



ELSEVIER

Transportation Research Part B 39 (2005) 65–79

TRANSPORTATION
RESEARCH
PART B

www.elsevier.com/locate/trb

Implementation and estimation of a combined model of interregional, multimodal commodity shipments and transportation network flows

Heejoo Ham ^a, Tschangho John Kim ^{b,c,*}, David Boyce ^d

^a *URS Corporation, Blue Bell, PA 19034, USA*

^b *Seoul National University, Seoul 151-742, South Korea*

^c *University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA*

^d *University of Illinois at Chicago, Chicago, IL 60607, USA*

Received 12 June 2002; received in revised form 25 July 2003; accepted 23 February 2004

Abstract

A combined model of interregional, multimodal commodity shipments, incorporating regional input–output relationships, and the associated transportation network flows is formulated as an alternative to the traditional four-step travel forecasting procedure of trip generation, distribution, mode choice, and assignment. The paper describes the formulation of the model, its solution using US interregional commodity shipment data, estimation of key parameters, and evaluation of the performance of the model with the observed data. The model was implemented to predict interregional commodity shipments by roads and railways among regions of the US when these networks were disrupted by earthquakes or other natural events.

© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Interregional commodity shipment modeling, as a subfield of research and practice, lags far behind urban travel modeling in its development and significance for policy analysis and planning. In spite of early interest and advances in the formulation of these models by Isard (1960), Leontief

* Corresponding author. Address: University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA. Tel.: +1-217-244-5369; fax: +1-217-244-1717.

E-mail address: tjohnkim@uiuc.edu (T.J. Kim).

and Strout (1963), and Wilson (1970), little progress on implementation and estimation of large-scale models has resulted.

Perhaps the overriding reason for the dearth of operational models is the lack of useable interregional commodity shipment and transportation network freight flow data. Although a census of commodity shipments has been conducted in the US since the 1960s, the data have been released in highly aggregated forms, making them less useful for model estimation than comparable data on urban travel. Likewise, road and rail network representations have generally not been available in a form suitable for freight network modeling.

Proposals for model formulations during the 1980s (Boyce and Hewings, 1981; Batten and Boyce, 1986; Kim et al., 1983) sought to maintain interest in this area. More recently, both Kim and Kockelman have led efforts to implement such models for policy analysis. The context of the former (Kim et al., 2002) was the impact of earthquakes on interregional trade, while the latter (Lin et al., 2002; Zhao and Kockelman, 2004) concerned relationships between Texas and its trading partners. Likewise, Fernandez et al. (2003) have proposed a comprehensive model representing the equilibrium of demand and supply of intercity freight flows.

Clearly, there is a need for such models. Regions (states, metropolitan regions, counties, etc.) faced with allocating scarce capital budgets to infrastructure improvements require forecasts to determine which links of the interregional road network are likely to be subjected to excessive freight flows in the future. Private carriers (motor carriers, railways, etc.) faced with long-term capital decisions also benefit from better forecasts. Seismic events in California and floods in the Mississippi Valley during the past decade illustrate that Nature can also have a devastating effect on operations of transportation networks. Answering questions about the priority of structural and other enhancements clearly requires better insights into network redundancy.

The model formulation is motivated by the following scenario. A large nation is divided into several economic regions, each comprising one or more urban subregions as well as rural agricultural, forestry or mining activities. Regional economic models encompassing both input–output and econometric relationships exist for each region, which describe the relation between regional production by sector for regional consumption, investment and government, as well as exports to other regions. We wish to augment the forecasts provided by these regional models with forecasts of interregional commodity shipments by sector and mode, and associated freight flows on the interregional road and railway systems. Depicting these interregional relationships not only helps to understand the implications of regional growth for transportation investment, but also improve the validity of the regional models themselves.

This contribution to interregional trade modeling has two principal objectives:

1. formulation and solution of an integrated model of interregional, multimodal commodity shipments and transportation network flows in a manner that the impact of network disruptions can be identified;
2. estimation of model parameters from such limited cross-sectional data as are available.

Although examples of similar model formulations for may be readily found in the urban travel modeling literature (Oppenheim, 1995; Boyce and Daskin, 1997), the paucity of examples of applications to interregional trade suggests the findings reported here may interest readers.

The paper is organized in the following way. First, the model requirements and data for the implementation are described. Next, the model formulation is presented and its properties analyzed. Then, the parameter estimation procedure and results are described, followed by a short presentation of model results. Finally, aggregate predictions from the model are compared with data used in the estimation process as an indication of the model's goodness of fit. A discussion of future research concludes the paper.

2. Model requirements and data availability

The commodity shipment model includes multiregional economic/input–output relationships concerning M industrial sectors producing a like number of commodities distributed over R regions. This input–output-based commodity shipment model is integrated with a multimodal model of transportation network flows to predict the monetary value (\$) of interregional commodity shipments and the weight (tons) of transportation network flows by sector and mode. Predicted commodity shipments between the regions are allocated to modes, routes and links according to a simple minimum distance criterion. The model is formulated as a constrained optimization problem, solved by the partial linearization algorithm of Evans (1976), and estimated with 1993 commodity flow data.

