Dollars and performance: treating alcohol misuse in Maine

Matilde P. Machado*

Department of Economics, Universidad Carlos III de Madrid and Institut d'Analisis Economico (CSIC), Calle Madrid 126, Getafe, 28903 Madrid, Spain

Abstract

If public funds are allocated efficiently, then an increase in funding should improve the performance of substance abuse treatment programs. In the data used in this paper, performance (measured as abstinence rates) and expenditures per patient are not positively correlated. One explanation is that funding is endogeneous, i.e. programs treating more difficult patients receive more funding. The data comes from all Maine's outpatient drug-free programs that received public funding between 1991 and 1994. After controlling for endogeneity, this paper concludes that the marginal impact of expenditures per patient on abstinence rates is small and statistically insignificantly different from zero.

JEL classification: I12; H51; L3

Keywords: Alcohol; Substance abuse; Non-profit; Performance; Cost-effectiveness

1. Introduction

Substance abuse treatment started receiving major funding from the US Federal Government around 1965 during the opioid epidemic of the 1960s. Since then, state governments as well as private sources have also contributed to the financing of these programs. Estimates for 1985 [1] say that 64% of the direct health costs from drug abuse was supported by state, federal, or local funds. With the Anti-Drug Abuse Act of 1986 the Federal Government increased its funding for substance abuse treatment substantially. [2] Despite the growth in

* Tel.: +34-91-624-9571; fax: +34-91-624-9875.
E-mail address: mmachado@eco.uc3m.es (M.P. Machado).

[1] See Rice et al. (1991).

[2] In the fiscal year 1990, for example, the Federal Government increased block grants for substance abuse treatment by 50%.

public funding, little is known about the cost-effectiveness of these programs. In this paper, the marginal impact of public funds on the abstinence rate of substance abuse (mainly alcohol abuse) outpatient drug-free treatment programs in the state of Maine from 1991 to 1994 is estimated.

The main motivation for this paper comes from the puzzling observation that the unconditional correlation between performance of treatment programs and expenditures per patient is non-positive in our dataset. If this anecdotal evidence proves to be generally true, the state of Maine would be better off transferring funds from alcohol abuse treatment [3] to other social programs.

A closer look at the allocation process of state and federal funds across providers of alcohol abuse treatment in Maine suggests that expenditures per patient are potentially endogenous. Funds are allocated in a centralized fashion by the Maine Office of Substance Abuse (OSA). Moreover, OSA collects very detailed information on patients' characteristics and performance at the patient level for every treatment agency. The use of this information in OSA's allocation decision is the potential source of endogeneity and may explain the non-positive unconditional correlation between expenditures per patient and treatment outcomes. For instance, if OSA allocates more funds per patient to those programs with more difficult patients, and we do not control for patient characteristics, then the estimated impact of funds on outcomes would be biased downwards. In this paper, instrumental variables (IV) are used to avoid this potential endogeneity bias.

The literature on treatment effectiveness has presented evidence (although many studies suffer methodological flaws [4]) that patients tend to do better after treatment. McLellan et al. (1997) in a review of effectiveness studies conclude that the "data reviewed (from controlled clinical trials or real-world settings) indicate that substance-abuse patients show major reduction in their alcohol and drug use following their treatment" as well as improvement in medical, psychological functioning, and other components of quality of life. From the point of view of a policy maker who has an interest in getting the most out of every dollar spent, effectiveness of alcohol abuse treatment is certainly a necessary condition for continuing public funding. Nevertheless, as McLellan et al. point out, their review "has also shown substantial variability in effectiveness of substance abuse treatment across different settings, modalities, and programs. Put simply, not all programs are effective".

Even if treatment, in general, proves to be effective, there is still scope for waste in the way public funds are being used. Waste may be caused by a bad allocation of public funds: programs that are not effective being funded, or effective programs being over-funded and spending the excess dollars in activities that do not contribute to the patients' recovery.

Cost-effectiveness studies, relating funds and outcomes, are rare, and do not always use multivariate regression analysis. The effect of funds on the performance of substance abuse treatment has been typically linked to the debate surrounding the choice of a less expensive outpatient treatment versus a more expensive inpatient treatment (Long et al., 1998; Walsh et al., 1991; Longabaugh et al., 1983). This paper sets out to analyze exclusively outpatient

[3] From this point onwards the terminology "alcohol abuse" instead of "substance abuse" given that around 90% of our sample are alcohol abusers is adopted.

[4] Apsler and Harding (1991) illustrate methodological mistakes using as an example the well-known Drug Abuse Reporting Program (DARP) study by Sells and Simpson (1976).

treatment programs, and estimate the marginal impact of expenditures per patient on their abstinence rate.

It is perhaps useful to draw some evidence of the relationship between performance and funding from other social programs. In particular, the earlier construction of standardized performance measures for students in the US has allowed a large number of such studies in the education field (Hanushek, 1986; Hanushek et al., 1994). Aggregate data show that spending per pupil has been growing steadily since 1890, while average student performance has not increased and has actually worsened since 1967 (Hanushek et al., 1994). The study *Equality of Educational Opportunity* (Coleman et al., 1966) concluded that student background was the major predictor of student performance while expenditures per student as well as other school inputs had no measurable impact. Researchers have not yet reached a consensus regarding the effect of funds on performance due to differences in the degree of data aggregation (Summers and Wolfe, 1977; Hanushek et al., 1996), the measurement of teacher's quality (Ferguson, 1991; Hanushek, 1986), and the treatment of endogenous expenditures (Lang and Somanathan, 1997). Like public schools, non-profit providers of alcohol abuse treatment are heavily financed by the federal and state governments, and are given great discretion in the use of their funding.

This paper's main conclusion is that the marginal impact of expenditures per patient is economically small and statistically not different from zero. The most optimistic estimation implies that, on an average, it would cost US$ 615,801.80 to produce an extra abstinent patient in Maine.

The paper is organized as follows. Section 2 characterizes the background of the study, Section 3 derives the estimation model, Section 4 describes the dataset, Section 5 shows the estimation results, Section 6 discusses the assumptions and caveats and finally, Section 7 concludes.

2. Maine's substance abuse treatment system

In 1988 the state of Maine spent US$ 7,304,928.00 on substance abuse treatment programs (mainly alcohol abuse) and this total has grown to US$ 10,085,716.00 by 1995. The size of these numbers and the fact that close to 1% of the state's population has received treatment from public programs has led the authorities to adopt two important measures to monitor the supply of treatment services. Firstly, in October 1989, the Maine Addiction and Treatment System (MATS) was introduced. Secondly, in fiscal year 1993, performance incentives were introduced in the contracts signed with treatment agencies. Both measures have received wide attention from state authorities throughout the US.

MATS consists of the requirement for all agencies receiving any public funding to fill out a standard admission and discharge form for each person who is treated for substance abuse related problems and to submit these forms to the relevant authority (after July 1990, the Office of Substance Abuse (OSA)). These forms give detailed information about each treatment program at the patient level. [5]

[5] See Commons et al. (1997) or Commons et al. (1994) for a description.

In July 1991,[6] OSA added performance standards to its contracts. Programs were supposed to meet these standards although no penalty existed in case of non-compliance. In the fiscal year 1993, however, OSA changed its contracts in an attempt to introduce incentives. These new contracts stated that "allocation of resources for the (next) contract year may be affected by agency performance in the previous year".[7] According to Commons and McGuire (1997), "to date, OSA has not found it possible to reward good performers by allowing them to retain surplus funds, due to legislative reduction in appropriations occurring as a result of Maine's current economic situation". Occasionally, however, good performers received OSA's recommendation for additional federal funding and poor performers were put on a fee-for-service and/or shorter contracts. These compensations and penalties were not explicitly stated in the contract and therefore we make the simplifying assumption that implicit incentives are non existent.[8]

OSA measured program performance in three categories: efficiency, effectiveness, and special populations. Efficiency was related to the degree of compliance in the provision of the amount of services contracted with OSA. Effectiveness and special populations were disaggregated into a set of specific indicators with their own standards and were measured for primary clients only.[9] Effectiveness intended to measure the quality of the services provided. Its indicators were obtained from the comparison of discharge and admission MATS data on every single patient and aggregated at the program-quarter level. Outpatient treatment programs had to satisfy at least 8 out of 12 indicators in order to perform on Effectiveness. The "percentage of abstinent patients 30 days prior to discharge" is one of these indicators which should reach at least 70%.[10] Finally, special populations intended to control for patient selection by guaranteeing that certain groups in the society (e.g. the elderly) received a given share of treatment.

Besides managing the data collection process, OSA was responsible for the allocation of state appropriations and federal block grants across non-profit substance abuse treatment agencies. The allocation of funds before PBC followed a historical pattern (Commons et al., 1997) meaning that each year's allocation was based on the previous year's. Nevertheless, there is no strong evidence that PBC changed the historical rules of funding allocation dramatically (Commons and McGuire, 1997). OSA's contract with the agencies established the amount OSA was assigning and its distribution across different pro-

[6] Beginning of fiscal year 1992. From now on all references to a year mean fiscal year.

[7] This is called "performance-based contracting" (PBC).

[8] Commons et al. (1997) run linear regressions of abstinence rates and a few covariates, among them a dummy variable that takes value 0 before PBC was first introduced and value 1 afterwards. The authors find a positive and significant impact of the PBC dummy although they recognize that the PBC effect cannot be distinguished from a shift in the trend. We have introduced time dummies as explanatory variables in our estimations and we concluded that, after controlling for a much wider and different vector of program and discharged patient characteristics (including fixed effects) and restricting ourselves to outpatient programs, the impact of time dummies on abstinence rates was not significantly different from zero. The results of these "informal tests" are not presented in this paper.

[9] Clients are either primary clients or "patients" (the ones who need treatment) or co-dependents of primary clients.

