J. ENVIRONMENTAL SYSTEMS, Vol. 7(1), 1977-78

# MODELING THE PUBLIC WELFARE SYSTEM: PART 2

**ATTIA I. SWEILLAM, PH. D.**
*Long Island Research Institute*
*State University of New York at Stony Brook*

**HELMAN I. STERN, PH. D.**
*Department of Industrial Engineering and Management*
*Ben Gurion University of the Negev*
*Beersheva, Israel*

**ABSTRACT**

A previous paper of the authors presented a pure markov chain model to forecast New York State public welfare cases. The model is comprised of three major components: transferred cases, closed cases, and opened cases. It was found that the opening and closing portions of the model contributed the majority of the forecasting error. This report compares the previous approach to a modified markov model with opening and closings determined by measures of policy change (version 1) or by linear regression on socio-economic and demographic factors (version 2). In this model, intercase transfer rates are conditional upon openings and closings. A validation procedure selected the modified model (version 1) to forecast total caseloads. The forecasts indicated mean absolute per cent errors for individual welfare categories and total caseloads of: 1.9 and 0.65 (1972), 1.3 and 1.0 (1973), respectively. The errors are well below the five per cent level required for departmental budget preparations. The model may be implemented with internally generated data.

## Introduction

The New York State Department of Social Services prepares a forecast of the total welfare caseload for the purpose of submitting its budget request for the coming fiscal year which begins in April. These requests are normally prepared in September of the prior fiscal year. Although no documentation of the techniques employed were prepared by the department, it was revealed that





doi: 10.2190/XK1E-T0FT-RXNF-NFC5
http://baywood.com

none of the many variables influencing caseload dynamics were taken into account other than a straight forward extrapolation of past caseload historical trends. Of prime importance were the forecasting errors of the *total caseload* which were found to be in the order of ten per cent. Officials stated that a five per cent error on total caseload forecasts would provide a marked improvement and be considered an acceptable level of error for their purposes. As such, it was decided to develop a forecasting model which included flows of cases into, out of, and between welfare caseload categories on a statewide level. This resulted in the markov chain model discussed in Part 1 of this study [1]. The model was comprised of three major components: opened cases, closed cases and intercategory transfers. These three components were combined to provide a projection of caseload dynamics on a quarterly basis. A preliminary forecast was prepared ex post facto for the fiscal year 1972. This forecast revealed a .2 per cent, four quarter mean absolute value error between observed and estimated welfare caseload totals. However, this error must be attributed totally to the transfer and closing portions of the model as actual observed opening data were used in the projection. Closed cases were roughly estimated on the basis of average closing rates and revealed a mean average absolute value error for total closed cases of 2.45 per cent. As total closings are approximately ten per cent of the total caseload this accounts for almost all of the total caseload error of .2 per cent. As total openings are also approximately ten per cent of the total caseload it was decided to devote further effort to the improvement of the opening and closing portions of the model.

Three alternative techniques were evaluated for forecasting caseload openings and closings. The first technique is based on the use of past opening and closing rates as a function of cases under care by each category. When combined with markov intercase transition rates it is termed the pure markov chain model (PMC). This model is the one investigated in part 1, and essentially represents a closed system. The second technique employs regression analysis to predict opened and closed cases as a function of socio-economic, demographic, and administrative policy factors. As implementation of this technique involves an appreciable amount of work in the collection and projection of these factors, it was decided to evaluate a simplified opening-closing model based only upon the anticipated effects of administrative policy decisions. As all administrative policy decisions for the coming fiscal year must be included in the budget proposal, the information required

for the implementation of this model is readily available from internal sources. The last two versions when combined with the intercase transition component requires a modification of the intercase transition rates, and are, therefore, referred to as modified markov chain models (MMC). To distinguish between the two methods of predicting openings and closings in the MMC model, we designate MMC1 and MMC2 as those using anticipatory policy change and linear regression, respectively.

In the next section of the paper, the mathematical basis of the MMC model is developed. This is followed by a description of the caseload opening and closing submodels. These submodels are then calibrated, using 1972 data, and the best predictive technique selected by a validation procedure. The last two sections of the paper provide projections for the years 1972 and 1973, and a summary and conclusion.

