Estimating Commodity Inflows to a Sub-State Region Using Input-Output Data: Accuracy Tests Using the Commodity Flow Survey

Louie Nan Liu
College of Management
Zhejiang Universoty, Hangzhou
People's Republic of China

Pierre Vilain
Vice President, Econsult Corporation
Philadelphia, Pennsylvania
USA

Paper Presented to the 13th International Conference on Input-Output Techniques
August 22, 2000

Abstract: *We describe a methodology to estimate current commodity inflows to a sub-state region using a supply-side, commodity-by-industry input-output model and commodity flow data for American states. Since the 1993 Commodity Flow Survey does not go below the state level, the estimation of commodity flows to a particular sub-state region in the United States has always proven difficult. By combining state-level commodity flow data with the supply-side commodity-by-industry input-output model, an estimate of commodity flows to smaller regions can be carried out entirely based on the regional industrial structure. Since the actual sub-state flows are typically unobserved, the accuracy of the methodology is unknown. However, by applying the same methodology to larger regions, with actual states used as the forecast region, the estimates can be compared to actual flows. We carry out this test of the methodology's accuracy, and report the results.*

INTRODUCTION

A typical problem faced by transportation planners is estimating freight flows between regions. This may be required to estimate the freight-bearing commercial traffic, or also infer the potential demand for a new type of transportation service such as air cargo. In the United States, estimates of freight flows exist between individual states[1], but little data exists for flows between areas below the state level, which we refer to as *sub-state regions*. With respect to freight *outflows* from sub-state regions, one can arguably produce relatively accurate estimates, given some adequate data on the region's industrial structure and the state-to-state trade data. By mechanically assigning freight commodity exports to the producing industries, one can estimate the share of a sub-state region's state exports based on its presence of these industries.

However, the estimation of the second category, air cargo *inflows,* is considerably more complicated, as pointed out by Ortúzar and Willumsen (1994). While we can roughly assign the production of commodities to certain industries, the consumption of commodities by various industries requires far more detailed knowledge of their input use. Fortunately, this type of information is readily available in input-output models, and some simple manipulations of standard input-output data yields a tool that can be then used to assign state-level commodity inflows to any sub-state region.

In the following sections we outline a methodology to estimate commodity inflows to smaller regions, which was initially described in Vilain et. al (1999). The methodology was devised specifically to regionalize inflows to sub-state regions in the United States, but it has also been used in other countries as well. In general, the methodology can be used in any country or region, given the availability of commodity-by-industry input-output accounts.

---

1 Existing data on commodity flows contained in the United States Department of Commerce's *1993 Commodity Flow Survey,* are rich in commodity-level detail but are not detailed to below the state level.

Having proposed a methodology to estimate freight inflows that is simple to use, it is of interest to examine the *accuracy* of the technique. In this paper, we carry out a series of simulations that we then test for their predictive accuracy. The key to being able to determine the accuracy of simulations is to carry them out for states, as if these were smaller sub-state regions. Since states are regions for which commodity flow data does exist, we can then compare the predicted inflows to the actual observed inflows. Our results show that, excluding inflows of mining, petroleum or coal products, the methodology leads to relatively accurate forecasts. Total inflows of all commodities to a state are typically predicted within 8 percent error, while the accuracy of forecasts for individual commodities is far more variable. Despite the mixed results, we argue that the methodology described here is a valid one, yielding predictions of commodity inflows that have an acceptable level of accuracy. The relative accuracy of the methodology must also be considered keeping in mind that, in the absence of expensive origin-destination surveys there are really no alternatives that yield reliable estimates of commodity inflows.

THE SUPPLY-SIDE INPUT-OUTPUT MODEL AND COMMODITY FORECASTING

A possible approach to estimating commodity inflows to a region could be through the estimation of a gravity-type model. Gravity models have been applied extensively to the analysis of passenger trip generation, and examples exist of their application to freight demand modeling (Ortúzar and Willumsen (1994))[2]. A gravity model is essentially estimating the importance of factors that *attract* freight inflows to a sample of areas, such as demand factors or accessibility. In terms of our problem of predicting commodity inflows to a sub-state area, this model could be estimated at the state level, and the estimated parameters used to predict inflows to a sub-state region. However, data requirements to calibrate such a model are significant, particularly with respect to the shipping costs between states. This same conclusion applies to other, closely related, models based on discrete choice analysis, including disaggregated freight generation models.

An alternative approach is proposed, one that bases estimates of actual commodity inflows to a sub-state region entirely on the region's industrial structure. The approach relies heavily on input-output data to estimate regional commodity inflows. In a form which is less familiar to users, namely in the commodity-by-industry format, input-output data is used to *share* state-level commodity inflows to sub-state regions by directing these inflows to the industries which use them as inputs. This procedure can be carried out fairly easily, relying entirely on published national input-output data, existing state-level commodity flow data from the *1993 Commodity Flow Survey*, and regional data on employment or earnings by industry.

