A Short Description of Candide Model 1.0†

RONALD G. BODKIN,*
with a Mathematical Appendix by Stephen M. Tanny

This article is printed here with our apologies to the author. The article was to have been included in the proceedings of the inaugural convention but was inadvertently omitted.

T. S. S.

CANDIDE Model 1.0 is a large scale, econometric model of the Canadian economy. Before describing the model itself, it may be appropriate to discuss the acronym.

The acronym CANDIDE represents, in English, CANadian Disaggregated Inter-Departmental Econometric (model or project). I claim that this acronym is bilingual (in both official languages of Canada) and can be rendered in French as: (modèle ou projet) CANadian Désagrégé Inter-Départemental Econométrique.[1] Moreover, the French word *candide* can be translated as: open, frank, sincere, ingenuous, or *candid*. Hence the acronym is bilingual (in my view), while the word is almost so.[2]

We may in turn to some of the general characteristics of CANDIDE model 1.0. This is a national model of Canada; in other words, it takes no explicit account of regional aspects of the Canadian economy. Secondly, this is a very large model, in almost any dimension. The model has 377 exogenous variables (variables not determined within the system), of which roughly 60 are the direct result of policy choices. According to my recent count, there are 1527 endogenous variables (variables determined within the model), of which 569 are the dependent variables in behavioural equations and 958 can be regarded as being explained by

*University of Western Ontario

†The bulk of the text of this paper was written in late 1973, when I was serving as Project Manager of the CANDIDE project and Stephen M. Tanny was a member of the project staff. Earlier versions of this paper appeared (in prepublished form) as Discussion Paper No. 1 of the Economic Council of Canada, "A Short, Nontechnical description of CANDIDE Model 1.0," and (as a publication in Spanish) as "Breve descripción del Modelo CANDIDE 1.0," pp. 653–658 of *comercio exterior*, Vol. XXIII, No. 7 (July 1973), which is the journal of the National Bank of Foreign Trade of Mexico, Mexico City. It should be observed explicitly that the description in this paper applies directly to only Model 1.0, although much of what is said applies with some modifications to the two later generations of the CANDIDE model currently extant, namely Model 1.1 and Model 1.2.

[1] The English word "interdepartmental" is generally translated into French as "interministériel," which unfortunately destroys the acronym's property of being bilingual. From this point of view, my translation as "interdépartemental" is simply bad French or *franglais*. However, I prefer to look at the matter in another light; poetic license is permitted in French as well as in English, and too much attention to the rules of language can cripple poetry.

[2] As the acronym implies, the model grew out of a project involving several departments and agencies of the Government of Canada with the Economic Council of Canada playing a leading role in the co-ordination of the project. A more complete history of the project appears in the "Foreword" of each of the CANDIDE Project Papers listed at the end of the text.

identities. (Of these 958 identities, roughly 400 are relationships emanating from a rectangular input-output system, which is utilized within the model itself.) Thirdly, the model has been estimated from annual data, usually from the postwar period (e.g., 1955-1968) and usually on the basis of ordinary least squares estimation techniques. A fourth technical characteristic is that the model is a dynamic one, in almost any interpretation of this concept: important use is made of both stock variables (cumulants of flows) and lags.[3]

Turning to the economic characteristics of the model, we may note that CANDIDE Model 1.0 is envisaged as a medium-term model, with an economic horizon 8 to 12 years into the future. In my view, the most interesting results from projecting the model into the future are not the estimates of the endogenous variables for the individual years (which may at times reflect idiosyncracies of the model's equations) but rather growth rates for these variables, over a decade or half a decade.[4] The model is an amalgam of conventional macro-econometric modelling and input-output analysis; the approach to macro-economic theory that is taken is definitely neo-Keynesian in spirit (in contrast to a monetarist approach). As implied above, important use has been made of a rectangular Input-Output system (on the side of both production and prices) in conjunction with the use of adjustment equations at the end of the process.[5] The CANDIDE Input-Output system has 51 industries, 84 commodities, and 166 categories of final demand by ultimate user. Indeed, after some aggregation, CANDIDE Model 1.0 can be regarded as a collection of submodels for 12 (or 13) major producing sectors.[6] For each of these sectors, there are, with some exceptions, equations describing sectoral demand (output), investment, capital stock, employment, industry prices (the implicit deflators of gross domestic product originating in the major producing sector in question), and capital consumption allowances.

A model of this size requires the use of the electronic computer to handle it efficiently or, indeed, merely to keep track of its myriad relationships.[7] The computer records are divided according to blocks, a set of no more than 99 equations that have some relationships to each other. However, for purposes of describing the model quickly, we may speak of sectors of the model or "Superblocks", which are collections of whole blocks and pieces of blocks.[8] (In one

[3]The lags employed are both discrete lags and distributed lags, with heavy use being made of polynominal (Almon) lag distributions in the case of distributed lags.

[4]The Economic Council of Canada, in its use of CANDIDE Model 1.0 in its Ninth Annual Review, *The Years to 1980* (Ottawa: Information Canada, 1972), made a great deal of use of decade and semi-decade average growth rates in the presentation of the results. This approach to the presentation of results was, in my view, something more than a mere expository device.

[5]The treatment could be regarded as a development of that of Franklin M. Fisher, Lawrence R. Klein, and Yoichi Shinkai, "Price and Output Aggregation in the Brookings Econometric Model," Chapter 17 of James S. Duesenberry, Gary Fromm, Lawrence R. Klein, and Edwin Kuh (eds.), *The Brookings Quarterly Econometric Model of the United States* (Chicago and Amsterdam: Rand McNally & Company and North Holland Publishing Company, 1965).

In this paper, I shall generally *not* give footnote references to the literature, especially when the work in question is either well known and/or extensively cited in our CANDIDE Project Papers. The exceptions to this rule (i.e., the few cases when literature is cited) will be cases like this footnote, where the reference seems (to me) particularly crucial to the discussion.

[6]The ambiguity arises because one has a choice whether to include the housing industry (the output of which is mainly paid and imputed residential rents) as a major producing sector.

[7]A brief description of the computer software utilized to keep track of the CANDIDE Model (and associated data and computer programs) may be found in Appendix A, "A System for Large Econometric Models—Management, Estimation, and Simulation," in M. C. McCracken, *An Overview of CANDIDE Model 1.0*, CANDIDE Project Paper No. 1 (Ottawa: Information Canada, 1973).

[8]The sectors of the model (the "superblocks") should not be confused with the producing sectors, which are simply the results of a regrouping (usually with some aggregation) of the basic 43 industries (or 51 Input-Output industries) of the model.