## The Modified Markov Chain Model

As the PMC model has been previously described in [1], it will not be repeated here. In that model it was determined that the intercase transition rates, $q_{ij}$, between all five caseload categories were found to be relatively stable. This finding will be also applied to the MMC model. In the MMC model the actual number of openings and closings are determined by either linear regression equations or estimates of anticipatory policy effectiveness. Equivalent closing rates may then be computed. It is then necessary to readjust the intercase transition rates to balance the system. The procedure is as follows:

Let $s_i^t$ be the estimated number of cases closed from category i in the interval of time (t, t + 1), and $p_i^t$ the total category i caseload at time t. Then the closing rates $q_{i6}^t$ may be computed as:

$$q_{i6}^t = s_i^t / p_i^t, \; i = 1, \ldots, 5 \tag{1}$$

Let $q_{ij}^o$ be the original PMC intercase transition rates, and $\bar{q}_{ij}$ the relative intercase transition rates which preserves the original intercase relationships within the system. Then,

$$\bar{q}_{ij} = q_{ij}^o \Big/ \sum_{j=i}^{5} q_{ij}^o, \; i = 1, \ldots, 5 \tag{2}$$

such that,

$$\sum_{j=1}^{5} \bar{q}_{ij} = 1, \; i = 1, \ldots, 5$$

Let $q_{ij}^t$ be the intercase transition rates adjusted for the prevailing closing rates in the MMC model. The MMC transition equations, where $n_j^t$ represents the openings in category j over the interval (t, t + 1), are:

$$p_j^{t+1} = \sum_{i=1}^{5} q_{ij}^t p_i^t + n_j^t, j = 1, \ldots, 5 \tag{3}$$

To insure that the intercase transition rates are balanced, the following identity must hold:

$$\sum_{j=1}^{5} q_{ij}^t = 1 - q_{i6}^t, i = 1, \ldots, 5 \tag{4}$$

Selecting

$$q_{ij}^t = \bar{q}_{ij} (1 - q_{i6}^t) \tag{5}$$

satisfies the required identity (4) and after substitution in (3) yields:

$$p_j^{t+1} = \sum_{i=1}^{5} \bar{q}_{ij} (p_i^t - s_i^t) + n_j^t, j = 1, \ldots, 5 \tag{6}$$

Thus, to operate the model, one need not actually calculate the closing rates $q_{i6}^t$ but may use the projected openings, $s_i^t$ directly in (6). The $\bar{q}_{ij}$ are determined from the PMC stable rates as in (2). An interpretation of (6) is that if an individual leaves the system in period (t, t + 1) he does so from the state he is in at time t. (Thus, predicting situations where an individual transfers within the system and then leaves during the interval (t, t + 1).) The first term in (6) represents the nonclosed cases (after subtracting out a prediction of closings $s_i^t$) that either remain in category j or are transferred from the other four categories within the system to j. To these remaining cases are added a prediction of the number of new case openings $n_j^t$. Openings and closings are predicted by either of two methods in the MMC equation. These are version 1: anticipatory policy change; and version 2: linear regression designated MMC1 and MMC2, respectively.

## Caseload Openings and Closings

In the PMC model, opening and closing rates are determined as an average of past history. In MMC1 the effects of internal policy decisions on opened and closed case ratios are estimated on a judgemental basis by concensus of departmental officials. In MMC2

linear multiple regression functions are constructed for newly opened and closed cases by category as the dependent variable. The independent variables are those which influence or explain the change in caseload behavior. They are classified into the following four major groupings: cases under care, administrative policy, economic influences, and demographic factors. The considerations taken into account in the selection of the independent variables are based on availability of numerical data and reasonable theoretical hypothesis on caseload behavior. These considerations are discussed in turn for each grouping.

Cases under care was selected under the attraction theory, whereby, it is hypothesized that increased case size tends to encourage applications resulting in increased levels of newly opened cases.

It is also hypothesized that administrative policy exerts a strong influence on the public welfare caseload. As this influence is difficult to assess in quantitative terms, for the purpose of this study, administrative policy is measured from available data collected monthly by the Department of Social Services [2] on reasons for opening and closing public assistance cases. These reasons are listed under the categories of: family, health, financial and legal reasons. The legal reason category is most reflective of administrative policy as it includes: changes in state law or agency policy, refusal after acceptance to comply with eligibility requirements, determination of no eligible child in the home, etc. The trend to legal reasons shows a strong and increasing influence of policy on closings. For example, over a ten year span in ADC there was an increase from 31.1 to 74.2 per cent of closings due to legal reasons [3]. To a lesser degree, although still significant, the per cent of cases opened for legal reasons has also increased for each form of assistance. Acceptance rate has also been selected as a potentially influencial measure of public policy in that it is believed to be a measure of how strictly eligibility criteria are being applied.

The economic factors which were assumed to influence the behavior of caseload openings and closings are: unemployment, unemployment insurance claims, business index, average weekly hours, average weekly earnings and working days in the month. Unemployment was hypothesized to increase openings and decrease closings. Unemployment insurance claims, lagged six months, allows us to capture the potential welfare population not included in the unemployment variable. The index of business activity is expected to have the reverse effect as unemployment.

Increases in the remaining three economic factors, reflecting increased income, should reduce the number of people meeting the financial criteria for assistance and have the reverse effect of increased unemployment.

Births and deaths were the only demographic factors selected to represent changes in the population characteristics of the state as data on family composition and migration are not readily available for short term periods. Births and deaths are expected to effect caseload, especially in ADC. An additional child may bring a two parent family into eligibility for aid to families with dependent children. Death of an individual recipient should close his case, but death of a breadwinner or parent might be reason for opening a case due to reduced family income.

