DOCUMENTATION FOR APPLICATION OF KOCKELMAN et al.'s RANDOM-UTILITY-BASED MULTI-REGIONAL INPUT-OUTPUT (RUBMRIO) MODEL

Last Update: October 2008

This document describes the RUBMRIO model and use of code. The code, written in both Visual Basic (an earlier version of the code) and C (most recent code version), is designed to allow practitioners, researchers, and others to make use of the RUBMRIO spatial input-output model of trade.

Document Contents:

1. Model Description

RUBMRIO is a transportation-economic model that simulates the flow of goods, labor, and vehicles across a multiregional area (e.g., across the 254 counties in the State of Texas, as well as domestic export shipments to 49 other U.S. states and foreign export shipments to 18 ports [both peripheral and internal to Texas]). RUBMRIO simulates trade across zones of a region, as motivated by foreign and domestic export demands, and computes this trade across numerous economic sectors (such as those in Table 1). Input-output relationships/tables are used to anticipate consumption needs of commodity producers, and multinomial logit models distribute commodity flows (across origin zones and shipment modes).

As the series of RUBMRIO papers (noted below) attests, the model has seen several stages of development. The first stage sought trade equilibrium across zones, as driven exclusively by foreign export demand. The second stage extended the model to recognize domestic export demands, convert money flows into vehicle flows and allow for transportation congestive feedback on the highway network (relying, originally, on TransCAD traffic assignment). Land use constraints and wages were reflected in extension papers (see Figure 2, below, and Ruiz-Juri and Kockelman 2004 & 2006), but these specifications require more data inputs and so are not incorporated in the web-provided code (but can be made available upon request). The most recent stage of the model's development allows for dynamic disequilibrium, where near-term predictions strongly reflect current population distribution (and associated household demands). The system evolves toward a long-run equilibrium, and will reach that if no population or other shocks occur along the way (which is unlikely in practice). The dynamic evolution in trade patterns reflects additions to labor supply, and congestive feedbacks can be accomplished directly, without relying on other, traffic assignment software. In addition, the latest code (shown here in C++) loads domestic import flows onto the transportation network, contributing to congestion.

Standard (BPR-type) link-performance functions are used with assumption of shortest path route choices (i.e., a deterministic user equilibria, or DUE), but other assumptions can be coded into the routine. The Frank-Wolfe algorithm is used in highway flow assignments, and the method of successive averages for the full-model iterative process (between the network and trade flow models).

For more information on the latest version and earlier installments of the RUBMRIO model specifications, please see the following pre-prints of published and presented papers (all bundled in the zipped Papers file, linked via the RUBMRIO website: http://www.ce.utexas.edu/prof/kockelman/RUBMRIO_Website/homepage.htm), as well as flowcharts and figures provided (as examples) at the end of this document.

- Evaluation of the Trans-Texas Corridor Proposal: Application and Enhancements of the Random-Utility-Based Multiregional Input-Output Model

  Journal of Transportation Engineering 132 (7): 531-539 (2006). With Natalia Ruiz-Juri.

- Tracking Land Use, Transport, and Industrial Production using Random-Utility Based Multiregional Input-Output Models: Applications for Texas Trade

  Journal of Transport Geography 13: 275-286 (2005). With Ling Jin, Yong Zhao, and Natalia Ruiz-Juri.

- The Random-Utility-Based Multiregional Input-Output Model: Solution Existence and Uniqueness

  Transportation Research 38B (9): 789-807 (2004). With Yong Zhao.

- Extending the Random-Utility-Based Multiregional Input-Output Model: Incorporating Land Use Constraints, Domestic Demand & Network Congestion in a Model of Texas Trade

  Proc'gs of the 83rd Annual Meeting of the Transportation Research Board (2004). With Natalia Ruiz-Juri.

- The Introduction of Dynamic Features in a Random-Utility-Based Multiregional Input-Output Model of Trade, Production, and Location Choice

Proc'gs of the 86th Annual Meeting of the Transportation Research Board (2007), and under consideration for publication in the *Journal of the Transportation Research Forum*. With Tian Huang.

Table 1. Description of Economic Sectors in Texas Applications of the RUBMRIO Model

	Sector Description	**IMPLAN**	**NAICS**	**SCTG**
1	Agriculture, Forestry, Fishing and Hunting	1~18	11	1,3,4,25
2	Mining	19~29	21	10~18
3	Construction	33~45	23	--
4	Food Manufacturing	46~84	311	2, 5~9
5	Chemicals Manufacturing	147~171	325	19~24
6	Primary Metals Manufacturing	203~223	331	32
7	Fabricated Metals Manufacturing	224~256	332	33
8	Machinery Manufacturing	257~301	333	34
9	Electronic and Electric Equipment	302~343	334,335	35,38
10	Transportation Equipment	344~361	336	36, 37
11	Other Durable & Non-Durable Manufacturing	85~111, 112~146, 362~373, 374~389	312~316, 339, 321~324, 337	26~31, 39~43
12	Transportation, Communications & Utilities	391~397, 398~400, 413~424, 30~32	48, 49, 51, 22	--
13	Wholesale trade	390	42	--
14	Retail trade	401~408, 409~412	44, 45	--
15	FIRE (Finance, Insurance & Real Estate)	425~436	52, 53	--
16	Services	437~509	54~56, 61~62, 71~72, 81, 92	--
17	Households			
18	Government			

Note: This table provides the corresponding sector code in different data sources which were used in Texas studies. IMPLAN stands for IMpact analysis for PLANning, NAICS stands for North America Industry Classification System, and SCTG stands for Standard Classification of Transported Goods.

2. Explanation of Inputs and Parameters

The following are descriptions of the various inputs and parameters the user must provide the RUBMRIO code.

Number of Zones (I and J): The total number of zones that comprise the region of interest. This is both the number of origin and destination zones, which are indexed by *i* and *j*, respectively.