[10] Notice that patients may be discharged either because they have completed treatment, or because they are referred to another treatment program, or even because they dropped out of treatment.

grams (e.g. detoxification, outpatient treatment, and residential treatment) within the same contract.[11]

The effort put into reforming the supply of alcohol abuse treatment shows the importance that the state of Maine attaches to these social programs, and the relevance of the question that this paper addresses, i.e. the impact of the marginal dollar on the performance of treatment programs.

3. The estimation model

We have in mind a model where the unit of analysis is the outpatient program i, $i = 1, \ldots, n$. Every quarter t, program i provides treatment to N_{it} patients. At the end of the quarter, D_{it} ($0 \leq D_{it} \leq N_{it}$) patients are discharged from treatment. We assume that all the patients discharged from program i in quarter t have the same probability p_{it} of performing successfully. We model p_{it} as a logistic function of expenditures per patient, $\bar{c}_{it}$, program and average discharged patient characteristics represented by the vector z'_{it}, and managerial effort or clinic specific productivity m_i, as follows:

$$p_{it} = \frac{\exp(\vartheta + \alpha\bar{c}_{it} + z'_{it}\beta + \varsigma m_i)}{1 + \exp(\vartheta + \alpha\bar{c}_{it} + z'_{it}\beta + \varsigma m_i)}. \tag{1}$$

In Eq. (1) ϑ is a constant term, and the total expenditures per patient $\bar{c}_{it} \equiv c_{it}/N_{it} \equiv x_{it}/N_{it} + o_{it}/N_{it}$ come partly from OSA's contribution (x_{it}), and partly from third party donations (o_{it}).

The assumption of a common probability of success for discharged patients implies that the total number of successes from program i in period t (Y_{it}) follows a binomial distribution with parameters (p_{it}, D_{it}). The range of values for $Y_{it} \in [0, D_{it}]$ is consistent with the frequency of *zero successes* and *all successes* that we observe in the data. In short, we interpret p_{it} as a reduced form production function where the output (the realized number of abstinent discharges) follows a random process and funds are used optimally in the acquisition of inputs, such as qualified staff.

From the consistent estimation of Eq. (1) we derive an estimate of the marginal impact of funds per capita on performance, given by

$$\left(\widehat{\frac{\partial p_{it}}{\partial \bar{c}_{it}}}\right) = \hat{\alpha}\hat{p}_{it}(1 - \hat{p}_{it}) \in [0, 0.25\hat{\alpha}] \quad \text{for} \quad \hat{\alpha} > 0. \tag{2}$$

In order to obtain consistent parameter estimates we must deal with the possibility of omitted variables that cause $\bar{c}_{it}$ to be endogenous. There are two good reasons to suspect that funding may be endogenous. First, funding decisions are made in a centralized fashion by a well informed decision maker, OSA. And second, as Figs. 1 and 2 show,[12] the unconditional correlation between expenditures per patient and abstinence rates is non-positive in the data.

[11] Interestingly the state does not support methadone maintenance programs despite their relative success (see Ball and Ross, 1991). Each program within the same contract is separately evaluated in terms of performance.

[12] In these pictures we use two alternative definitions of expenditures per patient defined later in the paper.

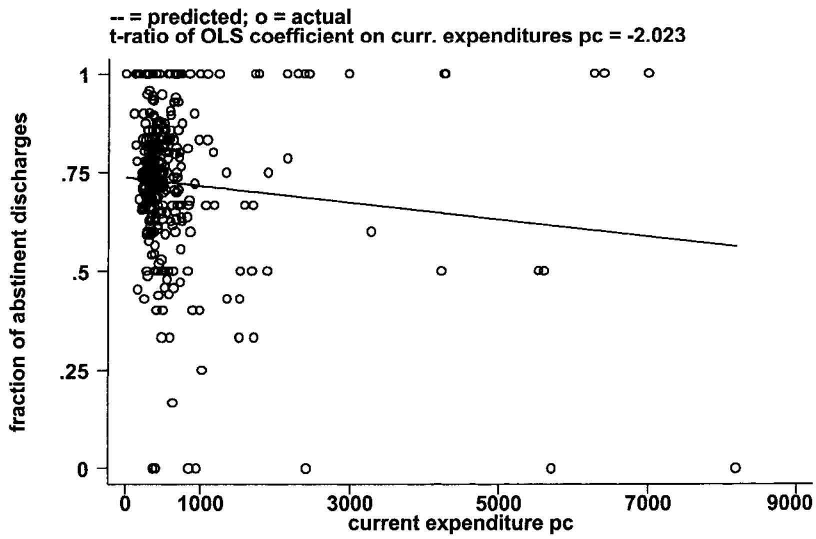

Fig. 1. Current expenditures per patient vs. abstinence rate.

This somewhat surprising empirical evidence suggests that OSA may be using funding allocations to compensate agencies with hard-to-treat patients. If this is indeed OSA's policy, we expect to observe a positive impact of funding on performance once we control for funding endogeneity.

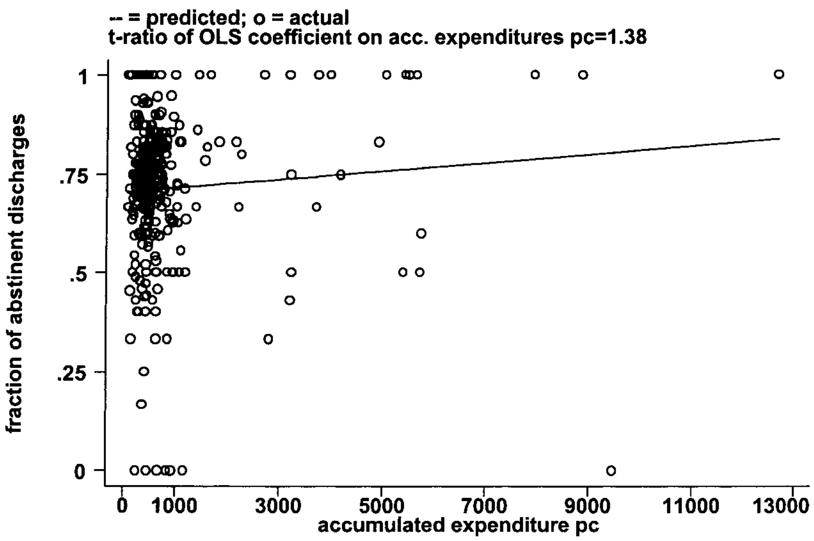

Fig. 2. Accumulated expenditures per patient vs. abstinence rate.

Eq. (1) can be estimated following two approaches: a non-linear maximum likelihood estimation based on the binomial distribution [13] and a non-maximum likelihood estimation of a linear version of Eq. (1).

A non-linear maximum likelihood estimation presents two problems. First, while the managerial effort m_i is not observable to the econometrician, it is likely to be observed by OSA officials due to a long-term relationship with these agencies. OSA is likely to take m_i into consideration when deciding on the allocation to agency i, in which case $\bar{c}_{it}$ is correlated with m_i and therefore, an endogenous variable. The omission of relevant variables correlated with $\bar{c}_{it}$ such as m_i, would yield biased estimates of α. Second, we might think that the problem of unobserved managerial effort m_i is solved with the introduction of fixed effects in its place. The introduction of fixed effects in a non-linear model, however, causes inconsistent estimates. Consistency requires the number of observations per agency to grow to infinity. [14] Hence, fixed effects may eliminate the endogeneity bias at the expense of an estimation bias.

A linear model has the advantage of both avoiding the inconsistency caused by the introduction of fixed effects and of dealing with any remaining endogeneity through the use of instrumental variables estimation. Before linearizing Eq. (1) replace the managerial effort variable m_i by a fixed effect parameter k_i to be estimated, and introduce an error term θ_{it} representing all time-variant omitted variables. The probability of success now takes the form

$$p_{it} = \frac{\exp(\vartheta + \alpha\bar{c}_{it} + z'_{it}\beta + k_i + \theta_{it})}{1 + \exp(\vartheta + \alpha\bar{c}_{it} + z'_{it}\beta + k_i + \theta_{it})}. \tag{3}$$

The linearization of Eq. (3) hinges on three basic steps: (1) the use of the equality $y_{it}/d_{it} = p_{it} + u_{it}$, where y_{it}/d_{it} is the realized success rate, and u_{it} is an error term with $E(u_{it}) = 0$, (2) the inversion of the logistic function and (3) a first-order linear approximation around $u_{it} = 0$. The resulting linear equation is the basic equation that we will estimate using quarterly data on abstinence rates per outpatient program, and is given by

$$\log\left(\frac{y_{it}/d_{it}}{1 - y_{it}/d_{it}}\right) = \vartheta + \alpha\bar{c}_{it} + z'_{it}\beta + k_i + \eta_{it}, \tag{4}$$

where η_{it} is a heteroscedastic error term with the following properties:

$$\eta_{it} = \theta_{it} + \frac{u_{it}}{p_{it}(1 - p_{it})}, \qquad E(\eta_{it}) = 0, \tag{5}$$

[13] Assuming that observations are independent across programs and time, the likelihood function is

$$L(\alpha, \beta, \vartheta, k_i | d, y, \bar{c}, z') = \prod_{i=1}^{n}\prod_{t=1}^{Q_i} C_{y_{it}}^{d_{it}} p_{it}^{y_{it}} (1 - p_{it})^{d_{it} - y_{it}},$$

where $C_{y_{it}}^{d_{it}} = d_{it}!/(y_{it}!(d_{it} - y_{it})!)$ is the combinatorial term of the realized number of discharges d_{it} by the realized number of successes y_{it}. Maximum likelihood results, not shown in this paper, can be found in Machado (1997).

