Draft for Review

Development and Calibration of the Statewide Land Use-Transport Model

Transportation and Land Use Model Integration Program
Phase II, Task 2.3

Prepared for

Transportation Development Branch
Oregon Department of Transportation
555 13th Street, NE
Salem, Oregon 97310

By

Parsons Brinckerhoff Quade & Douglas, Inc.
5801 Osuna Road, NE, Suite 200
Albuquerque, New Mexico 87109

And

Urban Analytics, Inc.
ECONorthwest
KJS Associates, Inc.

12 February 1999
(Third Revision)

Table of Contents

1. Introduction

The TRANUS package has been used to develop an integrated land use-transport model at the statewide level. It is based upon a specific modeling approach described by de la Barra (1989, 1995). For the prototype statewide model developed during this project, we used the model structure embodied within TRANUS, with little customization or deviation from standard practice. A fair amount of work in model specification was therefore already complete; a succinct description of the mathematical and algorithmic structure of the model is included in Appendices A and B. It is assumed that the reader is familiar with the paper and its contents.

This paper will outline the modelling details that are specific to the application of TRANUS in Oregon.

2. Design of the Statewide Application

This section outlines the key features of the Oregon statewide application of the Tranus model. The principal components that are defined in this TRANUS application are summarized in Table 1. A more extensive listing of model data elements, defined in terms of the principal components, is shown in Table 2. These data are described in greater detail in the following chapters of this report.

The period 1990 to 1995 was used as the base period upon which the model was calibrated. The simulation was carried out in five year intervals through the year 2030, although it is likely that ODOT will use 2020 as their planning horizon. The model is run over longer periods to test the temporal stability of the model and to identify emergent trends that might not be apparent in earlier time periods.

2.1 Model Schematic

The overall structure of the statewide model is shown schematically in Figure 1. This Figure depicts the interaction among the major model elements. Economic flows among industries and between households and industries are based on fixed demand coefficients each of which is indicated by an “f” in Figure 1. The demand coefficients, or technical coefficients, are derived from the I-O matrix of monetary transactions representing the Oregon economy. Derivation of the demand coefficients and the spatial disaggregation of economic flows is covered in Section 2.2.

Note that industries are grouped into 12 sectors for the initial statewide calibration. Originally, 27 sectors were proposed and employment data was developed for this level. The 27 sectors were aggregated to 12 sectors in response to software limitations (since eliminated) and to facilitate initial model development.

Demand for land of different types by industries and households is reflected through elastic demand functions indicated by an “e” in Figure 1. Development of data related to land market simulation has proved to be the most problematic and time consuming of the data efforts. The land market structure represented in Figure 1 is illustrative of the intent for the statewide model but the actual structure is still evolving as model and data-related development efforts progress.

Economic flows are translated into transport flows through the land use-transport interface. The economic flows contributing to each transport demand category are indicated by the numbers in the Transport Flows matrix in Figure 1. Transport categories 1-3 are commuter flows which are related to the flow of labor from households to industries. Non-commuter passenger demands are related primarily to the RETL, FIRE, SERV, and GOVT industry sectors whereas freight demands are related to the other industry sectors. Derivation of the land use-transport interface factors is covered in Section 2.4.

Only two modes have been defined for the statewide model – passenger and freight. This definition obviates the need for a separate mode choice model component. Instead, mode choice is combined with route choice in the assignment stage. Passenger demand theoretically is divided among several “operators” but in practical terms only auto exists for intercity travel in the calibration years. Freight demand is divided among light truck, heavy truck, and container truck in the

CHANGES

	AGFF	CONS	OMFG	WOOD	PRNT	TECH	TCPU	WLSE	RETL	FIRE	SERV	GOVT	HHIncLo	HHIncMi	HHIncHi	AGFF	IND	COM	SFR	MFR	RUR
Land supply																r	r	r	r	r	r
Exogenous demand	r	r	r	r	r	r	r	r	r	r	r	r	r	r	r						

ECONOMIC FLOWS

			Producing Sectors																				
			Industries												Households			Land					
			AGFF	CONS	OMFG	WOOD	PRNT	TECH	TCPU	WLSE	RETL	FIRE	SERV	GOVT	HHIncLo	HHIncMi	HHIncHi	AGFF	IND	COM	SFR	MFR	RUR
Consuming Sectors	Industries	AGFF	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	e					
		CONS	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f		e				
		OMFG	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f		e				
		WOOD	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f		e				
		PRNT	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f		e				
		TECH	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f		e	e			
		TCPU	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f		e	e			
		WLSE	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f		e	e			
		RETL	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f			e			
		FIRE	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f			e			
		SERV	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f			e			
		GOVT	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f			e			
	H'holds	HHIncLo	f	f	f	f	f	f	f	f	f	f	f	f							e	e	e
		HHIncMi	f	f	f	f	f	f	f	f	f	f	f	f							e	e	e
		HHIncHi	f	f	f	f	f	f	f	f	f	f	f	f							e	e	e

TRANSPORT FLOWS

		Producing Sectors															Modes	
		Industries												Households				
		AGFF	CONS	OMFG	WOOD	PRNT	TECH	TCPU	WLSE	RETL	FIRE	SERV	GOVT	HHIncLo	HHIncMi	HHIncHi	Passenger	Freight
Transport Categories	CmuteLo													1			1	
	CmuteMi														2		1	
	CmuteHi															3	1	
	Recreation	4								4		4	4				1	
	HBOther									5	5	5	5				1	
	NHBOther									6	6	6	6				1	
	NHBWork	7	7	7	7	7	7	7	7	7	7	7	7				1	
	VisitorBus																1	
	VisitorOther																1	
	LTFreight	8	8	8	8	8	8		8									2
	HTFreight	9	9	9	9	9	9		9									2
	CNFreight			10	10	10	10											2

		Passenger	Freight
Operators	Auto	1	
	Pass Rail	2	
	Intercity Bus	3	
	Lt Truck		4
	Hvy Truck		5
	Container		6
	Freight Rail		7

Figure 1: Structure of the statewide model

Table 1: Principal components of a TRANUS application

Component	Description	Primary Attributes[a]	Defined in file...
Scenario	A year/policy combination; at least one scenario will exist for each time interval.	Identifier, previous scenario	tranus.ctl
Sectors	Groups of homogeneous social or economic activities into which all activities are divided; households may be divided by size or wealth, while businesses are typically classified by their primary activity.	Identifier, disutility and price elasticities, scaling factor	L1E
Zones	Polygons into which the study area is divided to represent locii of socioeconomic activity; analogous to traffic analysis zones in traditional models.	Identifier, level (first, second, or external)	Z1E
Modes	A set of operators that users of a particular category can combine in order to perform their trips.	Identifier, path-building parameters	P0E
Operators	A homogeneous grouping of means of transport by capacity and other characteristics, such as automobile, bus, train, etc.; analogous to mode of transport in traditional models.	Identifier, mode, type, occupancy rate, minimum and maximum wait times, penalty[b]	P0E
Link types	Functional classification by which links are classified.	Identifier, administrator, penalty[b], maintenance cost	P0E
Administrators	Defines the agency responsible for the infrastructure; used only for reporting purposes.[c]	Identifier	P0E

a. Identifiers consist of a non-zero number (or number-character combination for scenarios) and a short string of characters describing the category.
b. Used to represent non-modeled characteristics such as comfort, reliability, safety, etc.
c. A single administrator (e.g., the government) is often used in TRANUS applications.

assignment process with rail handled exogenously in the initial model. The transport model is covered in Section 2.3.

2.2 Activity Model

The land use activity model is much broader in scope and has a much sounder theoretical basis than traditional approaches. It combines an input-output model of the Oregon economy with a spatial disaggregation model, also referred to as an activity distribution model, and a land market simulation model. The operation and mathematical formulation of the activity model are presented in the Tranus Reference Manual, Mathematical Formulation. Application of the model for the Oregon statewide implementation is outlined below.

2.2.1 Input-Output Matrix

The heart of the activity model is the input-output matrix representing the Oregon economy. The structure of this matrix for the statewide application and the treatment of the various economic flows is summarized in Figure 2. Note that the matrix in Figure 2 is the transpose of the corresponding matrix in Figure 1. The sense of Figure 2, with producing sectors on the left and flows going from top to left, is typically how I-O matrices are shown whereas the sense of Figure 1 is more typical of transport matrices. Each type of flow in the I-O matrix is represented by a differ-

Table 2: Statewide model data requirements

Category	Variables	Defined in terms of...							Primary attributes	Defined in file...
		Scenario	Sectors	Zone	Mode	Operator	Link type	Administrator		
Socioeconomic variables	Base year socioeconomic data		x	x					Exogenous production and consumption, induced production, unit price, value added, attraction factor	L0E
	Exports		x	x					Amount	L0E
	Imports		x	x					Unit price, amount, attraction factor	L0E
	Restrictions on internal production		x	x					Minimum and maximum amounts	L0E
	Location utility function parameters		x						Price elasticities by zone level, scaling parameter	L1E
	Demand function parameters		x						Minimum and maximum inputs, elasticity of demand to price	L1E
	Demand substitutions		x						Logit dispersion and scaling parameters, penalty	L1E
	Attractors of exogenous demand		x						Attraction function weights (zone level, production, price, and excess capacity)	L1E
	Attractors for induced production		x						Attraction function weight (zone level)	L1E
	Global increments of exogenous production and consumption		x						Production and consumption increments, minimum and maximum restrictions	L2E
	Increments of exogenous demand, production, and external zone exports and imports		x	x					Increment	L2E
	Increments of endogenous location attractors, production restrictions, and value added to production		x	x					Increment	L2E
Interface	Transport category		x						Time and volume conversion factors, directionality	F1E
	Intrazonal costs								Cost parameter by zone level	F1E
	Exogenous trips			x	x				Transport category, value, scaling factor	F1E

Table 2: Statewide model data requirements (Continued)

Category	Variables	Defined in terms of...							Primary attributes	Defined in file...
		Scenario	Sectors	Zone	Mode	Operator	Link type	Administrator		
Transport supply and demand	Transit lines					x			Identifier, frequency	P0E
	Transport demand categories								Available modes, value of travel and waiting time	P0E
	Energy and operating costs					x			Fixed and variable operating cost, energy consumption	P0E
	Operating characteristics					x	x		Maximum speed, operating cost	P0E
	User charges (fares and tariffs)					x			Time and distance cost and factors, transfer costs and prohibitions	P0E
	Links	x					x		endpoints, length, direction, capacity, transit lines, turn prohibitions	P1E
	Operator characteristics	x				x			Time factor, consolidation parameter	T1E
	Capacity restriction parameters	x					x		Speed-flow curve parameters	T1E
	Trip generation and mode split parameters	x							Elasticity, dispersion, and scaling factors	T1E

ent symbol to facilitate explanations. Following is a brief discussion of the flows in the input-output matrix all of which are annual dollar flows:

- Inter-Industry flows, indicated by "a" in Figure 2, are dollar flows from consuming sectors to producing sectors. Imports, produced by external zones and consumed by internal zones are included in these flows. Inter-industry flows, disaggregated to zone-to-zone flows, form the basis for freight demand and NHB passenger flows.
- The consumption of labor by industries and the production of labor by households is indicated by "l" in Figure 2. It is these flows which provide the basis for daily commuter travel demand.
- Consumption of goods and services by households, indicated by "p" in Figure 2, is represented in the I-O matrix by dollar flows from households to industries. These flows form the basis for home-based, non-work passenger demands.
- Other components of industry production are Exports, indicated by "x", and a residual category of final demand, OTHFD, represented by "o" in Figure 2. The latter category encompasses government payments, additions to inventory, and capital formation and is assumed to produce no significant transport demand. Exports, produced by internal zones and consumed by external zones, translate to freight flows and, potentially, to NHB passenger flows.
- Costs related to land consumption and related improvements are normally represented in I-O matrices as Value Added components. Since land is of special interest in the land use model, it is shown as separate production sectors with consumption indicated by "e" in Figure 2. No transport demand is generated by these flows.
- All value added components other than land and improvements are grouped in a single category, OTHVA, indicated by "v" in Figure 2. No transport demand is generated by OTHVA; it is included for accounting purposes only as is the OTHFD category.
- Total production and total consumption are represented by "T" in Figure 2. Theoretically these are equal to one another for each industry sector.

Flows represented in Figure 2 by "a", "p", and "l" are modeled through fixed demand functions usually referred to as technical coefficients. Derivation of these coefficients is covered below. Costs related to land and improvements can be represented in the model in several ways. Figure 2 reflects the most rigorous approach considered for the statewide model in which demand functions are elastic rather than fixed and substitution across land categories is possible for at least some consuming sectors. Options for land market simulation are discussed later in this section. The other component dollar flows shown in Figure 2, OTHFD, Exports, and OTHVA, are not modeled through demand functions per se but are input directly at zone level. OTHFD and Exports are represented as exogenous demand in internal and external zones, respectively. These values are what gets the demand-consumption cycle started. OTHVA is input as an average increment of value added to each unit of production.

2.2.2 Technical Coefficients

The fixed demand coefficients related to inter-industry flows (a), household production of labor (l), and household consumption (p) are calculated directly from the dollar transactions represented in Figure 2. If producing sectors are denoted by the subscript *m* and consuming sectors by the subscript *n*, then the coefficients are calculated as follows:

$$a_{mn} = \frac{IO_{mn}}{\sum_n IO_{mn}}$$

where: α_{mn} = amount of sector *m* required as input per unit of sector *n* production

IO_{mn} = dollar flow from sector *m* to sector *n*

$\sum_n IO_{mn}$ = total production for sector *m*

Although these coefficients may change over time due to technological or other changes, they are assumed to be fixed for the initial model implementation.

2.2.3 Spatial Disaggregation

Spatial disaggregation of the economic activity represented in the I-O matrix is made possible by input of zonal production measures. The key zonal inputs are employment by sector, households by income category, and imports and exports at the external zones. In order to maintain the accounting framework of the I-O matrix, all inputs are in terms of annual dollars in millions. Employment is expressed in dollar terms by factoring the number of employees in each sector for each zone by the sector average output per employee. Table 3 summarizes employment and production values for the statewide model. Note that the total production values in Table 3 correspond to the row totals in Figure 2 which are the base for the technical coefficients.

Households are represented in the model by labor production in annual dollars. Households are classified by income group (low, medium, high) in order to reflect spatial differences in labor production rates and household consumption rates by income group. Table 4 summarizes households by income and the factors to translate households into labor dollars.

Note that the calculation of labor production is done at zone level with zone-specific factors derived from census data. In order to match employee compensation totals contained in the I-O matrix, it was necessary to normalize the raw labor production values calculated at zone level. The labor production values in Table 4 are the sum of the normalized zonal values and correspond to the I-O matrix totals.

2.2.4 Land Market Simulation

Development of land area and price data has proved much more problematic, has involved much greater effort than anticipated, and has lagged behind other data development substantially as a result. Data development efforts and related problems are covered in Section 3.

Desired inputs to Tranus related to land market simulation include:

- Average land price/acre for 1990 & 1995/96 by category
- Average improvement price/acre for 1990 & 1995/96 by category
- Minimum and maximum lot sizes by category
- Elasticities of demand with respect to price by category

Table 3: Summary of 1990 employment and production

Economic sector	Employees	Annual production ($ million)	Annual prod/ empl ($ million)
AGFF	68,296	$4,458.2	$0.06528
CONS	68,311	$9,122.7	$0.13355
OMFG	71,171	$11,048.8	$0.15524
WOOD	78,202	$11,071.8	$0.14158
PRNT	14,704	$1,280.8	$0.08710
TECH	64,580	$8,169.3	$0.12650
TCPU	91,531	$9,672.2	$0.10567
WHSL	75,780	$4,311.6	$0.05690
RETL	248,581	$9,581.5	$0.03854
FIRE	92,434	$12,961.2	$0.14022
SERV	298,689	$19,577.5	$0.06533
GOVT	236,172	$7,574.7	$0.03207

Table 4: 1990 households by income and labor production factors

Parameter	Low income	Mid income	High income
Households	383,048	572,128	236,686
Average workers/household	0.672	1.279	1.771
Average income/worker	$7,301	$16,857	$29,831
Labor Production ($ million)	$2,447	$16,569	$16,552

Because of the data problems encountered, consideration has been given to a variety of possible approaches. For the statewide application, land market simulation is less critical than will be the case for sub-state or more detailed applications. Consequently, the range of options considered included ways to streamline the land market structure thereby reducing initial data requirements. Table 5 is indicative of the range of options considered. Listed for each option are the categories of land that would be defined.

Each of these options also has two sub-options which are:

1. Separate the cost of land and improvements

2. Combine the cost of land and improvements

Separation of land by urban and rural classification is designed to facilitate modeling of UGBs. Separation of urban and rural land anticipates major differences in price and density for a given use inside versus outside an UGB within a given zone. This division is important given the large zone size for the SW model and the fact that many zones will have both urban and rural land. If the need for urban-rural classification is accepted, Option 4 falls out. In the context of the initial statewide application, Option 3 is overly complex. Options 1 and 2 are considered more appropriate.

Table 5: Options for land market structure

Option 1	Option 2	Option 3	Option 4
Urban Rural Industrial Rural Commercial Rural Residential	Urban Ind/Comm Urban Residential Rural Ind/Comm Rural Residential	Urban Industrial Urban Commercial Urban SF Residential Urban MF Residential Rural Industrial Rural Commercial Rural Residential	Industrial Commercial SF Residential MF Residential

Yet another approach is to incorporate the costs of land and improvements in the model in the form of value added components but to not explicitly model the elastic consumption of land. Initial calibration runs are being made on this basis. This approach allows efforts to focus on other components of the activity model calibration and calibration of the transport model while further consideration and time is given to the land market and related data development. Since calibration is an iterative process and the activity distribution and transport models are critical in any event, this is an efficient way to proceed.

2.3 Land Use-Transport Interface

The activity-transport interface has a dual purpose. In the activity-to-transport direction, it translates economic flows in millions of dollars annually (disaggregated by sector and zone) to daily interzonal transport demand by category. In the transport-to-activity direction, it expands interzonal transport costs from daily to annual values consistent with the framework of the activity model.

The Tranus factors comprising the activity-transport interface are presented in Tables 6 and 7 for passenger demand categories and freight demand categories, respectively. A brief explanation of each column in Table 6 is given below:

1. Transport Category: These are the passenger demand categories defined for the transport model in Tranus.

2. Economic Sector: These are the selected economic flows to which each transport category is related.

3. TimeFac: This is the factor used to expand transport costs produced by the transport model to annual values which are added to production costs in the activity model. For passenger flows, TimeFac is set to the minimum value allowed by Tranus to minimize the impact of passenger demand on production costs.

4. VolFac Units: The units of the factor to convert activity flows to units appropriate to the trip generation model. For passenger flows, annual dollars are converted to equivalent households so that trip rates derived from the household surveys can be applied. Household/$mil factors are calculated based on total households and total production as shown in Table 6.

5. Cons>Pro & Pro>Cons: These factors indicate the proportion of trips in the Consumption-to-production direction and the reverse. Since the passenger trip generation rates represent daily demand, both factors are set to 0.5 so that half of the daily trips are in each direction.

6. Households (HHs): For Commute trips, the number of households is the total by income group; for all other categories total study area households are given.

7. Annual Production: This is the total annual production in $millions for the selected sectors. The values here correspond to sector row totals in the I-O matrix.

Table 6: Land use-transport interface: Passenger demand

Transport category	Economic sector	TimeFac	VolFac	Cons> Prod	Prod> Cons	House-holds	Annual production ($ million)	VolFac
1:CmuteLo	21:HHIncLo	0.0001	HH/$ million	0.5	0.5	383,048	2,447	156.5
2:CmuteMi	22:HHIncMi	0.0001	HH/$ million	0.5	0.5	572,128	16,569	34.5
3:CmuteHi	23:HHIncHi	0.0001	HH/$ million	0.5	0.5	236,686	16,552	14.3
4:Recreation	1:AGFF	0.001	HH/$ million	0.5	0.5		4,458	47.2
	9:RETL						9,582	29.3
	11:SERV						19,577	25.1
	12:GOVT						7,575	27.8
5:HBOther	9:RETL	0.0001	HH/$ million	0.5	0.5	1,191,862	9,582	34.1
	10:FIRE						12,961	19.9
	11:SERV						19,577	
	12:GOVT						7,575	28.8
6:NHBWrk	1:AGFF	0.0001	HH/$ million	0.5	0.5	1,191,862	4,458	11.0
	2:CONS						9,123	
	3:OMFG						11,049	
	4:WOOD						11,072	
	5:PRNT						1,281	
	6:TECH						8,169	
	7:TCPU						9,672	
	8:WLSE						4,312	
	9:RETL						9,582	
	10:FIRE						12,691	
	11:SERV						19,577	
	12:GOVT						7,575	
7:NHBOther	9:RETL	0.0001	HH/$ million	0.5	0.5	1,191,862	9,582	24.0
	10:FIRE						12,961	
	11:SERV						19,577	
	12:GOVT						7,575	

8. VolFac: The actual VolFac values in HHs/$mil units based on the preceeding columns. In the case of HBOther trips, several VolFac values are calculated to reflect more detailed trip purposes available from the survey data. Shop trips are related to the RETL sector; social and other miscellaneous trips are related to the FIRE and SERV sectors; and school trips are related to the GOVT sector which includes education.

The corresponding activity-transport interface factors are given for freight demand categories in Table 7. The bolded values are derived from survey data whereas assumed values are shown unbolded. Factors are shown only for the industry sectors which are primary freight produces. In contrast to passenger demand, freight demand is converted from dollars to tons and TimeFac represents true daily-to-annual factors. It is known from survey data what proportion of total sector flows are carried by each of the three freight categories (Lt Truck, Hvy Truck, and Container Truck). The average tons/$million factor for each sector is multiplied by the corresponding category proportion to obtain the VolFac values given in Table 7. Since freight flows from producers to consumers, the Pro>Cons factor is set equal to one and the reverse factor is zero.

Table 7: Land use-transport interface: Freight demand

Transport category	Economic sector	TimeFac (days/year)	Avg tons/ $ million	Category proportion	VolFac	Cons>Prod	Prod>Cons
8:LtTruck	1:AGFF	270	6,462.8	0.0387	250.1	0.0	1.0
	2:CONS		806.0	0.0826	66.6		
	3:OMFG		1,538.8	0.0801	123.3		
	4:WOOD		5,238.4	0.0407	213.2		
	5:PRNT		1000.0	0.1327	132.7		
	6:TECH		549.8	0.1712	94.1		
9:HvyTruck	1:AGFF	270	6,462.8	0.9613	6,212.7	0.0	1.0
	2:CONS		806.0	0.9174	739.4		
	3:OMFG		1,538.8	0.8061	1,240.4		
	4:WOOD		5,238.4	0.9564	5,010.0		
	5:PRNT		1000.0	0.7872	787.2		
	6:TECH		549.8	0.5955	327.4		
10:Container	1:AGFF	270	6,462.8	0.0	0.0	0.0	1.0
	2:CONS		806.0	0.0	0.0		
	3:OMFG		1,538.8	0.1138	175.1		
	4:WOOD		5,238.4	0.0029	15.2		
	5:PRNT		1000.0	0.0801	80.1		
	6:TECH		549.8	0.2333	128.3		

2.4 Transport Model

2.4.1 Transport modes and operators

The structure of the transport model is predicated upon the modes of transport to be studied. We have focused on person and freight movements which occur on the highway system. TRANUS defines modes of transport on two levels. The first level is called modes, and is defined as “a set of operators that users of a particular category can combine in order to perform their trips.” There may be one or more operators defined for each mode. It would appear from the documentation that the mode of transport denotes the ownership (public versus private) in most applications, and that the operators correspond to what we traditionally call modes in American modeling practice.

The modes of transport and operators that we have used in the model are listed in Table 8. The privately-owned automobile is the dominant mode of intercity person transport in Oregon, as it is in most Western states. A review of ODOT data on intercity public transport demand reveals a surprisingly small mode share, in part due to infrequent service. A separate mode for vanpools is included in the model, although there may be inadequate data to permits its inclusion in the Phase II work. The key characteristic of vanpools[1] which will set them apart from carpools (whether formally or informally formed) will be the ownership of the vehicle. The vanpool mode will exclusively use company or publicly owned vehicles dedicated to providing high occupancy commuting, and operated by the passengers themselves. While we could find no evidence of vanpools presently in operation in Oregon, vanpooling is a viable modal option that the model should be capable of evaluating. Specialized paratransit operations, such as on-demand transport for disabled persons, are included in the public transport mode.

Three modes of transport for freight are defined, as shown in the lower part of Table 8. The distinction between a light and heavy truck is based upon the powertrain configuration. An equivalency between the two modes and the 13 truck vehicle types used by ODOT and the FHWA is shown in Table 9. Heavy trucks are further divided between those transporting intermodal containers and those with all other types of trailers.

Table 8: Modes of transport

Operator	Mode	Description
Auto	Private	Privately owned passenger vehicle (all levels of occupancy)
Bus	Public	Intercity buses operated by common carriers.
Vanpool	Public	Vehicle used for high-occupancy commuting, operated by a passenger but owned or provided by an employer or public agency.
Light truck	Private	Light duty truck registered as a commercial vehicle. This mode can be used for commuting but otherwise cannot accommodate person transport.
Heavy truck	Private	Heavy commercial vehicle used exclusively for freight transport.
Container truck	Private	Heavy commercial vehicle which transports intermodal containers.

1. The term vanpool is taken in this context to denote any high-occupancy commuting vehicle, provided by a third party (such as an employer or traffic management agency).