Number of Export Zones (K): The total number of points of foreign export that the region trades with. Export zones are indexed by *k*.

Number of States (S): The total number of points of domestic export that the region trades with. States are indexed by *s*.

Number of Sectors (M): The total number of sectors that comprise the economy of the region of interest. Households and government are both treated as industrial sectors that buy from and sell to other sectors. The number of sectors is both the number of production and consumption sectors, which are indexed by *m* and *n*, respectively.

Sector Number/Index for Households: The household sector is treated specially: households are assumed to only trade via highway, and a separate logit model is used to compute the utility of trading between two zones for the household sector. Moreover, the trips generated by the household sectors' purchases (of retail goods, for example) are considered to be personal shopping trips, and those generated by industry sectors that purchase labor from the household sector are treated as work trips. The sector number (*m*) for households must be specified to account for its unique nature and distinctive behavioral models.

Sector Number/Index for Mining: The mining sector is also treated specially. Much of the mining sector's activity is the trade of crude petroleum and natural gas which is conducted by pipeline, so a special set of parameters is introduced to reduce this sector's trade by the percent of trade not done via highway or railway. The sector number (m) of mining must be specified to account for this industry's unique activity. If no mining sector is being modeled, specify this as 0.

Number of Nodes: The number of nodes in the highway network, including all zones, export zones, and states as nodes. *NOTE:* All nodes must be indexed. This indexing must include all O/D zones, then all export zones, followed by all states, and finally all other nodes (i.e. highway intersections). O/D zones, export zones, and states should be indexed as nodes in the order of their indexing as a zone. Proper indexing ensures that trips will be loaded onto the network correctly.

Number of Links: The number of links in the highway network.

Lambda (λ^m): Lambda is a parameter called by the nested logit (NL) sub-routine, which is used to determine the utility and flow of trades between all zone pairs (based on the utility of making purchases from each possible origin zone). RUBMRIO reads lambda as a vector, of 1 row and M columns.

Beta (β^m): This parameter also is utilized by the nested logit sub-model trade/purchase utilities. RUBMRIO reads this Beta term as a matrix of 3 rows and M columns. The three rows are for the β_0, $\beta_{Highway}$, and $\beta_{Railway}$ parameters corresponding to each of the M industry sectors. The β_0 term is the alternative specific constant for the highway mode, and the $\beta_{Highway}$ & $\beta_{Railway}$ terms are the distance or travel cost decay parameters for each of those modes.

HHBeta (β_{HH}^m): This parameter is called by the multinomial logit sub-model, to determine the utility of different industries' purchases of labor (across all potential origin zones, where households may reside). RUBMRIO reads HHBeta as a 1 x M row vector.

Delta (δ): This distance decay parameter reflects households' diminishing willingness to travel far for shopping trips, as distance (or travel cost) increases.

Technical Coefficients without Leakage (A_{0j}^{mn}): This matrix reflects the productive technology of all zones, for input needs across all industries. RUBMRIO reads the technical coefficients as a series of J matrices (i.e. corresponding to J zones), each with M rows by N columns (i.e. square matrices). The matrices should be tab-delimited, and no lines should be skipped between them.

Technical Coefficients with Leakages (a_j^{mn}): The matrix reflects the productive technology of all zones, with Regional Purchase Coefficients (RPCs) to account for "leakages" (i.e. purchases of goods outside the multizonal region [e.g., inputs provided by industries in other states or nations]). RUBMRIO-DUE reads the technical coefficients as a series of J matrices, each with M rows by N columns (i.e. square matrices). The matrices should be tab-delimited, and no lines should be skipped between them.

Truck Conversion Factors (TCF^m): These factors convert annual dollars (in year 2000 dollars) to *full* trucks for each industry (by first converting dollars to annual tons, then annual tons to annual trucks, and annual trucks to daily trucks). RUBMRIO reads truck conversion factors as a vector, with 1 row and M columns.

Empty Truck Percentage: This is the percentage of empty truck trips (e.g., those deadheading, and carrying no goods). It allows the model to reflect the underfilling or empty nature of many truck trips.

Passenger car Equivalency Factors (PCE^m): These convert trucks to an equivalent number of passenger cars for each industry. RUBMRIO reads this as an Mx1 vector. (The standard PCE is 2.0.)

Work Trip (vehicle) Occupancy: This is the number of workers per car for each sector. RUBMRIO reads work trip occupancy as an Mx1 vector. Work trip occupancy is often assumed to be 1.2 (for all sectors).

Truck Volume Factor: This is the percentage of daily truck trips that occur during the hour of interest. (This is assumed to be 0.08 in the applied papers, noted above.)

Household Trip Conversion Factor: The number of passenger-car trips per dollar expended by the household on shopping. (This is referred to as TPD in the papers.) The default value is 6.23E-05, or 1 trip for every $44 expended in a year (by a household), divided by 365 days of the year (allowing one to perform a day's traffic analysis).

Shop Volume Factor: This is the percentage of daily shopping trips that are expected to occur during the peak hour of interest. (Assumed to be 0.05 in the papers.)

Work Volume Factor: The percentage of daily work trips expected to occur during the hour of interest. (This is 0.20 in the papers.)

Labor Factors (wpd_j^m): These define the number of workers needed per dollar of output in each industry m (in each zone j). RUBMRIO-DUE reads this input as a matrix of J rows and M columns. (This is the "Workers per Dollar" parameter described in the papers.)

Gas Factors: These factors reduce highway & railway trade of the mining sector, since such trade is dominated by shipments of crude oil and natural gas, which are largely sent via pipelines. There is a factor for each zone and these factors must be loaded as a Jx1 vector. (These factors average 0.557 [a 56% reduction] in the Texas implementation of the model.)

