as permanent. With our specification, the temporary effects of services are captured by the θ_2^i terms and the permanent effects of services are captured by the θ_3^i terms.

Annual household income in period *t+1*, denoted W_{t+1}, is a known stochastic function of the child's mental health, the child's age, and the parents' period *t* mental health choices. This stochastic function is

(4) $$\ln W_{t+1} = \varphi_0 + \varphi_1 \ln W_t + \varphi_2\left(CBCL_t - \overline{CBCL}\right) + \varphi_3\left(age_t - \overline{age}\right) + \sum_{i=1}^{4} d_t^i \varphi_4^i + w_{t+1}.$$

w_{t+1} is a random shock to household income that is realized after all period *t* decisions have been made and before period *t+1* begins. We impose $\varphi_1 < 1$, implying in the absence of shocks and conditional on the child's age, CBCL score, and the parents' choices, log household income reverts to a mean of $\varphi_0(1-\varphi_1)^{-1}$. We include $CBCL_t$, age_t, and d_t^i for $i=1,\ldots,4$ in (4) to account for labor market outcomes we do not directly model that are influenced by a child's use of mental health services, a child's age, and a child's mental health status.[11] The random shocks to the child's CBCL score and household income, m_{t+1} and w_{t+1}, are jointly drawn after all period *t* decisions have been made. Although these shocks may be correlated, they are drawn independently over time, independently of the period *t* shocks to marginal utility, and independently of all other variables in the model.

We do not observe household assets in our data, so in our model parents do not have savings and cannot borrow against their future income to finance either their current consumption or their child's use of mental health services. Since parents of Demonstration children do not pay anything for mental health services their child uses, we specify they consume (in thousands of dollars) one half their annual income every six month period, implying $C_t = \left(\frac{1}{2}\right)\left(\frac{1}{1000}\right) W_t$. However, parents not enrolled in the Fort Bragg Demonstration pay a copayment of p^i at the beginning of period *t* for mental health service *i* (for $i = 1,2$) and all remaining income is spent on consumption.[12] These parents consume (in thousands of dollars)

(5)
$$C_t = \left(\frac{1}{2}\right)\left(\frac{1}{1000}\right)\left[W_t - \sum_{i=1}^{2} d_t^i p^i\right] \geq 0,$$

where C_t is non-negative because we do not allow parents to purchase services if the price of those services is larger than household income. Given our definition of FB, we define household consumption in each (six-month) period for all parents in thousands of dollars as

(6)
$$C_t = \left(\frac{1}{2}\right)\left(\frac{1}{1000}\right)\left[W_t - (1 - FB)\sum_{i=1}^{2} d_t^i p^i\right] \geq 0.$$

Parents assume the copayments on mental health services required by CHAMPUS is fixed over time.

Define the variable S_t as the vector of all relevant time-varying variables parents use at the start of period *t* to make their set of period *t* decisions. We have established through the structure of the problem that for a particular household this vector of "state variables" consists of[13]

$$\textbf{(7)} \qquad S_t = \left[W_t \quad M_t \quad \theta_{0,t}\right].$$

Note that there is a one-to-one mapping of the current period *t* and age_t: we define $t=1$ when the child is born, $t=2$ when the child is six months old, etc...[14] After all of the shocks to the marginal utility from choices have been drawn, parents make choices $d_t^1,\ldots,d_t^4$ in period *t* to maximize the present value of their expected remaining lifetime utility. This maximized value equals the value of the solution of

$$\textbf{(8)} \qquad \begin{aligned} V_t(S_t) &= \max_{\left(d_t^1,\ldots,d_t^4\right)\in D(FB)} \left[\ln\left[\left(C_t^{\beta}+\alpha M_t^{\beta}\right)^{\frac{1}{\beta}}\right]+\sum_{i=1}^{4} d_t^i b_t^i+\delta E_t\left[V_{t+1}\left(S_{t+1}\right)\right]\right] \\ s.t. \quad C_t &= \left(\frac{1}{2}\right)\left(\frac{1}{1000}\right)\left[W_t-(1-FB)\sum_{i=1}^{2} d_t^i p^i\right]\geq 0. \end{aligned}$$

In equation (**8**), $D(FB)$ is defined in Table 1 and $\delta=0.98$ (which is not estimated) is the factor by which parents discount future utility.[15] The expectations operator in (**8**) can also be written as $E_t\left[V_{t+1}\left(S_{t+1}\right)\mid S_t,d_t^1,\ldots,d_t^4\right]$. This expectations operator thus explicitly integrates $V_{t+1}\left(S_{t+1}\right)$ with respect to the distribution of S_{t+1} and the distribution of the period *t+1* utility shocks given the parents' period *t* choices and the parents' period *t* state variables. Parents know the technological processes that map their current choices and state variables into future state variables, and they use this knowledge to correctly calculate the expectation in (**8**): they use (**3**) and (**4**) to calculate the expectation in equation (**8**), and (**3**) and (**4**) are the true data-generating processes.

Let T be the period when the child is eighteen (i.e. $T=36$). When the child is eighteen, parents receive a “terminal” utility of

(9) $$V_T(S_T)=(1-\delta)^{-1}\ln\left[\left(W_T^{\beta}+\alpha M_T^{\beta}\right)^{\frac{1}{\beta}}\right],$$

As seen in equation (9), it is assumed that when the child is age eighteen parents no longer make choices for their child. For simplicity, it is also assumed that in all periods $t \geq T$ household income and the child’s mental health are permanently fixed at their period T levels.[16]

This model is defined as having a solution when, given any feasible value of the state space and any realization of the utility shocks in that period, the entire set of optimal decisions is known. We do not know how to calculate an analytic solution to the model, which is a function that maps the period, the state space, the utility shocks, and the entire set of parameters of the model to the set of optimal decisions. Instead, we condition on a particular set of parameters and computationally generate a solution to the model at that set of parameters using the solution algorithm that is similar to the algorithm described by Keane and Wolpin (1994). For an explanation of this solution algorithm as it is applied to this particular model, see Appendix A.

3. Data

Data on demographic information, household income, and the mental health of Demonstration and Comparison children were collected every six months for 984 children and adolescents ages 5 through 17 over two years.[17] These children entered the study between 1990 and 1993, with roughly half entering in 1991 (the first full year of the Demonstration). The children were recruited into the study if they had used services prior to the start of the Demonstration, and as a result over ninety percent were using services at the time of their initial interview. Many children stopped using services by the end of the study, so their experiences provide information on children that do not use services. Children who never received mental health services are not included in our data.

For each child in the sample, a care-giver was identified and in-person interviews were conducted with that individual every six months. In most cases, the care-giver was the child's mother and did not change across interviews. The interviews with the care-giver provide information on the child's mental health, household income, and demographics. We use the Child Behavior Checklist (CBCL) as our measure of the child's mental health. The CBCL total score is a measure of mental health symptomatology[18] and is commonly used in mental health research. To calculate a child's CBCL score, parents are asked a series of 112 questions about problems their children may be having at home, in school, or with peers during the past 6 months. For each behavior or characteristic, the respondent reports whether the characteristic is never, sometimes, or often true of their child. Responses to these items are summed, and a standardized score is calculated based on national norms for the child's age and gender. The resulting measure has a mean of 50 and a standard deviation of 10 in the population. As mentioned, children with many behavioral problems have high CBCL scores, while children with relatively few behavioral problems have low CBCL scores.