Table 9: Definition of freight modes by truck and trailer types

FHWA type[a]	Description	Trailer type: Regular	Trailer type: Container[b]
3	2 axle, 4 tire single unit truck	Light truck	Undefined
5	2 axle, 6 tire single unit truck		
6	3 axle single unit truck		Container truck
7	4 or more axle single unit truck		
8	4 or less axle single trailer truck	Heavy truck	
9	5 axle single trailer truck		
10	6 or more axle single trailer truck		
11	5 or less axle multi-trailer truck		
12	6 axle multi-trailer truck		
13	7 or more axle multi-trailer truck		

a. Types 1 (motorcycle), 2 (passenger auto), and 4 (bus) are defined in the FHWA *Traffic Monitoring Guide* but not included.
b. Either an international or domestic container.

Since very few intercity trips are completed using non-motorized modes of transport, and are not viable substitutes for other modes of transport, they are not be included in the statewide model. Other motorized modes, such as rail and air, are worthy candidates for inclusion in the model. The work and data requirements for developing such components, however, are beyond the scope of this project.

2.4.2 Transport Networks

The intercity roadway network within Oregon consists of all roadways with a functional classification of rural minor arterial and above. The emphasis in the network— as with other aspects of the statewide model— will be on intercity connectors. Most populated places in Oregon are represented as single zones, which implies that the roadway system within a community is not included in the network. The only exception is for Portland, Salem, and Eugene, where the major roadways in the within the urban area are included. Nodes have no attributes associated with them in TRANUS other than their coordinates.

Links are defined as one-way or two-way and by link type. The capacity of the link may optionally be specified, as well as downstream turn prohibitions and transit lines circulating on the link. With the exception of transit line definitions, these attributes are already defined in the Oregon Highway Management System (OHMS) and were directly imported into the node-link coding convention required by TRANUS. This maintains consistency between the administrative and modeling networks used by the Department, and will facilitate the exchange of data between them.

Operating characteristics of the links, such as operating cost per mile, tolls, and free flow speed, are separately coded for each combination of link type and operator (defined above in conjunction with mode of transport). Most of these are not attributes defined in the OHMS representation, and thus can be specified by the analyst. Values of these parameters will be defined and adjusted during the calibration process.

The conversion from GIS coverage to TRANUS network is carried out by a program developed by the consulting team. The program operates on a roadway (line) coverage in ESRI shapefile format, and carries out the following functions:

1. The program reads a point coverage of populated places (zone centroids) and a line coverage of roadways. The program optionally reads a second line coverage which defines the external network (the roadway system outside of Oregon) and joins it to the first coverage using an user-specified table of node equivalencies at the border.

2. Nodes are defined by the endpoints of each line segment (no shaping points between them are kept). The program combines very close nodes together within a specified search radius[2] (typically less than a few meters).

3. Each zone is connected to the closest node in the network, subject to a maximum length specified by the user. If no eligible node can be found within the specified search radius, the program attempts to break an existing line segment by anchoring the zone to the nearest shaping point (thereby promoting it to a regular node in the process).

4. Links that connect to only one other link at either end are successively combined with the adjacent link to form a new, single link.

5. The resulting network is subjected to a series of validity checks, including dangling links, one-way links with no access or egress, illogical attribute values (such as missing or illegitimate speeds or undefined capacities), etc.

If no errors are detected during the network conversion, a network file in TRANUS format (file P0E) is written. An equivalency table is written to disk that contains the network links and the line segment(s) that comprise it.[3]

2.4.3 Transport network attributes

As previously noted, nodes are defined only in terms of their coordinates. Links have a number of properties, to include their length, type (functional classification), direction of travel, capacity, transit lines which circulate on them, and administrator. The user is given the opportunity to choose the type of administrator at the time the network is created. Typically a single administrator within Oregon will be specified (e.g,. the Oregon DOT), although other groupings may be used (such as ODOT districts).

---

2. This feature is necessary to ensure that two consecutive line segments indeed will share the same node between them. This is not always the case in line coverages; a difference between the endpoint of the first line segment and the beginning point of the next line segment as small as one foot is interpreted with the GIS as separate points. Some GIS packages are smart enough to span points that close while building paths through the network. The approach above guarantees such continuity in the line layer.
3. This equivalency file is used to pass information from the model back to the GIS at the conclusion of the modeling process.

A table of operating characteristics by functional class has been specified which will provide the data necessary to define link types. Values of such parameters are freeflow speed and speed at capacity are defined using data in the Highway Capacity Manual (TRB, 1994). Separate values for rural and urban areas will be developed if necessary.

2.4.4 Transport demand categories (trip purposes)

The developers of TRANUS use the term transport demand categories to denote trip purposes. We originally specified five trip purposes: commuting, business travel, recreational travel, tourism, and all other trips. The distinction between recreational and tourist trips was based on residency: recreational trips are made by Oregon residents, whereas tourism trips were made by nonresidents. There was room for some confusion with these purposes, as both residents and nonresidents could make business and other trips. We eliminated the confusion, as well as better segmenting the trip purposes, by replacing the tourism purpose with two visitor purposes (business and other). The revised trip purposes and their definitions are shown in Table 10.

Table 10: Transport demand categories

Category	Definition
Commuting	Journeys from the home to usual place of work and back made by residents[a]
Business	All occupational travel made by residents other than for commuting.
Recreational	Nonrecurring recreational travel by Oregon residents.
Home-based other	All home-based other travel by residents not defined in the previous three trip purposes.
Non-home-based other	All non-home-based other travel by residents not defined in the previous three trip purposes.
Visitor business	Trips within Oregon made by nonresidents, where the primary reason for their travel to Oregon is related to their occupation.
Visitor other	All travel within Oregon by nonresident for all purposes other than business travel.

a. A resident is a member of a household in one of the internal zones of the model (all Oregon households plus those in Clark County, Washington).

3. Data Base Development

Data development has been a very large part of the Phase II efforts. This section outlines the approach to developing the key data required for the statewide TRANUS application and some of the problems encountered. Specific areas covered are:

- Data Sources
- Zone System
- Transport Network
- Input-Output Matrix
- Employment by Sector
- Households by Income
- Land Market Data

3.1 Data Sources

This section, to be provided, will summarize major data sources for each of the principal datasets.

3.2 Zone System

The zone system (TAZs) for initial statewide model implementation consists 145 TAZs. External TAZs, representing major roadways and railways crossing the state boundary, account for 23 of these and the remaining 122 TAZs are internal including TAZ 101 representing Clark County. ODOT has produced a version of the Oregon State Map with TAZ boundaries and numbers over-laid.

We initially postulated a compact network in which activity nodes along a route at which trips could begin, change modes, or end their journey. These activity nodes would have replaced more traditional traffic analysis zones (hereinafter simply called zones), as intercity modeling focuses on the interaction between communities. This approach was not be possible with TRANUS, which does not permit paths to be built through zone centroids. We will continue to represent almost all populated places as single points, and connect them to the closest node in the network. In a major change from our previously described approach, we've moved away from activity nodes back to the traditional definition of zones. Zones within Oregon (internal zones) are defined in one of two ways:

- Zone ensembles from the metropolitan areas are used to represent their modeled areas. These range in size from four to five for the Salem modeling area to around 20 for the Portland met-ropolitan region. Note that Clark County, Washington is included in the Portland modeling area, and is represented as several internal zones in our model.
- The remainder of the state is defined around populated places with a population greater than 500 inhabitants, more or less as originally envisioned. The zonal boundaries will be defined in terms of Census tracts. A variety of population and housing data are available at the tract level, the boundaries of which are relatively stable between the decennial Censes. If anything,

tracts are more likely to be subdivided than have their boundaries changed. This will enable us to maintain linkages as zone boundaries change over time; smaller zones (tracts) can be aggregated into their earlier levels for comparison purposes.

These zones will define first-level internal zones within TRANUS. The program permits two levels of zonal definitions, with the second level of internal zones nested beneath the first. It is envisioned that eventually the first-level internal zones used at the statewide level will be subdivided for use in substate modeling applications. We have not, however, defined these second-level zone structure in our Phase II work.

A rich times series of socioeconomic data has been developed for the first level zones from a variety of existing data sources. U.S. Census of Population and Housing data were assembled for each first level zone for 1970 to 1990, and supplemented with forecasts from state and federal sources for 1993 through 2015. Comparable time series data were compiled for economic activities from the County Business Pattern data, U.S. Department of Commerce forecasts, and state economic inventories.

Approximately 14 external zones were defined. These correspond to roadways and railways which cross the border of Oregon, and carry 2,500 vehicles per day (for roadways) or 50 rail cars per days (for railroads).

3.3 Transport Network

Description of network development procedures and related software to be provided as well as summary statistics for the base year network.

3.4 Input-Output Matrix

A paper by Carl Batten, describing development of the input-output model for Oregon (including Clark County), was presented at the last peer group meeting. It can be found on the ODOT web page. At that time, development was still underway. Subsequently, a transaction matrix for the full 27 industry sectors and three household sectors was developed. This matrix has been consolidated to the 12 industry sectors and three household sectors being used for the initial statewide model implementation. Elements of final demand and value added have also been consolidated as indicated in Section 2.

Figure 3 presents the consolidated transaction table. Figure 4 contains the technical coefficients calculated from the Figure 3 transactions as indicated in Section 2.

3.5 Employment by Sector

Development of estimated employment by sector for 1990 and 1995 is covered in a paper produced for the last peer group meeting. It can be found on the ODOT web page.

3.6 Households by Income

Development of estimated households by income category for 1990 and 1995 is covered in a paper produced for the last peer group meeting. It can be found on the ODOT web page.

3.7 Land Market Data

The land market was initially segmented into four sectors; Single Family Residential Dwellings (SFD), Multiple Family Residential Dwellings (MFD), Commercial Uses (COM) and Industrial Uses (IND). As the data analysis proceeded it became rapidly apparent that a second single family dwelling sector was need and the Rural Residential (RUR) category was included in the analysis. The RUR sector can be differentiated from the SFD by the large rural lots and low levels of public infrastructure service that are provided to these dwellings. In Oregon, this sector is also located outside of urban growth boundaries and land that are classed as exception lands. That is to say lands that have been "excepted" from the resource preservation goals (agriculture land and forest land) of the state wide planning process. Two rural resource land sectors were also identified: Agriculture (AGR) and Forest Land (FOR).

The results of the land cost estimation process behaved in a manner that is generally consistent with the values expected in a land rent model. Land values tended to decrease with distance from major activity centers in the largest metropolitan areas. Land extensive uses consumed larger amounts of land area and had lower units costs for land. Social and economic factors related directly to land prices for single family housing. Rural resource lands that could not be developed for other uses had very low land values as a reflection of the economic value of the return from a small areas of land. It is possible that additional work on the data will refine or improve the actual estimates of land values, but such work will do little to change the over all relation between the individual zone values.

3.7.1 Data Availability / Lot Sizes

The availability of data was the major challenge faced in this project. Some type of land and building value data is available at the County level for the entire state. However, these data were not always available in a digital form or in a consistent format. There are several counties where digital data were not available. These counties are generally the smaller rural counties located in eastern Oregon. Digital county assessor data were also not available for 1990 from County Assessor or the Department of Revenue. There were a variety of reasons for this, including changes to county computer systems that occurred during the early 1990's.

Despite the data availability problems in specific geographic areas and for specific time periods, there were adequate data available for most of the model zones to allow land values to be calculated and to allow the missing values to be estimated for the remaining zones using methods described below. The available land sales and county assessor data were structured in such a manner that was it possible to produce estimates of land costs by sector. It should be noted that nearly all of the variables in the land sales data set and the assessor data set have one-tailed distributions.

3.7.2 Sales Data (Land or Land and Improvements)

Sale transaction data were the preferred data for estimating the cost of land because it provided a record of market value transactions. County Assessors and several commercial data providers were contacted regarding the extent of their land sales records in digital form. It was determined that sales data were not widely available in a digital format in Oregon. Several commercial data services had digital data, but primarily for the three metropolitan areas (Portland, Salem and Eugene) in the Willamette Valley. The sales record generally contained the site address, land use

Table 11: Distribution of sales records

County	Metropolitan Area	Number of records	
		1989 & 1990	1994 & 1995
Multnomah	Portland	15,832	29,440
Washington		10,215	22,091
Clackamas		9,259	16,949
Clark (Washington)		8,210	17,474
Yamhill		2,945	6,313
Marion	Salem	7,203	12,393
Polk		3,094	4,621
Lane	Eugene	8,136	11,850
Jackson	Medford	0	2,935
Benton		22	4,982
Linn		0	1,756
Total		61,971	130,804

type, sale price and date, zoning, building square footage, number of bedrooms, baths and rooms and lot size. The value most often missing was the lot size.Approximately 167,000 sales records were purchased from DataQuick for the years 1989, 1990, 1994 and 1995. The geographic distribution of this data is shown in Table 11.

The sales records contained site address for most of the lots and/or buildings. These addresses were used to allocate records to individual mode zones using GIS address matching procedures. A portion of the available records were not matched to individual address within zones and these records dropped out of the pool of records available for the land cost estimation process. This reduction in the number of records was only a problem in the rural zones where some general land use categories have few, if any, records to begin with.

Sales data for vacant land was the preferred source of land costs. There were not enough vacant land sales records to estimate the average land cost at the county level by land sector. The majority of these sales records were for land that would be used for single family homes.

In order to have enough records to estimate the land cost for all sectors it was necessary to use the sales records for land and improvements. Once again the majority of land sales records were for the sales of single family homes. While these land and improvement sales records contained a land value field , the data in that field was the land value assigned to a parcel by the County Assessor and not the sale price of the land determined by a market transaction or estimated from a market transaction.

3.7.3 County Assessor Records

Copies of the County Assessor rolls were obtained from the Department of Revenue for 31 of the 36 counties in Oregon for FY 94-95. The assessor records provided information on the type of land use, real market value of land and improvements as determined by the Assessor and in many cases the lot area. Lot area was the most common missing variable. The assessor records generally did not include building size. The assessor records provided the basis of the land cost estimates for those counties where no land sales data was readily available. Unfortunately, a couple of these files did not contain usable cost information.

Most of the individual assessor records contained of a property class code that described the general land use type for the particular record. The P Class system was established by the Department of Revenue and is used by the assessors in more or less the manner proscribed by DOR. The land uses are generally classified as listed below:

100’s	Single Family Residential
200’s	Commercial Uses
300’s	Industrial Uses
400’s	Residential Tracts - (Generally Rural Residential)
500’s	Agriculture
600’s	Forest
700’s	Multifamily Residential
800’s	Recreation
900’s	Public Uses

This classification system allowed the assessor records to be segmented into the sectors used by the model. The assessor records did not have site addresses that could be used to assign individual records to model zones when a county contained more than one zone. They did however have tax code districts which are polygons that represent unique combinations of taxing jurisdictions. These districts are normally based on school districts and the polygons tend to cluster around urban areas where local governmental districts are the most numerous. Using the tax code maps obtained from the county assessors, it was possible to allocate the assessor records to individual model zone for further processing.

The lack of assessor records for 1990 presented a larger problem to the project goal of developing land values for 1990 and 1995. While 1990 tax rolls were not available, individual county assessors had printed copies of their annual tax code summary of total assessed value by jurisdiction for both FY 89-90 and FY 94-95. Using these total assessed values it was possible to estimate what the 1990 values were given the 1995 values. The 1990 data was developed using the following methodology.

- First, 1995 land values were estimate using the methodology described in Section 3.5.
- The next step was to remove all new taxable construction that occurred between 1990 and 1995, as described in FW Dodge construction summary, from the total change in assessed value from FY 89-90 to FY 94-95 to remove any growth in the assessed value associated with new construction.
- The next step was to convert FY 94-95 total assessed value in to 1990 dollars by using the GDP deflator to remove any growth in the assessed value associated with inflation.
- Finally, the average annual real growth in assessed value was calculated.
- This real growth factor was then used to deflate the 1995 land costs by sector to 1990 values.

3.7.4 Missing Land Area in Land Sales Data and Assessor Records

The data collected for this project represented several hundred thousand records. A substantial portion of these records were lacking one critical piece of information that limited their usefulness to the project: there was no land area data recorded. This situation is the result of a process that county assessors have historically used to track the land area in tax lots. The assessor is required to estimate the area of an individual tax lot when it is not known. All lots in platted subdivisions have a legal lot area that is defined by the subdivision plat. The assessor relies on the areas supplied by these plats if there are questions concerning the area of a platted lot. But for lots outside of platted subdivision, the assessor is required to estimate the lot area. This is not an easy process and it is subject to challenge by land owners who may disagree with the assessor results. Since the assessor is not required to tract the area of subdivision lots, they are commonly not entered into the assessor rolls and hence are not available in the digital copies of the assessor rolls. The assessor roll is one of the primary sources of land areas for the land sales data. As such both the land sales data and the assessor records have a substantial number of records that do not have a lot area.

This problem can be clearly in the tables below. Over one third of the single family lots did not have an lot acreage in the county assessor records. Generally speaking, lots outside urban subdivisions are larger that lots in platted subdivisions. Accordingly, the lots used to estimate land cost in for project tend to have larger average lot areas than would be the case if the platted subdivision lots were included in these estimates. Based on this assessment, it is assumed that the estimated land values will tend to be lower on a cost per square foot basis than they would have been if all of the lots had a land area. This biases in the land value appeared to occur in all land sector values in all zones and was assumed to have a fairly uniform impact on the estimated land values.

3.7.5 Data Distribution of Records by Tranus Land Sector

Metro's Regional Land Information System (RLIS) contains the largest single land parcel data set in Oregon, with more than 460,000 parcel level records for 1996. This data set was a convenient place to look at the distribution of lot sizes and the magnitude of the missing lot data problems. The RLIS data set has undergone considerable review in recent years and is considered to be one of the best data sources of current data in Oregon. The RLIS data set is primarily based on County Assessor data from Clackamas, Washington and Multnomah counties. Metro has access to the assessor and GIS information in Clark County, Washington, but these data are not routinely included in the RL:IS data. The RLIS project began about 1990. The original Metro parcels data records are not readily available and are of unknown condition.

The data in all of the tables have been categorized by the same set of lots size categories to provide a consistent distribution of data for all land sectors. Each land sector table contains two sets of estimates of average lot size. The first mean lot size estimate is based on the land area listed by the County Assessor. The second mean lot size estimate is based on the area of the map polygons in the RLIS base map. In many cases the polygon lot area is larger than the assessor lot area.

The Single Family Residential land market (Table 12) in the Portland area had a substantial number of lots (38.06%) with missing lot sizes. Analysis of RLIS polygons estimate that these lots had an average area of 12,884 square feet, which was 22% larger than the average lots area for SFD

Table 12: Distribution of lot sizes for single family residential land use in Metro, 1996

Lot area (ft^2)	# of lots[a]	% of total	Average Assessor lot area (ft^2)	Average RLIS polygon lot area (ft^2)
Missing Lot Values	122,904	38.06%	0	12,884
0 - 2,499	1,360	0.42%	1,738	3,075
2,500 - 4,999	62,208	19.26%	4,526	4,750
5,000 - 6,999	40,839	12.65%	6,155	6,197
7,000 - 9,999	44,149	13.67%	8,198	8,245
10,000 - 14,999	20,944	6.49%	12,158	12,212
15,000 - 19,999	7,640	2.37%	17,250	17,844
20,000 - 24,999	5,244	1.62%	21,941	22,025
25,000 - 43,499	2,165	0.67%	33,615	46,832
Greater than 1 Acre	15,491	4.80%	223,489	182,932
Total Records	322,944			
Mean Lot Size			6,882	10,584
Median Lot Size			4,791	7,637
Standard Deviation			42,515	27,012

a. Based on Assessor Lot Area from RLIS data.

Table 13: Distribution of lot sizes for rural residential land uses in Metro, 1996

Lot area (ft^2)	# of lots[a]	% of total	Average Assessor lot area (ft^2)	Average RLIS polygon lot area (ft^2)
Missing Lot Values	1,077	22.43%		186,476
0 - 2,499	2	0.04%	1,307	83,381
2,500 - 4,999	1	0.02%	4,356	6,281
5,000 - 6,999	1	0.02%	5,663	5,647
7,000 - 9,999	1	0.02%	8, 712	7,925
10,000 - 14,999	3	0.06%	11,182	25,643
15,000 - 19,999	0	0%	-	-
20,000 - 24,999	3	0.06%	23,666	25,983
25,000 - 43,499	13	0.27%	37,562	37,011
Greater than 1 Acre	3,701	77.07%	311,332	272,611
Total records	4,802			
Mean Lot Size			260,619	251,957
Median Lot Size			158,816	162,626
Standard Deviation			530,996	467,000

a. Based on Assessor Lot Area from RLIS data.

RLIS lots as a whole. The mean area for all SFD RLIS lots is slightly larger than the typical lot size (10,000 ft^2) expected for conventional single family development.

The RLIS SFD lots had a mean lots size 54% larger than the assessor mean lost size of 6,882 square feet. Generally speaking, the average lot size in each category was approximately equal for the assessor records and the RLIS records. Assuming that this relationship holds for all assessor records, the assessor data should produce a overestimate of the SFD land price in those area where not land sales data was not available. The comparison of the RLIS data and the assessor data in the Portland metropolitan area shows a lots size relationship that is contrary to the expected relationship discussed in the previous section. It is not known if this is a consistent data pattern through out the state or just the data pattern for the Portland metropolitan area.

The Rural Residential (RUR) land sector (Table 13) was dominated by the development of single family houses or mobile homes on lots larger than one acre. Within the Oregon land use planning framework this development is limited to small areas outside of UGBs. More commonly these lots tend to be in the 2 to 5 acre range and are served by minimal level of public infrastructure services such as sewer and water. While the number of lots in this land sector was relatively small in the metropolitan Portland area, this land use was a significant portion of the single family housing market in many other areas of the state.

The mean lot size for the assessor data and the RLIS data were approximately equal. The average RLIS lots size for lots with missing assessor land areas was approximately 30% smaller than the average lot size for an assessor lot. Based on the data in this table land cost estimates base on assessor data were expected to be similar to those estimated from land sales records.

The Multiple Family dwelling land sector (Table 14) consumed substantially less land than the single family land market in most urban areas. But it can be home for as much as half the population of a given urban area. This was especially true for some of the jurisdictions with large colleges or large scale suburban multiple family development patterns. Large MFD development, more than 300 units on a site, are still a relatively rare phenomena in Oregon.

The multifamily sector has the second highest percentage (28.8%) of lots with missing assessor lot areas. The average lots size of the MFD lots with missing assessor values was relatively small (7,201 sq. ft.) which is in keeping with the expected values discussed earlier. These small lots were expected to support small scale MFD projects that were more typical of older development patterns. The mean lot area for an RLIS polygon was approximately 10% less than the mean for an assessor polygon.

The Commercial land sector (Table 15) presented an interesting problem. For lots with an area of less than one acre, the mean values were comparable. However, for lots greater than one acre, the mean assessor lot area were approximately 65% larger than the RLIS lot area for the same lots. This difference was made more interesting by the fact that on the whole the sale prices for commercial land in the north Willamette Valley tends to be higher than the total assessed value for this land.

Commercial land contained a substantial percentage of lots without land area and the RLIS lot area for this lots averaged just over one acre (46,739 ft^2). Mean lots size for lots with assessor land areas were approximately 30% greater than the lots size for the RLIS polygons. If this distri-

Table 14: Distribution of lot sizes for multi-family residential lots in Metro, 1996

Lot area (ft^2)	# of lots[a]	% of total	Average Assessor lot area (ft^2)	Average RLIS polygon lot area (ft^2)
Missing Lot Values	4,860	28.81%		7,201
0 - 2,499	668	3.96%	1,681	1,655
2,500 - 4,999	4,384	25.99%	4,395	4,649
5,000 - 6,999	1,463	8.67%	5,933	5,980
7,000 - 9,999	1,905	11.29%	8,276	8,253
10,000 - 14,999	996	5.90%	12,415	12,326
15,000 - 19,999	419	2.48%	17,324	17,027
20,000 - 24,999	330	1.96%	22,020	21,947
25,000 - 43,499	596	3.53%	33,572	33,036
Greater than 1 Acre	1,247	7.39%	191,368	144,014
Total Records	16,868			
Mean Lot Size			19,664	17,746
Median Lot Size			4,791	5,098
Standard Deviation			102,540	58,313

a. Based on Assessor Lot Area from RLIS data.

Table 15: Distribution of lot sizes for commercial land use in Metro, 1996

Lot area (ft^2)	# of lots[a]	% of total	Average Assessor lot area (ft^2)	Average RLIS polygon lot area (ft^2)
Missing Lot Values	3,323	19.44%		46,739
0 - 2,499	475	2.78%	1,555	2,190
2,500 - 4,999	2,961	17.32%	4,147	4,452
5,000 - 6,999	1,223	7.16%	5,998	6,357
7,000 - 9,999	1,518	8.88%	8,520	10,744
10,000 - 14,999	1,734	10.15%	12,667	14,769
15,000 - 19,999	1,088	6.37%	17,441	17,505
20,000 - 24,999	946	5.53%	21,972	22,667
25,000 - 43,499	661	3.87%	33,707	39,138
Greater than 1 Acre	3,163	18.51%	456,474	275,785
Total Records	17,092			
Mean Lot Size			73,485	56,496
Median Lot Size			10,018	10,976
Standard Deviation			580,524	339,563

a. Based on Assessor Lot Area from RLIS data.

bution is matched in other parts of the state, values based on assessor records will tend to be lower than expected if sales data were available.

The Industrial Land (Table 16) had a lot values that were very comparable regardless of the source of the land area. The average lots sizes were approximately equal to each other. This was expected given the fact that substantial amounts of industrial land use occur outside of subdivisions. This urban land sector had the highest proportion of lots larger than one acre (49%). It also had the smallest percentage of missing lot area values (15.7%). The average lot size for lots with missing land areas was just under 3 acres (112,255 ft^2).

The Agriculture Land sector (Table 17) and the Forest Land sector (Table 18) were generally not available for future urban developed in the Oregon land use planning system. 98 percent of the agriculture lots and 98 percent of the forest lots had a lot size greater than one acre. The average lots sizes for these two sectors were approximately equal when comparing the assessor and the RLIS data.