Highway Distances between Zones ($d_{ij,\ highway}$): These must be loaded as a matrix of I rows and J columns.

Railway Distances between Zones ($d_{ij,\ railway}$): These must be loaded as a matrix of I rows and J columns.

Highway Distances to Export Zones ($d_{jk,\ highway}$): These must be loaded as a matrix of J rows and K columns.

Railway Distances to Export Zones ($d_{jk,\ railway}$): These must be loaded as a matrix of J rows and K columns.

Highway Distances to States ($d_{js,\ highway}$): These must be loaded as a matrix of J rows and S columns. (Note: The term "States" is used here, but any other zone that motivates/drives intra-area production and trade can be used. In this code, states essentially function as export zones, and are therefore treated as such. The only real difference is that they tend to be located further from the zones of interest than the foreign export zones.)

Railway Distances to States ($d_{js,\ railway}$): These must be loaded as a matrix of J rows and S columns.

Foreign Exports: These are the dollar value of output demanded by each foreign export zone for each sector. These inputs must be loaded as an M x K matrix.

Domestic exports: These are the dollar value of output demanded by each state (or other external, domestic zone), by sector. These inputs must be loaded as an M x S matrix.

Trade Tolerance: The allowable difference in total trade between iterations, stated as a percentage (e.g., 1%).

Maximum Trade Iterations: The maximum number of iterations of the trade loop to be run. This may be used to limit the run time, and convergence may not have occurred by the time this limit is reached. 100 maximum iterations may serve as a useful default for this value (assuming a 1 or 2% tolerance setting).

The above are the common inputs for both VB and C codes. The following parts are the additional inputs for VB and C codes, respectively. The inputs for VB code is described below, and then follows the inputs for C code.

Local final demand: These are the dollar values of intrazonal demand for each zone's output for each sector. While these are static in this VB version of RUBMRIO, Tian and Kockelman (2007) have adapted the model to incorporate dynamic local final demands. These inputs must be loaded as a J x M matrix.

Local congestion: The cost of traveling through a zone due to congestion from intrazonal trips (in units of time consistent with free flow travel times). Set as the average of n (can be defined, usually takes the 3 nearest (generally adjacent) zones' travel time to account for intrazonal congestion. These can be computed within the program once set.

Hourly link capacity: The hourly capacity (vehicles/hour) for each link in the highway network. These are required as part of a link input file.

Link free flow travel time: The uncongested travel time for each link in the highway network. Free flow travel times may be entered in any consistent unit of time. These are required as part of a link input file.

Link distance: The length of each link in the highway network. These are required as part of a link input file.

DUE Tolerance: The average allowable difference in link vehicle flows (between iterations), stated as a percentage (e.g., 1%).

Maximum DUE Iterations: The maximum number of iterations of the DUE loop to be run. This may be used to limit the run time, and convergence may not have occurred by the time this limit is reached.

Frank-Wolfe Algorithm Tolerance: The sensitivity desired when minimizing the objective function via the Bisection Method when performing the Frank-Wolfe algorithm. A tolerance of 1e-05 or less may be necessary to see significant congestive feedbacks.

The following inputs are for the ***Dynamic*** RUBMRIO, as coded in C.

Domestic Imports: These are the dollar value of input demanded by Texas from each state by sector. These inputs must be loaded as an M x K matrix.

Household Demands: These are the dollar value of commodities demanded by households from each county by sector. These inputs must be loaded as an M x J matrix.

Evolve Rate: This is the rate that labor supply shifts toward where labor demanded. It can be set as 0.05 per year, or other numbers according to expected local situations.

County Population: These are the population by county and can be loaded as a J x 1 vector.

Note: The network settings are in initial.txt. To change the DUE convergence criteria or other settings, please check those values in file: initial.txt.

3. **Using RUBMRIO**

A. Using the VB Code

a. Defining the Multizonal Region and Economy Size.

Users should enter the required information in the "Defining Region and Economy Size" box, according to the above definitions of parameters and inputs.

b. Defining Highway Network

Users should enter the required numbers in the "Define Network" box. A link input file must be loaded. Clicking on the "Load link file" button will bring up a Microsoft Common Open File Dialog box, in which the file containing link specifications can be found. The link input file should contain a line for each link in the network. For each link, the following must be provided in tab-delimited format in order: link tail, link head, hourly link capacity, link free flow travel time, and link distance. The link tail and link head should be the nodes according to the indexing described in section 2.

c. Loading Inputs

There are two ways to load the inputs:

i. Loading inputs separately: Click on the boxes under the "Load Separately" label to bring up the windows for loading parameters. These windows allow the user to load parameters, distances, and export demands. Clicking on the buttons in these windows will bring up a Microsoft Common Open File Dialog box, in which the file containing the parameters can be found. These files must be .txt files, and contain the parameters as specified in the Explanation of Inputs and Parameters. All matrices must be tab-delimited.

Within the "Load Parameters" window, there is a button for "Load Other Vehicle Parameters." This button requires a special file formatted as follows:

Empty Truck Percentage
Truck Volume Factor
HH Truck Conversion Factor
Shop Volume Factor
Work Volume Factor
Delta (Distance Decay Factor)

ii. Loading inputs as a single file: Instead of loading all the inputs separately, a single file containing all the inputs may be used. It must be formatted as follows:

Lambda
(skipped line)
Beta
(skipped line)
HHBeta
(skipped line)
Truck Volume Factor
(skipped line)
HH Truck Conversion Factor
(skipped line)
Shop Volume Factor
(skipped line)
Work Volume Factor
(skipped line)
Delta
(skipped line)
Ttuck Conversion Factor
(skipped line)
Truck Equivalency Factor
(skipped line)
Work Trip Occupancy
(skipped line)
Local Final Demand
(skipped line)
Foreign Dollars
(skipped line)
Domestic Dollars
(skipped line)
Highway Distances between zones
(skipped line)
Railway Distances between zones
(skipped line)
Highway Distances to export zones
(skipped line)
Railway Distances to export zones
(skipped line)
Highway Distances to states
(skipped line)
Railway Distances to states
(skipped line)
Labor Factors
(skipped line)
Gas Factors

(skipped line)
Technical Coefficients with RPC
(skipped line)
Technical Coefficients without RPC
(end of file)

NOTE: It is important that these inputs are internally consistent. If any of the parameters are missing or matrices contain the wrong number of elements, RUBMRIO-DUE WILL CRASH.

d. Setting model sensitivity

Enter the desired tolerance levels and maximum numbers of iterations. These will govern the run time & exit strategy of the program.

e. Running RUBMRIO-DUE.

