Updating Input-Output Tables: A Linear Programming Approach*

Ziad Ghanem
Input-Output Division, Statistics Canada

For presentation at the conference "Input-Output and General Equilibrium: Data, Modelling, and Policy Analysis", Free University of Brussels, September 2–4, 2004, Brussels, Belgium.

Abstract
We demonstrate an approach for updating supply and use tables based on the most current national accounts information and a linear programming balancing method. Detailed patterns from the latest available tables are complemented by growth rates and controls from the target year. Linear programming techniques are subsequently used to balance these large-scale tables. This balancing method allows for substantial flexibility in controlling results. Objective function weights can be used to account for the relative reliability of different data, while certain constraints can enforce consistency of the data with historically observed patterns and relationships. Unlike the traditional RAS procedure, this method does not require prior knowledge of row and column totals.

* This paper represents the views of the author and does not necessarily reflect the opinions of Statistics Canada.

1. Introduction

This paper describes a methodology for updating the benchmark Canadian national supply and use input-output tables, which are published annually by Statistics Canada with a lag of approximately 3 years. One of the main advantages of the input-output framework is the improvement in the quality of the final estimates generated by the integration of the industry and commodity accounts. It is also doubtless that an analytical approach by subject matter experts provides the best means of ensuring the most meaningful and well-thought out integration of these accounts. However, such an approach is both labour-intensive and time-consuming. The aim of this paper is to present a computationally efficient algorithm for estimating supply and use tables that makes full use of all current data sources at their most detailed level.

Following the approach of Cameron et al (1998), the estimation of the target year tables emulates the construction of the benchmarks by separating the process into the two distinct phases of data construction and data reconciliation. Section two of the paper discusses the first phase, in which estimates of the industry (or activity) account are made independently from considerations of their consistency with the commodity account. A description is provided of the data sources and methods used in building the initial estimate for each industry and final demand category for the target period under consideration. Section three of the paper discusses the second phase, which focuses on eliminating imbalances and ensuring a complete coherence between the estimates of the industry and commodity accounts. Section four provides concluding remarks.

The balancing algorithm implemented by Cameron et al (1998) is structured around a modified RAS procedure. However, modifying a RAS procedure to account for subtotals, different bounds, and proportionality relationships between different elements within a large-scale system, quickly degenerates into very cumbersome algorithms that lack the required flexibility in controlling the balancing process. The mathematical programming approach, in combination with modern computing technology, provides a more flexible environment for dealing with such problems. However, as demonstrated by Jackson and Murray (2002) in a wide survey of alternative formulations, the focus of the mathematical programming literature has traditionally been on minimizing variation from the benchmark coefficients subject to a set of marginal totals for the target year. The model presented in this paper, while opting for the mathematical programming approach, nonetheless, maintains the focus on two stage approach of first constructing and subsequently balancing the target period transactions.

2. Data construction

Although reflecting the same conceptual framework as the 1993 SNA supply and use tables, the Canadian tables are organized in separate tables of output, input and final demand. In this paper, we maintain the Canadian presentation, which we find more suitable for our present purpose[1].

The main data sources used as the starting point are the most recent benchmark input-output tables and the following set of estimates linking the benchmark to the target period under consideration: current and constant dollar estimates of total gross output by industry, current dollar estimates of the large aggregates of income- and expenditure-based GDP, detailed estimates of international trade by commodity, wage inflation by industry, and complete sets of commodity price indices for production, imports, and final demand.

[1] The Canadian tables are valued at “modified” basic prices, i.e., the valuation of output does not include subsidies receivable on products. All references to basic prices in this paper should be interpreted to mean “modified” basic prices. Also, the Canadian IO tables are sectored and do not follow the pure industry concept. To simplify the presentation, we abstract from this secondary complication. For more details on the Canadian IO tables see Statistics Canada (1989).

The starting point in estimating updated IO tables is the construction of complete initial estimates for each industry and final demand component of an input-output table for the target period under consideration. The general approach is to use any available data sources for the target year that are conceptually consistent with the IO tables either as controls or as indicators of growth, the remainder being filled in by the incorporation of ratios and proportions from the benchmark tables.

A few notes on notation are in order before the model equations are presented. For a given variable, the subscript *t=0* indicates the benchmark year, while the absence of this subscript indicates a reference to the target year. For the target year, a horizontal bar is used to distinguish the unbalanced estimate of a variable from the balanced estimate (without a bar), which will appear in section 3. Where it is sufficiently clear from the context, the absence of a subscript indicates summation over that subscript, for example v_j is the same as $\sum_i v_{ij}$.