---

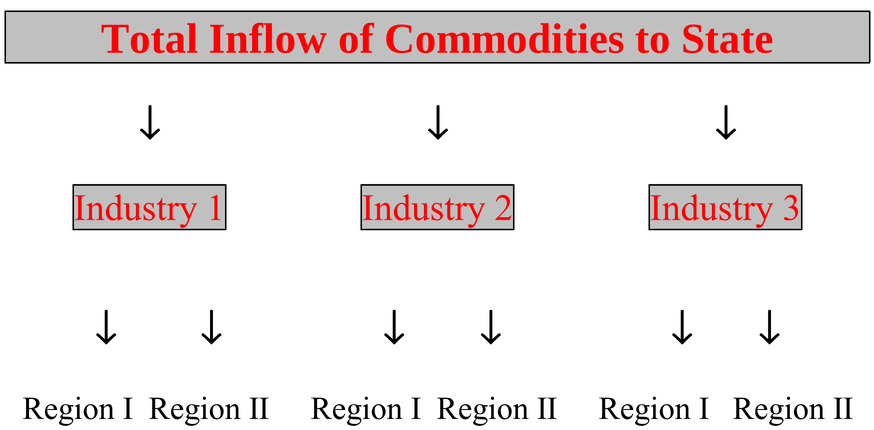
2 A gravity model for freight would be of the following general form:

$$T_{ij}^{k} = f(S_i^{k},\ D_j^{k},\ C_{ij}^{k})$$

Where T_{ij}^k are shipments of commodity *k* between regions *i* and *j*, S_i^k is the supply of commodity *k* in region *i*, D_j^k is the demand for commodity *k* in region *j*, and C_{ij}^k is a generalized cost of shipping commodity *k* between *i* and *j*.

The procedure described below involves two steps. First, using a modified regional input-output model, one defines the proportion of various commodities that are used by various industries in a region of interest. Then, one applies these proportions to existing state-level commodity inflow data from the *1993 Commodity Flow Survey* to "share down" the state-level flows to the region. One significant advantage of the methodology is that it takes into account the possibility that the input needs of a regional industry are met, in whole or in part, by regional suppliers. By accounting for existing patterns of regional inter-industry freight flows, the accuracy of estimated regional freight inflows is presumably greatly increased[3].

The procedure can be represented schematically. In essence, data on commodity inflows to a region (for example, a state) are divided into the various industries (including households) that are the likely users of these commodities as inputs. Once the inflows have been divided amongst the various inflow-consuming industries, they are then shared down to the appropriate sub-state regions based on their industrial structure. Let us suppose there are three industries and two sub-state regions, called *I* and *II*. This would then produce an assignment of commodity inflows that would follow the following pattern:



We now describe the supply-side, commodity-by-industry model. Consider first the usual national input-output accounts:

$$X = A\,X + Y \tag{1}$$

3 This aspect of the methodology contrasts with the approach suggested by Memmott (1983). While also based on input-output models, his suggested procedure for estimating regional freight flows does not account for the possibility of freight inflows being supplied regionally. As a result, the applicability of the approach for accurately estimating inflows from outside the region is limited.

Where X is an n by 1 column vector containing the output values for the n national industries, A is the n by n input requirements matrix, and Y is an n by 1 column vector of final demand purchases from each industry. Each a_{ij} element of A represents the dollar value of industry i' s output that is purchased by industry j to produce its own output.

The standard relationship outlined above is expressed in terms of dollar flows between industries. An alternative representation is based on *commodities* flowing to various industries, which is referred to in input-output accounts as the *use* or *absorption* matrix (Miller and Blair (1985)). This matrix can be represented as:

$$(2) \qquad U = [\, U_{ij} \,]$$

U is an m by n matrix, with each u_{ij} represents the amount of commodity i (expressed in monetary units) used by industry j as an input in its production. In other words, each of the m rows of U details the total industrial destinations of each of the m commodities represented in the accounts. In a manner similar to (1), we can define a set of commodity-based accounts:

$$(3) \qquad Q = B\,X + E$$

Where Q is a m by 1 vector of commodity gross outputs, E is a m by 1 vector of commodity deliveries to final demand and B is an m by n matrix where each element is defined as:

$$(4) \qquad b_{ij} = \frac{U_{ij}}{X_j}$$

The derivation of B in matrix terms is therefore:

$$(5) \qquad B = U\,(\hat{X})^{-1}$$

Where ^ indicates a diagonalized matrix. Each b_{ij} of matrix B represents the amount of commodity i required to produce a dollar's worth of industry j's output.

