

## USING EXISTING DATA BASES: A CASE STUDY

Errol Porter  
Ministry of Community and Social Services  
Toronto, Ontario

**Abstract** — In conjunction with a comprehensive program evaluation, a separate project using existing data was initiated to assess Ontario's Employment Support Initiatives (ESI) program. Its original purpose was to validate parts of the overall evaluation. However, it was found to have another major use: inexpensive assessment of the long-term impact of ESI on recidivism to income maintenance. This article describes the creation and use of a new data base that combined demographic and biographical data and the history on income maintenance over five years for the ESI clients and a randomly selected comparison group.

**Résumé** — Dans le cadre d'une évaluation globale du programme, un projet a été mis sur pied pour évaluer, à partir des données existantes, le programme ontarien d'Initiatives en matière de soutien à l'emploi (I.S.E.). Le but premier de l'étude était de valider certaines parties de l'évaluation globale. L'étude, toutefois, a également permis d'obtenir un deuxième résultat imprévu: une étude à bon compte de l'influence à long terme du programme d'I.S.E. sur la récidivité au recours aux programmes de maintien du revenu. Le présent article décrit la création et l'utilisation d'une nouvelle base de données réunissant des données démographiques et biographiques et des renseignements sur les antécédents relatifs au recours au maintien du revenu, glanés sur une période de cinq ans, de clients du programme d'I.S.E. et d'un groupe choisi au hasard pour fins de comparaison.

**DATA COLLECTION IS USUALLY** the most time-consuming and expensive part of social-program evaluation. This fact provides an incentive for doing secondary analysis of evaluation data and for seeking other data sources. A potential source of data for evaluation, one that has not been fully exploited in many areas (Burststein, 1984), can be found in administrative and management data banks and information systems. Such administrative data is collected by agencies in the normal course of their work, to assist in the administration of a program.

As the demand for evaluations increases, there will be greater pressure to utilize these existing data bases. This raises two issues:

1. The inadequacy of many existing data bases and information systems;
2. The need to design evaluations that complement and make appropriate use of existing data bases.

These complementary aspects of evaluation and the utilization of existing data need to be carefully considered. Failure to do so will result in evaluations that tend to be more expensive, laborious, and time-consuming, produce less conclusive information, and are less able to address the issues of concern (Chelimsky, 1985). Existing administrative data may be of use in all phases of most evaluations, and in some cases could provide the entire data set for an evaluation or allow a more realistic and efficient evaluation to be designed.

Researchers have only begun to realize the potential benefits of using existing data sources in the evaluation of employment programs. This study reports on a part of a much larger project—the evaluation of Ontario's Employment Opportunities Programs (EOP). EOP consists of seven distinct programs that provide employment services or job subsidies for youth or single parents on social assistance. Each program is delivered by numerous municipalities across Ontario. This article is concerned with one of these programs, the Employment Support Initiatives (ESI) Program.

## **ESI PROGRAM**

The ESI Program began in early 1983 as a pilot program operated by nine municipalities. Its main goal is to assist single parents on welfare or income maintenance to find full-time employment and achieve financial independence. Clients are provided with employment-related services such as upgrading of job skills, life-skill training, pre-employment services, and referral to educational and work-skill training programs, as well as funds for child care and employment-related costs.

There have been several studies of the ESI program and of the clients that entered the program in its first year of operation, including process and outcome evaluations (Community Concern Associates Ltd., 1986). Cooper's (1986) survey of the ESI clients who entered the program in the first full year of operation ending March 1984, was carried out in the spring of 1986 and will be referred to as the field survey. It described the first-year clients as being more likely to be employed full-time and to have been in receipt of social assistance for a shorter time than a comparison group, although the differences were not large. The study was limited by the low tracing rate and the even lower interview-completion rate for the two groups; these were thought to indicate that the most successful persons may not have been found in the telephone survey, and therefore that program effects might have been underestimated.

Our study was originally undertaken to assess the program with a more complete sample of the program's clients and to compare the results with this earlier study. We found that the approach can be useful in measuring long-term impact of employment programs and can provide information that might not be efficiently obtainable through the use of a client survey.