2.1 The Output table

The initial, unbalanced estimate of the output table is obtained from the current dollar total outputs by industry estimates and the benchmark commodity mix ratios by industry. To account for the impact of price changes on the distribution of commodity outputs, production price indexes are used to adjust the commodity mix ratios.

$$(1) \quad \bar{v}_{ij} = \left(\frac{v_{t=0,ij}}{v_{t=0,j}}\right)\left(\frac{p_i^V}{p_{t=0,i}^V}\right)v_j^C$$

where v_{ij} denotes the output of commodity *i* by industry *j*, p_i^V the output price index by commodity, and v_j^C the vector of industry gross outputs in current dollars for the target period. With *i* = 1,2,...,n, the commodity n-1 is set as the transport margin and the commodity n as the trade margin[2].

By converting equation (1) into matrix form, and after partitioning the output matrix to reflect the output of non-margin commodities $\bar{\mathbf{V}}^1$ and margin commodities $\bar{\mathbf{V}}^T$, we have:

$$\bar{\mathbf{V}} = \begin{bmatrix} \bar{\mathbf{V}}^1 \\ \bar{\mathbf{V}}^T \end{bmatrix}$$

2.2 The Input table

Let $\bar{\mathbf{U}}$ be a partitioned matrix of the initial, unbalanced estimate of the input table. We denote $\bar{\mathbf{U}}^1$ as a matrix of purchaser price intermediate inputs, of dimensions m x n, where m is the number of non-margin commodities and n the number of industries. Matrices $\bar{\mathbf{U}}^2, \bar{\mathbf{U}}^3$, and $\bar{\mathbf{U}}^4$ are of similar dimensions to $\bar{\mathbf{U}}^1$ and represent respectively transport, trade, and tax margins

[2] The Canadian tables include the production of several margin commodities which we've grouped into only two for the purpose of simplification. We have grouped under trade the retail and wholesale trade margins and under transportation, transportation services, gas distribution, pipeline transportation, and storage. Another non-produced margin, of course, is taxes on products.

associated with each cell in $\overline{\mathbf{U}}^1$. Matrix $\overline{\mathbf{U}}^T$ is a 2 x n matrix of direct purchases of transport services in the first row and trade services in the second row. Matrix $\overline{\mathbf{U}}^Y$ is the matrix of primary inputs of dimensions 8 x n. In order, from 1 to 8 respectively, these are taxes on products, subsidies on products, subsidies on production, taxes on production, net income of unincorporated businesses, wages, supplementary labour income, and other operating surplus.

$$\overline{\mathbf{U}} = \begin{bmatrix} \overline{\mathbf{U}}^1 & \overline{\mathbf{U}}^2 & \overline{\mathbf{U}}^3 & \overline{\mathbf{U}}^4 \\ \overline{\mathbf{U}}^T & 0 & 0 & 0 \\ \overline{\mathbf{U}}^Y & 0 & 0 & 0 \end{bmatrix}$$

Purchaser price intermediate inputs are calculated as the sum of their values in basic prices and their associated margins. Intermediate inputs in basic prices are estimated as a function of change in the constant dollar output by industry and the variation in the prices by commodity.

$$(2) \quad \overline{u}_{ij}^{(BP)} = u_{t=0,ij}^{(BP)} \left(\frac{v_j^K}{v_{t=0,j}^K} \right) \left(\frac{p_i^U}{p_{t=0,i}^U} \right)$$

The superscript (*BP*) denotes valuation in basic prices. Thus, $\overline{u}_{ij}^{(BP)}$ and $u_{t=0,ij}^{(BP)}$ denote the intermediate inputs in basic prices for the target and benchmark periods respectively; v_j^K are the constant dollar total industry output by industry *j*, and p_i^U are the commodity price indexes associated with demand, formed from a combination of output and import price indices.

Any changes in the prices of inputs are assumed to be outside the control of the industry and are borne by the residual item, surplus. While this will tend to under-represent the role of substitution effects and technological change, the paucity of easily usable, current, statistical data sources, makes it difficult to proceed in any other manner. However, the confrontation of supply and demand, in the ensuing balancing process, will tend to attenuate this weakness by incorporating elements of supply, which given their greater accuracy, implicitly force on the demand side any substitution effects that may have occurred.