Each sector produces one aggregate commodity. Therefore, the terms *sector* and *commodity* are used interchangeably. The level of aggregation is 11 sectors/commodities plus the construction and service sectors, which produce only for intraregional consumption and investment. The classification of sectors is shown in Table 1.

Commodity shipment data for the United States are collected periodically by the US Bureau of the Census in conjunction with the Bureau of Transportation Statistics (1997) of the US Department of Transportation. Commodity shipments are reported among the 50 states and among somewhat smaller regional economic units called national transportation analysis regions

Table 1
Sector descriptions

Sector	Description
1	Agriculture, forestry, and fisheries
2	Mining
3	Construction
4	Food and kindred products
5	Chemicals and allied products
6	Primary metals industries
7	Fabricated metal products
8	Industrial machinery and equipment
9	Electronic and electric equipment
10	Transportation equipment
11	Other non-durable manufacturing
12	Other durable manufacturing
13	TCU, services, and government enterprises

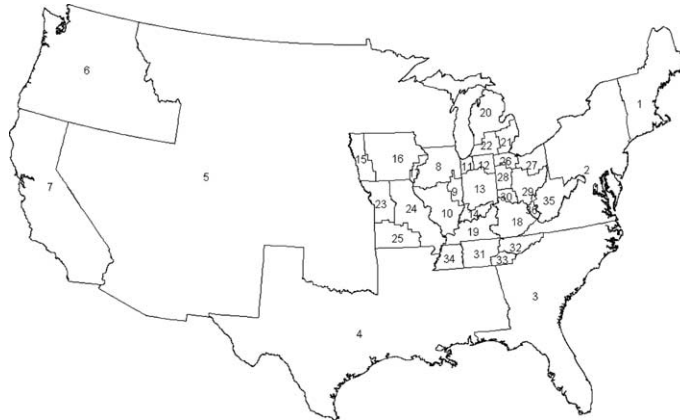


Fig. 1. Regions.

(NTAR). Interregional shipments by mode are also available, but not by commodity. For this model, a system of regions was devised based on NTAR, as shown in Fig. 1.

Transportation networks consist of the US interstate highway system and the US railway system. The data for the transportation networks, including roads and railways, were based on 1997 National Transportation Atlas Database (NTAD) of the US Department of Transportation. To analyze the truck commodity flows, a road network was constructed with 167 nodes and 532 links, based on the system shown in Fig. 2. The nodes and links are defined by road intersections and distances between intersections respectively.

The capacities of road links are assumed to be 11,000 trucks/lane/day based on observed flows on I-94 in northern Indiana, reputedly the most heavily used truck facility in the US. Railway links are considered to have unlimited capacity. The railway network is represented as a spider network connecting regional centroids with straight-line segments.



Fig. 2. Road network.

3. Model formulation, optimality conditions and solution algorithm

3.1. Model formulation

This model formulation may be regarded as integrating a Leontief–Strout–Wilson interregional commodity shipment model with a transportation network model. Exports and imports of commodities for each region are given by the regional input–output models for the period of analysis taken as one year. Our problem is to predict the shipments by sector/commodity between each pair of regions and the transportation network flows resulting from these shipments.

Let f_a^w denote the total flow (tons) on link a of mode w over an effective distance (miles) of $d_a^w(f_a^w)$. For this prototype model, link distances are either considered to be fixed, or to be effectively longer than the actual length in the case of congested road links, representing the additional operating costs of congestion. The traditional Bureau of Public Roads (BPR) link performance function is utilized for roads to represent this effective distance, which increases as a function of the flow to capacity ratio raised to a positive power. Sensitivity analysis results are reported for powers of 0 (no congestion), 4 and 8. For railways, link distances are fixed, corresponding to the assumption of unlimited capacity. The shipment distance d_{jj}^w for mode w within region j is also fixed. In applications of the model, links may be removed entirely from the network to represent the effects of catastrophic events, or in the case of roads, the number of lanes may be reduced.

Let h_{ijr}^{mw} denote the flow (tons) of output of sector m from region i to region j by mode w on route r , and x_{ij}^{mw} denote the shipment (\$) of output of sector m from region i to region j by mode w . To define the relationships among these variables and other quantities, consider the following definitions:

1. Link flow equals the sum of the route flows on all routes between all regions using link a of mode w :

$$\sum_m \sum_{ijr} h_{ijr}^{mw} \phi_{ijr}^{aw} = f_a^w \quad \text{for all links } a \text{ and modes } w \quad (1)$$

where ϕ_{ijr}^{aw} is 1, if route r from i to j uses link a of mode w , and 0, otherwise.