An alternative use of the data contained in U, Q and E is to derive a different matrix, generated by dividing each row element of U by its row sum. By dividing row elements by their total production for industrial or final demand uses, we obtain the commodity-by-industry equivalent of the "supply-side" input-output model (Augustinovics (1970)). Formally, we define:

$$(6) \qquad \beta = \left(\hat{Q}\right)^{-1} U$$

β has a different interpretation than B. Each β_{ij} element of β represents the share of commodity i sold to industry j. This matrix β is the key element in the methodology, allowing us to share state-level commodity inflows to the appropriate industries that use the commodities as inputs. Several further steps are required in order to do so.

The matrix β, which is a matrix representing national data, must be regionalized to the state level In order to share commodity inflows to the regional level, a procedure based on *location quotients* is used. We define a simple state-level location quotient as in Miller and Blair (1985):

$$(7) \qquad L_i = \frac{Ear_i^r / Ear^r}{Ear_i^N / Ear^N}$$

Ear_i is earnings in industry i in either state r (indicated by a superscript r) or in the nation (with superscript N). Ear^r is total regional earnings and Ear^N total national earnings. Note that the location quotient could also be based on employment data.

Constraining the location quotient to be equal to or less than 1, each L_j for the state is then used to regionalize the elements of β. Forming a vector of state-level location quotients, L_{state}, the following multiplication is carried out:

$$(8) \qquad \beta_{state} = \beta \; \hat{L}_{state}$$

Where ^ indicates the diagonalized matrix formed from the vector L_{state}. Each element of β_{state} adjusts the national values of β downward if the state contains a presence of the industry that is less than the national average. However, the *row sums* of β_{State} (as opposed to each $\beta_{ij\ (State)}$ element of β_{state}), should then be adjusted so as to be equal to 1. The reason for this is simple. As the matrix will be used to apportion freight flows to different industries, we are interested in the relative values of the elements of β_{State} rather than their absolute values. To ensure that row sums equal 1, we carry out a balancing procedure:

$$(9) \qquad c_{ij} = \beta_{ij\,(state)} \frac{1}{\Sigma_j \beta_{ij\,(state)}}$$

This balancing procedure now ensures that the row sums of a new matrix, C, sum to 1. This procedure is necessary in order to ensure that all commodity inflows to a state can be assigned an end user. It essentially reflects the following assumption: if an industry, say industry j, is not present in the state, the inflows of any commodity that it uses as an input are simply assumed to be used by other industries which are both present in the state and use the commodity as an input. This same procedure is also carried out if industry j is present in the state but its presence is below the national average: whatever inputs are not used by industry j are simply allocated to all the other industries that use the commodity and are present in the state.

Each c_{ij} element of matrix C now can be said to approximate the proportion of a commodity i which is shipped to a state which will be used by an industry j. In other words, C directs the commodities entering the state to the industries that can be expected to use the commodities as inputs. Mathematically, the operation involves a simple post-multiplication of the state-level commodity inflows by C, resulting in a disaggregation of these inflows into the industries that use them as inputs:

$$(10) \qquad \rho = \left(\hat{\Phi}\right) C$$

Where the vector Φ contains commodity inflows to the state, and ^ as usual indicates the vector is converted to a diagonalized matrix. The operation produces the matrix ρ, where elements ρ_{ij} detail the amount of commodity i flowing to industry j in the state. At this point, we have produced a matrix ρ that can be used to apportion freight inflows amongst the state industries that will use them as inputs. To further regionalize these flows to the sub-state level, another regionalizing procedure needs to be carried out. In a similar manner to the previous regionalization, we calculates a matrix of regional earnings shares, L_{reg} , which measure the relative representation of each industry in the sub-state region. Multiplying ρ by a matrix produced from diagonalizing the vector L_{reg} produces the matrix $\rho_{reg.}$:

$$(11) \qquad \rho_{reg} = \rho \hat{L}_{reg} = \left(\hat{Q}\, C\right) \hat{L}_{reg}$$

Each element of the matrix ρ_{reg} gives an approximation of the amount of a commodity shipped to the state which is used by a *regional* industry. In so doing, the state-level commodity inflows have been directed to a sub-state region, depending on the location of industries that use the commodities as inputs. Any row sum of ρ_{reg} gives an estimate of the total amount of a given commodity that is shipped to the region. The resulting vector of estimated regional inflows is denoted as Φ_{reg}, and the total inflow of any given commodity as $\Phi_{i(reg)}$:

$$(12) \qquad \Phi_{i(reg)} = \sum_j \rho_{ij(reg)}$$

An important assumption embodied in the use of ρ_{reg} is that each regional industry that uses a given commodity as an input will use it in the same proportions as the industry nationally. In other words, it is assumed that local industries use commodity inputs in relative proportions equal to the relative proportions in β, a standard assumption when regionalizing national input-output flows with location quotients.