TABLE 1
Sectors or "Superblocks" of CANDIDE Model 1.0

A. Final Demand by Ultimate User
   1. 406 equations (199 behavioural equations, 207 identities)
   2. Blocks in this sector:
      a. Block 1, Aggregate Consumption and Personal Savings
      b. Block 2, Disaggregated Consumption
      c. Block 3, Residential Construction
      d. Blocks 4 and 32, Fixed Business Investment
      e. Block 5, Inventory Investment
      f. Block 6, (Resource-Using) Government Expenditures
      g. Block 7, Export categories
      h. Block 8, Import categories

B. Industry Output Determination
   1. 246 equations (43 behavioural equations, 192 Input-Output identities, 11 other identities)
   2. Blocks in this sector:
      a. Block 9, Final Demand Conversion
      b. Block 25, Input-Output Estimates of Industry Gross Outputs
      c. Block 26, Input-Output Estimates of Levels of Real Domestic Product or Value-Added, by Industry
      d. Block 10, Adjustment Equations for Real Domestic Product, by Industry

C. Labour Supply and Requirements
   1. 86 equations (32 behavioural equations, 54 identities)
   2. Blocks in this sector:
      a. Block 11, Labour Supply (including Unemployment)
      b. Block 12, Labour Requirements
      c. Block 22, Demography (all identities)

D. Wages and Prices
   1. 572 equations (251 behavioural equations, 210 Input-Output identities, 111 other identities)
   2. Blocks in this sector:
      a. Block 13, Wages
      b. Block 14, Industry Prices (Implicit Deflators of Gross Domestic Product by Industry)
      c. Block 16, Export Prices
      d. Block 17, Import Prices
      e. Block 27, Input-Output Estimates of Commodity Prices
      f. Block 28, Input-Output Estimates of Consumption Prices
      g. Block 29, Input-Output Estimates of Prices of Government Expenditure
      h. Block 30, Input-Output Estimates of Prices of Expenditures on Machinery and Equipment
      i. Block 31, Input-Output Estimates of Prices of Construction Expenditures
      j. Block 15, Implicit Deflators of Consumption Expenditures
      k. Block 33, Implicit Deflators of Expenditures on Machinery and Equipment
      l. Block 34, Implicit Deflators of Expenditures on Construction
      m. Block 35, Implicit Deflators of Government Expenditures
      n. Aggregate Identities: Major Portion of Block 36 and Small Portions of Blocks 24, 37, and 38
   3. Alternative Grouping of this Huge Sector:
      a. Foreign Trade Prices
      b. Industry Wage Determination
      c. Industry Prices
      d. Input-Output Price Relationships
      e. Adjustment Equations for Deflators of Final Demand Categories
      f. Aggregate Identities Relating to Price Levels

E. Government and Private Revenues
   1. 60 equations (24 behavioural equations, 36 identities)
   2. Blocks in this sector:
      a. Block 18, Government Revenue (including Budget Balance Measures)
      b. Block 19, Private Revenues

TABLE 1 *continued*

F. Financial Flows
  1. 23 equations (16 behavioural equations, 7 identities)
  2. Blocks in this sector:
     a. Block 20, Money and Interest Rates
     b. Block 21, Financial Flows in the Balance of Payments

G. National Accounts Relationships
  1. 122 equations (1 behavioural relationships, 121 identities)
  2. Blocks in this sector:
     a. Block 24 (major portion)
     b. Block 36 (minor portion)
     c. Block 37 (major portion)
     d. Block 38 (major portion)

H. Linkages with the U.S. and other Foreign Economies
  1. 12 equations (3 behavioural equations, 9 identities)
  2. A single block sector, comprised of Block 23.

case, I choose to classify a computer block into simply one sector of the model.) My listing of the sectors of the model is given in Table 1 below; a perusal of this table gives the reader a basic idea of the structure of CANDIDE Model 1.0.

While a detailed description of the equations of the model is beyond the scope of this paper, I should like to point up to the reader some of the more interesting features of CANDIDE Model 1.0. This may be done by quickly reviewing the sectors of the model.

The first sector of the model in Table 1 is the fairly conventional one of effective demand by category of ultimate use, except that this sector is highly disaggregated in CANDIDE Model 1.0. We approach aggregate consumption through estimating personal savings and then subtracting personal savings (and several other items) from disposable income. One justification of this approach is that enables one to distinguish motives for saving; discretionary personal savings and contractual personal savings are distinguished in CANDIDE Model 1.0. In Block 2, expenditure functions for 55 categories of consumer expenditure are estimated, generally through a variant of the Houthakker-Taylor model of consumer expenditure. The demand for residential construction expenditures (Block 3) is generated through a starts-completions-expenditures' unit mechanism; income, price, and demographic factors play major roles in explaining the demand for residential construction. The demand of business firms for fixed capital formation in both plant and equipment (treated separately) is explained by a variant of Professor Jorgenson's "neo-classical" theory of investment; in CANDIDE Model 1.0, we distinguish some 38 industries which make separate investment decisions. The inventory investment demand functions (Block 5) are primarily oriented toward explaining the trends in this component of final demand, as would seem appropriate in a medium-term model. (In a quarterly model focussing on short-run fluctuations, more attention would have been paid to this category of final demand.) The demands of governments for resource-using expenditures, in Block 6, is a fairly novel feature of CANDIDE Model 1.0; we have taken the view that, in a medium-term context, government expenditures on economic goods and services are fairly responsive to economic factors like population and real income that hence there is relatively little discretion for political decision-makers to change expenditures autonomously.[9] Finally, the foreign trade demand functions (25

[9] If the user finds this approach objectionable, the software developed in conjunction with the model (see footnote 7) is flexible enough that he can override these government expenditure demand functions and take these variables as exogenous. Of course, in a projection exercise (i.e., a use of the model beyond the sample period), the user would have to furnish his own projections of the values of these variables.

categories of exports, 12 of imports) are explained in terms of variables such as relative prices, long-run income growth (or a variant), cyclical variables, and (in some cases) special circumstances.[10]

From the final demand sector of the model, we obtain (after some processing and aggregation) some 166 categories of final demand by ultimate use, which are then fed into our rectangular input-output system. Block 9 converts these 166 categories of final demand by ultimate use into a vector of final demands by the commodity classification (84 in total) of the Input-Output system utilized in CANDIDE Model 1.0.[11] A rectangular input-output system (one that distinguishes between the intermediate commodities and the producing industries) is then employed to generate estimates of the gross outputs of some 51 industries, in Block 25. In Block 26, these estimates are further processed and slightly aggregated to obtain input-output estimates of the real domestic product originating in some 43 industries. In Block 10, regression techniques are employed to estimate autoregressive correction equations, in which the discrepancy between actual real domestic product originating in a particular industry and its input-output estimate is generally estimated as a function of a constant term, a time trend, and the lagged value (or lagged values) of this discrepancy. One can interpret these autoregressive correction equations as attempting to correct cheaply for the rigidities entailed in the use of Input-Output tables pertaining to a single calendar year (1961). Thus our input-output system has made no explicit corrections for technological change, changing relative prices, and compositional changes within groupings; these are the sorts of rigidities that the adjustment equations attempt to overcome.

In the labour sector of the model, we explain labour force behaviour on the basis of demographic variables and also on the basis of participation rates. Block 22, which is comprised of only identities, keeps track of various age-sex groupings of the Canadian population, in terms of such ultimate determinants as birth rates, death rates, immigration rates, marriage and divorce rates, all of which are exogenous. In Block 11, participation rate equations are estimated for three "non-primary" labour force groups: females under 35, females 35 and over, and non-prime age males (males 14–24 and 55 and over). The participation rate for males 25–54 is taken as exogenous. The product of the participation rate and the relevant source population is the labour force for the age-sex group in question; the sum over the four groups yields the total labour force. The demand for labour (on the basis of both employees and man-hours) is obtained through a requirements approach: the outputs of the 43 industries of Block 10 are aggregated to a basis of 12 major producing sectors or major industries, and then equilibrium labour requirements are obtained by inverting a Cobb-Douglas production function.[12] Actual labour requirements are then obtained by permitting less than instantaneous adjustment.[13] Summing across the 12 major industries, one then obtains total employment; unemployment is then the discrepancy between the labour force and total employment, and

[10] For some foreign trade categories (particularly on the exports side), the category demand is taken as exogenous, because of the importance of special factors that are very difficult to capture in a fitted equation with a few explanatory variables. Often the projections of these foreign trade categories (into some future period) are made on the basis of the specialized knowledge of experts in the field.

[11] A more detailed description of the Input-Output system utilized in the CANDIDE Model appears in Ronald G. Bodkin, "A Large-Scale Input-Output Econometric Model of the Canadian Economy (CANDIDE)," pp. 27–44 of Karen R. Polenske and Jiri V. Skolka (eds.), *Advances in Input-Output Analysis* (Cambridge, Mass., Ballinger Publishing Co., 1976).