[14] Andersen (1973) and Neyman and Scott (1948) show inconsistency for the maximum likelihood logit estimator. Machado (1997, Chapter 7) derives a consistent estimator for the binomial distribution with a logistic probability of success in the presence of fixed effects. The latter estimator, however, demands a lot of computer time and is best used for smaller datasets.

$$\mathrm{var}(\eta_{it}) = \sigma_\theta^2 + \frac{\mathrm{var}(u_{it})}{d_{it}(p_{it}(1-p_{it}))^2} = \sigma_\theta^2 + \frac{p_{it}(1-p_{it})}{d_{it}(p_{it}(1-p_{it}))^2} = \sigma_\theta^2 + \frac{1}{d_{it}p_{it}(1-p_{it})}. \tag{6}$$

Lastly, we must point out that in spite of its advantages over the non-linear model, the linear model (4) is incompatible with 0 or 100% success rates ($y_{it}/d_{it} = \{0,1\}$). In our dataset, y_{it}/d_{it} attains zero in 1.7% of the observations and it attains 1 in 10% of the observations. [15] We apply the standard procedure of approximating the values of 1 with 0.999 and of 0 with 0.001. [16]

Next, we motivate the possibility of endogenous funding with a simple model where a central manager/planner allocates scarce resources to a given number of plants/agencies. The model shows that, under certain reasonable assumptions, a planner would optimally choose to compensate agencies with lower exogenous productivity by increasing their amount of funding.

3.1. A simple model of optimal allocation and implied endogeneity

Consider the problem of the optimal allocation of a fixed budget X by a central manager, across n plants. Assume the manager has complete information about the plants' characteristics and the characteristics of the production stochastic process.

Each plant i, $i = 1, \dots, n$ is characterized by three parameters N_i, ϕ_i, and θ_i all known to the manager. Plant i disposes of N_i homogeneous units of a quasi-fixed input from which it produces D_i units of output. N_i is exogenously given at price zero. [17] ϕ_i denotes the gross output-to-input ratio ($\phi_i = D_i/N_i$) which is plant specific. θ_i is a parameter that characterizes the probability that each of the D_i units produced is of minimum quality in which case it is considered a success. The budget X is interpreted as a variable input to the production function, which is allocated by the central manager as a function of N_i, ϕ_i, and θ_i.

The only source of uncertainty in the model is the number of successes. An example of this production technology is a batch production process, such as silicon chips. The size of the batch is the gross output given by D_i, from which a fraction will be defective at the end of the production process. Call the random variable $Y_i \in [0, D_i]$ the number of successes (net output) from plant i. We assume that Y_i follows a binomial distribution with parameters p_i and D_i. p_i represents the probability of success and is a function of the input

[15] These percentages are 3.7 and 16%, respectively, for the unclean dataset, that is before we drop observations with omitted values for variables in the estimations.

[16] Cox (1970, p. 33) shows that

$$E\left[\log\left(\frac{y_{it}/d_{it} + (2d_{it})^{-1}}{1 - y_{it}/d_{it} + (2d_{it})^{-1}}\right)\right] - \log\left(\frac{p_{it}}{1-p_{it}}\right) = o(d_{it}^{-1}).$$

We tried using this transformation but we do not report the results since the correlation between the adjustment term $(2d_{it})^{-1}$ and the set of instrumental variables used, produces invalid results.

[17] The simplest interpretation for N_i is a natural resource which supply differs by location and cannot be varied in the short run. This source of heterogeneity across plants has an obvious parallel in the OSA's optimizing problem with the number of patients at each agency.

ratio, $\bar{x}_i = x_i/N_i$, and the efficiency parameter θ_i,

$$p_i = f(\bar{x}_i, \theta_i). \tag{7}$$

The probability of success is non-decreasing in both $\bar{x}_i$ and θ_i ($f_{\bar{x}}(\bar{x}_i, \theta_i) \geq 0$, $f_\theta(\bar{x}_i, \theta_i) \geq 0$), and is concave in $\bar{x}_i$. This technology exhibits constant returns to scale on the two inputs x_i and N_i. If x_i and N_i are doubled, then D_i is also doubled, implying that $E(Y_i) = p_i D_i$ also doubles.

Finally, we assume that the manager's optimization problem is to allocate the budget across plants so as to maximize the total expected net output in the firm, which can be formalized as

$$\max_{\{x_1,\dots,x_n\}} E\sum_{i=1}^{n} Y_i \text{ s.t.} \begin{cases} \sum_{i=1}^{n} x_i = X, \quad x_i \geq 0, \\ Y_i \sim B(p_i, D_i), \quad p_i = f\left(\dfrac{x_i}{N_i}, \theta_i\right) = f(\bar{x}_i, \theta_i), \; i = 1, \dots, n, \end{cases} \tag{8}$$

where X, n, f, N_i, θ_i, and D_i for $i = 1, \dots, n$ are all known to the central manager. Since Y_i follows a binomial distribution, we rewrite the maximization problem as

$$\max_{\{x_1,\dots,x_n\}} \sum_{i=1}^{n} D_i f(\bar{x}_i, \theta_i) \quad \text{s.t.} \sum_{i=1}^{n} x_i = X, \quad x_i \geq 0, \; i = 1, \dots, n. \tag{9}$$

Denoting by λ the shadow value of public funds, the first-order condition when $x_i > 0$ is

$$D_i \frac{\partial f(\bar{x}_i, \theta_i)}{\partial x_i} = \lambda \Leftrightarrow \phi_i \frac{\partial f(\bar{x}_i, \theta_i)}{\partial \bar{x}_i} = \lambda. \tag{10}$$

Result 1. Conditional on being positive, the optimal $\bar{x}_i$, is a decreasing function of the efficiency parameter θ_i if $\partial^2 f(\bar{x}_i, \theta_i)/(\partial \bar{x}_i \partial \theta_i) < 0$.

Proof. $\mathrm{d}\bar{x}_i/\mathrm{d}\theta_i = -(f_{\bar{x}\theta}/f_{\bar{x}\bar{x}}) < 0$, trivial application of Roy's identity and second-order condition.[18] □

Result 1 tells us that conditional on receiving positive funding ($\bar{x}_i > 0$) and on the concavity of the probability of success with respect to $\bar{x}_i$, more efficient plants will have a lower input ratio then their more inefficient counterparts if their marginal productivity of success is decreasing in the efficient parameter θ_i.

[18] The result that $\mathrm{d}\bar{x}/\mathrm{d}\theta < 0$ when $f_\theta > 0$, hinges on the assumption that $f_{\bar{x}\theta} < 0$. The same conceptual result of lower resources ($\bar{x}_i$) for those with the highest probability of success also holds if $f_{\bar{x}\theta} > 0$ and $f_\theta < 0$. We restrict the model to these two cases because the theoretical prediction matches what we observe in the data, i.e. a non-positive or even negative correlation between expenditures (or funds) per patient and success rates. To see this, suppose that the opposite is true (w.l.o.g. we analyze only one case), i.e. $f_\theta > 0$ and $f_{\bar{x}\theta} > 0$, this implies that $\mathrm{d}\bar{x}/\mathrm{d}\theta = -(f_{\bar{x}\theta}/f_{\bar{x}\bar{x}}) > 0$ and $\mathrm{d}f/\mathrm{d}\theta_i = f_\theta + f_{\bar{x}}(\mathrm{d}\bar{x}/\mathrm{d}\theta) = f_\theta - f_{\bar{x}}(f_{\bar{x}\theta}/f_{\bar{x}\bar{x}}) > 0$. This alternative model predicts that we should observe a positive correlation in the data between resources per capita ($\bar{x}_i$) and success rates (f), which is not the case in our dataset.

We have shown, in a very simple context, that a well-informed planner will optimally use its information in its allocation decision. Machado (1997) extends this model and applies it to the OSA allocation problem assuming that OSA's objective is to maximize the number of abstinent discharged patients in the state.

3.2. Instrumental variables

Once we established the potential risk of a downward bias due to the endogeneity of $\bar{c}_{it}$, we need to look for instrumental variables. Instrumental variables have to be correlated with $\bar{c}_{it}$ but not directly correlated with the residuals η_{it} in Eq. (4).

The most natural set of instruments are time dummies (year and seasonal dummies $Y_1, Y_2, Y_3, Q_2, Q_3, Q_4$). These are naturally correlated with expenditures per patient because OSA makes its allocation decision annually. Furthermore, OSA pays the agencies on a quarterly basis and often makes adjustments to the amounts allocated in the form of amendments to the contracts. Contract amendments are usually the result of an unexpected increase or cut in its total budget. [19] Machado (1997) shows that an optimal allocation model produces time dummies as the natural set of instruments for $\bar{c}_{it}$.

This paper also considers the number of co-dependents per patient (CODEPS) as an instrumental variable candidate. CODEPS is likely to affect expenditures since it may increase costs related to case-management, e.g. the costs of setting up a schedule convenient to all members of the family, telephone calls, cancellations, etc. Our assumption is that it will not be directly correlated with performance but it may affect time in treatment. We will see that our results show some evidence of the latter effect.

The length of the contract with OSA (KYEAR) is also likely to be correlated with expenditures per patient since shorter contracts are usually paid by OSA on a fee-for-service basis which forces programs to cut costs or increase the number of patients to be able to break even.

Average medicaid funds per patient (MC) is a good candidate for an instrumental variable because medicaid funds are matched with OSA funds allowing for more expensive programs.

The presence of local representatives in political institutions at the state level may have an impact on the budget allocated to agencies within a particular region. Consequently, we created a dummy variable, LEG, which takes the value one if the agency is located in a city with a legislator or a representative in the Appropriations Committee.

Finally, we considered two variables related to the racial mix of patients (BLACK, WHITE) as instruments for $\bar{c}_{it}$. These characteristics are easily observable by OSA and are likely to affect its budgeting decisions. In fact, our results show that the racial mix variables are among the best instrumental variables.

4. The data

Our dataset is composed of 38 contracts between OSA and agencies of alcohol abuse treatment from fiscal year 1990 through 1994. We dropped fiscal year 1990 observations

[19] We do not believe that time dummies have an impact on the probability of success. As Footnote 8 indicates, time dummies were not significant as regressors in the probability equation.

from the estimations due to incompatibilities and missing values that occurred following a major change in the format of the individual patient admission and discharge forms (MATS) in April 1990. Moreover, when MATS was introduced, programs were not forced to fill out admission forms for clients who were already in treatment, which could lead to sizable biases in fiscal year 1990 observations. Our dataset is an unbalanced panel due to disruption of contracts, the emergence of new providers, or simply due to missing values for certain variables.