The values for transport margins are based on benchmark margin rates relative to basic prices for each cell times the basic price value for non-margin intermediate inputs for the target period. Since the variation in the prices of intermediate inputs should not affect the transportation margins, which are a function of volume, we deflate the estimate of intermediate inputs to remove the impact of price variations. Margin values are then inflated by the price movements of the transportation services to generate the current dollar estimates.

$$(3) \quad \overline{u}_{ij}^2 = \left(\frac{u_{t=0,ij}^2}{u_{t=0,ij}^{BP}} \right) \left(\frac{p_{t=0,i}^U}{p_i^U} \right) \left(\frac{p_{l=1}^T}{p_{t=0,l=1}^T} \right) \overline{u}_{ij}^{(BP)}$$

where p^T is the output price index for margin commodities of dimensions (2 x 1), with transport services as the first element and trade services as the second element.

Trade margins are assumed to be based on mark-ups after the production and transportation stages. Thus, they are based on benchmark margin rates relative to basic prices plus transport charges for each cell. This value is then inflated by the price movements of the trade services to generate the target year estimates.

(4) $$\overline{u}_{ij}^{3} = \left(\frac{u_{t=0,ij}^{3}}{u_{t=0,ij}^{BP} + u_{t=0,ij}^{2}} \right) \left(\frac{p_{l=2}^{T}}{p_{t=0,l=2}^{T}} \right) \overline{u}_{ij}^{(BP)}$$

The values of tax margins are equal to the benchmark tax rate as a function of the purchase price minus the taxes[3].

(5) $$\overline{u}_{ij}^{4} = \left(\frac{u_{t=0,ij}^{4}}{u_{t=0,ij}^{1} - u_{t=0,ij}^{4}} \right) (\overline{u}_{ij}^{(BP)} + \sum_{k=2}^{3} \overline{u}_{ij}^{k})$$

Having estimated the required components in equations (2) to (5), it is now possible to calculate the purchaser price input table as the sum of the basic price inputs and all their associated margins.

(6) $$\overline{u}_{ij}^{1} = \overline{u}_{ij}^{(BP)} + \sum_{k=2}^{4} u_{ij}^{k}$$

Some margin services are purchased directly. Similarly to other intermediate inputs, their values are determined as a function of movements in volume and prices.

(7) $$\overline{u}_{ij}^{T} = u_{t=0,ij}^{T} \left(\frac{v_{j}^{K}}{v_{t=0,j}^{K}} \right) \left(\frac{p_{i}^{T}}{p_{t=0,i}^{T}} \right)$$

The primary inputs, composed of taxes on production, subsidies on products and production, and net income of unincorporated businesses are based on their benchmark ratios to total output.

(8) $$\overline{u}_{ij}^{Y} = \left(\frac{u_{t=0,ij}^{Y}}{v_{t=0,j}} \right) v_{j}^{C} \qquad (\text{i} = 2,3,4,5)$$

Similarly to intermediate inputs, wages and supplementary labour income by industry move with volume and wage inflation, however, unlike commodity prices, wage inflation is available with industrial detail.

(9) $$\overline{u}_{ij}^{Y} = u_{t=0,ij}^{1} \left(\frac{v_{j}^{K}}{v_{t=0,j}^{K}} \right) \left(\frac{p_{j}^{L}}{p_{t=0,j}^{L}} \right) \qquad (\text{i} = 6,7)$$

where p_{j}^{L} denotes the index of wage inflation by industry *j*.

Surplus by industry is derived residually as the difference between total industry output and the sum of all calculated inputs.

[3] Again, in the Canadian IO tables further detail for tax margins is available. Taxes are estimated by type. Here we simplify again into one tax.

$$(10)\ \bar{u}^{Y}_{8j} = \bar{v}_j - \bar{u}^{1}_{j} - \bar{u}^{T}_{j} - \sum_{i=2}^{7} \bar{u}^{Y}_{ij}$$

2.3 The Final Demand matrix

Similarly to the input matrix, we set up the initial unbalanced estimate of the final demand table, $\overline{\mathbf{F}}$, as a partitioned matrix of purchaser price final expenditures, margin values associated with these purchases, and primary inputs (in this case, only indirect taxes have non-zero values), except that now final demand activities replace the industries as the columns.

$$\overline{\mathbf{F}} = \begin{bmatrix} \overline{\mathbf{F}}^1 & \overline{\mathbf{F}}^2 & \overline{\mathbf{F}}^3 & \overline{\mathbf{F}}^4 \\ \overline{\mathbf{F}}^T & 0 & 0 & 0 \\ \overline{\mathbf{F}}^Y & 0 & 0 & 0 \end{bmatrix}$$

Each matrix within the $\overline{\mathbf{F}}$ matrix is further partitioned into column vectors that reflect the distinct activities of final demand.