Click the "Run to trade and network equilibrium" button.

RUBMRIO writes all output to 2 files:

- The main output file is written to C:\Output.txt. This output consists of dollar flows between all zones by each mode and for each consuming sector.
- A second output file contains a table of PCE flows between all zones, and is written to C:\VehicleOutput.txt.

B. Using the C Code

a. Loading inputs as a single file: The C code requires that to the modeler writes all the inputs into a single file. It must be formatted according to the codes indicated. Different versions of code have a slightly different order of input. The input function within the codes offers a detailed explanation of variables and inputs. The modeler must make sure that all inputs are in the order of the codes. Distinct inputs can be separated using skipped lines.
NOTE: It is important that these inputs are internally consistent. If any of the parameters are missing or matrices contain the wrong number of elements, RUBMRIO WILL CRASH.

b. Setting model sensitivity
The desired tolerance levels and maximum numbers of iterations should be changed in the input .txt file. These will govern the run time & exit strategy of the program.

c. Running RUBMRIO
Run the model exactly like any other C programs.

d. Output
RUBMRIO will write the output files to the directory you defined in the C codes at the main function.

There is one main output file containing estimates of vehicle flows, worker flows and dollar value of trade flows by county and by sector. Other output files are intermediate files internal to the modeling system, in order to recognize congestive feedbacks and track the dynamic disequilibrium in locations and flows (from short- to longer-term conditions). Currently, these intermediate files consist of vehicle flows and long-term-equilibrium labor demands, near-term population, and near-term labor demands (all by county).

4. **Estimating Parameters in RUBMRIO**

There are several parameters in RUBMRIO to be estimated. This section explains briefly how to estimate these parameters, including example results. The estimation approach relies on maximum likelihood estimation (MLE), so any kind of statistical software that accommodates the MLE procedure can be used. Here, we explain the case of LIMDEP's NLOGIT command, with Matlab codes for data management and editing.

a. Model Formulation & Estimation Theory

The basic structure of RUBMRIO is 2-level nested logit (NL) model of origin and mode choice. Then, analyst should have basic knowledge of discrete choice modeling, NL model specifications, and MLE. There are typically two methods to estimate NL model: Simultaneous estimation and sequential estimation. Although simultaneous estimation is statistically preferred, here we explain sequential estimation method, which is simpler than simultaneous one.

b. Sequential Estimation Procedure

At first, parameters in mode choice model are estimated and then those in origin choice model are estimated based on log-sum variable that is computed by the parameters in mode choice model. This procedure is explained in more detailed below in the case using LIMDEP and MatLab software.

1) Data Preparation for Mode Choice Estimation: For each sector, create an Excel file formatted for input of LIMDEP using Matlab code "ForCostEstimationSetup.m". You can make assumptions on speeds, operating cost and values of time for each mode here.

2) Mode Choice Estimation: For each sector, estimate the mode choice (binary choice between truck and railway) model with Limdep. You can see an example of Limdep code for this estimation in upper part of "OriginChoice.lim".

3) Data Preparation for Origin Choice Estimation: Write the output of parameter estimation results in mode choice model in first part of "try3.m". Then, create another input file for estimation of origin choice model using the MatLab code. Again this procedure is repeated for each sector.

4) Origin Choice Estimation: For each sector, estimate the origin choice (among 254 zones in Texas application) model with Limdep. You can see an example of Limdep code for this estimation in latter part of “OriginChoice.lim”.

c. Notes

As usual, you have to check that the estimated parameters have reasonable signs and statistically significant. Overall goodness of fit should also be looked by several measures such as rho^2. A special note is posted here that in RUBMRIO, NL model is not usual one that represent individual travelers’ decision making but represent the aggregate trade flows (in dollar unit). Thus, usual interpretation of parameter results or statistical test may not be appropriate.

d. Estimation Examples (September 21, 2008 by Takuya Maruyama)

Here we show some examples of parameter estimation. Basically, we assume the following disutility function (of acquiring some commodity *m* from origin zone *i* and consuming it in zone j – or in export zone *k*).

$$U_{ij}^{m} = -p_i^m + \lambda_p^m \cdot Population_i + \lambda_{IV}^m \ln[\exp(\beta_0^m + \beta_{highway}^m \cdot d_{ij,highway}) + \exp(\beta_{railway}^m \cdot d_{ij,railway})]$$

where *p* is the price of purchasing $1 of commodity *m* in zone *i* (in units of utility), and λ^m and β are parameters to be estimated via a nested logit model (where subscript *IV* refers to inclusive value or *logsum* variable).

Table 2 provides example estimates when controlling for origin-zone *Population* (as in the above equation). The negative coefficient on population for Sector 2 is unexpected, but not disallowed.

