STUDY NO. I:

OVERVIEW OF THE REGIONAL CANDIDE MODEL

PRELIMINARY DRAFT
(TRANSLATION)

STUDY NO. I:

OVERVIEW OF THE REGIONAL CANDIDE MODEL

Analysis & Liaison Branch
Canada Department of Regional Economic Expansion
June 6, 1974

CONTENTS

An Overview of the Regional CANDIDE[1] Model

1. INTRODUCTION

The concept of regional models includes a wide variety of modelling exercises which differ not only with regard to their objectives but also with regard to their framework and specification. A distinction can be made between urban growth models, provincial and sub-provincial input-output models, etc., and multi-regional models such as CANDIDE-R. In the latter group, a continuum can be established between two extreme kinds of models. On the one hand, there are the models which view a region in the same manner as one would a country, thus applying to each region integrated general equilibrium frameworks open to external demand separated into other regions and the rest of the world. Examples of this king of model would be the Czamanski model of Nova Scotia, the econometric model of the Province of Ontario and the econometric model recently developed for the Province of Quebec[2].

---

1) The abbreviation CANDIDE comes from the words CANadian Disaggregated Inter Departmental Econometric. To date, two versions of the model have been developed: CANDIDE 1.0 and CANDIDE 1.1. The third version, the subject of this presentation, represents a partial regionalization of CANDIDE 1.1., which is why we call it CANDIDE-R.

2) See: - "An Econometric Model of Nova Scotia", S. Czamanski, Institute of Public Affairs, Dalhousie Univ., 1968.
- "An Econometric Model for the Ontario Economy" D. Harnitis (Ontario Economic Review, Special Supplement March 1971).
- An Econometric Model and Macro-economic Optimum for the Province of Quebec. L. Salvas Brinsard, R. Lacroix, G. Bélanger, R. Lévesque, C. Montmarquette, P. Outlas L'Actualité Économique, July-September 1973.

On the other hand, there are the models made to deal specifically with inter-regional flows without, however, giving an overview of the economy of the regions under study: such as the inter-regional input-output models and the multi-regional models centred on the concept of gravity[1]. CANDIDE-R represents a compromise approach to these two extremes: it disaggregates a certain number of national economic variables into five regions to provide us with a framework of feedback "from the nation to the regions" and "from the regions to the nation", and gives us a considerable number of inter-regional functional links, both direct and indirect.

The national versions of CANDIDE use several levels of aggregation which vary considerably from one sector to another. For example, in the CANDIDE Model 1.1., outputs are determined by an input-output system of 63 industries. Labour and wages are determined for a sub-sector of only 12 industries and corporate profits in all sectors are determined by only one equation in Block 19. So, we can say that the process of regionalization is a series of modifications to the levels of aggregation of certain sectors. Here is how Professor T. Matuszewski sums this up:[2]

---

1) "Multiregional I/O Analysis", an article written by W. Leontiel in collaboration with Alan Strout, published by T. Barna in the volume "Structural Interdependence and Economic Development."
2) Report to the Interim Committee on Medium Term Economic Models. T. Matuszewski, January 1970, p. 32-33.

"With regard to the regionalization a partial disaggregation of the Canadian system designed to show the connection between one region and the rest of the economy would seem to be the best thing. Studies would be conducted partly at the regional level, partly at the national level, using the most appropriate geographical frame work; it is not necessary that the level of aggregation which would be more appropriate as an analytical framework be the same as the level at which we are seeking results."

Generally speaking, geographical as well as sectoral framework could vary within a same model. This concept should be viewed as both positive and self-justified for the simple reason that certain economic sectors, such as money, interest rates, the rate of exchange and balance of payments, do in fact require a common ground for analysis on a national scale, while others, such as demography, labour supply, provincial government revenues and provincial government expenditures are best evaluated on a regional level because of their characteristics. From a model efficiency viewpoint, regionalization becomes more than just an exercise in futility. By choosing the appropriate sectors, the model can in fact be made more realistic and more coherent. This was one of our objectives.

2. DESIRED PROPERTIES FOR CANDIDE-R

Our principal objective was to provide a consistent framework for a wide range of impact studies. In the case of

CANDIDE-R, this kind of study consists in comparing two different solutions of the model: the first, a so-called "control" solution and a second containing some a priori modified parameters (or instruments), whose impact on the economy at large we wish to study. Such a study can be made on the estimation period or the forecast period. In other words, the model has been built in such as way that it can explore - ex ante - economic development opportunities which differ in each sector and each region. The process should, basically, help to classify these opportunities according to their relative contribution to:

1. The growth of the Canadian economy as a whole.
2. The resorption of regional disparities in personal income and employment opportunities, the latter being measured in participation rates, unemployment rates and migration rates.

So, these are the main objectives of this model-building exercise, although various subsidiary objectives have also been pursued. CANDIDE-R is, per se, a consistent framework for conducting medium term (12 to 15 years) regional forecasting. It represents a consistent framework for the study of the consequences of the broad trends of the Canadian economy on the regions. Furthermore, it helps to assess the consequences of national economic policies on the regions, either directly or in conjunction with appropriately defined sub-models.

3. CHARACTERISTICS OF CANDIDE-R

At the moment, it would be premature to outline the broad range of the model's potential applications, However, a better understanding of the model can be obtained by examining the main characteristics which were desired from the onset. Considering our understanding of the Department's objectives and programs, considering the data and conceptual limitations which we had to face, we have tried to develop an instrument which would embody the following five features:

3.1 A Comprehensive Model of Economic Activity

It is important that each of the five regions receive equal treatment within the model so that national disaggregated variables can be reconstructed as the sum of regional variables. It is even more important that the model be a "general equilibrium"-type model, as opposed to partial equilibrium models, such as input-output models, which are open to final demand.

It goes without saying that partial equilibrium projections are necessary and irreplacable in very many cases. Hence the need for a large model like CANDIDE-R, which complements such studies by integrating them into the whole of the economic system. Most partial equilibrium model-building exercises are oriented toward questions such as resource availability, comparative advantages, etc, in other words, they tend to be supply oriented. CANDIDE-R must be able to outline under what conditions, the economy of Canada and that of other countries can generate sufficient demand to sustain activities generated by specific regional economic development programs.

3.2 A General Model

For the case at hand, we felt the priority was to develop a model capable of answering, in part, a wide range of questions rather than one which would supply more definitive answers to a small number of problems. CANDIDE-R should, without modifications, be applicable to a large span of specific problems by serving as a bridge between micro-economic studies conducted outside the model and the macro levels of the economy.

3.3 A Disaggregated Model

Since a large portion of the Department's activities deals with specific sectors, industries and regions, this calls for considerations as to structure. Like most econometric models, CANDIDE-R is in no position to modify the economic structure on which it is based. However, a multi-sectoral structural base allows one to carry out impact studies embodying structural modifications to the economy, just as long as these modifications are properly defined a priori. This, we feel, is a definite advantage of disaggregated models.

3.4 A Medium Term Model

Most of the detailed regional data is only available on a yearly basis, with a lag in the release of information of one to two years, which further inhibits developing a short term regional model disaggregated by industry. Furthermore, regional development objectives tend to be prodominantly medium term.

We feel that some considerations to the distinction between short term and medium term models is called for here. The former, such as RDX-II, tend to be highly aggregated and much developed on the final demand side, because short term business cycle fluctuations depend largely on the evolution of the main expenditure categories. In view of their use and their useful horizon (8 to 16 quarters), they are evaluated mainly on the accuracy of the forecasts one draws from them. This is not the case for medium term models, particularly the CANDIDE models. These are comparatively more developed on the supply side: the main concern of a medium term study being the growth potential of the economy, i.e. the quantative, and possibly the qualitative evolution of the factors of production. Particular attention should be devoted the detailed structural elements and the internal consistancy of the model in order to elaborate, not mere forecasts, but realistic "scenarios" of the future. Under the circumstances, the accuracy of the forecasts becomes a secondary consideration for the simple reason that we cannot afford to wait 12 or 15 years to find out about it.

3.5 A Dynamic Model

By and large, the usefulness of a general model as a simulation tool is a function of the number and quality of the interaction between its component variables. In other words, regionalization of CANDIDE is worthless without the specification of feedback mechanisms from the nation to the regions, from the regions to the nation and between the regions. In the light of the model's anticipated use, it was felt essential that CANDIDE-R should be, first and foremost, an econometric model, i.e. it should contain the largest possible number of coefficients which have been empirically estimated on the period of observations. It also embodies some of the characteristics of a simulation model[1]. These components are parameters, which a basic econometric model would generally deal with as endogenous but which, because of difficulties encountered during the elaboration of the model, we must define a priori in order to complete the system.

4. THE PROCESS OF REGIONALIZATION IN CANDIDE- R

A comprehensive disaggregation of the national CANDIDE model into five or ten inter-related regional models of similar characteristics has been excluded, due to a lack of disaggregated regional data and for practical reasons such as software problems, resource limitations, difficulties in interpreting the results, etc.[2]. We felt it preferable to work on partial disaggregation to the extent considered desirable in terms of our original objectives. Therefore, we have made a choice of blocks, variables and feedbacks to be regionalized, at the exclusion of the rest of the national model. For example, consumption, prices and the

1) For a comparative analysis of these two model concepts see:
- "Simulation of Economic Systems", G.H. Orcutt, A.E.R. December 1960
- "Urban Dynamics", J. Forrester, Cambridge MIT Press, 1969

2) The basic idea of CANDIDE is to have an input-output model closed upon final demand through an appropriate number of equations (or exogenous variables) relative to each commodity. Let's say we wanted to disaggregate a model into a series of m regions, each with their own input-output core of n products; we would have to consider the import/export flows not only between the regions and the rest of the world but between the regions themselves as well. This means we would have to consider a possible total of m^2n, different flows a mind-boggling exercise!

external sectors have not been considered for regionalization. We have selected certain sectors where we can presume, implicitely, that economic behaviour is more adequately described at the regional level than on a national scale. Other sectors have been regionalized, mainly because of our original objectives, as well as to produce a minimum number of explanatory variables to be used in the regional equations.

