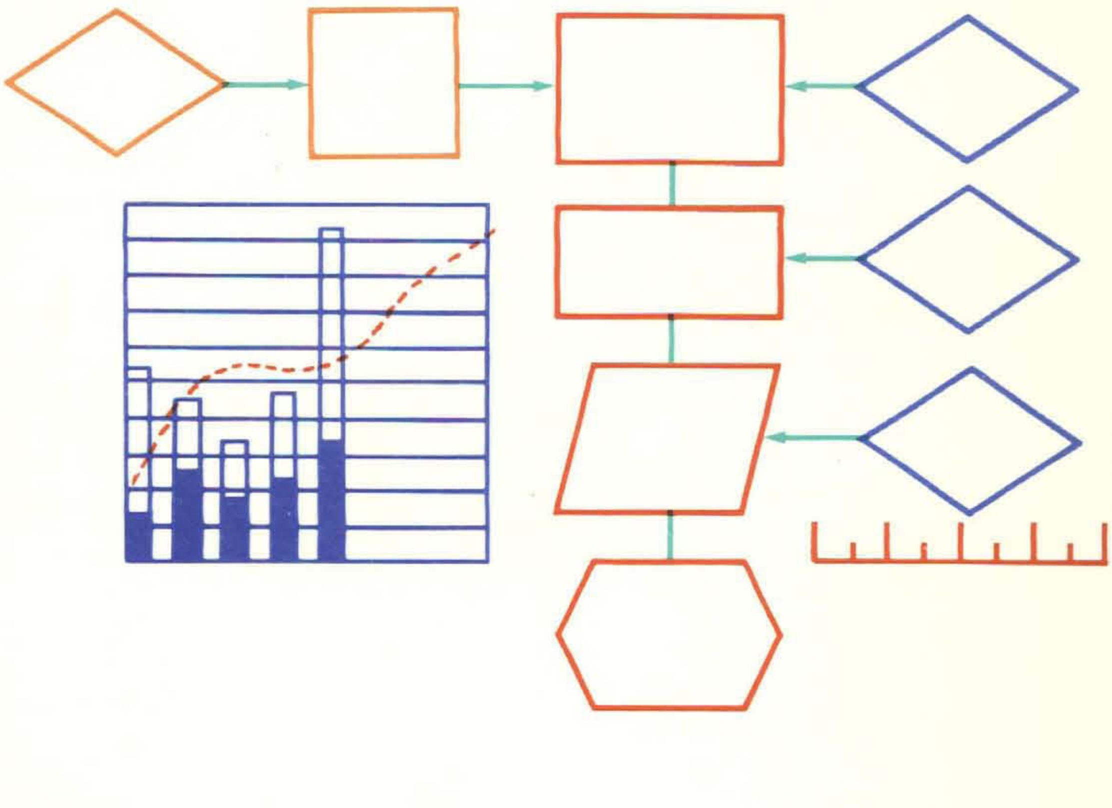
Regional Economic Expansion

Expansion Économique Régionale

Working Paper No. 3

MIGRATION IN CANDIDE-R

July 1975



ECONOMIC DEVELOPMENT ANALYSIS DIVISION

DIVISION DES ÉTUDES DE DÉVELOPPEMENT ÉCONOMIQUE

Quantitative Analysis Unit

Groupe D'Analyse Quantitative

Working Paper No. 3

MIGRATION IN CANDIDE-R

July 1975

This working document represents a partial regionalization of the CANDIDE 1.1 model. The acronym CANDIDE refers to the Canadian Disaggregated Interdepartmental Econometric model.

The CANDIDE-R version of the model outlined in this document is designed to help build an appreciation of the regional diversity of Canada. The authors draw attention to the tentative nature of the econometric work reported upon. So as to avoid attributing official status to the views expressed, prior consultation respecting quotation would be appreciated.

Economic Development Analysis Division,
Canada. Dept. of Regional Economic Expansion,
161 Laurier Avenue West,
Ottawa, Ontario,
K1A OM4.

Migration in CANDIDE-R

1.- Introduction

Until recently, it was more the exception than the rule to include migration flows in traditional econometric models. Model builders have often underestimated the importance of the effects of population movements on the economy, and vice-versa. However, migration flows are an important component of population growth, which if included in the model require a system of demographic projections within the model. The advantage of including both a demographic system and a migration mechanism in a econometric model is reflected in three considerations. First, the importance of migration varies from one country to another. In Canada, however, migration accounts for up to half of the annual increase in population and more in some years. Secondly, it is easy to visualize the importance of demographic characteristics and of migration for the potential growth of the economy, since such growth is dependenton the population structure. Finally, the importance of analyzing population movements increases with the degree of regional disaggregation of the model. As is shown in this paper the impact of regional migration flows on regional economics is not one which is easily ignored. International migration flows in CANDIDE-R are minimal as compared to the 5 inter-regional flows.[1]

These three considerations form the justification for the disaggregated, regional study of demography and migration in CANDIDE-R.

1. For example, see Table I. The five regions in CANDIDE-R are: the Atlantic region, Quebec, Ontario, the Prairies and British Columbia.

2.- Some Characteristics of Regional Migration Flow Data

Before examining the specification and estimation of the migration model, it is important to analyse briefly the main characteristics of population flows at the regional level. Since regional statistics for most of the migration flows are not readily available in government publications, some of the time series required for CANDIDE-R were constructed from indirect information. This section looks first at the importance of various types of flow measures of migration and then discusses the technique used to account for regional flows and the data problems encountered.

2.1 Regional Migration Rates

Table 1 depicts the inter-regional and international migration flows per region over 1961-1971. It is important to note both the volatility of the flows, and the relative magnitudes of inter-regional and international flows in the different regions. Table 2 shows regional migration rates 1961-1971, and facilitates comparison of migration rates between regions.

Table 1: Principal Regional Migration Flows (in 000 of persons)

	Inter-regional			International			
	Inflows	Outflows	Net	Inflows	Outflows	Net	Net Total
Atlantic							
1961-62	31.9	36.4	- 4.5	2.1	9.3	-7.1	-11.6
1962-63	31.9	44.2	-12.3	2.3	10.1	-7.7	-20.0
1963-64	34.5	49.8	-15.3	2.4	10.6	-8.2	-23.5
1964-65	34.8	53.3	-18.5	2.7	8.7	-6.0	-24.6
1965-66	38.8	56.8	-18.0	3.8	9.9	-6.1	-24.1
1966-67	42.5	57.1	-14.6	4.6	4.7	-0.1	-14.7
1967-68	44.2	48.2	- 4.0	4.3	5.3	-1.0	- 5.0
1968-69	41.0	46.7	- 5.7	4.4	3.4	1.0	- 4.7
1969-70	38.9	58.2	-19.3	4.2	2.2	1.9	-17.4
1970-71	43.8	50.1	- 6.3	4.1	1.1	3.0	- 3.3
Quebec							
1961-62	50.1	41.2	9.0	17.1	13.0	4.2	13.1
1962-63	49.9	43.8	6.1	20.5	14.3	6.2	12.3
1963-64	47.3	47.5	- 0.2	24.2	14.9	9.3	9.1
1964-65	47.7	48.3	- 0.6	26.8	15.5	11.3	10.7
1965-66	50.8	54.1	- 3.3	34.8	11.9	22.4	19.1
1966-67	50.3	59.4	- 9.1	44.5	20.8	23.7	14.6
1967-68	45.8	56.9	-11.1	39.7	23.9	15.8	4.7
1968-69	45.1	58.9	-13.8	32.5	18.5	14.0	0.2
1969-70	41.8	74.3	-32.5	26.8	15.8	11.0	-21.5
1970-71	40.2	75.8	-35.6	22.4	23.0	-0.6	-36.2
Ontario							
1961-62	85.3	90.6	- 5.3	34.6	18.1	16.5	11.2
1962-63	93.7	89.3	4.4	41.3	19.4	21.9	26.3
1963-64	102.5	92.2	10.3	54.7	17.1	37.6	47.9
1964-65	109.1	93.8	15.2	67.0	20.5	46.5	61.7
1965-66	117.7	102.0	15.6	90.3	15.9	74.5	90.1
1966-67	122.7	109.6	13.1	115.8	38.7	77.1	90.2
1967-68	109.5	107.4	2.1	105.6	43.6	62.0	64.1
1968-69	107.4	101.2	6.1	89.9	44.8	45.1	51.2
1969-70	143.8	96.4	47.4	87.7	44.7	43.1	90.5
1970-71	139.9	96.7	43.3	73.6	44.7	28.9	72.2

