TRACKING TRANSPORTATION AND INDUSTRIAL PRODUCTION ACROSS A NATION: APPLICATION OF THE RUBMRIO MODEL FOR U.S. TRADE PATTERNS

Xiaochuan Du
Graduate Student Researcher
Southeast University
Nanjing, China
dxc@seu.edu.cn

Kara M. Kockelman
(Corresponding author)
Professor and William J. Murray Jr. Fellow
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
6.9 E. Cockrell Jr. Hall
Austin, TX 78712-1076
kkockelm@mail.utexas.edu
Phone: 512-471-0210
FAX: 512-475-8744

The following paper is a pre-print and the final publication can be found in
Transportation Research Record No. 2269: 99-109, 2012
Presented at the 91st Annual Meeting of the Transportation Research Board, January 2012

ABSTRACT

This study describes and applies a random-utility-based multiregional input-output (RUBMRIO) model for U.S. transportation, production, and trade across over-3,000 contiguous counties, using the Freight Analysis Framework as its primary data source. Driven by foreign export demands, RUBMRIO simulates trade patterns of commodities among counties based on input-output expenditure shares and a nested-logit model for shipment origins and mode. A variety of network and export-demand scenarios are examined, for their effects on the distributions of trade flows and production.

Changes in export demands of different commodities highlight the importance of food and petroleum manufacturing sectors, in terms of production and labor-expenditure shifts. Transport cost reductions result in greater effects on total production than similar cost increases, with the most impacted U.S. counties centrally located. Changes in travel times along the Interstate Highway 40 corridor have ripple effects, affecting trade patterns everywhere, with the greatest changes observed around the corridor's midpoint.

Key words: spatial input-output model, nationwide trade flow patterns, integrated transportation-land use modeling

1. INTRODUCTION

The spatial structure and cost implications of transportation systems affect household and firm location choices, production levels, and trade patterns, in multiple ways. Such agents' choices manifest themselves in loads on the system, impacting network performance. To recognize this critical interaction and enhance planning, policy and investment decisions, integrated models of transportation and land use have been pursued.

Input-Output (IO) models are popular for simulating expenditure linkages across industries, and among producers and consumers. These models are demand driven, in the sense that production levels adjust to meet both final and intermediate demands. Traditional IO models have been extended to incorporate spatial disaggregation, leading to models like MEPLAN (*1, 2, 3, 4*), TRANUS (5), PECAS (6) and RUBMRIO (*7, 8*). These can be made dynamic, by allowing the travel costs associated with freight and person (labor and customer) flows to affect location and land use decisions in the model's next iteration, along with network system changes (e.g., roadway expansions) and exogenous economic shocks (e.g., increases in export demands).

Other spatial IO applications also exist. For example, Kim et al. (9) developed such a model for estimating interregional commodity flows and transportation network flows to evaluate the indirect impacts of an unexpected event (an earthquake) on nine U.S. states, with the US represented by 36 zones.. Canning and Wang (*10*) tested an IO program for international, inter-industry transactions across four regions and ten sectors using a global database documented in McDougall et al. (*11*). Rey and Dev (*12*) introduced a series of specifications for extra-regional trade, linking econometric and IO methods, and thus extending the traditional multiregional IO framework (which presumes fixed inter-zonal flow shares, much like an IO model, but with more rows and columns, for zone-pair trade dependencies). Ham et al. (*13*) estimated interregional, multimodal commodity shipments via an equivalent optimization, adding interregional and modal dispersion functions to their system's objective function.

This paper builds on Kockelman et al.'s work (*7*), which developed a Random-Utility-Based Multiregional Input-Output (RUBMRIO) model for Texas trade. Their study calibrated and applied a RUBMIRO model for Texas' 254 counties, across 18 social-economic sectors and two transportation modes, in order to meet foreign export demands at 31 key ports. This model simulates industrial production and trade patterns of labor and commodities, as driven by export demands, with specific trade pattern responding to travel-based prices, measured in utility units and based on expected minimum transportation costs (represented by distance on a two-mode highway/railway network). Their applications reflected links' congestion levels, and illustrated the multiplier effects that demand shifts can have, by port and sector. Zhao and Kockelman (*8*) then proved how the general RUBMIRO formulation will converge, on a unique flow solution.

In the present study, a U.S.-level RUBMRIO model is developed for trade patterns among the nation's 3,109 contiguous counties (excluding Hawaii and Alaska), across 20 social-economic sectors, and 2 transportation modes. The model relies on principles of random utility maximization or random cost minimization to anticipate domestic trade flows (among counties, as zones) and export flows (from counties to export zones), as well as import flows (via purchases abroad, or "leakages"). The utilities (or negative costs) of acquiring commodity *m* from every possible provider county *i*, and transporting it via highway or rail, to use as an input of production in county *j* (or export it from zone *k*) are expressed as a function of travel costs and endogenously estimated sales prices (in utility units) of commodities in each origin county.

The paper begins with a review of spatial IO models and related modeling practice, and then describes the RUBMRIO structure, data set acquisition, and parameter estimation. The applications anticipate trade and location choices resulting from a variety of scenarios, including changes in export demands and transport cost, and the paper concludes with a discussion of application results and further modeling opportunities.

2. STRUCTURE OF THE RUBMRIO MODEL

2.1 The Utility of Trade Choices

Applying random utility theory for cost minimization, both domestic trade flows (among counties, as zones) and export flows (from counties to export zones) are based on the utility (or disutility) of acquiring commodity *m* from every possible provider county *i*, and transporting it via highway or rail, to use as an input of production in zone *j* (or export it from zone *k*). In contrast to earlier applications of the RUBMRIO model (*7, 8, 26*), the utility expressions used here include both origin population as a “size” factor (to acknowledge current population’s [and employment’s] large role in trade patterns) and travel time and cost attributes between zones (rather than distance), as shown in Equations (1) and (2).

$$U_{ij}^{m} = -p_{i}^{m} + \gamma^{m}\ln(pop_{i}) + \lambda^{m}\ln\left[\sum_{\mathrm{t}}\exp\left(\beta_{0,t}^{m} + \beta_{1}^{m}\cdot time_{ij,t} + \beta_{2}^{m}\cdot cost_{ij,t}\right)\right] \quad (1)$$

$$U_{ik}^{m} = -p_{i}^{m} + \gamma^{m}\ln(pop_{i}) + \lambda^{m}\ln\left[\sum_{\mathrm{t}}\exp\left(\beta_{0,t}^{m} + \beta_{1}^{m}\cdot time_{ik,t} + \beta_{2}^{m}\cdot cost_{ik,t}\right)\right] \quad (2)$$

where p_{i}^{m} is the sales price of commodity *m* in county/zone *i* (in units of utility, as computed below), pop_{i} is the population of zone *i*, and $time_{ij,t}$ and $cost_{ij,t}$ represent the travel times and costs between zones *i* and *j* by mode *t*. Parameters γ^{m}, λ^{m}, and β^{m} were estimated using a series of industry-specific nested logit specifications (as described in *14*), as discussed later in the paper.

2.2 Production Function

Sales price is a key factor influencing purchase choices and production costs, and thus trade patterns. In the RUBMRIO model, sales price (the cost of producing 1 unit of commodity *n* in zone *j*) depends on the costs of purchasing raw materials, labor, and necessary services from other producers, including transport costs associated with the shipment of those inputs. The overall manufacture cost and ultimate sales price of industry *n*’s output from zone *j* is as follows:

$$p_{j}^{n} = \sum_{m}\left(A_{0j}^{mn}\times c_{j}^{m}\right) \quad (3)$$

where A_{0j}^{mn} are the set of technical coefficients for producing good type *n* in zone *j*. These are the shares of commodity *m* required to produce 1 unit of commodity *n* in zone *j* (in terms of dollar-per-dollar, and so are dimensionless), and c_{j}^{m} is the average cost of acquiring input *m* for productive use in zone *j*.