## **DEVELOPMENT OF THE DATA BASE**

Since it is well known that persons on income maintenance often leave for full-time employment even prior to the establishment of employment programs and incentive schemes (Porter, 1986), a comparison group was deemed necessary if measures of program effectiveness were to be derived. Comparison groups were derived for both the field survey and the study described herein.

The population sample for this study, like that of the field survey, consisted of two groups; all of the ESI clients who entered the program in the first year of operation, and a comparison group. The comparison group chosen was a random sample of single parents who were on income maintenance in October 1983, this point having been chosen as the mid-point in the first year of ESI operation. The comparison cases selected were similar to ESI clients in age, number of children, age of the youngest child, length of time that the recipient had been receiving assistance, and proportion of cases in each of the nine municipalities.

Tracking clients was not an issue in this study, since all persons in both groups were involved in tape matching. The IDs of ESI clients (as supplied by the municipalities) and comparison clients were matched to the computer tapes quarterly from September 1982 to December 1986. The ESI sample was 4,325; the comparison sample, 3,015. The field survey included 1,137 ESI and 444 comparison cases. This study involved 3,188 more ESI clients and 2,571 more comparison cases than did the field survey, which addition increased data reliability and reduced the possibility of biased results due to the difficulty of tracking clients.

The matching process resulted in a new data base consisting of two parts: biographical data (recipients' age, education, employment variables, family data, and location), and time-series data for each quarter (whether the person received assistance, the type of assistance received, the amount received, and the amount earned from part-time jobs while on assistance). The data base was downloaded to a microcomputer for analysis. Some tape matching of the list of program participants was not possible because of errors in the IDs submitted to the Ministry by the municipalities and partially incompatible data systems

in two of the municipalities. However, this affected only a small proportion of the ESI cases and none of the comparison cases, and did not bias the results.

Because of the large sample sizes used in this study and the fact that the ESI group consists of almost all the cases that were in the program in the first year of operation, sampling errors are generally very small. Most results are statistically significant, and highly so. Statistical significance does not, of course, necessarily imply policy significance.

## **FINDINGS OF THE STUDY**

The following findings reflect the kind of analysis that is possible with the derived data base.

### **Characteristics of the Sample**

There were no large differences in most of the characteristics of the two groups, because the comparison group was selected to be as similar as possible to the ESI client group. The main demographic characteristics of the ESI and comparison samples in the field survey were not significantly different: The average age was 33.5 years, the average age of the youngest child at home was 8.0, and the average number of dependent children per case was 1.7.

The characteristics of the total sample varied somewhat from those of the field survey. In particular, ESI cases were somewhat younger (34.7 years as compared to 36.3 for comparison cases), the youngest child of ESI cases was somewhat older, and the average number of dependents of the two groups was identical (1.61).

### **General Trends of ESI and Comparison Cases on Income Maintenance**

Persons who lose their jobs and run out of unemployment insurance or who have the responsibility of raising their children without the financial assistance of a spouse may apply for welfare or family benefits. The ESI program provides those wishing to enter or re-enter the labor market with a means to do so. Based on a time-series analysis of the persons who entered the ESI program in its first year of operation and of comparison cases, the rate of change from quarter to quarter for the two groups showed that persons from the two groups left income maintenance at about the same rate over the period of March 1983 to December 1986.

## **Overall Program Effectiveness**

The overall effectiveness of the ESI program was indicated by a reduction in clients' reliance on income maintenance. Other measures, such as achieving employment, were not available in the data base. A number of alternative time periods were examined, including month-to-month comparisons and comparisons averaged over a year.

The reduction in caseload varies between the two groups according to the time period chosen. Because of the small program effects, the apparent impact on caseload was sensitive to the time period chosen for the analysis. For example, the income maintenance caseload reduction from March 1984 (the end of the first year of ESI) to December 1986 was 37.6% for persons who participated in ESI, but 35.4% for the comparison group. (This points out an advantage of having a comparison group; before this study the media quoted only the ESI reduction.) Quarterly fluctuations make it possible to choose time periods whereby one of the groups seems to have a greater reduction than the other. However, after analyzing a number of periods, we concluded that program impact was small.