2. The materials balance constraint states the total shipment of commodity m into region j from all regions i is equal to the use of that commodity for producing all commodities (intermediate demand) plus regional consumption, investment and government expenditures (final demand).

$$\sum_i x_{ij}^m = \sum_n a_j^{mn} \sum_k x_{jk}^n + y_j^m \quad \text{for all sectors } m \text{ and regions } j \quad (2)$$

where a_j^{mn} is the quantity of inputs from sector m required to produce one unit of output of sector n in region j , the quantities being measured in value terms, and y_j^m is the final demand for the output of sector m in region j . The materials balance constraint represents the conservation of shipments between origin and destination regions and within regions. Regions are assumed to be indifferent regarding the source of the supply of commodities.

3. The sum of shipments of sector m from region i to region j by all modes equals the total shipment of sector m from region i to region j .

$$\sum_w x_{ij}^{mw} = x_{ij}^m \quad \text{for all sectors } m, \text{ and regions } i, j \quad (3)$$

4. The total flow (tons) of sector m on all routes r of mode w between regions i and j equals the total shipment (\$) of sector m from i to j by mode w divided by an exogenous ratio g^m (\$/ton) that converts commodity shipments from value of commodity shipments to weight of network flows.

$$\sum_r h_{ijr}^{mw} = \frac{x_{ij}^{mw}}{g^m} \quad \text{for all sectors } m, \text{ all modes } w, \text{ and regions } i, j \quad (4)$$

The ratio is derived from the 1993 Commodity Flow Survey by dividing the total value of shipments (\$) by total weight of the shipments (tons) throughout the US. The values of these ratios are shown in Table 3.

5. Finally, route flows must be non-negative, which implies that links flows are also non-negative:

$$h_{ijr}^{mw} \geq 0 \quad \text{for all sectors } m, \text{ and all routes } r \text{ of mode } w \text{ connecting regions } i, j \quad (5)$$

Having defined the constraints on interregional and route flows, we next consider the objective function. On the assumption that shippers collectively desire to minimize their total shipment distances, in the absence of shipment cost information, and that other factors causing the dispersion of shipments across origins and destinations and modes can be represented by interregional and modal dispersion (entropy) functions, the objective function is defined as follows:

$$\begin{aligned} \min_{(\mathbf{h}, \mathbf{x})} Z(\mathbf{h}, \mathbf{x}) = & \sum_{aw} \int_0^{f_a^w} d_a^w(\omega) d\omega + \sum_{mjw} \left(\frac{x_{jj}^{mw}}{g^m} \right) d_{jj}^w + \sum_m \frac{1}{\beta^m g^m} \sum_{ij} x_{ij}^m \ln \left(\frac{x_{ij}^m}{\bar{X}_i^m} \right) \\ & + \sum_m \frac{1}{\alpha^m g^m} \sum_{ijw} x_{ij}^{mw} \ln \left(\frac{x_{ij}^{mw}}{x_{ij}^m} \right) \end{aligned} \quad (6)$$

The exogenous parameters and variables are: α^m , the modal distance sensitivity parameter for sector m ; β^m , the interregional distance sensitivity parameter for sector m ; g^m , the ratio of monetary shipment value to weight used to convert the output of sector m from dollars to tons (\$/ton); and \bar{X}_i^m , the total output of sector m in region i (\$) in a prior year, a measure of its size, which is proportional to the a priori shipment before the consideration of shipment distances and other constraints.

The total distance to be minimized is represented by the sum of the link flows times link distances by mode plus intraregional shipments times intraregional modal shipment distances. The model solution determines the interregional and intraregional commodity shipments by mode, and the freight flows by routes and links within modes. Not all of these commodity shipments and flows, however, are uniquely determined by the solution.

3.2. Lagrangian analysis and optimality conditions

The Lagrangian function for deriving the optimality conditions is constructed using objective function (6), constraints (1)–(5), and Lagrangian multipliers as follows:

$$\begin{aligned}
 \min_{(\mathbf{h}, \mathbf{x}, \gamma, \theta, \mathbf{d})} L(\mathbf{h}, \mathbf{x}, \gamma, \theta, \mathbf{d}) &= \sum_{av} \int_0^{f_a^w} d_a^w(\omega) d\omega + \sum_{mjw} \left(\frac{x_{jj}^{mw}}{g^m} \right) d_{jj}^w + \sum_m \frac{1}{\beta^m g^m} \sum_{ij} x_{ij}^m \ln \left(\frac{x_{ij}^m}{\bar{X}_i^m} \right) \\
 &+ \sum_m \frac{1}{\alpha^m g^m} \sum_{ijw} x_{ij}^{mw} \ln \left(\frac{x_{ij}^{mw}}{x_{ij}^m} \right) \\
 &+ \sum_{mj} \gamma_j^m \left\{ \sum_n a^{mn} \left(\sum_k x_{jk}^n \right) + y_j^m - \sum_i x_{ij}^m \right\} \\
 &+ \sum_{mij} \theta_{ij}^m \left(x_{ij}^m - \sum_w x_{ij}^{mw} \right) + \sum_{mijw} d_{ij}^{mw} \left(\frac{x_{ij}^{mw}}{g^m} - \sum_r h_{ijr}^{mw} \right)
 \end{aligned} \tag{7}$$