3.7.6 Land Cost Estimation Methodology

Several methods for estimating land values were considered and a number were tested before the final estimates of land prices were undertaken. The testing was conducted using the most numerous type of records, SFD, in 33 model zones representing four counties in the northern Willamette Valley. As the result of this testing, the Residual Land Value method was chosen as the land cost estimation methodology. Each of the candidate methods are discussed in the following sections.

One aspect of the land cost coefficients that should be remembered when reviewing the results of this process is the fact that land and improvement prices are the result of the interaction of numerous market factors. The price of single family land was determined by a number of factors that make up the housing bundle. These included the relative attractiveness of a particular location in terms of external factors such as schools, distance to work and shopping, neighborhood amenities, availability and quality of public infrastructure services (sewer, water, streets, and storm drainage), and distance to the CBD, and internal factors such as household size, household income, vehicle ownership etc. The final estimated land prices produced by this process was affected by all of these factors and others.

- *Vacant Land Sales*

  The first option explored for estimating land cost by land sector was the use of vacant land sales data. These records were expected to provide the cleanest estimate of the average cost of land by sector. This cost would be estimated by the use of the following formula:

  $$LandCost = VacantLandCost / VacantLotArea$$

  Unfortunately there were not enough records of vacant land sales to estimate land cost in most of the test zones. In addition, most of the sales records were for single residential land and very few records were available for any of the other sectors. Consequently this method was dropped early in the process.

- *Residual Land Value*

  Several land cost estimation methodologies were developed after the vacant land method was dropped. The Residual Land Value method was eventually chosen as the preferred one for

Table 16: 1996 distribution of lot sizes for industrial land use in Metro

Lot area (ft^2)	# of lots[a]	% of total	Average Assessor lot area (ft^2)	Average RLIS polygon lot area (ft^2)
Missing Lot Values	766	15.73%		112,255
0 - 2,499	72	1.48%	1,612	2,283
2,500 - 4,999	608	12.49%	4,295	4,536
5,000 - 6,999	197	4.05%	6,033	8,910
7,000 - 9,999	255	5.24%	8,433	8,296
10,000 - 14,999	330	6.78%	12,698	18,222
15,000 - 19,999	205	4.21%	17,446	18,146
20,000 - 24,999	295	6.06%	21,693	20,649
25,000 - 43,499	518	10.64%	31,408	43,316
Greater than 1 Acre	1,623	33.33%	285,401	230,350
Total Records	4,869			
Mean Lot Size			87,556	87,920
Median Lot Size			15,246	20,243
Standard Deviation			357,453	338,145

a. Based on Assessor Lot Area from RLIS data.

Table 17: 1996 distribution of lot sizes for agricultural land use in Metro

Lot area (ft^2)	# of lots[a]	% of total	Average Assessor lot area (ft^2)	Average RLIS polygon lot area (ft^2)
Missing Lot Values	550	8.67%		723,286
0 - 2,499	1	0.02%	1,742	492,714
2,500 - 4,999	6	0.09%	3,411	44,594
5,000 - 6,999	3	0.05%	5,955	4,932
7,000 - 9,999	5	0.08%	8,538	34,572
10,000 - 14,999	11	0.17%	12,436	85,690
15,000 - 19,999	15	0.24%	17,716	182,790
20,000 - 24,999	15	0.24%	22,446	1,162,947
25,000 - 43,499	75	1.18%	35,096	353,202
Greater than 1 Acre	5,663	89.27%	1,122,532	990,715
Total Records	6,344			
Mean Lot Size			440,259	447,538
Median Lot Size			221,720	211,405
Standard Deviation			717,258	770,740

a. Based on Assessor Lot Area from RLIS data.

Table 18: 1996 distribution of lot sizes for forest land use in Metro

Lot area (ft^2)	# of lots[a]	% of total	Average Assessor lot area (ft^2)	Average RLIS polygon lot area (ft^2)
Missing Lot Values	43	1.88%		797,277
0 - 2,499	1	0.04%	1,742	296
2,500 - 4,999	2	0.09%	3,267	4,041
5,000 - 6,999	-		-	
7,000 - 9,999	7	0.31%	8,586	148,272
10,000 - 14,999	5	0.22%	13,242	412,662
15,000 - 19,999	10	0.44%	17,468	61,286
20,000 - 24,999	3	0.13%	22,072	182,341
25,000 - 43,499	16	0.70%	32,888	308,378
Greater than 1 Acre	2,203	96.20%	3,373,739	2,539,070
Total Records	2,290			
Mean Lot Size			277,738	282,881
Median Lot Size			183,170	209,157
Standard Deviation			256,481	247,028

a. Based on Assessor Lot Area from RLIS data.

areas where there were an adequate number of sales records. The Residual Land Value method uses the following equation to estimate residual land values:

$$LandCost = (SalesPrice - (BldgSize * ConstCost - ((Age/67)*(BldgSize * ConstCost))))$$

BldgSize is measured in square feet. The intent of this equation is to base land value on sale prices of land and improvement less the replacement value of the improvements. The residual value of the land reflects the value of an improved lot ready for development.

The replacement value of the improvement, primarily the value of the building, is deducted from the sale price. The per square foot average cost of structures by land market sector was calculated from FW Dodge construction estimates for the state of Oregon for 1990 and 1994. The building area in square feet was derived from the sales records..

Building values are depreciated using a straight line (accounting) depreciation of 1.5% per year or an assumed life span of 67 years for the structure. Metro noted in its Housing Needs Analysis that the typical housing unit depreciates at a rate of between 1.0% and 1.5% per year. A non linear depreciation model was also explored but it did not substantially effect the land vales estimated using this method in the testing phase of the project.

The depreciation rate has the greatest impact on estimation of land values associated with buildings older than 50 years. Housing units listed in the US Census as older than 50 years account for approximately 25% of all building in all of the model zones. These older units are more numerous in rural area and in smaller counties and less numerous in the larger metropolitan areas where much of the states growth has occurred in the last 40 years and where most of the sales data records are located. Land values for lots with buildings older than 50 year may tend to be somewhat overestimated, depending on the nature of the local real estate market and the condition of the specific building.

- *Simple Hedonic Pricing (Regression)*

  A set of simple hedonic price models were tested at the zone level. The model used simple linear regression with Sale Price as the dependent variable and Building Age, Building Size in Square Feet and Lot Size in Square Feet as the independent variables. The general form of this equation is as follows:

  $$SalesPrice = b_1 + b_2{*}Age + b_3{*}BldgSize + b_4{*}LotSize$$

  The regression equations were estimated with and without a constant value. The variables' coefficients were interpreted as the cost for each of the factors that contribute to the estimation of the sale price.

  The predictive power of the individual equations was generally low. However, regression coefficients were strongly correlated with the residual land values estimated (see Section 3.5) in the previously described methodology. The hedonic price methodology produced negative coefficients more frequently than the residual land value model.

- *Mortgage Equivalent*

  The discussion of alternate methodologies include the possibility of calculating an average monthly mortgage payment based on the average sale price. After some discussion, it was decided that the large number assumptions that would be necessary to estimate this value would increase the probably and the magnitude of any errors in the estimated land price. As a result this methodology was dropped from consideration.

3.7.7 Assessor Land Values

County assessor records had to be used in 22 counties to estimate either 1990 land values or 1995 and 1990 land values for the land sectors in model zones. The assessor records collected from the Department of Revenue did not contain building size data, so the residual land value method could not be used to estimate land costs. Instead, a simpler estimation methodology use the following equation:

$$LandCost = TotalAssessedValue / LotArea$$

During the test phase of this process, assessor land values were estimated and them compared with the values calculated through the residual land value and simple hedonic model. The correlation between the assessor land values and the other two was not as strong as the correlation between residual land value and the hedonic values as shown in Table 19. The process for deflating the FY 94-95 values to FY 89-90 values has been previously described.

Table 19: Correlation between residual and Hedonic land values

Pearson Correlation	Res Land Value	Assessor	Hedonic
Res Land Value	1.000		
Assessor Value	0.514	1.000	
Hedonic value	0.883	0.300	1.000

3.7.8 Other Measure - Unit Pricing

A number of other methods for estimating land prices were discussed early in the process. Two of these methods continued to be options for future use. One method was Unit Prices. The Unit Price approach used the prices for a housing or building unit instead of the cost of land on a per square foot basis. While the price of a unit of single family housing is fairly easy to explain,. it is harder to explain one unit of a commercial or manufacturing building. The other method was the Combined Land and Improvement Price, which is measured in total cost per square foot of land. This measure was easy to compute and combined the value of the land and the building into a single measure. These estimated values should be influenced by the size of the building as well as the size and the locational aspects of the lot.

3.7.9 Missing Values, Outliers and Other Data Problems

When the land price estimation process was completed, the results were reviewed to insure that they were reasonable. It was determined that a few of the land price values were not reasonable. Three distinct type of problems were identified:

- Negative Land Values
- Outlier Values
- Missing Values

Alternative value were substituted for the values derived from the model in each of these cases. The model zone and land sector cells that contain alternate values are identified in the individual excel spreadsheet by formatting them in BOLD and Italics. The values that were substituted to solve these problems were calculated from the same data sets as the problem values. The value replacement process followed a specific sequence that is unique to each problem. The value substitutions were made using the following decision rules:

- Negative Land Values - The Residual Land Value process can produce negative land values under the correct set of circumstances. This occurs when the replacement value of the building is greater that the price paid for the land and improvements. It is more likely to occur when there are a smaller number of case in an individual zone for a particular sector or when the sale prices in a zone is relatively low for a particular sector (i.e., that sector is not an attractive place for the particular land use). Replacement values were chosen for the negative land values in the following order:
    1. Median Land Value for the Zone and Land Sector
    2. County Mean Value for the Sector
    3. Mean value for the Sector from an adjoining zone that is similar.
- Missing Land Values - In some zones there were not sales in a particular sector or there was no data for a sector or zone. In these cases, the missing values were estimated by choosing in the following order
    1. County Mean Value for the Sector
    2. Mean Value for the Sector from an adjoining zone that is similar.

- Estimated Value is an Outlier - In general the data sets used to estimate the land value had a one tailed distribution. It is possible for a small sample or a very skewed data distribution to produce estimated values that are outliers when compare to with all adjoining value. This occurred in 5 cases. An example was an estimate value of $63 per square foot for commercial land in a suburban metropolitan zone. Replacement values were chose using the following order:
  1. Median Land Value for the Zone and Land Sector
  2. County Mean Value for the Sector
  3. Mean value for the Sector from an adjoining zone that is similar.

Copies of the estimated land values are found in Appendix D.

3.7.10 Land Consumption and Supply Issues

The last major data set required was for land supply. At the statewide model zone level land supply information is very generalized. It is possible to collect land information at the policy level (i.e., land that is zoned or planned for a particular use or model sector). This would require contacting 211 cities and 36 counties. 136 of the cities are smaller than 2,500 people and are not likely to have a planning staff to handle such a request.

Actual land use data was even harder to come by. In theory, this could have be done by using the Assessor P Class data field that has been discussed previously. However, it would have been necessary to have a computer mapping system or GIS to allocate these land uses and to clean up the missing lot size data. These systems do not exit statewide. In fact, they do not exist in two of the eleven counties that will be in the substate mode; and two of the mapping systems that do exist are relatively new and only have limited amounts of data. There are other approaches that could be used to approximate this data, such as using the habitat “GAP” mapping for endangered species, but each of these data sets have there own sets of problems.

The method put forward in this section should provide a generalized land supply estimate by policy category that will be sufficient for the first efforts at model calibration. Substantial additional work could be undertaken to improve this data set at the state wide level, but it is probable that this effort would only result in a marginal improvement in the data developed using the method outlined in this section. If additional effort is to be undertaken to improve the land supply estimates, it should be done in conjunction with the Substate Model where the land supply estimates are more critical to the operation of the model.

3.7.11 Urban Growth Boundary (UGB) Constraints and Location of Land Available For Development

At the state wide model level urban land is only somewhat constrained at this point in time. HB 2709 requires that each city / urban area have an UGB that contains a twenty year supply of land given the current rate and pattern of development inside the UGB. The UGB’s surrounding the 211 cities in the state generally meet this requirement. If a particular UGB does not have enough land to accommodate the projected growth for the next twenty years, the UGB must be expanded. Metro is going through this process at the moment. This process insure that there will always be adequate land for future urban development in Oregon urban area.

Under the policies set out in the Oregon planning system the first choice for expansion is “exceptions land’ (i. e., excepted from meeting the requirement of the resource preservation goals). These lands are generally denoted in the TRANUS model as the rural residential (RUR) sector. There is limited supply of non-residential land available in the exception areas. In general commercial and industrial uses are expected to stay within a UGB. It is also expected the UGB’s will grow by the inclusion of lands that are adjacent to UGB’s. The last choice for expansion is agriculture land (AGR) and forest land (FOR), but some of the land in these sectors may be included in a expansion if there are no other land sectors. For the purposes of the state wide model the urban land supply is unconstrained by the operation of state law and the urban land supply is expected to change to meet future growth requirements. It is possible to develop a constrained growth scenario as one of the future options tested by the model.

3.7.12 Land Supply By Plan Category

The land supply by sector was derived from the 1:100,000 zoning maps obtained from the State Service Center for GIS (SSCGIS). This map showed the land uses permitted by policy and zoning at a county level. The map was developed in the early 1980’s when the first round of comprehensive plan acknowledgments had been completed. The maps were developed by local Planning Departments and represent the best readily available maps. Outside of the UGB there has not been substantial change in the types of permitted land use since these maps were drawn. There are some differences between this coverage and the SSCGIS coverage for the UGB’s. Appropriate adjustment were made to the urban land areas to reflect the updated UGB polygons.

Generally speaking the land polygon areas shown on the state wide zoning map can be divided into four categories - Forest, Agriculture, Rural Residential (with minor amounts of commercial and industrial) and Urban. The total land area for each use were summarized for each zone. Then the urban land area was divided between the four urban categories - SFD, MFD, COM, and IND. The RLIS parcel database was analyzed in its entirety and average allocations of land in the urban polygon were estimated. These allocations were then compared with the housing unit and employee databases to develop a per unit and per employee allocation of gross acres of urban land. The average percentage of public land for schools, parks and public uses was also estimated, as was a percentage for public rights-of-way. When no better information was available, these percentages were used to allocate the land in the urban polygons to the urban land use sectors.

It was recognized that thesee allocations were gross urban acres that include vacant developable land. This vacant land was included in the total land allocated to each urban sector. Forest and Agriculture land was assumed to be not available for development. However, Rural Residential Land was assumed to be developable.

3.7.13 Interior Urban Model Zones and Land Supply Constraints

There were a small number of metropolitan model zones that did not contain land outside the UGB. These Interior Urban zones had limited growth capacity in terms of the amount of vacant developable land. These zones experienced growth through both the development of the remaining vacant land and through the redevelopment of existing developed lands at higher densities. LCOG has provided an estimate of the amount of vacant developable land in each of the interior zones for use in the model. It is possible to estimate similar information for the Oregon side of the

Portland metropolitan area from the RLIS data. It will be necessary to obtain a similar estimate from MVCOG for the Salem area or to use the Eugene or Portland Values as appropriate.

3.7.14 Land Consumption Functions - Building Space Consumption Functions

The initial calibration of the statewide model used a global set of land consumption functions. For each land sector a minimum and a maximum land consumption function was estimated. The minimum land consumption function represents that minimum amount of land that will be consumed by one household or the amount that will be consumed to support one employee. The maximum function likewise represents an upper limit on the amount of consumption per household and per employee by land sector. Average land consumption estimates (average lot sizes) were calculated during the land price estimation process and are available to check the land consumption estimates produced by TRANUS in the base year (1990) and in 1995.

The minimum and maximum land consumption functions were estimated using two sources. The minimums were estimated by assuming the maximum density allow in the proposed Metro Regional Framework Plan for each land sector. The maximum land consumption was estimated using existing data on residential development in Oregon and employment density, FAR, and employee per square foot of building data from the State of Michigan as presented in Fiscal Impacts of Alternatives Land Development Patterns in Michigan: The Cost of Current Development Versus Compact Growth, March 1997 which was conducted for SEMCOG by Rutgers University.

Building floor space consumption was estimated in a similar manner from the same sources. Average building size data is only available for the counties where there is sales data. In 1995, sales data was available for 10 counties that contained 70 of the 122 internal model zones. The single family residential sector has the most numerous set of records in this data base. The estimated land and building consumption function are shown in Table 20.

An alternate set of six consumption function tables has been proposed that make adjustments to the land and space consumption functions to better reflect the differences in land and building market that exist in different parts of Oregon. As of this time no decision has been made concerning the use of the alternative consumption functions.

Table 20: Land and building space consumption functions

Category	Use	Land consumption per household or employee (ft^2)		Space consumption per houshold or employee (ft^2)	
		Minimum	Maximum	Minimum	Maximum
Urban	SFD	3,500	90,000	1,000	50,000
	MFD	750	10,000	500	3,000
	COM	150	3,000	300	1,500
	IND	1,000	5,000	500	2,500
Rural	RUR	40,000	900,000	1,000	50,000
	AGR	850,000	28,000,000	500	2,500
	FOR	850,000	28,000,000	500	2,500

3.7.15 Lot Size and Building Size Data

Several model zones in the substate model area have estimated land prices that are based on a relatively small data set. There are two possible sources for this problem. First there may only be a few records because there have on been a few sales of land in a particular sector in a particular zone. There is little that can be done to improve the estimate with this type of problem as long as land sales records are the preferred source of data although a review of the records that did not address match successfully may result in an increase in the number of records in a particular zone.

The second set of problems exists when there are a number of records but many records are missing the data for the lot areas. This data can be improved by reviewing local assessor records and determining the lot size for specific land sale records. This project was limited in magnitude and offers the strong probability of improved land price estimates. This work would also provide benefit to the sub-state model work.

Building area by land sector is a desirable data set for TRANUS modeling. No definitive data set exists for this model variable. It is possible to construct a rough estimate of the SFD building area by model zone from a combination of Census data (number of units by year built before 1990), sales data (average building size by year built) and Center for Population Research and Census at PSU building permit data by jurisdiction (number and type of building built since 1990). In addition, the Metro's RLIS data contains data on building square feet. This data base is being updated and improved by the Natural Hazards Section of the Growth Management Department and is expected to be available in December 1997.

The Metro data base appears to be the best source for building size information for the MFD, COM and IND land sectors. FW Dodge data, the and Center for Population Research and Census at PSU building permit data and other construction data will provide a method of checking this information.

It is possible to build a generalized approach to estimating the total building square footage for the model zones from the Metro data and the Census data and then apply this method to produce estimates of building area in the other area of the state.

The residual land cost estimation model used a straight line depreciation method. Some consideration should be given to the question of whether or not to use a different method of depreciation for estimating the land values.

3.7.16 Land Supply by Generalized Plan Category

The methodology for estimating the total land supply for the state wide model can be improved upon. The estimates should be revisited during the sub-state modeling process where the supply of land in each zone is a key variable. The process of collecting this data may be more appropriately undertaken by ODOT. It will require contacting the cities and counties to obtain estimates of the amount of land in each of the land sectors, both built upon and available for development. This process is expected to take a fair amount of follow up work to insure that data is gathered in a timely manner.

4. Model Calibration

The construct of integrated land use and transport models dictates that model components are highly interrelated. Because of this, the process for calibrating such models typically is iterative in nature. The Oregon statewide application of the TRANUS model is no exception.

Figure 4-1 outlines the calibration process for the statewide model. It is both iterative and cyclical in nature. Three major cycles of calibration and testing were involved with numerous iterations within each. The major cycles were:

1. Initial base year (1990) calibration,
2. Initial 1990-1995 calibration/validation, and
3. Long-range application and sensitivity analysis.

Within each of the first two cycles it was necessary to iteratively apply the model with adjustments to activity model parameters and/or transport model parameters until results for both models were satisfactory. In the third cycle, the model was applied in five-year increments from 1990 through 2020 with various input variations to test the stability and reasonableness of results and, where necessary, to further refine the model parameters. Calibration results are outlined in the following sub-sections.

4.1 Base Year Calibration Model (Cycle 1)

A considerable amount of the effort involved in the initial calibration was simply getting all of the input data in the right form and in the right units and insuring that each model component was operating as intended. This entailed a greater amount of effort than is typical, even compared to previous TRANUS applications, for several reasons including:

- The degree of geographic detail
- The degree of sectoral detail
- Comprehensively modeling both freight and passenger flows
- The “full accounting” approach incorporating all input-output transactions
- Expressing all inputs in annual dollar terms

Having accomplished this, base year calibration of the activity model focused on three primary sets of parameters:

- Activity model distribution parameters
- Transport model passenger trip generation parameters
- Transport model freight trip generation parameters

A distribution or dispersion parameter must be developed for each of the twelve industry sectors and for each of the three household sectors. Elastic trip generation functions must be developed for each of the seven passenger transport demand categories and the three freight demand catego-

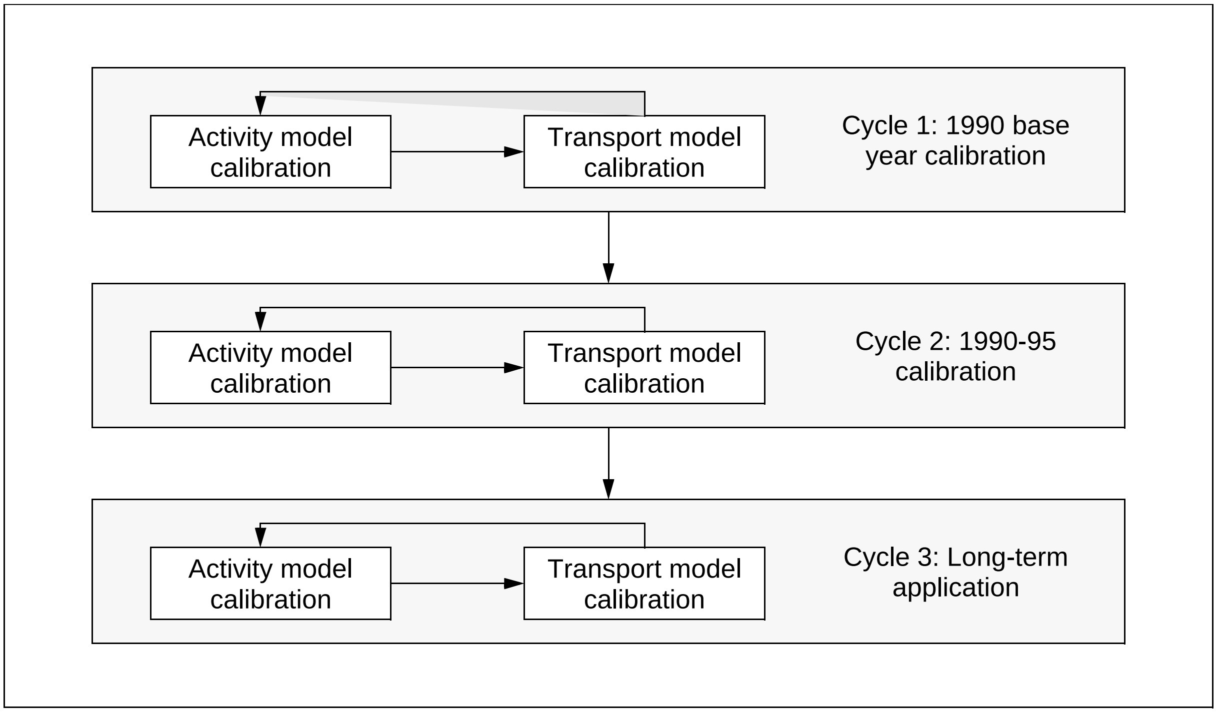

Figure 3: Statewide model calibration process

ries. Three parameters must be developed for each trip generation function: the minimum generation rate, the maximum generation rate, and the elasticity of demand. TRANUS trip generation functions control not only the number of trips generated but the trip length frequency distribution.

4.1.1 Base Year Activity Model Calibration

Primary criteria for base year calibration of the activity model are:

- Agreement between target and modeled production by sector, by zone
- Modeled prices by sector, by zone reasonably close to $1,000,000
- Model-produced shadow prices on the order +/- $200,000 (20%), randomly distributed

In base year calibration mode, the activity model iterates for a prescribed number of iterations or until closure criteria are met. Starting with input exogenous demand for internal zones and exports in external zones, the model allocates production to meet this demand across all zones based on the logit model formulation incorporating potential production and costs of production in each zone. Allocated production to satisfy exogenous demand generates additional demands for inputs which must be satisfied by additional production, and so on.

In order to match input production levels by zone in the base year, TRANUS develops what is termed a “shadow price” for each sector and zone. Since zones represent alternatives in the allocation of production activity, these shadow prices amount to alternative specific constants representing non-modeled effects. Since production is expressed in millions of dollars annually in the model inputs, the price per unit of production generated by the model is expected to be $1,000,000. In addition, the shadow prices created by the model to satisfy the target production

Table 21: Range of modeled production prices (current dollars)

Sector	Minimum	Maximum
AGFF	993,577	1,011,517
CONS	938,397	1,010,171
OMFG	941,023	999,578
WOOD	953,811	1,021,792
PRNT	918,712	1,000,999
TECH	919,992	999,161
TCPU	913,901	996,866
WLSE	887,173	1,031,235
RETL	1,003,226	1,099,731
FIRE	967,028	1,018,031
SERV	956,583	1,045,870
GOVT	926,858	1,119,932

inputs are ideally small relative to $1,000,000 and randomly distributed about zero. Unless something is drastically wrong, the activity model will replicate target production values by sector and zone very closely for the base year including target imports at the external zones.

For the base year, the activity model reaches a near equilibrium in less than 200 iterations with reasonably close matches to calibration targets. The model matches input production by zone precisely for every zone and every sector as expected. Modeled production prices are quite close to the expected value of $1,000,000 as indicated by the range of output production prices shown in Table 21.