Table 2. Origin and Mode Choice Parameters Using Distance & Population (as estimated by Tian Huang, July 2006)

Sector	Mode choice estimation				Origin choice			R^2	
	$\beta_{0,truck}$	β_{truck}	β_{rail}	N_{obs}	λ_p	λ_{IV}	N_{obs}	R_β^2	R_λ^2
1	2.85266	0.000431	0.000834	647	5.20E-08	-0.05202	49	0.55118	0.02214
2	1.42867	-0.0078	-0.00346	318	-1.85E-07	0.17106	49	0.15088	0.06094
4	3.73199	-0.00588	-0.00509	647	5.52E-08	0.056784	49	0.80020	0.02595
5	129.739	-0.36922	-0.34735	647	5.16E-08	0.001216	29	0.53474	0.02240
6	5.63033	-0.00251	-0.00126	555	5.91E-08	0.099023	49	0.93884	0.03119
7	5.56598	-0.00942	-0.00827	647	3.95E-08	0.041145	48	0.94483	0.01455
8	1.91121	-0.01584	-0.0735	514	9.29E-08	0.007972	47	0.91679	0.10563
9	2.24383	-0.00125	-0.00161	643	5.17E-08	0.440367	49	0.66955	0.03171
10	4.62545	-0.0034	-0.00239	104	1.10E-07	0.196367	35	0.86939	0.12909
11	2.89482	0.000947	0.000742	647	4.86E-08	-0.10195	49	0.73037	0.01821

Table 3 shows the original RUMBRIO model with only distance variable.

Table 3. Estimated parameters using CFS 1997 data (as appears in Kockelman et al.'s 2005 *J. Transport Geography* article)

	Mode choice estimation				Origin choice estimation	
	$\beta_{0,truck}$	β_{truck}	β_{rail}	$N_{obs.}$ & R^2	λ	$N_{obs.}$ & R^2
1	4.499	5.79E-10	-1.38E-04	830 & 0.8170	16212	49 & 0.0007
2	1.898	-1.07E-10	7.82E-10	288 & 0.6225	112.0	51 & 0.0004
4	3.856	-7.78E-04	8.91E-10	865 & 0.7360	6.926	49 & 0.4499
5	3.463	-6.97E-04	5.11E-10	1275 & 0.6711	5.507	49 & 0.3039
6	4.742	-1.16E-03	-1.38E-08	830 & 0.8170	3.096	48 & 0.2892
7	9.993	-2.15E-03	-1.16E-08	882 & 0.9898	1.723	49 & 0.3086
8	35.208	-2.79E-06	2.72E-12	919 & 1.0000	67293	48 & 0.0005
9	7.976	-1.77E-03	-9.35E-04	766 & 0.9877	155.7	49 & 0.0009
10	4.213	9.00E-03	1.14E-09	680 & 0.9877	3.976	48 & 0.2859
11	4.498	-1.12E-03	6.19E-10	1626 & 0.7487	3.334	49 & 0.3025

Results of Tables 2 and 3 are based simply on distance, so behavior is inelastic to transport costs. When re-estimating the model using inter-county time and cost estimates (both as a function of distance [with higher speeds assumed for longer trips], and mode), various unreasonable parameter estimates emerged, largely due to multicollinearity in these two variables. To avoid this issue, we collapsed time and cost into a single, generalized cost term (using value of travel time assumptions) and estimated a single coefficient for their combination.

$$U_{ij}^m = -p_i^m + \lambda_p^m \cdot Ln(Population_i) + \lambda_{IV}^m \ln[\exp(\beta_0^m + \beta_1^m \cdot g_{ij,highway}) + \exp(\beta_1^m \cdot g_{ij,railway})]$$

where $g_{ij,highway}, g_{ij,railway}$ represent generalized costs in dollar ($) units, and $g_{ij} = \tau\, time_{ij} + cost_{ij}$, where $\tau = 20$ $/hour (value of time).

The following assumptions were made to obtain estimates of intercounty travel times and costs for freight movements, as a function of distances. These times and costs should be revised based on detailed link and network data, by mode, where available.

- If highway distance <200 mile, truck speeds = 35 mile/hour.
- If highway distance >=200 mi, but <300 mi, truck speeds = 45 mi/hour
- If highway distance >=300 mi, truck speeds = 60 mi/hr.
- Railway speed is 40 mi/hr.
- Fixed operation time of trucks (at trip terminals) is 3 hours.
- Fixed operation time of rail cars (at trip terminals) is 5 hours.
- Operating cost for trucks is $10/mile (per truck).
- Operating cost for railways is $2/mile (per rail car).
- Additional value of travel time is $12/hr for both highway and rail cars.
- 1 rail car holds the same amount of goods as 1 truck (essentially 1 TEU container).

Ideally, the travel times used to estimate model parameters, the travel times used to apply the RUBMRIO model, and base-year traffic assignment travel-time results should be highly consistent. However, the above assumptions are simplistic and better network data are needed for more accurate planning models. Table 4's parameter estimates are based on the above assumption, and several unexpected parameter signs results, as highlighted in bold.

Table 4. Estimated Results using Generalized Costs & Ln (Population)

Sec#	Mode choice estimation			Origin choice			R^2	
	$\beta_{0,truck}$	β_1	N_{obs}	$\beta_{\ln(pop)}$	λ_{IV}	N_{obs}	R^2_β	R^2_λ
1	3.03181	-5.9E-05	653	0.285054	1.89734	49	0.61173	0.02962
2	2.89455	-0.000787609	318	-0.540616	0.073739	49	0.40854	0.03765
4	4.0562	-8.97E-05	838	-0.329318	-0.673765	49	0.81773	0.02031
5	3.15481	-0.000100146	929	0.284035	-0.082455	48	0.63552	0.00871
6	5.62768	-0.000150991	555	0.741003	0.357956	49	0.93917	0.06816
7	5.7638	-9.55E-05	666	0.302066	0.788258	48	0.95038	0.02670
8	10.596	-0.000160096	514	0.746167	-0.110355	47	0.99821	0.06824
9	2.75588	3.55E-05	643	0.420219	-1.04174	49	0 .72922	0 .02477
10	5.91692	-0.000182212	104	1.18956	0.341088	35	0 .92495	0 .14267
11	3.84303	-4.50E-05	1211	0.111783	-0.071504	49	0 .80102	0 .00164