s.t. $h_{ijr}^{mw} \geq 0$

The Lagrange multipliers are (γ_j^m) , (θ_{ij}^m) , (d_{ij}^{mw}) ; in vector form, the same symbols are used without subscripts or superscripts. The Karush–Kuhn–Tucker optimality conditions are obtained by taking partial derivatives with respect to the unknown variables as follows:

$$\frac{\partial L}{\partial h_{ijr}^{mw}} = \sum_a d_a(f_a^w) \varphi_{ijr}^{aw} - d_{ij}^{mw} \geq 0 \quad \forall mwrij \tag{8}$$

$$\frac{\partial L}{\partial x_{ij}^{mw}} = \frac{1}{\alpha^m g^m} \left(\ln \frac{x_{ij}^{mw}}{x_{ij}^m} + 1.0 \right) - \theta_{ij}^m + \frac{d_{ij}^{mw}}{g^m} \geq 0 \quad \forall mwij, i \neq j \tag{9}$$

$$\frac{\partial L}{\partial x_{jj}^{mw}} = \frac{1}{\alpha^m g^m} \left(\ln \frac{x_{jj}^{mw}}{x_{jj}^m} + 1.0 \right) - \theta_{jj}^m + \frac{d_{jj}^w}{g^m} \geq 0 \quad \forall mwj \tag{10}$$

$$\frac{\partial L}{\partial x_{ij}^m} = \frac{1}{\beta^m g^m} \left(\ln \left(\frac{x_{ij}^m}{\bar{X}_i^m} \right) + 1.0 \right) + \sum_l \gamma_l^l a^{lm} - \gamma_j^m + \theta_{ij}^m - \frac{1}{\alpha^m g^m} \sum_w \frac{x_{ij}^{mw}}{x_{ij}^m} \geq 0 \quad \forall mij \tag{11}$$

The complementary slackness conditions are:

$$h_{ijr}^{mw} \frac{\partial L}{\partial h_{ijr}^{mw}} = h_{ijr}^{mw} \left\{ \sum_a d_a(f_a^w) \varphi_{ijr}^{aw} - d_{ij}^{mw} \right\} = 0 \quad \forall mwrij \tag{12}$$

$$x_{ij}^{mw} \frac{\partial L}{\partial x_{ij}^{mw}} = x_{ij}^{mw} \left\{ \frac{1}{\alpha^m g^m} \left(\ln \frac{x_{ij}^{mw}}{x_{ij}^m} + 1.0 \right) - \theta_{ij}^m + \frac{d_{ij}^{mw}}{g^m} \right\} = 0 \quad \forall mwij, i \neq j \tag{13}$$

$$x_{jj}^{mw} \frac{\partial L}{\partial x_{jj}^{mw}} = x_{jj}^{mw} \left\{ \frac{1}{\alpha^m g^m} \left(\ln \frac{x_{jj}^{mw}}{x_{jj}^m} + 1.0 \right) - \theta_{jj}^m + \frac{d_{jj}^w}{g^m} \right\} = 0 \quad \forall mwj \tag{14}$$

$$x_{ij}^m \frac{\partial L}{\partial x_{ij}^m} = x_{ij}^m \left\{ \frac{1}{\beta^m g^m} \left(\ln \left(\frac{x_{ij}^m}{\bar{X}_i^m} \right) + 1.0 \right) + \sum_l \gamma_l^l a^{lm} - \gamma_j^m + \theta_{ij}^m - \frac{1}{\alpha^m g^m} \sum_w \frac{x_{ij}^{mw}}{x_{ij}^m} \right\} = 0 \quad \forall mij \tag{15}$$

Conditions representing route flows and distances may be derived from Eqs. (8) and (12) as follows:

1. if $h_{ijr}^{mw} > 0$, then $\sum_a d_a^w(f_a^w) \phi_{ijr}^{aw} = d_{ij}^{mw}$;
2. if $h_{ijr}^{mw} = 0$, then $\sum_a d_a^w(f_a^w) \phi_{ijr}^{aw} \geq d_{ij}^{mw}$.

The interpretation of these conditions is as follows: if the route flow h_{ijr}^{mw} is positive, then $\sum_a d_a^w(f_a^w) \phi_{ijr}^{aw}$, the length of route r is equal to the equilibrium shipment distance d_{ij}^{mw} from i to j by mode w for sector m ; if the route flow h_{ijr}^{mw} is 0, then the length of route r is not less than the equilibrium shipment distance. Hence, Lagrange multipliers d_{ij}^{mw} have an interpretation of equilibrium shipment distances. These optimality conditions are equivalent to the equilibrium conditions proposed by Wardrop (1952) for user-optimal route choice of drivers in road networks.