$$\overline{\mathbf{F}}^k = \begin{bmatrix} \overline{\mathbf{c}}^{Rk} & \overline{\mathbf{c}}^{Mk} & \overline{\mathbf{c}}^{Xk} & \overline{\mathbf{G}}^k & \overline{\mathbf{K}}^k & \overline{\mathbf{s}}^k & \overline{\mathbf{n}}^{Pk} & \overline{\mathbf{n}}^{Nk} & \overline{\mathbf{x}}^k & \overline{\mathbf{r}}^k & \overline{\mathbf{m}}^k \end{bmatrix} \qquad (\mathrm{k} = 1,2,3,4)$$

where $\overline{\mathbf{c}}^{Rk}$ is domestic consumption expenditures by residents, $\overline{\mathbf{c}}^{Mk}$ is consumption by residents abroad, i.e., travel imports, $\overline{\mathbf{c}}^{Xk}$ is consumption expenditure by non-residents, i.e. travel exports, $\overline{\mathbf{G}}$ are government current expenditures by level of government, $\overline{\mathbf{K}}$ are capital expenditures by industry, $\overline{\mathbf{s}}$ is scrap metal, $\overline{\mathbf{n}}^P$ is inventory additions, $\overline{\mathbf{n}}^N$ is inventory withdrawals, $\overline{\mathbf{x}}$ is domestic exports excluding the travel exports, $\overline{\mathbf{r}}$ is re-exports, and $\overline{\mathbf{m}}$ is international imports; $\overline{\mathbf{s}}, \overline{\mathbf{n}}^N$, and $\overline{\mathbf{m}}$ are negative values.

In the benchmark tables, consumption expenditures of non-residents are included with both domestic consumption expenditures and international exports. To remove the double counting and allow the calculation of personal expenditures by residents, a negative vector of the personal expenditures of non-residents is added in final demand. For the purposes of simplification, this duplication is eliminated. Domestic expenditures of residents and international exports are here shown net of those by non-residents and the expenditures of non-residents appear only once, as a positive value.

As distinct from other final demand categories, the complete commodity detail in purchaser prices for trade data, i.e. vectors $\overline{\mathbf{c}}^{M1}, \overline{\mathbf{c}}^{X1}, \overline{\mathbf{x}}^1, \overline{\mathbf{r}}^1$, and $\overline{\mathbf{m}}^1$, from both international merchandise trade and balance of payments statistics, is available for the target year.

In general, the final demand matrix is constructed based on the expenditure-based GDP aggregates for the target period. Within each aggregate, detail is derived from the benchmark commodity mix ratios. The table is estimated directly in purchaser prices.

Consumer expenditures are allocated based on the benchmark commodity mix and changes in consumer prices.

$$(11)\ \bar{c}_i^{R1} = \sum_j \left(\left(\frac{c_{t=0,ij}^{R1}}{\sum_i c_{t=0,j}^{R1}} \right) \left(\frac{p_i^C}{p_{t=0,i}^C} \right) c_j^* \right)$$

where c_j^* are the target year consumption expenditures by residents of which we have *j* commodity groupings and p_i^C is the consumer price index by commodity *i*.

$$(12)\ \bar{g}_{ij}^1 = \left(\frac{g_{t=0,ij}^1}{g_{t=0,j}^1} \right) g_j^*$$

where g_j^* denotes the target year total government current expenditures for the *j* levels of government.

Capital expenditures don't have an industrial distribution in the target year GDP aggregate. However, timely data are available from the Capital Expenditure Survey (CES). To derive an industrial distribution for capital expenditures, the growth rate by industry from the CES is applied to the benchmark industries.

$$(13)\ \bar{k}_j^S = \left(\frac{k_j^S}{k_{t=0,j}^S} \right) k_{t=0,j}^1$$

where k_j^S is the value of capital expenditures from the CES by industry *j*.

The capital expenditure aggregate is subsequently allocated based on the new industrial distribution available from equation (13). The benchmark commodity mixes by industry are applied to this estimate of total investments by industry.

$$(14)\ \bar{k}_{ij}^1 = \left(\frac{k_{t=0,ij}^1}{k_{t=0,j}^1} \right) \left(\left(\frac{\bar{k}_j^S}{\sum_j \bar{k}_j^S} \right) k^* \right)$$

where k^* denotes total capital expenditures.

$$(15)\ \bar{s}_i^1 = \left(\frac{s_{t=0,i}^1}{s_{t=0}^1} \right) s^*$$

where s^* is the total estimate for scrap metal.