[12] Actually, a slightly different approach is taken for two major industries, Agriculture and Public Administration (which includes national defense).

[13] One expects that the adjustment of the man-hours equations should be more rapid than that of the equations for employees, because in general it is easier to work short time or overtime than to hire or fire employees. In general, these expectations are borne out.

this figure can be expressed as a proportion of the total labour force, if desired.

Next, we may turn our attention to wage and price determination, in CANDIDE Model 1.0. The general approach (with some exceptions) to wage determination is that of a wage adjustment function or modified "Phillips Curve", applied (in Block 13) on the basis of the 12 major industries of the economy. In a typical industry wage adjustment equation, the explanatory variables would be: the rate of unemployment, the rate of change of the consumer price level, and perhaps the rate of change of the corresponding U.S. wage rate. In Block 14, the price of industry output (the deflators of gross domestic product) are explained on the basis of unit labour costs or unit total costs, at times supplemented by demand-type or U.S. variables. This is done for the twelve major industries, which are then expanded into estimates of deflators of gross domestic product for all 51 industries of the CANDIDE Input-Output system. Foreign trade prices are essentially exogenous; Blocks 16 and 17 merely keep track of identities and (in the case of the detailed import prices of Block 17) generate estimates of the prices of imported commodities for use in the price relationships of the Input-Output system. Given the deflators of gross domestic product for the 51 industries, the detailed import prices, and the price levels of some other primary inputs of the Input-Output system (such as indirect taxes and subsidies), one can compute, on the basis of the Input-Output system, estimates of the price levels of domestically produced commodities. In turn (Blocks 28, 29, 30, and 31), one can work through the expenditures matrix of Block 9 to obtain first estimates of the deflators for the categories of consumption expenditures, government expenditures, residential construction, and fixed business investment. Finally, these first estimates (on the basis of the Input-Output system) can be subjected to autoregressive correction equations, designed to adjust for the rigidities inherent in the use of Input-Output tables pertaining only to a single year and also to correct for other short-comings in the submodel of prices.

The revenues sector may be summarized briefly. In Block 18, government revenues (principally but not exclusively tax revenues) are generated, typically on the basis of exogenous tax rates. Block 18 also contains some identities for budget balance measures, and it should be explicitly noted that, in Model 1.0, all three levels of government (federal, provincial, and municipal) are combined, on the revenue side.[14] In Block 19, we calculate estimates of corporate profits (needed, in particular, as the base for the corporate income tax, the receipts of which are calculated in Block 18), capital consumption allowances, and some supplementary income items. Corporate profits are computed in essentially the following manner. From Block 12, one calculates total man-hours employed in the various industries of the economy and from Block 13, one calculates (for eight industries) the per man-hour rate of wage compensation. Hence the product is essentially wage and salary income by major industry, which can then be summed across industries to obtain the model's estimate of wages and salaries for the total economy.[15] Next, after calculating total nominal gross national product in Block 24, one can obtain an estimate of net property income of the Canadian economy by subtracting from nominal gross national product the estimated wage income, the capital consumption allowances, and two reconciliation items[16] and then adding in dividends to non-residents of Canada. Finally, corporate profits are linked to this concept of net property income by a rough-and-ready behavioural equation, which also shows the expected pro-cyclical behaviour of corporate profits.[17]

[14]On the side of government expenditures on resource-using activities, there is some disaggregation among levels of government in current account expenditures, although government capital expenditures are calculated for all three levels of government combined. On the side of government expenditures on transfer payments, these are accounted for in Block 19; most of the components of government transfer payments are exogenous. In all cases, the treatment of such transfer payments is for all three levels of government combined.

[15]The description in the text is oversimplified, but there is no need to go into exact details at this point. It should be noted, however, that the concept of labour income (wages and salaries) utilized in CANDIDE Model 1.0 does include an allowance for labour income originating in unincorporated businesses, which would be attributable to the efforts of the owner (or owners) and unpaid family workers.

Turning to the Financial Flows sector, we may comment that our approach to money and interest is rather strongly in the Keynesian tradition. The key interest rate, the average yield of three month treasury bills, is determined in a semi-reduced form; this variable, taken as dependent, is regressed against three explanatory variables: the monetary base (expressed in constant dollars by dividing by the implicit deflator of gross national expenditure), real gross national product (less the imputed items), and a distributed lag in the interest rate on prime commercial paper in the U.S.A. (utilized in the Wharton Model). Other interest rates (specifically, the yield on long-term government bonds, the yield on industrial bonds, and the conventional mortgage rate) are in part explained by this key interest rate. In the CANDIDE Model, interest rates primarily influence final demand through their effect on investment outlays. The principal effect is that on fixed business investment (Block 4) through the Jorgenson-type "cost of capital" variable; there are, however, effects on the demand for residential construction also. Block 21, financial flows in the balance of payments, is currently constructed on the assumption of a fixed exchange rate and so keeps track, in the model, of the movements of international reserves. Since Canada has had a floating exchange rate since June 1970, it would be highly desirable to model the current institutional structure. So far our efforts in this direction have not been overwhelmingly successful.

The penultimate sector of Table 1, National Accounts Relationships, can be regarded as a generalization of the simple national income accounting identity: $Y = C + I + G + (X - M)$, where Y is net national product, C is total consumption, I is net domestic investment, G is government expenditures on resource-using activities, and $(X - M)$ is net exports. The one behavioural equation of this sector is present to generate an estimate of the total of the imputed items in the National Accounts. The model generates its final estimate of gross national product in current dollars by taking a simple average of the estimate built up from the expenditures side and of the estimate built up from the side of production (gross domestic product, by major industry).[18] The final estimate of gross national product in constant dollars is constructed in a related manner.

The final sector is comprised of a single block, linkages with the U.S. and Other Foreign Economies. Under fixed exchange rates, this sector is anterior to the rest of the model; in other words, it can be solved by itself and the results fed into the solution of the main portion of the model. This sector takes exogenous variables and transforms them so that they are immediately usable in the solution of CANDIDE Model 1.0. Hence, one could have taken the results of transforming the exogenous variables in question as exogenous themselves; expressing the transformations as equations of the model is neater and more convenient, es-

[16]Specifically, these reconciliation items are net indirect taxes (indirect taxes less subsidies) and the inventory valuation adjustment item.

[17]Specifically, this is done by including the rate of unemployment as an explanatory variable in this relationship; the unemployment rate displays a regression coefficient that is significantly negative, by conventional tests of statistical significance.

[18]In a sense, the procedure in the model duplicates that of Statistics Canada in putting together the official estimates of gross national product in the national income accounts. Statistics Canada also employs a third approach, through the side of functional income shares received. This approach is not open to us, however, as the income receipts identity is already used (in Block 19) to generate estimates of corporate profits.