3.3. Derivation of mode choice and composite cost functions

Assuming interregional commodity shipments by mode x_{ij}^{mw} are positive, the following model may be obtained from Eq. (13):

$$x_{ij}^{mw} = x_{ij}^m \cdot \exp(\alpha^m g^m \theta_{ij}^m - 1.0) \cdot \exp(-\alpha^m d_{ij}^{mw}) \quad \forall mwij, i \neq j \quad (16)$$

By substituting Eq. (16) into Eq. (3), we obtain:

$$x_{ij}^m = \sum_w x_{ij}^{mw} = x_{ij}^m \cdot \exp(\alpha^m g^m \theta_{ij}^m - 1.0) \cdot \sum_w \exp(-\alpha^m d_{ij}^{mw}) \quad (17)$$

Therefore,

$$\exp(\alpha^m g^m \theta_{ij}^m - 1.0) = \frac{1}{\sum_w \exp(-\alpha^m d_{ij}^{mw})} \quad (18)$$

$$x_{ij}^{mw} = x_{ij}^m \cdot \frac{\exp(-\alpha^m d_{ij}^{mw})}{\sum_w \exp(-\alpha^m d_{ij}^{mw})} \quad \forall mwij, i \neq j \quad (19)$$

Based on Eq. (18), θ_{ij}^m can be interpreted in relation to the composite shipment distance, \tilde{d}_{ij}^m :

$$\theta_{ij}^m = -\frac{1}{\alpha^m g^m} \ln \left\{ \sum_w \exp(-\alpha^m d_{ij}^{mw}) \right\} + \frac{1}{\alpha^m g^m} = \frac{1}{g^m} \tilde{d}_{ij}^m + \frac{1}{\alpha^m g^m} \quad \forall mij, i \neq j \quad (20)$$

if the composite shipment distance \tilde{d}_{ij}^m is defined as:

$$\tilde{d}_{ij}^m = -\frac{1}{\alpha^m} \ln \left\{ \sum_w \exp(-\alpha^m d_{ij}^{mw}) \right\} \quad \forall mij, i \neq j \quad (21)$$

3.4. Interregional commodity shipment functions

Again assuming that $x_{ij}^m > 0$, from Eqs. (3), (15) and (21) we obtain:

$$x_{ij}^m = \bar{X}_i^m \exp \left(-\beta^m g^m \sum_t \gamma_t^l a_t^{lm} - 1.0 \right) \exp(\beta^m g^m \gamma_j^m) \exp(-\beta^m \tilde{d}_{ij}^m) \quad \forall mij \quad (22)$$

By defining the following balancing factors,

$$\delta_i^m = \exp\left(-\beta^m g^m \sum_l \gamma_i^l a_i^{lm} - 1.0\right) \quad \forall mi \quad (23)$$

$$\varepsilon_j^m = \exp(\beta^m g^m \gamma_j^m) \quad \forall mj \quad (24)$$

a simpler form for x_{ij}^m is:

$$x_{ij}^m = \bar{X}_i^m \delta_i^m \varepsilon_j^m \exp(-\beta^m \tilde{d}_{ij}^m) \quad \forall mij, i \neq j \quad (25)$$

By substituting Eq. (25) into Eq. (19), an interregional commodity shipment function by mode can be obtained:

$$x_{ij}^{mw} = \bar{X}_i^m \delta_i^m \varepsilon_j^m \exp(-\beta^m \tilde{d}_{ij}^m) \frac{\exp(-\alpha^m d_{ij}^{mw})}{\sum_w \exp(-\alpha^m d_{ij}^{mw})} \quad \forall mwij, i \neq j \quad (26)$$

In the same way, the intraregional commodity shipment function by mode is:

$$x_{jj}^{mw} = \bar{X}_j^m \delta_j^m \varepsilon_j^m \exp(-\beta^m \tilde{d}_{jj}^m) \frac{\exp(-\alpha^m d_{jj}^w)}{\sum_w \exp(-\alpha^m d_{jj}^w)} \quad \forall mwij, i = j \quad (27)$$

where

$$\tilde{d}_{jj}^m = -\frac{1}{\alpha^m} \ln \left\{ \sum_w \exp(-\alpha^m d_{jj}^w) \right\} \quad \forall mj \quad (28)$$

Notice the role of the ratios of commodity value to weight (g^m) in the functional form of Eq. (27). Had these ratios not been included in the denominator terms together with the Lagrange multipliers (α^m) and (β^m), then the deterrence functions in the interregional commodity shipment and mode choice functions would include this ratio. Such a result would appear to have no meaningful interpretation. A comparable result is found in the auto occupancy factors of multiclass travel forecasting models; cf. Boyce and Bar-Gera (2003).

3.5. Solution algorithm

A generalization of the algorithm proposed by Evans (1976) was used to solve the optimization problem. In this case each iteration of the algorithm uses Wilson's (1970) iterative balancing method to generate the subproblem interregional commodity shipments, and the all-or-nothing assignment method to find the subproblem network link flows.

The convergence of Wilson's iterative balancing method can be judged by the relative error between the observed and calculated final demands for each sector and each region, or the relative change of balancing factors. The former convergence criterion was used successfully with 0.001 as the stopping value. Unlike the experience of Rho et al. (1989), convergence was satisfactory.