## Data Calibration

The data calibration for the three predictive techniques used in this study: PMC, MMC1 and MMC2 are discussed in this section. Each technique is calibrated on quarterly data for each of the five public assistance categories from the period 1969 to 1971. The main differences between the three techniques are based on the method of estimating openings and closings. In the PMC and MMC1 models, caseload openings and closings are predicted through the use of caseload opening and closing rates. In the MMC2 model, openings and closings are predicted directly.

Table 1 of reference [1] provides the mean data used in the PMC model for all but the opening rates. Average mean quarterly opening rates for the PMC model were computed for OAA, AB, AD, HR, ADC as .073, .056, .185, .264, .116, respectively.

Data calibration for the MMC1 model are based on adjusting the opening and closing rates of the last year according to anticipatory

Table 1. Actual 1971 and Adjusted 1972 Opening and Closing Rates for MMC1

| | *Opening rates* | | *Closing rates* | |
|---|---|---|---|---|
| *Category* | *Actual* | *Adjusted* | *Actual* | *Adjusted* |
| OAA | .063 | .050 | .056 | .067 |
| AB | .056 | .050 | .047 | .047 |
| AD | .181 | .142 | .122 | .122 |
| ADC | .103 | .088 | .084 | .084 |
| HR | .312 | .312 | .245 | .282 |

Table 2. MMC1 Transition Rates for 1972 Forecasting

| | *To j* / *from i* | *1* *OAA* | *2* *AB* | *3* *AD* | *4* *HR* | *5* *ADC* | *6* *Close* |
|---|---|---|---|---|---|---|---|
| 1 | OAA | .933 | – | – | – | – | .067 |
| 2 | AB | .003 | .945 | .001 | .003 | .001 | .047 |
| 3 | AD | .007 | – | .856 | .002 | .013 | .122 |
| 4 | HR | .007 | – | .041 | .644 | .026 | .282 |
| 5 | ADC | – | – | .001 | .008 | .907 | .084 |
| 6 | Open | .050 | .050 | .142 | .312 | .088 | – |

policy changes for the forecasted year. The actual 1971 mean opening and closing rates are shown in Table 1 along with the adjusted rates for 1972. Administrative policy decisions planned for 1972 were substantial and included: a work reform policy, requirement of photo-identification cards, clarification and expansion of the definition of essential person, and a twenty per cent increase in social security and railroad retirement benefits. It is estimated that opening rates will be reduced by twenty per cent for the OAA, AB, AD, and ADC categories due to these policy changes. In addition, a fifteen per cent increase in closing rates for OAA and HR was estimated. On the basis of these rates, the original intercase transition rates $q_{ij}^t$ of the PMC model are adjusted according to equation (5) and shown in Table 2. Technically, it is not necessary to actually calculate these rates as calculating $s_i^t$ from the closing rates via (1) and $\bar{q}_{ij}^t$ from $q_{ij}^o$ found in Table 1 of reference [1] will suffice (see equation (6)).

Calibration of the MMC2 model requires the determination of the degree of association between caseload openings (closings) and the selected independent variables discussed previously. Data on the independent and dependent variables were collected monthly and aggregated into quarterly data. Cases opened, closed and under care along with legal reasons and acceptance rate were obtained from the records of the N.Y.S. Department of Social Services [2]. Demographic and economic data were obtained from N.Y.S. Departments of Health [4], Labor [5], and Commerce [6]. The step-wise multiple linear regression routine employed in this analysis was that of SPSS [7]. Regression equations were examined on the basis of the most recent two years of data (weaker relationships were found for data trends reaching farther back in time). The per cent of variation, $R^2$, ranged from 93 to 99 for the categories of ADC, AD, and HR as shown in Table 3. It was found

Table 3. Summary of the Relationship Between the Dependent and the Independent Variables as Revealed by the Regression Analysis

