DOCUMENTATION FOR APPLICATION OF DR. KOCKELMAN et al.'s RANDOM-UTILITY-BASED MULTIREGIONAL INPUT-OUTPUT (RUBMRIO) MODEL FOR U.S. TRADE PATTERNS

August, 2011

This document describes how to use the RUBMRIO model and code to track transportation and industrial production across a nation (e.g. U.S., or similar large-scale study area). The code, written in Visual C++ 6.0, is designed to allow practitioners, researchers, and others to make use of RUBMRIO spatial input-output model of trade.

Document Contents:

1. Model Description of RUBMRIO for the U.S. Application

RUBMRIO is an integrated transportation-economic model that simulates the flows of goods, labor, and vehicles across a multiregional area (e.g., across the 3109 counties in the U.S. continent, as well as foreign export shipments to 106 export zones, and other previous studies across 254 counties in the State of 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, 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 next 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 model loads domestic import flows onto the transportation network, contributing to congestion.

In this 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. As a preliminary study of apply RUBMIRO at a nationwide level, the current model only treats with trade equilibrium, without congestion feedback, land use constraints, or dynamic features. The model structure is shown in Figure 1, and a detailed description of the research can be found in

Tracking Transportation and Industrial Production across a Nation: Application of the RUBMRIO Model for U.S. Trade Patterns

To be presented at the November 2011 North American Meetings of the Regional Science Association International (RSAI), and under review for publication in *Transportation Research Record* (2011). With Xiaochuan Du

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)

- 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 Model: 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 Part B 38(9): 789-807 (2004). With Yong Zhao.

- Extending the Random-Utility-Based Multiregional Input-Output Model: Incorporating Land Use Constrains, 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.