## **Site Analysis**

An analysis was done of the income maintenance caseload changes for the nine ESI sites extant in the program's first year. Three of the nine municipalities had a greater caseload reduction for ESI clients than for cases in the comparison group, regardless of the time frame used in the analysis. Another three municipalities showed an ESI caseload reduction that was greater than that for the comparison group for particular time frames. When only cases on assistance in September 1983 were compared seven months later, seven of the nine municipalities had greater reductions for ESI clients than for comparison cases. The original field survey, for reasons of sample-size limitation, was unable to provide a site analysis.

## **Relation of the Results to the Field Survey**

The data collected in this study helped to clarify the results of the field survey. In particular, two issues relate directly to the field survey:

1. Only 43% of the sample of ESI clients and 57% of the comparison cases were tracked in the telephone survey preceding the interview phase. The interview completion rate was 32% and 28% for the ESI and comparison groups respectively. It was suggested that the most

successful ESI clients eluded the survey (because their most recent addresses were no longer included in income maintenance records, and some persons had changed addresses a number of times). If this is so, the survey would understate the success of the program.

2. The comparison sample in the field survey was screened in an attempt to ensure that this group would closely match the ESI group with respect to motivation to work. Specifically, in order to be included in the comparison group, an individual had to be either currently employed, actively seeking employment, currently enrolled in an educational/job training course, or planning to enrol in such a course within the next year. The screening process eliminated approximately one-third of the persons who were contacted from the original random sample. Although a previous study (Sheridan, 1983) had shown that more than one-third of the interviewed single parents on provincial social assistance wanted to care for their children rather than work, some critics believed that it also created a comparison group that was better qualified, more motivated, or more "job ready" than the ESI group. Again, this would tend to obscure the effects of participation in the ESI program.

Within the ESI group, 33.7% of those interviewed in the field survey were off assistance by December 1986, compared to 49.3% of those not interviewed. The comparison group was similar: 29.2% of those interviewed were off assistance, compared to 42.0% of those not interviewed. Non-interviewees in both the ESI and comparison groups were significantly less likely to be receiving social assistance than those interviewed. Thus, in *both* groups, those not interviewed were more successful, at least insofar as they were not in receipt of assistance. The observed program impact is somewhat greater if all cases are included in the analysis than if only interviewed cases are included.

The municipalities' ratings of employability, employment prospects, job skills, education, and health were available for a majority of clients. The analysis of this data showed that:

1. The ESI group was rated as more employable than the comparison group. Specifically, 58.8% of the ESI group was rated as employable, as compared with 47.9% of the comparison group. The ESI group also had more highly rated employment prospects.
2. The ESI clients had more job skills than the comparison cases. For example, 9.2% of the ESI group and 6.2% of the comparison group were rated as "skilled."

3. The level of education of the ESI clients was generally higher than that of the comparison group. In particular, 9.4% of the ESI clients and 5.4% of the comparison cases had post-secondary education; 37.5% of the ESI cases and 30.5% of the comparison cases had completed Grades 11 to 13. Significantly more of the comparison group than the ESI group had less than Grade 10 education.
4. ESI clients had significantly better health than the comparison cases. The health of 76.6% of the ESI group and only 67.0% of the comparison group was rated as "good."

Comparison of those interviewed and not interviewed in the two groups revealed that persons not interviewed in the field survey were less likely to be on income maintenance than those interviewed, although they were not necessarily employed. The comparison group was less employable, had poorer job prospects, had fewer job skills, had less education, and was in poorer health than the ESI group. These findings, based on municipal ratings, refute the contention that the comparison group was better qualified for jobs than the ESI group, or was more employable. In addition, persons in the comparison group who were interviewed in the field survey had lower ratings for employability, job prospects, job skills, education, and health than the ESI clients who were interviewed. This similarly rebuts the contention that the comparison cases interviewed were more employable or were better qualified. The screening approach used in the survey did not, in fact, screen out enough of the comparison cases.