4.1 Regionalization techniques

In the case of CANDIDE-R, disaggregation can be performed in different ways, which is why we feel we should be more explicit about what we call regionalization. Let us consider four possible approaches:

4.1.1 Complete regionalization

This means replacing a given national equation by five corresponding regional equations, the results of which are reaggregated by identities and fed back into the model. This requires that a sifficient number of equations be disaggregated in order to generate meaningful regional explanatory variables. The complete regionalization approach involves a "region to nation and vice-versa" feedback mechanism. If the right-hand side of the equation also includes variables from other regions, the model then involves interregional feedback linkages.

4.1.2 Fixed Share

This involves breaking down national variables with the help of fixed coefficients, the total of which is equal to 1. This is obviously the simplest approach since it doesn't involve any real feedback linkages.

4.1.3 Variable Exogenous Share

This is a way of improving the fixed share technique by designating the ratio of a regional variable over its national counterpact (known as "shares") as exogenous variables. For the forecast period, shares can be projected by extrapolation or, as in the case of other exogenous variables, determined a priori. There are no feedbacks with the rest of the model in this type of regionalization.

4.1.4 Limited Feedback

This technique specifies behavioural equations for the regional shares, thus integrating them into the model framework. In other words, the level of the variable is given by the national equation while the differences between the regions and the national average are endogenized through an additional series of equations. This is called a Limited Feedback technique: there is no feedback from the regions to the nation, yet there is feedback from the nation to the regions and within the region itself.

Most of the sectors in CANDIDE-R have been disaggregated on the basis of the complete regionalization or the limited feedback approaches, the only exception being investment which, for reasons given further on, has been broken down through a variable exogenous shares approach. We have tried to use the complete regionalization approach in the sectors where we felt the regional level better-suited to such economic analysis, from a theoretical point of view. In the remaining sectors, the limited feedback approach let us retain the explanatory power of the national formulation allowing for a minimal feedback at the regional level.

4.2 Estimation Techniques

In a disaggregated econometric model, it is generally assumed that the variables of a same sector are influenced by, if not common, at least corresponding variables. Which means, that the stochastic coefficients represent the specifics of each industry or region. From an estimation view-point, you run into two particular problems when building disaggregated regional models. On the one hand, there is a great deal of cross sectional data available by region and/or sector but on very few observations in most instances. On the other hand, estimation must be simplified by using common specification either in each equation of a same sector or in the equations related to the same industry in each region. One must then extract the most statistical information from the kind of "immitation effect" that link equations of the same type in order to increase the number of degrees of freedom and the stability of the estimated coefficients.

In CANDIDE-R, we have made a frequent use of two Generalized Least Squares techniques which were allowing us to "pool" time series of different equations[1]. Otherwise, the equations were estimated by ordinary least squares, with or without the usual treatment for autocorrelation and the "Almon" distributed lags technique.

5. STRUCTURE OF CANDIDE

One should remember, from the start, that CANDIDE-R is different from CANDIDE 1.1; it is neither a peripheral nor a satellite model of the latter. The regional elements are integrated into the model's base structure in such a way that one could not consider dissociating the two geographical levels for simulation purposes. However, large portions of the national model have been incorporated in full into the CANDIDE-R model and the similarities between them are quite obvious. Just as the 1.1 model is a modification of the 1.0 model, so CANDIDE-R is a modified version of the 1.1 model.

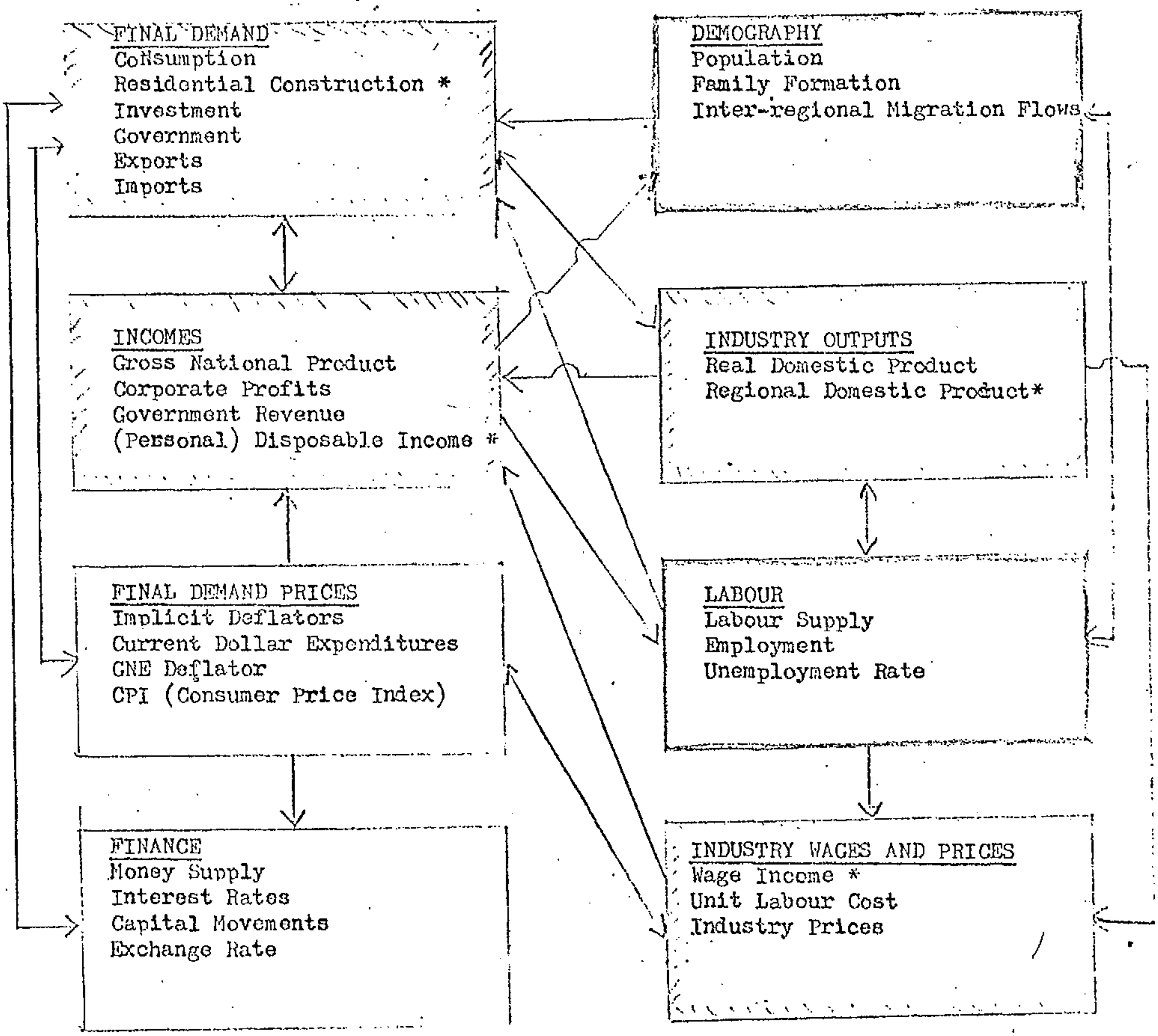
Table 1: Comparaison of the three CANDIDE models

	CANDIDE 1.0	CANDIDE 1.1	CANDIDE-R
Exogenous	377	453	608
Endogenous	1556	2063	2707
Total	1933	2516	3315
Blocks	38	46	52

1) For a more complete description of Generalized Least Squares techniques see:
- "Principle of Econometrics" Henri Theil, University of Chicago, 1971.
- "Elements of Econometrics", Jan Kmenta, Michigan State Univ., 1971.

CHART I

SECTORAL STRUCTURE OF THE CANDIDE-R MODEL



KEY

Complete Regionalization

Regionalization in Sectors *

The regionalization exercise focused on 11 blocks, 5 blocks from the 1.1 model plus 6 new blocks. The regional blocks are:

- Block 3 : Housing Starts
- Block 11: Labour Supply
- Block 12: Employment by Industries
- Block 13: Salaries and Wages in Industries
- Block 22: Demography
- Block 47: Population, by Age/Sex Groups
- Block 48: Population, by Age/Sex Groups and Labour Force Population
- Block 49: Migration Flows
- Block 50: Industry Outputs
- Block 51: Components of Personal Income
- Block 52: Investment by Industries

In this chapter, we try and give a broad outline of each sector as shown in Chart I, with an emphasis on the regional blocks (see sectors A to I in appendix, page 3). So as to give a genuine overview of CANDIDE-R, we will also take a brief look at the non-regionalized sectors. It goes without saying that a lot of the information in the following paragraphs comes from material already published on the national CANDIDE model[1].

1) For a more complete study of the basic structure of the national model see:
   - "AN Overview of CANDIDE Model 1.0", Project Paper No. 1, published by the Economic Council of Canada (M.C. McCraken), and subsequent Project Papers Nos. 2 to 16.