Another assumption implicit in the methodology is that firms all purchase locally produced commodity inputs in the same proportions. For example, if commodity *i* is produced in the state, and satisfies 10 percent of local state needs, it is assumed that all firms that use commodity *i* will purchase 10 percent of their input needs locally. This assumption can presumably create bias in estimates of regional inflows: to the degree that the local production of *i* is concentrated in certain sub-state regions, some local industries might purchase more than 10 percent of their needs from the local state suppliers.

Finally, besides assuming that firms all purchase locally produced inputs in the same proportions, the methodology also assumes that industries purchase their extra-regional inputs from any given region in the same proportion as well. Note that this assumption is the same as found in construction of the multi-regional input-output models (Miller and Blair (1985))

ACCURACY TESTS

Having described a relatively simple methodology to estimate freight inflows to a sub-state region,
Can we determine its accuracy? As mentioned previously, the approach suggested here is intended to be used for estimates of freight flows to sub-state regions, where by definition little or no data exists to permit validation of the estimates. This would imply that validating the results of the methodology would require actual survey data on freight inflows to the region. The lack of such surveys for small regions is precisely what motivated the elaboration of the supply-side commodity-by-industry methodology.

An alternative approach to determining the methodology is possible, however. This involves treating states as if they were sub-state regions, and creating larger regions comprised of a series of individual states. Then one uses the total freight inflows to these several states as if they were inflows to an individual state. In so doing, one must be careful to remove the freight flows between the various states which make up the larger region. The result is data detailing all inflows of commodities to the larger region from outside this region.

It should be pointed out that the *Commodity Flow Survey* is comprehensive in that all modes are covered. For the 1993 *Commodity Flow Survey*, the United States Bureau of the Census used a sample of 200,000 establishments in manufacturing, mining, wholesale, service and retail. Each establishment was asked to report shipments for two-week periods in each of the four calendar quarters, identifying domestic origin and destination, commodity type, weight, value and mode of transport. The *Commodity Flow Survey* does exclude certain commodities, notably crude petroleum. Also, while imports and exports are included, commodities that are shipped from a foreign location to another destination through the United States are excluded.

In carrying out our tests, we selected three regions in the United States which each contain a number of States. The first, which we refer to as the *Northeast Region,* contains the states of Massachusetts, Rhode Island, Connecticut, Maine, New Hampshire and Vermont. The second is referred to as the *Mideast Region*, which includes the states of Delaware, Maryland, New Jersey, New York and Pennsylvania. Finally, the *Great Lakes Region* is roughly in the country's middle and incorporates the states of Illinois, Indiana, Michigan, Ohio and Wisconsin.

For each of these three regions we estimate ρ for the entire group of component states, as defined in (10). For the purposes of the analysis, the states of Massachusetts, New York, Pennsylvania, Illinois and Ohio were treated as if it they were sub-state regions. For each of these states both ρ_{reg} and Φ_{reg} were calculated according to the definitions of equations (11) and (12), as if the detailed state-level data in the *Commodity Flow Survey* did not exist.

How do these estimates compare to the actual freight inflows to the states? Details of the estimates and the actual observed inflows are detailed for each of the states in Table 1 to Table 5. In general, the methodology performs well for total forecasts of commodities, but the forecast of specific commodities is variable. For example, in the case of Massachusetts, the forecast for total commodity inflows is within 8.7 percent of the actual observed inflow. This figure, however, obscures the fact that while some commodities are forecast with less than 5 percent error, others are forecast with as much as 67 percent error (as with forest and fishing products). This is due to the fact that a simple summation of the percent error of individual commodities will see negative and positive forecast errors canceling each other.

In general, two commodities tended to predict very poorly and were not even included in Table 1through Table 5. Mineral products and petroleum & coal products tended to be wildly inaccurate (in the case of New York, for example, the forecast was off by over 800 percent!). The fact that inflows of primary inputs such as mineral products and petroleum & coal products are predicted so poorly can be partly explained by the different patterns of energy consumption in various regions of the United States. In particular, the use of such energy sources as oil, coal, hydroelectric and nuclear power can vary across regions, regardless of industries[4]. Because of this consistently large error in predicting these commodities, we simply ignore them in our discussion, and point out that our methodology is inappropriate to forecast them.

As mentioned, summing the percentage error of individual commodities will tend to overstate the accuracy of the methodology. For example, this measure shows the forecast for the state of New York within 1.7 percent of the observed inflows, despite the fact that some individual commodities are forecast poorly. In order to evaluate the forecasts for the five states while accounting for the possibility of both negative and positive forecast error, we rely on the following measures: *Weighted average error (WAE)*, *mean absolute error (MAE)* and *Theil's inequality coefficient (TIC)*. The definitions of the measures for the *m* commodities are:

$$WAE = \sum_{i=1}^{m} \frac{\left| Estimated_i - Observed_i \right|}{Observed_i} \text{Re}\,lative\ Weight_i$$

$$MAE = \frac{1}{m} \sum_{i=1}^{m} \frac{\left| Estimated_i - Observed_i \right|}{Observed_i}$$

$$TIC = \frac{\sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(Estimated_i - Observed_i \right)^2}}{\sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(Estimated_i \right)^2} + \sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(Observed_i \right)^2}}$$

---

4 The bias of assuming national patterns of energy use or production to regions has been discussed by other authors, in particular Miller and Blair:

> Electricity produced in Eastern Washington by water power (Coulee Dam) represents quite a different mix of inputs from electricity that is produced from coal in the greater Philadelphia area or by means of nuclear power elsewhere (Miller and Blair (1985)).