TABLE 2
Classes of Exogenous Variables, in CANDIDE Model 1.0

A. Demographic Variables (Underlying Magnitudes)	55
B. Import and Export Prices	
1. Export Prices	23
2. Trade Price Relatives	6
3. Import Prices	12
C. U.S. Economy	
1. Wharton Model Variables	19
2. Others	3
D. Overseas Economies, Real Income Variables	4
E. Foreign Trade Variables	
1. Export Variables	16
2. Import Variables	6
F. Policy Variables	
1. Tax Rates and Government Revenues	31
2. Transfer Payments and Some (Resource-Using) Government Expenditures	20
3. Monetary Base and Other Financial Variables	6
G. Technical Exogenous Variables	
1. Time Trend	1
2. Dummies	48
3. Depreciation Rates and Scrappage Levels (for Industry Capital Stocks)	46
4. Three Pass Least Squares Variables	25
H. Exchange Rates and International Transactions	
1. Exchange Rates	2
2. International Transactions	25
I. Adjustment Items and Other Miscellaneous Exogenous Variables	29
J. Grand Total	377

pecially for projections beyond the sample period. As an illustration of the type of relationships in this block, consider Equation (5) of this block, in which the dependent variable is U.S. Personal Spending on Alcoholic Beverages, which is employed in explaining the demand for one of the export categories (Block 7). In the Wharton Model, one of the endogenous variables is U.S. Spending on Food and Alcoholic Beverages: this variable (exogenous from the point of view of the CANDIDE Model), along with the U.S. rate of unemployment, is employed in Equation (5) as an explanatory variable.

An econometric model is described, not only by the variables explained within it (the endogenous or jointly dependent variables), but also the set of variables which are taken as given to the system (the exogenous or independent variables). In Table 2, there appears a list of the exogenous variables of CANDIDE Model 1.0. In general, this list should be self-explanatory, although several comments may be offered. First, the Canadian economy (with levels of either total exports or total imports equal roughly to one-quarter of gross national expenditure[19] is very open to influences emanating from abroad, on both the side of prices and real incomes, and the list of exogenous variables correctly suggests that we have attempted to capture these influences in CANDIDE Model 1.0. Secondly, as in the case of a number of Latin American countries, the trading partner with the largest influence on Canada's domestic economy is the U.S.A. We have attempted to capture this influence by separating out U.S. variables explicitly, at many points of the model. Moreover, the treatment of U.S. exogenous variables is worthy of explicit comment. As has already been implied at several points in the discussion, 19 out of 22 U.S. exogenous variables are outputs (endogenous variables) of Professor Ross Preston's Wharton Annual and Industry Forecasting Model.[20] This has limited relevance for the sample period, as during the sample period the exogenous variables are what they are (except, possibly, for some counterfactual simulations like multiplier studies), and where they come from is a question of secondary importance. However, for projections into the future, the fact that these variables are outputs of an econometric model means that one can obtain a set of projections for all of these variables that are mutually consistent with each other and which come from a model of the U.S. economy that is both broadly consistent with the view of the Canadian economy taken in the CANDIDE Model and also regarded as utilizing the best practices currently available for this sort of interindustry, macroeconometric modelling.[21]

[19]Of course, net exports (exports minus imports) is a much smaller fraction of gross national expenditure.

This comment leads into the point that, for a number of cases, projections of the exogenous variables (of the sort that will satisfy the purposes of the exercise) is not a trivial task, and considerable thought and study must be devoted to the job of obtaining a satisfactory set of exogenous variables for some future period. In some cases, one can call upon the expertise of individuals who work in particularly defined fields; however, in many cases the generalist who is attempting to get sensible results from the model as an analytical tool cannot escape the task of making an intelligent set of judgments about the most probable values of some of the more critical exogenous variables. (In some contexts, the most probable set of values of these exogenous variables might be replaced by the most desirable, depending upon the purpose of the application.)

[20]Ross S. Preston, *The Wharton Annual and Industry Forecasting Model*, Studies in Quantitative Economics No. 7 (Philadelphia, U.S.A.: Economics Research Unit of the University of Pennsylvania, 1972). It may be noted that, in building the CANDIDE Model, the treatment of certain problems in the Wharton Annual and Industry Forecasting Model was a continuing source of inspiration.

[21]Of course, alternative methods (such as pure intuition, extrapolation of time trends, single equation relationships, or partial models) could be utilized for projecting these particular exogenous variables, without recognizing the relevance of their source. However, in our view, such an approach is likely to be inferior, with regard to both practical and intellectual considerations.

Candide Project Papers

1. M. C. McCracken, *An Overview of CANDIDE Model 1.0*, CANDIDE Project Paper No. 1, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).
2. Thomas T. Schweitzer and Tom Siedule, *CANDIDE Model 1.0: Savings and Consumption*, CANDIDE Project Paper No. 2, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).
3. H. E. L. Waslander, *CANDIDE Model 1.0: Residential Construction*, CANDIDE Project Paper No. 3, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).
4. C. Dewaleyne, *CANDIDE Model 1.0: Inventories*, CANDIDE Project Paper No. 4, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).
5. Derek A. White, *CANDIDE Model 1.0: Business Fixed Investment*, CANDIDE Project Paper No. 5, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1974).
6. Gilles Proulx and Thomas T. Schweitzer, *CANDIDE Model 1.0: Government Final Demand Expenditures*, CANDIDE Project Paper No. 6, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1974).
7. J. R. Downs (with Bobbi Cain), *CANDIDE Model 1.0: Foreign Trade*, CANDIDE Pro-

ject Paper No. 7, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).

8. L. Auer and D. Vallet, *CANDIDE Model 1.0: Industry Output Determination*, ed. Ronald G. Bodkin and Barbara A. M. Young, CANDIDE Project Paper No. 8, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1974).

9. Wolfgang M. Illing, *CANDIDE Model 1.0: Labour Supply and Demographic Variables*, CANDIDE Project Paper No. 9, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).

10. Wolfgang M. Illing, *CANDIDE Model 1.0: Labour Demand*, CANDIDE Project Paper No. 10, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).

11. Wolfgang M. Illing and C. Dewaleyne, *CANDIDE Model 1.0: Industry Wages and Industry Prices*, CANDIDE Project Paper No. 11, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1973).

12. L. Auer and D. Vallet, *CANDIDE Model 1.0: Final Demand Deflators*, ed. Ronald G. Bodkin, Stephen M. Tanny, and Barbara A. M. Young, CANDIDE Project Paper No. 12, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada 1975).

13. H. E. L. Waslander, *CANDIDE Model 1.0: Sector Incomes and Government Revenues*, CANDIDE Project Paper No. 13, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1974).

14. J. F. McCollum and Tom Siedule, *CANDIDE Model 1.0: Monetary Sector*, CANDIDE Project Paper No. 14, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1974).

15. J. F. McCollum, *CANDIDE Model 1.0: Capital Movements and Exchange Rates*, CANDIDE Project Paper No. 15, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1974).

16. H. E. L. Waslander, *Aggregate Identities and Other Functions Employed in CANDIDE Model 1.0*, CANDIDE Project Paper No. 16, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1974).

17. Ronald G. Bodkin, K. S. R. Murthy, and Tom Siedule, *The Tracking Properties of CANDIDE Model 1.0 over the Sample Period 1955-68*, CANDIDE Project Paper No. 17, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1975).

18. Ronald G. Bodkin and Stephen M. Tanny, (eds.), *CANDIDE Model 1.1*, CANDIDE Project Paper No. 18, Economic Council of Canada, for the Interdepartmental Committee (Ottawa: Information Canada, 1975).

Mathematical Appendix

On the Solution Algorithm Employed in CANDIDE Model 1.0 (by Stephen M. Tanny)

CANDIDE Model 1.0 contains over 1500 equations. A proper subset of approximately 1000 of these form the simultaneous core of the model. The majority of the equations in the simultaneous core are linear; however, there are a significant amount of non-linearities in the model. In this note, the Gauss-Seidel algorithm, which is used to solve this system of equations, is described. Some related solution techniques are also discussed.

We may first introduce some notation. Let y_i, $i = 1, 2, \ldots, n$, be the *i-th* endogenous (jointly dependent) variable, and let x_j, $j = 1, 2, \ldots, k$, be the *j-th* predetermined (independent) variable.[1] Hence, our system of equations has n jointly determined variables and k independent variables. The typical equation of CANDIDE Model 1.0 can be written:

$$y_i = f_i(y_1, y_2, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n; x_1, x_2, \ldots, x_k), \; i = 1, 2, \ldots, n. \qquad \text{(I)}$$

Here, the function f_i is specifically associated with the *i-th* endogenous variable y_i.