5.1 Final Demand

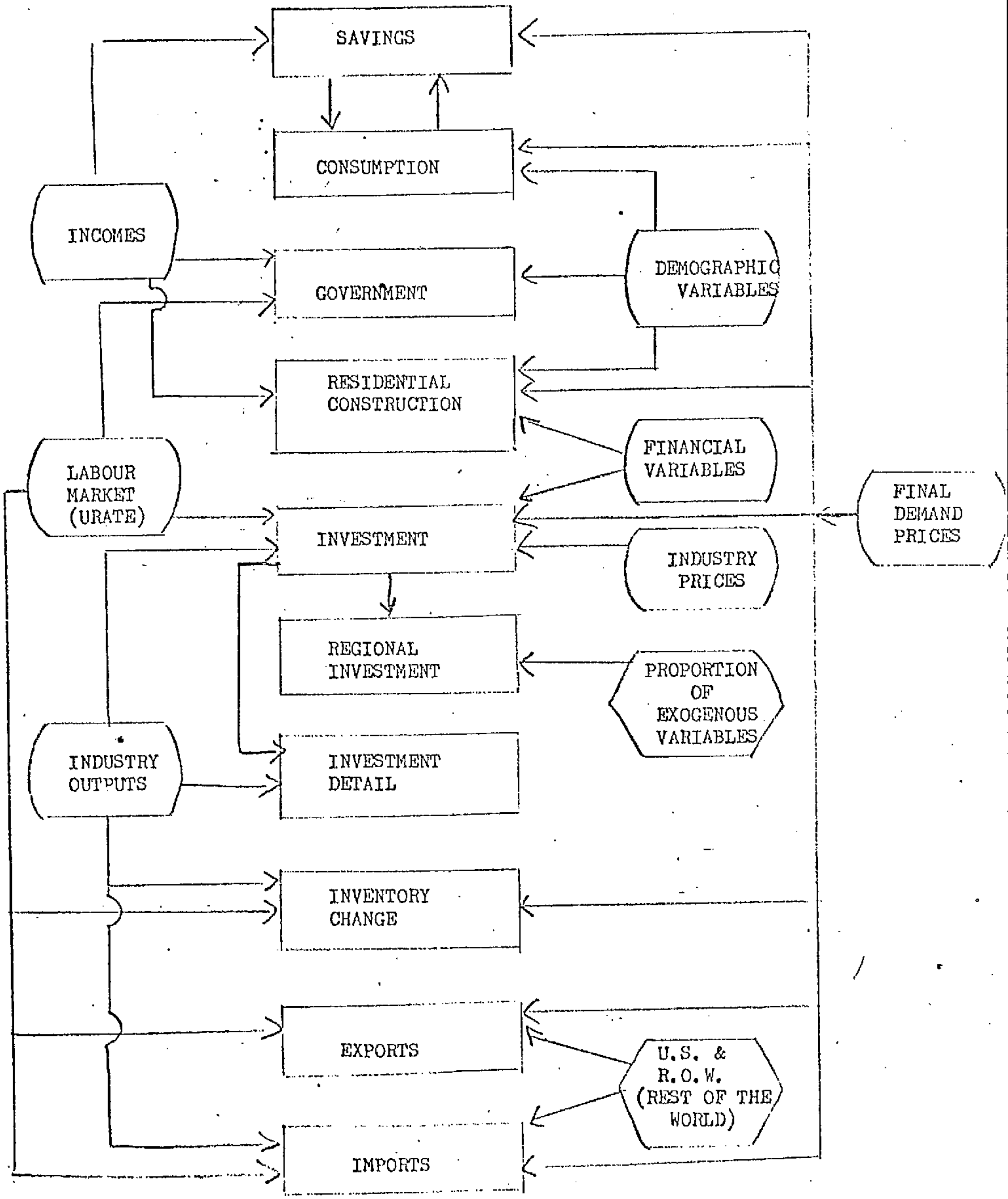
Final demand is probably the most developed sector of the CANDIDE model (see Chart II and Sectors A to I, Page 3, Appendix). It is broken down into expenditures made by persons, businesses, governments and non-residents on all outlays of non-intermediate goods and services. The sector can be broken down, at a fairly low level of aggregation, into: residential construction and personal expenditure on the consumption of goods and services, for the personal sector; fixed capital formation and inventories, for the business sector; current expenditures and gross capital expenditures, for the various levels of government; imports and exports of goods and services, for the non-resident sector. Final demand categories serve to determine the Gross National Expenditure (GNE) in current and constant dollars. It also provides an input to the Government Revenue and Balance-of-Payment blocks. Finally, the separate valuation of each category of goods and services in constant dollars feeds directly into the input-output model in the calculation of disaggregated outputs. With regard to the Final demand Sector, the work of regionalization has been carried out on two blocks only: residential construction and investment by industries.

5.1.1 Housing Starts (Block 3)

In the Residential Construction Sector, the determination of housing starts was completely regionalized (See above 4.1.1: Complete Regionalization). The model has been

CHART II

FINAL DEMAND SECTOR



MT	:	Average of mortgage interest rates
BLT	:	Average of government and corporate long term bond yields
HSS_r	:	Housing starts of single family units
$STHM_r$	:	Stock of multiple dwelling units
u	:	Error term.

The number of completion is calculated by distributed lags on housing starts. The sum of housing starts in each region is subsequently fed back in the model at the national level, to be used as a base for calculating residential construction investment expenditures.

5.1.2 Investment by industry (Block 52)

This block has been added to the model to obtain the information required for the Industry Output Block. At the moment, data such as capital stock, scrappage rates, price of investment goods etc... one needs for the construction of a typical "neo classical" investment model is not available at the regional level. Because of this, the block only contains a series of identities converting variable exogenous shares (See 4.1.3) into regional investment for the following 11 industry groups:

- Agriculture
- Forestry
- Fishing and Trapping
- Mines, Quarries and Oil Wells
- Manufacturing

- Construction
- Transportation, Storage, Communications and Public Utilities
- Trade
- Finance, Insurance and Real Estate
- Public Administration and Defense
- Commercial and non-commercial services

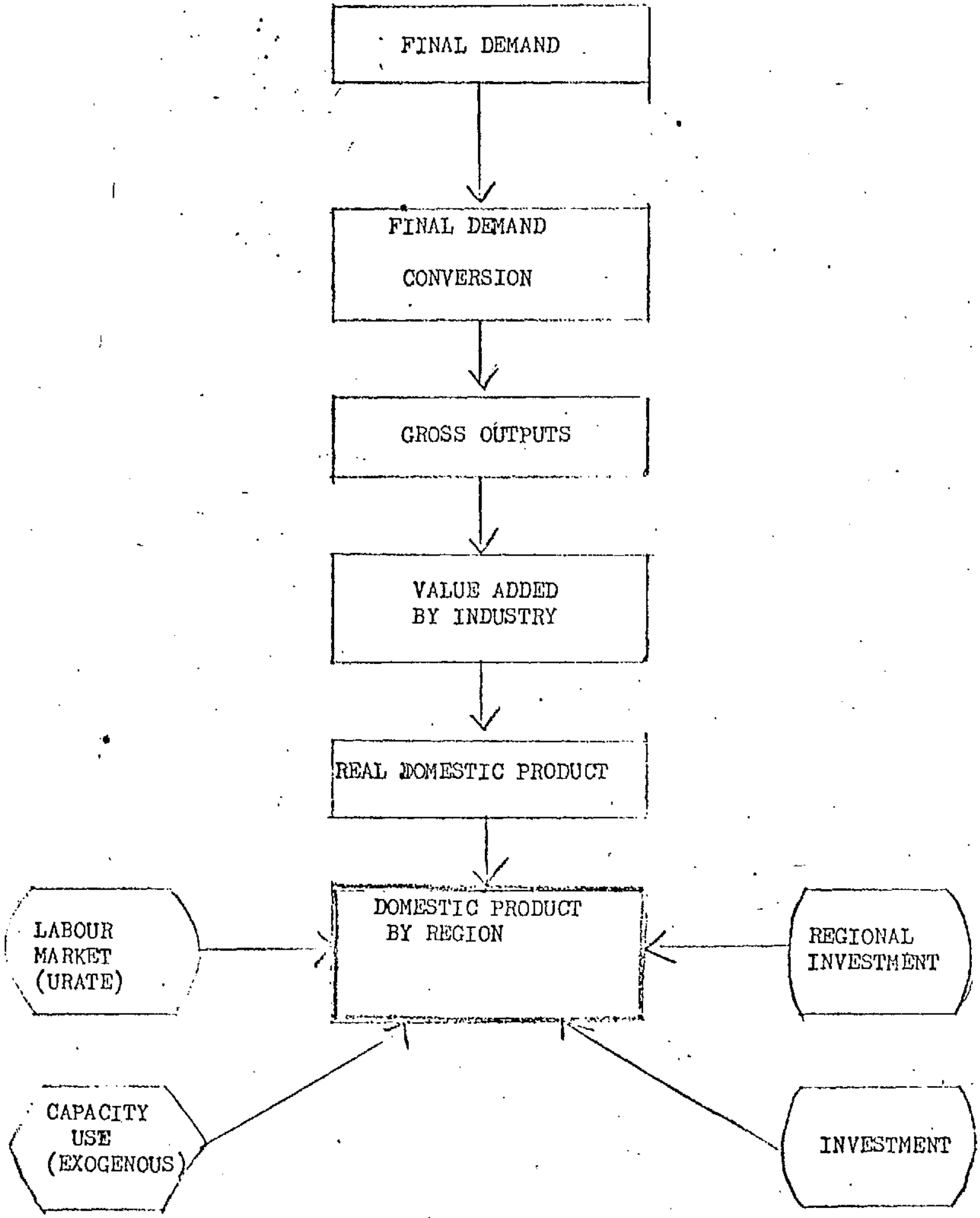
The structure of CANDIDE-R is such that disaggregated industry investment is endogenous at the national level, while its allocation among regions remains exogenous. A specification of this kind creates some problems in the use of the model as a forecasting tool. However, periodical surveys give information on medium term (5 years) investment intentions which can be used as a basis for regional allocation in the early stage of the forecast period. On the other hand, earlier experiences with the national model suggest that most impact studies require ad hoc adjustments to both autonomous investment and its regional allocation. In these cases, exogenous shares can be considered as genuine instruments for simulation purposes.

5.2 Industry Outputs

The determination of disaggregated industry outputs is without a doubt the main distinguishing feature of the CANDIDE model. First, the categories of expenditure, as calculated in

C H A R T III

INDUSTRY OUTPUTS SECTOR



the preceding sector, become a series of products compatible with the 1961[1] input-output model on which the sector is based. (See Chart III and Sector J, page 3, Appendix). Since this is a rectangular model, final demand by industries must also be calculated with the aid of another matrix which allocates commodities by industries; this is known as the market share matrix. Gross output is then calculated by using an inverted input-output matrix in the usual way $(I-DB)^{-1}$. Value added by industry is determined as a constant share of gross output which includes intermediary as well as final goods. Finally, real domestic product is reached through a series of adjustment equations fitted between industry outputs (calculated in the input-output model) and the level of output which has been observed during the estimation period.

5.2.1 Regional Industry Output (Block 50)

First we had to construct time series on regional output for the 11 types of industry mentioned in 5.1.2. Data for CANDIDE-R has been constructed from information on the components of Gross Domestic Product (current dollars) from Statistics Canada. In calculating the series in constant dollars, we have used the GDP deflators already present in CANDIDE 1.1 as common deflators for each region.

1) We hope to be able to use a 1966 input-output model in the near future.

We felt it was by far better to sacrifice some feedback linkages in order to keep intact the input/output core of CANDIDE. On the other hand, it would seem desirable to make use of a lot of the indirect effects which goes from output to employment and incomes. We have had to limit output regionalization because it is presently impossible to regionalize the input-output model in relation to final demand, i.e. with regard to regional consumption as well as inter-regional flows of goods and services. The variables of the allocation mechanisms do not then act as factors of production. They serve the purpose rather of assessing comparative regional advantages which are ultimately reflexted into cumulative output capacities and their rate of utilization.