For household income, respondents are asked to place their "total household income in the preceding tax year" in a series of categories: less than $5,000, between $5,000 and $9,999, between $10,000 and $14,999, between $15,000 and $19,999, between $20,000 and $29,000, between $30,000 and $39,000, between $40,000 and $59,000, and $60,000 or more. We assign "observed" household income as the midpoint of these categories, and households in the highest category are assigned income of $75,000. For demographics, interview data provide information on the child's age, gender, race (which we group into "white, non-hispanic" and "other"), household structure (a series of questions that identify individuals living in the household and their relationship to the child), and the education of the child's care-giver. We classify care-givers as being a high school graduate, having some post-secondary education, or having a college degree.[19]

We derive data on mental health service use from two sources. Our primary source of service use data for Comparison children are financial transactions from the CHAMPUS health insurance system records. Each CHAMPUS transaction notes the date and mental health service received. For Demonstration children, all services were either received at or arranged by a central clinic. That clinic maintained a management information system (MIS) that recorded the type of service, the number of service units received, and the date of service. The MIS and CHAMPUS records appear fairly comparable in terms of their accuracy (for details, see Foster, Summerfelt, and Saunders (1996)). Both the MIS and the CHAMPUS claims include an identification number for the child's military sponsor and the child's year of birth, and we use this information to link service use data to specific children in the sample.

We use the MIS and CHAMPUS records to measure service use during the six-month period between interviews. For each period, we identify whether a child received outpatient services, inpatient services (including care in residential treatment centers), intermediate non-residential services and intermediate residential services. We specify that any child that visited an outpatient service provider at least once in the six month period between interviews received "outpatient services" in that period, and we make similar classifications for inpatient, intermediate non-residential, and intermediate residential mental health services. Individuals that did not visit an outpatient, inpatient, or intermediate service provider between interviews were assigned as not having received any mental health services.[20]

The evaluation obtained interview data for 984 households, but our analyses are based on a subset of this data. First, we exclude 136 cases who lack CHAMPUS claims or MIS data in the original study. The difficulties involved with modeling within-household allocations of resources, and, the lack of data on the use of mental health services of siblings necessitates that we exclude multiple child households from our sample. This reduces our sample by 589 observations. We also drop 22 cases who were living with neither parent at the first interview. Finally, we drop individuals who were missing any demographic information, income, or CBCL score at the first wave of the study.[21] As a result, the sample we use to estimate our model consists of 201 only children followed over two years.[22]

Table 2 describes the initial characteristics of our sample. Consistent with the findings of previous research (Bickman, et. al. (1995)), the top panel of Table 2 shows that there do not appear to be any significant demographic differences between the Demonstration and Comparison children: the majority of both groups of children are male and white, the care-givers in the study are relatively well educated (which reflects the educational requirements for enlistment in the Army), and roughly half of the children are not living with their father. The middle portion of Table 2 shows the 25th, 50th, and 75th percentile of household income and the average CBCL score of the children in our sample. It appears that neither the initial distribution of household income nor the average CBCL score varies between Demonstration and Comparison children. The Wave 1 average CBCL scores show that the children start the study with relatively severe mental health problems. To get a better sense of the severity of the problems among these children, it is worth looking at the prevalence of specific CBCL items. In the six months prior to Wave 1, 55% of sample members complained of loneliness, 19% were sometimes or often cruel to animals, 17% had attempted to harm themselves, 10% heard voices, and 7% had set fire to property.

The bottom portion of Table 2 shows the mental health service use of children in our sample in the first observation period, i.e. Table 2 reports the fraction of children that use outpatient, inpatient, or intermediate mental health services at least once between Wave 1 and Wave 2. Due to the nature of the data collection, over ninety percent of Demonstration children and eighty percent of the comparison children use services in the first period. The use of inpatient services is much lower among Demonstration children, and this has been attributed to the availability of intermediate services at the Demonstration (Foster, Summerfelt, and Saunders (1996)).

Table 3 shows how CBCL scores and mental health service use vary across waves for both Demonstration and Comparison children. Three striking features of this table deserve mention. First, the use of mental health services for all children starkly fell from Wave 1 to Wave 4. Second, as shown by the large decrease in average CBCL scores from Wave 1 to Wave 4, the mental health of both Demonstration and Comparison children improved substantially.[23] Finally, at Wave 4 there is little difference in CBCL scores for the Demonstration and Comparison children. This is consistent with the findings in Bickman et al. (1995) and forms the basis for the conclusion that the Demonstration did not improve the mental health outcomes of children.

4. Likelihood and Functional Forms

We calculate the likelihood of our data at a given set of parameters using a simulation method that directly embeds the solution of our model into the likelihood function. Our simulation procedure also accounts for measurement error in household income, systematic differences among households in preferences and technology, unobserved variation in the mean of the child's CBCL score and in the mean of household log income, and missing income and mental health data.

To help explain the likelihood that we actually estimate, we start by explaining how we would estimate a simpler likelihood. Suppose that we observe household income without error and the child's CBCL score for an arbitrary household *j* in periods τ and $\tau+1$. Furthermore, suppose we observe the mean of the CBCL that household *j*'s child will revert to in the absence of treatment in period τ, $\frac{\theta_{0,\tau}^{j}}{1-\theta_1}$, and the mean of household log income $\frac{\varphi_0^{j}}{1-\varphi_1}$. At a given set of parameters to the model, the likelihood that household j's observed choices $\left(d_{\tau,j}^{1}, \ \dots \ ,d_{\tau,j}^{4}\right)$ occur in period τ and the likelihood of observing household j's income and CBCL is

$$\textbf{(10)} \qquad Pr\left(d_{\tau,j}^{1}, \ \dots \ ,d_{\tau,j}^{4} \ \middle| \ S_{\tau}^{j}\right) \, g\left(m_{\tau+1},w_{\tau+1}\right).$$

In the above equation, S_{τ}^{j} are household j's state variables in period τ (M_{τ}, W_{τ}, $\frac{\theta_{0,\tau}^{j}}{1-\theta_1}$, and $\frac{\varphi_{0}^{j}}{1-\varphi_1}$), $Pr(.)$ is the probability the observed choices occur given the observed state variables, $g(.)$ is the joint density of the CBCL ($m_{\tau+1}$) and log income ($w_{\tau+1}$) shocks, and $m_{\tau+1}$ and $w_{\tau+1}$ are given by equations (3) and (4).[24]

The model generates a probability distribution over choices because the shocks to the marginal utility of choices are not observed. We generate $Pr\left(d_{\tau,j}^{1}, \ \dots \ ,d_{\tau,j}^{4} \ \middle| \ S_{\tau}^{j}\right)$ using a smoothed simulator, as in Eckstein and Wolpin (1997). Given a solution to the model at the set of model parameters for which we wish to calculate a likelihood, we draw $l=1,\dots,L$ sets of utility shocks and compute

$$\textbf{(11)} \qquad Pr\left(d_{\tau,j}^{1}, \ \dots \ ,d_{\tau,j}^{4} \ \middle| \ S_{\tau}^{j}\right) = \sum_{l=1}^{L}\left[\frac{\exp\left(\frac{V_{\tau,l}^{j}-\max V_{\tau,l}}{\lambda}\right)}{\sum_{k}\exp\left(\frac{V_{\tau,l}^{k}-\max V_{\tau,l}}{\lambda}\right)}\right].$$

In the numerator of (11), $V_{\tau,l}^{j}$ is the value to household j of choosing observed choices $\left(d_{\tau,j}^{1}, \ \dots \ ,d_{\tau,j}^{4}\right)$ in period τ given states S_{τ}^{j} and the particular utility shock corresponding to draw l. This value is derived from equation (8). $\max V_{\tau,l}$ is the value to household j resulting from the optimal set of choices (also given by (8)) at the particular utility shocks of draw l. In the denominator of (11), k is an index for all the possible mental health choices facing household j in period τ. λ controls the degree of smoothing. In estimation, we set $\lambda=0.01$; consistency requires that $\lambda\rightarrow 0$ as the sample of children increases. Equation (11) yields choice probabilities that are smooth in the parameters of the model for computationally feasible L, allowing us to maximize the likelihood with respect to the model's parameters using derivative-based methods.