4. Parameter estimation

4.1. Mode choice

The mode choice function was specified with the following utility functions for each commodity:

$$U_t = b_t^v v_t + b_t^d d_t + C \quad (29)$$

$$U_r = b_r^v v_r + b_r^d d_r \quad (30)$$

where U_t is the utility by truck, v_t is the average value of shipments by truck (\$/ton), d_t is the average shipment distance by truck (miles), C is a constant (intercept), U_r is the utility by railway, v_r is the average value of shipments by railway (\$/ton), d_r is the average shipment distance by railway (miles), and (b_w^v, b_w^d) represent the coefficients to be estimated. To estimate the coefficients of the logit model, we applied linear regression analysis to the natural logarithm of the ratio of the probabilities for the road and railway modes as follows:

$$\frac{P_t}{P_r} = \frac{P_t}{1 - P_t} = \frac{e^{U_t}}{e^{U_r}}$$

Taking the natural logs of both sides, we obtain:

$$\ln \left(\frac{P_t}{1 - P_t} \right) = (C + b_t^v v_t + b_t^d d_t) - (b_r^v v_r + b_r^d d_r) \quad (31)$$

The coefficients of Eq. (31) were estimated using 1993 Commodity Flow Survey (CFS) data. The independent variables are average values of shipments and average shipment distances by truck or railway, and the dependent variable depends on the proportions shipped by truck. The available observations were for states of shipment origin to all destinations by commodity, the most modal detailed data released for the 1993 CFS. In view of the limited number of observations, the data were pooled across commodities by defining dummy variables for the 11 sectors of the 13 sectors used; sectors 3 (construction) and 13 (service) are excluded from the analysis. The descriptions of the sectors included in the analysis are shown in Table 1. Following the pooling of the observations across sectors, the equation to be estimated is:

$$\ln \left(\frac{P_t^m}{1 - P_t^m} \right) = \sum_m D_m \delta_m + (b_t^v v_t^m + b_t^d d_t^m) - (b_r^v v_r^m + b_r^d d_r^m), \quad m \in M \quad (32)$$

where $\delta_m = 1$ if the observation belongs to commodity sector m , and 0 otherwise. The results shown in Table 2 suggest that the share of the commodity shipment shipped by truck increases with mean value of shipment and decreases with shipment distance. The opposite effects are found for railway, although these coefficient values are not significantly different from 0. The regression analysis yielded a coefficient of determination R^2 of 0.83, indicating that the estimated commodity shipments by mode explain 83% of the variation in observed modal commodity shipments.

The resulting distance function inserted into the combined model is as follows:

$$d_{ij}^{mw} = \sum_m D_m \delta_m + b_w^v \bar{v}_w^m + b_w^d d_{ij}^w = C^{mw} + b_w^d d_{ij}^w$$

Table 2
Estimated values of mode choice coefficients, standard errors and *t*-values

Coefficient	Estimated value	Standard error	<i>t</i> -value
D_1	1.131	0.3699	3.06
D_2	1.067	0.2480	4.30
D_4	3.523	0.3057	11.53
D_5	2.503	0.2917	8.58
D_6	3.043	0.3447	8.83
D_7	4.038	0.4462	9.05
D_8	1.567	0.9910	1.58
D_9	1.661	0.6439	2.58
D_{10}	0.659	0.5134	1.29
D_{11}	2.612	0.2638	9.90
D_{12}	3.095	0.2377	13.02
C	Value merged into dummy variables		
b_t^{va}	+0.375	0.0533	7.04
b_t^d	-2.457	0.6818	-3.59
b_r^v	-0.0754	0.0758	-0.99
b_r^d	+0.290	0.1948	1.49

$R^2 = 0.83$; $F = 93.5$.

^a Note: Each coefficient value and standard error shown for the four shipment value and distances variables is multiplied by 1000 to conserve space.

where \bar{v}_w^m is the mean value of commodity shipments by sector and mode for the US. Therefore, in the model implementation, the modal distance sensitivity parameters α^m were set equal to 1.0, since the modal distance sensitivity is represented by the regression coefficients.

4.2. Interregional commodity shipments

Estimation of the interregional commodity shipment function consists of determining the values of the distance sensitivity parameters (β^m) associated with the interregional shipment entropy function in the objective function. Typically, (β^m) are estimated on observed values of interregional shipment costs or distances. For the CFS no interregional mean shipment distances were available that were considered sufficiently stable for the estimation procedure. Therefore, estimates of the interregional shipment distances were extracted from the model solution, and used with the interregional commodity shipments to compute mean shipment distances.

For this initial model implementation, only one distance sensitivity value β was estimated; the sensitivity values for the individual sectors were computed as $\beta^m = \beta/g^m$ using the ratio g^m for each sector. The quadratic interpolation method of Williams (1976) was applied iteratively to determine the value of β that best equated the predicted and observed mean shipment distance over all sectors.