Industry outputs have been regionalized following a limited feedback formulation (See 4.1.3). First, the disaggregated outputs are calculated at the national level, using the input-output model described above. Secondly, stochastic share equations are used to distribute, endogenously, industry outputs among the regions. However, the resulting industry outputs do not allow us to use the reaggregation process to carry out the direct impact of regional disparity on national industry outputs. This is because we used a share formulations which, be definition distribute the given national totals.

We formulate that the desired level of relative output $(X_{ij}/X_i)^*$ is a function of relative rates of utilization of accumulated capacity approximated by the cumulative sum of investment of the past four years.

$$\left(\frac{X_{ij}}{X_i}\right)^* = \alpha + \beta\left(\frac{T_{ij} \cdot C_{ij}}{T_i \cdot C_i}\right) + \mu_{ij} \quad \text{..............(1)}$$

where, X_{ij} : Output of the industry i in the region j

X_i : National outputs of the industry i

T_{ij} : Rates of utilization of accumulated capacity for industry i in region j. To get these rates, divide the known output by the potential output (which is obtained by intrapolating identified output peaks);

C_{ij} : Cumulated sum of investment in industry i, region j, during the past four years.

Then we add the partial adjustment mechanism:

$$\left(\frac{X_{ij}}{X_i}\right) - \left(\frac{X_{ij}}{X_i}\right)_{-1} = \gamma\left[\left(\frac{X_{ij}}{X_i}\right)^* - \left(\frac{X_{ij}}{X_i}\right)_{-1}\right] \quad \text{....(2)}$$

Equation's (1) final formulation:

$$\left(\frac{X_{ij}}{X_i}\right) = \alpha\gamma + \beta\gamma\left(\frac{X_{ij}}{X_i}\right)_{-1} + (1-\gamma)\left(\frac{T_{ij} \cdot C_{ij}}{T_i \cdot C_{ij}}\right) + \gamma\mu_{ij} \quad \text{.....(3)}$$

The regional allocation of output mechanism, as described in equation (3), is advantageous both because of its simplicity and because it implies a direct link between investment and output. Its simplicity helps avoid the pitfall one encounters when using estimated, as opposed to observed regional output data. Finally, because if makes output a function of investment, one can evaluate the impact of regional investment redistribution policies.

Equation (3) has been estimated with Zellner's Generalized Least Squares technique[1] for the above-mentioned five regions and eleven industries. The only changes made during estimation were to replace rates of utilization in some industries by $(1-U_j)$ and to omit the potential variable.

Results for Quebec manufacturing are:

$$MAYQ = MAY \left[\underset{(1.5)}{0.0446} + \underset{(4.9)}{0.7061} \; (MAYQ/MAY)_{-1} + \underset{(2.6)}{0.1619} \left(\frac{(1 - URATEQ)\; C_{ij}}{(1 - URATE)\; C_i} \right) \right]$$

MAYQ (M.C.G.)

$R^2 = 99.8$

SEE = 28.47

MAYQ/MAY (M.C.O)

$R^2 = 94.1$

SEE = 0.0023

D.W = 1.7

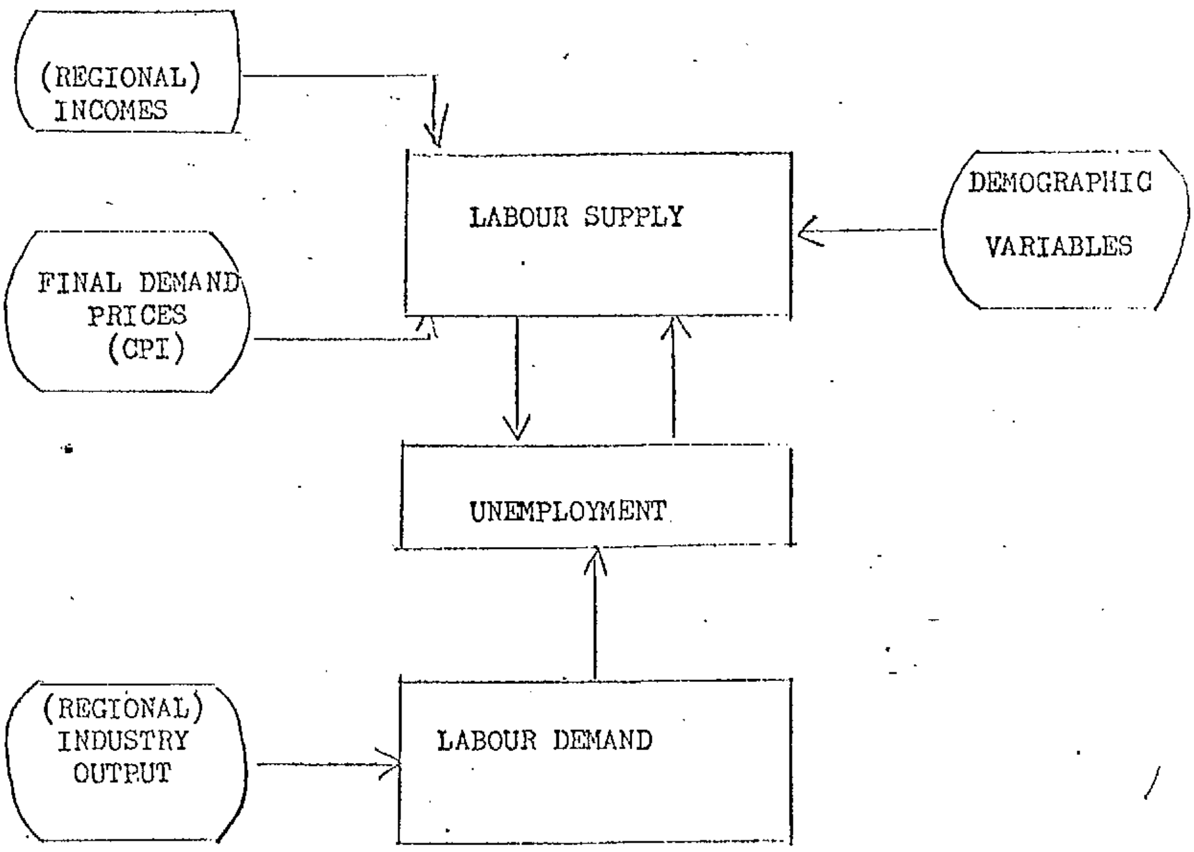
5.3 <u>Labour Market</u>

The CANDIDE-R model gives a complete regionalization of the Labour Market Sector (See Chart IV and Sectors K and L, Page 3, Appendix). The civilian labour force population is computed in the Demographic Sector. The following age groups have been isolated in most regions:

1) "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", A. Zellner, J.A.S.A. (1962) page 348 to 368.

CHART IV

LABOUR SECTOR



1 - Males, 14-24

2 - Males, 25-54

3 - Males, 55 and over

4 - Women, 14-24

5 - Women, 25 and over

To calculate the Regional Labour Force, we used behavioural equations of the participation rate in four of these five groupings, participation rates for the second group being exogenous as in the 1.1 model. Employment is also estimated through regional labour demand equations. So, the unemployment rate for each of the five regions is calculated residually from the difference between labour supply and labour demand. The national unemployment rate is entirely a function of regional variables.

5.3.1 Labour Supply (Block 11)

The participation rate of the secondary group was completely regionalized through a series of 16 stochastic equations disaggregating labour force into three of four age groupings per region. The specification of these equations is, in essence, similar to the national equations: the participation rate is a function of a distributed lag for earned income per person and of a short term adjustment variable such as the unemployment rate or the activity rate of a particular region. Here is an example of a typical participation rate equation:

$$PROFW = -1.3036 + 0.0035 \sum_{i=0}^{4} \left(\frac{1}{YEPPRW}\right)_{t-i} + 0.8202 \left(\frac{TEETW}{LFPW}\right)_{t-1}$$

S.E.E. 0.0064; $\bar{R}^2$ 0.9898; D.W 2.26

PROFW : Participation rate of women, 25 and over, in the Prairies region

YEPPRW : Real earned income per person, Prairies

TEETW : Total employment, Prairies

LFPW : Labour force population, Prairies

The most particular thing about the Labour Sector in CANDIDE-R is that the labour supply adjustment mechanism embodies fluctuations in unemployment, participation rates and of migration flows at the same time. A case in point is excess supply of labour in a given region. The lack of job opportunities will probably be reflected in a sudden rise of unemployment, unemployment being, for obvious reasons, the most flexible short term adjustment mechanism there is. Experience has shown us, however, that unemployment cannot increase indefinitely and that beyond a certain level or length of time some workers from the secondary group (and sometimes even from the primary group) retire from the labour force, thus creating a drop in the participation rate. If the job situation is comparatively better in other regions, additional adjustment will be carried out through migration flows as the labour force moves to the more prosperous regions. The relative importance of these three types of adjustment can vary in CANDIDE-R since they are an integral part of the model's structure.

5.3.2 Labour Demand (Block 12)

Block 12, Labour Demand, gives a complete regionalization in the above-mentioned eleven sectors of industry. Although we have gotten away from the corresponding block in CANDIDE 1.1, we have not set aside its original objectives; we have merely converted them into regional objectives.

For purposes of estimating regional unemployment rates, both the labour supply and the total employment of a region are required. Employment estimates are also required for calculating total labour income, one of the main components of personal income. After aggregating employment to each industry, the overall national employment can be related to a measure of aggregate labour supply in order to estimate the national rate of unemployment which is used elsewhere in the model.

The specification of the regional labour demand equations is based on the concept of effective demand as related to output. This approach is different from the one used in the CANDIDE 1.1 original block (where capital stock plays a major role). Since there is no data available on regional capital stocks, we decided to relate labour demand to output, which in turn is partially related to relative accumulated capacity in the form of utilized stock of capital. This means we can assume that 'ex ante' national investment decisions are made in keeping with standard, "neo-classical" analysis, while 'ex post' regional employment depends on the respective output of each region.

Despite these general postulations we were unable to isolate from the production all the elements which could explain either the short term fluctuations or the long term trends of each employment by industry. Which is why the specification we have adopted may include a time trend (for measuring the impact of technological change) as well as the squared output and/or the lagged output which would enable us to better grasp the true nature of each industry.