Table 1 (cont'd): Principal Regional Migration Flows
(in 000 of persons)

	Inter-regional			International			
	Inflows	Outflows	Net	Inflows	Outflows	Net	Net Total
Prairies							
1961-62	61.5	64.5	- 3.1	8.3	11.1	-2.8	- 5.9
1962-63	61.5	69.5	- 8.0	8.6	12.0	-3.4	-11.4
1963-64	60.9	72.8	-11.9	9.5	11.5	-2.0	-14.0
1964-65	62.1	80.0	-17.9	11.6	12.1	-0.5	-18.4
1965-66	64.0	99.0	-35.0	16.0	10.0	6.0	-29.0
1966-67	73.6	97.9	-24.3	22.5	11.4	11.0	-13.3
1967-68	75.2	84.4	- 9.2	26.5	12.8	13.7	4.5
1968-69	68.6	80.0	-11.4	23.3	11.5	11.8	0.4
1069-70	69.2	98.1	-28.9	20.3	9.2	11.1	-17.8
1970-71	66.6	93.1	-26.4	17.1	9.1	7.9	-18.5
B.. Columbia							
1961-62	47.0	43.3	3.6	7.2	3.3	3.9	7.5
1962-63	53.1	42.8	10.3	8.1	2.8	5.2	15.5
1963-64	60.3	42.4	17.8	10.4	3.5	6.9	24.7
1964-65	68.2	44.8	23.4	14.2	4.9	9.3	32.7
1965-66	87.5	46.4	41.2	21.4	2.2	19.3	60.4
1966-67	91.3	55.8	35.5	26.1	7.4	18.7	54.2
1967-68	80.6	58.9	21.6	24.5	4.5	20.1	41.7
1968-69	76.0	51.6	24.4	22.0	6.8	15.2	39.6
1969-70	90.2	57.6	32.6	22.6	5.9	16.7	49.3
1970-71	82.5	59.4	23.0	20.8	6.3	14.5	37.5
Canada							
1961-62			0.0	69.4	53.5	15.9	15.9
1962-63			0.0	80.9	58.4	22.5	22.5
1963-64			0.0	101.3	57.4	43.9	49.9
1964-65			0.0	122.4	62.8	59.6	59.6
1965-66			0.0	166.1	48.8	117.3	117.3
1966-67			0.0	213.7	82.9	130.8	130.8
1967-68			0.0	200.8	90.1	110.7	110.7
1968-69			0.0	172.2	85.9	86.3	86.3
1969-70			0.0	162.0	77.6	84.3	84.3
1970-71			0.0	138.2	82.4	55.7	55.7

SOURCE: Calculations based on data from Statistics Canada.

Table 2: Regional Migration Rates, 1961-1971 (1)
(Migrants per 1,000 persons)

	Inter-regional			International			
	Inflows	Outflows	Net	Inflows	Outflows	Net	Net Total
Atlantic	201.4	263.9	-62.5	18.3	34.4	-16.1	-78.6
Quebec	89.6	107.1	-17.5	55.3	32.8	22.5	5.0
Ontario	181.4	157.0	24.4	121.9	49.3	72.6	97.0
Prairies	208.6	264.0	-55.4	51.4	34.8	16.6	-39.4
British Columbia	452.2	308.7	143.5	108.8	29.2	29.6	173.1
Canada				78.2	38.3	39.9	39.9

(1) Rates are obtained by dividing the total number of migrants from 1961 to 1971 by the 1961 population.

SOURCE: Based on calculations from Statistics Canada.

2.1.1 Gross Flows and Net Flows

Inter-regional flows are substantially greater in volume than international flows. Annual internal migration[2] generally averages between 325,000 and 375,000 (for inflows and outflows) while immigration and emigration flows vary between 50,000 and 200,000 (see Table 1). Net flows are prone to much greater variation than gross flows, since the former includes variations in both inflows and outflows.

This relationship between net and gross flows, points out an important difference between internal and external flows. Net immigration to a region is generally the result of comparatively high

2. In order to maintain consistent terminology in this paper unless otherwise indicated, immigrants and emigrants refer to movement across international borders while migrants refer to people moving across regional boundaries. Net migration refers to the sum total of immigrants, emigrants, and regional inflows and outflows.

gross immigration, and a much lower, more constant gross emigration. In other words the variance of net immigration is dominated by the variance of its main component, gross immigration. However, net internal flows are the result of both very high gross inflows and outflows. The available data shows a relative stability in gross flows which is not the case for net flows (see Table 1).

2.1.2 The Importance of Back-Migration

Surprisingly in the case of internal flows the regions which generally showed a negative migration balance, the Atlantic and Prairie regions, usually also experienced relatively high gross inflows. The gross inflow rate (see Table 2) is in fact higher in these two regions than in Ontario. On the other hand, British Columbia, which has the highest gross inflow rate also has the highest gross outflow rate of all regions.

Comparative economic conditions in the source and receiving regions are not likely to explain these three strong gross flows. Such high inflows, for the Atlantic and Prairies regions can, however, be related to the importance of "back-migration", a factor which is often underestimated. It is highly possible that a great number of people are dissatisfied with their newly-adopted region of domicile, either for cultural, social or economic reasons, and that, after a certain period of time, they return to their region of origin. It is also possible that some number of the migrants included in the regional flows view their move to another region as strictly temporary and in the long run prefer to return to their region of origin.

Finally, the comparatively low internal migration rates in Quebec seem to confirm the generally accepted hypothesis that cultural and linguistic factors substantially reduce the population's mobility.[3]

2.1.3 Destination of Internal and External Migration

Table 2 indicates some of the differences in preference of the migrants and immigrants. Although the first Canadian destination of immigrants is specified they can and do subsequently change regions, thus becoming part of internal migration flows. The inter-regional flow data includes this movement of original immigrants, with the implicit hypothesis that the behaviour of new immigrants with respect to regional movement is the same as that of all other Canadians involved in inter-regional migration flows.

3. It is not that Quebecers are reluctant to leave their region, but rather that other Canadians are reluctant to go and settle in Quebec. See Tables 1 and 2 for a comparison of inter-regional inflows and outflows for Quebec.

It appears, over the sample period at least, that immigrants have a marked preference for urban areas, especially Toronto. This may explain the relatively high rate of immigration to Toronto and Montreal, while the internal migration flow for Quebec is, for the most part, negative.

Tables 3 and 4 show internal migration and internal migration rates for gross flows, per source and destination, 1961-1971.

Table 3: Internal Migration, per source and destination 1961-1971 (in 000 of persons)

Destination \ Source	Atlantic	Quebec	Ontario	Prairies	British-Columbia	Yukon & N.W.T.
Atlantic	0.0	80.0	230.2	40.2	30.5	1.5
Quebec	86.5	0.0	303.6	48.1	29.0	1.7
Ontario	318.6	382.1	0.0	282.7	140.4	8.2
Prairies	48.9	49.6	258.5	0.0	289.2	17.0
British-Columbia	44.9	46.5	181.0	450.4	0.0	13.7
Yukon & N.W.T.	1.8	2.3	6.1	17.9	14.0	0.0

SOURCE: Computed from data from Statistics Canada.

Table 4: Internal migration rates, per source and destination 1961-1971 (migrants per 000 persons) (1)

Destination \ Source	Atlantic	Quebec	Ontario	Prairies	British-Columbia	Yukon & N.W.T.
Atlantic	0.0	42.2	121.3	21.3	16.1	0.8
Quebec	16.4	0.0	57.7	9.1	5.5	0.3
Ontario	51.1	61.3	0.0	45.3	22.5	1.3
Prairies	15.4	15.6	81.3	0.0	91.0	5.3
British-Columbia	27.6	28.5	111.1	276.4	0.0	8.4
Yukon & N.W.T.	47.9	61.1	162.2	476.1	372.3	0.0

(1) These rates are obtained by dividing the total number of migrants from 1961 to 1971 by the 1961 population of the receiving region.

SOURCE: Computed from data from Statistics Canada.

Table 4 points out that the lure of Quebec for internal migrants is quite low, while the rate of migration from Quebec to Ontario, in particular, is considerable.

The largest internal flow (with the exception of the Yukon and NorthWest Territories) both in terms of volume and migration rate, is the migration of residents of the Prairie region to British Columbia. Moreover, considerable inter-provincial population movements destined for Alberta do not appear in the figures shown above, due to aggregation. On the whole it seems that residents of the Prairie region are more mobile than those in the rest of the country.