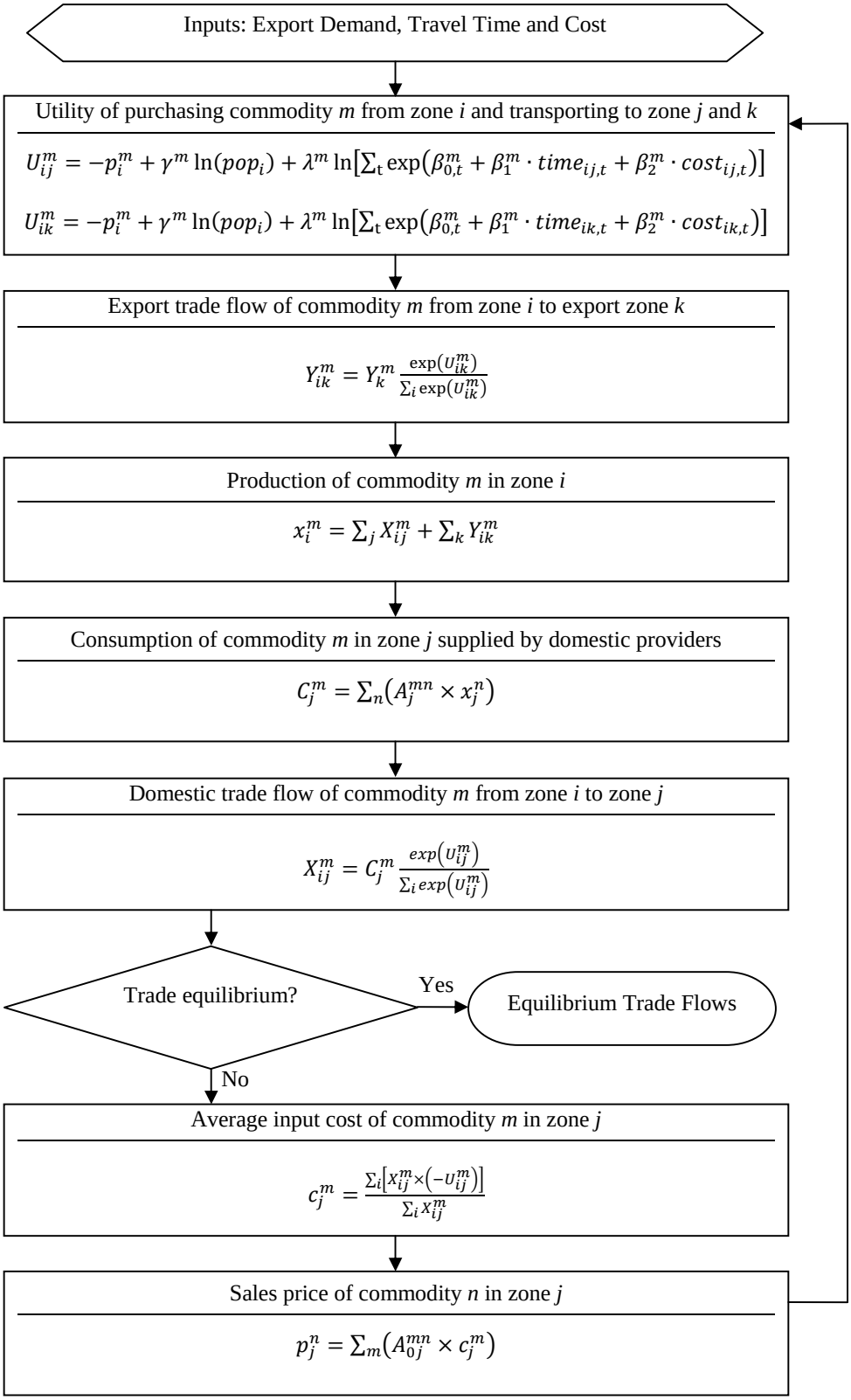

FIGURE 1 Model Structure and Solution Algorithm of RUBMRIO for the U.S. application

2. RUBMRIO Variables, Parameters, and Inputs for the U.S. Application

2.1 Variables

- Trade Utility (U_{ij}^{m}, U_{ik}^{m}): The disutility of purchasing commodity *m* in zone *i* and transporting it to zone *j* or export zone *k*.
- Export Trade Flow (Y_{ik}^{m}): The dollar value of foreign export trade flow of commodity *m* from zone *i* to export zone *k*.
- Domestic Trade Flow (X_{ij}^{m}): The dollar value of domestic trade flow of commodity *m* from zone *i* to zone *j*.
- Travel Time and Cost ($time_{ij,t}$, $cost_{ij,t}$): The travel time and cost from zone *i* to zone *j* (or export zone *k*) by mode *t*, which are calculated based on the corresponding transport distances.
- Zonal Production (x_{i}^{m}): The total production commodity *m* in zone *i*, which is represented by the total dollar value of both the export and domestic trade flow leaving zone *i*.
- Zonal Consumption (C_{j}^{m}): The total consumption of commodity *m* in zone *j* supplied by domestic zones (including local provider in zone *j*).
- Input Cost (c_{j}^{m}): The weighted average input cost of commodity *m* from all zones (include zone *j*) in zone *j*.
- Sales Price (p_{j}^{n}): The sales price of commodity *n* in zone *j*.

2.2 Parameters

- Number of Origin and Destination Zones (*i* and *j*): The total number of zones that comprise the study region. 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, which are indexed by *k*.
- Number of Economic Sectors (*m* and *n*): The total number of sectors that comprise the economy of the study region. This is both the number of production and consumption sectors, which are indexed by *m* and *n*, respectively. Table 1 shows the description of RUBMRIO sectors used in the U.S. application.
- Gama (γ^{m}): Gama is a parameter called by the upper level nested logit model, which reflects producer's attraction to an origin zone's size and prominence (approximated by its current population). The value of this parameter varies from different sectors.
- Lambda (λ^{m}): This is another parameter utilized by the upper level of nested logit model, which is used to determine the effect of transportation of the utility and trade flow between all zone pairs. The value of this parameter also varies across sectors.
- Beta (β^{m}): This is a vector of parameters that represents the sensitivity of travel times and costs of the two alternative modes (highway and railway) in the lower level of the

nested logit model. The $\beta^{m}_{0,highway}$ term is the alternative specific constant for the highway mode ($\beta^{m}_{0,railway}$ is set to zero to permit statistical identification of all parameters), β^{m}_{1} and β^{m}_{2} terms are the travel time and cost decay parameters for both of the two modes. All these parameters correspond to each of the industry sectors.