Our procedures can be illustrated by considering the special case in which all the equations are linear. We may include the effects of a given set of the predetermined variables in a set of constant terms, b_i, $i = 1, 2, \ldots, n$. (Thus, in the linear case, there is one and only one constant term for each equation.) With this simplification, the equation set (I) becomes:

$$\begin{aligned} y_1 &= b_1 - a_{12}y_2 - \cdots - a_{1i}y_i - a_{1n}y_n \\ &\vdots \\ y_i &= b_i - a_{i1}y_1 - \cdots - a_{i,i-1}y_{i-1} - a_{i,i+1}y_{i+1} - \cdots - a_{in}y_n \\ &\vdots \\ y_n &= b_n - a_{n1}y_1 - \cdots - a_{ni}y_i - \cdots - a_{n,n-1}y_{n-1}, \end{aligned} \qquad \text{(II)}$$

where a_{ij} is the coefficient of y_j in the *i-th* equation.[2]

We are now ready to discuss the Jacobi iteration, for linear systems. As an initial approximation to the solution of system (II), take:

$$y_1^0 = b_1, y_2^0 = b_2, \ldots, y_n^0 = b_n. \qquad \text{(III)}$$

Replacing the unknowns on the right-hand side of system (II) with these values, we obtain a new approximation $y_1^1, y_2^1, \ldots, y_n^1$ to the solution. This procedure can be expressed compactly as follows. Let y_i^m be the approximation to the solution value of y_i after m iterations. Then y_i^m is given by the expression:

$$y_i^m = b_i - \sum_{\substack{t=1 \\ t \neq i}}^{n} a_{it} y_t^{m-1}, \quad i = 1, 2, \ldots, n; \; m = 1, 2, 3, \ldots. \qquad \text{(IV)}$$

Iterating the procedure embodied in Equation (IV), we hope to converge to the solution of system (II) above. It is this procedure that is termed the *Jacobi iteration*.

The Gauss-Seidel algorithm is a slight modification of the Jacobi iteration. If the procedure outlined above does, in fact, converge to a solution, then (in general) y_i^m will be a better approximation to y_i than will y_i^{m-1}. One could thus take account of this fact by modifying algorithm (IV) above so that we utilize y_t^m as our estimate of y_t, $t = 1, 2, \ldots, i-1$, while for y_t, $t = i+1, i+2, \ldots, n$, we employ y_t^{m-1}, the estimates derived in the previous iteration. Thus, in forming our estimate of y_i on the *m-th* iteration, we use the most recent estimates available of the other jointly dependent variables. This procedure, which is termed the Gauss-Seidel algorithm, can be formally expressed as:

$$y_i^m = b_i - \sum_{t=1}^{i-1} a_{it} y_t^m - \sum_{t=i+1}^{n} a_{it} y_t^{m-1}, \quad i = 1, 2, \ldots, n; \; m = 1, 2, 3, \ldots. \qquad \text{(V)}$$

It is time to raise the issue of convergence of these algorithms. For the linear case, it is not difficult to show that these algorithms do in

[1] The predetermined variables of this appendix include both the exogenous variables discussed in the text and also lagged values of the endogenous variables (the y_i's). For the purposes of this appendix, this distinction is unimportant.

[2] We assume that this system of equations is linearly independent (i.e., the associated matrix of coefficients is nonsingular), so that a unique solution exists.

fact converge in most practical circumstances.[3] Moreover, it has been found that the Gauss-Seidel algorithm converges whenever the Jacobi iteration does. Since (as one might expect from the plausibility argument above) the former generally converges more rapidly and with smaller fluctuations, the Jacobi iteration is generally no longer used.

For the record, we might note that the Gauss-Seidel algorithm can be regarded as a special case of the following general relaxation algorithm:

$$y_i^m = w_i \left[b_i - \sum_{t=1}^{i-1} a_{it} y_t^m - \sum_{t=i+1}^{n} a_{it} y_t^{m-1} \right] + (1 - w_i) y_i^{m-1},$$

$$i = 1, 2, \ldots, n; m = 1, 2, 3, \ldots. \qquad \text{(VI)}$$

Here, the w_i's are termed "relaxation parameters" and generally vary between 0 and 2.[4] It is apparent that, with this algorithm, the estimate of y_i in the *m-th* iteration is a convex combination of its estimate according to the Gauss-Seidel formula and its estimate on the previous iteration. We shall return to the question of general relaxation algorithms in the discussion below of the solution of nonlinear systems of equations.

We may now consider the nonlinear case, which is the relevant case for the system of equations that constitutes CANDIDE Model 1.0. We may note that we can still apply the Gauss-Seidel algorithm outlined above, with some minor modifications.[5] To be explicit, we choose (on the basis of the economics of the problem under consideration) a given set of values for our k predetermined or independent variables, $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$. Thus the $\bar{x}_j$'s become simply constants, when the problem has been reduced to this level. Next, as the first set of approximations to the y_i's, to be denoted by the symbols y_i^0, $i = 1, 2, \ldots, n$, we may set:

$$y_1^0 = f_1(\hat{y}_1, 0, \ldots, 0; \quad \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)$$

$$\vdots$$

$$y_i^0 = f_i(0, \ldots, \hat{y}_i, \ldots, 0; \quad \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)$$

$$\vdots$$

$$y_n^0 = f_n(0, 0, \ldots, \hat{y}_n; \quad \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k). \qquad \text{(VII)}$$

Here, the symbol $\hat{y}_i$ means that y_i does not explicitly appear in the functional form (in this case f_i) under consideration. Next, the estimate of y_i after m iterations, to be denoted (as before) by the symbol y_i^m, is given by:

$$y_i^m = f_i(y_1^m, y_2^m, \ldots, y_{i-1}^m, y_{i+1}^{m-1}, \ldots, y_n^{m-1}; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k),$$

$$i = 1, 2, \ldots, n; \quad m = 1, 2, 3, \ldots. \qquad \text{(VIII)}$$

Thus, as in the linear case, we use the most recently available estimates for the set of jointly dependent variables $(y_1, y_2, \ldots, y_n)$, in calculating the estimate for the *m-th* iteration of a particular y_i. This is done on the view that, for a converging system, estimates from the most recently available iteration are likely to be more accurate. When two solutions differ by less than (small) positive constant ϵ[6], then we consider system (I) solved for the choice of values of the k predetermined variables of our problem.