2.2 Regional Migration Flow Data

2.2.1 Inflow Determination of the Migration Model

The migration flows used in the population algorithm are the total net annual migration in each of the five regions. These net flows are themselves a result of four gross flows included in the identity:

(1) TNMr = MINrT - MOrRT + MINrXT - MOrXT

Where: TNMr = Total net migration in regior r

MINrT = Total population inflow, to region r, from other regions in Canada.

MOrRT = Total population outflow from region r to other regions in Canada

MINrXT = Gross immigration to region r, from other countries

MOrXT = Gross emigration from region r, to other countries

In a migration system with five regions, there are then five gross international immigration flows per destination and an equal number of gross emigration flows. However, each region also has regional migration inflows and outflows with the other four regions. Therefore, there are 20 gross inflows and an equal number of gross outflows per source and destination.

$$(2)\quad MINrRT = \sum_{s=1}^{4} MINrRs$$

$$MOrRT = \sum_{s=1}^{4} MOrRs \qquad r \neq s$$

(where s refers to the source region, and like r is replaced in the mnemonic table by E, Q, O, W or C for the five regions, or by T for Canada. However, inter-regional flows as they are defined here are reversible, i.e. the total inflows to one region from a second correspond, by definition to the total outflows from the second region to the first. For example, in the Atlantic region:

(3) (a) MINEPT = MINERQ + MINERO + MINERW + MINERC

(b) MOERT = MINQRE + MINORE + MINWRE + MINCRE

Since, through this definition, the outflows are entirely determined by inflows the migration model needs only 20 inter-regional flows. Internal migration equations can be specified by either outflows or inflows; the second alternative has been chosen for CANDIDE-R.

2.2.2 Inflows from Other Regions

The 20 time series of inter-regional migration flows per source and destination were compiled from a Statistics Canada[4] tabulation done for the Department of Regional Economic Expansion. The tabulation consists of annual estimates based on the transfer of files of family allowance as compiled by the Department

4. Statistics Canada, Census Division, Population Estimates and Demographic Projections Section.

of National Health and Welfare and used by Statistics Canada to calculate the inter-census provincial populations.[5] Data was available only for the ten year period 1961-62 to 1970-71. This data on inter-regional migration flows is found in Table 1.

2.2.3 Gross Immigration

Statistics on immigration from other countries, per province of destination are available in the quarterly bulletin[6], "Immigration", published by the Department of Manpower and Immigration. The estimates for the census years were obtained by interpolation from the cumulative total of the four quarters closest to the census years (July 1-June 30). This data is used in determining the share of total immigrants to each region.

5. Statistics Canada foresees an improvement in the estimation of inter-regional migration flows, using data from Income Tax forms rather than family allowance.

6. "Immigration", Quarterly Statistics, Department of Manpower and Immigration.

2.2.4 Total Net Migration

The estimate of total net migration per region, which is implicitly included in the annual regional population projection, can be isolated easily from the following identity, as values exist for all right-hand side variables:

(4) $TNMr = POPr_t - POPr_{t-1} - NBr_t + NDr_t$

where: TNMr = total net migration in region r

$POPr_t$ = total population in region r, on June 1st of year t.

NBr_t = number of live births in region r between June 1st, year t-1 and May 30 year t.

NDr_t = number of deaths in region r between June 1st of year t-1 and May 30 of year t.

The results of the calculation are shown in the last column of Table 1.

2.2.5 Gross Emigration

As there already exist extimates for four of the five terms of identity 1, i.e. TNMr, MINrRT, MOrRT, and MINrXT, the last one, MOrXT, maybe calculated residually over the sample period. Although this procedure may appear arbitrary, at first glance, it permits indirect access to components of population growth used by Statistics Canada for inter-census projections. Determining gross emigration in this manner results in these figures reflecting the entirety of five year census revisions of population. That is, the figure for net emigration includes all the population revisions, from equation (4), while population estimates in CANDIDE-R do not include these revisions.[7]

3. The Theoretical Framework of the Model

The explanation of migration flows in CANDIDE-R must be based on the theory that economic conditions are the main reasons for population movements. The migration flow determination mechanism in a dynamic

7. Population is revised but inter-regional migration flows are not modified. Births and deaths cannot be revised since they are calculated at source each year.

model requires examination of both short and medium-term factors. The decision to move out of a region can be viewed as the function of medium term "push and pull factors" which are relatively constant over time, and which determine the medium run trend in inter-regional migration flows. However, the decision to move will also depend to a large extent on the availability of jobs and other short-term labour conditions, which may lead to wide fluctuations in migration flows about their trend values. The volume and direction of migration flows over the longer run, then, are functions of medium term social, economic, and demographic and cultural factors, while annual fluctuations may be attributed to short-term factors.[8]

Inter-regional migration flows are endogenously determined in CANDIDE-R. However, international immigration is viewed as a policy variable and the total level is therefore exogenous to the model although share distributions of immigration among the five regions are

8. For a similar viewpoint see: "Theodore P. Lianos The Migration Process and Time Lag", Journal of Regional Science, December 1972.

determined by stochastic equations. Emigration from each region is exogenous.

3.1 Medium Term Factors

Migration flows are positively related to per capita income of the receiving region and negatively related to that of the source region. This relation is relatively simple to specify in inter-regional flows, and in the equations for distribution of immigration is expressed by a ratio of regional to national personal disposable income per capita.

In the inter-regional flow equations, pooled time series - cross section data was used which permitted incorporation of a distance term in the constant term of each regression. This in no way affects the value of any coefficients other than the constant term and allows verification of the notion that the order of preference of the receiving region corresponds to the order of distances between regions.[9]

3.2 Short Term Factors

Turning now to the short term factors influencing migration flows, it is hypothesized that the volume of migration flows is positively related to the unemployment rate of the source region, and negatively related to that of the receiving region. Inter-regional migration flow equations include the unemployment rates of both source and receiving regions, while the share distribution equations for immigrants include the lagged ratio of regional to national unemployment rates.

As mentioned above, it was difficult to find an economic justification for the relatively high levels of internal migration to the Atlantic and Prairie regions. The theory proposed earlier was that

9. A similar coefficient on each of the four constant terms of the equations would indicate that the order of preference was the inverse of the order of distances.

these flows are largely the result of "back migration", or that a portion of earlier migration flows was returning to its original place of domicile. For this reason the inflow equation for the Atlantic and Prairies, from other regions, include lagged outflows from these regions. A positive ceofficient on these explanatory variables would give support to the above hypothesis. For example see equations (49.1) - (49.4) and (49.25) - (49.28) in Section 4 below.

3.3 Migration Equations: Implicit Form

The stochastic equations for migration are specified in terms of migration rates of gross flows. Migration statistics based on census years have been assigned, for use in the context of CANDIDE-R, to the year corresponding to the July census, (for example migration during the 1964-65 census year is defined as migration in 1965). The dependent variables have been related to the explained variable's observation of the initial year (1964, in the above example), which effectively embodies a six-month adjustment lag in the estimated functions.

Gross international emigration from the five regions has been left exogenous to the model, and gross international immigration is exogenous at the national level, but regional shares are determined by the following type of formulation.

$$(5) \quad \frac{MINrXT}{TMINXT} = \beta_1 + \beta_2 \left[\frac{YDr}{POPr} \Big/ \frac{YDC}{POP} \right] + \beta_3 \left[\frac{URATEr}{URATE} \right] + \beta_4 \left[\frac{PRATEr}{PRATE} \right] + \mu$$

where:

MINrXT	=	gross immigration to region r
TMINXT	=	total (national) gross immigration
YDr	=	personal disposable income in region r
POPr	=	population in region r
YDC	=	personal disposable income, national
POP	=	total population
URATEr	=	unemployment rate in region r
URATE	=	national unemployment rate
PRATEr	=	participation rate in region r
PRATE		national participation rate.

The coefficients β_2 , β_3 and β_4 represent distributed lags over two periods. In the mnemonic table "r" is replaced by the letter E, Q, O, W, or C (or BC) depending on the region in question.

The basic form of the equations for internal flows from region "s" to region "r" is:

$$(6)\quad \frac{MINrRs}{POPr_{-1}} = \beta_1\ (D) + \beta_2 \left[\frac{YPr}{POPr}\right]_{-1} + \beta_3 \left[\frac{YPs}{POPs}\right]_{-1} + \beta_4\ URATEr_{-1} + \beta_5\ URATEs_{-1} + \mu.$$

where: MINrRs = gross regional migration from region s to region r

D = distance between region s and region r

YPr = personal income in region r

YPs = personal income in region s

Since there are only ten annual observations for each inter-regional migration flow, it was hypothesized that the explanatory variables for a region's four inflows have the same influence on all four migration flows regardless of the source region. For example, a given level of unemployment in a region produces, ceteris paribus, an identical inflow from all four source regions. In other words the coefficient on URATEr is identical for all four inflow equations for region r. Not withstanding some of the conditions involved in the estimation technique, the constant term[9] accounts for differences in the migrants' preference. Thus the four inflows of each region are calculated at the same time, through the use of pooled time series - cross section data. The parameter D, the constant term, represents the distance in miles between the principal urban centres of each region (in the Prairies the geographical centre, Saskatoon was selected as the reference point).