Each contract may offer several programs on the same or different modalities.[20] Our unit of analysis is the outpatient treatment program within a contract that an agency signs with OSA.

The main data source is the MATS admission and discharge forms. These forms contain detailed information about the patient's alcohol/drug addiction, general demographics, involvement with the law, mental health, and social environment. Although information reported in the MATS forms is self-reported, they are filled out by a clinician, which we think increases the reliability of the data.

The second data source is the contracts between OSA and each of the agencies. The typical contract states the budget for the contract year as well as the break-down of the total budget across programs.

The third data source is the quarterly financial reports submitted to OSA by the agencies for each of their contracts. These reports state income and expenditures up to the contract-quarter.

Finally, we used the monthly consumer price index for the northeast region from the Bureau of Labor Statistics to deflate expenditures.

4.1. Performance measures

In our study, we use as the performance measure one of the effectiveness indicators considered by OSA — the percentage of discharged patients from a given program in a given quarter, that are 'abstinent 30 days prior to discharge'. On an average, 72% of the discharged patients were abstinent 1 month prior to their discharge (see Table 1).

4.2. Measures of spending

Quarterly financial reports show how much was received from OSA and other sources (income) and how much was spent (expenditures) on the contract as a whole, up to the date of the report.[21] Although this information has the advantage of being quarterly, and of reflecting real and not budgeted amounts, it is aggregated at the contract level instead of

[20] Usually, when a contract offers two programs of the same modality, e.g. outpatient treatment, they differ either by site (rural versus urban) or type of patients (adolescent, regular or elderly). In these cases we merged the data into a single observation.

[21] In the data, income and expenditures are highly correlated (correlation coefficient around 0.8). We decided to use expenditures rather than income because these are more representative of the actual investment in patients. Also, according to OSA officials, income may be seasonal or lagged. Medicaid payments, for instance, are usually made 1 year after expenses are incurred.

Table 1
Sample statistics for sample without fiscal year 1990

	Variable	N	Mean	S.D.	Min.	Max.	Median
Program characteristics							
Total number of clients measured for performance in terms of abstinence	PERACLI	470	79.76	88.26	0	354	36.50
Total number of clients	TOTCLI	470	110.40	120.31	1	500	55.00
Budgeted percentage of total contract income directed to outpatient	PERCO	405	0.82	0.27	0.03	1.00	0.97
Funds per patient (TEXPDF/PERACLI)	CPDS	334	891.9	1727.8	0.0	16229.7	437.1
Funds per patient taking average time in treatment into account	CUMCPDS	310	1537.8	5175.4	16.0	53206.8	490.0
Funds per patient (TEXPDF/PERACLI) taking outliers out	CPDS	322	755.0	1061.1	5.9	8173.3	436.6
Dummy for fiscal year 1991	Y_1	488	0.25	0.43	0	1	0
Dummy for fiscal year 1992	Y_2	488	0.26	0.44	0	1	0
Dummy for fiscal year 1993	Y_3	488	0.25	0.43	0	1	0
Dummy for fiscal year 1994	Y_4	488	0.24	0.42	0	1	0
Characteristics of the primary clients discharged							
Number of discharges measure for abstinence	ABCLI	422	38.50	42.25	1	201	19.00
Fraction of discharges that are abstinent	ABMET	422	0.72	0.21	0.00	1.00	0.75
Average time in treatment for discharges measured for abstinence	ABTIM	422	125.49	87.59	1.00	723	108.19
Fraction of discharges that dropped out	DROPOUT	422	0.42	0.23	0.00	1.00	0.43
Fraction of discharges that complete treatment	COMPLETE	422	0.39	0.23	0.00	1.00	0.38
Fraction of discharges that are white	WHITE	422	0.95	0.16	0.00	1.00	1.00
Fraction of discharges that are black	BLACK	422	0.01	0.03	0.00	0.33	0.00
Fraction of discharges that are homeless	HOMELESS	422	0.01	0.08	0.00	1.00	0.00
Fraction of discharges that depend on others for living	DEPENDNT	422	0.19	0.24	0.00	1.00	0.13
Fraction of discharges that are independent	INDEPDNT	422	0.79	0.24	0.00	1.00	0.87
Fraction of discharges that are in jail	JAILED	422	0.03	0.15	0.00	1.00	0.00
Fraction of discharges that are in parole	PAROLE	422	0.23	0.21	0.00	1.00	0.19
Fraction of discharges that are waiting trial	TRIAL	422	0.07	0.09	0.00	0.67	0.05
Fraction of discharges that are veterans of war	VET	422	0.14	0.14	0.00	1.00	0.13
Fraction of discharges that have psychiatric problems	PSYCH	422	0.11	0.14	0.00	1.00	0.07
Fraction of discharges that have less than the 12th grade	LESS12	422	0.41	0.25	0.00	1.00	0.36
Fraction of discharges that have between the 12th and 16th grade	BETW1216	422	0.58	0.25	0.00	1.00	0.62
Fraction of discharges that have more than 16th grade	MORE16	422	0.01	0.05	0.00	1.00	0.00
Fraction of discharges that had at least five prior treatment episodes	PRITX5	422	0.05	0.09	0.00	0.67	0.02

Discharges average age of first use of primary drug	DRGAGE1	403	15.04	2.14	9.00	27.80	15.07
Fraction of discharges with minor problems with spouse	LOWSO	404	0.79	0.23	0.00	1.00	0.85
Fraction of discharges with minor problems with family	LOWFAM	404	0.85	0.18	0.00	1.00	0.91
Fraction of discharges with minor problems at job/school	LOWJOB	404	0.90	0.16	0.00	1.00	0.95
Fraction of discharges whose main drug at admission is alcohol	D1ALCO	437	0.83	0.17	0.00	1.00	0.86
Fraction of discharges whose main drug at admission is marijuana	D1MARI	437	0.11	0.14	0.00	1.00	0.08
Fraction of discharges whose main drug at admission is cocaine	D1COCA	437	0.02	0.06	0.00	0.50	0.00
Fraction of discharges whose main drug at admission is tranquilizers	D1TRAN	437	0.01	0.03	0.00	0.50	0.00
Fraction of discharges whose main drug at admission is barbiturants	D1BARB	437	0.00	0.01	0.00	0.13	0.00
Fraction of discharges whose main drug at admission is methamphetamine	D1METH	437	0.00	0.01	0.00	0.14	0.00
Fraction of discharges whose main drug at admission is LSD	D1LSD	437	0.01	0.06	0.00	1.00	0.00
Fraction of discharges whose main drug at admission is heroin	D1HERO	437	0.00	0.02	0.00	0.17	0.00
Fraction of discharges whose main drug at admission is other	D1OTHR	437	0.01	0.04	0.00	0.50	0.00
Fraction of discharges whose second drug at admission is alcohol	D2ALCO	437	0.10	0.14	0.00	1.00	0.07
Fraction of discharges whose second drug at admission is marijuana	D2MARI	437	0.28	0.21	0.00	1.00	0.25
Fraction of discharges whose second drug at admission is methamphetamine	D2METH	437	0.01	0.05	0.00	1.00	0.00
Fraction of discharges with no use of main drug 1 month prior admission	F1NULM	437	0.47	0.27	0.00	1.00	0.50
Fraction of discharges using main drug once a month prior admission	F1ONEM	437	0.06	0.09	0.00	0.67	0.03
Fraction of discharges using main drug two to three times a month prior admission	F123M	437	0.07	0.09	0.00	1.00	0.05
Fraction of discharges using main drug once a week prior admission	F1ONEW	437	0.05	0.07	0.00	0.33	0.04
Fraction of discharges using main drug two to three times a week prior admission	F123W	437	0.11	0.12	0.00	1.00	0.08
Fraction of discharges using main drug four to six times a week prior admission	F146W	437	0.06	0.12	0.00	1.00	0.03
Fraction of discharges using main drug once a day prior admission	F1ONED	437	0.08	0.13	0.00	1.00	0.04
Fraction of discharges using main drug two to three times a day prior admission	F123D	437	0.04	0.09	0.00	0.67	0.00
Fraction of discharges using main drug more than three times a day prior admission	F13OVD	437	0.06	0.11	0.00	0.75	0.01
Fraction of discharges classified as casual users at admission	SEVCAS	417	0.08	0.14	0.00	1.00	0.03
Fraction of discharges classified as involved users at admission	SEVINV	417	0.21	0.19	0.00	1.00	0.18
Fraction of discharges classified as dependent users at admission	SEVDEP	417	0.36	0.23	0.00	1.00	0.34
Fraction of discharges classified as dysfunctional users at admission	SEVDYS	417	0.20	0.24	0.00	1.00	0.14
Fraction of discharges with undetermined severity at admission	SEVUDE	417	0.15	0.19	0.00	1.00	0.09
Fraction of discharges that never had IV drug use	IVNEVE	437	0.91	0.12	0.00	1.00	0.94
Discharges average weekly household income, deflated	HHINCD	391	597.94	361.76	0.00	5296.85	561.46

being disaggregated by program. This means that if a contract offers more than one program, we cannot disentangle the income received or expenditures made by each of them. To measure the expenditures made by outpatient programs, we assume that they are in accordance with the percentages stated on the contracts.

We use two alternative measures of expenditures per patient: *current* expenditures per patient defined by the quarterly expenditure on an outpatient program divided by the total number of patients who received treatment during that quarter and *accumulated* expenditures per patient, which is computed as the sum of the *current* expenditures per patient over the average time the discharged patients were in treatment. The formula for *accumulated* expenditures per patient is

$$\bar{c}_{it} = \frac{c_{it}}{N_{it}} + \sum_{s=1}^{T} \max\left\{\min\left\{\frac{\mathrm{AT}}{90} - s, s\right\}, 0\right\} \frac{c_{it-s}}{N_{it-s}}, \tag{11}$$

where AT stands for average time in treatment of the discharged cohort. The correction for AT may be important since, as can be seen in Table 1, the average time in treatment (AT) is 125.49 days and the median is 108.19 days, which is bigger than one quarter.