| *Dependent Variables* / *Independent Variables* | *ADC* | | | | *AD* | | | | *HR* | | | | *Frequency* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Open* | | *Close* | | *Open* | | *Close* | | *Open* | | *Close* | | |
| | *Seq.* | *Sign* | *Seq.* | *Sign* | *Seq.* | *Sign* | *Seq.* | *Sign* | *Seq.* | *Sign* | *Seq.* | *Sign* | |
| Cases Under Care | 3 | + | 5 | + | 0 | | 4 | - | 0 | | 0 | | 3 |
| Legal Openings | 1 | + | 0 | | 2 | + | 0 | | 1 | + | | | 3 |
| Legal Closings | 0 | | 1 | + | 0 | | 1 | + | 0 | | 1 | + | 3 |
| Births | 2 | + | 6 | - | 0 | | 0 | | 0 | | 0 | | 2 |
| Deaths | 0 | | 3 | + | 0 | | 0 | | 0 | | 0 | | 1 |
| Unemployment | 0 | | 4 | - | 5 | + | 0 | | 0 | | 0 | | 2 |
| Unemployment Ins. Claims | 5 | + | 0 | | 6 | - | 0 | | 5 | + | 0 | | 3 |
| Working Days in Month | 0 | | 0 | | 0 | | 6 | + | 0 | | 0 | | 1 |
| Business Index | 0 | | 0 | | 4 | + | 0 | | 4 | - | 0 | | 2 |
| Avg. Weekly Earnings | 0 | | 0 | | 0 | | 3 | + | 0 | | 3 | + | 2 |
| Avg. Weekly Hours | 4 | - | 7 | - | 1 | + | 5 | + | 2 | + | 0 | | 5 |
| Acceptance Rate | 0 | | 2 | - | 3 | + | 2 | - | 3 | + | 2 | - | 5 |
| No. of Contributing Variables | 5 | | 7 | | 6 | | 6 | | 5 | | 3 | | 32 |
| Per cent of Variation ($R^2$) | 93 | | 94 | | 96 | | 99 | | 97 | | 99 | | |

that $R^2$ for both OAA and AB varied from .24 to .72. As such, it was decided to use the method of PMC or MMC1 for estimating openings and closings for these cases. These findings are not unusual for AB as the eligibility criteria is blindness. It is surprising for closings of OAA as one would expect a correlation with deaths. A summary of the most contributing independent variables is also shown in Table 3, where Seq. represents the step in which the variable was added to the regression equation. A brief analysis of the results may be found in the Appendix. For further details, including the actual regression equations, see [3].

## Validation Process

The validation process for the three proposed models consists of two steps. Initially, a determination of the best forecasting technique is made on the basis of the best opening and closing tool as the relative intercase rates are assumed to be stable in all three models. Once this selection is made, projections of total cases under care are compared to observed data to assess the validity of the selected model. As the data used in the calibration are derived from the years 1969 to 1971, these years cannot be used to test the model as a forecasting tool. Observed caseload quarterly data for 1972 are used as an ex post facto test of the opened and closed cases. The best method was then used to project 1972 and 1973 total cases under care as if the observed data were not known in advance.

Tables 4 and 5 present an error analysis for each of the proposed opening and closing models for the year 1972. Upon examination of these results, it is found that MMC1 in general provides the best forecast of openings and closings. This model exhibited an average absolute value error of 7 per cent in forecasting total opened cases, while the remaining models gave almost an equal error of 18 per cent. On the single category level MMC1 performed better than the other two models in forecasting openings except for ADC where MMC2 was 2 per cent better. In forecasting closings, MMC1 exhibited the least absolute value error of 2.75 per cent although PMC and MMC2 had 5.25 and 4.0 per cent, respectively. On the single category level, MMC2 performed better than MMC1 for AD only. These results indicate that MMC1 is superior on the basis of predicting total caseload openings and closings. In addition, the projection errors are underestimated as known values for the independent variables were used. In any real projection, these values would not be known in advance, and

Table 4. Comparison Between Quarterly Forecasting Errors of Opened and Closed Cases by Category Using Pure Markov Chain, Modified Markov Chain, Version I and Version II, 1972

| | *Opened Cases* | | | *Closed Cases* | | |
|---|---|---|---|---|---|---|
| *Category* | *Pure Markov Chain* | *Modified Markov Chain Version I* | *Version II* | *Pure Markov Chain* | *Modified Markov Chain Version I* | *Version II* |
| AD | 6 | −19 | 8 | − 5 | −12 | 1 |
| | 29 | −04 | 29 | + 4 | − 7 | 2 |
| | 51 | + 8 | 38 | 21 | + 4 | − 5 |
| | 62 | +13 | 50 | 28 | + 7 | +10 |
| ADC | 21 | − 8 | 3 | −27 | + 3 | − 4 |
| | 42 | + 6 | 5 | −24 | + 4 | 3 |
| | 35 | − 1 | 4 | −17 | + 6 | 4 |
| | 54 | +11 | 7 | −13 | − 6 | 9 |
| HR | −25 | −11 | 13 | +10 | − 6 | − 4 |
| | −21 | − 8 | 18 | 12 | − 5 | 3 |
| | −13 | − 2 | 20 | 17 | + 1 | 9 |
| | + 1 | +11 | 39 | 6 | + 4 | 11 |
| Total | − 1 | −12 | 08 | − 8 | − 4 | − 3 |
| | 13 | − 2 | 16 | − 4 | − 2 | 3 |
| | 20 | + 1 | 21 | + 5 | + 4 | 4 |
| | 35 | +11 | 25 | + 4 | + 1 | 6 |

involve additional expense to project them. Since it was found that internal policy effects the major change in total openings and closings (as indicated by an analysis of the regression results in the Appendix), and operationally speaking, MMC1 is easier and cheaper to implement; it was selected over MMC2 in lieu of MMC2's seemingly superiority in predicting ADC openings and AD closings.