We can write (10) for an individual with observed choices from periods τ through $\tau+3$ (corresponding to 4 periods of observed data), observed CBCL scores and household income observed (without error) in periods τ through $\tau+4$, and $\frac{\theta_{0,t}^{j}}{1-\theta_1}$ and $\frac{\varphi_0^{j}}{1-\varphi_1}$ observed at $t=\tau$ as[25]

$$\textbf{(12)}\qquad \prod_{t=\tau}^{\tau+3} Pr\left(d_{t,j}^{1}, \ \dots \ ,d_{t,j}^{4} \ \mid \ S_t^{j}\right) \, g\left(m_{t+1},w_{t+1}\right).$$

We only need to observe $\frac{\theta_{0,t}^{j}}{1-\theta_1}$ and $\frac{\varphi_0^{j}}{1-\varphi_1}$ at the first observation period ($t=\tau$) because $\frac{\varphi_0^{j}}{1-\varphi_1}$ is fixed over time for each household and equation (3) determines $\frac{\theta_{0,t}^{j}}{1-\theta_1}$ for $t>\tau$. The likelihood of the whole sample is the product of (12) for all households in the sample.

For the remainder of this section, we drop the use of the j subscript and superscript. We do not directly maximize (12) with respect to the parameters of the model. As mentioned, income is categorically measured, and we include measurement error of income in the likelihood.[26] In an ad-hoc fashion, we assume that the logarithm of household j's "true" but unobserved period t income is normally distributed with mean equal to the logarithm of period t observed income and variance σ_{inc}^{2}. In other words, we assume

$$\text{(13)} \quad \begin{aligned} \ln(W_{t,o}) &= \ln(W_{t,r}) + inc_t \\ inc_t &\sim N(0, \sigma_{inc}^2), \end{aligned}$$

where $W_{t,r}$ is true household income and $W_{t,o}$ is observed household income (which is defined as the midpoint of the income category except for the highest category). We assume that the measurement error on income inc_t is serially independent and independent of all other shocks and variables in the model.[27]

We introduce population heterogeneity in preferences and technology into the likelihood by positing that there are two types of households in the world ("Type 1" and "Type 2" households), and types of households vary in preferences for mental health (α), their elasticity of substitution parameter (β), and the temporary and permanent effectiveness of mental health services (θ_2^i and θ_3^i for $i=1,\ldots,4$). When we say that there are two types of households in the population, we mean that there are two combinations of the parameter vector

$$\text{(14)} \quad \left[\alpha \quad \beta \quad \{\theta_2^i\}_{i=1}^4 \quad \{\theta_3^i\}_{i=1}^4\right]$$

in the population.[28] Since the parameters in (14) affect the costs, benefits, and outcomes of choices, different type households with identical state variables make systematically different decisions and have systematically different outcomes.

We do not observe a household's type, which does not change over time for any given household. However, in the likelihood we specify that the probability a household is a given type is a logistic function of variables of the first observation period:

$$Pr(TYP=1) = \frac{\exp(Z)}{1+\exp(Z)}$$

(15)

$$Z = \zeta_0+\zeta_1\ln(W_{\tau,r})+\zeta_2\ln(M_\tau)+\zeta_3 FB+\zeta_4(age_\tau-5)+ \zeta_5(WHITE)+\zeta_6(LOED)+\zeta_7(HIED)+\zeta_8(DADABS).$$

In this equation $W_{\tau,r}$ is "true" income in the first observation period, M_τ is the CBCL score of the child in the first observation period, FB denotes whether the household is in the Demonstration ($FB=1$) or not, and age_τ is the child's age in years in the first observation period. By including *FB* in the type equation, we can test for the presence of unobserved between-site differences. $WHITE$, $LOED$, $HIED$, and $DADABS$ are baseline demographic, education, and household structure dummy variables that have been shown in previous research to be significantly correlated with mental health decisions and outcomes (Foster, (1998)). These variables are not otherwise explicitly included in the model. $WHITE$ equals one if the child is white and non-hispanic, $LOED$ equals one if the care-giver only has a high-school diploma, $HIED$ equals one if the care-giver graduated college, and $DADABS$ equals one if the father does not reside in the household.

As noted, we do not observe $\frac{\theta_{0,t}}{1-\theta_1}$ for any *t* or $\frac{\varphi_0}{1-\varphi_1}$. For any given household we assume $\frac{\varphi_0}{1-\varphi_1}$ is fixed over time. As noted, if we observe $\frac{\theta_{0,t}}{1-\theta_1}$ in the first observation period ($t=\tau$), equation (3) determines this variable in all periods *t* subsequent to τ. Therefore, to generate a likelihood for each household we only need to specify $\frac{\theta_{0,t}}{1-\theta_1}$ in the first observation period τ and $\frac{\varphi_0}{1-\varphi_1}$. We posit that these variables are distributed in our sample according to

$$\textbf{(16)}\qquad \begin{aligned} \frac{\varphi_0}{1-\varphi_1} &= \iota_0 + \iota_1 \ln(W_{\tau,r}) + \iota_2 \ln(M_\tau) + \iota_3 FB + \iota_4(age_\tau - 5) + \iota_5 I(TYP=1) + \varepsilon_\varphi \\ \ln\left(\frac{\theta_{0,\tau}}{1-\theta_1}\right) &= \chi_0 + \chi_1 \ln(W_{\tau,r}) + \chi_2 \ln(M_\tau) + \chi_3 FB + \chi_4(age_\tau - 5) + \chi_5 I(TYP=1) + \varepsilon_\theta, \end{aligned}$$

where $I(.)$ is the identity function (which equals one if the argument inside the parenthesis is true, 0 otherwise) and ε_φ and ε_θ are random variables that are jointly normally distributed with mean zero and covariance matrix Σ_ε, but independently drawn of all other variables in the model. We include ε_φ and ε_θ to account for the fact that these variables vary in unobserved ways across households with the same observable characteristics. As before, FB refers to Demonstration participation, $W_{\tau,r}$ is "true" income in period τ, and M_τ is the 100 minus the CBCL score of the child in the first observation period. The distribution of ε_φ and ε_θ is bounded in such a way to ensure that $7.5 \le \frac{\varphi_0}{1-\varphi_1} \le 11.5$ and $1 \le \frac{\theta_{0,\tau}}{1-\theta_1} \le 99$.