Based on Labour Force Survey data, the equations were estimated by using Zellner's Generalized Least Squares technique. Here is the basic specification:

$$E_{ij} = \alpha_o + \alpha_1 X_{ij} + \alpha_2 T + \mu_1 \quad \ldots\ldots\ldots\ldots (1)$$

$$E_{ij} = \beta_o + \beta_1 X_{ij} + \beta_2 X_{ij}^2 + \beta_3 T + \mu_2 \ldots (2)$$

$$E_{ij} = \gamma_o + \gamma_1 (X_{ij})_{-1} + \gamma_2 T + \mu_3 \quad \ldots\ldots\ldots (3)$$

$$E_{ij} = \eta_o + \eta_o X_{ij} + \eta_1 (X_{ij})_{-1} + \mu_4 \quad \ldots\ldots\ldots (4)$$

where, E_{ij}: Employment in industry i, region j

ij: Output in industry i, region j

T : Time

i : Error term

Here are the results for the manufacturing industry in the Province of Ontario:

$$MAETO = 1020.85 + 0.1913 \; MAYO - 0.00000801 \; (MAYO)^2$$
$$\quad (3.34) \qquad (9.88) \qquad (-4.30)$$

$$-17.0298 \; TIME$$
$$(-3.29)$$

$\bar{R}^2 = 0.9829$

S.E.E. = 8.4761

D.W. = 2.79

5.4 Industry Wages and Prices

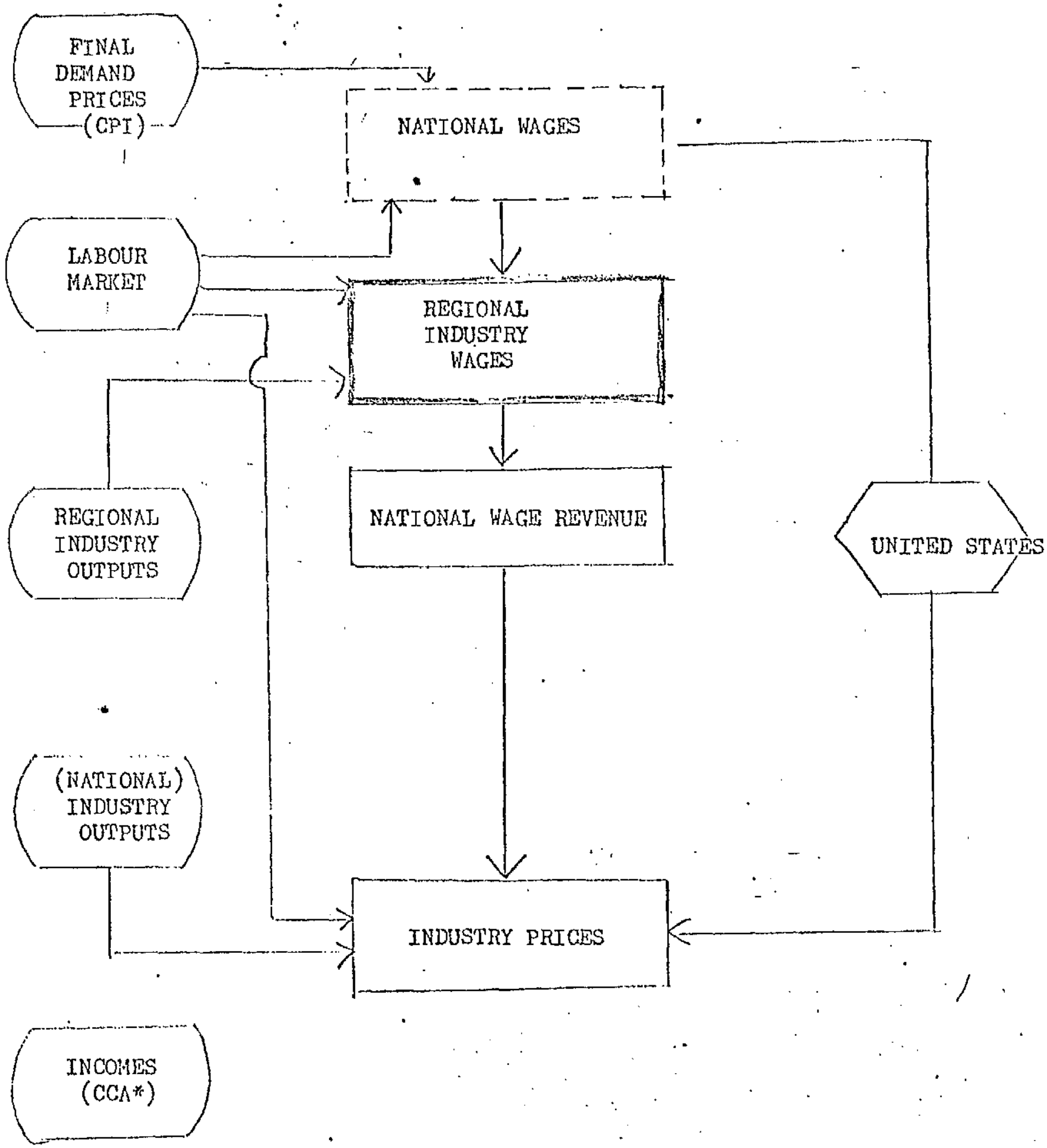
The determination of wage-rates at the national level follows a Phillips curve approach. Labour income is essentially a function of the Consumer Price Index and of an "inverted" unemployment rate. Industry prices, which are really implicit value-added price deflators, are used to evaluate current dollar industrial production and aggregated Gross Domestic Product which is itself used to calculate the current dollar Gross National Product. The typical specification for these price equations involve an estimate for unit labour cost per man-hour. Which is why Block 12 (Employment by Industries) had to keep a series of 13 equations from the 1.1 model which determine the total number of employed man-hours in each industry. U.S. price and wage variables were used in the equations when it was deemed appropriate (See Chart V and Sectors M and N, page 4, appendix).

5.4.1 Regional wages and salaries (Block 13)

Salaries and wages have been regionalized following a specification which involves both limited feedback and

CHART V

INDUSTRY WAGES AND INDUSTRY PRICES SECTOR



* Capital consumption allowances, National Accounts total.

complete regionalization. The regional block's objectives are the same as those of the Wages Block in the national model. It is used to estimate the major component of personal income and to calculate unit labour costs which are incorporated into the Industry Prices Block to explain implicit value-added deflators. This double objective relies on regional mechanisms which determine salaries. These mechanisms were elaborated outside of the national model and then incorporated into CANDIDE according to the widespread theory that the disparity is in itself, the appropriate variable, within the framework of a national study, for keeping track of regional differences.

To sum up, the original CANDIDE equations are used to estimate national averages in the first stage; in the second stage, regional mechanisms relate regional wages and salaries to the corresponding national average and regional differences in labour productivity; the third and final stage reaggregates the regional wages and salaries and substitutes them for those of the first stage. In other words, the process of reaggregation is actually utilized to convey the impact of regional disparity on the estimation of national wages and salaries which are fed into the overall model.

The creation of regional salary estimation mechanisms relies on the specification of a model of selective impulse[1].

1) To be published shortly.

This type of model allows us to complement the original Phillips curve approach with a framework more appropriate for long term labour income evolution. Our three-stage process not only recognizes the usefulness of the national equations in relation with short term "business cycle" types of fluctuations, it also allows regional equations to carry the influence of those variables which are more closely related to long term increases in wage rates.

The model recognizes the importance of productivity in a long term estimation of labour income[1], and retains relative labour productivity as the first impulse on regional wages (i.e. a shift in demand curves for labour curves). A second impulse is a result of real wage fluctuations at the national level which induce similar shifts in the regional labour supply curves. Finally, a third variable brings a notion of disequilibrium related to the unemployment rate into the estimation of the impulse coefficients.

This is what the equation for wages in all regions looks like:

$$W_{ij} = A \; U_j^{B_o} \; P_{ij}^{B_1} \; W_i^{*B_2} \; u_{ij}$$

1) See Kuh, E., "A productivity Theory of Wages", R.E. Studies, October 1967.

where, W_{ij} : Wages per man-year paid by the industry i in region j (in $)

A : Constant term

U_j : Unemployment in region j

P_{ij} : Ratio of labour productivity, industry i and region j to the national labour productivity, industry i

W^*_i : Wages per man-year in industry i (national) divided by the Consumer Price Index

μ_{ij} : Error term

We used annual time series for the period 1961-71. To help us overcome limitations due to the lack of observations and to get additional statistical information from the inter-regional dependancy which can be found in the numerous variables present in the error term, we have estimated the equation (logarythmic formulation) simultaneously for each industry and all the regions, using Zellner's Generalized Least Squares technique.

The following are the results for the Quebec manufacturing industry:

$$WSMAQ = Exp \Big[\underset{(57.3)}{7.915} + \underset{(3.3)}{0.047} \ln (URATEQ) + \underset{(8.0)}{0.853} \ln \left\{ \frac{(MAYQ / MAETQ)}{(MAY / MAET)} \right\} + \underset{(46.3)}{2.08} \ln (WSMAT/MAET/CPI) \Big] MAETQ$$

$\bar{R}^2$ = 99.5

SEE = 0.092

D.W. = 1.8

M.C.G. (1961, 1971)