The base IO-model parameters, A_{0j}^{mn}, represent direct backward linkages of an industry *n* to upstream industries (*m*), regardless of location, thereby constituting the "recipe" for production in industry *n*. They can be calculated based on a transactions table (input-output matrix of dollar flows between industries) by dividing each *m,n* cell's transaction by its corresponding column total.

The input costs, c_j^m, are a flow-weighted average of purchase and transport costs (in units of utility) of input commodity *m* from all origins *i* to zone *j*, as represented by -U_{ij}^m. The weights are domestic trade flows, X_{ij}^m, as defined below and shown in Equation (4).

$$c_j^m = \frac{\sum_i \left[X_{ij}^m \times \left(-U_{ij}^m\right)\right]}{\sum_i X_{ij}^m} \tag{4}$$

2.3 Trade Flows

Both domestic and export trade flows are calculated under an assumption of utility-maximizing/cost-minimizing behavior, which means consumers will choose producer(s) that can supply the lowest price (including transport cost) in order to maximize their utility/minimize their costs. The unobserved heterogeneity of this choice, across producers and consumers, introduces the random elements, which, under an iid Gumbel distributional assumptions, leads to a nested logit model for origin and mode choices. The domestic trade flow, X_{ij}^m, and export trade flow, Y_{ik}^m, are computed using Equations (5) and (6):

$$X_{ij}^m = C_j^m \frac{\exp\left(U_{ij}^m\right)}{\sum_i \exp\left(U_{ij}^m\right)} \tag{5}$$

$$Y_{ik}^m = Y_k^m \frac{\exp\left(U_{ik}^m\right)}{\sum_i \exp\left(U_{ik}^m\right)} \tag{6}$$

where $\mathrm{Y_k^m}$ is the demand for commodity *m*, as exported via zone *k*, and $\mathrm{C_j^m}$ is the total (dollar) amount of *m* consumed in zone *j*, which can be calculated as follows:

$$C_j^m = \sum_n \left(A_j^{mn} \times x_j^n\right) \tag{7}$$

Here, A_j^{mn} represents "local-purchase" technical coefficient for zone *j*. Unlike Equation (3)'s A_{0j}^{mn} values, Equation (7)'s *A* matrix relies only the amount of commodity *m* required from within the modeled region (the continental U.S. in this application). Any amount of *m* imported from foreign countries is excluded. Regional purchase coefficients (RPCs) bridge these two styles of *A* matrices by representing the proportion of total demand for a commodity that is supplied by producers within the study area, rather than imported from abroad (*15*). This relationship is shown in Equation (8). Finally, x_j^n is the total production of commodity *n* in zone *j*, which is the sum of domestic and export flows "leaving" zone *i* (though must also heads to zone *i* industries and consumers), as shown in Equation (9).

$$a^{mn} = \frac{X^{mn} \times RPC^{n}}{\sum_m X^{mn}} \quad (8)$$

$$x_j^m = \sum_j X_{ij}^m + \sum_k Y_{ik}^m \quad (9)$$

2.4 Solution Procedure

Equations 1 through 8 constitute the majority of the RUBMRIO model, and they are solved iteratively (using open-source C++ code) to achieve an equilibrium trade pattern, as described by Zhao and Kockelman (*8*), and shown in Figure 1. The iteration procedure begins with initial sales prices at zero, to quickly compute the utility of both domestic and export origin and mode choices. Then, (exogenously provided) export demands are distributed among production zones (according to the relative utilities of competing suppliers and modes). These export flows give rise to domestic demands and trade flows between counties (similarly distributed, on the basis of relative utilities). At each iteration, the total productions (outputs) of each zone *i* are multiplied by corresponding technical coefficient (following import/leakage considerations) in order to estimate the total consumption (set of input) required for purchased from domestic counties *j* (including zone *i* itself). Average input costs are computed as a flow-weighted average of utilities, and coupled with original technical coefficients to provide updated sales prices, which feedback for recalculating of all purchase utilities, and lead a new iteration, until consecutive trade flows stabilize (such that relative errors for each domestic flow value are less than l%), achieving system equilibrium.

3. DATA ACQUISITION AND PARAMETER ESTIMATION

Various data sets were used to calibrate and run the model for U.S. trade flows. The primary data source is the U.S. Department of Transportation's Freight Analysis Framework version 3 (FAF^3) database of networks and flows between FAF regions. IMPLAN's industry-by-industry transaction table and regional purchase coefficients for U.S. sectors in year 2008 were also used[1], along with TransCAD 4.0's railway network and demographic and county-boundary data.

FAF integrates data from a variety of sources to create a comprehensive picture of freight movement among states and major metropolitan areas by all modes of transport. Based on data from the U.S. 2007 Commodity Flow Survey and other sources, FAF^3 provides estimates for tonnage and value, by commodity type, mode, origin, and destination for year 2007 flows, and forecasts through 2040. Also included are truck flows assigned to the U.S. highway network for 2007 and 2040 (*16*). FAF^3's origin-destination-commodity-mode (ODCM) annual freight flows matrix was used to estimate RUBMRIO's nested logit model's origin and mode choice parameters, to calculate all export demands (by port and industry), and evaluate RUBMRIO model predictions.

Travel times and costs were computed for the 3,109 x 3,109 county-to-county matrix based on the shortest inter-county network distances for highway and railway modes, as were

---

[1] IMPLAN (Impact Analysis for Planning) is a social accounting and economic impact analysis software system, created by the Minnesota IMPLAN Group (MIG).

calculated by TransCAD on FAF^3's extensive (170,994-link) highway network and TransCAD's (16,552-link) railway network. All intra-county travel distances were assumed to be the radii of circles have the same areas as the original county (with an average county area of 966.3 square miles). Due to computer-memory limitations (relative the very large number of links in the FAF network), congested-travel time feedbacks are not tracked in the current model implementation. Moreover, travel times and costs are not available for this FAF network for trucking and rail models. As a result, travel times and costs used here are fixed and estimated based on shortest-path distances under a series of assumptions, as described in section 3.3.

3.1 Export Data

FAF^{3}estimates freight flows (in annual tons and dollar values) between FAF's 123 domestic analysis zones (averaging 25,035 square miles each), plus 8 foreign regions, across 43 commodity class, 8 transportation modes, and 3 trade types (export, import, and domestic) (*17*). FAF^3 shows foreign export flows exiting the U.S. via 106 of the 123 zones (with 3 zones in Alaska and Hawaii excluded here). The annual export dollar values of these 106 zones total to $1.10 trillion (with Sectors 9 and 10, for Machinery Manufacturing and Computer, Electronic Product and Electrical Equipment Manufacturing, respectively, enjoying the biggest shares of this total, at 17.54% and 18.74%, respectively) and drive the RUBMRIO model system, with production satisfied by outputs of industry across 3,109 counties (with zone centroids representing the locations of export. Figure 2 shows FAF3's 120 (continental) domestic zones and 106 export-zone centroids.

3.2 Estimation of Technical Coefficients

Technical coefficients A^{mn} reflect production technology or opportunities within counties and are core parameters in any IO model. Here, these coefficients are assumed stable over space and time, and provided exogenously, based on IMPLAN's transaction tables, as derived from U.S. inter-industry accounts. As shown in Table 1, IMPLAN's 440-sector transaction table (*18*) was collapsed into 18 industry sectors plus Household and Government sectors to represent the U.S. economy. Since FAF^3 uses the same 43 two-digit Standard Classification of Transported Goods (SCTG) classes (*19*) as the 2007 U.S. Commodity Flow Survey (CFS), IMPLAN's 440 sectors were bridged to a corresponding SCTG code based on the 2007 North American Industry Classification System or NAICS (*20*).