Estimates of inventories for the target year are available at a high level of aggregation. Since inventories can shift between withdrawals and additions, the allocation pattern for the required detail is based on the both positive and negative inventories.

$$(16)\ \bar{n}_i^P = \sum_j \left(\left(\frac{n_{t=0,ij}^P + \left| n_{t=0,ij}^N \right|}{n_{t=0,j}^P + \left| n_{t=0,j}^N \right|} \right) \mathrm{n}_j^{P*} \right)$$

$$(17)\ \bar{n}_i^N = \sum_j \left(\left(\frac{n_{t=0,ij}^P + \left| n_{t=0,ij}^N \right|}{n_{t=0,j}^P + \left| n_{t=0,j}^N \right|} \right) \mathrm{n}_j^{N*} \right)$$

where n_j^{P*} and n_j^{N*} are the target year inventory additions and withdrawals of which we have *j* commodity aggregates.

The values for trade and tax margins are based on benchmark margin rates relative to purchaser price for each cell times the purchaser price estimate for the target period.

$$(18)\ \bar{f}_{ij}^k = \left(\frac{f_{t=0,ij}^k}{f_{t=0,ij}^1} \right) \bar{f}_{ij}^1 \qquad (\mathrm{k} = 3,4)$$

The transportation margins are based on the margin rates relative to the purchaser price after accounting for any variation in the price of the commodity being transported and the price variation in transportation services.

$$(19)\ \bar{f}_{ij}^2 = \left(\frac{f_{t=0,ij}^2}{f_{t=0,ij}^1} \right) \left(\frac{p_{t=0,i}^F}{p_i^F} \right) p_{l=1}^T \bar{f}_{ij}^1$$

where p^F is the price index in purchaser for final demand commodities.

The only occurrences of direct purchases of margin commodities appear in the trade vectors as imports and exports of transportation services. Thus, target year estimates of $\overline{\mathbf{F}}^T$ are available from the balance of payments data. Similarly for $\overline{\mathbf{F}}^Y$, the only data elements in this matrix are the indirect taxes on imports, i.e., the import duties, and these also are available for the target year in a timely fashion from international trade data.

3. Balancing the tables

While our primary focus is on achieving a quick and efficient solution, practical experience suggests that it is a worthwhile effort to invest some time in analyzing and adjusting some of the large commodity imbalances that one observes in the initial estimate prior to applying the balancing algorithm. Some of the large inconsistencies may uncover easily correctable errors in the raw data such as obvious misclassifications, illegitimate births or deaths, etc.

The following linear programming approach balances an input-output table while minimizing differences from the initial unbalanced estimate. The system is balanced, i.e., the values are forced to comply with all input-output identities, while respecting a series of aggregates, a range of allowable bounds for the variations in each cell, and a set of predetermined proportionality relationships among different data points. We avoid non-linear relationships mainly for reasons of computational efficiency.

3.1 The objective function

We wish to estimate a balanced version of matrices $\overline{\mathbf{V}}, \overline{\mathbf{U}}, \text{and } \overline{\mathbf{F}}$, which we will denote as $\mathbf{V}, \mathbf{U}, \text{and } \mathbf{F}$. In a similar manner to Matuszewski et al (1964), we set a linear objective function whose aim is to ensure that the balanced matrices "differ as little as possible" from the unbalanced IO tables. However, as distinct from Matuszewski et al (1964), our focus is on differences in the transactions of the target year and not on coefficients relative to the benchmark.

(20) Minimize $Z = \mathbf{i}'(\mathbf{\Phi} \bullet (\mathbf{V}^P + \mathbf{V}^N))\mathbf{i} + \mathbf{i}'(\mathbf{\Psi} \bullet (\mathbf{U}^P + \mathbf{U}^N))\mathbf{i} + \mathbf{i}'(\mathbf{\Omega} \bullet (\mathbf{F}^P + \mathbf{F}^N))\mathbf{i}$

where the *P* and *N* superscripts represent the positive and negative variations respectively associated with each element in the matrices $\overline{\mathbf{V}}, \overline{\mathbf{U}}, \text{and } \overline{\mathbf{F}}$, and $\mathbf{\Phi}, \mathbf{\Psi}, \text{and } \mathbf{\Omega}$ are exogenously set weights associated with each element; the **i**'s are identity vectors of appropriate lengths, used to sum the matrices.

While all values in the objective function are positive, the variation in each element of the IO tables is decomposed into its positive and negative components. This is done in order to generate a linear objective function; a necessary precondition for solving large-scale problems given the limitations of available software.