Miller and Blair allude to a problem inherent in using national input-output data regionalized on the basis of non-survey techniques. This issue also affects the procedure we have suggested for estimating commodity flows. Since the methodology presented here relies on national input-output data, it will tend to assume that energy sources reflect the national "average".

WAE and MAE are meant to address the issue of negative and positive errors canceling each other out. The TIC does so as well, but is scaled to lie between 0 and 1, with a lower TIC indicating a more accurate forecast. The summary measures for the five states are included in Table 6.

Table 6: Accuracy Measures for State Forecasts

Measure:	Massachusetts	New York	Pennsylvania	Illinois	Ohio
WAE	**13.7%**	**19.6%**	**24.4%**	**24.0%**	**19.6%**
MAE	**17.5%**	**30.3%**	**23.4%**	**22.6%**	**30.1%**
TIC	**0.065**	**0.121**	**0.132**	**0.171**	**0.107**

Another measure of accuracy is the simple correlation coefficient between the observed and estimated commodity inflows. For all five states, including all commodities other than mining and petroleum produces (in other words, a sample of 106 forecasts), the correlation coefficient is 0.96.

It is of interest to note that the size of the observed inflows to the states will be negatively correlated with the absolute and relative percent error of the forecasts. While this is perhaps expected, another fact is surprising: If a state accounted for large proportions of the total commodity inflows to a region5, this tended to increase, rather than decrease, the forecast error of individual commodities. A correlation coefficient between the proportion of a state's share of regional inflows and the commodity-level forecast error was equal to –0.6. The importance of this fact is that it implies the methodology can be applied to smaller sub-state regions without necessarily expecting that forecast errors will greatly increase.

CONCLUSIONS

The procedure described above offers a relatively easy tool to estimate sub-state commodity inflows, one that can be used by transportation planners for relatively accurate "back of the envelope" predictions of aggregate commodity inflows to smaller regions. Further, the procedure has the important advantage of using the appropriate *observed* state-level commodity inflows as a starting point to estimate sub-state flows, something which cannot be claimed by econometric or gravity models which generalize inflow patterns observed in one region to another region. Though somewhat laborious, the calculations are relatively simple, using data that is widely available and low cost, at least in the United States and European Union countries.

While estimates of total freight inflows were in some cases surprisingly accurate, the estimate errors of individual commodities were often significantly greater. In particular, commodities,

---

5 States accounted for as high as 42 percent (in the case of Pennsylvania) and as low as 25 percent (in the case of Illinois) of inflows to their respective regions.

such as energy inputs, whose use could vary significantly across regions in the United States, were predicted very poorly. Excluding these commodities, 55 percent of individual commodity forecasts were predicted with less than 15 percent error, arguably acceptable imprecision for the suggested uses of the approach.

As discussed, the method does entail two crucial assumption: First, all firms at the state and regional level are assumed to display the same input use as their counterparts nationally, a necessary assumption in non-survey regional input-output modeling. This assumption appears to be a significant flaw in the estimate of inflows of energy inputs, as mentioned. Second, all firms in a regional industry are assumed to purchase locally-produced commodity inputs in the same proportions. This could introduce bias, particularly in the case of large states where firms located near a local supplier of a given commodity could consume significantly more inputs produced locally than those located further away in the state. Our method cannot differentiate these differences among firms of a same industry. This could in turn lead to overestimates of inflows of a given commodity to regions with an important local producer of that commodity. Conversely, it could also lead to underestimates of inflows to the other regions.

Despite this imperfection, it is argued that the method of estimating sub-state inflows using input-output data is sound. We argue that, in the absence of detailed and costly surveys, our approach estimates the most elusive component of regional trade estimates, commodity inflows, with acceptable levels of accuracy.