As introduced in section 2.3, RPCs represent the share of local demand that is supplied by domestic producers. These RPCs are generated by IMPLAN automatically, using a set of econometric equations (*21*). As shown in Equation (8), the original industry transaction tables were multiplied by RPCs to recognize the effects of imports, which lead the "leakages" or consumption losses from counties across their borders. For U.S. purchases as a whole, leakages average 46.96% in the highest-leakage/biggest-importing industry (Sector 14: Transportation, Communication and Utilities) and are zero percent in the lowest three (Sector 4-Food, Beverage and Tobacco Product Manufacturing, Sector 15-Wholesale Trade, and Sector 16-Retail Trade). Of course, counties/zones closer to international borders are more likely to leak than those located centrally, ceteris paribus (and production technologies will vary across counties), but these variations are not known, so a constant RPC value was used in all counties.

3.3 Estimation of Origin and Mode Choice Parameters

Transport cost can be critical to choice of an input's origin and shipping mode. As introduced in Equations 1 and 2, parameters γ^m, λ^m, and β^m reflect producers' and shippers' attraction to an origin zone's size and prominence (proxied by its current population) and sensitivity to travel times and costs of the two alternative modes (highway and railway), for each commodity m. To estimate such parameters for the nested logit model structure (with lower level for mode choice and upper level for origin choice), FAF^3's dollar values of freight flows between 120 domestic zones were used, for the 12 SCTG codes closest to the model's final 20 economic sectors (as shown in Table 1). FAF's commodity-based categories were matched to IMPLAN's industry-based categories by anticipating each commodity's final industrial producer. Since the SCTG codes do not match all NAICS and IMPLAN codes and not all industries ship commodities (e.g., Construction), there are not enough categories in FAF^3 data sets to match this application's 18 industry sectors. Here, sector 3 (Construction) was assumed to share sector 2's (Mining) parameters. And parameters from sectors 14 (Transportation, Communication and Utility), 15 (Wholesale Trade), and 16 (Retail Trade) were assumed to be the average of all other sectors' parameter values. Household and Government purchases were assumed to be strictly local in the calculations. Each FAF record was used as an data point or "observation", and its dollar value used as the "weight" factor in the logit's log-likelihood function.

In the lower layer of the nested logit model, mode choices were first estimated for each of the 12 sectors *m*. Recognizing that heavy-truck and rail modes carry 40.1% and 40.2% of the U.S.'s 3,344 billion ton-miles of traded commodities (according to the 2007 Commodity Flow Survey (*22*)), the RUBMRIO model used here includes just two modes: truck and rail, and other modes such as water, air, and multiple mode (with shares of 4.7%, 0.1%, 12.5%, respectively) are excluded. The explanatory variables are travel time and cost between counties (and from counties to export zones), based on shortest-path distances over TransCAD's highway and railway networks. For sector *m*, the probability of choosing transport mode *t* between origin *i* and destination *j* is as follows:

$$P_{t|ij}^m = \frac{\exp\left(V_{ij,t}^m\right)}{\sum_s \exp\left(V_{ij,s}^m\right)} \qquad (10)$$

The systematic (non-random) conditional indirect utility $V_{ij,t}^m$ is given by:

$$V_{ij,t}^m = \beta_{0,t}^m + \beta_1^m \cdot time_{ij,t} + \beta_2^m \cdot cost_{ij,t} \qquad (11)$$

where $time_{ij,t}$ and $cost_{ij,t}$ are the travel time and cost from *i* to *j* by mode *t*, and β's are mode choice parameters to be estimated (with $\beta_{0,rail}^m$ was set to zero, to permit statistical identification of all parameters). To compute $time_{ij,t}$ and $cost_{ij,t}$, the following assumptions were used:

1. Highway travel time is 3 hours of load-and-unload time, plus 1 hour of delay due to local-zone navigation, plus en-route travel time assuming an average truck speed of 45 miles/hour.

2. The highway network's shortest-path distances were used to compute highway travel costs by simply assuming an average marginal cost per truck-mile of $1.73 as estimated by the American Transportation Research Institute for year 2008 heavy-truck movements (*24*).

3. Railway travel times assumed 22 hours of terminal dwell time, plus an in-transit average train speed of 25 miles/hour, plus truck-based transshipment times to and from the nearest rail terminals locations, from and to the shipment's origin and final destination (assuming transship distances equal half the radii of each county's equivalent circle areas). While 22 hours may sound long, and 25 mi/hr may sound slow, both were estimated using the industry-shared Railroad Performance Measure (RPM)'s dataset (*23*).

4. The average cost of rail shipments was assumed to be $0.6 per mile, as implied by an American Association of State Highway and Transportation Officials' (AASHTO) report (*25*), which suggested that railway transport costs approximate one-third of highway costs, per mile. And the trans-shipment cost was included in total railway travel costs, following earlier assumptions.

In the upper layer, the probability of a producer in zone *i* choosing input *m* from firms in zone *j* is:

$$P_{ij}^{m} = \frac{\exp\left(V_{ij}^{m}\right)}{\sum_{i} \exp\left(V_{ij}^{m}\right)} \tag{12}$$

where V_{ij}^{m} is the expected maximum utility across mode alternatives plus the origin-size attractiveness term, as shown in Equation (13).

$$V_{ij}^{m} = \gamma^{m} \ln(pop_{i}) + \lambda^{m} \ln\left[\sum_{\mathrm{t}} \exp\left(V_{ij,t}^{m}\right)\right] \tag{13}$$

Table 2 shows all parameter estimates for the origin and mode choice models by sector. Since the SCTG codes do not match all NAICS and IMPLAN codes and not all industries ship commodities (e.g., Construction), there are not enough categories in FAF^3 data sets to match this application's 18 industry sectors. Here, sector 3 (Construction) was assumed to share sector 2's (Mining) parameters. And parameters from sectors 14 (Transportation, Communication and Utility), 15 (Wholesale Trade), and 16 (Retail Trade) were assumed to be the average of all other sectors' parameter values. Household and Government purchases were assumed to be strictly local in the calculations. Analysts can revise such assumptions, using the RUBMRIO open-source code and example data sets at http://www.caee.utexas.edu/prof/kockelman/RUBMRIO_Website.

4. SIMULATION RESULTS FOR APPLICATION SCENARIOS

4.1 Simulation Result

Using the data sources and estimated parameters described above, both the export trade flows from the 3109 county zones to the 120 export zones, and the domestic trade flows between the 3109 counties, were computed by the RUBMRIO model, Figure 1's iterative process. In order to meet the $11.1 trillion dollars of FAF^3 export demand, $10.1 trillion dollars of domestic trade flows were generated between (and within the 3109 counties), which is 77.6% of the 13.0

trillion domestic trade flows shown in FAF3. This gap in the magnitude of domestic flows between RUBMRIO and FAF^3 is expected, considering that import flows were not modeled explicitly, and all counties were assumed to have the same production process (as conveyed by technical coefficients), due to data limitations.