The values of the ratio of value to weight and the commodity-specific distance sensitivities for no capacity limit and for two forms of the link performance function are shown in Table 3. The sector specific sensitivities show a reasonable variation in relation to the sector shipment value to weight ratios. Little difference is noted for the assumption of unlimited capacity vs. the effective distance function that is increasing with flow, since the network is relatively uncongested. The

Table 3

Beta values by sector and goodness-of-fit measures for three values of the link performance function parameter P : $d_a^i = d_a(1 + 0.15(f_a/k_a)^P)$; k_a = link capacity

Sector	g^m (\$/ton)	β/g^m ($P = 0$)	β/g^m ($P = 4$)	β/g^m ($P = 8$)
1	243.20	2.944	3.021	3.012
2	15.821	44.69	45.86	45.73
4	1006.4	0.7026	0.7210	0.7188
5	1048.1	0.6746	0.6922	0.6902
6	842.98	0.8387	0.8607	0.8581
7	2828.1	0.2500	0.2566	0.2558
8	12,729	0.0555	0.0570	0.0568
9	13,684	0.0517	0.0530	0.0529
10	6808.5	0.1039	0.1066	0.1063
11	392.40	1.802	1.849	1.844
12	14,638	0.0483	0.0496	0.0494
β		707.0	725.6	723.38
ρ^2		0.84	0.85	0.85
R^2		0.62	0.63	0.63

values of R^2 and ρ^2 indicate a reasonable level of goodness of fit for a model based on highly aggregated data.

5. Predicted commodity shipments

The model predicts value of commodity shipments by sector (\$) for the originating and terminating shipments by region, interregional and intraregional shipments, and mode shares by sector; physical flows (tons) are predicted for road and railway links, as well as interregional shipment distances. The commodity shipments for sector 3 (construction) and sector 13 (services) are excluded from the following tables because these sectors only produce for local final demand in each region. The observed and predicted total commodity shipments for the US by sector are shown in Table 4. Although the predicted shares by sector for sectors 5, 6, 7, 9 and 12 are quite similar to the observed shares, there are major discrepancies in sectors 1, 2, 4, 10 and 11. The overall predicted value of commodities is about 8% too low.

The model also predicts shares of commodity shipments by mode of transportation, road and railway. As shown in Table 5, the correspondence of the observed and predicted mode shares varies substantially by sector. The overall share of commodity shipment by railway is about double the observed share, corresponding to substantial overpredictions by sector. Still, given the paucity of data, and its reliability, and the simplicity of the model, the results may be considered as promising.

6. Summary and future research

A combined input–output and transportation network model was implemented and estimated to predict interregional commodity shipments and transportation network flows by sector and mode. Thus, the model integrates the traditional steps of generation, distribution, mode choice,

Table 4
Total observed and predicted commodity flows by sector

Sector	Observed		Predicted		Difference in shares
	Shipments (billion \$)	Sectoral share (%)	Shipments (billion \$)	Sectoral share (%)	
1	121.6	2.9	264.1	6.7	-3.8
2	39.6	0.9	206.6	5.3	-4.4
4	814.7	19.1	521.1	13.3	5.8
5	460.7	10.8	398.6	10.1	0.7
6	185.2	4.4	153.5	3.9	0.5
7	211.0	5.0	149.1	3.8	1.2
8	338.9	8.0	200.0	5.1	2.9
9	323.4	7.6	340.3	8.7	-1.1
10	460.6	10.8	303.0	7.7	3.1
11	801.9	18.8	992.0	25.3	-6.5
12	496.7	11.7	396.8	10.1	1.6
Total	4254.1	100.0	3925.0	100.00	0.0

Table 5
Total commodity shipments by sector and mode

Sector	Road				Railway			
	Shipments (billion \$)	Sectoral share (%)	Mode shares (%)		Shipments (billion \$)	Sectoral share (%)	Mode shares (%)	
			Predicted	Observed			Predicted	Observed
1	166.7	5.4	63.1	72.2	97.3	11.4	36.9	27.8
2	144.2	4.7	69.8	49.0	62.4	7.3	30.2	51.0
4	455.7	14.8	87.5	96.0	65.4	7.7	12.5	4.0
5	307.1	10.0	77.0	87.4	91.5	10.7	23.0	12.6
6	131.1	4.3	85.4	89.6	22.4	2.6	14.6	10.4
7	132.0	4.3	88.5	96.9	17.1	2.0	11.5	3.01
8	142.8	4.6	71.4	90.0	57.3	6.7	28.6	10.0
9	278.3	9.1	81.8	97.8	62.1	7.3	18.2	2.2
10	207.7	6.8	68.6	81.7	95.2	11.2	31.4	18.3
11	844.0	27.5	85.1	90.1	148.0	17.4	14.9	9.9
12	262.5	8.5	66.2	94.0	134.3	15.7	33.8	6.0
Total	3072.0	100.0	78.3	90.3	853.0	100.0	21.7	9.7

and assignment into a single procedure. The model was implemented using input–output coefficients, final demands, and the total commodity shipments for 13 sectors and 36 regions as well as the US road and railway networks for.