TABLE 1: Annual Domestic Commodity Inflows (Excluding Energy*) to the State of Massachusetts: Estimated vs. Observed (Thousands of Metric Tons)

Commodity Description	Observed Inflows	Estimated Inflows	% of Total Observed	Estimated - Observed	(Est-Obs) Obs	Weighted Avg. Relative Error	Mean Absolute Error
Food or kindred products	6,931	6,716	32.9%	-215	-3.1%		3.1%
Primary metal products	1,051	1,313	5.0%	262	24.9%		24.9%
Chemicals or allied products	3,332	2,527	15.8%	-805	-24.2%		24.2%
Pulp, paper, or allied products	2,713	2,420	12.9%	-293	-10.8%		10.8%
Clay, concrete, glass, or stone products	1,512	1,776	7.2%	264	17.5%		17.5%
Lumber or wood products, excluding furniture	1,340	1,013	6.4%	-327	-24.4%		24.4%
Waste or scrap materials	32	36	0.2%	4	12.5%		12.5%
Farm products	259	228	1.2%	-31	-12.0%		12.0%
Fabricated metal products	644	576	3.1%	-68	-10.6%		10.6%
Transportation equipment	903	579	4.3%	-324	-35.9%		35.9%
Rubber or miscellaneous plastics products	557	479	2.6%	-78	-14.0%		14.0%
Machinery, excluding electrical	221	194	1.0%	-27	-12.2%		12.2%
Electrical machinery, equipment, or supplies	333	266	1.6%	-67	-20.1%		20.1%
Textile mill products	391	340	1.9%	-51	-13.0%		13.0%
Furniture or fixtures	244	219	1.2%	-25	-10.2%		10.2%
Miscellaneous products or manufacturing	276	273	1.3%	-3	-1.1%		1.1%
Apparel or other finished textile products	238	205	1.1%	-33	-13.9%		13.9%
Forest and fishing products	3	1	0.0%	-2	-66.7%		66.7%
Instruments, photographic goods, optical good	33	26	0.2%	-7	-21.2%		21.2%
Leather or leather products	41	40	0.2%	-1	-2.4%		2.4%
Total, All Commodities	**21,054**	**19,227**	**100.0%**	**-1,827**	**-8.7%**	**13.7%**	**17.5%**

Note: Observed inflows are obtained from the 1993 Commodity Flow survey; estimated inflows are derived using the input-output approach From the inflows to the six-state Northeastern Region (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont).
*Energy commodities include mining, petroleum and coal products.

Source: Estimates by the authors using 1993 Commodity Flow Survey.

TABLE 2: Annual Domestic Commodity Inflows (Excluding Energy*) to the State of New York: Estimated vs. Observed (Thousands of Metric Tons)

Commodity Description	Observed Inflows	Estimated Inflows	% of Total Observed	Estimated - Observed	(Est-Obs) Obs	Weighted Avg. Relative Error	Mean Absolute Error
Food or kindred products	14,785	15,430	25.9%	645	4.4%		4.4%
Lumber or wood products, excluding furniture	7,962	4,093	14.0%	-3,870	-48.6%		48.6%
Pulp, paper, or allied products	7,241	6,843	12.7%	-398	-5.5%		5.5%
Chemicals or allied products	6,508	7,551	11.4%	1,044	16.0%		16.0%
Primary metal products	4,358	4,119	7.6%	-239	-5.5%		5.5%
Clay, concrete, glass, or stone products	3,079	2,763	5.4%	-317	-10.3%		10.3%
Transportation equipment	2,872	2,437	5.0%	-436	-15.2%		15.2%
Fabricated metal products	2,264	1,895	4.0%	-369	-16.3%		16.3%
Rubber or miscellaneous plastics products	1,381	1,308	2.4%	-73	-5.3%		5.3%
Farm products	1,222	3,879	2.1%	2,657	217.5%		217.5%
Textile mill products	1,161	951	2.0%	-210	-18.1%		18.1%
Electrical machinery, equipment, or supplies	1,067	1,030	1.9%	-37	-3.5%		3.5%
Machinery, excluding electrical	848	829	1.5%	-19	-2.2%		2.2%
Furniture or fixtures	570	654	1.0%	84	14.7%		14.7%
Miscellaneous products or manufacturing	528	443	0.9%	-85	-16.1%		16.1%
Apparel or other finished textile products	429	437	0.8%	7	1.7%		1.7%
Waste or scrap materials	408	1,052	0.7%	643	157.5%		157.5%
Instruments, photographic goods, optical good	202	179	0.4%	-22	-11.0%		11.0%
Forest and fishing products	113	116	0.2%	3	2.5%		2.5%
Leather or leather products	43	39	0.1%	-4	-9.3%		9.3%
Tobacco products, excluding insecticides	17	27	0.0%	9	54.0%		54.0%
Total, All Commodities	**57,060**	**56,074**	**100.0%**	**-986**	**-1.7%**	**19.6%**	**30.3%**

Note: Observed inflows are obtained from the 1993 Commodity Flow survey; estimated inflows are derived using the input-output approach from the inflows to the five-state Mideast Region (Delaware, Maryland, New Jersey, New York, and Pennsylvania).

*Energy commodities include mining, petroleum and coal products.

Source: Estimates by the authors using 1993 Commodity Flow Survey.