The likelihood for a given individual in our sample that we calculate can be written as

$$\textbf{(17)}\qquad \sum_{k=1}^{100}\left(h(inc_{\tau,k}) \sum_{TYP=1}^{2} Pr(TYP)\, f(\varepsilon_{\varphi,k},\varepsilon_{\theta,k}) \left[\prod_{t=\tau}^{\tau+3} Pr\left(d_t^1,\ldots,d_t^4\right) h(inc_{t+1,k}) g(m_{t+1},w_{t+1})\right]\right)$$

divided by 100, where $h(.)$ is the distribution of the measurement error to income and $f(.)$ denotes the joint distribution of the initial observed endowment of mental health and household ε_φ and ε_θ. A full explanation of how we calculate (17) follows. First, for a given household we draw measurement error on income in the first observation period τ, which determines true income in period τ. Given true income in τ, we know the probability distribution over types from (15), denoted $Pr(TYP)$. Then, repeating for both types, we draw endowment shocks ε_φ and ε_θ from $f(.)$ (and with equation (16) this determines $\frac{\theta_{0,\tau}}{1-\theta_1}$ and $\frac{\varphi_0}{1-\varphi_1}$). Given these variables, household type, and "true" income in the first observation period, the following is repeated for the four periods of the sample: the probability over choices is determined as in (11), a measurement error on future "true" income is drawn, and the density of the implied CBCL and log income shocks is calculated. This whole process is repeated 100 times; the average of the resulting likelihoods is taken as the likelihood of a given household. We multiply the likelihood of all households together to derive the likelihood of our whole sample.

There are some households for which we do not observe the child's CBCL score and/or household income in later waves. When the child's CBCL score in period *t+1* is not observed, we draw m_{t+1} from the appropriate distribution (which generates a simulated CBCL in period t+1) and continue our likelihood calculations. Similarly, when the household income is not observed, we draw w_{t+1} from the appropriate distribution and generate simulated true income in period *t+1*. In the periods for which we simulate either income or CBCL, we do not include the density of m_{t+1} and w_{t+1} in (12), since we already have drawn CBCL and income from the appropriate density.

5. Analysis

We maximize the log-likelihood of our sample using the BHHH Quasi-Newton algorithm, which produces an estimate of the asymptotic covariance matrix of our estimated parameters. Our maximized log-likelihood value is -2575.047. The maximum likelihood estimates of all of the parameters of the model with standard errors are listed in Table 4.

Given the estimates of the type-specific utility function parameters in Table 4, Table 5 displays the range of per-period utility different parents receive when their child does not use mental health services. From Table 5, we see that Type 1 parents with income of $75,000 a year and a child with a CBCL score of 75 will not forego more than 0.424 units of utility for a 5 standard deviation (50 point) one-period improvement of their child's CBCL. Table 5 also shows that Type 2 parents are willing to forego even fewer units of utility for a one period 50 point improvement in their child's CBCL. The average additive utility all parents receive when their child uses outpatient services in a given period is -1.447 (see Table 4). The immediate implication is that on average, parents will not use outpatient services if these services have only one period effects on their child's CBCL score. That said, when administered together, outpatient, intermediate non-residential, and intermediate residential services have the largest estimated temporary effects on a child's symptomatology. Combined, these services lower the child's CBCL by 4.798 points for children of Type 1 parents and 8.655 points for children of Type 2 parents in the period immediately following use.[29] While these effects are rather large in practical terms, they are small relative to the typical CBCL shock. In particular, the standard deviation of the per-period shock to the child's CBCL is estimated to be 9.14 CBCL points, which is almost one standard deviation of the CBCL measure itself.[30]

Over 90% of parents in our sample enroll their children in mental health services in the fist observation period of our data. From our preceding utility function analysis, we infer that parents believe services have multi-period effects. Given the assumption that parents are rational, we accordingly estimate that the permanent effects of outpatient and inpatient services are large and statistically significant. Table 6 displays the percentage reduction of the long-run mean of a child's CBCL ($\frac{\theta_{0,t}}{1-\theta_1}$) resulting from the one-period use of various mental health services. This table shows that outpatient and inpatient services have large and beneficial effects on the long-run mean of a child's CBCL score. For example, a child of Type 2 parents that receives both outpatient and inpatient services reduces $\frac{\theta_{0,t}}{1-\theta_1}$ by almost 20 percent after one period. Table 6 also shows the marginal permanent effects of both non-residential and residential intermediate services are negligible in comparison to the marginal permanent effects of outpatient and inpatient services.

Our model thus fits the observed choice distribution by wave by specifying that parents are forward looking when making mental health choices for children: they enroll children in mental health services because these services have multi-period (permanent) effects on the child's mental health. Simulations show, however, that our model cannot fit the dramatic decrease in service use across waves. We simulate our model for a large sample of people with initial characteristics identical to the characteristics of our first wave of data. The results of this simulation, reported in Table 7, show our model underpredicts the use of services in Wave 1 (63.64% compared to 90.55% in the data) and overpredicts the use of services in Wave 4 (37.41% compared to 24.23% in the data).

Simulations (not reported in a table) also show that the use of services accounts for approximately half of the dramatic decrease in the observed CBCL of children from

Wave 1 to Wave 4. The average child in our sample at Wave 1 has a CBCL of 64.5. Simulations of our model imply the average value of $\frac{\theta_{0,t}}{1-\theta_1}$ of the children in our Wave 1 sample is 53.4. Since 50 is the population mean of the CBCL, at Wave 1 the children in our sample were essentially "normal" and without treatment, the average CBCL score of these children would eventually have returned to 53.4. The use of services, however, lowers the average value of $\frac{\theta_{0,t}}{1-\theta_1}$ of children in our sample from 53.4 in Wave 1 to 42.6 by the start of Wave 4. As a result of this decrease, the average observed CBCL of children falls from 64.5 to 51.7 at the start of Wave 4. In simulations where children receive no mental health services, the average long-run CBCL of children stays fixed at 53.4 and the average CBCL of children falls from 64.5 to only 57.5 at the start of Wave 4.

Our model cannot fit the differences in use and outcomes between Demonstration and Comparison children to any aspect of the Demonstration itself. Our estimates imply that service use is price inelastic, and, the presence of intermediate services does not affect use. Correspondingly, simulations show that the choices of Demonstration parents and the outcomes of Demonstration children were completely unaffected by the Demonstration itself. This result coincides with previous analyses of the Demonstration (Bickman, et. al. (1995)).

Our finding that the use of mental health services of the children in our study is price inelastic contradicts previous research in this area (Leibowitz et. al. (1985), Tsai et. al. (1988), Patrick et. al. (1993)). Simple observation of the pattern of use over time helps verify our result. For both Comparison and Demonstration children, service use falls dramatically over waves, even thought the price of services for a given groups is fixed across waves. It is hard to imagine, then, that prices play much of a role of determining service use of these children (or differences in service use of Comparison and Demonstration children) in any given wave.

Our utility function estimates also help explain this result. As shown in Table 5, in relation to the average disutility parents experience when their child uses services, parental per-period utility does not change much with small changes in income. For example, Type 1 households with a child with CBCL score of 75 earning $10,000 annually (row 1 of Table 5) only get an increase in per-period utility of 0.07 if their annual household income increases to $11,000, all else held the same. Given the large disutility associated with service use and the high variance of the per-period shock to disutility from using services (equal to 0.463 for outpatient services, see Table 4), our utility function estimates directly imply the income elasticity of service use is essentially zero, a result which is verified by simulation analysis.[31] In our model, an increase in prices of services exactly corresponds to an income decrease for parents using services. Since our measured income elasticity is essentially zero, and services are relatively inexpensive ($101.48 for outpatient services and $268.32 for outpatient and inpatient services), it is not surprising that for this sample of children we find service use is independent of price, even for very large changes in prices.