Table 5. Comparison Between the Mean Absolute Value Forecasting Errors of Pure Markov, Modified Markov Chain Version I and Version II, 1972

| | *Opened Cases* | | | *Closed Cases* | | |
|---|---|---|---|---|---|---|
| *Category* | *Pure Markov Chain* | *Modified Markov Chain Version I* | *Version II* | *Pure Markov Chain* | *Modified Markov Chain Version I* | *Version II* |
| AD | 37 | 11 | 31 | 15 | 7.5 | 4.5 |
| ADC | 38 | 7 | 5 | 20 | 4.75 | 5.0 |
| HR | 15 | 8 | 23 | 11 | 4.0 | 6.75 |
| Total | 17 | 7 | 18 | 5.25 | 2.75 | 4.0 |

Table 6. Quarterly 1972 Projections and Error Analysis of Caseload by Category Using Modified Markov Chain Model (Version I)

| *Category* | *Quarter* | *Estimated* | *Observed*[a] | *Residual* | *Per cent error* |
|---|---|---|---|---|---|
| OAA | 1 | 111592 | 111584 | + 8 | 0.0 |
| | 2 | 111361 | 112392 | -1031 | -0.9 |
| | 3 | 111153 | 111917 | - 764 | -0.7 |
| | 4 | 110970 | 108617 | +2353 | +2.2 |
| AB | 1 | 3990 | 4011 | - 21 | -0.5 |
| | 2 | 3971 | 4085 | - 114 | -2.8 |
| | 3 | 3952 | 4164 | - 212 | -5.4 |
| | 4 | 3933 | 4213 | - 280 | -7.1 |
| AD | 1 | 135456 | 137055 | -1599 | -1.2 |
| | 2 | 139668 | 143401 | -3733 | -2.7 |
| | 3 | 143818 | 147333 | -3515 | -2.4 |
| | 4 | 147909 | 149993 | -2084 | -1.2 |
| ADC | 1 | 345805 | 347376 | -1571 | -0.4 |
| | 2 | 348470 | 349551 | -1081 | -0.3 |
| | 3 | 351140 | 352620 | -1480 | -0.4 |
| | 4 | 353818 | 352577 | +1241 | +0.4 |
| HR | 1 | 100836 | 101522 | - 686 | -0.6 |
| | 2 | 99440 | 102091 | -2651 | -2.7 |
| | 3 | 98135 | 101366 | -3231 | -3.3 |
| | 4 | 96918 | 99308 | -2390 | -2.5 |
| Total | 1 | 697679 | 701548 | -3869 | -0.5 |
| | 2 | 702910 | 711520 | -8610 | -1.2 |
| | 3 | 708198 | 717400 | -9202 | -1.3 |
| | 4 | 713548 | 714708 | -1160 | -0.1 |

[a] These figures were adjusted from those of Table 4 in reference [1] after corrections of faulty reporting.

Using MMC1, a forecast of total caseload by category for the year 1972 was attempted. This forecast is based on the adjusted opening and closing rates shown in Table 1, and the use of equations (1), (2) and (6). Table 6 presents the resultant errors for each category over the four quarters of 1972. The caseload errors by category fluctuate between -7.0 and +2.0 per cent with an average absolute error of 1.9 per cent. It is also of interest to observe that the total caseload projection (the sum of all categories) exhibited an error of at most 1.0 per cent below the observed values. The mean absolute value of the total caseload projection for 1972 is .65 per cent.

Table 7. Quarterly 1973 Projections and Error Analysis of Caseload by Category Using Modified Markov Chain Model (Version I)

| *Category* | *Quarter* | *Estimated* | *Observed* | *Residual* | *Per cent error* |
|---|---|---|---|---|---|
| OAA | 1 | 108527 | 108070 | + 457 | +0.4 |
| | 2 | 108406 | 106680 | +1726 | +1.6 |
| | 3 | 108257 | 106381 | +1876 | +1.8 |
| | 4 | 108081 | 106326 | +1755 | +1.6 |
| AB | 1 | 4192 | 4202 | - 10 | -0.2 |
| | 2 | 4171 | 4186 | - 15 | -0.3 |
| | 3 | 4151 | 4184 | - 33 | -0.8 |
| | 4 | 4131 | 4165 | - 34 | -0.8 |
| AD | 1 | 151021 | 152014 | - 993 | -0.6 |
| | 2 | 151882 | 149441 | +2441 | +1.6 |
| | 3 | 152365 | 151111 | +1254 | +0.8 |
| | 4 | 152729 | 155520 | -2791 | -1.8 |
| HR | 1 | 93792 | 98119 | -4327 | -4.4 |
| | 2 | 88742 | 88295 | + 447 | +0.5 |
| | 3 | 84119 | 81774 | +2345 | +2.8 |
| | 4 | 77109 | 75059 | +2050 | +2.7 |
| ADC | 1 | 350675 | 350832 | - 157 | 0.0 |
| | 2 | 348681 | 345776 | +2905 | +0.8 |
| | 3 | 346608 | 342472 | +4136 | +1.2 |
| | 4 | 344465 | 339310 | +5155 | +1.5 |
| Total | 1 | 708207 | 713237 | -5030 | -0.7 |
| | 2 | 701882 | 694378 | +7504 | +1.0 |
| | 3 | 695500 | 685922 | +9578 | +1.4 |
| | 4 | 686515 | 680380 | +6135 | +0.9 |