Shadow price adjustments vary considerably on a zone-by-zone basis by sector but exhibit quite similar overall patterns. For each sector, the magnitude of the zonal shadow price adjustment is highly correlated with and inversely related to zonal production. This is illustrated in Figures 4 through 11 related to sectors AGFF, TECH, RETL, and SERV, respectively. Each of these figures is a scatterplot of the zonal shadow price adjustment (in percentage terms) versus zonal production where each point represents a single zone. All twelve of the industry sectors exhibit patterns similar to those in Figures 4 through 7. Adjustments are predominately positive in sign and decline rapidly in percentage terms as the value of production increases. Adjustments are within 20 % (+ or -) except for zones with very small production values. This is a healthy result.

Figures 8 and 9 are the corresponding scatterplots for low-income and high-income households, respectively. While broadly similar the scatterplots for industry sectors, these show a more balanced distribution of positive and negative adjustments. There is still a very strong trend to smaller percentage adjustments as zonal production increases, especially for low-income and medium-income households (latter not shown). High-income households exhibit more pronounced adjustments suggesting they are influenced more by non-modeled factors than are lower-income households - a result consistent with other empirical evidence. Nevertheless, the adjustments rarely exceed the 20% criteria.

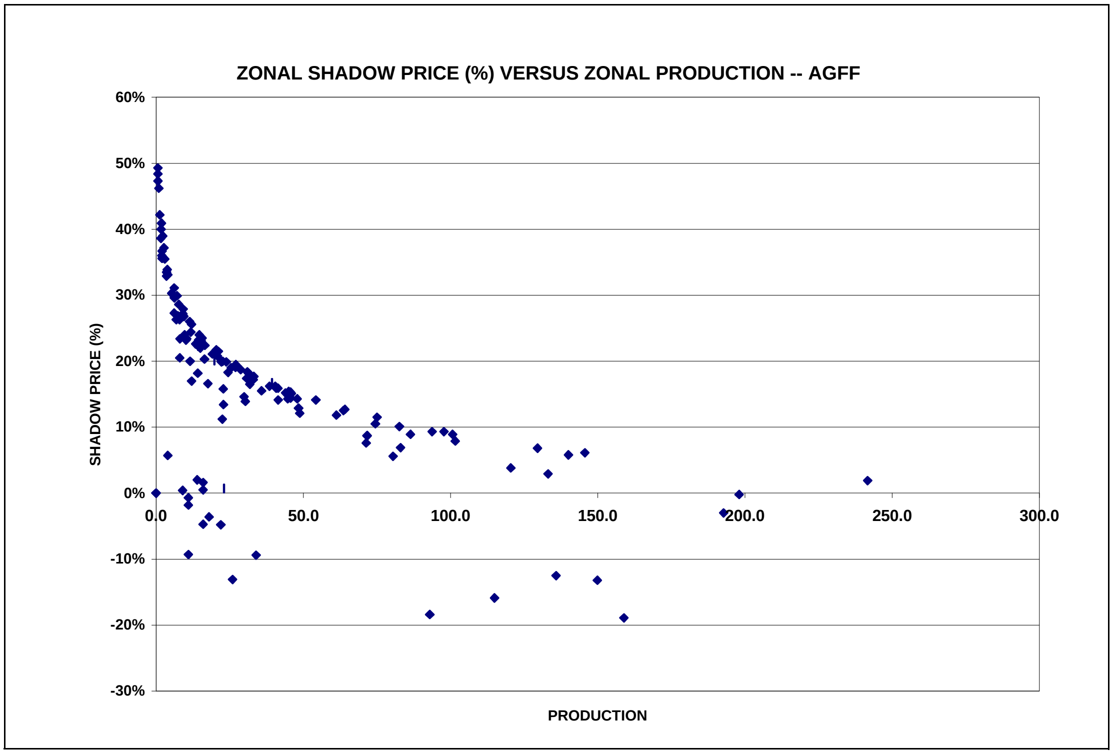

Figure 4: Scatterplot of AGFF price adjustments

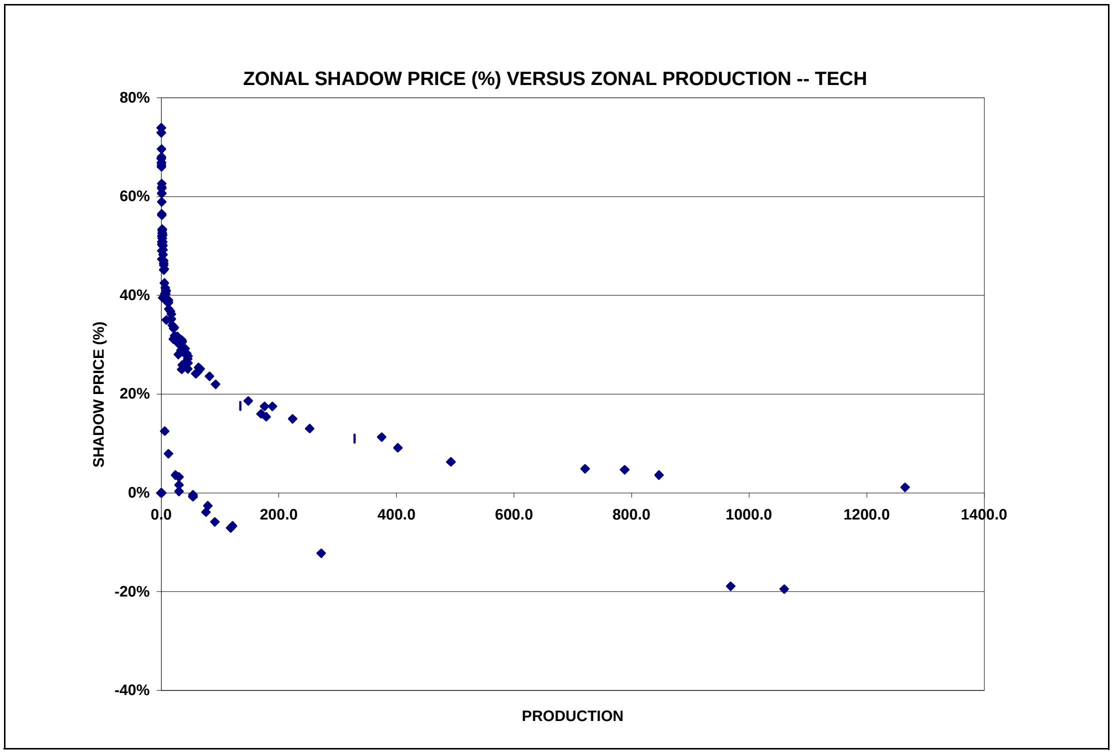

Figure 5: Scatterplot of TECH price adjustments

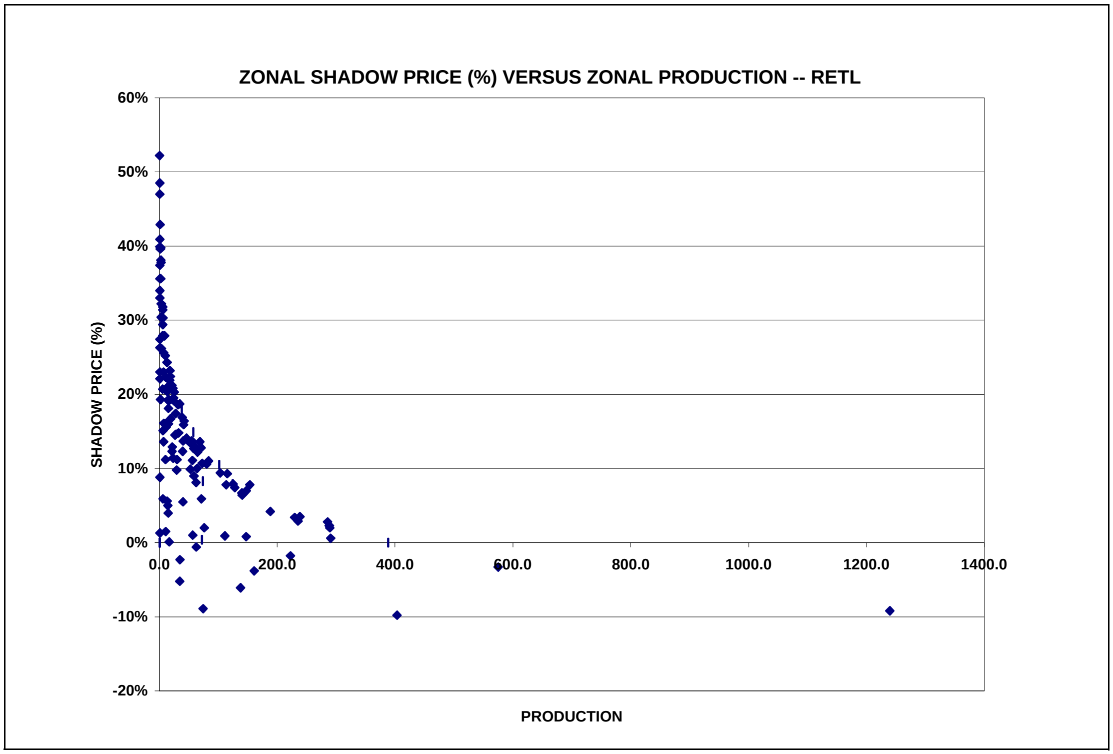

Figure 5: Scatterplot of RETL price adjustments

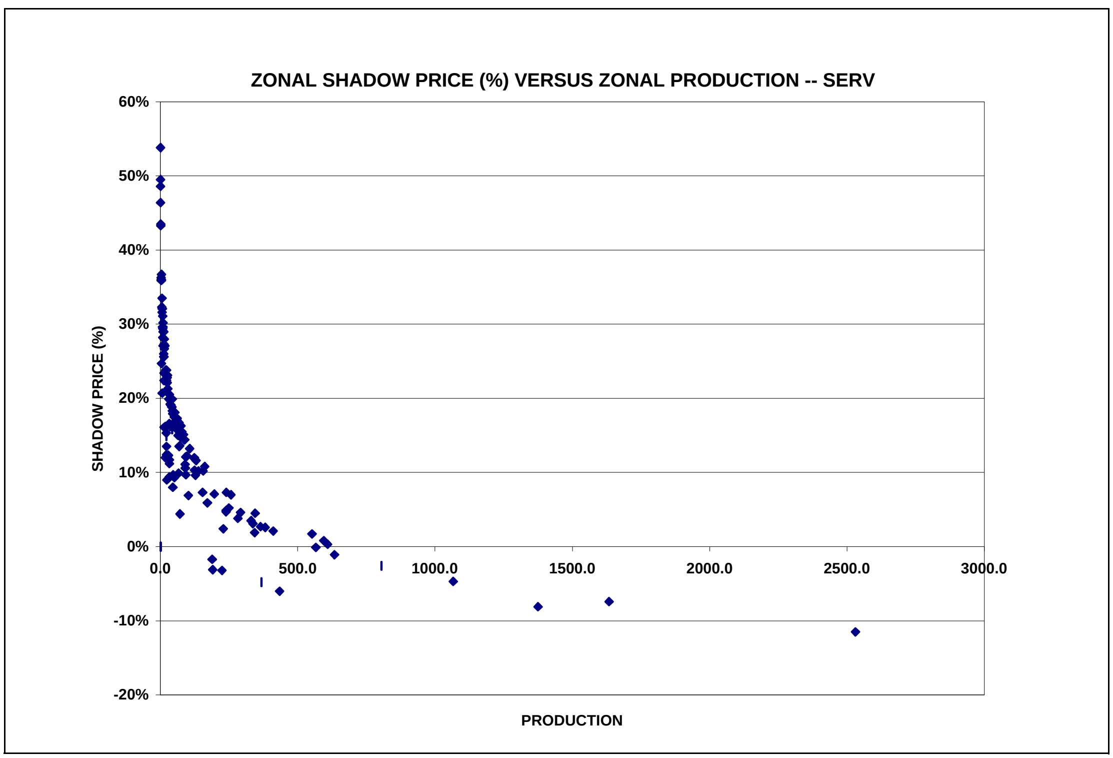

Figure 7: Scatterplot of SERV price adjustments

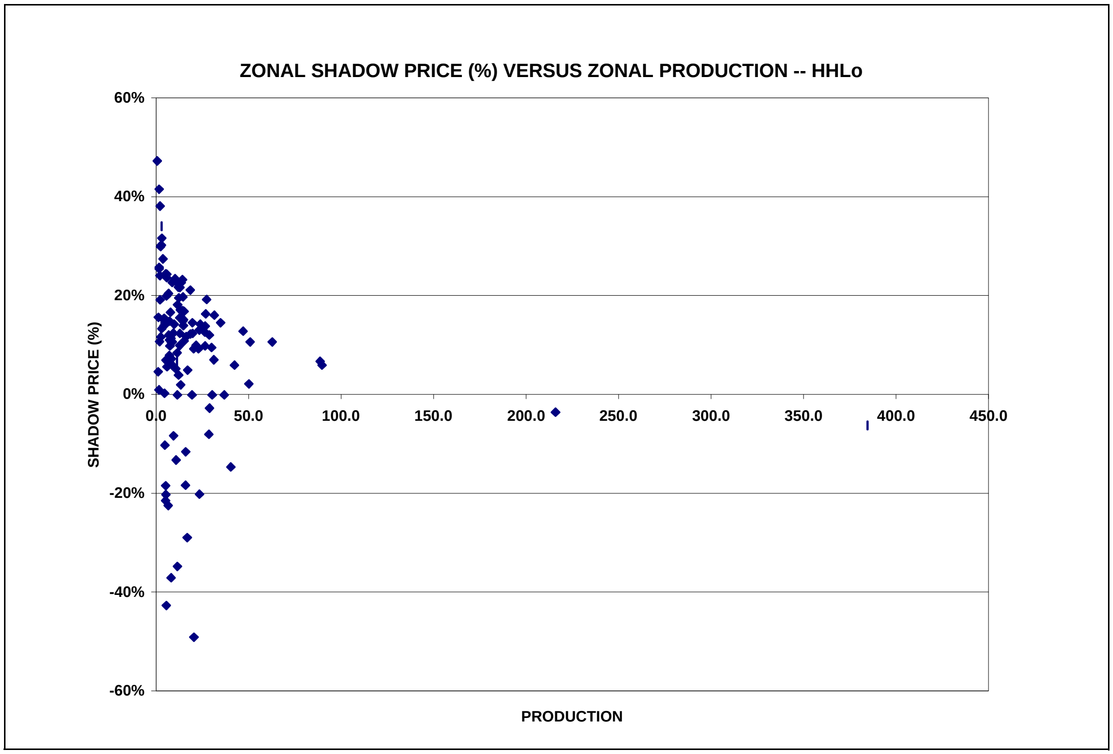

Figure 6: Scatterplot of HHLo price adjustments

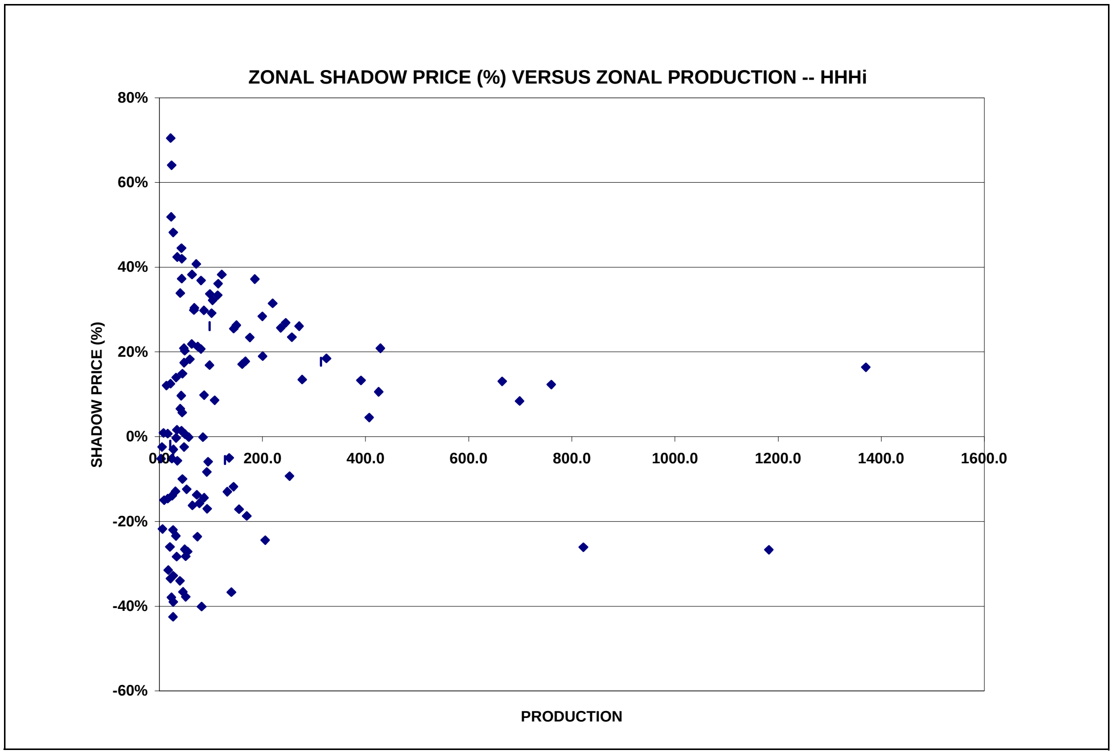

Figure 9: Scatterplot of HHHi price adjustments

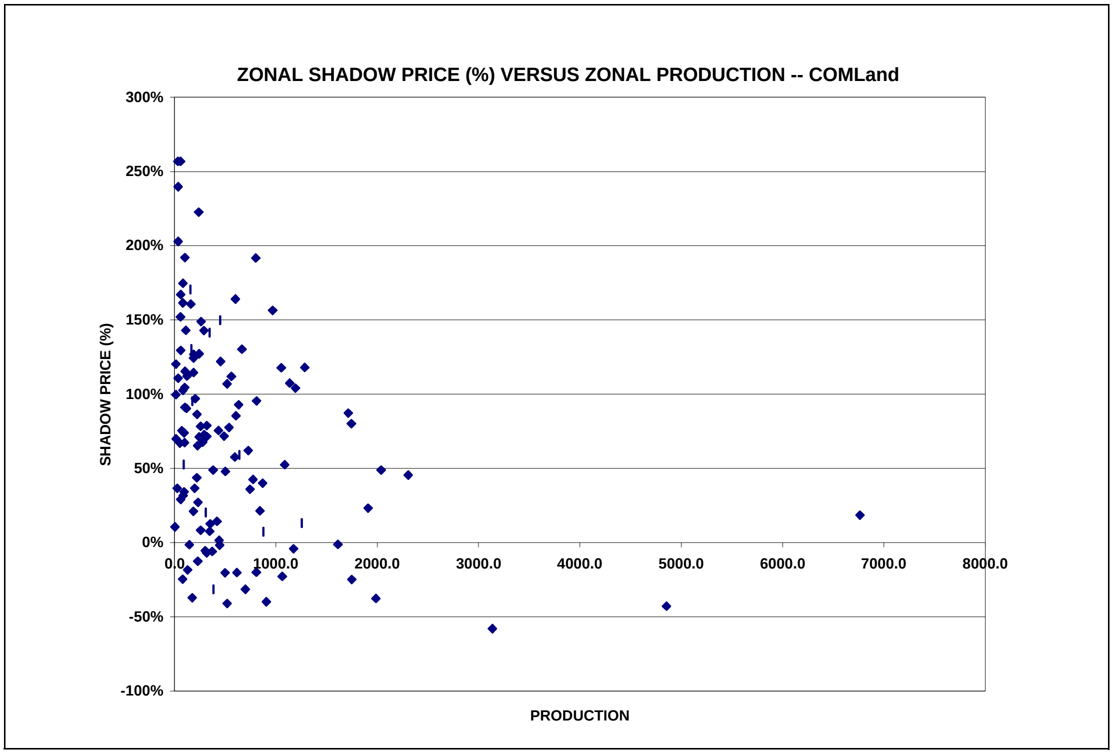

Figure 7: Scatterplot of COMLand price adjustments

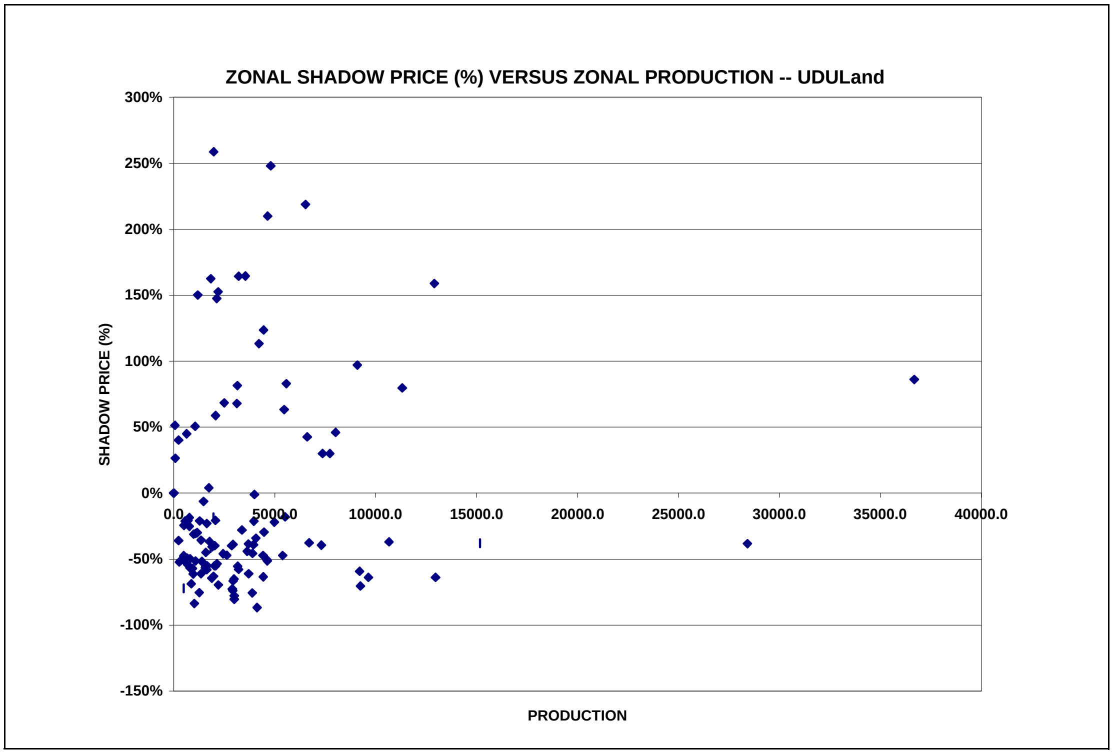

Figure 11: Scatterplot of ULULand price adjustments

Whereas shadow price adjustments rarely exceed the 20% criteria for industry and household sectors, land sectors are another matter. Relatively large adjustments have been necessary to reproduce base year land utilization patterns as shown in Figures 10 and 11 related to commercial land (COMLand) and urban dwelling unit land (UDULand), respectively. There are several possible explanations for this result, including:

- There are only four categories of land spanning a very wide range of densities across the state leading to high potential for aggregation error,
- The current model design does not distinguish between blue-collar and white-collar demands for land which in effect lumps farms, factories, and offices in the same category,
- The land quantity and price data are not nearly as robust as other zonal data leading to higher expected margins of error,
- There are numerous important factors affecting land prices which are not reflected in the current statewide model.

The high land-price adjustments found to be necessary for the prototype statewide model indicate this is an area that needs careful rethinking in ongoing efforts.

4.1.2 Base Year Transport Model Calibration

Before results of the activity model can be scrutinized too closely, economic flows from the activity model must be translated to transport demand categories, the transport model run for at least several iterations, and the results checked for reasonableness and for consistency with calibration targets. Many rounds of feedback and adjustment were necessary within and between the activity and transport models before results begin to fall into line. Initial runs of the activity model were made based on free-flow times and costs from the transport model. Early runs of the transport model resulted in trips lengths and demand levels grossly in excess of calibration targets and network capacities. It was found necessary to drastically modify the starting values of the activity model parameters which is not surprising since the starting values were simply “guesstimates.” Data on zone-to-zone economic flows that might be used to rigorously estimate activity model parameters are not available. In addition, there is little experience elsewhere to draw upon. The circumstances dictated a trial-and-error approach. In the process it was found necessary to make several modifications to the TRANUS software as well as to model parameters. Modifications to initial trip generation functions were also necessary to match the targets for number of trips by category and average trip lengths.

Base year calibration of the transport model involved relatively familiar targets and criteria. Primary criteria for base year calibration of the transport model were:

- Matching target total trips and average trip length for each transport demand category within +/- 20%
- Matching target passenger trips by sub-state area within +/- 20%
- Matching average weekday counts along major intercity corridors within +/- 20%

Table 22: Substate regions for survey expansion

Designation	PUMAS in PUMAGRP	Corresponding Survey Areas
RURAL EAST	0100, 0200, 0300	Deschutes, Klamath, Malhuer, Umatilla Surveys
RURAL WEST	0400, 0500, 0600	Clatsop, Coos, Josephine, Lincoln Surveys
LCOG/RVCOG	0700, 0800, 0900	Eugene/Lane County Survey
MWVCOG	1000, 1100	Salem & 3-County Surveys
METRO	1200, 1300, 1400, 1500	Metro Survey

Calibration targets for 1990 passenger trips were developed from results of the household surveys conducted by ODOT throughout the state. Survey trip generation rates and average trip lengths by trip category, cross-classified by household size and income, were derived from the household survey data for each survey area. The state was divided into regions that would enable expansion of the survey data to represent the entire state. Regions were based on 1990 Census PUMAS and the distribution of household surveys as indicated in Table 22. Census PUMS data for 1990 was used to derive estimates of households for each sub-state region. Estimates of households for each region were cross-classified by household size and income categories corresponding to those used for tabulation of household trip generation rates and average trip length by trip category. These data were used to estimate the total trips and weighted average trip lengths by trip category for the entire state providing targets against which TRANUS model results could be compared.