Our estimates of the permanent and temporary effects of intermediate services help explain why the presence of intermediate services does not account for differences in use between Demonstration and Comparison households. Similar to inpatient services, we estimate that parents experience high disutility from their child's use of intermediate non-residential and residential services. Unlike inpatient services, however, intermediate services have negligible permanent effects on a child's mental health for both household types. Given the high disutility and low effectiveness of intermediate services, it is unlikely that the availability of intermediate services affect parents' choices.

Since neither aspect of the Demonstration appears to have affected the use of services, the model fits observed differences of Demonstration and Comparison children and households by specifying that selection into the study was not identical at the Demonstration and Comparison. Although the long-run mean of the (logarithm of) household income ($\frac{\varphi_0}{1-\varphi_1}$) is nearly identical across sites, and the Wave 1 average $\frac{\theta_{0,t}}{1-\theta_1}$ of Demonstration children (53.97) is nearly identical to that of Comparison children (52.28), we estimate that 60.75 percent of Demonstration households are Type 2 households, while only 8.62 percent of Comparison households are Type 2 households. It appears that this last result is not an artifact of "overfitting" the data. Although the type-specific parameters of the model appear to be imprecisely estimated, they are jointly significant: there are 21 type-specific parameters in the model, and a likelihood ratio test (a χ^2_{21} statistic) of a one-type model is rejected at the 1% level.[32]

Conclusions

In this paper, we specify a model where parents care about their child's mental health. We estimate the structural parameters of the model using a procedure that corrects for missing data, measurement error in income, unobservable variation in the child's long-run mental health and (the log of) household income, and distinct differences in the population in preferences and technology. We find that this model can account for both the dramatic decline in the use of services of the children in the study over time, and, the measured improvement of the mental health of children in the study over time.

Our estimates imply that parents receive utility from the mental health of their children and are therefore willing to invest in their child's mental health. The non-monetary cost of this investment, which is the average disutility of enrolling a child in mental health services, is large relative to the monetary costs. This feature of parental utility partially explains two of our results, namely that parents only enroll their children in services if services have multi-period effects on their child's mental health, and, the use of services is price-inelastic. This last result directly conflicts with conclusions drawn from a different experiment, the RAND Health Insurance Experiment (Leibowitz et. al., 1985).

Over 90 percent of parents enroll their children in services in the first wave of data. Given our estimates of parental utility and the assumption that parents are rational, we correspondingly find that outpatient and inpatient services have large and permanent effects on a child's mental health. We simulate that the permanent effects of service use accounts for approximately half of the measured improvement over time of the average CBCL score of children in our sample. We also find the marginal permanent impact of intermediate services are negligible, explaining why simulations show that elimination of intermediate services from the choice set of Demonstration parents does not affect the use of services.

Finally, given the irrelevance of intermediate services and the estimated price inelasticity of service use, we find that neither the presence of intermediate services, nor the fact that outpatient and inpatient services were free, affected the choices of

Demonstration parents or outcomes of Demonstration children. The model therefore attributes differences in choices and outcomes of Demonstration and Comparison households to unobservable aspects of selection into the study: the average type of household in the Demonstration is different than the average type of household in the Comparison. This result, which does not appear to be the result of overfitting the data, casts doubt on either the structure of our model or the usefulness of inference using the quasi-controlled experiment methodology.

Appendix A: Solution Algorithm

As noted in the model section of our paper, for any given household in all periods of the model the state variable vector S_t consists of three elements, $S_t=\left[W_t \quad M_t \quad \theta_{0,t}\right]$. Note that the potential values of all elements of this state space are bounded. M_t (100 minus the child's CBCL score) is bounded between 0 and 100 by construction of the CBCL. We also restrict that true (but unobserved) household income, W_t, be no greater than $100,000. This assumption seems reasonable given that all parents are employed by the U.S. Army. Finally, we redefine S_t as $S_t=\left[\frac{W_t}{1000} \quad M_t \quad \frac{\theta_{0,t}}{1-\theta_1}\right]$ so that the range of all three state variables for all households runs from 0 to 100: the third element of the vector, $\frac{\theta_{0,t}}{1-\theta_1}$, is the mean of the child's CBCL score, and by construction this mean lies between 0 and 100. In solving the model, we augment S_t with $\frac{\varphi_0}{1-\varphi_1}$, the mean of log income. Even though this variable does not change during the decision problem for a given household, it varies across households. With $\frac{\varphi_0}{1-\varphi_1}$ as a state variable, our solution procedure (see below) yields a solution to the model for all households in our sample.

The solution algorithm we use to computationally solve the model, similar to that described in Keane and Wolpin (1994), solves the model recursively. We repeat the following procedure for $FB=1$ (Demonstration) and $FB=0$ (Comparison). Given our definition of utility in the terminal decision period we know $E_T\left[V_T(S_T)\right]=V_T(S_T)$ for all feasible values of the vector S_T; $E_T\left[V_T(S_T)\right]$ is the expected value of having the vector of state variables S_T in period *T*. First, we randomly draw *J* sets of the four period *T-1* state variables, i.e. we randomly draw S_{T-1}^j for $j=1,\ldots,J$. At a given draw of the four variable state space S_{T-1}^j, *K* sets of the m_T and w_T are drawn and equations (3) and (4) are used to calculate the expectation in equation (8),

$$E_{T-1}\left[E_T\left[V_T(S_T)\right] \mid S_{T-1}^{j}, d_{T-1}^{1}, \ldots, d_{T-1}^{4}\right], \tag{18}$$

at all possible combinations of $d_{T-1}^{1} \ldots d_{T-1}^{4}$. This expectation is calculated using Monte-Carlo Integration. Equation (18) gives the expected period *T* value (taken at period *T-1*) of having state space S_{T-1}^{j} and choosing a particular choice combination $d_{T-1}^{1} \ldots d_{T-1}^{4}$.

At this point, *L* sets of period *T-1* utility shocks are drawn. At a given draw of the utility shocks, the optimal set of mental health choices and corresponding value of these optimal choices $V_{T-1}(S_{T-1})$ is calculated using equation (8) for both Demonstration and Comparison children. This process is repeated over all *L* sets of the randomly drawn utility shocks, and $E_{T-1}\left[V_{T-1}(S_{T-1}^{j})\right]$ is calculated via Monte Carlo integration. $E_{T-1}\left[V_{T-1}(S_{T-1}^{j})\right]$ is the period *T-1* expected value of having period *T-1* state variables S_{T-1}^{j}; this expectation is taken before the random values of the period *T-1* utility shocks have been realized.

These steps are repeated at all sets of randomly drawn state variables (all $j=1,\ldots,J$), yielding $E_{T-1}\left[V_{T-1}(S_{T-1}^{j})\right]$ for $j=1,\ldots,J$. This set of expected values are regressed on a quadratic function of the randomly drawn state space, and this regression yields predicted values of $E_{T-1}\left[V_{T-1}(S_{T-1})\right]$ at all S_{T-1}, not just the set of S_{T-1}^{j} that were directly computed. We assume that the value of $E_{T-1}\left[V_{T-1}(S_{T-1})\right]$ predicted by this regression is correct for all feasible S_{T-1}.[33] At this point, *J* sets of the period *T-2* state variables are randomly drawn, and the entire process as just described continues recursively until the first decision period, yielding a full solution to the model.