### **Employment-Related Characteristics and Reliance on Income Assistance**

As expected, persons in either group who were described as more educated, in good health, having job skills, or being employable were more likely to be off income maintenance. In fact, the more education the person had, the less likely it was that that person would be on income maintenance by the end of the follow-up period. Persons in the comparison group who were rated as "employable" or as having "good" employment prospects or as having "good" health were more likely to have left income maintenance in the period September 1983 to December 1986 than ESI clients with the same rating. A large number of persons in both the ESI and comparison groups who were on social assistance in December 1986 were considered to be "employable," to have "good" health, or to have "good" employment prospects, and could be considered as candidates for employment or training programs.

## Cost Savings Analysis

Between December 1984 and December 1986, 58.6% and 47.5% of the ESI and comparison cases respectively had left social assistance. Although more of the ESI cases than comparison cases left social assistance in this period, of those that left, 28.8% of the ESI leavers and 18.6% of the comparison group leavers had returned one or more times by December 1986. Of the original samples, 45.6% of the ESI group and 40.0% of the comparison group were off by December 1986.

The following demonstrates a simplified cost-savings analysis done with the available data. The Ministry of Community and Social Services (MCSS) cost reduction for the ESI and comparison group was based on the number of months off income maintenance in the period December 1984 to December 1986, the number of persons off assistance during that period, and the average monthly cheque of these cases just prior to their leaving. The cost reduction was defined as the amount of money that MCSS would have paid the person if that person had been on income maintenance. Because group sizes differed, it was necessary to pro-rate the comparison group's cost reduction to represent an amount that would be equivalent to a comparison group of the same size as the ESI group. The savings generated, based on the difference in cost reduction between the two groups, are as follows:

- |   |               |
|---|---------------|
| • Total cost reduction to MCSS<br>for the ESI group (A)   | \$ 21,401,858 |
| • Total cost reduction to MCSS<br>for the comparison group if it<br>were the same size as the ESI group (B) | \$ 19,997,677 |
| • Gross savings (A - B)   | \$ 1,404,181  |
| • Less the cost of the ESI program<br>for the first year of operation                                       | \$ 3,259,657  |
| • Net cost to MCSS  | \$ 1,855,476  |

The cost of the first year's operation of the ESI program was used in this calculation because it was not known how much time the clients in the ESI program's first year of operation spent in the program in second and subsequent years. It is, however, known that 49% of the first year's cases were carried over into the second year of operation (Community Concern Associates Ltd., 1986). The cost of the program shown in the preceding calculation should be substantially increased to take this into account. The fact that the ESI cases,



having better education and health, were more likely to obtain jobs would also tend to increase the net cost. On the other hand, the calculation only takes into account nine quarters or 27 months of savings — a longer follow-up period should yield a larger gross savings. It is important to note that the cost-savings obtained from a later cohort of persons in the ESI program was substantially better (A.R.A. Consultants, 1987), indicating either a program improvement over the first year of operation or a relative change in the mix of clients.

## RESULTS OF ANALYSIS

The following are the main results of the study:

1. The ESI program did not have a large effect on first-year clients' reliance on income assistance. In seven out of nine municipalities, there was a small reduction in the number of ESI clients on income maintenance, relative to comparison cases.
2. There was a bias in the field survey — the more successful clients were less likely to have been interviewed. The demonstrated program impact is somewhat greater if all cases are included in the analysis than if only interviewed cases are included. Although the absolute impact on ESI is underestimated in the field survey, the impact relative to the comparison group described in that study is still a reasonable indicator of program effect.
3. Persons in the comparison group were less employable, had poorer job prospects, had fewer job skills, had less education, and were in poorer health than the ESI cases. This refutes the contention that the comparison group was better qualified or more employable than the ESI group. In fact, the opposite is true. More educated or employable persons in both the ESI and comparison groups were more likely to be off income maintenance.
4. The data base provided a way to analyze basic cost-savings. When the cost of the *first year's* operation of the ESI program and the case cost reductions generated over the December 1984 to December 1986 period are taken into account, there is a net cost to the Ministry.