Interestingly, the $\beta_{0,truck}$ estimates and goodness of fit statistics (R^2) are similar to those estimated by Huang (2006), as shown in Table 2. Unexpectedly, sector 9 (electronics) has a positive β_1 estimate, suggesting that higher travel costs are somehow "preferred" when selecting an origin for such inputs. Several sectors also are estimated to have negative inclusive value coefficients (λ^m_{IV}), while one such estimate exceeds 1.0 (all in violation of strict random utility-maximizing/cost-minimizing theory). Questionable coefficients are highlighted in green. All parameter values are, of course, up to the modeler and the CFS data set obscures a variety of important attributes in individual buyers' mode and origin choice decisions (including actual region [rather than state] of origin, by sector and mode). Use of stated preference survey results and micro data at the level of individual firms should prove very useful in enhancing parameter estimates for application.

Questions & Suggestions: Please contact Dr. Kara Kockelman (kkockelm@mail.utexas.edu) for more information on the code, and to offer suggestions for enhancements.

We welcome all suggestions! And we look forward to trying to incorporate them.

5. **Flowcharts & Figures (Model Specifications & Example Application Results)**

FIGURE 1. RUBMRIO's Main Model Structure (Spatial Input-Output Specification for Flow Estimates)

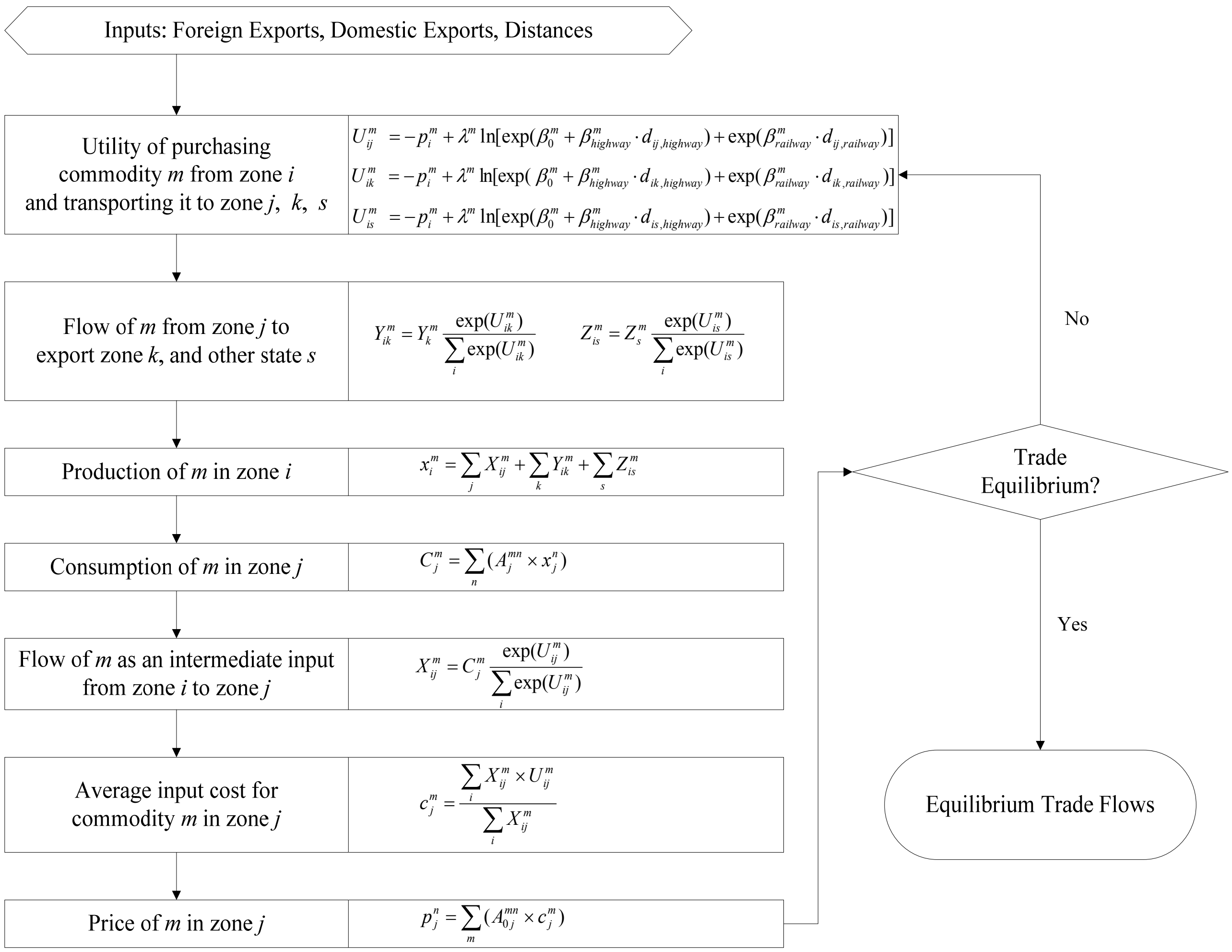

FIGURE 2. Dynamic RUBMRIO Model Structure (Huang & Kockelman 2007)