REFERENCES

ALTONJI, JOSEPH G., FUMIO HAYSHI, AND LAURENCE J. KOTLIKOFF (1997): "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, 105, 1121-1166.

BECKER, GARY S. AND NIGEL TOMES (1976): "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84, S143-S162.

BECKER, GARY S. (1991): *A Treatise on the Family, 2nd Edition.* Cambridge: Harvard University Press.

BEHRMAN, JERE R., ROBERT A. POLLACK, AND PAUL TAUBMAN (1982): "Parental Preferences and Provision for Progeny," *Journal of Political Economy*, 90, 52-73.

BEHRMAN, JERE R. (1997): "Intrahousehold Distribution and the Family," in *The Handbook of Population and Family Economics*, ed. by M.R. Rosenzweig and O. Stark. Elsevier Science.

BICKMAN, L., W. BREDA, E. MICHAEL FOSTER, P.R. GUTHRIE, C.A. HEFLINGER, AND E.W. SUMMERFELT (1995): *Managed Care in Mental Health: The Fort Bragg Experiment.* New York: Plenum.

BROWNING, MARTIN (1992): "Children and Household Economic Behavior," *Journal of Economic Literature*, 30, 1434-1475.

DAVIS, MORRIS A. (1998): "The Health and Financial Decisions of the Elderly." Ph.D. Dissertation, University of Pennsylvania.

ECKSTEIN, ZVI, AND WOLPIN, KENNETH I. (1997): "Youth Employment and Academic Performance in High School," mimeo, March.

FOSTER, E. MICHAEL (1998): "Dose-Response of Outpatient Therapy for Children and Adolescents: Instrumental Variables Estimates," mimeo, April.

FOSTER, E. MICHAEL, SUMMERFELT, WM. THOMAS, AND SAUNDERS, ROBERT C. (1996): "Mental Health Service Utilization in a Children's Mental Health Managed Care Demonstration." *The Journal of Mental Health Administration*, 23, 80-91.

JENSEN, P.S., K. HOAGWOOD, AND T. PETTI (1996): "Outcomes of Mental Health Care for Children and Adolescents: II. Literature Review and Application of a Comprehensive Model," *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 1064-1077.

KEANE, MICHAEL P. AND KENNETH I. WOLPIN (1994): "The Solution and Estimation of Discrete Choice Dynamic Programing Models by Simulation and Interpolation: Monte Carlo Evidence," *Review of Economics and Statistics*, 76, 648-672.

KEANE, MICHAEL P. AND KENNETH I. WOLPIN (1997): "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," mimeo, June.

LAZEAR, EDWARD P. AND ROBERT MICHAEL (1988): *Allocation of Income Within the Household.* Chicago: University of Chicago Press.

LEIBOWITZ, ARLEEN, ET. AL. (1985): "Effect of Cost Sharing on the Use of Medical Services by Children: Interim Results from a Randomized Controlled Trial." *Pediatrics*, 75, 942-951.

PATRICK, C., PADGETT, D.C., BURNS, B.J., SCHELSINGER, H.J., AND J. COHEN (1993): "Use of Inpatient Services by a National Population: Do Benefits Make a Difference?" *Journal of the American Academy of Child and Adolescent*

Psychiatry, 32, 144-154.

PITT, MARK M., MARK R. ROSENZWEIG, AND MD. NAZMUL HASSAN (1990): "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries," *American Economic Review*, 80, 1139-1156.

ROSENZWEIG, MARK R. AND KENNETH I. WOLPIN (1995): "Sisters, Siblings, and Mothers: The Effect of Teen-Age Childbearing on Birth Outcomes in a Dynamic Family Context," *Econometrica*, 63, 303-326.

ROSENZWEIG, MARK R. AND KENNETH I. WOLPIN (1988): "Heterogeneity, Intrafamily Distribution, and Child Health," *Journal of Human Resources*, 23, 437-461.

SHESHINSKI, EYTAN AND YORAM WEISS (1982): "Inequality Within and Between Families," *Journal of Political Economy*, 90, 105-127.

SMYRNIOS, KOSMAS X. AND ROBERT J. KIRKBY (1993): "Long-Term Comparison of Brief Versus Unlimited Psychodynamic Treatments With Children and Their Parents," *Journal of Consulting and Clinical Psychology*, 61, 1020-1027.

STROUL, B. AND R. FRIEDMAN (1986): *A System of Care for Children and Youth with Severe Emotional Disturbances*. Washington, DC: Georgetown University Child Development Center, National Technical Assistance Center for Children's Mental Health.

TARGET, MARY, AND FONAGY, PETER (1994): "The Efficacy of Psychoanalysis for Children: Prediction of Outcome in a Developmental Context." *Journal of American Academic Child and Adolescent Psychiatry*, 33, 1134-1144.

TSAI, S.P., REEDY, S.M., BERNACKI, E.J., AND E.S. LEE (1988): "Effects of Curtailed

Insurance Benefits on Use of Mental Health Care." *Medical Care*, 26, 430-440.

WEISZ, J.R., G.R. DONENBERG, S.S. HAN, AND D. KAUNECKIS (1995): "Child and Adolescent Psychotherapy Outcomes in Experiments Versus Clinics: Why the Disparity?" *Journal of Abnormal Child Psychology*, 23, 83-106.

1. See Browning (1992) and Behrman (1997) for two different reviews of this literature.

2. The Army implemented the Continuum of Care philosophy starting in 1990 order to reduce costs, the idea being that parents would replace inpatient therapy for their children, which is costly to provide, with some form of intermediate therapy provided at the Demonstration, which was thought to be less costly to provide and potentially more effective (Bickman, et. al. (1995)).

3. For a full literature review of the estimated impact of mental health services on a child's mental health, see Jensen, Hoagwood, and Petti (1996).

4. Research from the HIE shows that when parents paid no costs, per-child expenditures were 10% greater than when families covered 25% of the costs of care (Leibowitz et. al., (1985)). This result is generally consistent with a small body of research on cost-sharing and the use of mental health services by children and adolescents (Tsai et. al. (1988), Patrick et. al. (1993)).

5. In our data, we rarely observe inpatient services administered without outpatient services. As a result, we impose that outpatient services accompany inpatient services in our model.

6. Our model may apply to children 4 and younger, but we do not have these children in our data set.

7. We specify consumption in thousands of dollars so the two arguments in the parental utility function have approximately the same scale and range.

8. The data section of this paper contains more information on the CBCL and its relation to mental health.

9. When $\beta = 0$, this utility function collapses to $\ln(C_t) + \alpha \ln(100 - M_t)$.

10. Note that this composite has constant elasticity of substitution and constant returns to scale.

11. 58.15 is the average CBCL score in the sample ($\overline{CBCL}$) and 12.15 is the average age of children in our sample ($\overline{age}$).

12. CHAMPUS has a yearly deductible, so in reality parents face copayments only if their yearly expenditures on mental health services are greater than the deductible. We ignore this deductible in our model.

13. For reasons discussed in the likelihood section, it is useful to consider φ_0 as a state variable, even though its value does not change during the decision horizon of any given household.