TABLE 3: Annual Domestic Commodity Inflows (Excluding Energy*) to the State of Pennsylvania: Estimated vs. Observed (Thousands of Metric Tons)

Commodity Description	Observed Inflows	Estimated Inflows	% of Total Observed	Estimated - Observed	(Est-Obs) Obs	Weighted Avg. Relative Error	Mean Absolute Error
Food or kindred products	12,683	10,141	25.7%	-2,543	-20.0%		20.0%
Primary metal products	7,860	6,467	15.9%	-1,393	-17.7%		17.7%
Chemicals or allied products	7,125	5,940	14.4%	-1,185	-16.6%		16.6%
Pulp, paper, or allied products	5,724	4,537	11.6%	-1,187	-20.7%		20.7%
Clay, concrete, glass, or stone products	3,046	2,598	6.2%	-448	-14.7%		14.7%
Lumber or wood products, excluding furniture	2,618	4,083	5.3%	1,466	56.0%		56.0%
Waste or scrap materials	1,873	709	3.8%	-1,164	-62.2%		62.2%
Farm products	1,523	3,792	3.1%	2,268	148.9%		148.9%
Fabricated metal products	1,511	1,563	3.1%	51	3.4%		3.4%
Transportation equipment	1,305	1,434	2.6%	129	9.9%		9.9%
Rubber or miscellaneous plastics products	973	940	2.0%	-33	-3.4%		3.4%
Machinery, excluding electrical	677	560	1.4%	-117	-17.2%		17.2%
Electrical machinery, equipment, or supplies	675	612	1.4%	-63	-9.4%		9.4%
Textile mill products	553	533	1.1%	-20	-3.5%		3.5%
Furniture or fixtures	407	380	0.8%	-27	-6.7%		6.7%
Miscellaneous products or manufacturing	281	246	0.6%	-34	-12.3%		12.3%
Apparel or other finished textile products	254	232	0.5%	-22	-8.6%		8.6%
Forest and fishing products	93	66	0.2%	-27	-28.7%		28.7%
Instruments, photographic goods, optical good	92	99	0.2%	7	7.4%		7.4%
Leather or leather products	28	23	0.1%	-5	-18.1%		18.1%
Tobacco products, excluding insecticides	14	15	0.0%	1	6.5%		6.5%
Total, All Commodities	**49,315**	**44,970**	**100.0%**	**-4,345**	**-8.8%**	**24.7%**	**23.4%**

Note: Observed inflows are obtained from the 1993 Commodity Flow survey; estimated inflows are derived using the input-output approach from the inflows to the five-state Mideast Region (Delaware, Maryland, New Jersey, New York, and Pennsylvania).

*Energy commodities include mining, petroleum and coal products.

Source: Estimates by the authors using 1993 Commodity Flow Survey.

TABLE 4: Annual Domestic Commodity Inflows (Excluding Energy*) to the State of Illinois: Estimated vs. Observed (Thousands of Metric Tons)

Commodity Description	Observed Inflows	Estimated Inflows	% of Total Observed	Estimated - Observed	(Est-Obs) Obs	Weighted Avg. Relative Error	Mean Absolute Error
Food or kindred products	20,585	13,618	31.3%	-6,967	-33.8%		33.8%
Chemicals or allied products	12,893	10,025	19.6%	-2,868	-22.2%		22.2%
Pulp, paper, or allied products	7,506	6,006	11.4%	-1,500	-20.0%		20.0%
Primary metal products	6,013	4,116	9.1%	-1,898	-31.6%		31.6%
Clay, concrete, glass, or stone products	4,594	4,151	7.0%	-443	-9.6%		9.6%
Lumber or wood products, excluding furniture	3,655	2,985	5.6%	-670	-18.3%		18.3%
Farm products	2,306	2,566	3.5%	260	11.3%		11.3%
Fabricated metal products	1,483	1,355	2.3%	-128	-8.6%		8.6%
Transportation equipment	1,265	1,686	1.9%	421	33.3%		33.3%
Rubber or miscellaneous plastics products	1,138	1,151	1.7%	13	1.1%		1.1%
Waste or scrap materials	1,117	932	1.7%	-185	-16.5%		16.5%
Machinery, excluding electrical	918	715	1.4%	-203	-22.1%		22.1%
Electrical machinery, equipment, or supplies	707	737	1.1%	30	4.2%		4.2%
Furniture or fixtures	440	378	0.7%	-62	-14.1%		14.1%
Miscellaneous products or manufacturing	323	327	0.5%	4	1.4%		1.4%
Textile mill products	323	293	0.5%	-30	-9.3%		9.3%
Apparel or other finished textile products	300	300	0.5%	-1	-0.2%		0.2%
Instruments, photographic goods, optical good	161	91	0.2%	-70	-43.3%		43.3%
Tobacco products, excluding insecticides	43	14	0.1%	-29	-66.8%		66.8%
Leather or leather products	23	34	0.0%	11	46.1%		46.1%
Forest and fishing products	23	12	0.0%	-12	-50.0%		50.0%
Ordnance or accessories	15	10	0.0%	-5	-32.5%		32.5%
Total, All Commodities	**65,831**	**51,501**	**100.0%**	**-14,330**	**-21.8%**	**24.0%**	**22.6%**

Note: Observed inflows are obtained from the 1993 Commodity Flow survey; estimated inflows are derived using the input-output approach from the inflows to the five-state Great Lakes Region (Illinois, Indiana, Michigan, Ohio, and Wisconsin).