The MMC1 model was again used to provide a forecast of total caseloads for 1973. These forecasts were made in January, 1974 when actual 1973 caseload statistics were still in the process of being compiled. Actual 1973 observed data became available in April, 1974 after which the projections were compared and resulted in the forecast errors documented in Table 7. These projections are based on 1972 closing and opening rates modified according to expected administrative policy changes for the year 1973. The adjusted rates are shown in Table 8 and reflect the anticipated impact of administrative actions. This impact was judged to decrease opening rates of AD, ADC and HR categories by ten per cent, and increase closing rates by five per cent for all

Table 8. Actual 1972 and Adjusted 1973 Openings and Closings Rates

| | *Opening* | | *Closing* | |
|---|---|---|---|---|
| *Category* | *Actual* | *Adjusted* | *Actual* | *Adjusted* |
| OAA | . 05 | . 05 | . 06 | .067 |
| AB | . 05 | . 05 | .047 | .047 |
| AD | .142 | .128 | .122 | .128 |
| HR | .312 | .281 | .282 | .296 |
| ADC | .088 | .079 | .084 | .088 |

categories. The analysis leading to these judgements is based on a restrictive administrative policy plan to control the growth of the 1973 caseload. This plan included the actions of: introducing on February 1, 1973 a detailed welfare application form to replace the old "declaration" method of eligibility, the introduction on July 1, 1973 of periodic, face-to-face recertification interviews to replace the old "mail-order" procedure which contributed heavily to the error rates state auditors were finding, the implementation of a new intake screening and track down of missing fathers which resulted from the fact that more than four in every ten ineligibility errors involved the absence of a father actually residing at home, the deployment, beginning in December, 1973, of up to 200 members of the state audit staff to take up residence in local welfare agencies. The resulting 1973 forecasting errors by category ranged from -4.4 to +2.8 per cent with an average absolute value error of 1.3 per cent. It is also of interest to observe that the total caseload forecasts for 1973 varied between -.7 and +1.4 per cent with an average absolute error of 1.0 per cent. Figure 1 provides a plot of total caseloads over time by quarter for fiscal years 1972 and 1973 contrasting the actual caseload history with the forecast. Basically, we underestimated the trend by 1.0 per cent for 1972 and overestimated it by 1.4 per cent for 1973.

## Summary and Conclusions

Modeling the public welfare system is approached in parts 1 and 2 of this study by developing a mathematical model to explain and forecast the potential New York State Public Assistance caseload changes. Public assistance is disaggregated into five major categories: Old Age Assistance, Aid to the Blind, Aid to the Disabled, Aid to

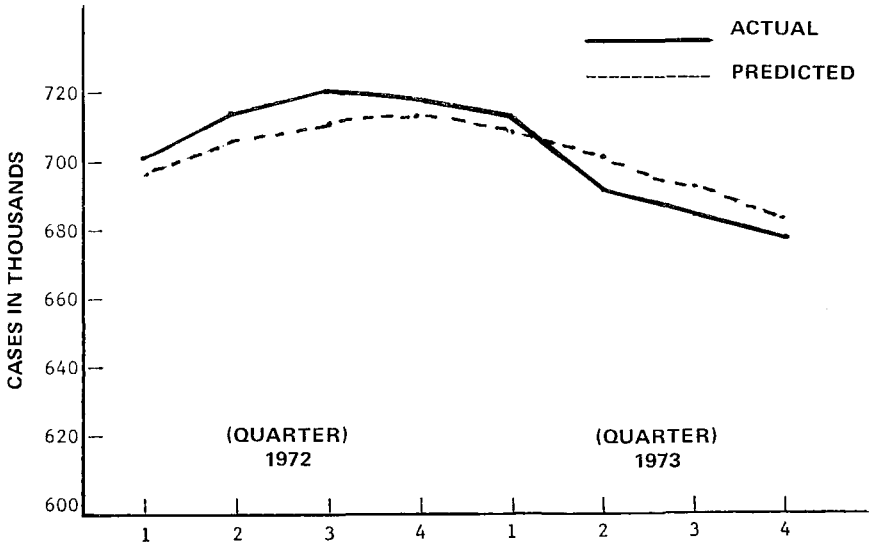


Figure 1. Prediction of quarterly caseload by Modified Markov Chain Model (Version I) compared with the actual caseload for 1972 and 1973.

the Dependent Children and Home Relief. The caseload is composed of three major components: case openings, case closings, and case transfers. These components represent the dynamic behavior of public assistance case movements over a particular period of time.