9. A similar coefficient on each of the four constant terms of the equations would indicate that the order of preference was the inverse of the order of the distance.

4.- Migration Equations: Empirical Results

The 20 stochastic equations for inter-regional migration flows and the 5 for shares of international immigration are found below. An explanation of the mnemonics is found in the Appendix. Looking first at the immigration share equations, these equations include in 3 cases as explanatory variables relative (regional/national) per capita income, the relative unemployment rate and the relative participation rate. Two equations include only the relative unemployment rate as an explanatory variable. The signs on the coefficients are all in accordance with a priori expectations. The positive coefficient on per capita income indicates a stronger attraction for immigrants the higher is regional income relative to the national level. Similarly, a relatively high unemployment rate discourages immigration, while relatively high participation rates draw more immigrants, in anticipation of good labour market conditions.

Turning to the inter-regional flow equations, it is worth noting once more that the lagged outflow variables in the inflow equations for the Atlantic region and the Prairies have very significant positive coefficients, supporting the concept of back-migration. The other explanatory variables included in these and other regional inflow equations are per capita income (source and receiving regions) and the unemployment rate (source and receiving regions). The coefficients on the per capita income variables are positive and negative respectively, and the signs of the unemployment coefficients are negative and positive respectively. Where the unemployment rate enters the equation in relative form the coefficients have negative signs. A discussion of the estimation techniques follows below in Section 4.6, and in Appendix A.

4.6 A Brief Analysis of Empirical Results

In comparing the results of the inter-regional migration flows which were estimated with the Generalized Least Squares[11] (GLS) technique with the results obtained using Ordinary Least Squares (OLS), the advantages of GLS are not clearcut if only the S.E.E. and $\bar{R}^2$ values are examined. According to Table 5, which compares GLS and OLS results for the Atlantic region, most of the coefficients, $\hat{\beta}$, are considerably modified with improved t-statistics in most cases. This demonstrates that GLS tends to yield more robust estimators. Treating the autocorrelation separately for each region seemed an appropriate method, since the $\hat{\rho}$ coefficients have generally very different values as is seen in the estimation results above. The precision of the values predicted by the pooled time series-cross section equations has been markedly improved through the use of GLS over OLS.

11. For a discussion of Generalized Least Squares see Appendix A: Estimation Techniques.

Table 5: Comparison of Ordinary and Generalized Least Squares equations for MINERs [1]

	Ordinary Least Squares $\hat{\beta}$	t	Generalized Least Squares $\hat{\beta}$	t
Constant (512)	0.0000011	1.0	0.0000016	2.6
Constant (800)	0.0000033	3.6	0.0000036	4.6
Constant (1952)	0.0000001	1.0	0.0000002	2.3
Constant (2784)	-0.000001	0.9	-0.0000002	1.4
(MINrRE/POPE)-1	0.2814680	5.5	0.3386734	10.9
(MINrRE/POPE)-2	0.2801510	5.1	0.2078984	5.0
(URATEs/URATEE)-1	0.0010801	1.8	0.0009356	3.3
S.E.E.	0.0004		0.0003	
$\bar{R}^2$	0.9924		0.9935	

1 s =sending region, r = receiving region

The results of the migration model for gross flows are on the whole quite satisfactory (see Table 6). It is of course difficult to reach a comparable level of precision for net flows, since errors are accumulated in both inflow and outflow calculations.

From a theoretical point of view it is felt that given the inaccuracy of national migration flow data, there is a tradeoff between the robustness of the coefficients and the quality of specification as obtained from the explanation of gross flows on the one hand and the accuracy of the model in terms of net flows on the other.[12]

12. The first migration model tested was the type which explained net flows (TNMr) directly by stochastic equations. Even though the values predicted over the sample period seemed better, this approach was discarded for theoretical reasons (wrong signs on some coefficients) and because of unstable behaviour of the equations during a multi-period simulation.

Table 6: Results of the migration model[1] for the sample period (in 000 of persons).

	MINrRT		MOrRT		MINrXT	
	Calculated	Observed	Calculated	Observed	Calculated	Observed
Atlantic						
1963	31.5	31.7	43.3	44.0	2.3	2.3
1964	34.5	34.4	45.6	49.6	2.4	2.4
1965	37.9	34.6	47.7	53.2	2.8	2.7
1966	38.6	38.6	55.9	56.6	3.8	3.8
1967	41.5	42.2	54.7	56.9	5.2	4.6
1968	42.9	44.0	54.3	48.0	4.7	4.3
1969	40.8	40.9	53.1	46.6	4.1	4.4
1970	38.1	38.8	58.0	58.1	4.2	4.2
1971	43.3	43.8	50.5	49.9	3.9	4.1
Quebec						
1963	48.7	49.7	45.1	43.6	19.5	20.5
1964	47.5	47.1	48.8	42.3	24.0	24.2
1965	47.9	47.4	49.7	48.2	27.8	26.8
1966	48.7	50.6	55.5	53.9	35.6	34.3
1967	49.4	50.1	57.3	59.2	44.4	44.5
1968	48.6	45.7	58.7	56.3	42.6	39.7
1969	43.9	44.9	61.8	58.7	33.4	32.5
1970	42.1	41.7	69.2	74.2	27.7	26.8
1971	40.5	40.0	67.4	75.5	24.4	22.4
Ontario						
1963	89.8	92.9	88.1	88.6	41.6	41.3
1964	98.7	101.4	90.9	91.5	53.4	54.7
1965	103.6	107.8	95.9	93.3	65.0	67.0
1966	119.9	116.8	101.2	101.4	90.2	90.3
1967	115.4	121.9	105.4	108.9	114.8	115.8
1968	117.4	109.0	104.2	100.9	104.3	105.6
1969	119.9	106.4	99.7	100.6	90.0	89.9
1970	133.6	143.2	99.3	95.9	86.9	87.8
1971	136.9	139.5	97.3	96.2	73.8	73.6

Table 6 (Cont'd) Results of the migration model for the sample period (in 000 of persons.)

	MINrRT		MOrRT		MINrXT	
	Calculated	Observed	Calculated	Observed	Calculated	Observed
Prairies						
1963	58.7	59.7	64.5	68.0	9.2	8.6
1964	61.7	59.3	69.4	71.3	10.8	9.5
1965	60.2	60.1	81.2	78.4	11.9	11.6
1966	63.3	62.4	89.1	97.3	15.8	15.9
1967	73.1	71.5	90.9	96.2	22.7	22.4
1968	71.0	73.6	92.7	82.7	23.2	26.5
1969	68.8	66.9	84.7	78.2	22.4	23.3
1970	66.6	67.5	93.9	96.0	20.9	20.3
1971	62.7	65.0	92.4	90.4	15.8	17.1
British Columbia						
1963	52.2	51.8	39.9	41.6	8.1	8.1
1964	56.3	58.7	44.0	41.3	10.6	10.4
1965	69.6	66.6	44.6	43.5	14.6	14.2
1966	78.3	85.9	47.1	45.2	20.4	21.4
1967	84.2	89.6	55.2	54.4	26.4	26.1
1968	85.2	79.3	55.2	57.6	25.8	24.5
1969	78.3	74.7	52.4	49.9	22.0	21.9
1970	91.4	88.9	51.3	55.9	21.9	22.6
1971	80.0	81.1	55.8	57.3	20.1	20.8

1) The calculated values were taken from a simulation of the entire CANDIDE-R model in which observed values were used for lagged endogenous variables (single period simulation).