5.5 Final Demand Prices

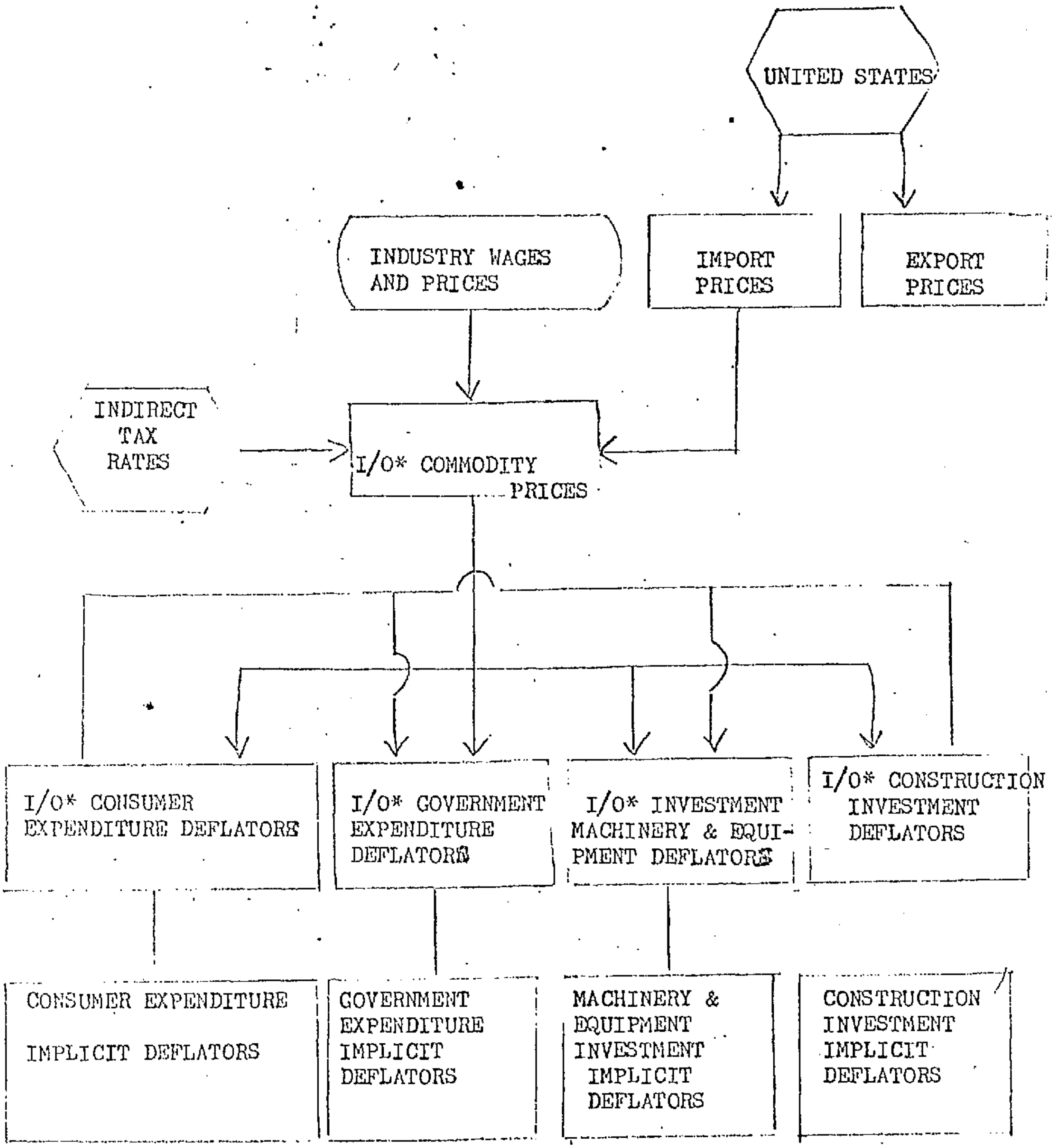
Because of its many equations, the Final Demand Prices Sector is the largest and without a doubt, the most extensive sector of the CANDIDE model. These prices are used to convert final demand categories into current dollars. They serve as measures of price effects in a lot of demand equations. One of its resulting aggregates, the Consumer Price Index, influences a great number of national as well as regional equations (See Chart VI and Sectors O to R, page 4, Appendix).

Without a doubt the most interesting aspect of the Prices Sector in CANDIDE is the way in which the model converts industry prices into final demand prices. What the model does is use the input/output table to produce a complex weighting pattern for price changes involving the use of commodities as intermediary inputs, cost of raw materials and import prices as well as indirect tax rates at each stage of the production. In other words, the model adjusts prices at whatever level there is value added.

It would have been impossible to regionalize the Final Demand Prices Sector without, on the other hand, regionalizing the national model's core, its input-output tables, a possibility which we have already excluded. Yet it is not altogether certain that price determination is truly regional enough to justify the amount of work involved in a geographical disaggregation of the kind. The little information available on prices per region tends to prove that only the housing sector shows clear medium term regional differences. Which is why construction cost variables in Block 3 were exogenous.

CHART VI

FINAL DEMAND PRICES SECTOR



* Input-output.

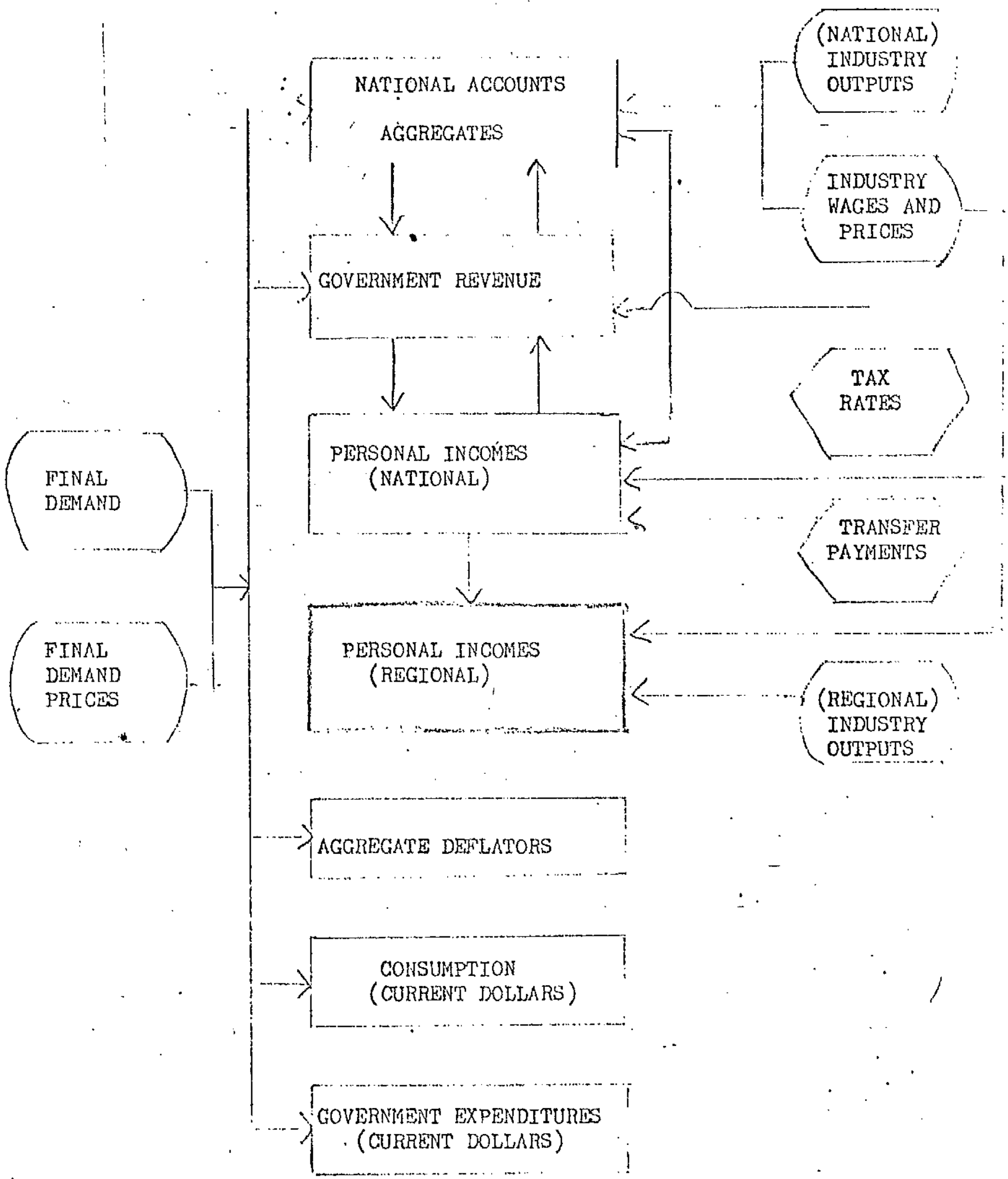
We assume, in the CANDIDE-R model, that prices are determined exclusively at the national level, which in no way excludes their use in regional equations.

5.6 Personal Income and Government Revenue

This sector of the CANDIDE model determines government revenue, personal income and personal disposable income. The present CANDIDE-R model does not regionalize government revenue but it does completely regionalize personal income. (See Chart VII and Sectors S and T, page 4, Appendix). Personal income variables are involved in a great number of regional as well as national equations and are used, along with government revenue and other forms of revenue, to determine national income components. Tax revenues are generally obtained by applying a tax rate to a tax base for each separate direct and indirect taxes. CANDIDE 1.1 contains separate equations explaining both tax revenue of the federal government and that of other levels of government (this also applies to the Government Expenditures Sector). Lastly, national income is divided into wages and net property income. Various taxes and redistributions of income are then introduced, after which personal disposable income and retained corporate earnings are calculated.

C H A R T VII

INCOMES SECTOR



5.6.1 Personal Income (Block 51)

Personal income is one of the key variables of CANDIDE-R because it influences the regional equations pertaining to participation rates, migrations flows and residential construction and a great many national demand equations as well. Wages and salaries the main component of personal income, have been regionalized in Block 13 as described in 5.4.1. To fully regionalize personal income, other equations pertaining to its other components must be brought in. Three of these items have been completely regionalized: accrued net income of non-farm unincorporated business (including rents) as well as interest, dividends and miscellaneous investment income. Output, wages, prices and occasionally interest rates are the main explanatory variables in these equations. Here is an example:

$$AGUNY_r = B_o + B_1\ AGY_r + B_2\ AGP + B_3\ WSAG_r$$

$AGUNY_r$: Accrued net income from farm production in region r

AGY_r : Gross domestic product in agriculture (in constant dollars) in region r

AGP : Agriculture implicit price deflator, 1961=1.0

$WASAG_r$: Wages and salaries in agriculture, region r

The other items of personal income, military allowances, government transfers to persons and indirect taxes, have been regionalized by a variable exogenous share approach (See 4.1.3).