Note that the target number of trips and average trip lengths developed from the survey data exclude intrazonal trips. This is consistent with the focus of the model on intercity trips but it is also essential to exclude intrazonal trips since TRANUS ignores intrazonal flows.

An overall comparison of transport model results to targets is given in Table 23. This comparison is for the entire model area including Clark County. The total number of trips by transport category produced by the model is very close to the target values. Average trip lengths in miles produced by the model are higher than target values for all categories. Model trips lengths are as much as 45 percent higher than the target values; however, there is reason to believe the target values are low. The model is focused on longer, intercity trips which are made infrequently com-

Table 23: Overall comparison of transport model results

Transport category	Zone-to-zone trips			Average trip miles		
	Target	Model	% difference	Target	Model	% difference
CmuteLo	129,531	127,244	-1.8%	12.9	14.5	12.3%
CmuteMid	436,076	433,088	-0.7%	14.1	17.2	22.1%
CmuteHi	210,485	209,364	-0.5%	15.7	18.6	18.3%
Recreation	457,644	453,623	-0.9%	17.8	24.6	38.1%
HBOther	1,577,152	1,564,168	-0.8%	16.4	21.7	32.1%
NHBWrk	433,584	445,496	2.7%	15.5	20.3	30.8%
NHBOth	768,171	779,190	1.4%	18.1	26.2	44.8%
Total	4,012,643	4,012,173	0.0%	16.4	21.9	33.3%

pared to other household trips. Traditionally, such trips are subject to higher than average under-reporting in household surveys. In addition, such trips are subject to higher coding errors or incomplete coding where distant destinations are involved. Such problems could lead to significant error in the target average trip lengths. The overestimate of trip lengths by the model is of less concern for these reasons; however, it does highlight what has been quite problematic in the calibration of the current statewide model.

A more detailed comparison of transport model results is presented in Table 24. This table compares estimated trips by category and sub-state region to target values.

This comparison indicates substantial error in the distribution of trips throughout the state. The model substantially overestimates trips generated by the Portland Metro area and substantially underestimates trips elsewhere. This result is related to the trip length overestimation problem identified above which in turn is closely related to the results of the activity model. Since there are

Table 24: Comparison of trips by substate region

	Category	Metro	MWV	L/RV	RurEast	RurWest	Total
Target passenger trips by substate area	CmuteLo	75,328	25,366	17,867	7,343	3,626	129,531
	CmuteMid	240,328	97,095	59,360	18,562	20,730	436,074
	CmuteHi	111,164	51,901	25,836	13,029	8,556	210,485
	Recreation	261,749	80,837	63,818	14,272	36,967	457,644
	HBOther	806,643	307,533	224,409	60,549	178,020	1,577,153
	NHBWrk	262,775	76,912	68,034	13,465	12,398	433,584
	NHBOth	390,226	110,950	114,242	24,330	128,421	768,170
	Total	2,148,214	750,593	573,566	151,550	388,718	4,012,641
Model passenger trips by substate area	CmuteLo	95,654	9,082	18,506	3,123	879	127,244
	CmuteMid	325,858	31,632	56,915	14,377	4,306	433,088
	CmuteHi	165,501	12,156	24,601	5,226	1,880	209,364
	Recreation	357,665	45,029	27,493	21,726	1,710	453,623
	HBOther	1,228,750	132,711	139,468	54,243	8,996	1,564,168
	NHBWrk	383,750	20,819	29,219	8,751	2,957	445,496
	NHBOth	610,454	77,559	47,471	38,987	4,719	779,190
	Total	3,167,632	328,988	343,673	146,433	25,447	4,012,173
Percent difference	CmuteLo	27%	-64%	4%	-57%	-76%	-2%
	CmuteMid	36%	-67%	-4%	-23%	-79%	-1%
	CmuteHi	49%	-77%	-5%	-60%	-78%	-1%
	Recreation	37%	-44%	-57%	52%	-95%	-1%
	HBOther	52%	-57%	-38%	-10%	-95%	-1%
	NHBWrk	46%	-73%	-57%	-35%	-76%	3%
	NHBOth	56%	-30%	-58%	60%	-96%	1%
	Total	47%	-56%	-40%	-3%	-93%	0%

no observed zone-to-zone economic flows to which activity model output can be directly compared, the primary check on the activity model is through the related transport flows. The combination of errors in trip lengths and trip distribution indicated by Tables 23 and 24 point to problems in the economic flows produced by the activity model. These errors are also influenced by the trip generation rates and elasticities of the transport model. It is clear that zone system design and the widely disparate zone sizes also contribute to the problem. Further, the use of scaled utilities by TRANUS in both the activity and transport models clouds interpretation of model behavior.

The current results are far better than earlier results and much has been learned in the arduous course of getting to this point. While some further improvement to model results is no doubt possible, the ultimate solution is to provide greater calibration flexibility in the design and implementation of the model. Some modifications have been made to TRANUS to facilitate calibration and Modelistica has been helpful in this regard as well as in providing technical assistance throughout the calibration process. The modifications required to deal with an application as large and diverse as the Oregon statewide application are more far reaching. Substantially greater flexibility is required in model structure and formulation and there is a need for flexibility in parameter specification on a geographic basis to recognize fundamentally different area types.

One last comparison is presented for the base year transport calibration. Figure 12 presents daily vehicle traffic flows produced by the statewide model compared to ADT flows from traffic counts. The flows included in Figure 12 are concentrated in the upper Willamette Valley with each point representing a specific location. In spite of the weaknesses in the prototype model noted above and the coarseness of the statewide network, the model does a reasonably good job of reproducing observed flows in this area. Note also that these results do not reflect any network-level calibration or any post-processing of results. A regression of model versus count flows yields an r^2 value of 0.87 indicating a very strong relationship.

4.2 1990 - 1995 Calibration/Validation (Cycle 2)

Upon satisfactory completion of the base year calibration, focus turned to application of the model for the 1990 to 1995 increment of time and the adjustment of model parameters as necessary. Global increments of exogenous production were input, the activity and transport models run, and the model results compared to target values. Primary criteria for evaluation of the model were:

- Closely matching the actual increments of change in households and employment by sector for each zone
- Matching target passenger trips by sub-state area within +/- 20%
- Matching the expected change in passenger trips by sub-state area within +/- 20%
- Matching 1995 truck and auto average weekday counts along major intercity corridors within +/- 20%
- Matching the 1990 to 1995 change in truck and auto average weekday counts along major intercity corridors within +/- 20%

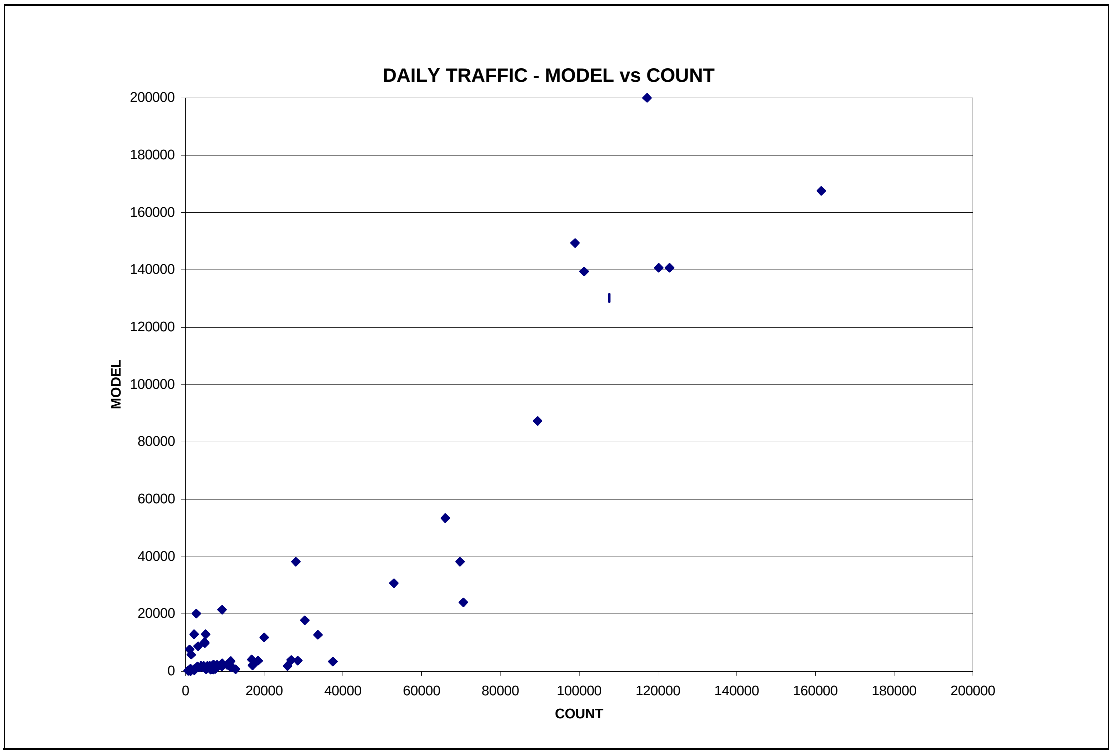

Figure 12: Modeled traffic flows versus count ADT

Where necessary, zonal attractor values were adjusted to bring model allocations by zone into line with observed values. Refinements were also made to other model parameters where necessary to satisfy both the 1990 and the 1990-1995 calibration criteria.

4.3 Long-Range Application and Sensitivity Analysis (Cycle 3)

Once agreement is reached on the most appropriate structure of the land market for the statewide application, the necessary input data and relationships will be assembled and entered into TRANUS. The base year (1990) and 1990-1995 calibrations will then be revisited incorporating the agreed land market simulation. In addition to maintaining the criteria previously specified, the model will be calibrated to satisfy additional criteria related to the land market, namely:

- Matching the target amount of land consumed by type, by zone
- Matching the target prices for land plus improvements by type, by zone
- Matching the increments of change in land area consumed and related prices by type, by zone

Consistent with the greater variability inherent in land market characteristics, problems of data availability and consistency, and the resulting wider margins of confidence in target values, criteria related to the land market will be somewhat less ambitious than those discussed earlier for the activity and transport models.

Appendix A: Overview of the TRANUS Modeling System

TRANUS is an integrated land use and transport model which can be applied at an urban or regional scale. The program suite has a double purpose: firstly, the simulation of the probable effects of applying particular land use and transport policies and projects, and secondly the evaluation of these effects from social, economic, financial and energy points of view.

The advantages of integrating land use and transportation are well known and have been documented extensively in the literature. However, there are very few such integrated modeling packages in practice. For the transport planner, land use and transport integration provides a means of making medium and long-term demand estimates which are impossible with transport-only models where demand is a given input. The integrated approach can also be very useful as an alternative method for constructing realistic estimates of origin-destination matrices; large surveys can be very expensive, and even with a generous sample size it is very difficult to obtain a complete estimate of the matrices. The alternative is to carry out a much smaller survey and use the data to calibrate an integrated land use and transport model, obtaining complete and realistic matrices.

For the land use planner, whether at an urban or regional scale, integrated modeling makes possible an assessment of the implications of transport policies on the location and interaction of activities. But it is in consistent land use and transport planning that the TRANUS system shows its full potential. However, it is possible to apply TRANUS as a stand alone transport model from given data about demand, should this be required for short term policy appraisal.

The TRANUS system has been developed since 1982 by MODELISTICA, and has been applied to many real cases and research projects. Through the years several versions have been released, incorporating many improvements as a result of theoretical developments and practical requirements. The versions released have been widely applied in practical applications, and in research and teaching activities. The current version 4.4 belongs to a new series which incorporates a large number of innovations and extended facilities for the user.

TRANUS was developed for microcomputers, and in this sense it pioneered the area. The use of microcomputers facilitates the application of the system and considerably reduces the cost of a study. Only a few years ago it was the only integrated package that could run efficiently in very small machines. Nowadays, with the increasing capacity and reduced cost of personal computers, these advantages are more evident. But the fact that the internal structure of the programs was designed for small PC´s, a fact that required a strict discipline for making the most efficient use of limited resources, allows TRANUS to operate smoothly in most types of equipment.

The programs are provided with this Reference Manual. All the programs have been designed and developed entirely by MODELISTICA, and are the intellectual property of MODELISTICA's directors Beatriz Pérez, Tomás de la Barra and Juancarlo Añez. The name TRANUS is a registered trademark. Spanish and English versions are maintained. A special agreement has been made with Rickaby Thompson Associates in Milton Keynes, UK who are MODELISTICA's European agents for distribution of the package.

A.1 The theoretical framework of TRANUS

In the last two decades, important theoretical and empirical developments have been achieved in the field of activity location and transport analysis. Many of these achievements have occurred in the area of systems analysis, based on a structural and quantitative analysis to explain urban and regional phenomena. The list of authors and corresponding literature is too vast to describe here, and only the most significant ones will be mentioned.

The origins of spatial analysis go back as far as Von Thünen in 1826, but in recent times the work of Hansen (1959) and Lowry (1964) can be considered as starting points. The first generation of gravity-type models of the 1960s, and the first operational transport models, were followed by the important work of Wilson (1970), which not only introduced entropy-maximizing techniques to spatial modeling, but also led the way towards integrated land use and transport analysis. Wilson showed that land use and transport could be represented by means of a common theoretical framework. The work of Echenique et al (1977) introduced micro-economic principles and input-output techniques, and a achieved a good degree of implementation. In input-output analysis, the work of Leontief and Strout (1966) is of major importance, both at national and regional scales.

TRANUS is rooted in this tradition, but also draws heavily on the work of Domencich and McFadden (1975) on discrete choice analysis and random utility theory. Although these authors did propose a general model, most of their work and the work of their followers is centered on the problem of modal choice in transport, and no specific models are proposed for the rest of the elements of the urban or regional system. In TRANUS, this theoretical backbone has been extended to all decision levels, from modal split to assignment, trip generation and the location of activities. A detailed explanation and a good complement to this description of the theory underlying the TRANUS system can be found in De la Barra (Integrated Land Use and Transport Modelling, Cambridge University Press, 1989).

A.1.1 Individual and aggregated choices

In general terms, decision theory describes social processes as a set of decisions made by individuals. The main assumption is that individuals choose rationally between the options available to them. Each individual, faced with a number of options, will rank them according to the degree of satisfaction or utility perceived in each case, and will choose the one which provides the greatest utility. Utility, on the other hand, is a subjective concept - its perception will vary from one individual to another, and from one choice to another.

Mathematically, utility can be represented as a utility function for a particular individual, which contains variables describing measurable attributes of each option. Faced with a particular set of options, an individual may be assumed to evaluate each one with the same utility function, and will choose the option which yields the greatest utility.

This is the basis of micro-economic theory. However, for the urban analyst it is of little practical value, since it would be impossible to keep track of utility functions for each individual living in a city or region, and because the number of options can be very large. There is, then, the need for aggregation. Individuals may be grouped according to common socio-economic characteristics, and options into groups of similar types. Spatial aggregation is important: point location of individuals or firms must be replaced by location in larger discrete areas or zones.

Aggregation introduces sources of variability, because individuals within a group are different and perceive utility in different ways. The same can be said about aggregated options and zones. Naturally, if groups are small, variations will be small also.

In order to solve this problem, random utility theory assumes that, since we have no grounds for choosing a particular distribution of the way utility is perceived, all we can do is to assume that it is randomly distributed. The same will apply to options and spatial location. Utility functions will no longer apply to a particular individual: instead, they will apply to a population of individuals, related to groups of options. A population-related utility function must not only contain the aggregate measurable characteristics of each option group, but also random elements. Alternatively, we might assume that the utility function itself is random.

In the individual case, the utility function is deterministic and produces a unique result: the selection of a particular option, i.e. the one of greatest utility. In the aggregate case, since there are random elements, utility functions are probabilistic, producing a distribution of individual choices among the available groups of options.

The perception of utility within a group will have a particular distribution, which will be the same for all option groups, giving rise to a joint distribution. Mathematically, the probabilistic model is obtained by integrating the joint distribution. Hence, several models may be derived from the general one, according to the particular shape of the distribution. Domencich and McFadden (1975) explored several possible shapes, showing that the most appropriate was the Weibull distribution, which after integration yields a Multinomial Logit Model (MNL), because of its simplicity and superior statistical properties.

An interesting corollary to this, described by Williams (1977), is that if MNL is the chosen model, then there is one and only one way of measuring the average utility of the population, which is the logarithmic average of the distribution, also called composite cost. Furthermore, if such a model is applied in the context of two different scenarios of future conditions, the difference in utility will be equivalent to the consumers' surplus in traditional economic theory. In TRANUS, this general formulation has been improved in several ways, and as a result, the whole structure of the models can be described as a large chain of nested multinomial logit models with an improved formulation.

A.1.2 Decision chains and hierarchies

The paragraphs above have dealt with one particular choice situation. In an urban or regional system, however, a long and complex decision chain may be established. For example, a typical chain would be:

place of work --> residence --> shopping --> transport mode

Each link along the chain is clearly conditioned by the preceding link. For instance, where to go shopping is a decision conditioned by the place of residence. In order to represent such a decision chain in a set of models, each component must precede the next in the correct order. If each link along the chain has a corresponding model, producing probabilities such as P(w), P(r), P(s) and P(m) in the example above, the number of people traveling by bus mode from an origin to a destination and who work in a particular zone could be calculated as the number of people that work in that zone by P(w) x P(r) x P(s) x P(bus). This is a comfortable solution, because it is possible to model each link in the decision chain separately, thus avoiding very large computations.

However, the problem is more complex than this, because each link in the chain may influence the preceding ones. In the above example, it could well be that people decide where to go shopping precisely because there is a good bus service. Thus, the choice of transport mode affects the choice of shopping. Similarly, the choice of residential location may have been influenced by the availability of local shopping facilities.

In order to accommodate this, the process of calculation must begin from the other end, that is, from the last link in the chain, proceeding backwards. In the example, we would have to calculate the overall availability of transport from residence to shopping. The way to do this is by calculating the corresponding composite costs, producing the overall aggregate utility of all shopping facilities around each residential area. In this way the transport element has been transferred into the shopping utility, and eventually into the residential utility. On reaching the top of the decision chain, the model must reverse direction and proceed again in the original direction, calculating the probabilities. This way of calculating composite costs and models is known as nested multinomial logit model (NMLM).

If it was not for variable costs and elasticities, the calculation process would end here. In the above example, if the bus service is used beyond its capacity, the cost of travel (or time) may increase until residents eventually choose other options. This effect may be further transferred to the residential choice. Thus the calculation process becomes iterative, aggregating utilities and estimating probabilities backwards and forwards several times until a state of equilibrium is reached. Demand elasticities also influence the process; in the example, if bus services to shopping facilities become congested, some people may travel less, perhaps once each week instead of once each day, thus generating fewer trips.

Sometimes a particular choice is divided into sub-choices, and this gives rise to the concept of hierarchies in the decision-making process. For example, residential choice might be divided into two hierarchical levels: first, the choice of a district within the city, and then the choice of a neighborhood within a district. Similarly, a hierarchy of transport modes can be established: first the choice of public rather than private transport, and then the choice of buses rather than other options such as trains or the subway. These hierarchical levels are related to each other in the same way as different links in the decision chain, and the computing process is identical.

The concept of hierarchical analysis is a powerful one. The use of hierarchies permits the representation of some parts of the urban system in great detail, while other parts are dealt with in less detail. For instance, it may be of interest to investigate residential location within a particular district: in this case the residential model can determine the probable location of residences in all districts at a first hierarchical level, and apply a second level of analysis only to the district in question. It is important, however, that these levels of analysis are not confused. In the literature, the use of hierarchies is a solution to a well known problem related to MNL models, and called ‘the independence of irrelevant alternative options’.

A.2 The mathematical and algorithmic structure

In the following sections the general structure of the land use and transport model is described. This is followed by sections describing each one of the main components: activity location and interaction, an interface module, the transport system, and the evaluation procedure.

A.2.1 General structure of the model

An explicit dynamic structure relates the two main components of TRANUS: land use and transport. The way in which land use relates to transport through time is shown in Figure 1, where discrete time intervals are represented as t1, t2, t3, and so on.

In general terms, the land use and transport systems influence each other trough time. Economic activities in space interact with each other, generating flows; these flows determine transport demand within the same time period, and are assigned to the transport supply in the transport system. In turn, the supply-demand equilibrium at a transport level determines accessibility, which is fed back to the land use system, influencing the location and flows between activities. This feedback, however, does not occur instantly in the same time period, but is lagged; hence, transport accessibilities in period t1 affect the distribution of flows in the following period t2. Since there are also elements of inertia in land use from one period to the next, the effects of changes in transport might well take several periods to consolidate.

As a result, a change in the transport system, such as a new road, a mass transit system or changes in tariffs, will have an immediate effect on travel demand, but will only affect activity location and interaction in the following time period. Changes in land use, on the other hand, such as growth in the production of particular economic sectors, a new supply of land, buildings or investment, will result in modified interactions and in turn in changes in the transport demand within the same time period.

A.2.2 Activity location and interaction: the land use model

The land use model is basically a spatial input-output model with a very general formulation. Within this structure the user may define a complex model of a region with a detailed representation of the economic and social system, or a very simple one with only the main elements represented. The analyst may determine the structure of the model he thinks most suitable to the emphasis and purpose of the study.

In order to define the system, the study area must be divided into zones, which may be internal or external; internal zones, in turn, may be defined within a two level hierarchy. Then the economy of the area must be divided into sectors; these may correspond to economic categories (agriculture, mining, industry, services, etc.), to factors (labor, capital), or to physical elements (land or floorspace), and may be represented in different units (money, physical units, jobs, people, etc.). Each sector in the model has a number of associated attributes, the main ones being production (exogenous or induced), production cost, demand, consumption cost, value added, restrictions, and equilibrium price

Production in one sector and zone requires inputs from other sectors and zones, and sectors are related to each other through demand functions which vary according to the cost of the inputs plus the cost of transporting them from the production zone to the consumption zone. This is a very broad definition, and in fact it allows for the representation of economic transactions, generation of employment, demand for services, and the consumption of land and floorspace.

The results of these transactions are flows, represented in origin-destination matrices, which are later transformed into transport demand. For example, the generation of employment from production may lead to trips to work, or the consumption of one productive sector by another may lead to movements of commodities..

The model keeps track of the costs involved in the transactions. The cost of producing one unit of a sector is the sum of the cost of all inputs, plus the cost of transport and any possible value added (including taxes and subsidies). If there are, however, restrictions on the amount of production which may take place in a particular zone and sector, the model simulates an equilibrium price. Typically land and floorspace are restricted to existing stock, and in this case equilibrium prices will represent land values or rents.

The model also makes an explicit representation of imports and exports. These may only occur between internal and external zones. Exports represent demands from outside the study area and generate induced production. By contrast, imports represent demands within the study area which originate from the external zones, and thus generate no induced production.

The spatial distribution of inputs from demand zones is determined by a two-level hierarchical MNL model, in which the utility function includes transport disutilities and consumption costs. The analyst may add to the model a set of attractor functions, which may include ‘modeled variables’, that is, variables which have already been included in the model definition, or exogenous variables to account for unexplained elements which may affect the distributions (typically environmental elements).

The spatial input-output model is complemented by an incremental model. From one time period to the next, the analyst may define changes to the main variables of the model, such as changes in the exogenous production and consumption, imports and exports, and restrictions. The user may specify a global increment (positive or negative) for any variable, together with a distribution function which will allocate the increments to zones, but he may also specify particular changes in individual zones.

The logical sequence of calculations for the activity location model is shown in Figure 2. The model begins by reading in a set of user defined parameters and previous time period land use results. The model will check that the values are consistent, and that it is plausible to honor the constraints as specified.

Next, the incremental models are applied to the data. Studywide increments are allocated to zones according to the distribution functions, and zone specific increments are added. This results in a current time data set which governs the rest of the calculations, and allows for the calculation of the attractor variables which will influence the spatial distributions in the main input-output model. The attractors are calculated from the user specified attractor functions.

At this point an iterative process is started. Each loop begins by calculating the demand for inputs required to achieve the production allocated in the previous iteration. This is done by considering the demand functions, which may be elastic with respect to consumption costs. Once total demand for inputs has been determined for each demand zone and sector, it is distributed to production zones and sectors, according to the utility function and the attractors imbedded in the MNL model. It is important to keep in mind that the distribution process must take into account the possibility of imports from external zones. Once the distribution has been determined, it is then possible to calculate production costs as the sum of the cost of the inputs, plus the transport costs, plus possible value added costs.

The next step is to check whether total production by zone and sector satisfies possible constraints. If total production in a constrained sector and zone is greater than the maximum, an equilibrium price is increased. Otherwise, an equilibrium price is decreased. It follows that if a particular zone and sector has not been constrained, the equilibrium price will be equal to the consumption cost.

At the beginning of the first iteration only exogenous production (final demand plus exports) will be present, and at the end there will be a certain amount of induced production, that is, the production that is necessary to satisfy the demands from the start up production. In the second iteration, induced production from the final demand plus that of the first iteration must be determined and allocated. This process continues in subsequent iterations, but in each one smaller increments will be required. The iterative process continues until the increments become sufficiently small that the system has acceptably converged.

In most applications production converges quite quickly, and if no constraints have been specified, prices should converge in resonance. However, if constraints have been specified, the system must perform further iterations in order to guarantee that production is reasonably equal to the constraints, the definition of reasonable being determined by a user specified convergence criterion.

Once convergence has been achieved, the model is ready to output the results. These will consist of a set of flow matrices, i.e. the results of the distribution model, and the final characteristics of each sector: production, consumption and prices.