In general, the weights in Equation (20) are set in accordance with the perceived relative quality of the different elements. For example, intermediate inputs may be given a weight of 1, while outputs and international trade, which may be considered more reliable, are given a weight of 1.1, and inventories, which are considered very weak, a value of 0.9. This allows the system to resolve imbalances while respecting the relative strengths of output and trade data, while inventories are used relatively freely, within their allowable bounds, in solving commodity imbalances. The weights of margin cells are set equal to the weights of their associated purchaser price cells. This is necessary to ensure that no bias is introduced into the balancing process through differing treatments between the margins and the purchaser price value of commodities whenever adjustments based on the supply-demand identity are required.

The objective function is minimized subject to the constraints listed below in equations (21) to (44).

3.2 The relationship between starting values and the values to be estimated

(21) $\mathbf{V} = \overline{\mathbf{V}} + \mathbf{V}^P - \mathbf{V}^N$

(22) $\mathbf{U} = \overline{\mathbf{U}} + \mathbf{U}^P - \mathbf{U}^N$

(23) $\mathbf{F} = \overline{\mathbf{F}} + \mathbf{F}^P - \mathbf{F}^N$

(24) $\mathbf{V}, \mathbf{V}^P, \mathbf{V}^N, \mathbf{U}, \mathbf{U}^P, \mathbf{U}^N, \mathbf{F}, \mathbf{F}^P, \mathbf{F}^N \geq 0$

In equations (21) to (23), each non-zero variable to be estimated by the model is set equal to the initial estimate plus an amount of positive variation and minus an amount of negative variation. Since these variations appear in the objective function, this ensures that the distance between the initial estimates and the variable to be estimated is maintained at a minimum. This minimization, of course, is subject to the weighting assigned in the objective function and the model constraints. Equation (24) ensures that all variables have a non-negative value.

3.3 Industry balance

Let $\breve{\mathbf{U}}$ be all the inputs into each industry, a partitioned matrix of the inputs in purchaser price, the inputs of margins, and all primary inputs.

$$\breve{\mathbf{U}} = \begin{bmatrix} \mathbf{U}^1 \\ \mathbf{U}^T \\ \mathbf{U}^Y \end{bmatrix}$$

For each industry, the sum of total outputs must equal the sum of total inputs.

(25) $\mathbf{i}'\mathbf{V} - \mathbf{i}'\breve{\mathbf{U}} = 0$

This implies that total industry outputs are endogenous, i.e. they are allowed to deviate from the initial estimate.

3.4 Commodity balance

For each non-margin commodity, supply is defined as the sum of its outputs, imports, inventory withdrawals, scrap, and margins, while demand is equal to the purchaser price value of the sum of its intermediate uses, current expenditures of the personal sector and government, capital expenditures, inventory additions and re-exports.

For convenience, a subset of final demand is defined as final expenditures by residents, $\overline{\mathbf{F}}^{Rk}$.

$$\overline{\mathbf{F}}^{Rk} = \begin{bmatrix} \overline{\mathbf{c}}^{Rk} & \overline{\mathbf{c}}^{Mk} & \overline{\mathbf{G}}^{k} & \overline{\mathbf{K}}^{k} & \overline{\mathbf{n}}^{k} \end{bmatrix} \qquad (\text{k} = 1,2,3,4)$$

(26) $$\begin{aligned} &\mathbf{V}^1\mathbf{i} - \mathbf{m}^1 - \mathbf{n}^{N1} - \mathbf{s}^1 + (\sum_{k=2}^{4}\mathbf{U}^k)\mathbf{i} + (\sum_{k=2}^{4}\mathbf{F}^{Rk})\mathbf{i} + \sum_{k=2}^{4}(\mathbf{c}^{Xk} + \mathbf{x}^k + \mathbf{r}^k) \\ &= \mathbf{U}^1\mathbf{i} + \mathbf{F}^{R1}\mathbf{i} + \mathbf{c}^{X1} + \mathbf{x}^1 + \mathbf{r}^1 \end{aligned}$$

3.5 The bounds on variables

Let $\breve{\mathbf{F}}$ be the matrix of all final demand expenditures in purchaser prices, direct expenditures on margin commodities and primary inputs.

$$\breve{\mathbf{F}} = \begin{bmatrix} \mathbf{F}^1 \\ \mathbf{F}^T \\ \mathbf{F}^Y \end{bmatrix}$$

(27) $\underline{\alpha}_{ij} v_{ij} \le v_{ij} \le \overline{\alpha}_{ij} v_{ij}$

(28) $\underline{\beta}_{ij} u_{ij} \le \breve{u}_{ij} \le \overline{\beta}_{kj} u_{ij}$

(29) $\underline{\gamma}_{ij} f_{ij} \le \breve{f}_{ij} \le \overline{\gamma}_{ij} f_{ij}$

where α, β, and γ represent lower and upper bound scaling factors. The bounds are exogenously set based on two factors: the perceived strengths of the data and the starting commodity imbalances. For example, the manufacturing industries are treated more restrictively than other industries due to the perceived higher quality of their industry outputs and price indices. By setting the bounds on the variables as a function of the starting imbalance, we also force the solutions to avoid perturbing well behaved elements of the system.