Regarding the estimation of model parameters, an overall distance sensitivity value β was determined using the observed interregional commodity shipments. The model with a single β value showed a goodness-of-fit between the observed and predicted interregional shipments of more than 80% for ρ^2 and more than 60% for R^2 . The model includes a logit mode choice function based on shipment value and distance. The regression model derived for this functions has an R^2 of 83%.

The model was successfully solved using Wilson's iterative balancing method and partial linearization algorithm of Evans. The model was implemented with simplified transportation networks to estimate intraregional and interregional shipments, link flows, and shipment lengths for each sector and mode. The predicted commodity shipments appear to compare reasonably well with the observed commodity shipments, considering the aggregate nature of the regions and sectors.

Suggestions for future research are the following:

1. Experience with various regional systems strongly suggests that the regions should be similar in geographic size. That is, very small and very large regions should not be included in the same model. If predictions are desired for small regions, then a system of large regions divided into smaller subregions should be devised, and the model refined accordingly to represent both regional and subregional shipments. The use of prior total shipments helps to mitigate this problem, but does not solve it.
2. To represent transportation generalized costs better, transit time, in addition to shipment distance, should be included in the transportation network representation. In this way the conceptual notion of an effective shipment distance, which is increasing with flow, can be eliminated.
3. Distance (or cost) sensitivity parameter values corresponding to the entropy functions should be determined for each sector. While the present approach produced reasonable initial results, it should not be regarded as a satisfactory procedure in general.

Acknowledgements

The support of the Earthquake Engineering Research Centers Program of the National Science Foundation, Award Number EEC-9701785 through the Mid-America Earthquake (MAE) Research Center, and the National Institute of Statistical Sciences, under National Science Foundation award DMS 9313013, is gratefully acknowledged. We are grateful for the critical comments and helpful suggestions of two anonymous referees.

References

- Batten, D.F., Boyce, D.E., 1986. Spatial interaction, transportation, and interregional commodity shipment models. In: Nijkamp, P. (Ed.), *Handbook of Regional and Urban Economics*, vol. 1. North-Holland, Amsterdam, pp. 357–406.
- Boyce, D., Bar-Gera, H., 2003. Validation of multiclass urban travel forecasting models combining origin–destination, mode, and route choices. *Journal of Regional Science* 43, 517–540.
- Boyce, D.E., Daskin, M.S., 1997. Urban transportation. In: ReVelle, C., McGarity, A. (Eds.), *Design and Operation of Civil and Environmental Engineering Systems*. Wiley, New York, pp. 277–341 (Chapter 7).
- Boyce, D.E., Hewings, G.J.D., 1981. Interregional commodity flow, input–output and transportation modeling: an entropy formulation. Presented at the Conference on Interregional Models, First World Regional Science Congress, Cambridge, MA, 1980.
- Evans, S.P., 1976. Derivation and analysis of some models for combining trip distribution and assignment. *Transportation Research* 10, 37–57.
- Fernandez, J.E., de Cea, J., Soto, A., 2003. A multi-modal supply-demand equilibrium model for predicting intercity freight flows. *Transportation Research* 37B, 595–614.

- Isard, W., 1960. *Methods of Regional Analysis*. MIT Press, Cambridge.
- Kim, T.J., Boyce, D.E., Hewings, G.J.D., 1983. Combined input–output and commodity flow models for interregional development planning: insights from a Korean application. *Geographical Analysis* 15, 330–342.
- Kim, T.J., Ham, H., Boyce, D.E., 2002. Economic impacts of transportation network changes: implementation of a combined transportation and input–output model. *Papers in Regional Science* 81, 223–246.
- Leontief, W.W., Strout, A., 1963. Multiregional input–output analysis. In: Barna, T. (Ed.), *Structural Interdependence and Economic Development*. Macmillan, London.
- Lin, J., Kockelman, K., Zhao, Y., 2002. Tracking land use, transport, and industrial production using random-utility based multizonal input–output models: applications for Texas trade. Working Paper, Department of Civil Engineering, University of Texas, Austin.
- Oppenheim, N., 1995. *Urban Travel Demand Modeling*. Wiley, New York.
- Rho, J.H., Boyce, D.E., Kim, T.J., 1989. Comparison of solution methods for Wilson’s interregional commodity flow model. *Geographical Analysis* 21, 259–267.
- Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. In: *Proceedings of the Institution of Civil Engineers, Part II*, vol. 1, pp. 325–378.
- Williams, I., 1976. A comparison of some calibration techniques for doubly constrained models with an exponential cost function. *Transportation Research* 10, 91–104.
- Wilson, A.G., 1970. Interregional commodity flows: entropy maximizing procedures. *Geographical Analysis* 2, 255–282.
- Zhao, Y., Kockelman, K.M., 2004. The random-utility based multiregional input–output model: solution existence and uniqueness. *Transportation Research, Part B*, in press, doi:10.1016/j.trb.2003.10.005.