If very simple applications are being considered, the model described above may seem too complicated. Most of what have been described are possibilities offered to the ambitious user, but if that are not required the model may be “stripped down” to a very straightforward definition and many of the concepts described above may be ignored. The developers of the model hope that very simple applications may be made with ease, but they also hope that the model may cater for very demanding applications.

A.2.3 Activity location/transport interface

As a result of the activity location process, a set of matrices of flows is produced, from which potential transport demand may be derived. This is the purpose of the activity location/transport interface, a set of programs which act as intermediaries. This can be a very simple process; for example, if a typical input-output transaction has been simulated, resulting in a flow of money units from origin zones to destination zones, multiplying by a value-to-volume ratio would produce the weight of commodities which would require transport.

But there might be other elements to take into consideration. It is quite common that the timescale of the allocation model is different from that of the transport model; for instance, production in a typical input-output model is represented in millions of monetary units per year, but in the transport model it is probably more comfortable to work in tons per day, so the activity location/transport interface must be able to account for these timescale transformations.

Furthermore, the user might wish to work with different land use and transport categories. For instance, a sophisticated activity system might include many economic sectors such as agriculture, mining, many different types of manufacturing industries and several tertiary activities, but in terms of transport it might be preferable to speak in terms of light, heavy, and refrigerated cargo, or of containers and liquid bulks. The interface allows for such specific transformations; land use categories need not match transport categories, and they may be transformed from one to the other with predefined functions.

In typical urban relationships there are further details to consider. Take, for instance, a flow matrix which relates employment to residence; these flows must be transformed into potential home-work flows as well as the return work-home flows, if the transport model is to work with daily trips.

Finally, the analyst might wish to introduce exogenous transport demand which does not correspond to the modeled system. This is very common when through traffic is relevant in the study area. The interface module provides a simple mechanism for introducing exogenous trips by transport category, and even by mode for specific origin-destination zones. It allows for unrelated land use flows to be introduced; if taken to the extreme, all of the transport demand may be defined as unrelated to land use interactions, and in this case the TRANUS system becomes a transport only model.

This description of the interface component of the TRANUS system has explained the transformations that are made possible from land use categories to transport ones. But once the transport module has performed its functions, opposite transformations must be made of transport costs and disutilities into relevant categories for activity interactions. The reverse transformations are performed automatically by the interface module, without the need for further specifications by the analyst. The program simply reverses the logic and performs the reverse timescale transformations.

A.2.4 Transport calculations

The activity location/transport interface will supply the transport model with a set of matrices of flows representing potential transport demand, and possible exogenous trips. The purpose of the transport model is to transform potential demand into actual trips, and to assign them to the many transport supply options. The calculation sequence of the transport model is shown in Figure 3.

The first task of the transport model is to search for a set of paths connecting each origin to each destination by each transport mode. Each mode, in turn, may be divided between several transport operators, and a trip maker may freely transfer from one operator to another within a mode (these transfers are transshipments in the case of commodities). For example, three modes might be defined: cargo, public transport and private transport; in turn, cargo might be divided into light and heavy trucks and freight trains, and public transport might be divided into buses, metro and passenger trains. In this example, commodities may transfer from light trucks to trains and back to heavy trucks, or any other combination; public transport passengers may take a bus to a metro station and then take the metro to a passenger train. Modern transport systems increasingly provide opportunities of this kind.

In order to determine the paths, the transport model must analyze the transport network, defined as a set of interconnected directional links in the form of a graph. Each link is described in terms of its main characteristics: link type, link length, physical capacity, capacity of each public transport operator and prohibited turns. Link type, in turn, defines a number of general characteristics which are common to all links of the same type, such as speed, circulation charges (tolls), maintenance costs and the transport administrator in charge of the link.

Each path is represented as a sequence of operator and link pairs which connect an origin centroid to a destination one. When along a path there is a change of operator, a transfer is assumed, adding a transfer cost and a waiting time to the cost of the path. The algorithm may turn out that two paths follow exactly the same sequence of physical links but make use of different operators, (perhaps minibuses instead of buses or walking). The algorithm also allows for the representation of integrated fares.

The method of finding paths is unique to the TRANUS system and has been called multidimensional path search. It takes as a basis Dijkstra's (1959) original method for minimum path search which avoids path enumeration, adding to it several dimensions to account for a number of transport specific characteristics. The algorithm may be viewed as a model of the options available to users when traveling from an origin to a destination.

The multi-path search is subject to two parametric restrictions: a maximum number of paths and a dispersion factor with respect to the minimum. If for a given mode the user has defined a maximum number of paths = 5 with a dispersion factor = 1.3, the algorithm will search for up to 5 paths provided that these paths are of a generalized cost less than 1.3 times the cost of the minimum path. Generalized cost includes the out-of-pocket costs (tariffs), plus the value of time, plus operator dependent penalties (equivalent to mode specific constants). The dispersion factor is also called overlapping factor, because it controls the degree of overlapping among competing paths. The result of this process is a set of distinct travel options, as will be seen below.

The representation of the network has been modified to simplify the coding of prohibited turns. In many other algorithms, nodes represent real street intersections and links represent road sections. This method of describing a network is maintained only for the purpose of coding link data. The program, however, transforms it into a dual graph representation, where the vertices of the graph represent road sections and links represent permitted connections. This results in a simple way of coding turn prohibitions, and avoids the traditional method of introducing fictitious nodes and links. It is thus more economical in computing terms, and produces a clean output free of unnecessary links which do not exist in reality. The dual graph representation is completely transparent to the user, because when the path search process is finished, the program translates the coding of nodes and links back to the original representation.

Finally, a new feature called overlapping control has been added to the algorithm. This feature solves at a network level the well known problem of the independence of irrelevant alternatives in MNL models. The method actually measures the degree of overlapping between competing paths and transfers this information to the assignment algorithm (which is described below) in order to find different options and correct the probability of each path been chosen.

The resulting algorithm is very powerful indeed, because it takes into account most of the features which are unique to transport systems. It can be shown that the algorithm will produce realistic results both at a regional or urban level, which is not the case in some other methods such as incremental or equilibrium assignment. The path search algorithm, together with the assignment model described below, is fully consistent with random utility theory, while many other methods are not. The method also facilitates the task of coding and checking the network, which is often one of the most time consuming tasks in model calibration; special facilities are provided for inspecting the resulting paths. The method also guarantees that if the network is symmetrical the resulting paths will also be symmetrical.

Once all paths have been found, the process of calculation enters an iterative cycle. It begins by calculating both monetary and generalized costs along each path. Generalized costs add to the monetary costs the value of travel and waiting times and penalty factors by operator. Penalty factors represent subjective elements that affect each operator, such as comfort, reliability, and so on, which are not included in any of the other variables. Monetary costs will not change from one iteration to the next, but generalized costs must be recalculated to account for changes in travel and waiting times due to congestion.

A weighted arithmetic average cost over all paths is calculated for monetary costs, but composite costs are aggregated from a path level to a mode level through a logarithmic average. Similarly, costs are aggregated over all modes to obtain the average monetary and composite cost of travel from an origin to a destination for a given user category.

The next step in the iterative process is trip generation, which transforms the potential travel demand calculated by the activity/transport interface into actual trips at particular time of the day (peak hour, 24 hours, etc.). Trip generation determines the number of trips from an origin to a destination by a particular transport category, as an elastic function of the corresponding composite cost. Elasticity in travel demand means that for a given functional relationship, more trips will be made if there is a reduction in the cost of travel. For example, if a certain number of people live close to their place of work, they might travel more times in a day because, perhaps, they go home for lunch. If a family lives close to shopping facilities, they might travel several times a day to do their shopping, while if they live in a more isolated location, they might go only once a week. In peak hour analysis, it may be that heavy congestion forces people to travel before or after peak hour. As a result, in each iteration, the number of trips will be reduced as congestion builds up. Elastic trip generation also means that if a new transport facility is introduced a certain amount of induced demand will be generated.

Trips for each category are distributed to modes by means of a MNL logit model in which the utility function is determined by the composite cost of travel by mode. As with all MNL models in TRANUS, diminishing marginal perception of utilities is taken into account. Modal choice is made over all modes available to each category. For instance, a trip category representing commodities will only choose from cargo modes, such as trucks, ships, and so on. Modal choice for passenger categories is restricted by car availability, which is explicitly built into the modal split model.

Trips by mode must be assigned to the different paths connecting origins to destinations by that mode. Since each path implies a particular sequence of operators and transfers, trips are simultaneously assigned to operators, as well as to links of the network. This is carried out by means of another MNL model which includes both diminishing marginal perception of utilities and overlapping, and where the utility function is determined by the composite cost of each path. The combination of the MNL modal split and assignment models is equivalent to the two-level hierarchical modal split model commonly found in the literature. Both models are nested through composite costs.

By applying vehicle occupancy rates, trips are transformed into vehicles by operator in each link of the network. In the case of public transport operators, vehicles subject to fixed routes are assigned directly to the network; alternatively, the user may ask the model to determine the number of vehicles as a function of demand. Finally, the number of vehicles by operator are transformed into standard vehicles by applying appropriate rates.

The model can check for the possibility of vehicles returning empty. A parameter called percentage of consolidation controls the maximum percentage of return vehicles which may attract cargo on their way back. If, however, very asymmetric flows are present in the system a large proportion of vehicles might have to return empty.

In the final stage of the iterative process a capacity restriction procedure is performed, whereby travel speeds are reduced and waiting times are increased in every link for each operator as a function of demand/capacity ratios. Waiting times take into consideration the frequency of transit services as well as possible delays due to congestion and the demand/capacity ratio in the vehicles themselves. The latter means that people waiting for buses will have to wait longer if the vehicles are too full.

At this point the convergence of the model is estimated as the change in volumes, speeds and waiting times by link. The user specifies a convergence criterion, and the model is said to have converged if in the current iteration volumes, speeds and waiting times have changed with respect to the previous iteration values below the convergence criterion in all links of the network. If the system has not converged, the calculation returns to the estimation of composite costs, because the new travel and waiting times will affect them at a path level, and consequently, at all other levels. In turn, the new costs will affect trip generation, modal split and assignment, and even the location of activities in a next time period. This is one of the main advantages of the structure of TRANUS, because the effects of congestion are felt all through the decision chain, not just in route assignment as in most traditional transport models.

It must be noted that these iterations do not change the path set but its costs and disutilities, and therefore, the order in which paths were found by the path search algorithm might change due to congestion; in other words, what was the minimum path in the first iteration might well end as the worst path of the choice set. If the analyst thinks that congestion might also change the choice set itself, the path search algorithm may be operated again based on congested travel times, and the whole process may be repeated as many times as felt necessary. In theory, paths should be recalculated in every iteration, but this procedure would be too costly in terms of computer time.

A.2.5 The evaluation procedure

For the simulation and evaluation of land use and/or transport policies, the model must be applied throughout the projection period at discrete time intervals for a base case scenario, where the policies are not included, and for an alternative scenario which includes an explicit definition of the policies. The differences in the results will represent the net effect of introducing the policies. At the end of the process an evaluation procedure may be applied, which compares the results of the base case and alternative scenarios, and estimates a number of socio-economic, financial, and optionally, energy consumption indicators.

Current costs and benefits are arranged in the form of accounts. There are three main accounts: users, operators and administrators. Users perceive benefits (positive or negative) from land use and transport. Operators of transport services perceive income from the fares they charge to users, and must pay operating costs. Administrators are the organization which look after the infrastructure; they may perceive benefits from possible toll or parking charges to vehicles, and they must pay maintenance costs.

If energy evaluation option is selected, the evaluation procedure calculates the amount of energy required for domestic heating or cooling as a function of dwelling size and insulation levels. In the case of transport related energy, the transport model estimates consumption by operator as a function of average vehicle speed in each link. The results are read by the evaluation procedure which transforms consumption into standard energy units.

For the economic and financial evaluation it is necessary to provide the evaluation procedure with details such as the capital costs of the alternative scenario through time, discount rates, and shadow prices for some elements, such as for different types of fuel. With this data, the procedure will estimate indicators such as cost/benefit ratios, net present values and internal rate of return, considering economic or financial costs.

A.3 Operational structure of TRANUS

From an operational point of view, TRANUS consists of a set of computer programs linked to each other through data files. There are several types of programs, which are described very briefly in the following sections. The general operative structure of TRANUS is shown in Figure 4.

A.3.1 The main programs

LOC This program performs all the land use related calculations. It requires as inputs the location of activities and prices in a previous time period, a description of current land use policy, transport costs for a previous time period, and a set of parameters. The program outputs the resulting location of activities and prices, and a file containing the resulting matrices of functional flows.

FLUJ This program acts as an interface between the activity location model LOC and the transport model TRANS, transforming socio-economic flows into transport categories, and adjusting for different timescales. FLUJ reads the matrices produced by LOC and a parameter file where the optional transformations are specified, and produces a new file containing the resulting matrices by transport categories. Exogenous trips may be introduced by category and mode.

PASOS This program performs multi-path search from the transport network. It reads a network file and a parameter file, and searches for the first n paths connecting each origin to each destination by mode.

TRANS This program contains the transport model. It reads the network file, the path descriptions, and the functional flows by transport category. With this data, the program performs all transport-related calculations: trip generation, modal split, assignment and capacity restriction. After supply-demand equilibrium, the program outputs the results of the assignment process, the costs and trip matrices, and the evaluation indicators calculated during the simulation.

COST The purpose of this program is to act as a transport-to-activity location interface. Thus, its function is the opposite of FLUJ. It reads the matrices of cost by transport category produced by the transport model, and turns them back to the, but in reverse. It also estimates intrazonal costs.

EVAL Contains the evaluation procedure. It reads the land use and transport results from two scenarios and several time periods. With this, and additional data from a parameter file, the program organizes an accounting system around users, operators and administrators, calculating present values of current costs and benefits on a yearly basis, present values of capital costs, consumers' surplus, cost/benefit indicators and rates of return, together with energy evaluation.

A.3.2 Complementary programs

TDAT Verifies transport related files, and check the symmetry of the network.

LCAL Used for calibrate the activity model. This program proceeds by constraining the location of production with real values and calculate a set of attractors which guarantee that the model simulates production values correctly. The resulting attractors may then be introduced as data to the activity location model LOC.

Interactive Data Base IDB

The interactive data base for TRANUS is a Windows interface that makes it easy to create a project, providing facilities to introduce and edit input data and to view and analyze the results of the models. The IDB manages all scenarios in the project. Data is carried automatically to all dependent scenarios in the scenario tree. Many windows are provided to display and edit the data. When a particular data item has changed IDB shows it in a green color, and the scenario in the tree is marked with a green square. Results of simulations can be copied and pasted in any Windows application in the usual way of the Windows environment. The IDB provides a context sensitive Help. Two tipical windos of the IDB are shown in Figure 6.

A.3.3 Display programs

As an alternative to the IDB, results of simulations can be displayed and/or printed with the following display programs from the DOS environment.

MATS Used for printing output matrices from the main programs. The user may choose from a menu of options, such as trip matrices by mode and category, total trips, composite or monetary transport costs by category, activity flows, and some others. An option produces histograms showing the distribution of trips by cost for each category. Optionally, matrices may be aggregated in various ways.

MATESP Used for printing spetial output matrices from the transport model. The user may choose from a menu of options, such as matrices of trips making transfers by mode; matrices of trips using an specific link (or a set of links) by transport category; travel time, travel distance and travel costs matrices by category and mode; trip matrices by transport operator. Optionally, matrices may be aggregated in various ways.

IMPLOC This program reads land use outputs from different runs, and compiles tables for one or more user-defined variables.

IMPAS The purpose of this program is to print the results of the path search program PASOS with several output options.

IMPTRA This program prints the results of the assignment from the transportation model in a number of optional formats.

A.3.4 Graphic system GUS

This is the interactive graphic end of the system, in the form of a Windows based Graphic User Shell (GUS). It complements the IDB displaying full color graphics to show the results of the transport simulations: paths, assigned traffic by mode, operator and transit lines, level of service, volume/capacity ratios, desire lines, and so on. Because it is a Windows program, the resulting images can be copied to most word processors, spreadsheets or paint programs for further enhancements. Figure 5 shows some graphics generated by GUS.

Appendix B: Mathematical and algorithmic structure of TRANUS

This chapter describes the mathematical structure of TRANUS to serve as a reference material for those interested in internal structure. The main components of the system are the activity model, the activity-transport interface, the transport model, the transport-activity interface, and the evaluation procedure. Each component is discussed in greater detail below.

B.1 The activity model

The purpose of the activity location model is to simulate a spatial economic system. Given a region or city divided zones, the model estimates the activities that locate in each zone and the interactions that they generate.

B.1.1 Basic concepts

The central element in the activity model is a spatial input-output procedure defined by economic sectors and their production and consumption relationships.

The starting point is the classical structure of an input-output model. The main elements are final demand, intermediate demand and primary inputs. The vector of final demands represents the final destination of production. In input-output models final demand usually includes private consumption, government consumption, exports and investments. The economic system must produce the quantities demanded in each sector; to achieve this, intermediate inputs are required, generating a chain of productions and consumptions. In addition to intermediate inputs there are primary inputs; they include salaries, imports, profits and taxes. The sum of final demand plus all intermediate demands is equal to total production in the system. Similarly, the sum of all intermediate production plus primary inputs is equal to total production.

In TRANUS, the basic concepts of the input-output model have been generalized and a spatial dimension has been added. The term sector is much more general than in the traditional concept; it may include the classical sectors in which the economy is divided (agriculture, manufacturing, mining, government, etc.), factors of production (capital, land labor), as well as population groups, employment, floorspace, energy, or any other that is thought relevant to the spatial system being represented. The number and types of sectors must be defined according to the requirements of each application; the units in which each sector is represented (money, production, jobs, people, hectares, etc.) can also be defined to suit each case. This makes it possible to apply the model to urban or regional areas alike.

A first distinction can be made between transportable and non-transportable sectors. The main difference is that transportable sectors may be consumed in places different to those in which they were produced. For example, the demand for coal by a steel industry located in a particular place may be satisfied by mining industries located in other regions. Similarly, the demand for labor in a central area may be satisfied by population living in the outskirts. A typical example of a non-transportable sector is land and buildings; these must be consumed in the same place where they produced. As a consequence, transportable sectors generate trades or, in general, economic flows of goods, money people; such flows are later turned into demand for trips. The transport system, in turn, must make such flows possible, and imposes transportation costs on them. By contrast, non-transportable sectors do not require transport and do not generate flows.

Another distinction is made between internal and external zones, and receive a different treatment in the model. All economic relations occur between internal zones. External zones are only used to represent imports and exports. However, it is possible to define external zones only for the transport model to represent external or through trips.

In turn, internal zones can be of two types: first or second hierarchical level. A first level zone may consist of two more second level zones, thus affecting the spatial and sectoral distributions. A sector may be transportable at both and second level zones, may be non-transportable, or may be only transportable at a second hierarchical level. The of all activities in second level zones will always equal the activities in the first level zone to which they belong.

Each sector located in specific zones is characterized by a number of specific associated variables:

- Exogenous production: It is the production not generated or demanded by other internal sectors. It is equivalent to final demand in input-output models. The location of exogenous production does not depend on the internal logic implicit in the model; rather, it depends on political or historical considerations or on elements not modeled or external to the system. Exogenous production is not subject to the locational and sectoral distribution procedures of the model. It is a given input, to be added to endogenous, induced, production. The growth of exogenous production from one period to the next is taken care of by an incremental model; in this way increments are assigned to zones as a function of attractor variables. Alternatively or complementary, specific growth in exogenous production can be exogenously determined for particular zones.
- Induced Production: It is production generated by internal or external demands. It is allocated to zones with the spatial and sectoral distribution model. The growth of induced production depends on the growth of those sectors that demand it.
- Exogenous demand: It is an additional demand to that generated internally. If this additional demand takes place in external zones, it is termed exports. The term exogenous demand, then, will always refer to that taking place in internal zones only. Exogenous demand is distributed together with induced demand. The growth of exogenous demand from one period the next is dealt by the incremental model.
- Induced demand: It is determined by the consumption requirements of final demand sectors or by the intermediate activities.
- Exports: Exports are defined as internal production of the study area consumed in external zones It is represented as exogenous demand in external zones. The model allocates the required production among internal zones only.
- Imports: Imports are defined as a demand in an internal zone satisfied by production in external zones. It is distributed together with the rest of induced demand. Imports compete with internal production, but may be restricted to fixed amounts given exogenously or estimated by the incremental model.
- Capacity of production: Total production (exogenous+induced) in a sector and zone may be subject to restrictions. Maximum, minimum or maximum and minimum restrictions may be imposed for particular sectors and zones.

- Consumption cost: It is the cost of a unit of output at the consumption zone (CIF). The consumption cost is, hence, equal to the production cost plus the transport cost from the production zone to the consumption zone. Because consumption in a particular zone may be satisfied by production from several zones with different production and transportation costs, consumption cost is calculated as a weighted average.
- Production cost: It is the unit cost of production at the production zone (FOB). It is calculated as the sum of the consumption costs of required inputs, plus value added.
- Equilibrium price: It is the value of a sector in a particular zone when production is constrained. It represents a scarce commodity with supply limited to a production restriction. If there are no restrictions, then the price is equal to the production cost. If demand exceeds production capacity, the price increases, generating an excess profit to the producer or opposite case is that of demand being less than a minimum constraint; in this case the equilibrium price drops below production costs, generating a loss to producers.
- Value added: It is the value of capital and labor that is added to all other input commodities in order to obtain a unit of production. Typically value added includes payments to capital (rent), to labor (salaries), taxes or subsidies, capital payments on equipment, and so on.
- Consumption utility: It is the logarithmic average of the utility values used in the probabilistic distribution of demand to production zones. Uses transport utility instead of transport monetary cost.
- Transport cost: It is the monetary expenditure needed to transport one unit of production from the production zone to the consumption zone. It is calculated by the transport model. In the case of commodities, transport cost is a unit cost. For such as residents, this value represents transport expenditure, and depends on the number of trips in a given time period, that is, the time period being represented in the activity model (month, year, etc.). It has an influence on the cost of production.
- Transport disutility: Also calculated by the transport model. It includes the monetary cost but other elements as well, such as the value of time, subjective elements, and others. Transport disutility is always entered to the activity model as a unit cost, i.e. trip. It influences the spatial-sectoral distribution of production.

B.1.2 Demand and distribution of production

In principle, every sector requires inputs from other sectors. Hence, a part of total production goes to intermediate consumption, and the rest goes directly to final consumption, whether internal or external (exports). Given a certain amount of final demand in one or more sectors and zones, the model estimates induced production through demand functions. The model then allocates this induced production to zones through spatial distribution functions. In turn, induced production requires further inputs, thus generating a production chain and the corresponding location of activities.

From the above relations, economic transactions are derived, giving rise to functional flows if production and consumption take place in different zones. These are defined as transportable flows involving people, goods, services money. From these flows, transport demand is derived at a later stage. In some transactions, non-transportable commodities may be involved, such as land or floorspace. In this case economic trades are involved, but no flows are generated. Each sector may generate different transactions and different types of flows. A manufacturing industry, for example, requires non-transportable land and buildings as well as transportable raw materials, other manufactured goods and labor, thus generating a demand for transport.

As shown in Figure 1, a transaction involves a consumption zone and one or more production zones; it may be that consumption zone is, at the same time, a production zone. The model distributes the goods and services purchased among production zones in a probabilistic way. In the diagram, the arrows point in the direction of the flows of goods; money for such purchases will flow in the opposite direction.

B.2 Structure of the activity model

The activity location model performs the calculation steps described in Figure 2:

incremental location of exogenous variables
calculation of attractors for induced production
estimation of induced demand
estimation of production costs
location of induced production
calculation of consumption costs and disutilities
restriction checks and adjustment of equilibrium prices

B.2.1 Increment and location of exogenous variables

The first stage in the sequence of calculations of the activity location model is the estimation of the growth of exogenous variables in each sector and zone. By definition, exogenous variables depend on elements not simulated in the model: they are given inputs. Consequently, any increment of these variables in the future must be given to the model in the corresponding time period.

The exogenous variables that may be modified for each time period with the incremental model are:

exogenous production
exogenous consumption
capacity of production (restrictions)
exports
imports
initial attractors

If H denotes any of the above exogenous variables, the increment between period *t-1* and *t* (positive or negative) in specific zone i may be added exogenously as:

$$H_i^{n,t} = H_i^{n,t-1} + \Delta H_i^{n,t} \tag{1}$$

However, for the first three exogenous variables, that is, production, consumption and restrictions, it is possible to specify a study-wide global increment. In this case an attractor function must also be specified. The model will allocate the global increments to zones in the proportions that result from the attractor functions. Here the distribution of increments of total production is described; exogenous consumption and restrictions use a similar formulation.

If $X_i^{*n,t}$ is the exogenous production of sector *n* in zone *i*, time period *t*, then:

$$X_i^{*n,t} = X_i^{*n,t-1} + \Delta X_i^{*n,t} \rho_i^{n,t} + \Delta X_i^{*n,t} \tag{2}$$

where: $X_i^{*n,t-1}$ is the exogenous production of sector *n* in zone *i* for time period *t-1*,

$\Delta X^{*n,t}$ is the global increment of exogenous production of sector *n* between *t-1* and *t*,

$\Delta X_i^{*n,t}$ is the given increment of exogenous production of n in zone i for time period t (user defined), and

$\rho_i^{n,t}$ is the proportion of the increment of *n* allocated to zone *i* for period *t*.

The proportion of the global increment assigned to each zone is a function of attractor functions:

$$\rho_i^{n,t} = \frac{A_i^{n,t}}{\sum_i A_i^{n,t}} \tag{3}$$

where $A_i^{n,t}$ is the attractor of sector *n* in zone *i* for period *t*, defined in the next section.

