Modeling a Public Employment Office: Report on a Pitfall

Raul S. Andrade
Pontifícia Universidade Católica do Rio de Janeiro

ABSTRACT

How and why an attempt to build a system dynamics model of the operation of a public employment office resulted in an econometric model.

INTRODUCTION

In 1972-1973, when this work was done, unemployment was high and public employment offices at some states were expected to play a part in alleviating this social problem. At the same time, they were criticized for their performance. The original idea of this research was to develop an operational measure of employment office performance which would take into account the special restraints to which such office is subjected. Besides the obligation to help all those that walked in looking for a job, a public employment office was sometimes the agent of welfare programs which ranged from foodstamps to worker retraining.

While recognizing that, because of its social function, a public employment office has in fact a multitude of different objectives, the focus of the study was on its ability to produce placements, i.e., to successfully match an applicant and a job opening. The evaluation of placement performance is complicated because when the economy recedes job openings decline while applicants increase, creating an extended backlog of unemployed clients in the office. The job of matching an applicant and an opening becomes more difficult and it may appear that placement performance degrades when in fact the office personnel may be doing a better job.

MODELING

Such performance measure was never developed but the model described in Figure 1 was built as an instrument for its posterior development. It includes what seemed to be the most important components of the system: job openings, applicants, and staff time. Andrade (1973) worked as a volunteer at a public employment office and interviewed its personnel in order to enhance his understanding of the system. (Experienced system dynamicists will promptly verify that this model violates some good modeling rules.)

Monthly data for a period of two years was available aggregating a small number of offices at community level from July 1970 to July 1972. Econometrics was used to estimate the parameters in the model and to check the significancy of each relationship. Two results were particularly remarkable. The first has to do with the number of job referrals (JOBREF) made by the office, an auxiliary variable of paramount importance. This variable was modeled as a production function. Job referrals were seen as consequence of three factors of production: openings, applicants and amount of staff time spent by the office