To further examine the base-case application's results, all export and domestic trade flows were summed within the corresponding (and more aggregate) FAF^3 zones. Table 3 shows some statistics for the ordered and then summed RUBMRIO trade flows and the corresponding FAF^3 data. These 10 categories of values indicate how many of the lowest-value 14,400 (120 x 120) domestic trade flows (aggregated across industry types) and lowest-value 12,720 (120 x 106) export flows are needed to hit the first ten percent of the respective total trade flow values, and then how many of the next-lowest set are needed to hit the 20 percent mark, and so forth. These types of cumulative counts can then be compared across the actual (FAF) flows and model-predicted (RUBMRIO) flows, and they show how RUBMRIO results in more low-value flows, resulting in somewhat (but not significantly) lower counts for the RUBMRIO Table 3 cell values in the second and higher (20%+) categories. Fortunately, the logit modeling approach still results in many high-value trades, so there is a reasonable mix.

As is typical of logit models (and regression models in general), RUBMRIO model distributes trade flows everywhere, somewhat smoothly over space, based on the relative utilities of purchase and transport (as described in section 2.3), rather than, say, existing trade relationships between big market players in specific locations. Thus, all counties are assigned some RUBMRIO trade flows, with many very low flows, in contrast to FAF^3 data, which offer fewer low-value flow pairs. If trade flows were micro-simulated to represent real trade agreements between individual market agents, the flows discretized, and the runs randomized (rather than producing average/expected-flow results), more variations in flow volumes would be evident in (each run of) the RUBMRIO model outputs. This type of work may make a nice extension to RUBMRIO, and is similar to some of the work being done with the PECAS model.

4.2 Foreign Export Demand Effects

Here, the foreign exports are assumed to be the only source of final demand, which must be satisfied by the U.S. counties. To forecast the effects of different export demands on the U.S. economy, a series of scenarios were carried out by changing the export demands in each of the 12 export-related sectors. A flow multiplier and a value-added multiplier were used as indicators of the marginal differences in domestic trade values and labor expenditures (as purchased from the household sector) due to a unit change in each export type, as shown in Equations 14 and 15.

$$\text{Flow Multiplier} = \frac{\text{change in trade flows (\$)}}{\text{change in export of specific commodity (\$)}} = \frac{\Sigma_{i,j,m} X_{ij}^{m'} - \Sigma_{i,j,m} X_{ij}^{m^0}}{\mathrm{Y}_{\mathrm{k}}^{\mathrm{m}'} - \mathrm{Y}_{\mathrm{k}}^{\mathrm{m}^0}} \qquad (14)$$

$$\text{Value Added Multiplier} = \frac{\text{change in purchases from household sector (\$)}}{\text{change in export of specific commodity (\$)}} \qquad (15)$$

$$= \frac{\sum_{i,j,m} X_{ij}^{m'} - \sum_{i,j,m} X_{ij}^{m^0}}{Y_k^{m'} - Y_k^{m^0}}$$

The simulated results show that the multiplier effects on both domestic trade flows and labor expenditures, with respect to foreign export of different sectors, vary substantially by sector, as shown in Figure 3. The multiplier values for trade flows ranged from 7.6 to 12.3 across export sectors, while the value-added multipliers exhibit a similar distribution among commodity types. Export demands for Commodities 4 and 5 (Food, Beverage and Tobacco Product Manufacturing and Petroleum and Coal Product Manufacturing, respectively) provide the greatest impacts/multipliers. These industry sectors are estimated to be the most sensitive to export demand changes, where decreases (or increases) could be most harmful (or beneficial) to the U.S. economy (and household incomes).

4.3 Highway Congestion Effects

As a key component of the utility functions, transport time is expected to affect trade flow patterns and local productions. In this study, Interstate Highway 40 (IH40), which runs east-west across the length of the U.S., from Atlantic to Pacific Coasts and is considered one of the nation's most important and busy freight corridors, was selected to examine congestion effects. Its travel times were increased 10% (and TransCAD's shortest path algorithm re-run) to represent added congestion or reduced capacity, and then reduced 10% to simulate added capacity.

Following a 10% increase in IH40 travel times, total production levels among the nation's top-ten most-affected counties fell 3.68 to 14.14%. In contrast, the 10% IH40 travel-time decrease top production increases of 3.32 to 5.23%, as shown in Table 4. Overall, the model results suggest that the increase in IH40 travel time had a larger impact on production, largely by reducing final export demands calling on counties along the IH40 corridor, as illustrated in Figure 4. These reductions in servicing export demands cause a chain reaction in intermediate input demands, which are met mainly by intra-county (rather than inter-county) production. As shown in Figure 4, central U.S. counties near the IH40 corridor's mid-section therefore experienced the greatest impacts, as may be expected. However, some production reductions (in the case of a travel time rise) center on the northwest region, far from the IH40 corridor. This is understandable when one recognizes that most high-value exports locate on the southeast region (as shown in Figure 2), and IH40 is one of the most critical corridors connecting the northwest and southeast. Thus, the rise in IH40 travel times strongly increased the disutility between them, resulting in a significant loss of production for the northwestern counties.

It is interesting that the effects are not symmetric when one considers the case of a 10% fall in IH40 travel times. While a mid-US cluster of counties emerged as most affected (in percentage terms), the northwest region showed no significant signs of benefit. Perhaps the trade routes were already most competitive and no obvious substitutions in trade emerge as the utilities to cross-US trade rise.

4.4 Transport Cost Effects

Transport cost is another key component of most any trade model, and can rise and fall relatively quickly in response to changing energy prices, labor costs, shipping regulations, and interest rates (which affect the real price of vehicle capital). Here, the marginal average cost of trucking (original cost is $1.73 per mile) was raised and lowered 20% (to $2.08 and $1.38 per mile, respectively) to examine its effects on U.S. trade and production patterns.

The nation's total production and consumption levels remain constant in this application, as expected, since the total export demand and production technologies are held constant. Thus, some counties loss production due to the rise of travel cost, and these losses are distributed to other counties who benefit from the rise of travel cost. As shown in Table 5, total production was predicted to rise sharply in many counties (with the top ten county-level increases ranging from 57.6% to 16.9%) in the case of transport cost reduction, and to fall somewhat less abruptly (with the top ten drops ranging from 36.2% to 9.98%) when operating costs rise. Figure 5 shows where these production changes take place. As expected, central U.S. counties are more affected by changes in trucking cost, since they generally have the longest distance to cover in meeting export-zone demands (since export-zones are primarily on the U.S. border). Such impacts raise the question of whether central-U.S. states should work harder to improve their networks (both rail and highway, as well as waterways, pipelines, and airport terminals) in order to better meet potential inter-regional trade demands, and thereby relatively dramatically improve their production levels (and their populations' employment and income levels). Interestingly, border states presumably have less incentive to improve their transportation systems.

5. CONCLUSIONS

This study establishes an open-source multiregional input-output model and associated inputs for trade forecasting at the national level, county to county, based on the principle of random trade-cost minimization. It also provides detailed trade predictions following model parameter estimation and application to the U.S. context. The simulated scenarios (of export demand changes, travel time changes, and trucking cost changes) offered reasonable, detailed, and meaningful estimates of production responses to shifts in various important inputs. They highlighted valuable dependencies, including the benefits of central-U.S. network investments. Such models should be able to assist nations, states and regions in appreciating their role in facilitating or hindering trade flows, interregional interactions, economic vitality, more sustainable mode choices, and energy-efficient trade patterns.

This study is an initial attempt to apply RUBMRIO to the complexity of a national-scale, 3109-county setting. The model specification (and associated code) for this large network context should be enhanced in several ways, including congested network assignment for travel time feedbacks (with dollar flows expressed as vehicle flows, similar to work by Ruiz-Juri et al. (*26*)), inclusion of import flow volumes, and the introduction of dynamic features to pivot off current trade relationships and move labor and capital across space in a reasonable fashion (as pursued by Huang and Kockelman (*27*)). Other possible extensions include use of market-clearing prices across industries and sites, using computable general equilibrium (CGE) concepts and techniques, recognition of other modes of travel and interzonal household and government travel patterns. Another term for the attractiveness of different origins, to better reflect the supply power of existing centers, should be pursued.