- Technical Coefficients without Leakage (A^{mn}_{0j}): This *m* * *n* matrix is generated from the original industry-by-industry transaction table, which reflects the productive technology of all zones, for inputs needs across all industries.
- Technical Coefficients with Leakage (A^{mn}_{j}): This is an *m* * *n* matrix that reflects the productive technology of all zones, with Regional Purchase Coefficients (RPCs) to account for "leakages" (i.e. purchases of commodity outside the study area, e.g., foreign import).

TABLE 1 Description of Economic Sectors in U.S. Application of the RUBMRIO Model

Sector Description		IMPLAN	NAICS	SCTG
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, 327, 321~323, 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	--	--	--

Note: This table provides the corresponding sector code in different data sources which were used in the U.S. studies. IMPLAN stands for Impact Analysis for Planning, NAICS stands for North America Industrial Classification System, and SCTG stands for Standard Classification of Transported Goods. Household and government are both treated as industrial sectors that buy from and sell to other sectors.

2.3 Inputs

- Foreign Export Demand (Y_k^m): This is an $m * k$ matrix of dollar value of output demanded by each foreign export zone by each sector.
- Population (pop_i): The population of each zone.
- Zone Radii: The radii of the circles with the same area of each county, which are assumed to be the intra-zonal travel distances.
- Highway Distances between Zones: This is an $i * j$ matrix of the highway distance between all zones, which is used to calculate the highway travel time and cost.
- Railway Distances between Zones: This is an $i * j$ matrix of the railway distance between all zones, which is used to calculate the railway travel time and cost.
- Highway Distances to Export Zones: This is an $i * k$ matrix of the highway distance from all zones to all export zones, which is used to calculate the travel time and cost.
- Railway Distances to Export Zones: This is an $i * k$ matrix of the railway distance from all zones to all export zones, which is used to calculate the travel time and cost.

3. Using the RUBMRIO code for the U.S. Application

3.1 Defining the study area and economy size

Users should define the number of the study area and economy size in the "user option" section, which locates at the beginning of the RUBMIRO code. The required setting includes the number of origins, destinations, export zones, and industry sectors.

3.2 Preparing and defining the inputs files

All of the input data and parameters must be written in corresponding ".txt" files as described below:

- Parameters.txt: This file contains the parameters for the origin and mode choice model (γ^m, λ^m, and β^m), the original technical coefficients matrix, and the regional purchase coefficients. The origin and mode choice parameters must be written as an $m * 5$ matrix, parameters for the same sector must be written in the same line with a column order of "γ^m, λ^m, $\beta^m_{0,railway}$, β^m_1, and β^m_2". The technical coefficients must be written as an $m * n$ matrix. Distinct parameters must be separated by skipped lines, and titled by a single string. And all of the data must be tab-delimited. Figure 2 provides a sample of this file.

```
GAMA_LAMBDA_BETA0_BETA1_BETA2
[Value – m*5]
(Skipped line)
Original_Technical_Coefficients
[Value – m*n]
(Skipped line)
Regional_Purchase_Coefficients
[Vaule – 1*m]
(end of file)
```

FIGURE 2 Sample of Parameters Input File

- Zone Info.txt: This file contains the radius and population of each zone, and should be written in the format described above. Figure 3 provides a sample of this file.

```
Zone_Radii
[Value – 1*m]
(Skipped line)
Zone_POP
[Value – 1*m]
(end of file)
```

FIGURE 3 Sample of Zone Info Input File

- Highway Distance between Zones.txt, Highway Distance to Export Zones.txt, Railway Distance between Zones.txt, and Railway Distance to Export Zones.txt: These 4 files contain the corresponding distance matrix described above, which must be written as a $i * j$ or $i * k$ matrix, tab-delimited, WITHOUT the title string.

The names of these input files are flexible, which should be defined in the "user option" section in the code. Users should also pre-define the storing location of these inputs before running the model.

3.3 Setting model sensitivity

The desired tolerance levels and maximum numbers of iterations should be set in the "user option" section in the code, which will govern the run time and exit strategy of the program.

3.4 Running the RUBMRIO model

Run the model exactly like any other C/C++ programs.

3.5 Output

RUBMRIO will write the output files to the directory defined in the "user option" section in the code. There are two main output files that contain the sector-combined export trade flows and

domestic trade flows separately. RUBMRIO will also generate files to output the export and domestic trade flows by sectors.