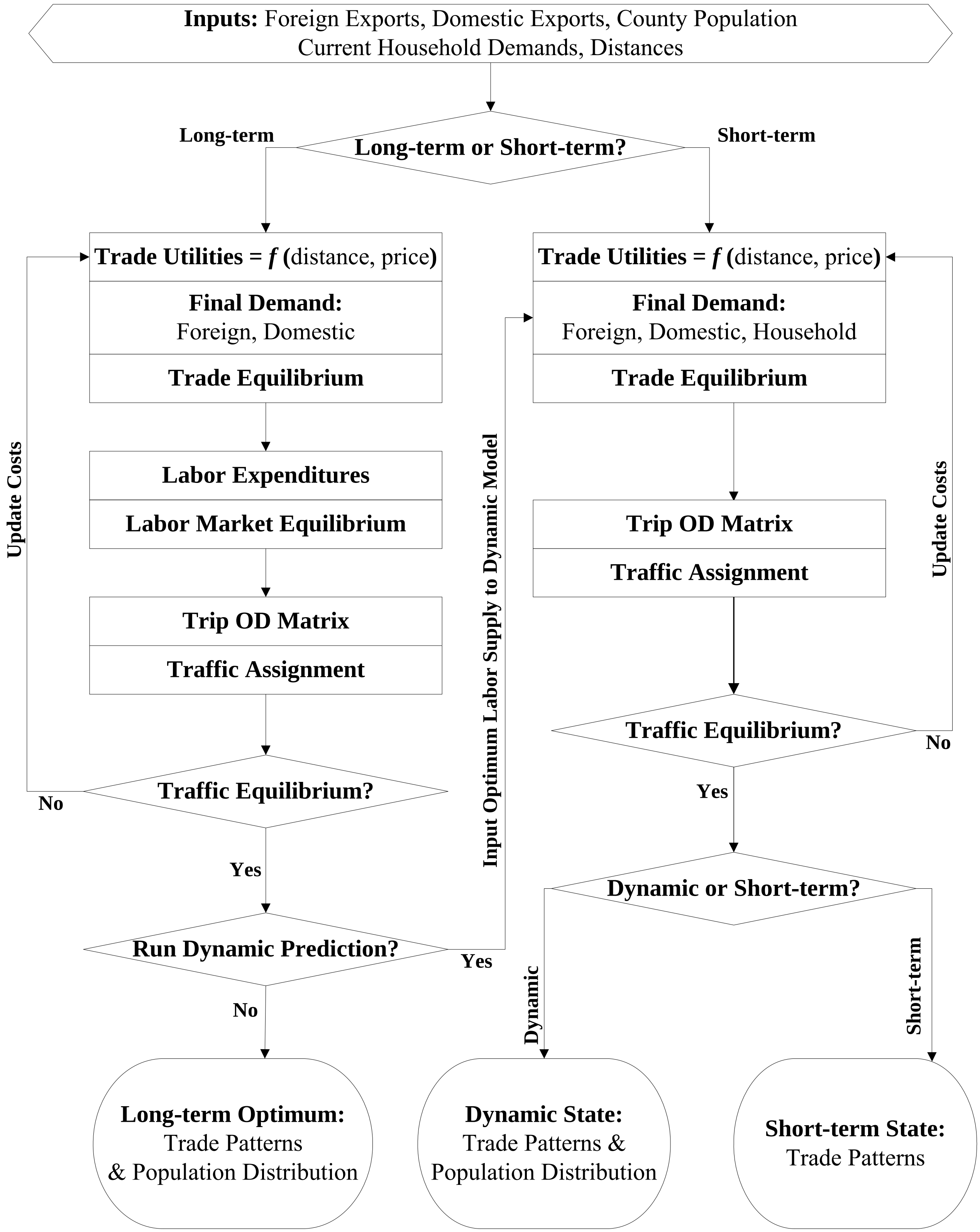

FIGURE 3. RUBMRIO Model Structure, with Land Rent & Labor Market Extensions (as described in Ruiz-Juri & Kockelman 2006)