In summary, the nationwide RUBMRIO model offers a valuable set of relationships to predict trade flows, location choices/production levels, and relative market prices. Predictive models of this type and their quantification of effects are very valuable for assessing both national and regional transportation, land use, productive technology, and trade policies.

ACKNOWLEDGEMENTS

This material is based upon Dr. Kockelman and teammates' earlier work, as partly supported by U.S. National Science Foundation Grant No. 9984541. The authors thank the China Scholarship Council's State Scholarship Fund for the financial support of the first author's visit and research, and they appreciate those who have provided software, data, and technical consultation, including Drs. Ming Zhang, Frank Southworth, and Avinash Unnikrishnan.

REFERENCES

1. Echenique, M. H. The Use of Integrated Land Use Transportation Planning Models: The Cases of Sao Paolo, Brazil and Bilbao, Spain, in *The Practice of Transportation Planning*, M. Florian (ed.), The Hague: Elsevier, 1985.
2. Hunt, J. D. and Echenique, M. H. Experiences in the Application of the MEPLAN Framework for Land Use and Transportation Interaction Modeling. *Proc. 4th National Conference on the Application of Transportation Planning Methods*, Daytona Beach, FL, 1993, pp. 723-754.
3. Hunt, J. D. and Simmonds, D.C. Theory and Application of an Integrated Land-Use and Transport Modeling Framework. *Environment and Planning,* 20B, 1993, pp.221-244.
4. Abraham,J.E. and Hunt, J.D. Firm location in the MEPLAN model of Sacramento. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1685, Transportation Research Board of the National Academies, Washington, D.C., 1999, pp. 187-198.
5. De la Barra, T. *Integrated Land Use and Transport Modeling: Decision Chains and Hierarchies*. Cambridge University Press, New York, 1995.
6. Hunt, J.D., Abraham,J.E. Design and Application of the PECAS Land Use Modeling System. *Presented at the 8th Computers in Urban Planning and Urban Management Conference, Sendai, Japan*, 2003. Available at http://www.ucalgary.ca/~jabraham/Papers/pecas/8094.pdf. Accessed Jan. 20, 2011.
7. Kockelman, K.M., Jin, L., Zhao, Y., and Ruiz-Juri, N. Tracking Land Use, Transport, and Industrial Production Using Random-Utility Based Multiregional Input-Output Models: Applications for Texas Trade. *Journal of Transport Geography*, Vol. 13, No. 3, 2005, pp. 275-286.
8. Zhao, Y., Kockelman, K.M. The Random-Utility-Based Multiregional Input-Output Model: Solution Existence and Uniqueness. *Transportation Research Part B: Methodological*, Vol. 38, No. 9, 2004, pp. 789-807.
9. Kim, T.J., Ham, H. and Boyce, D.E. Economic impacts of transportation network changes: implementation of a combined transportation network and input-output model. *Papers in Regional Science*, Vol. 81, 2003, pp. 223-246.
10. Canning, P. and Wang, Z. A flexible mathematical programming model to estimate

interregional input-output accounts. *Journal of Regional Science*, Vol. 45, No. 3, 2005, pp. 539-563.

11. McDougall, R.A., Elbehri, A. and Truong, T.P. Global Trade Assistance and Protection: The GTAP 4 Database. Center for Global Trade Analysis, Purdue University. https://www.gtap.agecon.purdue.edu/databases/v4/default.asp . Accessed Jan. 15, 2011.
12. Rey, S.J. and Dev, B. Integrating econometric and input-output models in a multiregional context. *Growth and Change*, Vol. 28, 1997, pp. 222-243.
13. Ham, H. J., Kim, T. J. and Boyce, D. E. Implementation and estimation of a combined model of interregional, multimodal commodity shipments and transportation network flows. *Transportation Research Part* B 39, 2005, pp. 65–79.
14. Ben-Akiva, M., and Lerman, S.R. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, Massachusetts, 1985.
15. MIG, Inc. *The Controlled Vocabulary of IMPLAN-Specific Terms.* http://www.implan.com/v4/index.php?option=com_glossary&id=198. Accessed June. 30, 2011.
16. Federal Highway Administration, U.S. Department of Transportation. *Freight Analysis Framework.* http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm. Accessed Nov. 25, 2010.
17. Southworth, F., Davidson, D., Hwang, H., Peterson, B.E., and Chin, S. *The Freight Analysis Framework, Version 3: Overview of the FAF3 National Freight Flow Tables.* http://faf.ornl.gov/fafweb/Data/FAF3ODCMOverview.pdf. Accessed Nov. 25, 2010.
18. MIG, Inc. 440 IMPLAN Sector Scheme (2007 to current). http://implan.com/V4/index.php?option=com_docman&task=doc_dow. Accessed March. 19, 2011.
19. Bureau of Transportation Statistics, U.S. Department of Transportation. *2007 Commodity Flow Survey Standard Classification of Transported Goods (SCTG).* CFS-1200 (10-24-2006). http://www.bts.gov/publications/commodity_flow_survey/survey_materials/pdf/sctg_booklet.pdf. Accessed March, 19, 2011.
20. U.S. Census Bureau. *2007 North American Industry Classification System.* http://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2007. Accessed March. 19, 2011.
21. MIG, Inc. *Calculate the IMPLAN RPCs.* http://implan.com/V4/index.php?option=com_docman&task=doc_download&gid=125&Itemid=7. Accessed March. 19, 2011.
22. Bureau of Transportation Statistics, U.S. Department of Transportation. *2007 Commodity Flow Survey: United States.* EC07TCF-US, April, 2010. http://www.bts.gov/publications/commodity_flow_survey/final_tables_december_2009/pdf/entire.pdf . Accessed Sep. 3, 2010.
23. Railroad Performance Measures. http://www.railroadpm.org/Graphs/Terminal%20Dwell%20Graph.aspx. Accessed May. 14, 2011.
24. The American Transportation Research Institute. *An Analysis of the Operational Cost of Trucking.*

http://www.atri-online.org/research/results/economicanalysis/Operational_Costs_OnePager.pdf. Accessed May. 16, 2011.

25. American Association of State Highway and Transportation Officials. *Freight-Rail Bottom Line Report*. http://rail.transportation.org/Documents/FreightRailReport.pdf. Accessed May. 14, 2011.
26. Ruiz Juri, N., Kockelman, K.M. Extending the Random-Utility-Based Multiregional Input-Output Model: Incorporating Land-Use Constrains, Domestic Demand and Network Congestion in a Model of Texas Trade. Proceedings of the 83rd Annual Meeting of the Transportation Research Board, 2004.
27. Huang, T., Kockelman, K.M. The Introduction of Dynamic Features in a Random-Utility-Based Multiregional Input-Output Model of Trade, Production, and Location Choice. *Journal of the Transportation Research Forum*. Vol. 47, No. 1, 2008.