B.2.2 Calculation of attractors

The model calculates the attractor functions, both for the allocation of induced production as well as for the increments in exogenous variables as described in the previous section. In the case of induced production, attractors are calculated before the iterative sequence starts. The attractor functions are defined as follows:

$$A_i^{n,t} = \left(\sum_k b_k^n \left(X_i^{*k,t-1} + X_i^{k,t-1}\right)\right) W_i^{n,t} \tag{4}$$

where: $X_i^{*k,t-1}$ is the exogenous production of a sector *k* attracting *n* in zone *i*

$X_i^{k,t-1}$ is the induced production of a sector *k* attracting *n* in zone *i*

b_k^n is the relative weight of sector *k* as an attractor to sector *n*

$W_i^{n,t}$ is the initial attractor of zone *i* that takes into account non-modeled elements that attract the location of *n*

The set of relative weights of each sector *k* in the attractor function of sector *n* may be different for first and second zones.

B.2.3 Generation of induced demand

The amount of inputs that a unit of production of a sector requires from another sector is determined by a function. The model includes as options a fixed demand (equivalent to technical coefficients in an input-output model), variable (elastic) demand and the possibility of specifying substitutes. Land is a typical example of a substitute when different types of land are present in the system, such as low density, high density, industrial, commercial, etc.

The general form of a demand function is as follows:

$$a_i^{mn} = \min^{mn} + \left(\max^{mn} - \min^{mn}\right) \cdot \exp(-\delta^{mn} U_i^n) \tag{5}$$

where: a_i^{mn} is the amount of the production of sector n demanded by a unit of sector m in zone i

$\min^{mn}$ is the minimum amount of n required by a unit production of m

$\max^{mn}$ is the maximum amount of n required by a unit production of m

δ^{mn} is the elasticity parameter of m with respect to the cost of input n

U_i^n is the consumption disutility of n in i.

The resulting form of the demand function is shown in Figure 3; in this example a maximum consumption of 50 and minimum of 10 is applied for different values of δ^{mn}.

Figure 3: Examples of a demand function

Next, the proportion applied to the demand function to take into account the presence of substitutes is estimated with multinomial logit model of the form:

$$S_i^{mn} = \frac{\exp(-\delta \tilde{U}_i^{mn})}{\sum_k \exp(-\delta \tilde{U}_i^{mn})} \tag{6}$$

where K^n is the set of all substitute sectors of n.

The amount of inputs n demanded by sector m in zone i is, then:

$$D_i^{mn} = \left(X_i^{*\mathrm{m}} + X_i^m\right) a_i^{mn} S_i^{mn} \tag{7}$$

The total demand for inputs n in a particular zone i is obtained as the sum of the consumption of n by all sectors possible exogenous demands:

$$D_i^n = \sum_m D_i^{mn} + D_i^{*n} \tag{8}$$

where: D_i^n is the total demand for n in zone i,

D_i^{*n} exogenous demand for n in zone i.

In the first iteration the system will only have exogenous production and the induced production directly derived from In successive iterations induced demand from the production of all sectors in the previous iteration are added.

B.2.4 Calculation of production costs

The production cost is calculated as the consumption cost of all necessary inputs to produce one unit of m in zone the value added:

$$c_i^m = \left(\sum_n D_i^{mn} \tilde{c}_i^n\right) + VA_i^m \tag{9}$$

where VA_i^m is the value added to the production of *m*, and $\tilde{c}_i^n$ is the consumption cost of input *n* in zone *i*.

B.2.5 Location of induced production

Once the amount of production demanded in each zone has been estimated, it must be distributed to production zones. a sector is non-transportable, all production is assigned to the zone in which it is demanded. If the sector is transportable, demand is distributed to production zones with a multinomial logit model, in which the utility function each zone is determined by:

$$U_{ij}^n = \lambda^n \left(p_j^n + h_j^n\right) + t_{ij}^n \tag{10}$$

where: p_j^n is the price of sector *n* in the production zone *j*,

h_j^n is the shadow price of sector *n* in the production zone *j*,

t_{ij}^n is the transport disutility for sector *n* from the production zone *j* to the consumption zone *i*, and

λ^n is a parameter that regulates the relative importance of prices versus transport disutility in the utility function

The shadow price of production is estimated at calibration stage. If zone *j* is a second level zone, the calculation of utility function is identical, except that a different parameter is used $\lambda''^n \neq \lambda^n$.

The results of the above calculation are divided by the utility of the best option to obtain the marginal utility:

$$\tilde{U}_{ij}^n = \frac{U_{ij}^n}{\left(\min_j U_{ij}^n\right)^{\theta^n}} \tag{11}$$

These marginal utilities are entered into the multinomial logit model to estimate the probability that the production sector *n* demanded in zone *i* is located in zone *j*:

$$P_{ij}^n = \frac{A_j^n \cdot \exp(-\beta^n \tilde{U}_{ij}^n)}{\sum_j A_j^n \cdot \exp(-\beta^n \tilde{U}_{ij}^n)} \tag{12}$$

$$X_{ij}^{n} = D_j^{n} P_{ij}^{n} \quad (13)$$

where: X_{ij}^{n} production of *n* located in the production zone *j* induced by activities in the consumption zone *i*,

A_j^n attractor term for the production of *n* in *j*,

U_{ij}^{n} marginal utility of location of n in zone j to satisfy the demand in zone i,

β^n dispersion parameter of the multinomial logit model.

If the consumption zone i is an internal zone, the distribution is applied to all zones internal and external. If the consumption zone is external (exports), the distribution is applied to internal zones only. In other words, the model not allow that the demand for exports be satisfied by imports.

Equation 12 is applied for first level production zones. For second level zones a conditional probability is applied:

$$P_{ij''}^{n} = P_{ij}^{n} \cdot \frac{A_{j''}^{n} \cdot \exp(-\beta''^{n} \tilde{U}_{ij''}^{n})}{\sum_{J} A_{j''}^{n} \cdot \exp(-\beta''^{n} \tilde{U}_{ij''}^{n})}, \forall j'' \in j \quad (14)$$

The term j'' denotes a second level zone that belongs to the first level zone *j*. The first probability corresponds to macrozone *j*, and the second probability corresponds to a subzone j'' that belongs to *j*. Note that $\beta^n \neq \beta''^n$, and that $U_{ij}^{n*} \neq U_{ij''}^{n*}$. Production assigned to a subzone is:

$$X_{ij''}^{n} = D_i^{n} P_{ij''}^{n} \quad (15)$$

Finally, total induced production allocated to a zone (macrozone or subzone) is obtained by adding over all demand zones:

$$X_j^n = \sum_i X_{ij}^n \quad (16)$$

B.2.6 Consumption costs

Once demand has been assigned to production zones, consumption costs are calculated, that is, the amount that a sector *m* located in *i* has to pay for the consumption of one unit of input *n*. Because purchases are spatially distributed, an average is calculated, weighted by the price paid in each production zone plus the corresponding transport costs:

$$\tilde{c}_i^n = \frac{\sum_j X_{ij}^n \left(P_j^n + tm_{ij}^n\right)}{\sum_J X_{ij}^n} \quad (17)$$

where: X_{ij}^{n} amount of production of sector *n* demanded in *i* and produced in *j*,

P_j^n unit price of *n* in the production zone *j*,

$tm_{ij}^{\ n}$ monetary cost of transporting a unit of sector *n* from the production zone *j* to the consumption zone *i* (different from transport disutility).

B.2.7 Consumption disutility

The disutility of consuming *n* in *i* is the logarithmic average of the disutilities used in the distribution to the production zones:

$$U_i^n = -\frac{\ln Pg^n}{\beta^n}(\min_j U_{ij}^n)^{\theta^n} \tag{18}$$

Note that the expression is multiplied by the minimum disutility, because in the distribution the marginal disutility used. *Pg* is defined as a series of the following form:

$$Pg^n = \sum_{j=1}^{Z} G_j \prod_{h=1}^{j-1} (1 - G_h) \tag{19}$$

where the term G_j is the numerator of the probability that the production demanded in *i* locates in zone *j*; from equation (12):

$$G_j = A_j^n \exp(-\beta \tilde{U}_{ij}^n) \tag{20}$$

B.2.8 Checking for restrictions and adjustment of equilibrium prices

The production of a sector in a zone may be limited to the minimum and/or maximum capacity of production. production assigned to a zone after the distribution lies within the established limits, the price is equal to the production costs plus value added. If, however, production is above the maximum or below the minimum, then the price is determined by demand-supply equilibrium. At the end of each iteration, the model checks for restrictions and adjusts prices accordingly; the price is increased if the maximum restriction is violated, and is reduced if the minimum restriction is violated. These variations in price affect the distribution of production in subsequent iterations, until an equilibrium is reached. Prices are adjusted as follows:

$$P_j^{n,t}\left\{\begin{array}{c} < P_j^{n,t-1}, (X_j^{*n} + X_j^n) < R\min_j^n \\ P_j^{n,t-1}, (X_j^{*n} + X_j^n) > R\max_j^n \\ c_j^{n,t}, R\min_j^n = 0, R\max_j^n = \infty \end{array}\right\} \tag{21}$$

where: $P_j^{n,t+1}$ is the unit price of sector *n* in zone *j*, to be used in the next iteration,

$P_j^{n,t}$ is the unit price of sector *n* in zone *j* in the current iteration,

$R\min_j^n$ and $R\max_j^n$ are the minimum and maximum restrictions to the production of sector *n* in zone *j*.

B.2.9 Convergence

In each iteration the convergence in prices and production is evaluated. Both are calculated for each zone and sector the percentage variation with respect to the previous iteration. These indicators are calculated separately for each sector, and adopt the value of the worst zone, that is, the zone that varied the most:

$$Cp_j^{n,\,it} = \max_j \frac{p_j^{n,\,it} - p_j^{n,\,it-1}}{p_j^{n,\,it-1}}, \quad CX_j^{n,\,it} = \max_j \left[\frac{X_j^{n,\,it} - X_j^{n,\,it-1}}{X_j^{n,\,it-1}} \right] \tag{22}$$

The model ends the iterative process when both convergence indicators are smaller than a prespecified convergence criteria, or when a maximum number of iterations is reached.

B.3 Activities-transport interface

The activity location model produces as output, among others, a set of matrices of economic flows by (transportable) sector. These flows are turned into trips by the transport model, but before doing this, it is necessary to perform transformations. The purpose of the activities-transport interface is to perform such transformations, including possible changes in categories, time period, direction of flows and others. The result will be a set of matrices of flows by transport categories. The following transformations are made possible by the interface:

- transformation from economic categories to transport categories,
- time factor,
- volume/value factor, and
- direction of flows.

Each one of these possibilities is described in the following sections.

B.3.1 From economic categories to transport categories

The interaction between transportable economic activities generate flows of goods or people. In a regional application, for instance, the interaction between industries such as agriculture, manufacturing and so on, generate movements of commodities; the interaction between employment and residents generate movements of commuters. Each transportable sector generates a corresponding matrix of flows. It may be that these categories coincide with those in the transport model, but this may not be convenient for some applications. For example, the analyst might find it convenient for purpose of the study to transform these economic categories into categories that are relevant to transport. Instead of having economic sectors in the transport model, it might be preferable to work with categories such as heavy bulk, general cargo, containers, and so on. In the case of movements of people, several trip purposes might be merged into trips by car owners and non-car owners. For the sake of saving computer time, the analyst may want to reduce the number of transport categories if he or she feels that this will not affect the results seriously.

In other words, the number and types of categories in the activity model does not have to match transport categories. The activity-transport interface allows for such transformations, applying given proportions.

B.3.2 Time factors

It is common that the activity model is set on a different time scale as the transport model. In a regional application, production by sector usually refers to annual amounts, while the transport model will probably work on a daily basis. an urban application, monthly units are probably the most convenient way to represent salaries and expenditure, but could be the case that a peak hour simulation is required for the transport system.

To make time units compatible, the interface makes a distinction between two types of flows: normal flows or commuter-type flows. Commodity movements are typically represented as normal flows. For instance, the annual flow of Tons of agricultural produce will generate a certain daily amount, and for this the interface will apply a time factor. Movements of people are typically represented as habitual flows. A flow between jobs and residents will generate trips that take place every day, even if the activity model is in terms of a month. Hence, for this type of flows time factors are not applied. A flow does not correspond necessarily to a trip, but the transport model will make the necessary adjustments with the trip generation model described further below.

B.3.3 Volume/value factor

The units in which activities are represented in the activity model does not necessarily correspond to transport units. instance, manufacturing industry might be represented in money units or even employment, while the transport model will probably work with physical units such as Tons. For this purpose, the interface applies value/volume factors.

B.3.4 Direction of flows

The activity model always generates economic flows from the consumption zone to the production zone, that is, in direction that purchases take place. This, however, might not be convenient to the transport model. In an urban application, residents generated by employment will be in the direction workhome; if the transport model is going to derive peak hour trips from these flows, it will be preferable to reverse the direction of the flows. If total day trips are being considered, then both directions will be relevant to represent return trips. These transformations are also dealt by the interface.

B.3.5 Transformation of flows equation

All the transformations described in the preceding paragraphs may be represented in a single equation as follows:

$$F_{ij}^{S} = \sum_{n}\left(X_{ij}^{n}\frac{pr^{ns}vol^{ns}pc^{ns}}{tiem^{ns}} + X_{ij}^{n}\frac{pr^{ns}vol^{ns}cp^{ns}}{tiem^{ns}}\right) \tag{23}$$

where: F_{ij}^{S} is the flow of transport *s* from origin *i* to destination *j*,

X_{ij}^{n} is the production of the transportable sector *n* located in *j* and consumed in *i*,

pr^{ns} is the proportion of the economic flow n that forms part of transport category s,

vol^{ns} is the value/volume factor for the economic flow n that forms part of the transport category s,

$tiem^{ns}$ is the time factor for the economic flow n that forms part of the transport category s,

cp^{ns} is the proportion of the economic flow that moves in the direction of consumption to production,

pc^{ns} is the proportion of the economic flow that moves in the direction of production to consumption.

Summation is made over all economic flows n that form part of the transport category s. It is essential that every transportable factor forms part of a transport category, however small the proportion. There are no restrictions on the proportions, although it is common that:

$$\sum_s pr^{ns} = 1$$

The are no restrictions on the proportions that determine the direction of flows. For example, if the direction of the economic flows is the same than the transport flows, then:

$cp^{ns} = 1$ and $pc^{ns} = 0$

Otherwise:

$cp^{ns} = 0$ and $pc^{ns} = 1$

If a two-way directionality is needed, as for total day trips, then:

$cp^{ns} = 1$ and $pc^{ns} = 1$.

B.4 The transport model

B.4.1 Basic concepts

The main purpose of the transport model is to estimate travel demand and assign it to transport supply, such that an equilibrium is reached. Travel demand is estimated from the economic flows turned into flows by transport category the interface. In turn, transport costs and disutilities resulting from equilibrium are used by the activity model to simulate a subsequent time period.

The two main elements of the transport model are, then, supply and demand. Users represent demand, that is people goods that require a transport service for freight or passengers. On the supply side, it is possible to distinguish between physical supply (roads, parking, stations, ports, etc.) and the operative supply. An administrator is usually in charge the physical supply, providing maintenance and, in some cases, charging the users. The operative supply, represented private or public organizations that operate vehicles of several types, usually charge fares and must pay for operating costs and possible charges to administrators. Figure 4 shows a diagram relating the various entities of the transport system and their main economic relationships.

Private transport is a special case in which the user is, at the same time, the operator. It is also common that railways both an operator and administrator of their own infrastructure. The transport model, however, will always consider as separate entities for accounting purposes.

B.4.2 Transport demand

Users are classified by transport categories; among other things, this allows for a separate treatment of passengers and freight. Passengers, in turn, may be classified by income group, trip purpose, or combinations of both. Each person category has an associated car availability, that limits the selection between public and private modes of transport. From each travel option, users perceive a disutility, that includes the monetary cost of travel, the value of travel time waiting time, and subjective elements, such as comfort, reliability, safety, and so on.

B.4.3 Operative supply

Transport supply is organized hierarchically in three levels: modes, operators and routes. Modes represent a set of operators that provide a service for a particular kind of user. In TRANUS, modes represent broad categories such as public, private, light or heavy commodities. Each trip category may choose among specific modes, such that commodities can only choose from freight modes and people can only choose from passenger modes. For each transport category a choice set must be defined, consisting of a list of modes available to that particular category. Each mode, in turn, may consist of several operators. An operator is characterized by a service of some general characteristics, such as vehicle type, tariffs, operating costs, transfer costs, energy consumption, and so on. Many operators may belong to a common mode. A user can freely transfer from one operator to another, provided such operators belong to the same mode. A typical application may consider three modes: freight, public and private. The freight mode may contain light and heavy trucks, railways, barges, and so on, as operators, such that a consignment combine them to reach its destination. The private mode usually considers a single operator: cars, but parking or HOV cars may be included as well. Public transport probably provides the most complex structures, and a large number of different operators may be specified, such as buses, minibuses, light and heavy subways, feeder buses, jitneys, etc. Certain rules might be imposed on transfers; integrated fares is one example, and some transfers may be prohibited altogether, even if the source and destination operator belong to the same mode. For instance, a transfer between a normal car and a HOV car may be prohibited.

At a more detailed level, a public transport operator may be organized as a set of routes. The definition of a route is service that must follow a specific sequence of links in the network. If a user wants to transfer from one route to another, he or she will have to pay a transfer cost and waiting time, except for possible integrated fares.

B.4.4 Physical supply

A transport network represents physical supply in the model. The network is defined as a directed graph, or a one-way links and nodes, like the example of Figure 5. Nodes may represent road junctions, points at which the characteristics of a road may change, stations, bust stops, ports, and the like. A subset of nodes are called Centroids represent zones, and for the transport model all trips either start or finish at centroids. Each centroid must connected to one or more nodes of the network. Links, on the other hand, may represent street sections, highways, rural roads, railways, airways, waterways or any other kind of relevant infrastructure. Links have specific characteristics, as distance, capacity, speed, and so on. Some of these characteristics are link-specific (distance, capacity), but others defined generically, and for this purpose link types are defined. All links of the same type share common characteristics (speeds, toll charges, maintenance costs, etc.). Link types also define which operators can use them; for instance, a highway link type may be used by trucks, cars and buses, but not by trains or vessels.

The following characteristics define a link:

origin node
destination node
link type
link length
physical capacity
transit routes
prohibited turns

The link type attribute specifies the following generic characteristics:

free flow speed by operator
car-equivalent units by operator
distance-related operating cost of vehicles per operator
charges or tolls by operator
administrator in charge
fixed and marginal operating costs
capacity restriction function

Usually, the physical capacity of a link is measured in car-equivalent per hour or daily. In special cases other units may be used for convenience, such as trains, coaches, Tons, or any other. In very dense parts of a city, the capacity of a link may be determined by the capacity of intersections.

Transit routes are coded directly in each link. Each route has a specific frequency (vehicles per time unit). The model assigns the corresponding vehicles to the network. It is also possible to specify transit routes with an undefined frequency, in which case the model will adjust the frequency to demand.

Prohibited turns are also coded for each link, to indicate nodes towards which vehicles cannot turn. This a simple and error-free way of coding turn prohibitions, without having to resort to fictitious nodes and links.

The transport model represents the transport network internally as a dual graph. The network is coded in the traditional way, with links representing road sections and nodes representing intersections. Then the program automatically turns the external links into nodes of an internal dual graph, and creates internal links to represent possible connections. Prohibited turns are simply skipped during the process of creating the internal dual graph. This method is completely transparent to the model user, because once all calculations have been performed, the model translates back the results in the original form. Figure 6 shows an example of a simple network and the way in which it is coded by the user. Figure 7 shows the corresponding internal dual graph created by the model.

Formally, the original direct graph may be defined as:

$N=(V,L)$

where V is a set of vertices (nodes) $V=\{v_1, v_2, ..., v_n\}$, and L is a set of links $L=\{l_1, l_2, ..., l_n\}$. The definition of links goes beyond the concept of a physical link; a link is defined as:

$$l_i = \left(v, w, r, Q^{vw}\right), v, w \in V, r \in R^{vw}$$

where v and w are the origin and destination nodes of the physical link; r represents a transit route that belongs to the R of all routes that offer a service along physical link (v,w). Q is the set of attributes of the link (distance, capacity, Furthermore, a subset $S \subset R$ may be defined to represent all those routes that actually stop at link (v,w). In TRANUS, this way of representing the network with multiple operators and routes associated with each physical link is called a multidimensional network.

Possible connections in this original network are:

$$T = \left\{(v, w, x) \mid v, w, x \in V, \left(v, w, Q^{VW}\right) \in L, \left(w, x, Q^{wx}\right) \in L\right\}$$

Prohibited turns GP are a subset of T, $GP \subset T$.

The dual graph $\mathsf{D}(N)$ corresponding to the original graph N is defined by the following pair:

$\mathsf{D}(N) = (\mathsf{V},\mathsf{L})$

where $\mathsf{V} = L(N)$. In other words, the vertices V of the dual graph are the links of the original graph. Dual links L are defined as follows:

$$\mathsf{L} = \{(f, g, Q^{fg}) \mid f, g \in \mathsf{V}, f = (v, w, r, Q^{vw}), g = (v, x, t, Q^{wx}), (v, w, x) \notin GP,$$
$$r = t \vee r \in S^{vw}, t \in S^{wx}\}$$

The first condition, $f, g \in \mathsf{V}$, means that all vertices L of the dual graph must correspond to links in the original direct graph. Dual links are then created to connect origin vertices with destination vertices.

B.4.5 Structure of costs

The transport model makes a distinction between three types of costs, corresponding to the three main entities in the transport system: users, operators and administrators.

- user costs: include monetary and non-monetary components, also called generalized costs; these are accounted terms of demand units (per trip);
- operating costs: strictly monetary, in terms of vehicle units;

- administrative costs: include maintenance costs, also strictly monetary; these are accounted in terms of distance.

User and operative costs influence travel demand and assignment; administrative costs are only taken into account evaluation purposes.

B.4.6 Monetary costs to users

The monetary component of user costs is termed tariffs. Two broad types of tariffs may be specified, depending on whether they are dependent or independent of operating costs. It is common that public transport operates with tariffs that are quite independent of their operating costs. Freight operators, by contrast, tend to link tariffs with their actual operating cost, (including a profit margin). In the case of private cars, the user and operator are the same, so that tariffs should strictly correspond to operating costs; however, there is considerable evidence that car users do not perceive full costs, but only a small proportion. For convenience, the term tariff refers to monetary costs to users, even as in case of cars it might sound inadequate.

The model offers the following elements to specify a tariff function:

$$t_0 = f(tf_0, tt_0, td_0, tc_0) \tag{24}$$

where: tf_0 is a fixed tariff when operator o is boarded; if there are integrated tariffs, then tf will depend on the previous operator,

tt_0 is a tariff per unit of time,

td_0 is a tariff per unit of distance,

tc_0 is a factor that multiplies the operating cost to be added to the tariff.

For each application of the model, different tariff functions may be specified. In the case of a freight service, it is common to specify tariffs as a function of operating costs; hence only tc_0 will have a value (e.g. 1.2), leaving the of the elements as zero. In the case of transit, it is common to have tariffs as a function of distance and a fixed component. In general, tariffs either have fixed and distance and time related elements, or are related to operating costs; a combination of both is very rare.

B.4.7 Operating costs

Operating costs per vehicle, of a particular operator *o* and link type τ; includes the following elements:

$$c_0^{\tau} = f\left(cf_0, ct_0, cd_0{}^{\tau}\right) \tag{25}$$

where: cf_0 is the fixed operating cost of a vehicle of operator *o* to be applied only when the vehicle is boarded, that is, once for every trip made; usually refers to administrative costs and loading/unloading in the case of goods vehicles.

ct_0 is the operating cost per unit of time; usually includes drivers' salaries and capital payments.

$cd_0{}^{\tau}$ is the operating cost per unit distance of a vehicle of operator *o* when traveling along a link of type τ; usually includes tires, spares, maintenance, oil, and others.

The cost of energy may be simply included as part of the distance related element , using an average consumption figure. Alternatively, the model can estimate energy consumption in each link as a function of speed:

$$Ed^{lo} = ed^n_{\min} + \left(ed^0{}_{\max} - ed^0_{\min}\right) \times \exp\left(-\delta^0 V^0_l\right) \tag{26}$$

where: Ed^{lo} is the energy consumption per unit distance of a vehicle of operator *o* traveling along link *l*

$ed^0{}_{\min}$ is the minimum consumption of energy per unit distance when a vehicle of operator *o* travels at free flow speed

$ed^0_{\max}$ is the minimum consumption of energy per unit distance when a vehicle of operator *o* travels at a speed close to zero

V^0_l is the speed of vehicle of operator *o* in link *l*, after capacity restriction

δ^o is a parameter regulating the steepness of the energy consumption curve

The function is negative, because as speed increases energy consumption is reduced. Next the model must be supplied with the unit cost of energy pe^0, and calculates the cost of energy as $Ed^{lo} * pe^0$. The result is added to the rest the distance-related elements $cd_0{}^{\tau}$. Figure 8 shows examples of typical energy consumption curves for two different operators

B.4.8 Maintenance costs

Administrators pay for fixed and marginal maintenance costs per unit distance. Fixed costs include routine maintenance and any other, assuming that no vehicles use the infrastructure. Marginal costs represent the maintenance costs attributable to each additional vehicle traveling along the link per unit distance. Each administrator *a* is in charge of particular set of link types T^a . If *lt* is a link of the set L^{τ} of links of type τ, the cost of maintenance of administrator

$$cm^a = \sum_{\tau \in T^a} \sum_{lt \in L^t} [mf\tau \cdot d^{\tau}_l + \sum ma_{\tau}{}^0 \cdot Ve^0_l] \tag{27}$$

where: $mf\tau$ is the fixed maintenance cost per unit distance of links type τ

d^{τ}_l is the distance of link *l* of type τ

$ma_{\tau}{}^0$ is the marginal cost of maintenance of links type τ per vehicle of operator *o*

$Ve_l{}^0$ is the number of vehicles of operator *o* traveling along link *l*

B.4.9 Structure of the transport model

The transport model follows a calculating sequence as described in Figure 9.