For margins, we are more concerned with the stability of the rates rather than the actual values. Thus, we define margin rates for the input and final demand matrices in equations (30) and (31). For each element in both matrices, the transport margin rate is a ratio of transport to basic price values. The trade margin rate is a ratio of trade to basic prices plus transport margin values. The tax margin rate is a ratio of tax to the sum of basic price, transport and trade values.

(30) $$\tilde{u}_{ij}^{l} = \frac{\bar{u}_{ij}^{l}}{\bar{u}_{ij}^{1} - \sum_{k=l}^{4} \bar{u}_{ij}^{1}} \qquad (l = 2,3,4)$$

(31) $$\tilde{f}_{ij}^{l} = \frac{\bar{f}_{ij}^{l}}{\bar{f}_{ij}^{1} - \sum_{k=l}^{4} \bar{f}_{ij}^{1}} \qquad (l = 2,3,4)$$

We now proceed to define the bounds on the margins based on the derived rates. The bounds on the value of each margin are defined as a function of the initial margin rate times the appropriately valued estimate of intermediate use and final demand and a scaling factor to allow for some adjustment in the rates.

(32) $$\underline{\delta}_{ij}^{l} \tilde{u}_{ij}^{l} (u_{ij}^{1} - \sum_{k=l}^{4} u_{ij}^{l}) \leq u_{ij}^{l} \leq \bar{\delta}_{ij}^{k} \tilde{u}_{ij}^{l} (u_{ij}^{1} - \sum_{k=l}^{4} u_{ij}^{l}) \qquad (l = 2,3,4)$$

(33) $$\underline{\xi}_{ij}^{l} \tilde{f}_{ij}^{l} (f_{ij}^{1} - \sum_{k=l}^{4} f_{ij}^{l}) \leq f_{ij}^{l} \leq \bar{\xi}_{ij}^{k} \tilde{f}_{ij}^{l} (f_{ij}^{1} - \sum_{k=l}^{4} f_{ij}^{l}) \qquad (l = 2,3,4)$$

where $0 < \underline{\delta}_{ij}^{l}, \underline{\xi}_{ij}^{l} < 1$ and $\bar{\delta}_{ij}^{l}, \bar{\xi}_{ij}^{l} > 1$ are exogenously determined scaling factors meant to allow small departures from the initial rates. The bounds on the tax rates, however, must be treated much more conservatively than other margin rates, since by their very nature they are extremely stable. The margin values are therefore endogenously determined in the balancing process, while the margin rates are allowed to vary by up to a certain percentage in either direction. This might go too far in allowing each cell-specific margin to vary, the benchmark tables tend to maintain the same margin rate across whole categories of commodities.

Giving each margin rate an upper bound that exceeds the benchmark rate creates the possibility that the sum of all margins may equal or even exceed the purchaser price value for a given cell. To safeguard against this possibility, equations (34) and (35) state that the sum of all margins cannot exceed some portion of the purchaser price value.

(34) $$(1 - \eta_{ij}^{U}) u_{ij}^{1} > \sum_{k=2}^{4} u_{ij}^{k}, \quad 0 < \eta_{ij}^{U} < 1$$

(35) $$(1 - \eta_{ij}^{F}) f_{ij}^{1} > \sum_{k=2}^{4} f_{ij}^{k}, \quad 0 < \eta_{ij}^{F} < 1$$

where the η_{ij} 's are either arbitrarily set very small values or a number which is set as a function of historically observed lower limits of the basic price as a proportion of the purchaser price for each cell.

3.6 Supply and demand of margins

The total output of the trade and transport margin commodities must equal the direct purchases of margin commodities by all activities as well as the margins incorporated in the purchaser price value of all other non-margin commodities.

$$(36)\ \sum_j v_{1j}^T = \sum_j u_{1j}^T + \sum_j \sum_i u_{ij}^2 + \sum_j f_{1j}^T + \sum_j \sum_i f_{ij}^2$$

$$(37)\ \sum_j v_{2j}^T = \sum_j u_{2j}^T + \sum_j \sum_i u_{ij}^3 + \sum_j f_{2j}^T + \sum_j \sum_i f_{ij}^3$$

3.7 International trade

The model must also respect certain basic relationships between domestic supply and demand and international trade.