LIST OF TABLES AND FIGURES

TABLE 1 Description of Economic Sectors in RUBMRIO Model

Sector	Description	IMPLAN Code	NAICS Code	SCTG Code
1	Agriculture, Forestry, Fishing and Hunting	1~19	11	1
2	Mining	20~30	21	10~15
3	Construction	34~40	23	
4	Food, Beverage and Tobacco Product Manufacturing	41~74	311, 312	2~9
5	Petroleum and Coal Product Manufacturing	115~119	324	16~19
6	Chemicals, Plastics and Rubber Product Manufacturing	120~152	325, 326	20~24
7	Primary Metal Manufacturing	170~180	331	32
8	Fabricated Metal Manufacturing	181~202	332	33
9	Machinery Manufacturing	203~233	333	34
10	Computer, Electronic Product and Electrical Equipment Manufacturing	234~275	334, 335	35, 38
11	Transportation Equipment Manufacturing	276~294	336	36, 37
12	Other Durable & Non-Durable Manufacturing	75~114, 153~169, 295~304	313~316, 321~323, 327, 337	25~31, 39
13	Miscellaneous Manufacturing	305~318	339	40, 41, 43
14	Transportation, Communication and Utilities	31~33, 332~353	22, 48, 49, 51	--
15	Wholesale Trade	319	42	--
16	Retail Trade	320~331	44, 45	--
17	FIRE (Finance, Insurance and Real Estate)	354~366	52, 53	--
18	Services	367~440	54~56, 61, 62, 71, 72, 81, 92	--
19	Household	--	--	--
20	Government	--	--	--

TABLE 2 Estimated Parameters for Nested Logit Models of Origin and Mode Choice

Sector	Origin Choice Parameters			Mode Choice Parameters			
	γ^m	λ^m	Rho-Square	$\beta^m_{0,tuck}$	β^m_1	β^m_2	Rho-Square
1	0.0496	0.448	0.403	5.640	-4.010	-4.040	0.999
2	0.414	-3.830	0.262	1.850	0.857	0.0761	0.109
4	0.858	-1.430	0.242	5.600	1.810	0.464	0.772
5	0.103	1.010	0.493	1.670	-1.560	-3.410	0.755
6	0.790	0.801	0.206	1.420	-1.010	-1.120	0.486
7	0.753	1.690	0.130	1.430	-0.823	-1.280	0.817
8	0.904	0.173	0.16	3.180	-0.478	-0.741	0.936
9	0.775	0.339	0.224	-3.610	-8.500	-6.980	0.934
10	1.000	0.097	0.288	-1.590	-6.000	-4.160	0.613
11	1.020	-0.840	0.130	-3.470	-6.090	-5.270	0.825
12	0.888	1.090	0.081	5.540	1.540	0.575	0.562
13	0.921	0.805	0.272	2.830	-1.900	-1.960	0.926

Note: The correlated nature of cost and time variables, and use of assumed (rather than actual) results, is presumably causing the negative coefficient estimates for several sectors. Such situations appear more common for high-weight, low-time-value goods, with long-distance transport relying on rail, rather than the faster mode of trucking.

TABLE 3 Cumulative Distribution of RUBMRIO Trade Flows and Corresponding FAF3 Flows

Cumulative Percentage of Trade Flows	Number of Domestic Flows		Number of Export Flows	
	RUBMRIO	FAF3	RUBMRIO	FAF3
0 ~ 10%	10723	8409	12074	10841
10% ~ 20%	1899	2382	350	870
20% ~ 30%	895	1350	157	408
30% ~ 40%	456	836	74	242
40% ~ 50%	232	556	34	149
50% ~ 60%	105	354	16	94
60% ~ 70%	46	238	8	57
70% ~ 80%	25	153	3	34
80% ~ 90%	13	88	2	16
90% ~ 100%	6	34	2	9

Note: The count values indicate how many of the ordered (from lowest to highest) trade flows (aggregated across industry types) are needed to hit the first ten percent of the associated total (domestic or export) trade flow value, and then how many are needed to fill the next decile, reaching the 20 percent mark, and so forth. 14,400 flows exist for domestic shipments (120 x 120 zones), and 12,720 for export (120 x 106 zones).

TABLE 4 Ten Counties with Largest Falls and Rise in Total Production when IH 40 Travel Times Rise and Fall by 10%

County Name	Total Production (million dollars)		Percentage Change	County Name	Total Production (million dollars)		Percentage Change
	Under Original IH40 Travel Times	After IH40 Travel Times Rise by 10%			Under Original IH40 Travel Times	After IH40 Travel Times Fall by 10%	
Waller, TX	$13,798M	$11,847M	-14.14%	Clinton, OH	$1,762M	$1,854M	5.23%
Island, WA	12,996	11,560	-11.05%	Linn, MO	512	533	3.97%
Klickitat, WA	5,153	4,695	-8.88%	Harding, NM	162	168	3.56%
Somervell, TX	399	374	-6.34%	Roger Mills, OK	254	263	3.53%
San Luis Obispo, CA	31,157	29,369	-5.74%	Carson, TX	469	485	3.40%
Bourbon, KY	1,166	1,108	-4.97%	Collingsworth, TX	270	279	3.40%
Nicholas, KY	639	610	-4.48%	Wheeler, TX	353	365	3.38%
Harrison, KY	1,087	1,043	-4.07%	Morton, KS	308	319	3.36%
Robertson, KY	365	351	-3.86%	Donley, TX	333	344	3.35%
Calaveras, CA	11,148	10,738	-3.68%	Texas, OK	898	928	3.32%

TABLE 5 Ten Counties with Largest Rise and Falls in Total Production when Operational Costs of Trucking Fall and Rise by 20%

County Name	Total Production (million dollars)		Percentage Change	County Name	Total Production (million dollars)		Percentage Change
	Under Original Transport Costs	After Transport Costs Fall by 20%			Under Original Transport Costs	After Transport Costs Rise by 20%	
San Juan, NM	$7,819M	$12,322M	57.57%	San Juan, NM	$7,820M	$5,774M	-26.17%
La Plata, CO	3,838	5,748	49.74%	La Plata, CO	3,838	2,905	-24.33%
Montezuma, CO	2,461	3,602	46.32%	Aroostook, ME	16,872	13,068	-22.54%
Kane, UT	1,014	1,379	35.97%	Montezuma, CO	2,461	1,930	-21.61%
Dolores, CO	440	579	31.66%	Curry, OR	9,648	8,109	-15.95%
Hinsdale, CO	258	329	27.58%	Dolores, CO	440	374	-14.91%
San Juan, CO	223	279	25.00%	Hinsdale, CO	258	220	-14.42%
Island, WA	12,996	16,091	23.81%	San Juan, CO	223	193	-13.56%
Worth, MO	173	203	16.93%	Whatcom, WA	22,803	19,918	-12.65%
Mercer, MO	226	264	16.87%	Guthrie, IA	406	365	-9.98%

FIGURE 1 RUBMRIO Structure and Solution Algorithm.