We may sketch the general relaxation algorithm in the nonlinear case, also. The starting point remains the same as in system (VII) above, but the estimate of y_i after m iterations is:

$$y_i^m = w_i f_i(y_1^m, y_2^m, \ldots, y_{i-1}^m; y_{i+1}^{m-1}, \ldots, y_n^{m-1}; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k) + (1 - w_i) y_i^{m-1},$$

$$i = 1, 2, \ldots, n; \quad m = 1, 2, 3, \ldots. \qquad \text{(IX)}$$

Again, a similar criterion for convergence may be employed. It is worth noting that general relaxation algorithms have been used to solve very large systems (comprised of some 20,000 to 50,000 equations) that arise from problems in the physical sciences, while in CANDIDE Model 1.0 we have a simultaneous core of approximately 1000 equations. Moreover, when the general relaxation algorithm is employed, generally the bulk of work focuses on calculating optimal values for the relaxation parameters (the w_i's); indeed, these optimal values may even vary by the stage of the iteration procedure. The method, while interesting, has not been applied to the CANDIDE system, nor would it appear to be a useful computational procedure at this stage of the CANDIDE Project. However, the general relaxation algorithm may become of interest to us if the development of our large scale econometric model would be to pursue a path of even further disaggregation.[7]

Finally, we may turn to the question of the existence of a solution to the system of equations (I), using the Gauss-Seidel algorithm and the convergence criterion outlined above.[8] Although the technique is known to converge for a number of special cases (which is growing over time), there still remains a wide area of ignorance. This is so because it would seem that one cannot prove existence in general for all kinds of nonlinearities[9], but must instead focus on particular cases. In particular, while we have not proved that a system of equations with the characteristics of CANDIDE Model 1.0 must converge (to a unique solution), in practice it appears to do so, with the use of the Gauss-Seidel algorithm. A similar experience has been noted for the large scale Brookings econometric model (of the U.S. economy).[10]

---

[3]We may note, in passing, that necessary and sufficient conditions for the convergence of these algorithms are known; in particular, they can be formulated in terms of a property of certain related matrices known as quasi-nilpotence. On these matters, a useful reference is: E. K. Blum, *Numerical Analysis and Computation: Theory and Practice* (Reading, Mass., U.S.A.: Addison-Wesley Publishing Company, Inc., 1972).

[4]The case in which $w_i = 1$, $i = 1, 2, \ldots, n$, is of course the special case of the Gauss-Seidel algorithm. Let us denote the Gauss-Seidel type of estimate of y_i at the *m-th* iteration by the symbol $y_i^{GS,m}$. Then Equation (VI) above can be rewritten:

$$y_i^m = y_i^{GS,m} + (w_i - 1) \left[y_i^{GS,m} - y_i^{m-1} \right], \qquad \text{all } i, \text{ all } m. \qquad \text{(VIA)}$$

Thus the general relaxation algorithm makes some use of the numerical evolution of the successive estimates of y_i, in order to produce more accurate estimates (hopefully) at the *m-th* iteration and so shorten the process of convergence. From Equation (VIA) above, one might expect intuitively that values of the w_i parameters between 1 and 2 will generally be most helpful in shortening the process of convergence. (This might be expected to be the case, because a value of w_i in excess of unity entails pushing the estimate of y_i at the *m-th* iteration in the direction of the trend of its development, over the process of iterating.) While this is generally true, Gary Fromm and Lawrence R. Klein have pointed up a case in which it is better to choose w_i (in our notation) to lie between 0 and 1. (See Gary Fromm and Lawrence R. Klein, "Solutions of the Complete System," Chapter 11 of James S. Duesenberry, Gary Fromm, Lawrence R. Klein, and Edwin Kuh (eds.), *The Brookings Model: Some Further Results* (Chicago, Amsterdam, and London: Rand McNally & Company and North Holland Publishing Company, 1969), pp. 375–382, especially p. 382.)

[5]We do not discuss the Jacobi iteration for the nonlinear case because, here also, it is rarely used at present, having been superseded almost universally by the Gauss-Seidel algorithm.

[6]Specifically, we consider the system solved when we reach a particular iteration (call it m^*) for which the following condition holds:

$$\max_i \left| y_i^{m^*} - y_i^{m^*-1} \right| < \epsilon.$$

In this case, the solution is said to require m^* iterations.

[7]While absolving him of any errors or inaccuracies in the above discussion, we should like to thank Dr. David Hill of Temple University for his comments on the issues discussed in this paragraph.

[8]Secondarily, we should be interested in the possibility of multiple solutions; but in practice multiple solutions do not appear to be a problem with large scale econometric models.

[9]Indeed, it is easy enough to "cook up" special nonlinearities for which the standard methods will *not* converge.

[10]For a further discussion of the application of methods like these to the system of equations of the Brookings econometric model, see Fromm and Klein, "Solutions of the Complete System."

This suggests that the nonlinearities encountered in large scale econometric models are sufficiently "well-behaved" that they can be reasonably approximated by linear functions, at least over the range of interest. Thus, in practical terms, it would appear that we can continue to use computational methods (which can be justified rigorously only in the linear case and for some special nonlinearities) to obtain solutions for our large scale, nonlinear econometric models. In this regard, there would appear to be an analogy to the possible problem of "cycling", when applying the well-known simplex method to solve a standard linear programming problem in the "degenerate" case.[11]

Appendix 2

Optimization in the Context of a Large Scale Econometric Model, Such as CANDIDE

As this is a conference on optimal control theory, it behooves me to raise the issue of optimization in the context of the CANDIDE Model. Optimization has a long history as a topic of great interest in economics. But how does one optimize in the context of a descriptive model?

Obviously, one doesn't, until a few more rabbits are put into the hat. First, we may distinguish two classes of predetermined or independent variables: policy variables (such as tax rates, high-powered money, the exchange rate [in a regime of fixed exchange rates], and the rate of interest charged on N.H.A. mortgages) and other predetermined variables (such as the rate of unemployment in the U.S.A., the rate of economic depreciation on machinery and equipment in agriculture, past values of some of the dependent variables and a dummy variable representing strike conditions in the steel industry). In principle, policy independent variables are under the control of Canadian policy-makers, while non-policy predetermined variables are not.[1] Let z_t, $t = 1, 2, \ldots, k_1$, be the set of policy independent variables, while the notation of Appendix 1, x_j, $j = 1, 2, \ldots, k_2$, is retainted for the other predetermined variables of the model.[2] Then we may rewrite the typical equation of CANDIDE Model 1.0 (Equation (I) of the preceding appendix) as:

$$y_i = f_i(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n; \quad z_1, z_2, \ldots, z_{k_1}; \quad x_1, x_2, \ldots, x_{k_2}),$$
$$i = 1, 2, \ldots, n.^3 \qquad \text{(X)}$$

The equations of the econometric model can then be viewed as objective conditions that serve as constraints on the process of formulating suitable economic policies.

Next, we note that, if we wish to interpret "suitable economic policies" as optimizing economic policies, we must introduce a criterion function. Let the variable U denote the economic component of social well-being; this variable is generally regarded as an ordinal variable or as being measured only up to a monotonic transformation. We shall suspend disbelief and assume that a typical Canadian policy-maker is able to specify the economic component of social well-being (sometimes termed "social welfare" for short) as a "well-behaved" function of the variables of the system. In particular, it is typically assumed that social welfare depends principally on a few key endogenous variables of the system, such as constant-dollar gross national expenditure, the rate of unemployment, and some price index (possibly in the form of a rate of change). In particular, without loss of generality, we shall assume that U depends upon the first n^* endogenous variables, where $n^* \leqslant n$.[4] One can argue that the level of social welfare depends secondarily on how intensely one utilizes the policy variables themselves, e.g., politicians may feel that the electorate will hold it against them if the rate of personal income taxation is too high. Again without loss of generality, assume that U depends on the first k^* policy variables, where $k^* \leqslant k_1$. Then we can write down an explicit social welfare function (as viewed by a typical Canadian policy-maker) as:

$$U = \phi(y_1, y_2, \ldots, y_{n^*}; \quad z_1, z_2, \ldots, z_{k^*}) \qquad \text{(XI)}$$

The assumption that the social welfare function ϕ is assumed to be "well-behaved" means (among other things) that it may be assumed to possess continuous partial derivatives.[5] With this by way of introduction, we may state our optimization problem very simply: for a particular year, we seek to maximize U, as given in equation (XI), subject to the n equations of constraint (X). Hence, in this view of the problem, the equations of the model serve as objective factors (constraints) limiting the ability of the policy-makers to raise social welfare indefinitely by manipulating the policy variables.