More than 80% of the programs spend less than US$ 1000.00 per patient under both definitions of expenditures. For *current* expenditures per patient, we left out of the estimation observations where expenditures per patient were above US$ 11,000.00 or zero. The outliers belonged to two small programs in the dataset. We deleted the whole year of data for the outlier program.

4.3. Patient characteristics data

In our estimation, we control for discharged patients characteristics, most of which were measured at the time of the patients' admission to the program. These data are taken from the MATS forms and aggregated at the program level. Some of these variables may be classified as frequency of use of primary drug, severity of alcohol abuse problem, type of primary and secondary drug, marital status, professional status, intravenous drug user, etc. Most of the variables have values between [0, 1] because they represent the fraction of discharged patients who fall into a particular category. Other variables tell us an average value; for example, the variable HHINC tells us the average household income of the discharged population at the time of their admission. Table 1 provides basic descriptive statistics for relevant variables.

5. Results

5.1. The first-stage estimation

Tables 2 and 3 show the OLS estimation results of the first-stage regression for *current* and *accumulated* expenditures per patient, respectively, with and without fixed effects. The tables only show the estimated coefficients of the sets of instrumental variables used in

Table 2
First-stage regression of current expenditure per patient using time dummies (TD), demographics (DM) and contract variables (CV) as instruments[a]

Description	Variable	TD				TD + DM				TD + DM + CV			
		WFE[b]		WTFE[c]		WFE[b]		WTFE[c]		WFE[b]		WTFE[c]	
		EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]
Intercept	C	4700.93	3.16	719.01	0.35	4700.93	3.16	719.01	0.35	4479.08	3.25	2284.78	1.22
Fiscal year 1991	Y_1	287.48	3.58	171.70	1.47	287.48	3.58	171.70	1.47	535.37	5.40	327.48	2.38
Fiscal year 1992	Y_2	167.93	2.18	193.54	1.66	167.93	2.18	193.54	1.66	409.77	4.01	279.91	1.91
Fiscal year 1993	Y_3	93.00	1.28	56.17	0.54	93.00	1.28	56.17	0.54	257.90	3.32	112.82	1.02
Second quarter	Q_2	76.16	1.08	−1.55	−0.01	76.16	1.08	−1.55	−0.01	68.26	1.07	12.52	0.12
Third quarter	Q_3	−84.33	−1.28	−64.94	−0.62	−84.33	−1.28	−64.94	−0.62	−91.92	−1.54	−64.03	−0.67
Forth quarter	Q_4	−45.69	−0.64	−41.88	−0.36	−45.69	−0.64	−41.88	−0.36	−33.27	−0.51	−32.45	−0.31
Percentage of white patients	WHITE					−1387.65	−1.90	−8.93	−0.02	−1730.52	−2.59	−221.56	−0.48
Percentage of black patients	BLACK					−666.59	−0.53	87.04	0.06	−1292.02	−1.13	654.66	0.48
Contract length in years	KYEAR									17.65	4.35	12.99	2.37
Average number of co-dependents per patient	CODEPS									848.99	5.75	802.98	6.56
Legislator or representative in Appropriations Committee	LEG									−258.60	−1.32	−38.21	−0.18
Goodness-of-fit													
Number of observations		288		288		288		288		287		287	
R^2		0.88		0.64		0.88		0.64		0.91		0.71	
Adjusted-R^2		0.83		0.55		0.83		0.55		0.86		0.63	
F-statistic (zero slopes)		17.39		6.93		17.39		6.93		20.91		8.93	
		Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
F-test (joint significancy of all "instruments")		2.753	0.014	0.653	0.688	2.336	0.020	0.494	0.860	4.700	0.000	4.225	0.000

[a] Regressions contain the following variables: D1ALCO, D1MARI, D1METH, D1HERO, D1LSD, D1BARB, D1TRAN, F146W, F1ONED, F123D, F13OVD, F1ONEM, F123M, F1ONEW, F123W, SEVDEP, SEVINV, SEVCAS, D2ALCO, D2METH, D2MARI, RISK, HOMELESS, DEPENDNT, JAILED, PAROLE, TRIAL, VET, PSYCH, LESS12, BETW1216, HHINCD, PRITX5, DRGAGE1, LOWSO, LOWFAM, LOWJOB, ACCTGNEW, MISACCTG, JSTANDCR, F1DMARI, F1DTRAN, F11ALCO, F11HERO, F11MARI, F14HERO, F14ALCO, F1DALCO, F1OTRAN, F11COCA and F11LSD.

[b] With fixed effects.

[c] Without fixed effects.

[d] Estimated coefficient.

[e] t-statistic.

Table 3
First-stage regression of accumulated expenditure per patient using time dummies (TD), demographics (DM), contract variables (CV) and county variables (CYV) as instruments[a]

Description	Variable	TD				TD + DM				DM			
		WFE[b]		WTFE[c]		WFE[b]		WTFE[c]		WFE[b]		WTFE[c]	
		EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]	EC[d]	t-stat.[e]
Intercept	C	8040.63	2.86	475.27	0.15	8040.63	2.86	475.27	0.15	8811.71	3.20	862.05	0.28
Fiscal year 1991	Y_1	264.21	1.77	−89.52	−0.50	264.21	1.77	−89.52	−0.50				
Fiscal year 1992	Y_2	31.84	0.23	−14.99	−0.09	31.84	0.23	−14.99	−0.09				
Fiscal year 1993	Y_3	−8.62	−0.07	−8.57	−0.05	−8.62	−0.07	−8.57	−0.05				
Second quarter	Q_2	−175.98	−1.38	−191.67	−1.14	−175.98	−1.38	−191.67	−1.14				
Third quarter	Q_3	−202.90	−1.63	−110.87	−0.68	−202.90	−1.63	−110.87	−0.68				
Forth quarter	Q_4	−87.76	−0.66	14.88	0.09	−87.76	−0.66	14.88	0.09				
Percentage of white patients	WHITE					−6539.76	−4.69	−1872.60	−2.36	−6120.61	−4.46	−1788.60	−2.31
Percentage of black patients	BLACK					−6623.87	−2.84	12.45	0.01	−5835.37	−2.52	213.89	0.09
Goodness-of-fit													
Number of observations		277		277		277		277		277		277	
R^2		0.78		0.53		0.78		0.53		0.77		0.53	
Adjusted-R^2		0.68		0.41		0.68		0.41		0.68		0.42	
F-statistic (zero slopes)		7.81		4.21		7.81		4.21		8.28		0.91	
		Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
F-test (joint significancy of all "instruments")		1.094	0.368	0.42	0.865	2.930	0.004	1.033	0.412	9.018	0.000	2.985	0.053
Intercept	C	7943.35	2.86	1937.33	0.63	7943.35	2.86	1937.33	0.63	8864.19	3.22	1917.78	0.63
Percentage of white patients	WHITE					−6088.27	−4.45	−2088.40	−2.67	−6061.75	−4.41	−2075.95	−2.67
Percentage of black patients	BLACK					−3725.37	−1.44	2538.81	0.86	−5700.71	−2.46	314.19	0.14
Contract length in years	KYEAR	8.11	1.57	13.61	2.16	8.11	1.57	13.61	2.16	5.25	1.07	12.98	2.15
Medicaid funds per patient	MC	−0.24	−1.75	−0.17	−1.27	−0.24	−1.75	−0.17	−1.27				
Legislator or representative in Appropriations Committee	LEG	−614.70	−1.56	−25.61	−0.08	−614.70	−1.56	−25.61	−0.08				

	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Goodness-of-fit												
Number of observations	275		275		275		275		277		277	
R^2	0.78		0.54		0.78		0.54		0.77		0.54	
Adjusted-R^2	0.68		0.42		0.68		0.42		0.68		0.43	
F-statistic (zero slopes)	8.13		4.59		8.13		4.59		8.19		4.78	
F-test (joint significancy of all "instruments")	2.166	0.093	1.980	0.118	4.681	0.000	2.331	0.043	6.323	0.000	3.446	0.018

[a] Regressions contain the following variables: D1ALCO, D1MARI, D1METH, D1HERO, D1LSD, D1BARB, D1TRAN, F146W, F1ONED, F123D, F13OVD, F1ONEM, F123M, F1ONEW, F123W, SEVDEP, SEVINV, SEVCAS, D2ALCO, D2METH, ssD2MARI, RISK, HOMELESS, DEPENDNT, JAILED, PAROLE, TRIAL, VET, PSYCH, LESS12, BETW1216, HHINCD, PRITX5, DRGAGE1, LOWSO, LOWFAM, LOWJOB, ACCTGNEW, MISACCTG, JSTANDCR, F1DMARI, F1DTRAN, F11ALCO, F11HERO, F11MARI, F14HERO, F14ALCO, F1DALCO, F1OTRAN, F11COCA and F11LSD.

[b] With fixed effects.

[c] Without fixed effects.

[d] Estimated coefficient.

[e] t-statistic.

the "production function" estimation. The tables also show goodness-of-fit statistics and an F-test of the joint significancy of the instruments used.

Our baseline model is the estimation with fixed effects for which our instruments perform better as first-stage regressors, as we can see from the comparison of the P-values of the F-test, at the bottom of the tables.

For *current* expenditures per patient most candidates for instrumental variables are good first-stage regressors. The lowest P-value of the F-test (in Table 2) is obtained with the largest number of instruments.

For *accumulated* expenditures per patient we had more difficulty in finding good instruments. In general, the time dummies, KYEAR and CODEPS are not good predictors of *accumulated* expenditures per patient while the variables WHITE, BLACK, MC and LEG perform quite well.