14. Following convention, the period *t* shocks to marginal utility are not explicitly included in S_t, even though their values are used to determine the set of optimal decisions.

15. In future work, we will estimate the discount factor.

16. Although these restrictions may seem severe, they keep the computation of the model feasible and do not require estimation of additional parameters.

17. Individuals ages 4 and under were excluded from the study.

18. "Symptomatology" involves counts of behaviors or feelings that serve as signs or indicators of emotional and behavioral problems.

19. The first category does include a few high school dropouts, but these are few in number because of education requirements for military personnel.

20. We recognize that our assumption that one “dose” of mental health services has the same effect as multiple doses for all mental health services is quite severe (see Foster (1998)). That said, we believe an interesting (but complicated) extension of our model

would allow parents to choose the dose of services their child receives.

21. We drop one Comparison group child from our sample who appeared to use intermediate services. We also drop nine children from our sample who use either inpatient or intermediate services without also using outpatient services.

22. Seven of these children drop out of our sample once they turn 18.

23. The full extent of the reduction in the CBCL, however, is likely understated in Table 3 because the availability of a CBCL score after Wave 1 is correlated with a child's use of services. Wave 4 CBCL scores are recorded for 69% of those children that had mental health therapy between Wave 3 and Wave 4, while only 52% of those that had no therapy between Wave 3 and Wave 4 have Wave 4 CBCL scores reported.

24. This likelihood incorporates the assumption that the utility shocks are realized before and independent of the log income and CBCL shocks.

25. Although we observe the use of services between waves 4 and 5 (period $\tau+3$), in our data we do not observe CBCL scores and household income in the fifth wave (period $\tau+4$). For simplicity, we ignore this detail in exposition of the likelihood.

26. Similar to income, we also include measurement error on out-of-pocket costs of outpatient and inpatient services for Comparison parents in the likelihood.

27. When we simulate true income for our likelihood calculations (described later in this section), we restrict the distribution of inc_t by imposing that true household income be no larger than $100,000.

28. The rest of the parameters of the model are assumed to be the same for both types of households in the population.

29. The child's CBCL does not immediately revert to its mean, so temporary effects last

longer than one period.

30. As an aside, this implies that either the CBCL is not a good measure of behavior, or, across six month periods a child's behavior is subject to large random variation.

31. The simulated income elasticity depends on whether parents know the change in income is permanent or temporary. The elasticity of permanent income changes is larger than that of temporary income changes, however both elasticities are essentially zero.

32. The likelihood ratio test value is 48.55.

33. See Keane and Wolpin (1994) for a justification of this assumption.

TABLE 1

PARENTS' SET OF MENTAL HEALTH CHOICES

Choice Set	$\left(d_t^1, d_t^2, d_t^3, d_t^4\right)$	Outpatient	Inpatient	Intermediate Non-Residential	Intermediate Residential
	(0,0,0,0)	No	No	No	No
$D(0)$	(1,0,0,0)	Yes	No	No	No
	(1,1,0,0)	Yes	Yes	No	No
	(0,0,0,0)	No	No	No	No
	(1,0,0,0)	Yes	No	No	No
	(1,1,0,0)	Yes	Yes	No	No
	(1,0,1,0)	Yes	No	Yes	No
$D(1)$	(1,1,1,0)	Yes	Yes	Yes	No
	(1,0,0,1)	Yes	No	No	Yes
	(1,1,0,1)	Yes	Yes	No	Yes
	(1,0,1,1)	Yes	No	Yes	Yes
	(1,1,1,1)	Yes	Yes	Yes	Yes

TABLE 2

CHARACTERISTICS OF INITIAL WAVE OF DATA

Wave 1 Characteristic		Full Sample 201 children	Demonstration 137 children	Comparison 64 children
Percent White		70.6	69.3	73.4
Percent Male		59.2	55.5	67.2
Caretaker: Some College[1]		84.6	88.3	76.6
Caretaker: College Grad		26.9	28.5	23.4
Percent Father not in HH[2]		51.2	49.6	54.7
Income: Median		$25,000	$25,000	$25,000
Income: 25th Percentile		$17,500	$25,000	$17,500
Income: 75th Percentile		$35,000	$35,000	$35,000
Child CBCL: Mean (Standard Deviation)		64.6 (9.9)	64.5 (10.1)	64.7 (9.6)
Percent Using Services:	Outpatient	90.5	94.2	82.8
	Inpatient	19.9	12.4	35.9
	Intermediate Non-Residential	10.0	14.6	0.0
	Intermediate Residential	7.5	10.9	0.0

1. The percentage of households where the primary care-giver has had some college education.
2. The percentage of households where the father is not a household member.

TABLE 3

CBCL AND MENTAL HEALTH SERVICE USE BY SITE AND WAVE[1,2]

Variable	Site	Wave 1	Wave 4
Average CBCL	Demonstration	64.5	53.3
	Comparison	64.7	54.1
Percent Use No Services	Demonstration	5.8	72.4
	Comparison	17.2	83.3
Percent Use Outpatient	Demonstration	94.2	27.6
	Comparison	82.8	16.7
Percent Use Inpatient	Demonstration	12.4	1.5
	Comparison	35.9	3.3
Percent Use Intermediate Non-Residential	Demonstration	14.6	3.7
	Comparison	0.0	0.0
Percent Use Intermediate Residential	Demonstration	10.9	3.7
	Comparison	0.0	0.0

1. CBCL scores and service use were recorded for 137 Demonstration children and 64 comparison children in our Wave 1 sample. Service use was recorded for 134 Demonstration children and 60 Comparison children in our Wave 4 sample while CBCL scores were recorded for only 73 Demonstration children and 39 Comparison children in our Wave 4 sample.

2. The CBCL score in the Wave x column refers to the child's CBCL at the Wave x interview. The percent that use services in the Wave x column refers to the percent of children that use services between Wave x and Wave x+1.

TABLE 4

MODEL PARAMETER ESTIMATES, EQUATIONS (1) AND (2)

Variable		Estimate	SE
Eqn. (1) - Utility Function Parameters			
α (type 1)		0.660	0.209
β (type 1)		-0.357	0.279
α (type 2)		0.187	0.158
β (type 2)		0.062	0.525
Eqn. (2) - Marginal Utility of Choices Parameters [1]			
$\bar{b}^1$	average outpatient utility	-1.447	0.319
$\bar{b}^2$	average inpatient utility	-1.834	0.481
$\bar{b}^3$	average int. non-residential utility	-1.698	0.561
$\bar{b}^4$	average int. residential utility	-1.516	1.036
$\mathrm{Var}(e^1)$	variance of outpatient utility	0.463	0.205
$\mathrm{Var}(e^2)$	variance of inpatient utility	0.462	0.239
$\mathrm{Var}(e^3)$	variance of int. non-residential utility	1.121	0.796
$\mathrm{Var}(e^4)$	variance of int. residential utility	0.881	1.042
$\mathrm{Cov}(e^1,e^2)$		-0.071	0.105
$\mathrm{Cov}(e^1,e^3)$		0.131	0.108
$\mathrm{Cov}(e^2,e^3)$		0.160	0.084
$\mathrm{Cov}(e^1,e^4)$		-0.315	0.311
$\mathrm{Cov}(e^2,e^4)$		0.245	0.221
$\mathrm{Cov}(e^3,e^4)$		0.498	0.414

1. Inpatient and intermediate services are never administered without outpatient services. Thus, parents with children using outpatient and inpatient services receive an average utility of -1.447 + -1.834 = -3.281, etc...