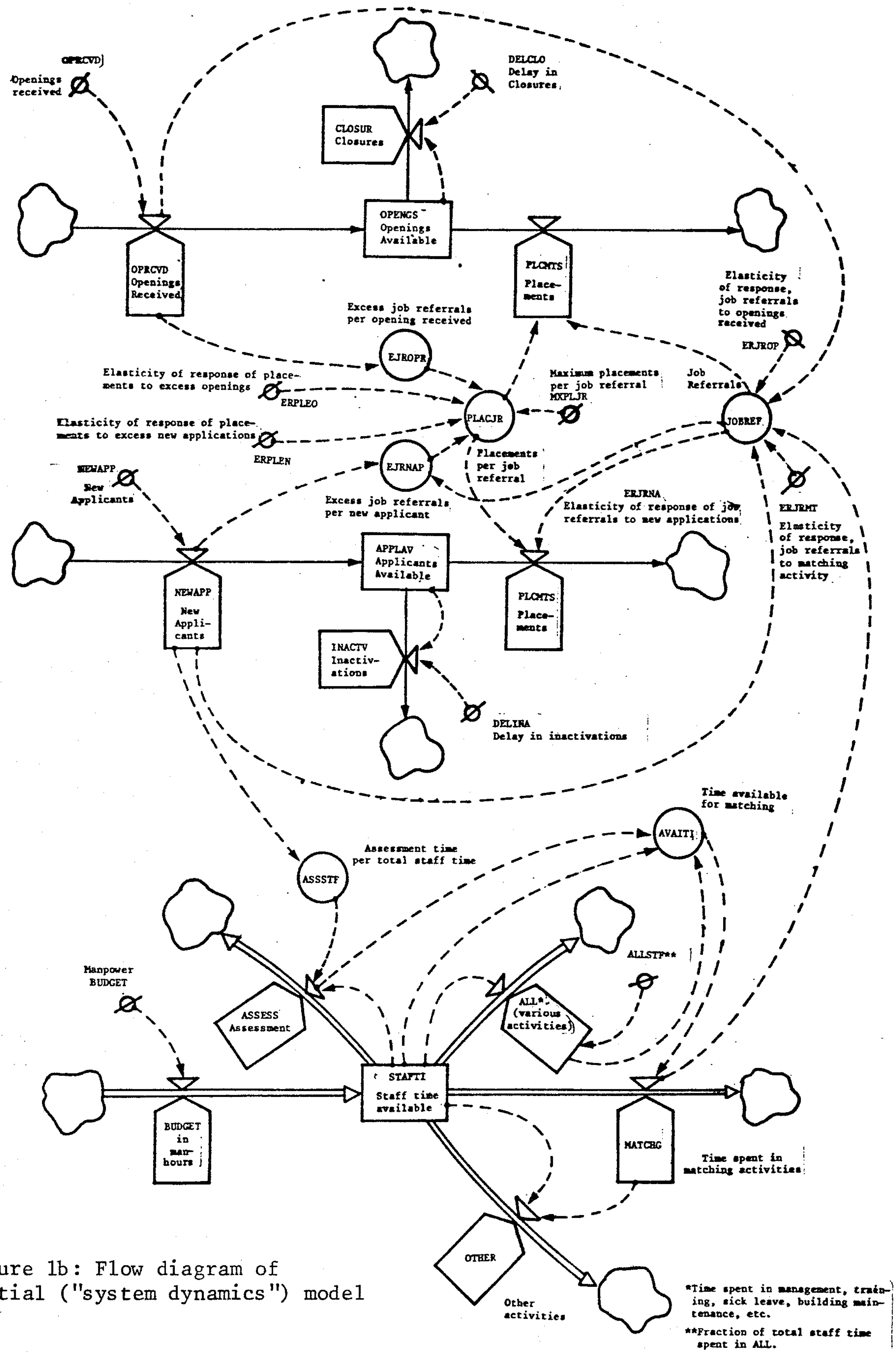


Figure 1b: Flow diagram of initial ("system dynamics") model

Figure 1b: Equations of the initial ("system dynamics") model:

```
 (1) PLCMTS.JK = (PLACJR.JK)(JOBREF.JF)

 (2) PLACJR.JK = (MXPLJR)*EXP(ERPLEO)(EJROPR.JK)*EXP(ERPLEN)(EJRNAP.JK)

 (3) EJROPR.JK = (JOBREF.JK - OPRCVD.JK)/(OPRCVD.JK)

 (4) EJRNAP.JK = (JOBREF.JK - NEWAPP.JK)/(NEWAPP.JK)

 (5) JOBREF.JK = (SRFCTR)(OPRCVD.JK)**ERJROP*NEWAPP.JK**EJRNAP*MATCHG.JK**
                                                                  ERJRMT

 (5) NEWAPP.JK = CONNAP+(NAPOPR)(OPRCVD.JK)+(NAPPLC)(PLCMTS.IJ)+
                  (NAPNAP)(NEWAPP.IJ)+(NAPCMS)(COMMSV.IJ)+(NAPWUI)(UNEMPL.JK)

 (6) OPRCVD.JK = CONOPR+(OPRPLC)(PLCMTS.IJ)+(OPRSTF)(DT)(STAFTI.J)+
                                                   (OPRVAC)(VACANC.JK)

 (7) APPLAV.J  = APPLAV.I + (DT)(NEWAPP.IJ - INACTV.IJ - PLCMTS.IJ)

 (8) OPENGS.J  = OPENGS.J + (DT)(OPRCVD.IJ - CLOSUR.IJ - PLCMTS.IJ)

 (9) INACTV.JK = (DT/DELINA)(APPLAV.J)

(10) CLOSUR.JK = (DT/DELCLO)(OPENGS.J)
```

staff trying to match openings and applications (MATCHG). Surprisingly, the effect of this last variable on the number of job referrals was statistically non-significant. This means that the number of job referrals made did not depend on how much effort to match applicants and openings is made by the office staff. Employment service activities seemed to be irrelevant to placement performance. For a moment, Parkinson's Law seemed to be reigning. Or else, the staff time recording system was unreliable.

The econometric analysis also showed that the number of placements did depend directly on the number of job referrals, the success of which depended on the number of openings available. As the number of job referrals depart from the number of openings received, the productivity of job referrals decreases and vice-versa.

The staff time spent in matching activities correlates with the number of new applications and with the total staff time availble only. The remaining activities depend on total staff time only. The number of openings received was directly related to the number of placements made during the previous month. The number of new applications was related to the number of openings received during the same period.

If the results of the econometric analysis are taken seriously and a revised model is built on their basis, Figure 2 is the result. Only two variables influence the number of placements in this model, the number of job referrals and the number of openings received. It turns out that this is an econometric model with all relevant parameters estimated from monthly data. There are no delays in this model except for the processes of inactivating applicants and closing openings.