WHITE and BLACK have a negative impact in both *current* and *accumulated* expenditures per patient. This probably reflects the negative correlation with the number of American Indians, which are concentrated in areas such as Indian reservations that may require more expensive programs.

The length of the contract with OSA (KYEAR) has a significant positive impact on both *current* and *accumulated* spending per patient probably because short-term contracts force programs to be more cost-conscious.

CODEPS also affects positively and very significantly the *current* spending per patient although it had no significant impact on *accumulated* spending per patient. We suspect this difference has to do with two distinct effects that offset each other, causing CODEPS not to be correlated with *accumulated* expenditures per patient. First, conditional on time in treatment, family therapy sessions involve costs that increase both *current* and *accumulated* spending. Second, it is likely that family therapy affects the patient's time in treatment and therefore the *accumulated* spending. For example, patients may succeed faster (decreasing the average time in treatment), and some of the failures may drop out later (increasing the average time in treatment). We think the net effect on the average time in treatment is negative offsetting the increase in costs per period.

Average medicaid funds per patient (MC) had no significant effect on *current* expenditures per patient and an almost significantly negative impact on *accumulated* expenditures per patient contrary to our belief. It is possible that medicaid selects providers on the basis of cost.

Surprisingly, LEG has a negative impact on both *current* and *accumulated* expenditures per patient probably because we are also controlling for another political variable, the city's representation in the Human Resources Committee (JSTANDCR). JSTANDCR is not used as an instrument because it is also correlated with performance.

5.2. The instrumental variable estimation of the "production function"

This subsection describes the results of the "production function" estimation. Tables 4 and 5 show the estimated α and intercept term for the IV and OLS estimation for *current* and *accumulated* expenditures per patient, respectively. These tables also provide measures of goodness-of-fit and tests of endogeneity and overidentifying restrictions. The IV estimates reported are obtained from the standard IV estimator $\hat{\beta} = (X'Z(Z'Z)^{-1}$

Table 4
"Production function" estimation using current expenditures per patient and also using time dummies (TD), demographics (DM) and contract variables (CV) as instruments[a]

Description	Variable	OLS results (R.S.D.[b])				IV estimation results (R.S.D.[b])											
						TD				TD + DM				TD + DM + CV			
		WFE[c]		WTFE[d]		WFE[c]		WTFE[d]		WFE[c]		WTFE[d]		WFE[c]		WTFE[d]	
		EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]
Intercept	C	−4.766	−0.934	−6.432	−1.111	−4.121	−0.612	−7.271	−1.285	−4.093	−0.674	−10.052	−1.825	−4.905	−1.011	−8.521	−1.592
Expenditures per patient	CPDS	−0.0002	−0.591	−0.0002	−0.616	−0.0004	−0.415	0.0004	0.249	−0.0004	−0.418	0.0006	0.377	−0.0002	−0.370	−0.0007	−1.880
List of instruments																	
Fiscal year 1991	Y_1						Y_1				Y_1			Y_1			
Fiscal year 1992	Y_2						Y_2				Y_2			Y_2			
Fiscal year 1993	Y_3						Y_3				Y_3			Y_3			
Second quarter	Q_2						Q_2				Q_2			Q_2			
Third quarter	Q_3						Q_3				Q_3			Q_3			
Fourth quarter	Q_4						Q_4				Q_4			Q_4			
Percentage of white patients	WHITE										WHITE			WHITE			
Percentage of black patients	BLACK										BLACK			BLACK			
Average number of co-dependents per patient	CODEPS													CODEPS			
Contract length in years	KYEAR													KYEAR			
Legislator or representative on the Appropriations Committee	LEG													LEG			
Goodness-of-fit																	
Number of observations		288		288		288		288		288		288		287		287	
R^2		0.70		0.49		0.70		0.47		0.70		0.44		0.70		0.46	
Adjusted-R^2		0.58		0.37		0.57		0.34		0.58		0.32		0.58		0.34	
F-statistic (zero slopes)		5.95		4.11		5.73		3.76		5.93		3.55		5.91		3.84	
Mroz version of Hausman test — all coefficients																	
Value of the test (IV against OLS estimation)						0.057		0.291		0.00		0.20		0.00		0.08	
P-value (Chi-square 81 d.f.)						1.000		1.000		1.00		1.00		1.00		1.00	

Table 4 (*Continued*)

Description	Variable	OLS results (R.S.D.[b])		IV estimation results (R.S.D.[b])					
				TD		TD + DM		TD + DM + CV	
		WFE[c]	WTFE[d]	WFE[c]	WTFE[d]	WFE[c]	WTFE[d]	WFE[c]	WTFE[d]
Testing the overidentifying restrictions									
Sargan test				8.933	5.341	8.934	12.413	13.241	17.561
P-value				0.112	0.376	0.257	0.088	0.278	0.092
F-test of first-stage estimation and d.f.				2.753 (6,200)	0.653 (6,228)	2.336 (8,200)	0.494 (8,228)	4.700 (11,196)	4.225 (11,224)
P-value of *F*-test of first-stage estimation				0.014	0.688	0.020	0.860	0.000	0.000

[a] Regressions contain the following variables: D1ALCO, D1MARI, D1METH, D1HERO, D1LSD, D1BARB, D1TRAN, F146W, F1ONED, F123D, F13OVD, F1ONEM, F123M, F1ONEW, F123W, SEVDEP, SEVINV, SEVCAS, D2ALCO, D2METH, D2MARI, RISK, HOMELESS, DEPENDNT, JAILED, PAROLE, TRIAL, VET, PSYCH, LESS12, BETW1216, HHINCD, PRITX5, DRGAGE1, LOWSO, LOWFAM, LOWJOB, ACCTGNEW, MISACCTG, JSTANDCR, F1DMARI, F1DTRAN, F11ALCO, F11HERO, F11MARI, F14HERO, F14ALCO, F1DALCO, F1OTRAN, F11COCA and F11LSD.

[b] Robust standard deviation.

[c] With fixed effects.

[d] Without fixed effects.

[e] Estimated coefficient.

[f] *t*-statistic.

Table 5
"Production function" estimation using accumulated expenditures per patient and also using time dummies (TD), demographics (DM), contract variables (CV) and county variables (CYV) as instruments[a]

Description	Variable	OLS results (R.S.D.[b])				IV estimation results (R.S.D.[b])							
						TD				TD + DM			
		WFE[c]		WTFE[d]		WFE[c]		WTFE[d]		WFE[c]		WTFE[d]	
		EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]	EC[e]	t-stat.[f]
Intercept	C	−1.913	−0.320	−2.484	−0.439	−4.010	−0.515	−1.957	−0.331	−3.699	−0.772	−2.999	−0.613
Expenditures per patient	CPDS	0.0000	0.257	0.0003	2.217	0.0003	0.396	−0.0003	−0.355	0.0002	0.570	0.0003	0.553
List of instruments													
Fiscal year 1991	Y_1						Y_1				Y_1		
Fiscal year 1992	Y_2						Y_2				Y_2		
Fiscal year 1993	Y_3						Y_3				Y_3		
Second quarter	Q_2						Q_2				Q_2		
Third quarter	Q_3						Q_3				Q_3		
Forth quarter	Q_4						Q_4				Q_4		
Percentage of white patients	WHITE										WHITE		
Percentage of black patients	BLACK										BLACK		
Number of observations		277		277		277		277		277		277	
R^2		0.68		0.53		0.68		0.47		0.68		0.53	
Adjusted-R^2		0.55		0.42		0.54		0.34		0.55		0.42	
F-statistic (zero slopes)		5.13		4.69		5.03		3.64		5.24		4.90	
Mroz endogeneity test (IV against OLS estimation)						0.000		0.014		0.000		0.000	
P-value (Chi-square 82 d.f.)						1.000		1.000		1.000		1.000	
Sargan test of the overidentifying restrictions						10.871		7.319		10.944		11.176	
P-value						0.054		0.198		0.141		0.131	
F-test of first-stage estimation and d.f.						1.094 (6,190)		0.420 (6,217)		2.930 (8,190)		1.033 (8,217)	
P-value of F-test of first-stage estimation						0.368		0.865		0.004		0.412	

Table 5 (*Continued*)

Description	Variable	DM				CV + DM + CYV				CV + DM			
		WFE[c]		WTFE[d]		WFE[c]		WTFE[d]		WFE[c]		WTFE[d]	
		EC[e]	*t*-stat.[f]	EC[e]	*t*-stat.[f]	EC[e]	*t*-stat.[f]	EC[e]	*t*-stat.[f]	EC[e]	*t*-stat.[f]	EC[e]	*t*-stat.[f]
Intercept	*C*	−3.647	−0.756	−2.734	−0.565	−2.818	−0.587	−3.094	−0.633	−3.508	−0.742	−2.835	−0.584
Expenditures per patient	CPDS	0.0002	0.487	0.0005	0.953	−0.0001	−0.205	0.0003	0.813	0.0002	0.408	0.0004	0.945
List of instruments													
Percentage of white patients	WHITE		WHITE				WHITE				WHITE		
Percentage of black patients	BLACK		BLACK				BLACK				BLACK		
Contract length in years	KYEAR						KYEAR				KYEAR		
Legislator or representative in Appropriations Committee	LEG						LEG						
Medicaid funds per patient	MC						MC						
Number of observations		277		277		275		275		277		277	
R^2		0.68		0.52		0.68		0.53		0.68		0.53	
Adjusted-R^2		0.55		0.41		0.55		0.42		0.55		0.42	
F-statistic (zero slopes)		5.25		4.69		5.23		4.86		5.28		4.82	
Mroz endogeneity test (IV against OLS estimation)		0.000		0.002		0.001		0.000		0.000		0.001	
P-value (Chi-square 82 d.f.)		1.000		1.000		1.000		1.000		1.000		1.000	
Sargan test of the overidentifying restrictions		0.351		0.808		7.481		1.661		0.871		0.862	
P-value		0.553		0.369		0.113		0.798		0.647		0.650	
F-test of first-stage estimation and d.f.		9.018 (2,196)		2.985 (2,223)		4.681 (5,191)		2.331 (5,218)		6.323 (3,195)		3.446 (3,222)	
P-value of *F*-test of first-stage estimation		0.000		0.053		0.000		0.043		0.000		0.018	

[a] Regressions contain the following variables: D1ALCO, D1MARI, D1METH, D1HERO, D1LSD, D1BARB, D1TRAN, F146W, F1ONED, F123D, F13OVD, F1ONEM, F123M, F1ONEW, F123W, SEVDEP, SEVINV, SEVCAS, D2ALCO, D2METH, D2MARI, RISK, HOMELESS, DEPENDNT, JAILED, PAROLE, TRIAL, VET, PSYCH, LESS12, BETW1216, HHINCD, PRITX5, DRGAGE1, LOWSO, LOWFAM, LOWJOB, ACCTGNEW, MISACCTG, JSTANDCR, F1DMARI, F1DTRAN, F11ALCO, F11HERO, F11MARI, F14HERO, F14ALCO, F1DALCO, F1OTRAN, F11COCA and F11LSD.