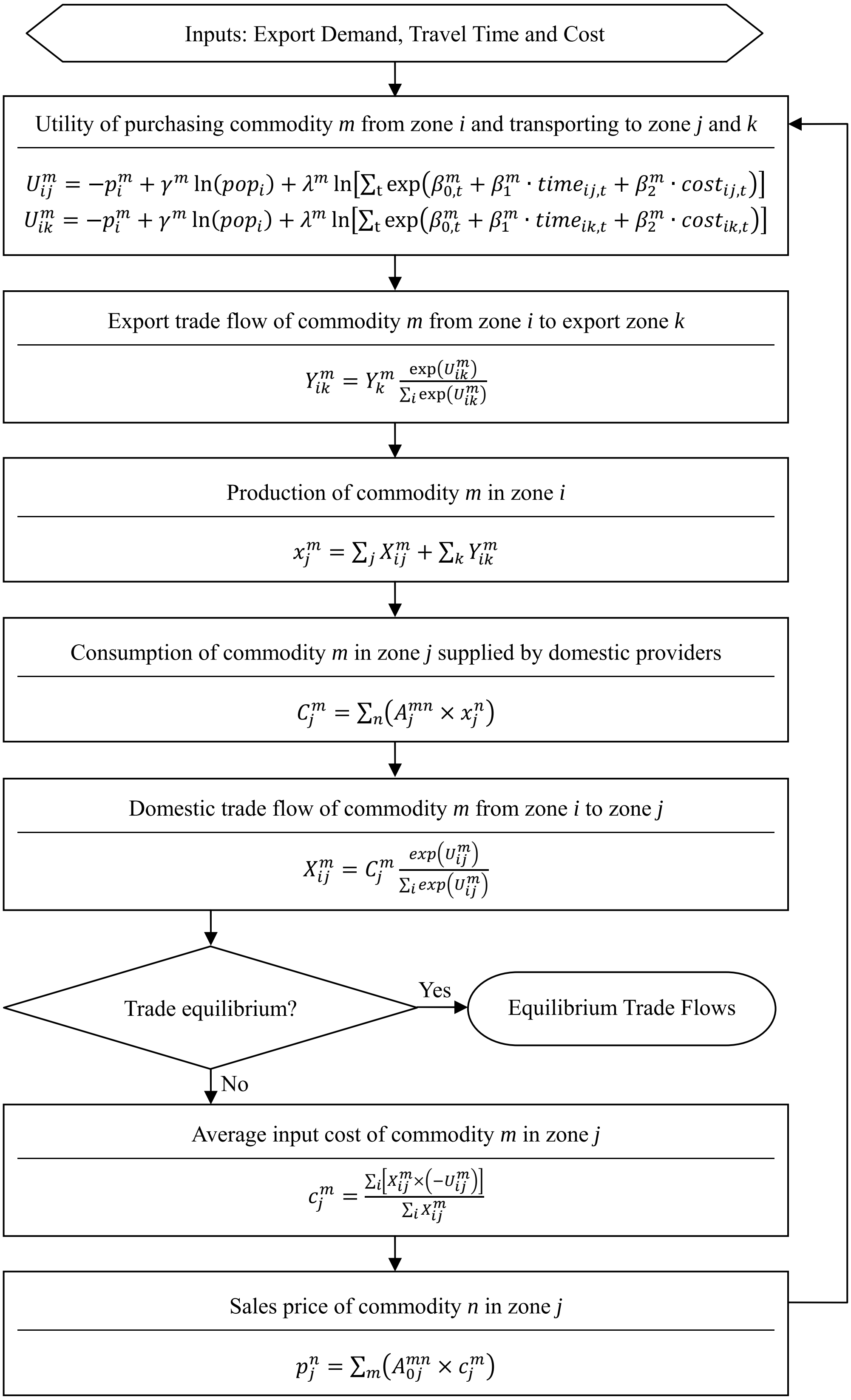
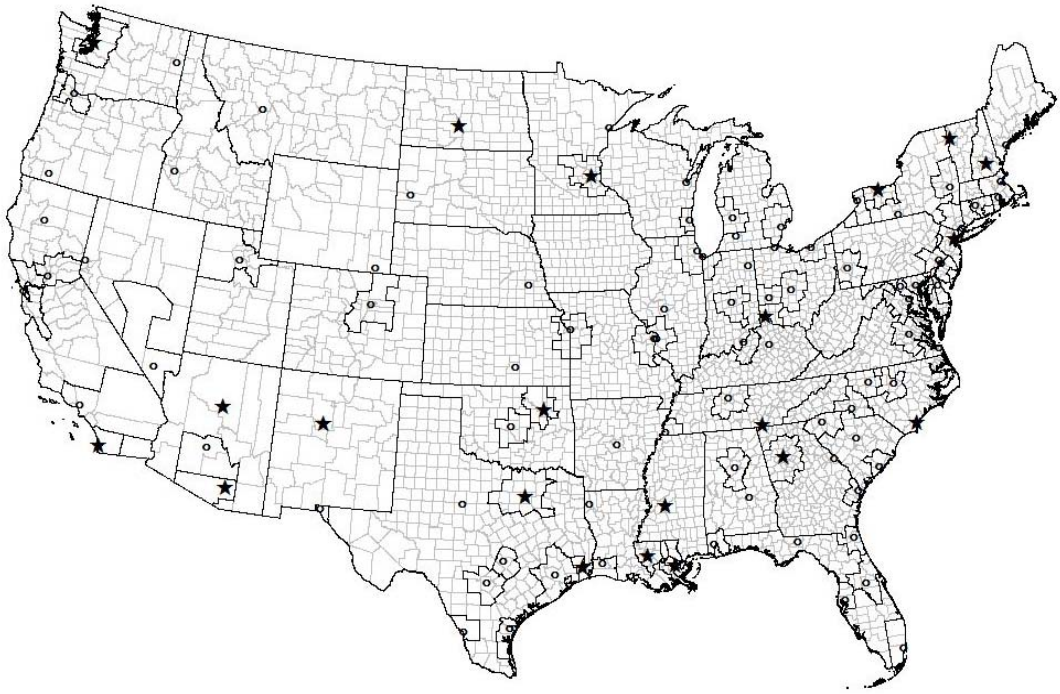


FIGURE 2 Centroids of Export Zones and Continental FAF3 Zones.



★ Annual exports above $100 Million

○ Annual exports below $100 Million

FIGURE 3 Multipliers for Trade Flow and Labor Expenditure Changes, due to Changes in Export Demands, by Commodity Type.

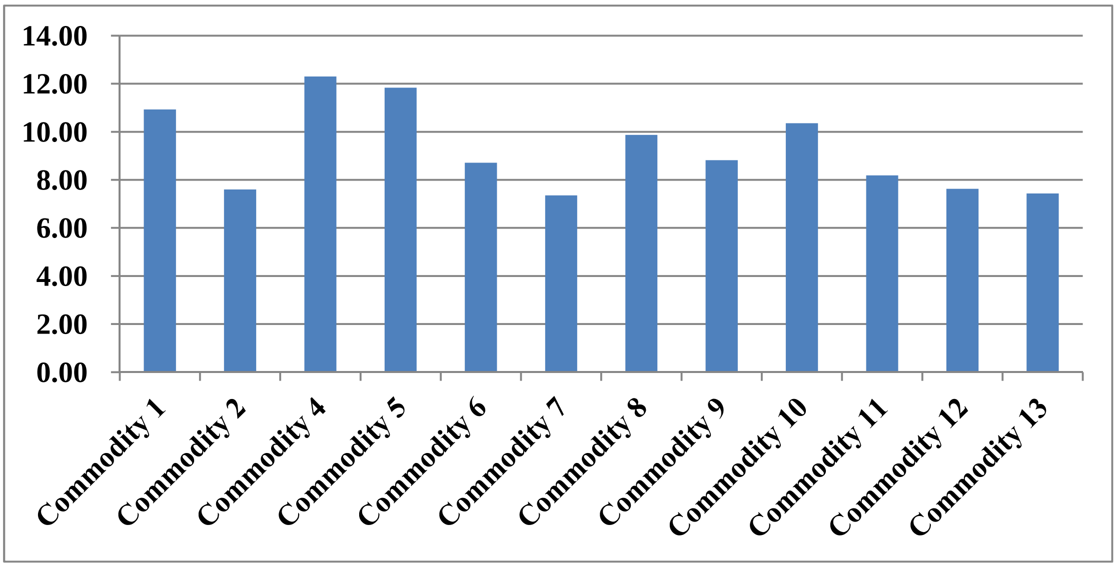


Figure 3 (a) Multipliers for total U.S. transactions by export commodity type.