Before solving this problem (formally) in general, we may present a solution (which will also be a formal one) in a special case. In particular, consider the case in which the f_i's of system (X) are linear equations, at least in the jointly dependent variables.[6] Then it is well known that, provided the matrix of coefficients on the jointly dependent variables is nonsingular, we may solve for each of the dependent variables in terms of all of the predetermined variables:

$$y_i = g_i(z_1, z_2, \ldots, z_{k_i}; \quad x_1, x_2, \ldots, x_{k_2}),$$
$$i = 1, 2, \ldots, n. \qquad \text{(XII)}$$

Substituting these equations (termed "reduced form equations" in the literature of econometrics) into the criterion function ϕ, we may express U as a function of all k_1 of our policy variables:

$$U = \phi(g_1, g_2, \ldots, g_{n^*}; \quad z_1, z_2, \ldots, z_{k^*})$$
$$= \psi(z_1, z_2, \ldots, z_{k_1}). \qquad \text{(XIII)}$$

[11] In the usual development of the simplex method (e.g., in David Gale, *The Theory of Linear Economic Models* (New York, Toronto, London: McGraw-Hill Book Company, 1960), especially pp. 97–121), the technique is developed under a "nondegeneracy hypothesis." The assumption of nondegeneracy is not very restrictive and in any case is not a necessary condition for obtaining a solution to the original linear programming problem. Thus, Gale develops (pp. 124–127) a generalized simplex method, based on lexicographical ordering, which will solve any feasible linear programming problem. However, he points out, in real life this sophisticated technique is rarely used. This is so because a "cycling" (the appearance and reappearance of a given set of bases without attaining an optimum) under the use of the standard replacement rules almost never occurs in practice, and so the standard simplex method will do for the overwhelming majority of cases, even when there is degeneracy and so (in principle) there could be trouble. (In his book, Gale points out that, to date, the only linear programs with degeneracy that were found to cycle were examples specially constructed to do so. More recently, in the past several years, a case of a degenerate linear programming problem that cycled when the simplex algorithm was applied to it has been discovered "in nature," according to Professor Herbert E. Scarf of Yale University.)

[1] Of course, as Professor Dobell has pointed out, the borderline may well be fuzzy in practice. It is equally true that some of the uncontrollable predetermined variables (e.g., the U. S. unemployment rate) may be under the control (at least indirectly) of some foreign policy-makers. In the case of lagged values of the endogenous variables, these are clearly uncontrollable, in the context of current policy. Every decision-maker in the world is a prisoner of history, both his own and everyone else's!

[2] In terms of Appendix 1, we should have: $k_1 + k_2 = k$.

[3] The definition of terms are the same (aside from the exceptions noted) as in Appendix 1.

[4] In general, we should expect that n^* will be considerably smaller than n. Thus, if n is equal to 1527, it would be surprising if n^* exceeded 20.

[5] One may inquire whether social welfare depends on the other predetermined variables, the x_t's. In general, this may well be the case. However, since the other predetermined variables are under neither the direct nor the indirect control of Canadian policy-makers, there is nothing that can be done about them, in the present context. Accordingly, the effects of these other predetermined variables may be impounded in the social welfare function ϕ, rather than being represented explicitly.

[6] The critical consideration is that the system of equations be linear in the y_i's, the set of jointly dependent variables. Nonlinearities in either the z_t's or the x_t's are quite unimportant for these purposes, so long as the functions remain "well-behaved."

(Again, we may impound the effects of the noncontrollable predetermined variables into our derived criterion function ψ.) We may note that we have now reduced optimization problem to one of obtaining an unconstrained maximum of the derived ψ function. Because we have assumed that all of our functions are "well-behaved," we may solve our problem by writing down the necessary conditions for a relative maximum at an interior point of the domain of definition of ψ.[7] These conditions are:

$$\frac{\partial \psi}{\partial z_t} = 0, t = 1, 2, \ldots, k_1. \qquad \text{(XIV)}$$

Since we have roughly 57 policy variables in CANDIDE Model 1.0, this problem is a fairly tractable one, especially when compared to those discussed below.

Next, consider the case in which the equations of the model, system (X) above, are nonlinear, which of course is the situation for CANDIDE Model 1.0. In this case, we cannot, in general, obtain reduced form equations. Nevertheless, the optimization problem will have a formal solution, provided our functions are well-behaved. We may introduce n Lagrange multipliers λ_i, $i = 1, 2, \ldots, n$, and seek to maximize U of Equation (XI) subject to system (X) as the equations of constraint. Hence, we set up the Lagrangian function L as:

$$L = \phi(y_1, y_2, \ldots, y_{n^*}; \quad z_1, z_2, \ldots, z_{k^*}) + \sum_{i=1}^{n} \lambda_i (y_i - f_i). \qquad \text{(XV)}$$

Necessary conditions for an interior maximum are obtained by differentiating L partially with respect to the $2n + k_1$ variables of the problem and then setting these partial derivatives equal to zero:

$$\frac{\partial \phi}{\partial y_t} - \sum_{\substack{i=1 \\ i \neq t}}^{n} \lambda_i \frac{\partial f_i}{\partial y_t} + \lambda_t = 0, t = 1, 2, \ldots, n;$$

$$\frac{\partial \phi}{\partial z_t} - \sum_{i=1}^{n} \lambda_i \frac{\partial f_i}{\partial z_t} = 0, t = 1, 2, \ldots, k_1; \qquad \text{(XVI)}$$

$$y_i - f_i = 0, i = 1, 2, \ldots, n.$$

(In the first $n + k_1$ equations, some [indeed most, in a usual problem of this sort] of the partial derivatives of ϕ with respect to the y_t's and the z_t's will be zero, as most of the jointly dependent variables and also most of the policy variables do not enter directly into the social welfare function.) We must, of course, check for a global maximum, both among interior points found by solving system (XVI) above and also among permissible boundary solutions.[8] Moreover, since $2n + k_1$ is much larger in general than k_1 (roughly 3100 contrasted to 57 in CANDIDE Model 1.0), the inability to obtain reduced forms in general in the nonlinear case has pushed us into a much larger computational problem.

Having solved the problem in a formal sense, we may indicate why the solution is at best a formal one, rather than one capable of implementation. On the side of the constraints (the equations of the econometric model), there are several major difficulties. First, the non-policy predetermined variables (the x_j's, $j = 1, 2, \ldots, k_2$) include lagged values of the jointly dependent variables.[9] This makes no difference to the solution of the single-period optimization problem but it does mean that a multi-period optimization is not simply a collection of single-period optimization problems for each of the years comprising the totality of periods under consideration. Moreover, it can be argued that economic policy should focus on more than an optimization for a single year; the performance of the economy five and ten years into the future is (or should be) important to policy-makers, also.[10] For example, it is sometimes argued that a rapid return to full employment starting with an economy in recession is not desirable, because one might then be sowing the seeds of future inflation. But if we wish to solve a multi-period optimization problem, we must take time interdependencies into account; this would presumably be done by solving a dynamic programming problem. But each year that one adds in dynamic programming problem adds to the dimensions of the problem; for example, if we wish to optimize over the twelve years running from 1974 through 1985, the scale of the problem to be solved would be roughly twelve times as large as that of the single-period problem. But a problem with approximately 37,000 nonlinear equations is a big problem in anyone's books. Nevertheless, there are still further difficulties. Up to this point, we have treated the mathematical equations that represent the constraints as though they were exact (non-random). In point of fact, they are stochastic (random), and optimal policy should presumably take this into account. In the approach outlined above, this is in fact done in an elementary way, namely we set each random variable equal to its expected value. Nevertheless, since instability and disappointed expectations themselves have welfare costs, ideally a better method of taking account of the random nature of the constraints should be utilized.[11] Finally, we may note that the true values of the parameters of the equations of constraint are not known; all we have are estimates based on certain statistical techniques.[12] Again, policy that is fully optimal should take this consideration into account. In particular the possibility that imprecise estimates give rise to erroneous prescriptions of policy should be taken into account, at least ideally.