4.1 Atlantic Region

Migration in Atlantic from Quebec

(49.1) $$MINERQ = POPE_{-1} * [0.0000016 \;(512) \;[2.63] + 0.3387\,(MINQRE/POPE)_{-1} \;[10.87] + 0.2079\,(MINQRE/POPE)_{-2} \;[5.01] + 0.00094\,(URATEQ/URATEE)_{-1}] \;[3.35]$$

ρ = 0.0128

Migration In Atlantic from Ontario

(49.2) $$MINERO = POPE_{-1} * [0.0000036 \;(800) \;[4.63] + 0.3387\,(MINORE/POPE)_{-1} \;[10.87] + 0.2079\,(MINORE/POPE)_{-2} \;[5.01] + 0.00094\,(URATEQ/URATEE)_{-1}] \;[3.35]$$

ρ = 0.3302

Migration in Atlantic from Prairies

(49.3) $$MINERW = POPE_{-1} * [0.0000002 \;(1952) \;[2.28] + 0.3387\,(MINWRE/POPE)_{-1} \;[10.87] + 0.2079\,(MINWRE/POPE)_{-2} \;[5.01] + 0.00094\,(URATEW/URATEE)_{-1}] \;[3.35]$$

ρ = 0.1593

Migration in Atlantic from British Columbia

(49.4) $$MINERC = POPE_{-1} [0.0000002 \;(2784) \;[1.42]$$

$$+ \underset{[10.87]}{0.3387} \ (MINCRE/POPE)_{-1}$$

$$+ \underset{[5.01]}{0.2079} \ (MINCRE/POPE)_{-2}$$

$$+ \underset{[3.35]}{0.00094}((URATEC/URATEE)_{-1}]$$

$\rho = 0.7240$

(49.1) - (49.4) estimated using pooled time-series cross - section data.

$\bar{R}^2 = 0.99$
S.E.E. = 0.0003
(*GLS*, 1961- 1971) Hildreth-Lu

Immigration in Atlantic

$$(49.7) \quad MINEXT = POLIMM *[\underset{[3.55]}{-0.2684} + \sum_{i=1}^{2} \alpha_i((YDE/POPE)/(YDC/POP))_{t-i}$$

$$+ \sum_{i=1}^{2} \beta_i \ (URATEE/URATE)_{t-i} + \sum_{i=1}^{2} \gamma_i (PRATEE/PRATE)_{t-i} \ ]$$

PDL, Degree 1, $\alpha_3 = 0$

α_1	=	0.0994	[2.78]
α_2	=	0.0497	[2.78]

PDL, Degree 1, $\beta_3 = 0$

β_1	=	-0.0071	[1.69]
β_2	=	-0.0035	[1.69]

PDL, Degree 1, $\gamma_3 = 0$

γ_1	=	0.1573	[2.80]
γ_2	=	0.0786	[2.80]

$\bar{R}^2 = 0.62$
S.E.E. = 0.0018
D.W. = 1.58
(*OLS*, 1961-1971)

$\rho = 0.7840$

4.2 Quebec

Migration in Quebec from Atlantic

$$(49.9) \quad MINQRE = POPQ_{-1} *[\underset{[7.77]}{0.000004} \ (512) + \underset{[1.19]}{0.0003} \ (YPQ/POPQ)_{-1}$$

$$\underset{[2.17]}{-0.0005} \ (YPE/POPE_{-1} \ \underset{[2.41]}{-0.00008} \ (URATEQ)_{-1}$$

$$+ \underset{[1.51]}{0.00004} \; (URATEE)_{-1}]$$

ρ = 0.3107

Migration in Quebec from Ontario

$$(49.10) \quad MINQRO = POPQ_{-1} * [\underset{[20.75]}{0.00002} \; (320) + \underset{[1.19]}{0.0003} \; (YPQ/POPQ)_{-1}$$

$$- \underset{[2.17]}{0.0005} \; (YPO/POPO)_{-1} - \underset{[2.41]}{0.00008} \; (URATEQ)_{-1}$$

$$+ \underset{[1.51]}{0.00004} \; (URATEO)_{-1}]$$

ρ = 0.5261

Migration in Quebec from Prairies

$$(49.11) \quad MINQRW = POPQ_{-1} * [\underset{[9.22]}{0.000001} \; (1472) + \underset{[1.19]}{0.0003} \; (YPQ/POPQ)_{-1}$$

$$- \underset{[2.17]}{0.0005} \; (YPW/POPW)_{-1} - \underset{[2.41]}{0.00008} \; (URATEQ)_{-1}$$

$$+ \underset{[1.51]}{0.00004} \; (URATEW)_{-1}]$$

ρ = 0.1909

Migration in Quebec from British Columbia

$$(49.12) \quad MINQRC = POPQ_{-1} * [\underset{[5.88]}{0.0000007} \; (2304) + \underset{[1.19]}{0.0003} \; (YPQ/POPQ)_{-1}$$

$$- \underset{[2.17]}{0.0005} \; (YPC/POPC)_{-1} - \underset{[2.41]}{0.00008} \; (URATEQ)_{-1}$$

$$+ \underset{[1.51]}{0.00004} \; (URATEC)_{-1}]$$

ρ = 0.6509

(49.9) - (49.12) estimated using pooled time-series cross-section data.

$\bar{R}^2$ = 0.99
S.E.E. = 0.0002
(*GLS*, 1961-1971) Hildreth-Lu

Immigration in Quebec

(49.15) $$MINQXT = POLIMM * [\underset{[5.63]}{0.4528} + \sum_{i=1}^{2} \beta_i (URATEQ/URATE)_{t-i}]$$

PDL, Degree 1, $\beta_3 = 0$			
β_1	=	-0.1252	[3.30]
β_2	=	-0.0626	[3.30]

$\bar{R}^2$ = 0.79 ρ = 0.9
S.E.E. = 0.0097
D.W. = 0.85
(*OLS*, 1961-1971) Hildreth-Lu

4.3 Ontario

Migration in Ontario from Atlantic

(49.17) $$MINORE = POPO_{-1} * \exp. [\underset{[21.52]}{-0.9638} \ln (800)$$

$$+ \underset{[2.16]}{1.3095} \ln (YPO/POPO)_{-1}$$

$$- \underset{[1.34]}{0.8252} \ln (YPE/POPE)_{-1} \underset{[4.55]}{-0.4966} \ln (URATEO)_{-1}$$

$$+ \underset{[2.80]}{0.3772} \ln (URATEE)_{-1}]$$

ρ = 0.1181

Migration in Ontario from Quebec

(49.18) $$MINORQ = POPO_{-1} * \exp. [\underset{[26.41]}{-1.0375} \ln (320)$$

$$+ \underset{[2.16]}{1.3095} \ln (YPO/POPO)_{-1}$$

$$- \underset{[1.34]}{0.8252} \ln (YPQ/POPQ)_{-1}$$

$$- \underset{[4.55]}{0.4966} \ln (URATEO)_{-1}$$

$$+ \underset{[2.80]}{0.3772} \ln (URATEQ)_{-1}]$$

ρ = 0.4499

Migration in Ontario from Prairies

$$(49.19) \quad MINORW = POPO_{-1} * \exp. [\underset{[29.72]}{-0.8470}\; ln\,(1280)$$

$$+ \underset{[2.16]}{1.3095}\; ln\,(YPO/POPO)_{-1}$$

$$- \underset{[1.34]}{0.8252}\; ln\,(YPW/POPW)_{-1}$$

$$- \underset{[4.55]}{0.4966}\; ln\,(URATEO)_{-1}$$

$$+ \underset{[2.80]}{0.3772}\; ln\,(URATEW)_{-1}]$$

$\rho = 0.6945$

Migration in Ontario from British Columbia

$$(49.20) \quad MINORC = POPO_{-1} * \exp. [\underset{[30.52]}{-0.8834}\; ln\,(2112)$$

$$+ \underset{[2.16]}{1.3095}\; ln\,(YPO/POPO)_{-1}$$

$$- \underset{[1.34]}{0.8252}\; ln\,(YPC/POPC)_{-1}$$

$$- \underset{[4.55]}{0.4966}\; ln\,(URATEO)_{-1}$$

$$+ \underset{[2.80]}{0.3772}\; ln\,(URATEC)_{-1}]$$

$\rho = -0.1703$

(49.17) - (49.20 estimated using pooled time-series cross-section data.