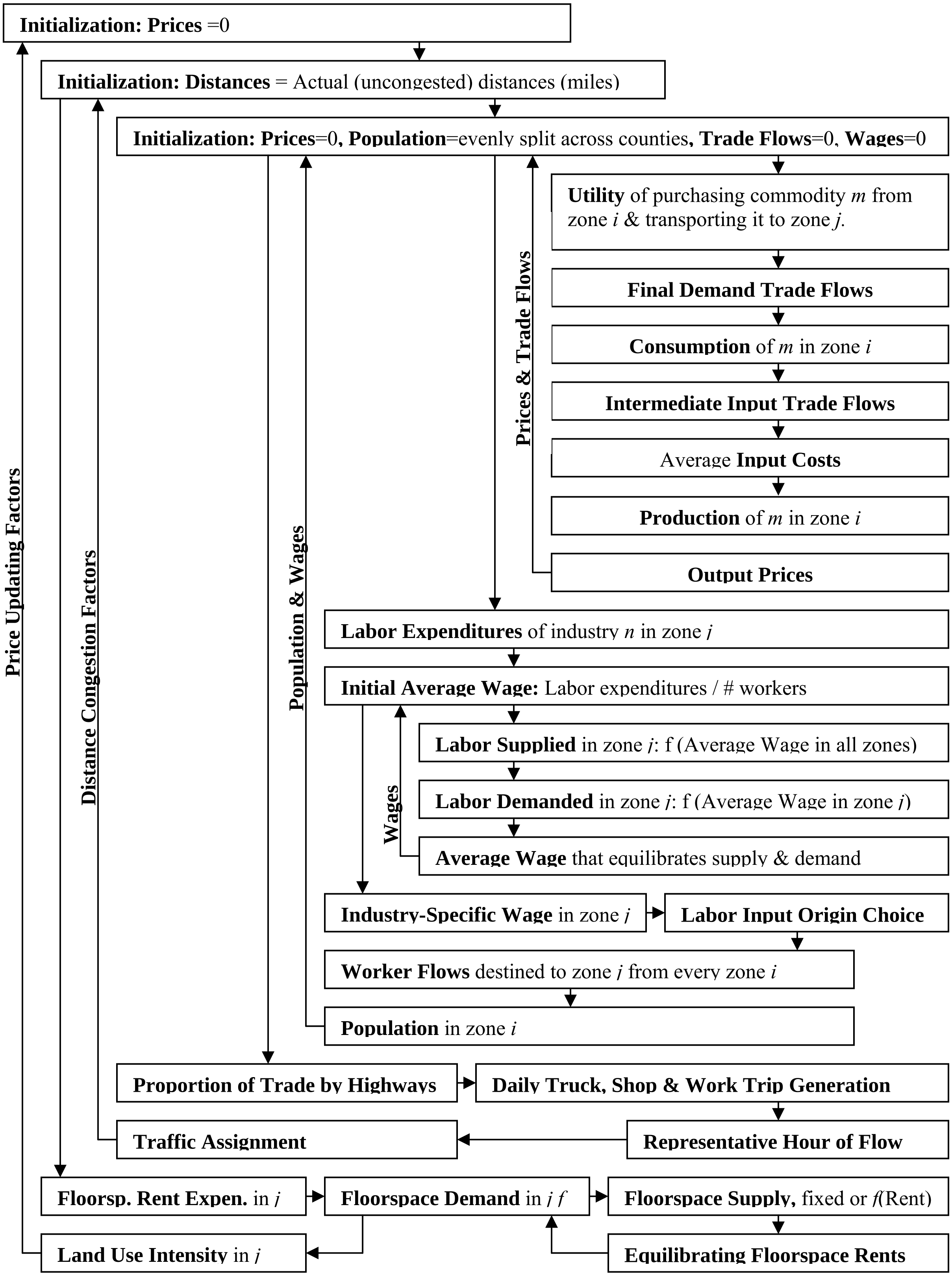

FIGURE 4: Estimates of Percentage Changes in Production Levels (for Priority [P] and Complete [C] Trans Texas Corridor [TTC] Networks) (as described in Ruiz-Juri & Kockelman 2006)

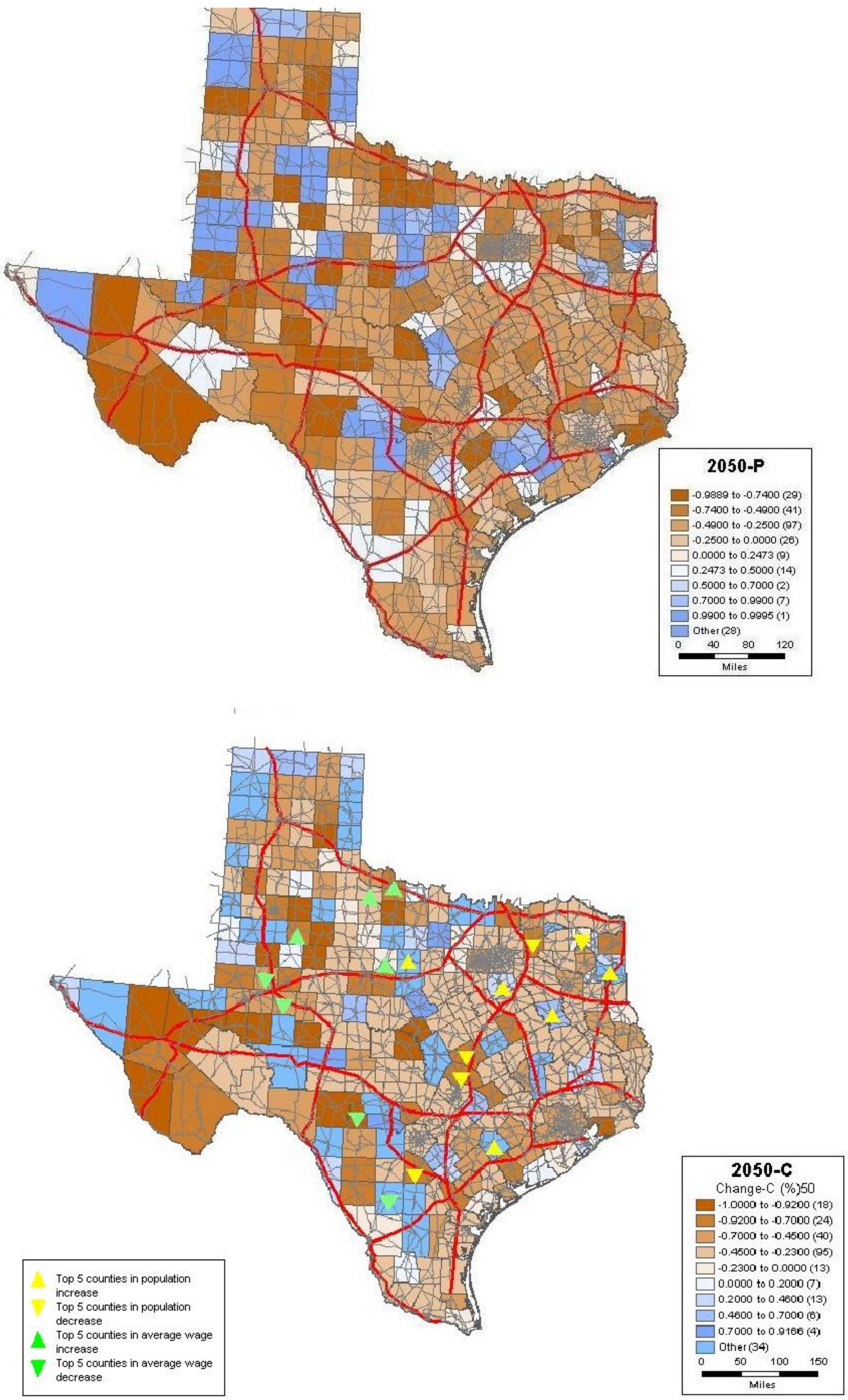

FIGURE 5: Estimates of Changes in Intrazonal Trade Percentages (for Priority [P] and Complete [C] Trans Texas Corridor [TTC] Networks)