[b] Robust standard deviation.

[c] With fixed effects.

[d] Without fixed effects.

[e] Estimated coefficient.

[f] *t*-statistic.

$Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y$, where Z denotes the matrix of instruments and X the matrix of regressors.[22]

In our empirical model, the standard Hausman test (Hausman, 1978) is not suitable for testing the endogeneity of expenditures per patient because of the inexistence of an efficient and consistent estimator under the null hypothesis of exogeneity. On the one hand, the maximum likelihood estimates are not consistent due to the presence of fixed effects. On the other hand, feasible generalized least squares (FGLS) estimates are not efficient (see Greene, 1993, p. 365) because the covariance matrix of the residuals depends on parameters that also explain the probability of success (see Eq. (6)). Mroz (1987) devised an endogeneity test that does not require efficiency under the null. We computed the Mroz endogeneity test based on the distance between a consistent estimator under the null (the OLS estimator) and a consistent estimator under the alternative (the IV estimator).

To test the overidentifying restrictions we have used the Sargan specification test based on the correlation between instruments and residuals.[23]

Table 4 shows a negative, although not significantly different from zero, marginal impact of *current* expenditures per patient on performance for the baseline model with fixed effects. The Sargan test does not reject the overidentifying restrictions at the 10% confidence level, and the instruments are strongly correlated with *current* expenditures per patient as the low P-values of the F-test indicate. All sets of instruments used are good instruments, although the most solid Sargan and F-tests are obtained when all instruments are used simultaneously. The Mroz endogeneity test does not reject the null hypothesis of exogeneity of *current* expenditures per patient. The exogeneity of funding is probably related to the rich set of regressors already included in the estimations, in particular the fixed effects. When fixed effects are not included in the estimations, we see that the fit of the regressions drops considerably and the instruments are not valid. The estimates of the model without fixed effects are, therefore, shown for purposes of comparison only. We conclude that the impact of *current* expenditures per patient on performance is small and non-positive.

For *accumulated* expenditures per patient, Table 5 shows much more optimistic estimates. For the most part, the coefficient on accumulated expenditures per patient is now positive, although not significantly different from zero, and quantitatively small. The sets of valid

[22] Although this paper does not report the results, we have also tried taking into account the heteroscedasticity of the error term by estimating the residual covariance matrix Ω. Assuming Ω to be a diagonal matrix, the estimation of the variance terms was based on the fitted values from the following regression:

$$\hat{\eta}_{it}^2 = \sigma_\theta^2 + \chi \frac{1}{d_{it}\hat{p}_{it}(1-\hat{p}_{it})} + \xi_{it}.$$

Once we estimate $\hat{\Omega} = \mathrm{diag}\{\widehat{\hat{\eta}_{it}^2}\}$ we can derive the IV–OLS analogue and IV–GLS analogue (see Bowden and Turkington, 1984). However, the estimated standard deviations of both the IV–OLS analogue and IV–GLS analogue were very large due to the low fit of the $\hat{\eta}_{it}^2$ regression which R^2 was never above 0.01. This result may be due to omitted variables, or heteroscedasticity in the other component of the error term θ_{it}.

[23] The Sargan test takes the form

$$S = (Z'(y - X\hat{\beta}))' \left(\sum \hat{u}_i^2 z_i z_i'\right)^{-1} ((y - X\hat{\beta})Z) \to \chi^2_{(s)},$$

where s is the difference between the number of instruments and the number of parameters to estimate.

instruments are now different from the sets used for *current* expenditures per patient. In particular, time dummies are no longer strongly correlated with *accumulated* expenditures per patient and do not show a good Sargan test, as can be seen in the second column of Table 5. WHITE and BLACK, on the other hand, perform very well as instrumental variables. The only case where the baseline model with fixed effects shows a negative coefficient on $\bar{c}_{it}$ is when the P-value on the Sargan test is as low as 0.113, and therefore, the instruments are somewhat more correlated with the residuals than in other cases. In all the cases, the Mroz endogeneity test does not reject the null hypothesis of exogeneity of the *accumulated* expenditures per patient.

Overall, the instrumental variable approach was unable to find a positive significantly different from zero relationship between *current* or *accumulated* expenditures per patient and performance measured in terms of fraction of abstinent discharges.

6. Discussion

This section discusses policy implications, caveats, options and assumptions made in this paper that we think are relevant when interpreting results.

An important limitation of this study is the restriction to a single performance measure. The traditional performance measures, based on substance consumption such as abstinence or simply reduction in consumption levels, are widely accepted as necessary for a good evaluation of a patient's performance. [24] More recently, other measures of performance such as reduction in criminal activities, general health, and behavior at work or school have become popular among researchers and health practitioners (McLellan et al., 1997). In the case of alcoholics, however, abstinence as a measure of performance is a good choice since a transition to moderate drinking is usually difficult to sustain. However, it is quite feasible that marginal dollars may have a significant positive effect in attaining other treatment goals.

Another limitation of this study is its forced reliance on performance data collected at the time of discharge from treatment, due to the inexistence of follow-up data. This will not be a source of bias if the correlation between performance at discharge and at follow-up is independent of expenditures per patient. Yet, it is reasonable to think that programs that invest more in their patients also have a stronger correlation between performance at discharge and at follow-up. This is clearly a problem that we can only hope is unimportant in this particular dataset.

Next, the use of aggregated data for the purposes of this paper is discussed at length, as well as the reasons why we think that if aggregation introduces any bias on the estimated impact of marginal funding on performance, it is more likely to be an upward bias.

Firstly, it must be stressed that this paper does not attempt to assess treatment effectiveness, i.e. whether treatment improves patients' condition, which, of course, requires a patient-level analysis. Secondly, it does not attempt to determine which type of patients

[24] One exception to this rule is the "Harm-Reduction" movement in certain European countries (originated in The Netherlands and UK), which aims at decreasing the devastating consequences of substance abuse through, for instance, controlled consumption, needle sharing, or even liberalization of drugs, and not necessarily through the reduction of consumption and abstinence (see Marlatt, 1996).

are more cost-effective[25] but rather focuses on the performance of entire programs. This paper attempts to determine whether the marginal funding received by treatment agencies is being used to promote abstinence among their patients. This question legitimizes the use of aggregate data.

Notice that although the purpose of this paper differs from treatment effectiveness evaluation, the two ideas are linked. Treatment effectiveness matters for our results in the sense that it is a necessary (but not sufficient) condition for a positive impact of funding per patient on performance. An estimate of α that is not statistically significantly greater than zero is consistent with a situation where treatment is effective but funds are inefficiently allocated, as well as with a situation where treatment is not effective at the margin. The idea is that even if treatment is effective there could still be waste caused by a misuse of public money, either because of a bad allocation across agencies, or a bad allocation across patients within an agency, or even because the state is overfunding treatment programs in general.

This analysis is not able to determine the reason for the low value of α. The literature on substance abuse is, however, rich in studies of treatment effectiveness and overall, despite problems with many studies, has concluded that treatment is effective at least for some patients. A relevant study is Lu and Mcguire (1997). They use a patient-level dataset, also from outpatient programs in Maine, and ask whether more units of treatment lead to a reduction in the frequency of use of the preferred substance. Units of treatment can be regarded as an approximation to funds at the patient level, although one unit of treatment does not cost the same for all patients neither in the same programs nor across programs. They claim that there is some evidence of marginal treatment effectiveness for the more severe patients, although there is no marginal impact of treatment on the least severe. However, the positive effect on the more severe patients disappears when they control for the interaction of "units of treatment" with "time in treatment".

This paper's results in conjunction with the inconclusive results of Lu and McGuire raise a number of policy issues. In the first place, OSA should assess the added value (e.g. through a comprehensive cost-effectiveness study) of the outpatient drug-free treatment programs that it supports. Secondly, it should investigate whether agencies are implementing treatment adequately. And third, OSA should review and monitor agencies' expenditures to make sure that scarce public funds are not being devoted to x-inefficiencies. As an alternative to monitoring, OSA could minimize any waste by implementing incentives (such as PBC) more effectively.

Regarding the possibility of aggregation bias, we will argue that if aggregation causes any bias on the estimate of α, it is more likely that it is an upward bias. An important reference is Hanushek et al. (1996) study of the impact of aggregation in the size of omitted variables biases. Although the authors caution about the lack of a clear prediction for most models, they show in a simple two-variable model that if the omitted variables are at the same level of aggregation as the data used (e.g. in our case program characteristics or average patient characteristics), aggregation will bias estimates upwards. Furthermore, they show evidence of this effect by comparing the estimated impact of school resources on school performance

[25] To study which patients are more cost-effective we would need reliable information on (at least) treatment costs at the patient level. Costs at the patient level can be computed from information reported in the MATS discharge forms. We have found, however, that these data are unreliable.

using both data aggregated at the state level and data disaggregated to the school level. Finally, they claim that it is plausible that missing state level variables are relevant for performance since "the key policies are made at the state level".