Analysis and forecasting of caseload components are approached through two models: a PMC model with stationary transition rates and opening and closing rates as a function of the number of cases under care. The second model is a MMC model with intercase transitions conditional upon closings, and opening and closing rates determined by anticipatory policy changes (version I) or linear regression analysis (version II). MMC1 is based on the assumption that public assistance is a "closed system" where opened and closed cases are a function of administrative policy actions only. MMC2 is based on the assumption that public assistance is a "partially open system" where opened and closed cases are a function of socio-economic factors, demographic changes as well as administrative policy effects.

The PMC model and MMC1 have been calibrated on three years of quarterly data from 1969 to 1971. MMC2 was examined on the basis of four years and two years of data. It was found that stronger relationships were obtained (i.e., better fits) by using the most recent two years of data in lieu of four years of data.

The exogenous variables which are assumed to influence the public assistance newly opened and closed cases are: number of cases under care at the end of the previous period, opening rates of each category, number of cases opened or closed for legal reasons, number of working days in the month, number of unemployed persons, unemployment insurance claims, total activity index, average weekly hours (mfg.), average weekly earnings (mfg.), number of births and number of deaths.

The regression analysis technique was employed in the MMC2 model to explain and predict the number of newly opened and closed cases by reference to the values of one or more of the associated variables. The prediction was based on the degree of association between newly opened and closed cases, and the influencing factors. The regression technique was used in a step-wise process to provide a means of choosing the fewest number of the influencing variables which will provide the best prediction possible of the newly opened and closed cases.

The data analysis and data calibration reveal the following findings:

a. Both the OAA and AB categories are independent of inter-case movement. Therefore, we may assume that the major increase in their caseloads is attributed mostly to newly opened cases rather than to transfers from other categories.
b. There is significant interchange between AD and HR while ADC receives most of its transfer cases from the HR category and does not contribute to any other category in the system.
c. Public assistance categories may be grouped into the three independent subgroups of (OAA), (AB), and (AD, HR, ADC). According to this structure, OAA and AB can be forecasted independently without any effect on AD, HR and ADC categories.
d. Administrative actions, as determined by the number of closed and opened cases for legal reasons, were found the most common influencing factor in determining opening and closing cases.
e. The analysis showed that the economic and demographic factors did not reflect much influence on the caseload changes in spite of the strong associations obtained between these variables as shown by $R^2$ in the regression equations.

The models used in the study have been tested ex post facto on actual opened and closed cases for 1972 and 1973 data. It was found that the MMC1 model gives better results than the other two models in forecasting total openings and closings. As such,

MMC1 is selected to forecast the public assistance caseload by category level. The findings of the analysis and the choice of the MMC1 model in forecasting the potential caseload affirmed the hypothesis that public assistance is primarily a "closed system." Caseload changes seem to be more a function of welfare administrative policy changes rather than a function of a socio-economic and demographic factors.

The validation procedure of the MMC1 model for forecasting the caseload of 1972 indicated mean absolute errors of 1.9 and .65 per cent for individual categories and total welfare caseloads, respectively. The testing procedure of 1973 projections indicated mean absolute errors of 1.3 and 1.0 per cent for individual categories and total welfare caseloads, respectively. The accuracy of forecasting the caseload by using MMC1 depends on the accuracy of estimating the impact of the anticipatory policy changes on case openings and closings. This might seem to be a difficult factor to estimate, but it is the choice of the policy maker to determine a range of rates for openings and closings. Then the function of the model is to evaluate the impact of each decision on the future total caseload changes by category. Based on this assumption, the model can be used as an administrative tool in evaluating a set of options before making a course of action. With the recursive structure of the MMC1 model, it is much more convenient to be able to use the outputs of the current time period, with the anticipatory policy changes, to generate the size of potential caseload for the subsequent intervals.

When implementing MMC1, it is suggested to use the most recent time series data for calibration of the intercase transitions. When internal factors (policy decisions) change they seemingly overshadow the effects of the outside environments and MMC1 should be used. Since MMC2 showed very good associations between opening and closings and socio-economic factors, it may be worth pursuing when it is believed that the internal policy environment has not changed drastically as it should pick up some of the effects of the outside environment on the caseload. On the other hand, although the regression equations provided good correlations between internal and outside factors with caseload openings and closings it cannot be used in the causal sense when the internal environment changes drastically as the caseload is particularly sensitive to new forms of internal policy decisions.