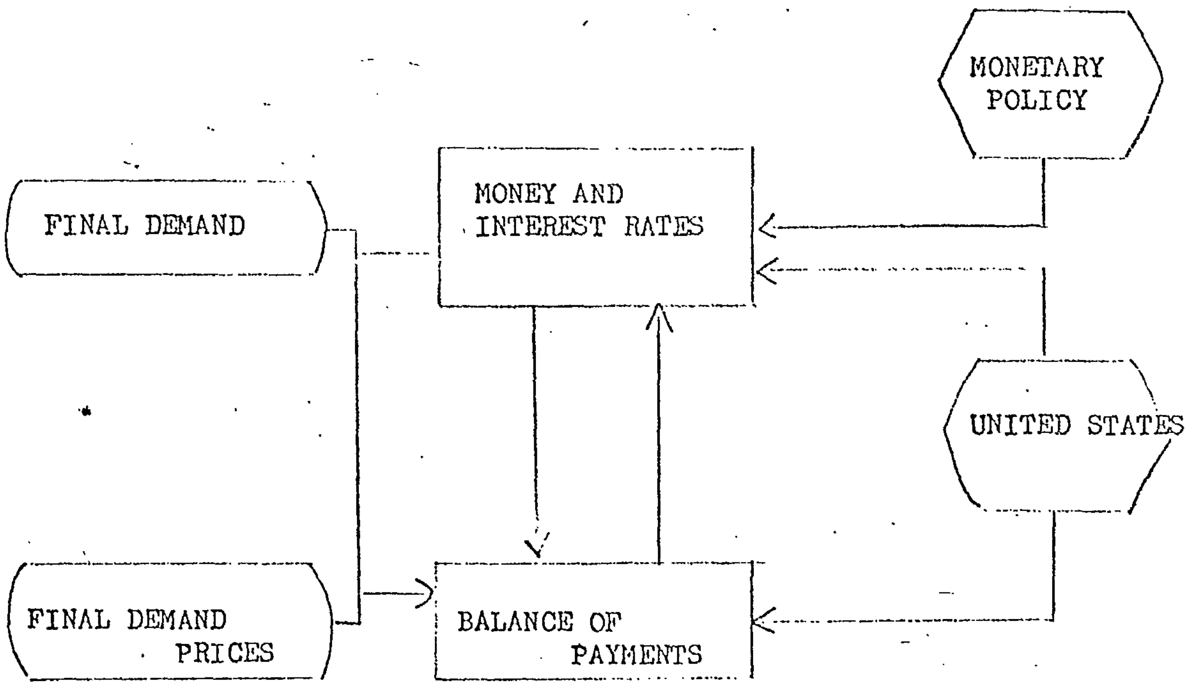
5.7 Finance and Balance of Payments

We should mention that the study of capital flows and balance of payments in CANDIDE is relatively simpler than some other sectors (See Chart VIII and Sectors U and V, page 4, Appendix). On the financial side, money supply is exogenous to the model, while demand for money is related to interest rates; these rates in turn affect certain components of final demand, particularly corporate investment and residential construction. In the Balance-of-Payments Sector, international capital flows react, more or less residually, by contraction or expansion without the intervention of a corrective feedback mechanism between them and the rest of the model. For all intents and purposes, the rate of exchange should be considered as an exogenous variable. However, disaggregated international prices have been specified in American dollars and are converted into Canadian dollars when multiplied by the appropriate exchange rate variable ($CAN/$US). This formulation makes it possible to assess directly the impact of exchange rate movements right down to the model's lowest echelons of aggregation.

The Finance and Balance of Payments Sector did not call for a regionalization exercise.

C H A R T VIII

FINANCIAL SECTOR



5.8 Demography

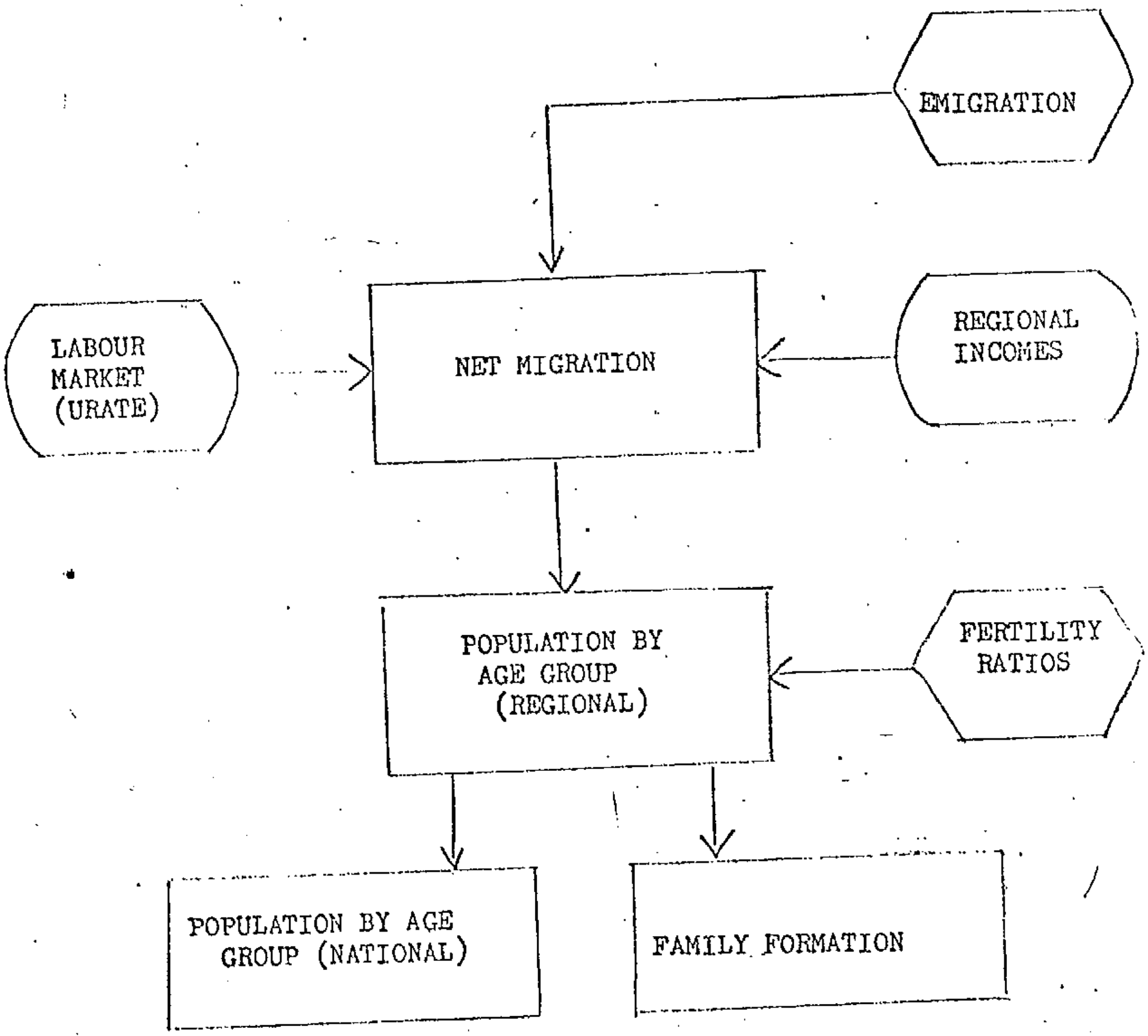
Population and Migration Flows are the subjects of the most elaborate regionalized sector in CANDIDE-R. The sector contains 298 equations in 4 of the model's 52 blocks (See Chart IX and Sector W, page 4, Appendix). The calculation of the main demographic aggregates has been completely regionalized and the total canadian population in CANDIDE-R is obtained by summing regional population figures. The system can be described as a three stage process. In the first stage, net regional migration flows are calculated using stochastic equations of gross flows for each region. Then population, broken down into age-sex groups, is determined through a population projection algorithm. In the third and last stage, the model's necessary aggregates are totalled and fed back into the other sectors. Disaggregated population is used to determine labour force population, for the labour supply block and to calculate the number of families per region for the residential construction block. Net migration flows are also related to family migration flows and the family formation mechanism as well. Population also figures in a lot of regional as well as national equations where variables are on a "per capita" basis.

5.8.1 Migration Flows (Block 49)

The model gives a complete regionalization of the four gross flows which figure in the calculation of total net migration per region as utlized in the population algorythm. The process is explained in the following example, which applies to the Atlantic region;

C H A R T IX

DEMOGRAPHIC SECTOR



$$TNME = MINERT - MOERT + MINEXT - MOEXT \quad (1)$$

$$MINERT = MINERQ + MINERO + MINERW + MINERC \quad (2)$$

$$MOERT = MINQRE + MINORE + MINWRE + MINCRE \quad (3)$$

TNME : Total net migration in Atlantic region

MINERT:: Total inflows from other regions of Canada

MOERT : Total outflows from Atlantic to other regions of Canada

MINEXT : Total emigration from Atlantic region

MINERQ : Immigration to Atlantic region from Quebec. (The fourth letter of the mnemonic code stands for the receiving region, the sixth represents the sending region.)

Each of the right-hand variables of equations (2) and (3) and MINEXT in (1) are determined through stochastic equations. MOEXT in (1) is left exogenous to the model. The Stochastic equations have been formulated to keep account of the comparative economic situation in the sending region (push factors) as well as the receiving region (pull factors). Income variables, together with variables measuring short term employment opportunities, such as unemployment rate, can be found in most equations. The following example illustrates gross immigration per receiving region, in this case the Atlantic region:

$$\frac{MINEXT}{POPE<-1>} = \underset{[3.84]}{0.02092} + \underset{[6.19]}{0.00959}\left(\frac{1.0}{URATEE<-1>}\right) + \sum_{i=1}^{3} \beta_i \left(\frac{YPE}{POPE}\right)_{t-i}$$

$$-0.00385 \quad (0.75 * OIP<-1> + 0.25 * USIP<-1>) \; +$$
$$[-1.93]$$

$$0.00789 \left(\frac{(MAY + MIY + UTY)<-1>}{11547.40381} \right) + \sum_{i=1}^{3} C_i \; LFPE_{t-i}$$
$$[5.36]$$

$$+ \rho \, (e_{t-1})$$

$\bar{R}^2 = 0.9131$

SEE = 0.00018

DW = 2.08

OLSQ (1952-1971)

$\rho = -0.39382$

R.E.P., le Degré, $B_3 = 0$

$B_0 = 0.000784$ [2.09]
$B_1 = 0.000523$ [2.09]
$B_2 = 0.000261$ [2.09]

R.E.P., le Degré, $C_3 = 0_Z$

$C_0 = 0.0001104$ [-3.76]
$C_1 = 0.00000736$ [-3.76]
$C_2 = 0.00000368$ [-3.76]

Due to the highly volatile nature of inter-regional migration flows we were faced with the problem of estimating robust coefficients from ten observations only without seriously deteriorating the accuracy of the predicted values. The estimation results obtained from a Generalized Least Squares technique suggested by J. Kmenta were satisfactory. (See Note 1. page 12). This technique allows us to estimate coefficients common to each internal migration flow, destined for the same region, by pooling information from both time series and cross section data.