## Advantages

Our approach, which can be used either for a stand-alone project or to complement or validate a large data-collection effort, has a number of advantages:

1. Data provided can be made more accurate and reliable than that obtained through costly client surveys by inexpensively increasing the sample sizes or even by using the total population.
2. The resulting data would be more representative of the programs, since all clients would be included, not just clients with telephones or those in the sites that cooperated with the data collection. As well, there would be no tracing loss of clients if the needed data were on the system.
3. Information is more valid because it does not depend on the clients' memory of when they were in the program. Interpretation is often straightforward, because the data available is usually behaviourally anchored and relatively simple (unfortunately, often too simple).
4. Site-by-site analysis can be done so that the effectiveness of specific program models can be assessed. As well, clients with specific characteristics can be assessed, making the approach useful for defining appropriate target groups.
5. Longer term impact can be readily assessed.
6. By forming longitudinal records of each client, researchers can establish their quarterly or monthly history. Accurate measures of their leaving the program and their recidivism can then be developed for the program.
7. The information produced would be complementary to that produced in client surveys, filling some of the gaps that might exist in the latter type of study (such as those due to small sample sizes and data at only one or possibly two points in time), and yielding more comprehensive and reliable information for planning purposes. This is one example of the advantages possible in using multiple data sources, which can often be used to reinforce conclusions.
8. This approach helps overcome the problem of finding eligible non-recipients of the program when forming comparison groups.
9. Researchers can evaluate employment programs more cost-effectively in the future by optimally combining this and other evaluation methods.

## Disadvantages

1. The analysis is limited to the data in the administrative information system. Existing data bases usually do not have the full range of variables that could be collected in a uniformly implemented client survey. (In this study, the main indicator of program effectiveness was the proportion of cases that were off income maintenance, since it was not possible to obtain information on whether these persons were employed, in training programs, or financially independent for other reasons.)
2. One needs to exercise caution in using administrative data for evaluation and to consider its weaknesses. For example, some sites or offices do not collect certain non-compulsory data. Other problems could include misidentification of individuals, incomplete data, and double counting of individuals.
3. Variation from site to site could exist (there is often little communication among sites, as well as lack of standardization, especially in decentralized organizations).
4. Difficulties with client-matching from period to period can occur when suitable client identifiers are not used.

## CONCLUSIONS

If the overall needs for information were more carefully considered in the development of information systems, the cost of evaluations could be reduced and evaluation needs could be better met. As well, greater care in designing and implementing information systems to ensure consistent definition of variables and categories would provide more reliable information.

The study described in this article provided reliable and inexpensive information. Since the effects of program improvements and staff experience were not reflected in the first-year results, further studies will include an analysis of later cohorts of the ESI program as well as a similar analysis of the Youth Employment Preparation Program and the Social Services Employment Program. Besides being useful for validating client surveys, this study demonstrated an approach for assessing long-term impact of a program based on existing management information systems.

## REFERENCES

- A.R.A. Consultants. (1987). Summary of evaluation results and cost-benefit analysis of employment opportunities program (Employment Opportunities Program Evaluation Report, Vol. 10). Toronto: (Emmott et al.)
- Burstein, L. (1984). The use of existing data bases in program evaluation and school improvement. *Educational Evaluation and Policy Analysis*, 6, 307-318.
- Chelimsky, E. (1985). Program evaluation and the use of extant data. In L. Burstein, H. E. Freeman, & P. H. Rossi (Eds.), *Collecting evaluation data: Problems and solutions*. Beverly Hills, CA: Sage.
- Community Concern Associates Ltd. (1986). *ESI: Update descriptions of the original projects* (Employment Opportunities Program Evaluation Report, Vol. 1). Toronto: (Haggarty et al.)
- Coopers & Lybrand Consulting Group (1986). ESI: Follow-up survey of clients who entered in its first year. (Employment Opportunities Program Evaluation Report, Volume 2), Toronto: (L. Bozinoff et al.)
- Porter, E. (1986). *Follow-up report of cases that left the FBA program in the period September 1, 1978 to August 31, 1979: An analysis of existing data*. Research and Program Evaluation, Ministry of Community and Social Services. Toronto.
- Sheridan College of Applied Arts & Technology (1983). Employment Readiness of Selected FBA Recipients. Toronto: (G. Kohler et al.)