$\bar{R}^2 = 0.95$
$S.E.E. = 0.0908$
(*GLS*, 1961-1971) Hildreth-Lu

<u>Immigration in Ontario</u>

$$(49.23) \quad MINOXT = POLIMM * [\underset{[3.47]}{0.7502} + \sum_{i=1}^{2} \alpha_i ((YDO/POPO)/(YDC/POP))_{t-i}$$

$$+ \sum_{i=1}^{2} \beta_i \,(URATEO/URATE)_{t-i}]$$

PDL, Degree 1, $\alpha_3 = 0$		*PDL*, Degree 1, $\beta_3 = 0$	
$\alpha_1 = 0.0092$	[0.07]	$\beta_1 = -0.2174$	[5.18]
$\alpha_2 = 0.0046$	[0.07]	$\beta_2 = -0.1087$	[5.18]

$\bar{R}^2$ = 0.97
S.E.E. = 0.0086
D.W. = 2.18
(*OLS*,1961-1971) Hildreth-Lu

4.4 Prairies

Migration in Prairies from Atlantic

$$(49.25)\quad MINWRE = POPW_{-1} * [\underset{[1.14]}{0.0000004}\ (1952) + \underset{[7.99]}{0.2562}\ (MINERW/POPW)_{-1} + \underset{[4.31]}{0.0023}\ (YPW/POPW)_{-1} - \underset{[4.31]}{0.0022}\ (YPE/POPE)_{-1} - \underset{[5.50]}{0.0004}\ (URATEW)_{-1} + \underset{[3.90]}{0.0002}\ (URATEE)_{-1}]$$

$\rho = 0.0079$

Migration in Prairies from Quebec

$$(49.26)\quad MINWRQ = POPW_{-1} * [\underset{[1.07]}{0.0000003}\ (1472) + \underset{[7.99]}{0.2562}\ (MINQRW/POPW)_{-1} + \underset{[4.31]}{0.0023}\ (YPW/POPW)_{-1} - \underset{[4.31]}{0.0022}\ (YPQ/POPQ)_{-1} + \underset{[5.50]}{0.0004}\ (URATEW)_{-1} + \underset{[3.90]}{0.0002}\ (URATEQ)_{-1}]$$

$\rho = 0.2692$

Migration in Prairies from Ontario

$$(49.27)\quad MINWRO = POPW_{-1} * [\underset{[18.30]}{0.00005}\ (1280) + \underset{[7.99]}{0.2562}\ (MINORW/POPW)_{-1}$$

$$+ \underset{[4.31]}{0.0023} \; (YPQ/POPW)_{-1} - \underset{[4.31]}{0.0022} \; (YPO/POPO)_{-1}$$

$$- \underset{[5.50]}{0.0004} \; (URATEW)_{-1} + \underset{[3.90]}{0.0002} \; (URATEO)_{-1}]$$

ρ =-0.3823

Migration in Prairies from British Columbia

(49.28) $MINWRC = POPW_{-1} * [\underset{[10.93]}{0.000007}$ *(832)*

$$+ \underset{[7.99]}{0.2562} \; (MINCRW/POPW)_{-1}$$

$$+ \underset{[4.31]}{0.0023} \; (YPW/POPW)_{-1} - \underset{[4.31]}{0.0022} \; (YPC/POPC)_{-1}$$

$$- \underset{[5.50]}{0.0004} \; (URATEW)_{-1} + \underset{[3.90]}{0.0002} \; (URATEC)_{-1}]$$

ρ = -0.3187

(41.25) - (41.28) estimated using pooled time-series cross-section data.

$\bar{R}^2$ = 0.99
S.E.E. = 0.0003
(*GLS*, 1961-1971) Hildreth-Lu

<u>Immigration in Prairies</u>

(49.31) $MINWXT = POLIMM * [\underset{[5.99]}{0.1914}$

$$+ \sum_{i=1}^{2} \beta_i (URATEW/URATE)_{t-i}]$$

<u>*PDL*, Degree.1, $\beta_3 = 0$</u>

β_1 = -0.0749 [2.33]
β_2 = -0.0375 [2.33]

$\bar{R}^2$ = 0.74
S.E.E. = 0.0081
D.W. = 1.61
(*OLS*, 1961-1971) Hildreth-Lu

ρ = 0.7918

4.5 British Columbia

Migration in British Columbia from Atlantic

(49.33)
$$MINCRE = POPC_{-1} * \exp.\ [\underset{[19.46]}{-0.7448}\ ln\ (2784) + \underset{[2.39]}{1.0562}\ ln\ (YPC/POPC)_{-1} \underset{[1.50]}{-0.6206}\ ln\ (YPE/POPE)_{-1} - \underset{[7.30]}{0.4784}\ ln\ (URATEC)_{-1}]$$

$\rho = 0.5117$

Migration in British Columbia from Quebec

(49.34)
$$MINCRQ = POPC_{-1} * \exp.\ [\underset{[1.10]}{0.6716}\ ln\ (2304) + \underset{[2.39]}{1.0562}\ ln\ (YPC/POPC)_{-1} \underset{[1.50]}{-0.6206}\ ln\ (YPQ/POPQ)_{-1} \underset{[7.30]}{-0.4784}\ ln\ (URATEC)_{-1}]$$

$\rho = 0.9944$

Migration in British Columbia from Ontario

(49.35)
$$MINCRO = POPC_{-1} * \exp.\ [\underset{[29.50]}{-0.5471}\ ln\ (2112) + \underset{[2.39]}{1.0562}\ ln\ (YPC/POPC)_{-1} - \underset{[1.50]}{0.6206}\ ln\ (YPO/POPO)_{-1} - \underset{[7.30]}{0.4784}\ ln\ (URATEC)_{-1}]$$

$\rho = -0.3394$

Migration in British Columbia from Prairies

(49.36)
$$MINCRW = POPC_{-1} * \exp.\ [\underset{[16.99]}{-0.5054}\ ln\ (832) + \underset{[2.39]}{1.0562}\ ln\ (YPC/POPC)_{-1} - \underset{[1.50]}{0.6206}\ ln\ (YPW/POPW)_{-1} - \underset{[7.30]}{0.4784}\ ln\ (URATEC)_{-1}]$$

$\rho = 0.5003$

(49.33) - (49.36) estimated using pooled time-series cross-section data.

$\bar{R}^2$ =0.99
$S.E.E.$ =0.0881
(GLS, 1961-1971) Hildreth-Lu

<u>Immigration in British Columbia</u>

(49.39)

$$MINCXT = POLIMM * [\underset{[5.37]}{-1.2574} + \sum_{i=1}^{2} \alpha_i ((YDBC/POPC)/(YDC/POP))_{t-i} + \sum_{i=1}^{2} \beta_i (URATEC/URATE)_{t-i} + \sum_{i=1}^{2} \gamma_i (PRATEC/PRATE)_{t-i}]$$

PDL, Degree 1, α_3 =0			PDL, Degree 1, β_3 = 0		
α_1	= 0.2005	[3.53]	β_1	= -0.0066	[0.50]
α_2	= 0.1003	[3.53]	β_2	= -0.0033	[0.50]

PDL, Degree 1, γ_3 =0		
γ_1	= 0.7093	[7.80]
γ_2	= 0.3546	[7.80]

$\bar{R}^2$ = 0.90
$S.E.E.$ = 0.0049
$D.W.$ = 1.62
(OLS, 1961-1971) Hildreth-Lu

ρ = 0.2797

5.- How the Migration Block Works

As mentioned earlier the migration data was arranged according to census year thus enabling use of variables lagged over one period as explanatory variables for all stochastic equations in the migration model. Since only exogenous variables or lagged endogenous variables are used as input to Block 49 (the migration block), it can be specified as a recursive block which is an appreciable advantage in terms of the operation cost and stability of CANDIDE-R.

It would have been possible, in principle, to constrain the coefficients of the inter-regional migration equations in such a way as to produce a sum total of internal flows which would be zero at the national level during the sample period. However, this condition could not be maintained during the forecast period without the addition of another adjustment mechanism. Since internal flows per source and destination are used only to calculate total gross inflows and outflows (see 2.2.1, equations (3a) and (3b)), it was possible to devise an adjustment mechanism affecting only the MOrRT variables, these being constrained by the necessary condition that: Σ MOrRT = Σ MINrRT. The adjustment mechanism actually used in Block 49 is:

$$(7)\quad MOrRT = \sum_{s} MINsRr + \left[\left(\sum_{r} MINrRT - \sum_{r} \sum_{s} MINsRr \right) \left(\frac{MINrRT}{\sum_{r} MINrRT} \right) \right]$$

$$r, s = 1, 2, \ldots 5$$
$$r \neq s$$

Finally, the equation for net Canadian immigration, in Block 22 of CANDIDE 1.1, has been replaced in CANDIDE-R by an identity to compute the sum of total net regional migration (TNMr) and the net migration of the Yukon and North-West Territories, TNMYNW. As shown in Table 3, the balance in each region of migration to and from the Yukon and N.W.T. is generally insignificant. This is why it is assumed to be zero for all cases except foreign migration which is taken into account in the exogenous values of TNMYNW.