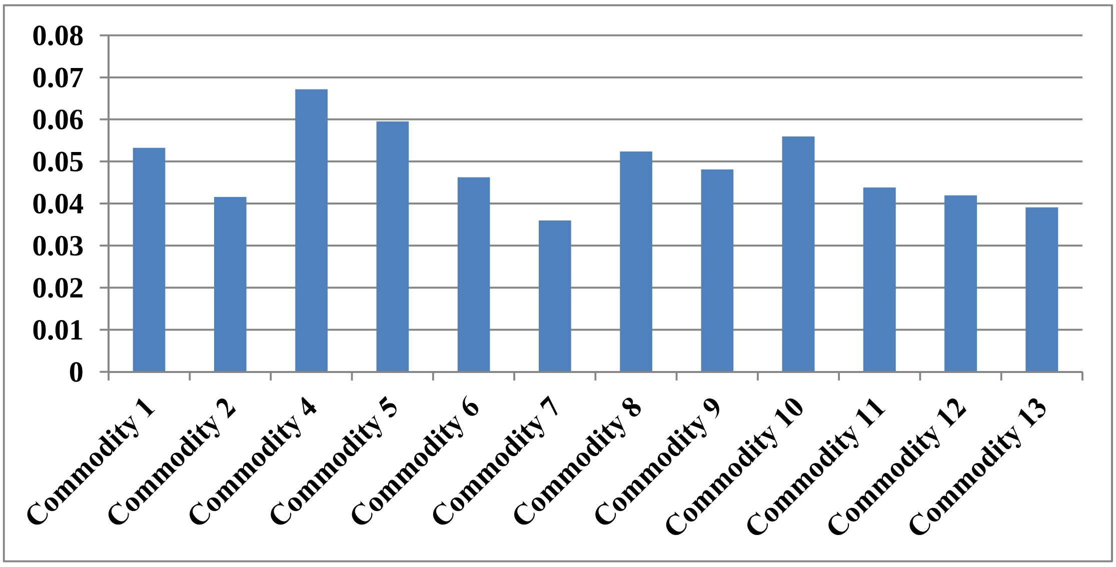
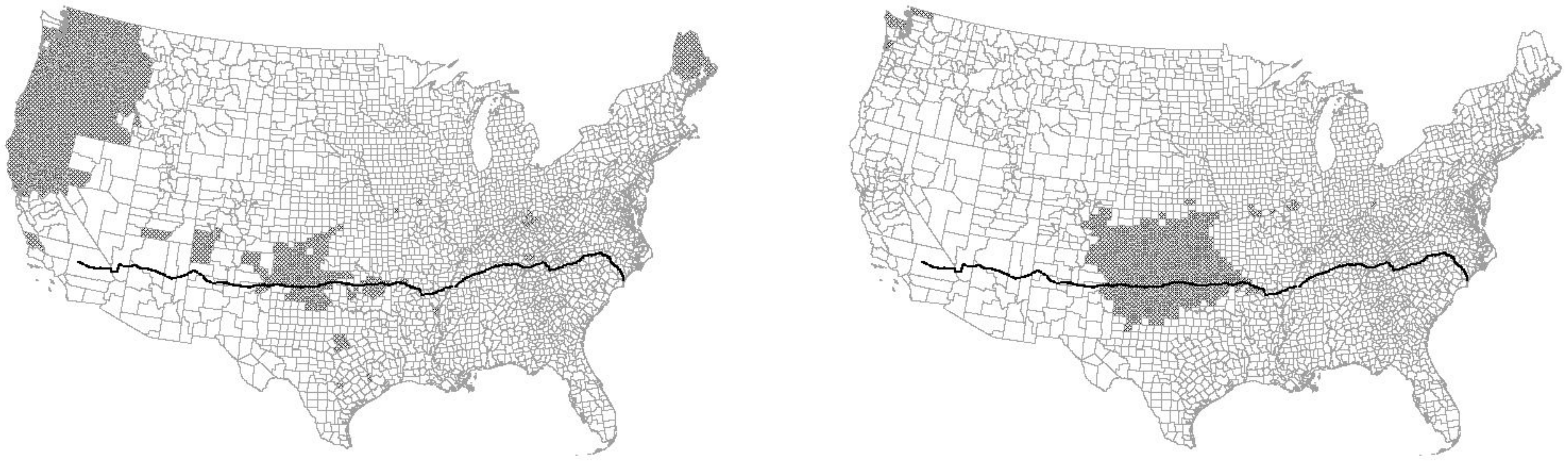


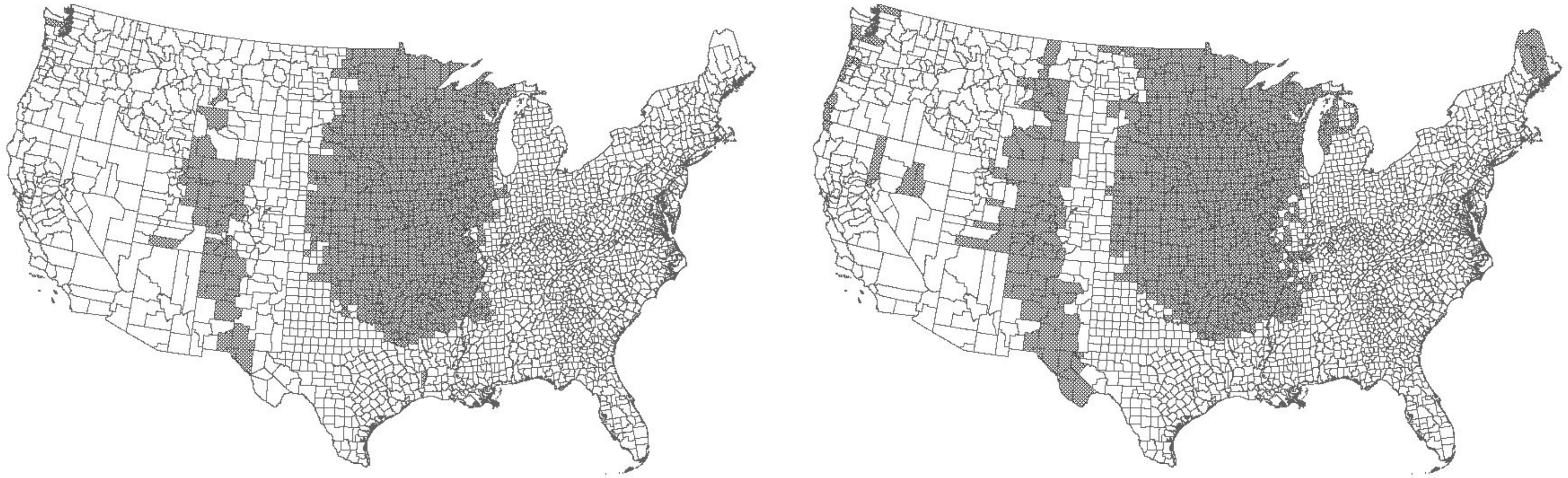
Figure 3 (b) Multipliers for total U.S. labor expenditures by export commodity type.

FIGURE 4 Counties with Largest Reductions and Increases in Total Production when IH 40 Travel Times Rise and Fall by 10%



Left: Counties with a 1% or Higher Reduction in Total Production when Travel Times Rise
Right: Counties with a 2% or Higher Rise in Total Production when Travel Times Fall

FIGURE 5 Counties with Largest Increases and Reductions in Total Production when Operational Costs of Trucking Fall and Rise by 20%



Left: Counties with a 10% or Higher Rise in Total Production when Trucking Costs Fall

Right: Counties with a 5% or Higher Reduction in Total Production when Trucking Costs Rise