Even after consideration of possible problems with the constraints, I should be quite clear that I regard the most serious difficulties as arising from the social welfare function ϕ.

[7]Of course, we must check to be certain that we have really found the maximum over the domain of definition. In particular, we should check the matrix of second partial derivatives; if it is negative definite at the relevant point (points), we have obtained a relative maximum. If more than one relative maximum is obtained, the criterion function should be evaluated to find the *maximum maximorum*. Finally, the boundary of the domain of definition should be examined, in order to guard against possible "corner solutions" (maximum positions on the boundary).

[8]In this regard, it may be noted that a sufficient condition for a local maximum is that the matrix of second partial derivatives of the function ϕ be negative definite at an interior point satisfying system (XVI). Of course, it is a fairly stringent sufficient condition in this constrained maximization problem and in any case we are interested in the global maximum (which conceivably could occur on the boundary) rather than in local maxima per se.

[9]For that matter, several of the policy variables make an appearance in lagged form also, and this further increases the complexity of the multi-period optimization problem discussed in this paragraph. (Of course, this consideration is irrelevant to the single-period optimization problem discussed above.

[10]Thus, one might argue that the social welfare function should be reformulated to include dynamic phenomena, such as sustained rates of growth of real output, rates of inflation (rates of change of particular price indexes), and the like. This naturally increases the complexity of the problem.

[11]As an illustration of this point, one might consider a recent article by William C. Brainard, "Uncertainty and the Effectiveness of Policy," *American Economic Review, Papers and Proceedings*, Vol. LVII, No. 2 (May, 1967), pp. 411-425. Brainard shows, in the context of a fairly simple model, that the lessons from a model with complete certainty cannot be simply carried over to a model with random disturbances by merely replacing random variables by their expected values (or, more generally, by their "certainty equivalents"). Instead, at times a complete reorientation of policies will be required, and theorems derived from a world of certainty may be invalid.

[12]The fact that the equations of the system are, in general, subject to random disturbances is not unrelated to the fact that the universe parameters are in general only estimated, rather than known exactly. (If the relationships in question were exact, then we could estimate parameters precisely, with the passage of time which presumably would give us enough independent observations.) Moreover, the situation could be even worse than that outlined in the text; the relationships appearing in the model could well be misspecifications in some sense. (Thus we might have included some variable that was better excluded, or excluded some variable that was better included, or the functional form estimated might be inappropriate, from some view-point.) With actual misspecification of the relationships of a model, the possibilities for erroneous solutions for presumably optimal settings of the policy variables mount considerably.

First, as noted above (footnote 10), the appropriate social welfare function is probably dynamic (multi-period) in character. Secondly, most policy-makers (politicians and others) untrained in mathematics or economics just don't think of the world (or the impact that they wish to have on it) in this manner. In a large problem, it would be difficult enough to get a policy-maker to specify in advance what variables should enter into the social welfare function, let alone what the ratios of partial derivatives (the absolute values of which are termed by economists marginal rates of substitution) ought to be. Indeed, policy-making is often concerned with threshold effects that call into question our assumption of well-behaved social welfare functions: certain problem areas (deficits in the balance of payments, shortages of fuel oil and other energy sources, environmental effects, etc.) become problems only in certain ranges of the solution space, and their appearance may be quite precipitous.[13] Finally, in a democratic society, one must face the fact that different individuals have different views of the relative importance of the various social goals, and some method of resolving this conflict of view-points must be adopted.[14] But, if some method of compromise is used, this may well call into question the concept of a social welfare function. Even waiving aside the other difficulties, the resulting actions may well be suboptimal under the views of the social welfare function held by every member of the society.

If the problem, as posed, is incapable of a practical solution, is there a related problem which one can solve without too much difficulty? I should argue that indeed there is. Instead of setting up a social welfare function, one can simply query policymakers about their tastes at the very end of the procedure to be suggested. The procedure would be as follows. Select a time period, beginning with the next year and going as far into the future as the model is deemed reliable. Project as realistically as possible all of the non-policy independent variables; to allow for uncertainty, multiple possibilities with regard to the path of development for some of the key non-policy independent variables should be considered. Consider a number of settings of the totality of the policy variables of the model; one might possibly wish to consider, in a realistic problem, some 20 to 100 separate policy packages. Next, solve the model for all of the policy packages in the alternative settings, for all of the years up to the final year on the economic horizon. Next, present the results (probably, in a large model, with summary measures requested by the policy-maker) to a policy-maker, asking him to choose the policy package that he deems optimal, for each of the alternative settings. It may be noted that this approach to the theory of policy merely requires the policy-maker to say which path of development of the economy ("scenario"), as depicted in the model, that he prefers; in particular, he is not required to fuss with the specification of a social welfare function. Of course, he is required to have some faith that the model is an adequate description of the entity (the Canadian economy, in the case of CANDIDE) being modelled.[15] I should argue that freeing policy-makers of the necessity to specify their goals, quantitatively and qualitatively, in advance is an important step toward formulating a problem that can be solved in a practical sense. Most individuals, when confronted with a situation about which they have some knowledge and some feeling, are able to state which combination of events they prefer, at least relative to some other combination of events. Moreover, to some extent, this is the method that is used in practice by departments of the Government of Canada, in their applications of CANDIDE Model 1.0 to particular issues of economic policy.[16]

[13]The fact that policy-makers do not think in these terms is well illustrated by the following story, possibly apocryphal, that has been told about an eminent economist, since deceased, and the late Gamal Abdel Nasser, who was the Premier of Egypt at the time of his death. According to the story, the eminent economist was engaged as a consultant by the Egyptian government and, upon meeting Premier Nasser, he promptly asked him, "What is your social welfare function?" Premier Nasser immediately excused himself, called his officials to come to him, and asked in turn who in the world this fellow was. Upon being told that the gentleman was a distinguished European economist, Premier Nasser is reported to have said, "Well, never mind. That fellow can't possibly be of much use to us!" Furthermore, in my view, it would be a rash economist who could contradict the late Premier of Egypt (on this point, anyway).

[14]Anyone with an experience of family life will probably agree, upon brief reflection, that the question of who decides what the family's social welfare function looks like can be a terribly important question, in particular contexts.

[15]Of course, in the face of uncertainty (about the development of the independent variables not under the control of the policy-maker, or for that matter any uncertainty emanating from sources outlined in the preceding paragraph), the solution outlined in this paragraph is incomplete. In all likelihood, the policy package implied by the preferred path of development of the economy will differ depending upon the alternative settings of the development of the non-policy independent variables. Thus, in this case, the policy-maker will be left with the problem of reconciling these alternative packages, each of which is optimal under some circumstances. In particular, the policy-maker may implement that set of policies deemed to be most probable to be optimal; he may average these various optimal packages in some way; or he may be cautious and choose that package that is at least minimally satisfactory in all the circumstances examined. However, it must be admitted that, at this point, he is deserted by his expert and told to find his own method of selecting a policy package. (Of course, in a medium-term context, the possibility that future policy can correct, to some extent, for some unexpected surprises in the past [or even actions that are retrospectively deemed to have been mistakes] mitigates the importance of the policy decision taken at the present moment.)

[16]These ideas are not original. A similar point of view has been expressed by Thomas H. Naylor in "Policy Simulation Experiments with Macro-Econometric Models: the State of the Art," pp. 211-227 of M. D. Intriligator (ed.), *Frontiers of Quantitative Economics* (Amsterdam: North Holland Publishing Company, 1971).