6.- Concluding Comments

International immigration accounts for a considerable portion of national population growth and it is unevenly distributed throughout the regions. Inter-regional flows play on equally important role in the spatial allotment of resources between regions, and in the regional adjustment of the labour markets. This is why, during the course work on CANDIDE-R a complex system was developed for determing migration flows, to complete the population algorithm. The net migration flows for each region has been broken down according to gross emigration and gross immigration; as well as gross inter-regional inflows and outflows. The first flow is exogenous at the regional level, while gross immigration is exogenous at the national level but endogenously distributed among the regions, by five share equations. All inter-regional flows are endogenous and have been explained by equations using a Generalized Least Squares technique and pooled time series-cross section data.

Appendix A Estimation Techniques[1]

The traditional linear regression model is characterized by its two main hypotheses for error terms, ie. normality and homoscedasticity and non-autocorrelation.

(1) Normality: $\varepsilon \sim N\ (O, \Sigma)$

(2) Homoscedasticity and non-autocorrelation:

$E\ (\varepsilon\varepsilon) = \sigma^2\ I_t$ or $E\ (\varepsilon_i \varepsilon_j) = \sigma^2$ for $i = j$, and zero for $i \neq j$

In these two hypotheses O is a column vector of dimension T, Σ the variance-covariance matrix and I_t a TxT identity matrix. The other related hypotheses for a traditional linear regression model are that the component variables of X are non-stochastic, the rank of X is equal to K<T and there is no linear relationship between the explanatory variables. If all these conditions are adhered to, the Ordinary Least Squares estimator $\hat{\beta}$ is obtained from the following equation: (3) $\hat{\beta} = (X'X)^{-1}\ X'Y$

1. This section draws heavily from the following sources: Henri Theil, Principles of Econometrics John Wiley and Sons, Inc. New York 1971, pp. 106-129 and Jan Kmenta Elements of Econometrics, MacMillan Co., New York, 1971, Chapter 12.

Under the above conditions we can prove that the estimator $\hat{\beta}$, is the best linear unbiased estimator (BLUE) and that it is also asymptotically unbiased efficient and consistent.

The distribution of this estimator is:

(4) $\hat{\beta} \sim N\ (\beta,\ \sigma^2\ (X^1X)^{-1})$

If σ^2 is unknown:

(5) $\hat{\beta} \sim N\ (\beta,\ S^2\ (X^1X)^{-1})$

where $S^2 = e^1e/t-k$ is an unbiased estimator of σ^2 and has the same properties.

If we drop the hypotheses on homoscedasticity and non-autocorrelation, (2) above, the resulting model may be called a general linear regression model. Where r is the new variance-covariance model, conditions (1) and (2) become:

(6) $\varepsilon \sim N\ (O,r)$

(7) $$\Omega = \sigma_{ij} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \ldots & \sigma_{1t} \\ \sigma_{21} & & \vdots \\ \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} \ldots & \sigma_{nn} \end{bmatrix}$$

The estimator $\hat{\beta}$ of the general form of the Generalized Least Squares model is given by the following equation (comparable to (3)):

$$(8) \quad \hat{\beta} = (X^1\Omega^{-1}X)^{-1} \; X^1\Omega^{-1}Y$$

It is simple to prove that Ordinary Least Squares is merely a special case of equation (8), where $\Omega = \sigma^2 I_t$ and $\Omega^{-1} = (^1/\sigma^2)\, I_t$. Similarly, a heteroscedastic model is the special case where Ω is a diagonal matrix, but the diagonal elements are not necessarily the same, for example

$$(9) \quad \Omega = \begin{bmatrix} \sigma_{11} & \sigma & \cdots & \sigma \\ \sigma & \sigma_{22} & \cdots & \sigma \\ \vdots & \vdots & \ddots & \vdots \\ \sigma & \sigma & \cdots & \sigma_{tt} \end{bmatrix} \qquad \Omega^{-1} = \begin{bmatrix} ^1/\sigma_{11} & \sigma & \cdots & \sigma \\ \sigma & ^1/\sigma_{22} & \cdots & \sigma \\ \vdots & \vdots & \ddots & \vdots \\ \sigma & \sigma & \cdots & ^1/\sigma_{tt} \end{bmatrix}$$

where $\sigma_{11} = \sigma_1^2$, $\sigma_{22} = \sigma_2^2$ etc. In our example, as in many other applications of Generalized Least Squares, the problem is really to find the appropriate estimators for the elements of the matrix Ω.

Estimation with pooled time-series cross-sectional data

The Generalized Least Squares technique used in the estimation of migration equations is found in Kmenta[2]. This approach consists of making specific assumptions on the behaviour of the error term, when observations on time-series and cross-section are used simultaneously, in order to incorporate these in the matrix Ω of (8) above. With regard to cross-sectional data, for example, the most commonly used assumptions are those of mutual independence of the error term, and heteroscedasticity. For time-series data it is often assumed that errors are correlated yet not heteroscedastic. In the framework of this analysis, the geographical regions are the cross-sectional units, making the assumption of mutual independence of errors an implausible one. The model chosen for the estimation of internal migration flows is a cross-sectionally correlated and time-wise autoregressive

2. Jan Kmenta, op.cit. p. 512-514.

model. The assumptions on the error term are as follows:

(10) Heteroscedasticity: $E\ (\varepsilon_{it}^2) = \sigma_{ii}$

(11) Mutual Correlation: $E\ (\varepsilon_{it}\ \varepsilon_{jt}) = \sigma_{ij}$

(12) Auto-regression: $E_{it} = \rho_i\ \varepsilon_{i,t-1} + \mu_{it}$

where: $\mu_{it} \sim N\ (0,\ \emptyset_{it})$

$E\ (\varepsilon_{i,t-1}\ \mu_{jt}) = 0 \quad i,j = 1,2, \ldots N$

$E\ (\mu_{it}\ \mu_{jt}) = \emptyset_{ij} \quad t,s = 1,2, \ldots T$

$E\ (\mu_{it}\ \mu_{js}) = 0 \quad t \neq s$

The initial value of ε has the following properties, by assumption:

(13) $\varepsilon_{io} \sim N \quad 0,\ \dfrac{\emptyset i_1}{1-\rho_i^2}$

(14) $$E\,(\varepsilon_{io}\,\varepsilon_{jo}) = \frac{\phi\,ij}{1-\rho_i\,\rho_j}$$

Thus the matrix Ω for this model is:

(15)
$$\Omega = \begin{bmatrix} \sigma_{11} P_{11} & \sigma_{12} P_{12} & \cdots & \sigma_{1n} P_{1n} \\ \sigma_{21} P_{21} & \sigma_{22} P_{22} & \cdots & \sigma_{2n} P_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{n1} P_{n2} & \sigma_{n2} P_{n2} & & \sigma_{nn} P_{nn} \end{bmatrix}$$

where

$$P_{ij} = \begin{bmatrix} 1 & \rho_j & \rho_j^2 & \cdots & \rho_j^{t-i} \\ \rho_i & 1 & \rho_j & \cdots & \rho_j^{t-2} \\ \rho_i^2 & \rho_i & 1 & & \rho_j^{t-3} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \rho_i^{t-1} & \rho_i^{t-2} & \rho_i^{t-3} & \cdots & 1 \end{bmatrix}$$

In practice, the procedure used here can effectively be broken down into five stages. First, we applied Ordinary Least Squares (equation (3)), to the matrix notation equation $Y = X\beta + \varepsilon$, in order to calculate the e_{it} residuals, which serve as estimators for ε_{it}. Secondly the estimators for ρ_i were obtained from the following formulation:

$$\hat{\rho}_i = \frac{\sum_i \sum_t e_{it}\, e_{i,t-1}}{\sum_i e^2_{i,t-1}} \qquad t = 2,3, \ldots t \qquad i = 1,2,3,4$$

In the third place we applied Ordinary Least Squares to the new system:

$$(17) \quad Y^* = X^*\beta + \mu^*$$

where:

$$Y^* = Y_{it} - \hat{\rho}_i\, Y_{i,t-1}$$

$$X^* = X_{it} - \hat{\rho}_i\, X_{i,t-1}$$

$$\mu^* = \varepsilon_{it} - \hat{\rho}_i\, \varepsilon_{i,t-1}$$

In the fourth stage the variance-covariance matrix was obtained from:

(18)

$$\hat{\Phi} = \begin{bmatrix} \hat{\Phi}_{11} \; I_{t-1} & \hat{\Phi}_{12} \; I_{t-1} & \cdots & \hat{\Phi}_{1n} \; I_{t-1} \\ \hat{\Phi}_{21} \; I_{t-1} & \cdots \cdots \cdots & & \\ \cdot & & & \\ \cdot & & & \\ \hat{\Phi}_{n1} \; I_{t-1} & \hat{\Phi}_{n2} \; I_{t-2} & \cdots & \hat{\Phi}_{nn} \; I_{t-1} \end{bmatrix}$$

where I_{t-1} is a (T-1) x (T-1) identity matrix and $\hat{\Phi}_{ij}$ is defined by: $S_{ij} = \dfrac{\hat{\Phi}_{ij}}{1-\hat{\rho}_i \; \hat{\rho}_i}$

where $\hat{\Phi}_{ij} = \dfrac{1}{T-K-1} \sum_{t=2}^{T} \hat{\mu}^{*}_{it} \; \hat{\mu}^{*}_{jt}$

Due to the small number of observations and the large number of explanatory variables in each equation $\hat{\Phi}_{ij}$

was defined as: $\hat{\Phi}_{ij} = \frac{1}{T-1} \sum_{t=2}^{T} \hat{\mu}^{*}_{it} \hat{\mu}^{*}_{jt}$. This in no way affects the asymptotic properties of the estimators even though the significance tests and the variance of the estimators are slightly modified.