This model has been simulated and validated by comparison of its output with the actual time data that generated it. Theil's Inequality Coefficients (U) and R-squares were computed and the results were good.

In summary, a model initially contrived as a system dynamics model had its equations parameters estimated by econometric techniques. Some of the relationships turned out to be statistically non-significant and were dropped out of the model. But the resulting model lost the distinctive features of a system dynamics model. It became an econometric model.

DISCUSSION

In good system dynamics practice the number of elements per unit of time leaving a level from a first-order exponential delay must be a function of the number of elements in the level at the beginning of the period. It must not be a function of the rate at which new elements enter it. In order for an element to leave a box, it must already be in that box. The instantaneous depletion rate R2 must not be formulated as a function of the speed, R1, at which new elements enter the box (Figure 3). If the speed at which new elements enter the box influences the speed at which the level is depleted, then another level must be created to represent the perceived speed. The speed of depletion R1 can only depend on the perceived rate of accumulation upon a time lag, that is, a delay.

Figure 2a: Flow diagram of the revised ("econometric") model

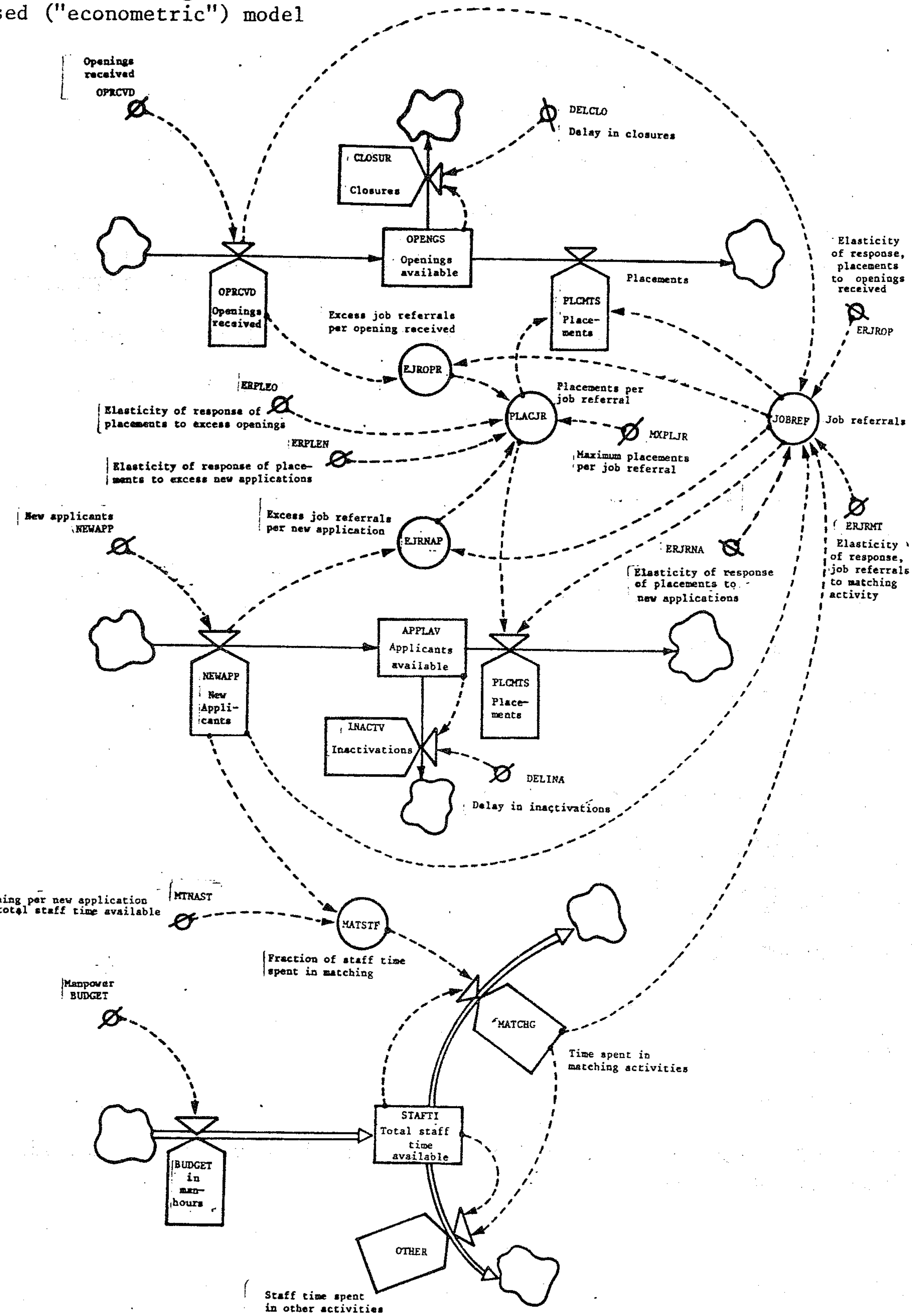


Figure 2b: Equations of the revised ("econometric") model:

(1) PLCMTS.JK = (PLACJR.JK)(JOBREF.JK)

(2) PLACJR.JK = (MXPLJR) * EXP(ERPLEO)(EJROPR.JK)

(3) EJROPR.JK = (JOBREF.JK - OPRCVD.JK)/OPRCVD.JK

(4) EJRNAP.JK = (JOBREF.JK - NEWAPP.JK)/NEWAPP.JK

(5) JOBREF.JK = JRFCTR * OPRCVD.JK**ERJROP * NEWAPP.JK**ERJRNA * MATCHG.JK**ERJRMT

(6) APPLAV.J = APPLAV.I + (DT)(NEWAPP.IJ - INACTV.IJ - PLCMTS.IJ)

(7) OPENGS.J = OPENGS.I + (DT)(OPRCVD.IJ - CLOSUR.IJ - PCMTS.IJ)

(8) INACTV.JK = (DT/DELINA) APPLAV.J

(9) CLOSUR.JK = (DT/DELCLO) OPENGS.J

(10) MATCHG.JK = (MATSTF.JK) (DT) (STAFTI.J)

(11) MATSTF.JK = (MTNAST)(NEWAPP.JK)

(12) MSPLJR = EXP(-0.5629+0.3429*COT)

(13) ERPLEO = -0.6628 + 0.5391*COT

(14) ERPLEN = 0

(15) JRFCTR = EXP(3.492+1.460*COT)