*Energy commodities include mining, petroleum and coal products.

Source: Estimates by the authors using 1993 Commodity Flow Survey.

TABLE 5: Annual Domestic Commodity Inflows (Excluding Energy*) to the State of Ohio: Estimated vs. Observed (Thousands of Metric Tons)

Commodity Description	Observed Inflows	Estimated Inflows	% of Total Observed	Estimated - Observed	(Est-Obs) Obs	Weighted Avg. Relative Error	Mean Absolute Error
Chemicals or allied products	10,705	9,610	23.5%	-1,095	-10.2%		10.2%
Food or kindred products	7,956	10,770	17.5%	2,814	35.4%		35.4%
Primary metal products	6,710	5,394	14.8%	-1,316	-19.6%		19.6%
Clay, concrete, glass, or stone products	4,372	3,938	9.6%	-434	-9.9%		9.9%
Pulp, paper, or allied products	4,117	4,983	9.1%	866	21.0%		21.0%
Lumber or wood products, excluding furniture	3,309	2,922	7.3%	-387	-11.7%		11.7%
Transportation equipment	1,460	1,628	3.2%	168	11.5%		11.5%
Fabricated metal products	1,269	1,268	2.8%	-1	-0.1%		0.1%
Rubber or miscellaneous plastics products	1,190	1,049	2.6%	-141	-11.9%		11.9%
Waste or scrap materials	1,115	802	2.5%	-313	-28.1%		28.1%
Electrical machinery, equipment, or supplies	677	601	1.5%	-76	-11.2%		11.2%
Farm products	646	1,716	1.4%	1,070	165.5%		165.5%
Machinery, excluding electrical	610	666	1.3%	56	9.1%		9.1%
Furniture or fixtures	338	326	0.7%	-13	-3.7%		3.7%
Textile mill products	336	274	0.7%	-62	-18.6%		18.6%
Miscellaneous products or manufacturing	278	261	0.6%	-17	-6.0%		6.0%
Apparel or other finished textile products	276	259	0.6%	-17	-6.2%		6.2%
Instruments, photographic goods, optical good	55	76	0.1%	21	38.6%		38.6%
Leather or leather products	43	28	0.1%	-15	-34.8%		34.8%
Ordnance or accessories	15	8	0.0%	-6	-42.7%		42.7%
Tobacco products, excluding insecticides	6	12	0.0%	6	105.0%		105.0%
Forest and fishing products	5	8	0.0%	3	60.9%		60.9%
Total, All Commodities	**45,488**	**46,599**	**100.0%**	**1,111**	**2.4%**	**19.6%**	**30.1%**

Note: Observed inflows are obtained from the 1993 Commodity Flow survey; estimated inflows are derived using the input-output approach from the inflows to the five-state Great Lakes Region (Illinois, Indiana, Michigan, Ohio, and Wisconsin).
*Energy commodities include mining, petroleum and coal products.

Source: Estimates by the authors using 1993 Commodity Flow Survey.

REFERENCES

Augustinovics, Maria, "Methods of International and Intertemporal Comparisons of Structure". In *Contributions to Input-Output Analysis,* edited by A. P. Carter and A.Brody. Amsterdam: North-Holland, 1970.

Memmott, Frederick W. Application of Statewide Freight Demand Forecasting Techniques, *National Cooperative Highway Research Program Report 260*. Washington, DC: Transportation Reaserch Board and National Research Council, 1983.

Miller, Ronald. E., and Peter D. Blair, *Input-Output analysis: Foundations and Extensions*. Englewood Cliffs, New Jersey: Prentice-Hall, 1985.

Ortúzar, Juan de Dios and Luis G. Willumsen, *Modelling Transport.* Chichester: John Wiley & Sons, 1994.

United States Department of Commerce, Bureau of Census, *1993 Commodity Flow Survey*, CD-CFS-93-1, December 1993.

United States Department of Commerce, Bureau of Economic Analysis, *Input-Output Accounts of the US Economy, 1992 Benchmark*, November 1997.

Vilain, Pierre, Louie Nan Liu and David Aimen, "Estimation of Commodity Inflows to a Substate Region: An Input-Output Based Approach", *Transportation Research Record 1653.* Washington, DC: Transportation Reaserch Board and National Research Council, 1999.