The last stage of this estimation process consists of applying Aitken's method to the transformed data:

$$(19) \quad \hat{\beta} = (X^{*1} \hat{\Phi}^{-1} X^{*})^{-1} X^{*1} \hat{\Phi}^{-1} Y^{*}$$

$$(20) \quad \hat{Y} = X\hat{\beta} + \hat{\rho}_{i} e_{t-1}$$

The last equation allows us to obtain the vector of calculated values, $\hat{Y}$, which corresponds to the Generalized Least Squares algorithm.

Appendix B **Mnemonic Table**

MNEMONIC LIST -- BLOCK 49 -- LISTE DES MNEMONIQUES

MINCBE	EB	49033	3	MIGRATION IN -- BRIT, COLUMBIA FROM ATLANTIC
MINCBO	EB	49035	1	MIGRATION IN -- BRIT, COLUMBIA FROM ONTARIO
MINCBQ	EB	49034	1	MIGRATION IN -- BRIT, COLUMBIA FROM QUEBEC
MINCBT	EI	49037		MIGRATION IN -- BRIT, COLUMBIA TOTAL INTERNAL FLOWS
MINCBW	EB	49036	2	MIGRATION IN -- BRIT, COLUMBIA FROM PRAIRIES
MINCXT	EB	49039	2	IMMIGRATION IN BRITISH COLUMBIA
MINEBC	EB	49004	1	MIGRATION IN --ATLANTIC -- FROM BRITISH-COLUMBIA
MINEBO	EB	49002	1	MIGRATION IN --ATLANTIC -- FROM ONTARIO
MINEBQ	EB	49001	1	MIGRATION IN --ATLANTIC -- FROM QUEBEC
MINEBT	EI	49005		MIGRATION IN --ATLANTIC -- TOTAL INTERNAL FLOWS
MINEBW	EB	49003	2	MIGRATION IN --ATLANTIC -- FROM PRAIRIES
MINEXT	EB	49007	2	IMMIGRATION IN ATLANTIC
MINOBC	EB	49020	1	MIGRATION IN --ONTARIO -- FROM BRITISH COLUMBIA
MINOBE	EB	49017	3	MIGRATION IN --ONTARIO -- FROM ATLANTIC
MINOBQ	EB	49018	1	MIGRATION IN --ONTARIO -- FROM QUEBEC
MINOBT	EI	49021		MIGRATION IN --ONTARIO -- TOTAL INTERNAL FLOWS
MINOBW	EB	49019	2	MIGRATION IN --ONTARIO -- FROM PRAIRIES
MINOXT	EB	49023	2	IMMIGRATION IN ONTARIO
MINQBC	EB	49012	1	MIGRATION IN --QUEBEC -- FROM BRITISH COLUMBIA
MINQBE	EB	49009	3	MIGRATION IN --QUEBEC -- FROM ATLANTIC
MINQBO	EB	49010	1	MIGRATION IN --QUEBEC -- FROM ONTARIO
MINQBT	EI	49013		MIGRATION IN --QUEBEC -- TOTAL INTERNAL FLOWS
MINQBW	EB	49011	2	MIGRATION IN --QUEBEC -- FROM PRAIRIES
MINQXT	EB	49015	2	IMMIGRATION IN QUEBEC
MINWBC	EB	49028	1	MIGRATION IN --PRAIRIES -- FROM BRITISH COLUMBIA
MINWBE	EB	49025	3	MIGRATION IN --PRAIRIES -- FROM ATLANTIC
MINWBO	EB	49027	1	MIGRATION IN --PRAIRIES -- FROM ONTARIO
MINWBQ	EB	49026	1	MIGRATION IN --PRAIRIES -- FROM QUEBEC
MINWBT	EI	49029		MIGRATION IN --PRAIRIES -- TOTAL INTERNAL FLOWS
MINWXT	EB	49031	2	IMMIGRATION IN PRAIRIES
MOCBT	EI	49038		MIGRATION OUT -BRIT, COLUMBIA TOTAL INTERNAL FLOWS
MOCXT	XP	342	1	EMIGRATION FROM BRIT, COLUMBIA
MOEBT	EI	49006		MIGRATION OUT -ATLANTIC TOTAL INTERNAL FLOWS
MOEXT	XP	338	1	EMIGRATION FROM ATLANTIC
MOOBT	EI	49022		MIGRATION OUT -ONTARIO -- TOTAL INTERNAL FLOWS
MOOXT	XP	340	1	EMIGRATION FROM ONTARIO
MOQBT	EI	49014		MIGRATION OUT -QUEBEC -- TOTAL INTERNAL FLOWS
MOQXT	XP	339	1	EMIGRATION FROM QUEBEC
MOWBT	EI	49030		MIGRATION OUT -PRAIRIES -- TOTAL INTERNAL FLOWS
MOWXT	XP	341	1	EMIGRATION FROM PRAIRIES
POLIMM	XZ	622	3	UPOLITICAL,DUMMY FOR IMMIGRATION
POP	EI	22030	3	TOTAL POPULATION CANADA
POPC	EI	4806104		POPULATION TOTALE -BRITISH COLUMBIA
POPE	EI	4805704		POPULATION TOTALE -ATLANTIC
POPO	EI	4805904		POPULATION TOTALE -ONTARIO
POPQ	EI	4805804		POPULATION TOTALE -QUEBEC
POPW	EI	4806004		POPULATION TOTALE -PRAIRIES
PRATE	EI	11075	3	TOTAL PARTICIPATION RATE,RATIO
PRATEC	EI	11072	3	TOTAL PARTICIPATION RATE BRITISH COLUMBIA
PRATEE	EI	11068	3	TOTAL PARTICIPATION RATE ATLANTIC
TNMC	EI	49040	5	TOTAL NET MIGRATION BRITISH COLUMBIA
TNME	EI	49008	5	TOTAL NET MIGRATION ATLANTIC
TNMO	EI	49024	5	TOTAL NET MIGRATION ONTARIO
TNMQ	EI	49016	5	TOTAL NET MIGRATION QUEBEC
TNMW	EI	49032	5	TOTAL NET,MIGRATION PRAIRIES

MNEMONIC LIST -- BLOCK 49 -- LISTE DES MNEMONIQUES

URATE	EI	11062	6	TOTAL UNEMPLOYMENT RATE	
URATEC	EI	11057	6	UNEMPLOYMENT RATE IN %	BRITISH COLUMBIA
URATEE	EI	11053	6	UNEMPLOYMENT RATE IN %	ATLANTIC
URATEO	EI	11055	6	UNEMPLOYMENT RATE IN %	ONTARIO
URATEQ	EI	11054	6	UNEMPLOYMENT RATE IN %	QUEBEC
URATEW	EI	11056	6	UNEMPLOYMENT RATE IN %	PRAIRIES
YDBC	EI	51037	3	PERSONAL DISPOSABLE INCOME	BRIT. COLUMBIA
YDC	EI	19016	3	DISPOSABLE INCOME,MILL.OF CURRENT $	
YDE	EI	51033	3	PERSONAL DISPOSABLE INCOME	ATLANTIC
YDO	EI	51035	3	PERSONAL DISPOSABLE INCOME	ONTARIO
YPC	EI	51042	4	PERSONAL INCOME	B.C.
YPE	EI	51038	4	PERSONAL INCOME	ATLANTIC
YPO	EI	51040	4	PERSONAL INCOME	ONTARIO
YPQ	EI	51039	4	PERSONAL INCOME	QUEBEC
YPW	EI	51041	4	PERSONAL INCOME	PRAIRIES