TABLE 4

MODEL PARAMETER ESTIMATES (CONTD.)

EQUATION (3): CBCL PROCESS PARAMETER ESTIMATES

Variable		Estimate	SE
θ_1	lag CBCL	0.732	0.028
Temporary Effects of Services, Type 1			
θ_2^1	outpatient	-2.025	0.661
θ_2^2	inpatient	1.886	1.258
θ_2^3	intermediate non-residential	-1.567	1.500
θ_2^4	intermediate residential	-1.206	1.798
Permanent Effects of Services, Type 1			
θ_3^1	outpatient	-0.075	0.017
θ_3^2	inpatient	-0.101	0.026
θ_3^3	intermediate non-residential	-0.034	0.030
θ_3^4	intermediate residential	-0.005	0.018
Temporary Effects of Services, Type 2			
θ_2^1	outpatient	-2.213	1.219
θ_2^2	inpatient	3.028	3.316
θ_2^3	intermediate non-residential	-3.576	2.544
θ_2^4	intermediate residential	-2.876	3.645
Permanent Effects of Services, Type 2			
θ_3^1	outpatient	-0.144	0.048
θ_3^2	inpatient	-0.052	0.040
θ_3^3	intermediate non-residential	0.012	0.024
θ_3^4	intermediate residential	0.015	0.025
$\mathrm{Var}(m_{t+1})$	variance of CBCL shock	83.454	5.386

TABLE 4

MODEL PARAMETER ESTIMATES (CONTD.), EQUATIONS (4), (5), AND (13)

Variable		Estimate	SE
Eqn. (4) - Log Income Parameters			
φ_1	lag log income	0.872	0.023
φ_2	CBCL effects	-0.003	0.000
φ_3	age effects	0.008	0.010
φ_4^1	effect of outpatient services	-0.003	0.007
φ_4^2	effect of inpatient services	0.005	0.022
φ_4^3	effect of int. non-residential services	-0.009	0.018
φ_4^4	effect of int. residential services	0.000	0.029
$\mathrm{Var}(w_{t+1})$	variance log income shock	0.103	0.024
$\mathrm{Cov}(m_{t+1}, w_{t+1})$	covariance of CBCL, log income shock	-0.141	0.371
Eqn. (5) - Budget Constraint Parameters			
p^1	price of outpatient services	$101.48	$15.86
p^2	price of inpatient services	$166.84	$39.24
Eqn. (13) - Measurement Error Parameters			
σ_{inc}^2	variance of log income msmt. error	0.032	0.017
σ_p^2	variance of log prices msmt. error	1.133	0.230

TABLE 4

MODEL PARAMETER ESTIMATES (CONTD.), EQUATION (15)

	Variable	Estimate	SE
	Eqn. (15) - Type Parameters		
ζ_0	constant	-1.366	17.008
ζ_1	wave 1 “true” income	0.316	1.084
ζ_2	wave 1 CBCL score	-0.517	3.459
ζ_3	Demonstration status	-3.406	1.587
ζ_4	wave 1 age	0.106	0.182
ζ_5	care-giver white dummy	1.809	1.180
ζ_6	care-giver lo-education dummy	-0.029	1.755
ζ_7	care-giver high-education dummy	1.519	1.120
ζ_8	father absent dummy	1.530	1.151

TABLE 4

MODEL PARAMETER ESTIMATES (CONTD.), EQUATIONS (16) AND (17)

Variable		Estimate	SE
Eqn. (16) - Income Endowment Parameters			
ι_0	constant	1.470	4.954
ι_1	wave 1 "true" income	0.915	0.378
ι_2	wave 1 CBCL score	0.428	1.167
ι_3	Demonstration status	-1.227	0.663
ι_4	wave 1 age	-0.064	0.076
ι_5	Type=1 dummy	-1.611	0.846
$Var(\varepsilon_\varphi)$	variance of unobserved component	0.359	0.570
Eqn. (16) - Mental Health Endowment Parameters			
χ_0	constant	-0.089	1.091
χ_1	wave 1 "true" income	0.062	0.072
χ_2	wave 1 CBCL score	0.780	0.203
χ_3	Demonstration status	-0.004	0.109
χ_4	wave 1 age	0.032	0.015
χ_5	Type=1 dummy	-0.113	0.169
$Var(\varepsilon_\theta)$	variance of unobserved component	0.063	0.027
$Cov(\varepsilon_\varphi,\varepsilon_\theta)$	covariance of unobserved components	-0.060	0.048

TABLE 5

CONSUMPTION, MENTAL HEALTH, AND PARENTAL UTILITY[1], BY TYPE

	Household Consumption[3] in \$000s	Mental Health of Child[2] M_t=25.0	M_t=75.0
Type 1	C_t=5.0	0.724	0.982
	C_t=37.5	2.036	2.460
Type 2	C_t=5.0	4.639	4.832
	C_t=37.5	6.326	6.501

1. Parental utility is defined as $(\beta^{-1})\log\left(C_t^{\beta}+\alpha M_t^{\beta}\right)$.

2. The mental health of a child is defined as as 100 minus the child's CBCL score.

3. When the child does not use mental health services, household consumption equals one-half of annual household income.

TABLE 6

THE PERMANENT EFFECTS OF VARIOUS MENTAL HEALTH SERVICES, BY TYPE

Type	Mental Health Service	Effect on $(1-\theta_1)^{-1}\theta_{0,t}$
Type 1	Outpatient	7.47% decrease
	Outpatient and Inpatient	17.58% decrease
	Outpatient and Intermediate Non-Residential	10.87% decrease
	Outpatient and Intermediate Residential	7.97% decrease
Type 2	Outpatient	14.40% decrease
	Outpatient and Inpatient	19.56% decrease
	Outpatient and Intermediate Non-Residential	13.20% decrease
	Outpatient and Intermediate Residential	12.90% decrease

TABLE 7

CHOICE DISTRIBUTION BY WAVE, ACTUAL VS. PREDICTED (IN PARENTHESIS)

	Wave[1]			
	1	2	3	4
no services	9.45 (36.36)	42.71 (45.78)	68.88 (55.20)	75.77 (62.59)
outpatient only	63.68 (45.54)	43.22 (40.98)	23.47 (35.23)	18.04 (29.19)
outpatient and inpatient	13.93 (8.40)	3.02 (4.57)	2.55 (2.38)	1.55 (1.65)
outpatient and non-residential	3.48 (4.40)	5.03 (4.30)	1.53 (3.90)	1.55 (3.61)
outpatient, inpatient, and non-residential	1.99 (1.10)	0.50 (0.62)	1.02 (0.34)	0.52 (0.27)
outpatient and residential	1.99 (1.39)	1.51 (1.38)	0.51 (1.00)	2.06 (0.86)
outpatient, inpatient, and residential	1.00 (0.42)	1.01 (0.21)	0.51 (0.16)	0.00 (0.10)
outpatient, non-residential, and residential	1.49 (1.87)	1.01 (1.82)	1.53 (1.60)	0.52 (1.60)
outpatient, inpatient, non-residential, and residential	2.99 (0.54)	2.01 (0.34)	1.02 (0.19)	0.52 (0.14)

1. Refers to use in the period between Wave x and Wave x+1.