(16) ERJROP = 0.4483+1.339*COT

(17) ERJRNA = 0

(18) ERJRMT = 0

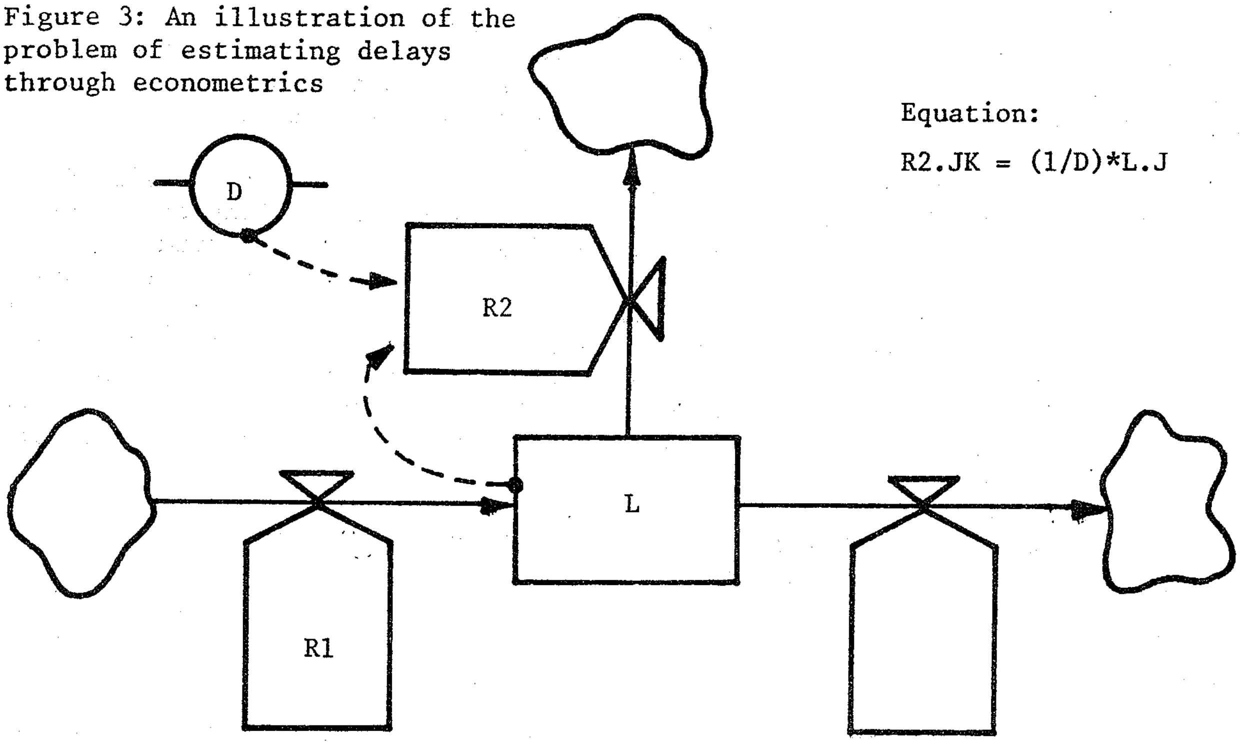
(19) MTNAST = 0.0003546

(20) DELINA = 0.500

(21) DELCLO = 0.500

(22) COT = 0 or 1, depending on order-taking organization, whether decentralized or centralized, respectively.

Figure 3: An illustration of the problem of estimating delays through econometrics



Equation:

R2.JK = (1/D)*L.J

These delays were neglected in the second model (Figure 2). Rates were defined as functions of other rates. This procedure is acceptable in econometric modeling, where the objective is to produce output that closely ressembles the actual data. In system dynamics modeling this procedure does not make sense, as a depletion rate must depend on the level of elements available but not on the rate at which thy become available. But if this statement is true, then depletion rates should correlate poorly with the accumulation rates. Such was not the result of the econometric analysis of the archival data.

One interpretation is that the delays in the system are probably small compared to the period of data aggregation, one month. For example, the mean lifespan of an opening was probably very short, in the order of one week. The mean lifespan of an opening is the delay in filling or closing it. A very large number of openings became available and were filled or closed in an interval much shorter than one month. Thus, the number of openings received or closed during a month was much larger than the number of openings available at the end of the month.

Many of the placements were made on openings that had appeared during the same month. This is why PLCMTS regressed well with OPRCVD.

A consequence of using monthly data to estimate the delays in this system was bias in the estimation. It implies, for example, that the average time a new applicant or opening remains in the neighborhood of a half month. If the data had been collected more often these durations would be shorter and not necessarily equal. The determination of proportionality constants by statistical regression must be made with data collected more often than one month if the delays in these subsystems are shorter than one month. Or else, unaggregate data must be used. Although most parameters in system dynamics studies are estimated on the basis of descriptive information obtained from participants in the system being modeled (Graham 1980), excessive preoccupation with posterior validation completely ignored this important source of data.

Perhaps the greatest mistake was to select model variables on the basis of statistical tests. Mass and Senge (1980) are clarifying: "failing the t-tests tells the modeler that the estimated standard deviation of the estimate is large relative to the estimated parameter value. However neither outcome tells the modeler whether the underlying parameter... is in any sense 'important' for the model being estimated... Failure to pass the t-test means that the available data do not permit accurate estimation of the parameter". It does not mean that the variable in question is unimportant for the model. Thus, it would have been wise to maintain all variables included in the logical a priori model of Figure 1 and to run a model-behavior test.

It can also be remarked that the use of table functions and multipliers would have greatly facilitated modeling.

CONCLUSION

A model initially contrived as a system dynamics model had its equation parameters estimated from aggregate data by econometric techniques. Variable selection was made on the basis of these tests and the result was that the model lost all the distinctive features of a system dynamics model becoming a sheer econometric model instead. The pitfall has been interpreted and some mistakes identified.

REFERENCES

Andrade, R.S. "Employment Service Dynamics: Placement-Oriented Econometric Analysis and Dynamic Modeling of Local Office Operations". Ph.D. dissertation. Department of Industrial Engineering and Operations Research, University of California, Berkeley, 1973.

Graham, A.K. "Parameter Estimation in System Dynamics Modeling". in Randers, J. (ed.) Elements of the System Dynamics Method. Cambridge Mass.: The MIT Press, 1980, pp. 143-161.

Mass, N.J., and Senge, P.M. "Alternative Tests for Selecting Model Variables", in Randers, J. (ed.) Elements of the System Dynamics Method. Cambridge, Mass.: The MIT Press, 1980, pp. 205-225.