Here's an example of this in the case of the internal migration flow from Ontario to Quebec:

$$\frac{MINQRO}{POPQ^{<-1>}} = \underset{[22.61]}{0.00001985}\ (320.0) + \underset{[1.27]}{0.0002814} \left(\frac{YPQ}{POPQ}\right)^{<-1>}$$

$$-\underset{[-2.31]}{0.000551} \left(\frac{YPO}{POPO}\right)^{<-1>} - \underset{[-2.56]}{0.0000783}\ URATEQ^{<-1>}$$

$$+ \underset{[1.60]}{0.0000382}\ URATEO^{<-1>} \qquad pe_{t-1}$$

$$p = 0.5261347$$

5.8.2 Population (Blocks 47 and 48)

For each region, the model makes a separate calculation of the population for a breakdown of 14 age groupings per sex, consisting of a cohort group for every 5 years of age with an open category for those 65 years and over. Population growth is determined by its usual three components, births, deaths, and net migration flows. The algorythm works according to a standard demographic projection model wherein each of the cohort groups goes through an aging-attrition process every five years. The number of births is computed using a set of five exogenous fertility ratios (one per region) applied to the group of women 15 to 49 years. Net migrants are allotted to the appropriate age groups by a matrix of fix coefficients based on the structure observed from 1966 to 1969. The model gives five-year projections for each year and proceeds iteratively on the calculated values. In other words, the population for 1969 gives the population for 1974 which in turn gives the population for 1979; the population for 1970 gives that of 1975 and so on. Most of the 140 equations are as follows:

$$POP_{i,j,t} = S_{i,j} \left(POP_{i-1,j,t-5} \right) + \sum_{n=o}^{4} W_{ij,t-n} \quad M_{j,t-n}$$

$POP_{ij,t}$	:	Population of group i in region j, at time t
S_{ij}	:	Five-year survival rate for group i in region j
$W_{i-1,j,t-n}$	:	Proportion of net migrants, in region j, in the year t-n, who will be the same age as POP_{ij} in t
$M_{j,t-n}$	:	Net migration in region j during t-n

Block 47 also contains 26 identities to compute the total population by region and the 21 groupings of the labour force population, as utilized in Block 11. The latter are calculated using exogenous adjustments for the non-civilian population[1].

5.8.3 Family Formation Mechanism (Block 22)

In CANDIDE-R, the family formation mechanism, used as an input for the housing starts equations, has been completely regionalized. We start with the following identities for each region:

1) The two population blocks have been used on several occasions as a separate model for supplying population projections, some of which up to the year 2041.

$$NTFAM_r = MAR_r + NIMFE_r - DEATH_r - DIVOR_r \quad (1)$$

$$FAM_r = FAM_r <-1> + NTFAM_r \quad (2)$$

where, $NETFAM_r$: Net family formation in region r

MAR_r : Number of marriages

$NIMFE_r$: Net family migration

$DEATH_r$: Number of deaths of married persons

$DIVOR_r$: Number of divorces

FAM_r : Number of families

The five divorce variables are introduced exogenously into the model while the number of deaths of married persons is endogenized through an identity which assumes an exogenously determined ratio for the number of deaths over total population. The model can calculate the number of marriages by using national marriage rates with regional age-sex population groups. The marriage rates are then realigned by multiplying them by a different constant for each region. Finally, net family migration flows are endogenized by pooled time series-cross sections equations (See 5.8.1) tying in these variables with net inter-regional and international movements of persons in each region.

The number of families obtained in equation (2) is multiplied by the observed ratio household/family in order to calculate the number of family households. Non-family households is also enter into the block by way of an identity consisting of the product of the population and the ratio of non-family households to the population.

6. CONCLUSION

The development of CANDIDE-R already represents 9 man-years of research during a period of two and a half years[1]. The estimation of the 11 regional blocks described in Chapter 5 is now completed, including their integration into a separate, machine-readable model[2]. Up until now, the model has been simulated over the estimation period and the results so far are quite encouraging. The next step will involve assembling a set of exogenous variables up to the year 1985 in order to produce an acceptable control solution for the forecast period. The model's complete documentation is already underway but there is still a considerable amount of work left to be done (See Appensix page 5). The development of subroutines for generating easily readable reports will also be looked into in the near future.

The regionalization of revenue and expenditures in the junior, i.e. provincial and municipal government levels (Blocks 6 and 19of the CANDIDE-R and 1.1 models) has also been considered

1) The principal authors of CANDIDE-R are G. Fortin and G. Simard of DREE, Professor A. D'Amours, Director of the Department of Economics at Sherbrooke University. The following also participated: (Ms.) S. Simard, S. Hébert, R. Corbeil and J.P. Rioux also of DREE, and J. Hodgson who is now with the Department of Finance. The main writer of this overview was G. Fortin.

2) This integration was done under the supervision of G. Simard.

lately. A self-evident choice, since the main expenditure and revenue sectors have already been divided up into federal (on the one hand) and provincial/municipal (on the other hand) in the CANDIDE 1.1 model. The next step, in the case of the latter, will be to replace aggregated equations by genuine provincial equations, using as explanatory variables the regional variables which have already been endogenized in CANDIDE-R. The work on these blocks is presently at the data-gathering stage[1].

1) See: "A proposed Outline for the Regionalization of the Junior Government Revenue and Expenditure Sectors", G. Fortin, May 28, 1973 (report for internal circulation, DREE).

APPENDIX

REGIONAL CANDIDE

As of February 28, 1974

Total number of Variables	3315
Total number of Exogenous Variables	608
Total number of Endogenous Variables	2707
Regional Variables	803
- Exogenous	163
- Endogenous	640
National Variables	2512
- Exogenous	445
- Endogenous	2067

Types of Exogenous Variables

Type	Description	Number of Variables
I	Dummies (including Time Trends)	50
II	Foreign Trade Variables	
	1. Exports, Flows and Adjustment Items	27
	2. Foreign Trade Prices	58
	3. Imports (Fuel Flows and Adjustment Items)	11
III	International Transactions (Incl. Exchange Rate)	8
IV	Capital Stock Characteristics (Scrappage Levels)	84
V	Policy Variables	
	1. Government Expenditures on Transfers and Goods and Services	32
	2. Monetary Variables	2
	3. Tax Rates and Government Revenue	62

.../2

APPENDIX

Type	Description	Number of Variables
VI	Demographic Variables	
	1. Regional	63
	2. National	32
VII	Real Income Variables, Overseas Economics	3
VIII	U,S. Economy	
	1. Wharton Model Variables	21
	2. Other	2
IX	Residential Construction - Regional	15
X	Regional Shares of Investment	60
XI	Miscellaneous Exogenous Variables	
	1. Adjusting Items	28
	2. Regional Output, Employment, Wages for very small Sectors and Potential Output	32
	3. Other	18
Total Exogenous Variables		608

Endogenous Variables

Total Endogenous Variables			2707
Regional Level	- Total		640
	- Identities	389 *	
	- Behavioural	251	
National Level	- Total		2067
	- Identities	1459**	
	- Behavioural	608	

* 140 for Population Algorithm
** 460 Dealing with Input-Output Table

... /3

APPENDIX

List of Sectors, Regional CANDIDE

Sector	Name of Sector	No. of Equations	Blocks from which Equations are Drawn
A	Personal Savings and Aggregate Consumption	9	1
B	Consumption Categories	126	2, 37, 41, (Eq. 1-3) 19, (Eq. 47-49)
C C	Residential Construction		
	National Variables	25	3
	Regional Variables	30	
D	Fixed Capital Formation in Business.		
	National	273	4, 32, 44, 45
	Regional	55	52
E	Inventory Investment	38	5, 38, (Eq. 40 to End)
F	Government Expenditures on Goods and Services	79	6, 38 (up to 39)
G	Export Flows	164	7 (in part) 16 (in part) 18 39 (in part) 42 (in part)
H	Import Flows	92	8, 17 (in part) 43 (in part)
I	Trade Flows in Automobiles (Detail)	52	41 (Eq. 4 to end) 46 (except final Eq.)
J	Derivation of Real Domestic Product by Industry		
	National	271	9, 10, 25, 23
	Regional	79	50
K	Labour Supply		11
	National	6	
	Regional	67	
L	Labour Demand		12
	National	27	
	Regional	60	

... /4

APPENDIX

Sector	Name of Sector	No. of Equations	Blocks from which Equations are Drawn
M	Wage Rates and Labour Incomes		
	National	48	13
	Regional	60	14 (Eq. 1-12)
N	Industry Prices (GDP or Deflators)	79	14 (13 to end)
O	Export Prices	20	16 (in part) 42 (in part)
P	Deflators of Import Trade Flows	40	17 (in part) 43 (in part)
Q	Prices of Competing Import Commodities	80	40
R	Derivation of Final Demand Deflators	420	26, 27, 28, 29, 30, 31, 15, 33, 34, 35, 17 (Eq. 40), 24 (Eqs. 24, 25, 27, 28) 36 (Eq. 1-4)
S	Personal Disposable Incomes		
	National	46	19 (in part)
	Regional	43	51
T	Government Revenue and Transfer Payments	39	19 (in part)
U	Money and Interest Rates	12	20
V	Balance of Payments	18	21
W	Demography		
	Regional - Population by Age Group	171	47, 48
	- Inter-regional Migration	40	49
	National	87	22
X	National Income Accounting Identities including Savings and Government Budget Balance	37	24 (in part) 36 (in part)
Y	Linkage with U.S. and other Foreign Economies	14	7 (in part) 29 (in part) 46 (in part)
	Total Number of Equations	2707	

CANDIDE-R Project Papers

1. G. Fortin, G. Simard, A. D'Amours, "Overview of the CANDIDE-R model".

2. S. Hébert, "Labour Supply".

3. R. Corbeil, "Residential Construction".

4. A. D'Amours, S. Simard, "Industry Outputs".

5. A. D'Amours, S. Simard, "Wages and Salaries".

6. A. D'Amours, J.P. Rioux, "Employment".

7. G. Fortin, J.P. Rioux, "Migration Flows".

8. G. Fortin, G. Simard, "Population".

9. J.P. Rioux, R. Corbeil, S. Simard, "Estimation Techniques".