In the case of outpatient treatment in Maine, it is also likely that the missing relevant variables are at the program level rather than at the patient level. As the more recent literature on treatment effectiveness has shown, program/agency characteristics (e.g. location, facility, staff, director, funding, staff enthusiasm and opinion about the program, treatment philosophy, etc.) are very important in explaining performance. An interesting example is Ball and Ross (1991) study of six methadone treatment programs. Furthermore, OSA is likely to observe these relevant program/agency variables and to take them into account in the allocation decisions, which is the potential source of omitted variable (endogeneity) bias.

In addition to the arguments above, we can prove that in the context of our logistic "production function" without endogeneity, aggregation would, most likely, lead to an upward bias in the impact of funding on the average probability of success. [26]

Ultimately, we think that the concern with aggregation boils down to the estimation of an average marginal effect that may overshadow large individual marginal effects. In our results, however, the estimated marginal effect is so small that if large individual marginal effects exist it must be that programs are offering counterproductive treatment to a large proportion of the patient population, which is unlikely.

7. Conclusions

This paper estimates the marginal impact of public funds on the abstinence rates of non-profit outpatient treatment programs for alcohol abusers in the state of Maine, from 1991 through 1994. The premise is that, given the scarcity of public funds, a marginal increase in the allocation of funds to these treatment programs should bring an increase in their performance; otherwise, the state would be better off by reallocating money to other state programs with positive marginal returns.

This paper used an IV methodology to deal with the potential endogeneity of expenditures per patient. Endogeneity of expenditures is likely if the authorities use the funding allocation to compensate agencies for particularly hard situations, in which case programs that treat more difficult patients receive a larger allocation per patient.

This paper's results indicate that the marginal impact of expenditures per patient on the abstinence rates of outpatient programs is not significantly different from zero. More importantly, the estimates are so small that we may say that they are not economically significant. To illustrate this point, take our most reliable and optimistic estimate of α ($\hat{\alpha} = 0.0002$) plus 2.5 times its standard error (i.e. $\hat{\alpha} = 0.001331$), this is the highest value of $\hat{\alpha}$ that we are not able to reject at the 1% confidence level. Next, suppose that all programs are equal to the average program. In this case, the representative program treats 79.76 patients per period, discharges 38.5 patients at the end of the period, and, on average, has

[26] As an illustration, take programs 1 and 2: $\partial\bar{p}/\partial\bar{c} = (\partial((p_1 + p_2)/2)/\partial((c_1 + c_2)/2)) = (\partial p_1/\partial c_1) + (\partial p_2/\partial c_2) = \alpha_1 p_1(1 - p_1) + \alpha_2 p_2(1 - p_2) > \max\{\alpha_i p_i(1 - p_i),\ i = 1, 2\}$.

a probability of success equal to 0.72. With this set of values, the average cost of obtaining one more abstinent patient in the state of Maine is US$ 615,801.80, which is substantially greater than the average *accumulated* expenditures per patient of US$ 1537.80.[27]

In conclusion, although the results are quite striking, we recommend further research to determine the use that treatment agencies are making of the public funds before deciding on a budget cut to alcohol abuse treatment programs, since it is possible that the marginal dollar is having a positive impact on other treatment goals that the state of Maine considers worthy of public funding.

Acknowledgements

I would like to thank Michael H. Riordan, Kevin Lang, and Thomas G. Mcguire for their suggestions, Margareth Commons and Mingshan Lu for their cooperation with the data, Jereal Holley from the Maine Office of Substance Abuse, and Daniel A. Ackerberg, Stephen Donald, Samuel Kortum, Alwyn Young, Juan Ruiz, Maite Martinez and Berthold Herrendorf for their helpful comments. I would also like to thank two anonymous referees. The author benefited from the financial support of a Alfred P. Sloan Dissertation fellowship, and from programs CIÊNCIA and PRAXIS (JNICT, Portugal). The author also acknowledges the support from the National Institute on Drug Abuse (Grant 1 R01 DA08715-01).

References

Andersen, E.B., 1973. Conditional Inference and Models for Measuring. Publikation nr. 5, Mentalhygiejnisk Forskningsinstitut.

Apsler, R., Harding, W.M., 1991. Cost-effectiveness of drug abuse treatment: current status and recommendations for future research. In: Drug Abuse Services Research Series. 1. Background Papers on Drug Abuse Financing and Service Research. DHHS Publication no. (ADM)91-17777, National Institute on Drug Abuse, Washington, DC, pp. 58–81.

Ball, J.C., Ross, A., 1991. The Effectiveness of Methadone Maintenance Treatment: Patients, Programs, Services, and Outcome. Springer, New York.

Bowden, R., Turkington, D., 1984. Instrumental Variables. Cambridge University Press, Cambridge.

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., York, R., 1966. Equality of Educational Opportunity. Government Printing Office, Washington, DC.

Commons, M., McGuire, T.G., 1997. Some economics of performance-based contracting for substance abuse services. In: Egertson, J.A., Fox, D.M., Leshner, A.I. (Eds.), Treating Drug abusers Effectively. Milbank Memorial Fund and Blackwell Publishers, Oxford, pp. 223–249.

Commons, M., McGuire, T.G., Riordan, M.H., Spanfelner, K., 1994. Review of the Outcome Measures in the Maine Performance-Based Contracting System for Substance Abuse Services. Mimeo, Boston University.

[27] Taking the average program as representative and using 'd' for total derivative:

$$1 = \mathrm{d}Y = nE(\mathrm{d}Y_i) = nD\frac{\mathrm{d}p}{\mathrm{d}\bar{c}}\mathrm{d}\bar{c} = nD\frac{\alpha}{N}p(1-p)\,\mathrm{d}\bar{c} = nD\frac{\alpha}{N^2}p(1-p)\,\mathrm{d}c \Rightarrow n\,\mathrm{d}c = \frac{N^2}{\alpha Dp(1-p)}.$$

Replace the average number of patients $N = 79.76$, the average number of discharges $D = 38.5$, and the average probability of success $p = 0.72$, and using $\alpha = 0.0002 + 2.5 \times 0.0004525 = 0.001331$, we get $n\,\mathrm{d}c = \text{US\$}\,615{,}801.80$.

Commons, M., McGuire, T.G., Riordan, M.H., 1997. Performance-based contracting for substance abuse treatment. Health Services Research 32 (5), 631–650.
Cox, D.R., 1970. The analysis of binary data. In: Methuen's Statistical Monographs. Methuen & Co. Ltd., London.
Ferguson, R.F., 1991. Paying for public education: new evidence on how and why money matters. Harvard Journal on Legislation 28 (2), 465–498.
Greene, W.H., 1993. Econometric Analysis, 2nd Edition. Macmillan, New York.
Hanushek, E., 1986. The economics of schooling: production and efficiency in public schools. Journal of Economic Literature 24 (3), 1141–1177.
Hanushek, E., et. al., 1994. Making Schools Work. The Brookings Institution, Washington, DC.
Hanushek, E., Rivkin, S.G., Taylor, L.L., 1996. Aggregation and the estimated effects of school resources. Review of Economics and Statistics 78 (4), 611–627.
Hausman, J., 1978. Specification tests in econometrics. Econometrica 46 (6), 1251–1271.
Lang, K., Somanathan, R., 1997. Estimating the Effects of School Expenditure on Student Achievement: An Instrumental Variable Approach. Mimeo, Boston University.
Long, C.G., Williams, M., Hollin, C.R., 1998. Treating alcohol problems: a study of programme effectiveness and cost-effectiveness according to length and delivery of treatment. Addiction 93 (4), 561–571.
Longabaugh, R., McCrady, B., Fink, E., Stout, R., McAuley, T., Doyle, C., McNeill, D., 1983. Cost-effectiveness of alcoholism treatment in partial versus inpatient settings — six-month outcomes. Journal of Studies on Alcohol 44 (6), 1049–1071.
Lu, M., Mcguire, T.G., 1997. The Productivity of Outpatient Treatment for Substance Abuse. Mimeo, Boston University.
Machado, M.P., 1997. Cost-effectiveness of the public provision of substance abuse treatment in Maine. PhD Dissertation, Boston University.
Marlatt, G.A., 1996. Harm reduction: come as you are. Addictive Behaviors 1, 779–788.
McLellan, A.T., Woody, G.E., Metzger, D., McKay, J., Durell, J., Alterman, A.I., O'Brien, C.P., 1997. Evaluating the effectiveness of addiction treatments: reasonable expectations and appropriate comparisons. In: Egertson, J.A., Fox, D.M., Leshner, A.I. (Eds.), Treating Drug Abusers Effectively. Milbank Memorial Fund and Blackwell Publishers, Oxford, pp. 7–40.
Mroz, T., 1987. The Sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. Econometrica 55 (4), 765–799.
Neyman, J., Scott, E.L., 1948. Consistent estimates based on partially consistent observations. Econometrica 16 (1), 1–32.
Rice, D.P., Kelman, S., Miller, L.S., 1991. Economic costs of drug abuse. In: Cartwright, W.S., Kaple, J.M. (Eds.), NIDA Research Monograph 113. US Department of Health and Human Services, Rockville, MD, pp. 10–32.
Sells, S.B., Simpson, D.D. (Eds.), 1976. Effectiveness of drug abuse treatment. Vol. 3. Ballinger, Cambridge, Massachusetts, USA
Summers, A.A., Wolfe, B.L., 1977. Do schools make a difference? American Economic Review 67 (4), 639–652.
Walsh, D., Hingson, R., Merrigan, D., Morelock Levenson, S., Cupples, A., Heeren, T., Coffman, G., Becker, C., Barker, T., Hamilton, S., McGuire, T.G., Kelly, C., 1991. A randomized trial of treatment options for alcohol-abusing workers. The New England Journal of Medicine 325 (11), 775–782.