The model is able to answer and justify a much broader range of questions than can be answered from the existing stationary status. Pre-knowledge of the answers to these questions will help in welfare

administrations and program planning. For example, does the distribution of opened and closed cases rates vary from period to period, and how are these variations related to changes in the categorical eligibility provisions, welfare policy changes and administrative procedures? What is the interrelationship among the five public assistance categories, and also between public assistance categories and other social services programs within and outside the welfare system? The model is designed to be flexible and to be employed in any welfare jurisdiction at any period in time. It is also flexible enough to include or delete any variable without major changes in the total structure of the model. The model will help create and organize a permanent data base which will identify the data elements required to be collected, retrieved and processed continuously and regularly. This will help increase the accuracy level of management reporting system, eliminate redundancy in data acquisition, and standardize the definition of data elements. Also, the output of the model will interface with other departmental activities such as: program planning, coordination, evaluation, administration, and budgeting. The modeling approach seems to be presently and potentially feasible to be applied to different functions of public welfare systems such as: public assistance expenditure, medical assistance and social services delivery systems. This approach is recommended as a management tool for better program planning and administration.

## ACKNOWLEDGEMENT

Grateful acknowledgement is made to Professor Paul Zuber, Director of the Center for Urban and Environmental Studies at Rensselaer Polytechnic Institute, Troy, New York, for his encouraging interest and helpful comments, and the staff of the New York State Department of Social Services for its most generous support and contribution of valuable time and information.

## REFERENCES

1. H. I. Stern and A. I. Sweillam, Modeling the Public Welfare System: Part 1, *Journal of Environmental Systems*, *3*:4, pp. 277-285, Winter, 1973.
2. New York State Department of Social Services, Albany, New York, Monthly Statistical Report on Reasons for Opening (Closing) Public Assistance Cases, Forms DSS-848, DSS-881, 1962-1972.
3. A. I. Sweillam, Modeling the Public Welfare System: A Public Welfare

Caseload Forecasting Model, Ph.D. Thesis, Urban-Environmental Studies, Rensselaer Polytechnic Institute, Troy, New York, August, 1974.
4. New York State Department of Health, Office of Biostatistics, Albany, New York.
5. New York State Department of Labor, Division of Research and Statistics, New York State Population, Income and Employment, 1950-1972.
6. New York State Department of Commerce, Quarterly Summary of Business Statistics, 1969-1972.
7. H. H. Nie, D. H. Bent and H. Hyll, *Statistical Package for the Social Sciences*, McGraw-Hill Book Co., New York, 1970.

## APPENDIX

### Summary of the Linear Regression Results

a. Public policy changes, as determined by the number of closed and opened cases for legal reasons are the most common influencing factor in determining opened and closed cases in most public assistance categories.

b. The opening rates are the second strongest factor influencing the caseload. It is associated positively with openings except for ADC, and negatively with closed cases for all categories in the system. This result was expected and the study verifies it. The opening rate is another measure of policy change.

c. Average weekly hours is the next variable in sequence associated with all opened and closed cases except for HR closed cases where there is no association. The association was unexpectedly positive for both opened and closed AD cases while negative for ADC cases. It showed positive relationship with HR openings while it has no effect on HR closings.

d. Cases under care is associated positively with ADC and negatively with AD openings. Unexpectedly, no association has been found with the other variables. Also the positive association with ADC closings is unexplained.

e. Demographic factors are the least influencing factors associated with opened and closed cases. Births are only associated positively with opened cases and negatively with closed cases for ADC, as was expected. Although mortality was hypothesized to result in increased openings in ADC and increased closings in all programs, this was not verified by the regression results.

f. Economic factors are less associated as influencing factors on the caseload component. Unemployment insurance claims associated positively with opened cases of ADC and HR and negatively with opened cases of AD. This relationship was expected. Unemployment is associated negatively with ADC closings and positively with AD opened cases. No association was found with the other independent variables, which was unexpected. Business Index associated positively with AD openings, which was unexpected, negatively with HR opening, which was expected and had no association with the other variables. Average weekly earnings associated positively with closed cases of AD and HR which was expected and had no association with the other variables.

g. In summary, the regression analysis reveals different percentage variations in caseloads explained by different independent variables as shown in the last row of Table 3. Ninety-three per cent of the variation of ADC openings was explained by five variables: cases under care, legal openings, unemployment, unemployment insurance claims, average weekly hours, and opening rate. Ninety-nine per cent of the variation of AD was explained by six variables: cases under care, legal closings, working days in month, average weekly earnings, average weekly hours, and opening rates. Ninety-seven per cent of the variation of HR openings was explained by five variables: legal openings, unemployment insurance claims, business index, average weekly hours, and opening rates. Ninety-nine per cent of the variation of HR closings was explained by three variables: legal closings, average weekly earnings, and opening rate.

Direct reprint requests to:

Dr. Attia I. Sweillam
State of New York
Department of Mental Hygiene
Long Island Research Institute
P.O. Box Q
Central Islip, New York 11722