Each commodity export, net of margins, must be less or equal to domestic supply (output, inventory withdrawals, and scrap metal).

$$(38)\ \mathbf{x}^1 + \mathbf{c}^{X1} - \sum_{k=2}^{4} (\mathbf{x}^k + \mathbf{c}^{Xk}) \le \mathbf{V}^1\mathbf{i} - \mathbf{n}^{N1} - \mathbf{s}^1$$

Imports must be less than or equal to the sum of domestic demand at basic prices and re-exports.

$$(39)\ -\mathbf{m}^1 \le (\mathbf{U}^1 - \sum_{i=2}^{4} \mathbf{U}^i)\mathbf{i} + \mathbf{F}^{D1} - \sum_{k=2}^{4} \mathbf{F}^{Dk} + \mathbf{r}^1 - \sum_{i=2}^{4} \mathbf{r}^i$$

Re-exports must be constrained as a function of imports. Therefore, as a first step, the benchmark basic price ratio of re-exports to imports is calculated in equation (40).

$$(40)\ \tilde{r}_{t=0,i}^1 = \frac{r_{t=0,i}^1 - \sum_{k=2}^{4} r_{t=0,i}^k}{m_{t=0,i}^1}$$

The benchmark ratio in equation (40) is subsequently used to define lower and upper bounds for re-exports in equation (41).

$$(41)\ \overline{\theta}\tilde{r}_{t=0,i}^1 \le r_i^1 - \sum_{k=2}^{4} r_i^k \le \underline{\theta}\tilde{r}_{t=0,i}^1$$

where the $\overline{\theta}$'s are exogenously set parameters, with $\underline{\theta}$ denoting a lower bound scalar, slightly lower than one and $\overline{\theta}$, an upper bound scalar, slightly greater than one.

3.8 The GDP aggregates

The sum across industries of each component of primary inputs, excluding the taxes on products, is constrained to the target year income-based GDP totals.

$$(42)\ \sum_j u_{ij}^{Y} = y_i^{*} \qquad (\mathrm{i} = 2,3,...,8)$$

where y_i^* is the target period, income-based GDP aggregate and the elements in the rows of y_i^* are in the same order as the rows in u^Y.

Taxes on products are the sum of the margins on commodities in input, final demand, and the duties on imports that appear in the purchaser price final demand. These are set equal to the aggregate from the target year income-based GDP.[4]

$$(43)\ \sum_j \sum_i u_{ij}^{4} + \sum_j \sum_i f_{ij}^{4} + \sum_j f_{1j}^{Y} = y_1^{*}$$

The sum of all the components of final demand must add-up to the target year expenditure-based GDP aggregates.

$$(44)\ \mathbf{iF}^{1}\mathbf{A} = \mathbf{f}^{*}$$

where $\mathbf{f}^*$ is a vector of target year expenditure-based GDP aggregates and **A** is an aggregation matrix, composed of ones and zeros, that maps the detailed final demand activities to their totals in $\mathbf{f}^*$.

4 Concluding remarks

The model presented in this paper has been implemented on the Canadian, working level input-output tables, which have dimensions of approximately 300 industries, 170 final demand categories, and 725 commodities. Using a commercially available optimization software on a P4 PC, the solution time was around 15 minutes.

The main advantages of the approach presented are the capacity to integrate all detailed target year information, the ability to direct and control the solution in accordance with qualitative judgements on the quality of the different elements, and the speed of solution time.

[4] Some small values of direct taxes appear in purchaser price tables both in the input tables. We abstract from that detail.

References

Cameron, G., D. Hayes, and A. Loranger, (1998). *A Computer-Intensive Approach to Input-Output Table Construction: Initial Experiments and Results*, Input-Output Division, Statistics Canada, Ottawa, mimeo.

Jackson W. R. and A. T. Murray (2002). *Alternate Formulations for Updating Input-Output matrices*, http://www.rri.wvu.edu/pdffiles/jackson2002-9wp.pdf.

Matuszewski, T., P. R. Pitts and J. A. Sawyer (1964). “Linear Programming Estimates of Changes in Input Coefficients,” *Canadian Journal of Economics and Political Science*, **30**, pp. 203-210.

Statistics Canada (1984). *A User Guide to the Canadian System of National Accounts*, Catalogue no. 13-589E Occasional, Ottawa, Statistics Canada.