Figure 9: Structure of the transport model

B.4.10 Path building

The purpose of the path building procedure is to derive a set of travel options from an origin to a destination by a particular mode. A path is not just a sequence of links, but a sequence of link/operators (or routes) combinations. There might be two paths with identical physical links, but with different operators or transit routes.

Formally, a path may be described as:

lo1, lo2, lo3, ... loz

mi = (li, oi)

m_i denotes a particular combination between a physical link l_i and an operator (or route) o_i along the path sequence. The origin node of l_1 must be the centroid of the origin zone, where operator o_1 is boarded. Along the path sequence, a change from one operator to another may occur, and this introduces the possibility of transfers. Finally, the destination node of l_z must correspond to the centroid of the destination zone of the trip.

During path search, the model calculates the generalized cost for each path, accumulating the following elements for each link/operator combination m that forms part of the sequence:

$$c_{ijp}^{ks} = \sum_{m=1}^{z} RT_m^s + RD_m + TR^s{}_{m-1,m} \quad (28)$$

where c_{ijp}^{ks} is the generalized cost of path *p* from *i* to *j* by mode *k* for the category *s*.

RT are time-related costs that the trip maker needs to cross link *l* by operator *o*; *RD* are distance-related costs in link *TR* is the cost of boarding a new operator or route; this can take place either at the beginning of a trip or when there transfer somewhere along the path, that is, when $o_i \neq o_{i-1}$. In the following paragraphs each one of these cost elements is described; most of the elements used to calculate these costs have already been defined when describing operating costs and tariffs.

Time-related costs:

$$RT^s{}_m = tv_m(tt_0 + vv^s + ct_0tc_0 + p_m + pg_0) \quad (29)$$

where: tv_m is the travel time of operator o in link l, a function of the length of the link and the speed of the operator;

vv^s is the value of travel time of the trip category s;

p_m is a penalizing factor of the link/operator combination *m(l,o)*

pg_0 is a global penalizing factor of operator *o*, similar to mode constants

Distance-related costs:

$$RD_m = d_l(td_0 + cd_0tc_0) \quad (30)$$

where d_l is the length of link l

Transfer costs:

$$TR^s_{m-1,m} = tf_0 + te_m ve^s + cf_0 tc_0 \; ; \quad tf, te, cf \begin{Bmatrix} = 0 \text{ si } m-1 = m \\ != 0\text{si } m-1 \neq m \end{Bmatrix} \tag{31}$$

where: te_m is the waiting time for a vehicle of operator o in link l; applies only to transit

ve^s is the value of waiting time for trip category s

The maximum waiting time for a unit of operator or route o in link l in free flow conditions is the inverse of the frequency. The model uses, as a first approximation, the average waiting time, that is, half the inverse of the frequency:

$$te_m = \frac{1}{2f_m} \tag{32}$$

where f_m is the frequency of the route or operator o (vehicles per unit of time); if the route has an undetermined frequency the operator is of a transit type without specific routes, the model calculates the frequency as a function of demand the average occupancy rate.

In the initial path search by mode, the model uses the average free flow waiting time as indicated above. However, iterative sequence of the transport model, waiting times are adjusted as a function of congestion and the demand/capacity ratio of the units themselves. These adjustments are carried out together with the capacity restriction procedure explained further below. Also in path search the values of travel and waiting times are taken as the average all trip categories that use a particular mode, while in the iterative sequence of the transport model the actual values time of the corresponding categories are used.

The path search algorithm, exclusive to TRANUS, is a multidimensional and multipath procedure. The method also includes a procedure to control for the independence of options, to avoid highly correlated paths with small variations separate distinct options. This method is called overlapping control, and solves the well known problem of attribute correlation among options in multinomial logit models. The method keeps track of the degree of coincidence among competing paths in terms of link/operators combinations; paths with a high degree of coincidence are discarded, thus resulting in clearly distinct trip options.

Overlapping control is achieved through a penalizing factor called the O k factor, a positive number that regulates the degree of accepted dispersion for paths of mode k. With this factor, the path search model proceeds in the following steps:

a. search for the minimum path from i to j by mode k and stores it;

b. penalize by O k the time components of all link/operators m that form part of the path;

c. go back to step a) and iterate until,

d. the minimum path found in a) is identical to any of the paths that got stored.

If O_k =1 the minimum path found in the first search will immediately re-emerge in the second search; in this case the model yields a single path for each O-D pair. As the value of O_k increases, the number of paths that succeed in getting stored also increases. Note that a same operator/link combination may emerge several times in different paths and is penalized over and over again. Also note that no paths will be stored with a generalized cost grater than the cost of minimum path multiplied by O_k.

B.4.11 Disutilities and probabilities

Transport disutilities are a measure of accessibility that influence both location decisions and transport choices. Transport monetary costs form part of disutilities, and are used directly to calculate production costs in the activities model. The transport model keeps track of both disutilities and monetary costs separately.

The generalized cost of each path is calculated as described previously. From these, a composite cost or disutility calculated over all paths first, and then over all modes. If c_{ijp}^{ks} is the generalized cost of travel from an origin *i* to a destination *j* by mode *k*, path *p* for the trip category *s*, the probability of choosing such path is given by the following multinomial logit model:

$$P_{ijp}^{ks} = \frac{\exp(-\gamma^s \tilde{c}_{ijp}^{ks})}{\sum_p \exp(-\gamma^s \tilde{c}_{ijp}^{ks})} \tag{33}$$

where γ^s is a dispersion parameter in the logit path choice model; if big, the path with least generalized cost will have a high probability over all other options; if small, trips will be assigned evenly to all available options.

The marginal generalized cost of path *p*, $\tilde{c}_{ijp}^{ks}$ differs from the normal or straight generalized cost in two ways: first is multiplied by a factor I_p that represents the degree of independence of the path with respect of all other competing paths (overlapped cost), and second, it is divided by the overlapped cost of the best path:

$$\tilde{c}_{ijp}^{ks} = \frac{I_p c_{ijp}^{ks}}{\left(\min_p I_p c_{ijp}^{ks}\right)^{\theta^s}} \tag{34}$$

The value of I_p depends on the number of link/operators *m(l,o)* that path *p* shares with all other competing paths; hence, its value will be bigger as more link/operators are shared:

$$I_p = \sum_p lo_p \tag{35}$$

The composite disutility of mode *k* for trip makers of category *s* is estimated by aggregating over all paths in following way:

$$c_{ij}^{ks} = -\frac{\ln Pg^{ks}}{\gamma^k}\left(\min_p c_{ijp}^{ks}\right)^{\theta^s} \tag{36}$$

This equation is multiplied by the overlapped cost of the minimum option because the marginal utility has been used. Pg^{ks} is defined as a series in the following way:

$$Pg^{ks} = \sum_p G_p \prod_{h=1}^{p-1}(1 - G_h) \tag{37}$$

where function G_p is the numerator of the logit model of equation (33):

$$G_p = \exp\left(-\gamma^s \tilde{c}_{ijp}^{nk}\right) \tag{38}$$

The probability that trip makers s choose a specific mode k to travel from i to j is calculated from the composite disutilities of each mode c_{ij}^{ks} with the following multinomial logit model:

$$P_{ij}^{ks} = \frac{\exp - \lambda^s \tilde{c}_{ij}^{ks}}{\sum_k \exp - \lambda^s \tilde{c}_{ij}^{ks}}, k \in K^s \tag{39}$$

where: λ^s is a dispersion parameter of the logit mode choice model. If big, the majority of trip makers s will choose the mode least disutility; if small, they will spread out in even proportions.

K^s is the set of modes k available to category s; e.g. goods only choose from freight modes, people from passenger modes, etc.

$\tilde{c}_{ijp}^{ks}$ is the marginal disutility of mode k, the result of dividing its disutility by that of the best mode:

$$\tilde{c}_{ij}^{ks} = \frac{c_{ij}^{ks}}{\min_k c_{ij}^{ks\theta^s}}, k \in K^s \tag{40}$$

Finally, the composite disutility for all trip makers s traveling from i y j is estimated aggregating over all modes:

$$c_{ij}^{s} = -\frac{\ln Pg^s}{\lambda^s} \min_k c_{ij}^{ks\theta^s}, k \in K^s \tag{41}$$

This equation is multiplied by the utility of the best mode, because the marginal disutility was used in the calculation.

Pg^s is defined as a series in the following way:

$$Pg^s = \sum_k G_k \prod_{h=1}^{k-1} (1 - G_h) \qquad (42)$$

where function G_k is the numerator of the logit model of equation (39):

$$G_k = \exp\left(-\lambda^s \tilde{c}_{ij}^{ks}\right) \qquad (43)$$

B.4.12 Trip generation

The purpose of the trip generation model is to calculate the number of trips derived from a functional flow estimated the activity location model and transformed into flows by transport category by the land use/transport interface. The number of trips generated by a category s for an O-D pair for a particular time period is a function of the composite disutility calculated in equation (41). It is estimated with an elastic trip generation model that simulates a demand curve:

$$T_{ij}^s = F_{ij}^s \left[v_{\text{min}}^s + \left(v_{\text{max}}^s - v_{\text{max-min}}^s \right) \exp\left(-\eta^s c_{ij}^s \right) \right] \qquad (44)$$

where: v_{min}^s is the minimum number of trips made by category s independent of the composite disutility,

v_{max}^s is the maximum number of trips made by category *s* when the composite disutility tends to zero,

η^s is the elasticity of category *s* with respect to the composite disutility.

In each iteration of the transport model, disutility increases because of congestion. As a result, the number of trips decreases, depending on the elasticity of the trip category. When the system converges to an equilibrium, the difference between the number of trips estimated in the first and last iterations is called repressed demand, that is, the number trips that were not made because of congestion.

B.4.13 Modal split

The modal split model estimates the number of trips of category s that choose mode k, from the modal probabilities equation (45) and the number of trips by category calculated in (44):

$$T_{ij}^{ks} = T_{ij}^s P_{ij}^{ks} \varphi^s + 1 - \varphi^s B^k, \quad B^k = \left\{ \begin{array}{l} 1 \text{ if } k \text{ is public} \\ 0 \text{ if } k \text{ is not public} \end{array} \right\} \qquad (45)$$

where φ^s is the car availability rate for transport category *s*.

Note that the probability applies only to trip makers that have a car available, while transit captive population only chooses between public modes. Also note that the term car availability is used, and not car ownership; a trip maker not own a car but still have a company car or share the trip with others.

B.4.14 Trip assignment

Trips by category and mode are assigned to paths with a multinomial logit model, with the path choice probabilities calculated in equation (33) applied to the trips by mode calculated in (45):

$$T_{ijp}^{ks} = T_{ij}^{ks} \times P_{ijp}^{ks} \quad (46)$$

Once the assignment process has finished for all O-D pairs, categories and modes, the model calculates and displays following results:

> T_m is the demand in the link/operator combination *m(l,o)*, in proper units (e.g. Tons, passengers)
>
> V_m is the number of vehicles traveling along link/operator *m*, applying occupancy rates, except for transit routes with given fixed frequencies
>
> VE_l is the number of vehicles in equivalent units on link *l* estimated as:

$$VE_l = \sum_m V_m eq_m \quad (47)$$

> where V_m are car equivalent rates by operator *o* and link *l*.

Transit operators get a special treatment; if a frequency has been defined for a particular route, the number of corresponding vehicles is a given data that the model assigns directly to the links involved. If, however, a frequency been left undefined, the model calculates the required frequency from demand figures, applies given average occupancy rates and assigns the resulting vehicles to links:

$$f_m = \max_l \frac{T_m}{to^0}, \text{ if the frequency is undefined,} \quad (48)$$

where *to* is the occupancy rate for operator *o*. The capacity *q* of an operator or route in a link is:

$$q_m = f_m \times to^0 \quad (49)$$

The demand/capacity ratio for each operator in a link is:

$$dc_m = \frac{T_m}{q_m} \quad (50)$$

The overall demand/capacity ratio for the link is calculated dividing the equivalent vehicles that share it by the given physical capacity of the link:

$$DC_l = \frac{VE_l}{Q_l} \quad (51)$$

Speeds and waiting time for each operator or route are also presented as a result of the assignment process after capacity restriction. This is described in the following section.

B.4.15 Capacity restriction

The traveling speed of each operator in each link is adjusted at the end of each iteration as a function of the level of congestion. For those links with an undefined capacity, speeds remain constant, equal to the initial free flow speeds. links with a finite, defined capacity, speeds are adjusted with a hyperbolic secant model of the form:

$$V_m^{it} = V_i^0 - \text{sech}[\rho(DC_l)^{\beta}] \quad (52)$$

$$\rho = \text{sech}^{-1}(1 - \alpha) \quad (53)$$

$$\beta = \frac{\ln\left[\frac{\text{sech}^{-1}(\nu)}{\rho}\right]}{\ln\gamma} \quad (54)$$

where: V_m^{it} is the speed of operator *o* in link *l*, for the iteration *it*

V_m^0 is the initial (free flow) speed of operator *o* (iteration 0) in link *l*

DC_l is the demand/capacity ratio for link *l*

α is the given proportion in which the initial speed is reduced when the demand/capacity ratio is =1

υ is the minimum percentage to which the initial speed is reduced; the model uses a pre-set value of 1%

γ is the demand/capacity ratio at which the initial speed is reduced to the minimum value (> 1)

Figure 10 shows an example of a speed reduction function for an operator with an initial speed of 80 Km/hr, with a value of $\gamma = 1.25$. The full line corresponds to a value of $\alpha = 0.7$ and the dotted line to a value of $\alpha = 0.95$. As can seen, when *DC*=0 there is no reduction to the initial speed; as *DC* increases, the initial speed is reduced, until it reaches a value of αV_m^0 when *DC*=1. Beyond *DC*=1, the reduction continues until speed reaches the value of 1% of the initial speed (vV_m^0) when $DC > \gamma$.

Once all links have been processed, the speed to be used in the following iteration is the average between the current speeds and those of the previous iterations:

$$V_m^{it+1} = \frac{V_m^{it} + V_m^{it-1}}{2} \quad (55)$$

The capacity restriction procedure also adjusts waiting times for transit operators. Waiting times may increase for two reasons: the reduction in frequency as a result of reduced speeds, and the demand/capacity ratio on transit units. If the operating speed of a transit operator is reduced because of congestion, it is expected that the time interval between should also increase. Also, as the demand/capacity ratio on the units increases, the probability that awaiting passengers have to wait more also increases. To allow for these adjustments, the model first calculates a congestion factor result of the reduction in operating speed calculated above:

$$f_{c_m} = \frac{V_m^{it}}{V_m^0} \tag{56}$$

The congestion factor will always have a value between zero and one, because the speed after capacity restriction will always be smaller than the initial free flow speed. Waiting time for an operator o due to congestion is adjusted as follows:

$$te_m^{it} = \left[\frac{1}{2f_m fc_m}\right] \times \left[\Theta + \exp(\omega dc)\right], \quad \begin{matrix} dc \leq 1\{\theta = 0, \omega = 0.2\} \\ 1 \leq dc\{\theta = 3, \omega = 0.5\} \end{matrix} \tag{57}$$

The first term of the equation adjusts the frequency of the operator as a function of the congestion factor. The second term adjusts the waiting time as a function of the demand/capacity ratio dc on the units of the operator.

Finally, the resulting waiting times are checked against given minimum and maximum waiting times. For urban applications it is common that the average waiting time is linked to the frequency of the services; for regional applications, when frequencies tend to be very low, say two or three buses per day, waiting times are not related to frequencies because users are aware of timetables.

Like speeds, waiting times are also averaged at the end of each iteration:

$$te_m^{it+1} = \frac{te_m^{it} + te_m^{it-1}}{2} \tag{58}$$

These last calculations end the current iteration. A fresh iteration starts with new estimates of operating costs, generalized costs and disutilities as a result of the adjusted travel and waiting times; these in turn affect trip generation, modal split and assignment, causing a new set of adjustments in capacity restriction. Once all necessary iterations have been performed, the resulting costs and disutilities will affect the location and interaction of activities in a new time period.

Convergence

In the transport model convergence is checked for all links as the percentage difference between the current iteration the previous one, considering two variables: operating speeds and traffic flows. The iterative process ends when such differences are both below a pre-defined convergence criteria. The model reports the worst case links in terms of speeds and traffic flows.

Appendix C: Definition of Socioeconomic Sectors

Table C-1 lists the sectors defined in the statewide and substate models, as well as equivalent commodity, industry, and input-output account classifications. The latter are from the IMPLAN model of the Oregon economy developed as part of this project.

Table C-1: BEA industry code equivalencies for STCC and sectors

Sector	SIC	Iocode[a]	Description	STCC
Agriculture (AGFF)	*02	1	Livestock and livestock products	01
	01	2	Other agricultural products	
	*08,*09	3	Forestry and fishery products	08,09
	*02,07, *08,*09	4	Agricultural, forestry, and fishery services	
Mining & construction (CONS)	10	5+6	Metallic ores mining	10
	12	7	Coal mining	11
	13	8	Crude petroleum and natural gas	13
	14	9+10	Nonmetallic minerals mining	14
	15-17	11	New construction	
		11+12	Construction	
		12	Maintenance and repair construction	
*Other manufacturing (OMFG)	34,37	13	Ordnance and accessories	19
	20	14	Food and kindred products	20
	21	15	Tobacco products	21
	22	16	Broad and narrow fabrics, yarn and thread mills	22
		17	Miscellaneous textile goods and floor coverings	
		18	Apparel	23
	23	19	Miscellaneous fabricated textile products	
Wood products (WOOD)	24	20+21	Lumber and wood products	24
	25	22+23	Furniture and fixtures	25
	26	24	Paper and allied products, except containers	26
		25	Paperboard containers and boxes	
Printing & publishing (PRNT)	27	26A	Newspapers and periodicals	27
		26B	Other printing and publishing	
*Other manufacturing (OMFG)	28	27A	Industrial and other chemicals	28
		27B	Agricultural fertilizers and chemicals	
		28	Plastics and synthetic materials	
		29A	Drugs	
		29B	Cleaning and toilet preparations	
		30	Paints and allied products	
	29	31	Petroleum refining and related products	29
	30	32	Rubber and miscellaneous plastics products	30
	31	33+34	Footwear, leather, and leather products	31
	32	35	Glass and glass products	32
		36	Stone and clay products	
	33	37	Primary iron and steel manufacturing	33
		38	Primary nonferrous metals manufacturing	
	34	39	Metal containers	34

Table C-1: BEA industry code equivalencies for STCC and sectors (Continued)

Sector	SIC	Iocode[a]	Description	STCC
*Other manufacturing (OMFG) (continued)	34 (continued)	40	Heating, plumbing, and fabricated structural metal products	34 (Continued)
		41	Screw machine products and stampings	
		42	Other fabricated metal products	
Technology manufacturing (TECH)	35	43	Engines and turbines	35
		44+45	Farm, construction, and mining machinery	
		46	Materials handling machinery and equipment	
		47	Metalworking machinery and equipment	
		48	Special industry machinery and equipment	
		49	General industrial machinery and equipment	
		50	Miscellaneous machinery, except electrical	
		51	Computer and office equipment	
		52	Service industry machinery	
	36	53	Electrical industrial equipment and apparatus	36
		54	Household appliances	
		55	Electric lighting and wiring equipment	
		56	Audio, video, and communication equipment	
		57	Electronic components and accessories	
		58	Miscellaneous electrical machinery and supplies	
	37	59A	Motor vehicles (passenger cars and trucks)	37
		59B	Truck and bus bodies, trailers, and motor vehicles parts	
		60	Aircraft and parts	
		61	Other transportation equipment	
	38	62	Scientific and controlling instruments	38
		63	Ophthalmic and photographic equipment	
*Other manufacturing (OMFG)	39	64	Miscellaneous manufacturing	39
*Transportation, communications, and public utilities (TCPU)	40,41,*47	65A	Railroads and related services; passenger ground transportation	
	42,*47	65B	Motor freight transportation and warehousing	
	44	65C	Water transportation	
	45	65D	Air transportation	
	46,*47	65E	Pipelines, freight forwarders, and related services	
	48	66	Communications, except radio and TV	
		67	Radio and TV broadcasting	
	49	68A	Electric services (utilities)	
		68B	Gas production and distribution (utilities)	
		68C	Water and sanitary services	
Wholesale (WLSE)	50,51	69A	Wholesale trade	
Retail (RETL)	52-57,59	69B	Retail trade	
*FIRE	60-62,67	70A	Finance	
	63,64	70B	Insurance	
*Households		71A	Owner-occupied dwellings	

Table C-1: BEA industry code equivalencies for STCC and sectors (Continued)

Sector	SIC	Iocode[a]	Description	STCC
*FIRE	65	71B	Real estate and royalties	
Services (excludes education [SIC 82])[b] (SERV)	70	72A	Hotels and lodging places	
	72,76	72B	Personal and repair services (except auto)	
	*73	73A	Computer and data processing services	
	81,*87,89	73B	Legal, engineering, accounting, and related services	
	*73,*87	73C	Other business and professional services, except medical	
	*73	73D	Advertising	
	58	74	Eating and drinking places	
	75	75	Automotive repair and services	
	78,79	76	Amusements	
	80	77A	Health services	
	82-84, 86,*87	77B	Educational and social services, and membership organizations	
*TCPU	43	78	Federal Government enterprises	
*Government (GOVT)	Adjust-ments[c]	79	State and local government enterprises	
		80	Noncomparable imports	
		81	Scrap, used and secondhand goods	40
*Government (GOVT)		82	General government industry	
		83	Rest of the world adjustment to final uses	
*Households		84	Household industry	
		85	Inventory valuation adjustment	
	Income[c]	88	Compensation of employees	
		89	Indirect business tax and nontax liability	
		90	Other value added	
*Households	Final demand[c]	91	Personal consumption expenditures	
		92	Gross private fixed investment	
		93	Change in business inventories	
		94	Exports of goods and services	
		95	Imports of goods and services	
*Government (GOVT)		96	Federal Government purchases: National defense[d]	
		97	Federal Government purchases: Nondefense	
		98	State and local government purchases: Education	
		99	State and local government purchases: Other	

a. Based on the 1987 benchmark input-output accounts; these codes may change in the 1992 and subsequent accounts.

b. There is a bit of slippage here: Education (SIC 82) is grouped with government, but in the input-output accounts it is lumped into Organizations. Furthermore, government (SIC 91-99) isn't called out explicitly in the input-output accounts. Government is assumed to be subsumed under Iocodes 79, 82, and 96-99.

c. As defined in the benchmark input-output accounts; not necessarily comparable to the definitions used in de la Barra (1989) and in the Tranus documentation.

d. Not included in the IMPLAN model or the input-output accounts for Oregon.

Appendix D: Zonal Estimates of 1990 and 1995 Land Values

Table D-1: 1990 and 1995 average land values by zone ($\$/ft^2$)

Zone	1990 average land values ($/SQfoot)							1995 average land values (in 1990 dollars)						
	SFD	RUR	MFD	COM	IND	AGR	FOR	SFD	RUR	MFD	COM	IND	AGR	FOR
100	11.902		35.635	26.016	18.353			7.520	0.000	23.943	28.065	16.327	0.000	0.000
101	1.044		5.176	2.562	1.793			4.530	0.000	10.414	7.078	3.961	2.612	0.019
102	0.680	0.929	1.126	4.075	0.854	0.247	0.382	0.458	0.366	2.891	0.760	2.501	0.291	0.332
103	0.233	0.233	8.395	1.237	2.200	0.021	0.023	1.598	0.366	5.279	0.760	2.501	0.291	0.332
104	1.876	0.830	0.391	4.075	0.555	0.210	0.830	1.152	0.510	2.891	3.372	0.391	0.286	0.333
105	0.750	0.740	1.126	8.343	0.854	0.233	0.740	1.209	0.424	2.891	3.372	0.582	0.270	0.333
106	1.987	1.418	1.126	4.075	0.854	0.292	0.206	1.088	0.468	2.891	3.372	0.582	0.290	0.179
107	3.238	1.631	8.449	4.075	0.854	0.343	0.356	1.946	1.213	7.519	3.372	0.582	0.298	0.310
108	7.522		13.278	12.025	9.628			10.734	0.000	17.588	6.990	4.606	0.000	0.000
109	7.933		16.783	10.202	8.738			9.023	0.000	14.547	6.990	4.606	0.000	0.000
110	2.107		10.819	1.232	2.200			5.973	0.000	9.990	6.990	4.606	0.000	0.000
111	1.232	1.232	4.643	0.603	2.200	0.021	0.023	5.509	0.000	9.403	3.253	3.135	0.030	0.010
112	0.161	0.161	8.395	1.237	2.200	0.021	0.023	0.611	0.611	9.403	3.253	3.135	0.030	0.010
113	2.522	1.510	1.126	3.865	1.969	0.427	0.687	1.308	0.510	2.891	3.288	1.327	0.371	0.598
114	1.746	1.243	1.126	4.075	0.336	0.195	0.302	0.979	0.125	2.891	3.372	0.218	0.170	0.263
115	4.460		2.886	4.787	1.947			7.492	1.536	7.505	7.865	1.099	0.000	0.000
116	0.395	0.246	2.886	5.524	1.947	0.021	0.023	2.521	0.412	7.505	7.865	1.099	0.120	0.144
117	1.563		0.879	3.052	1.947			5.340	1.536	7.903	9.299	1.099	0.000	0.000
118	0.789	0.270	2.886	4.787	1.947	0.021	0.023	4.287	0.810	1.992	7.865	1.142	0.007	0.127
119	1.522	1.522	2.886	4.787	1.282	0.021	0.023	4.597	1.536	7.505	12.314	1.099	0.030	0.147
120	0.246	0.246	2.886	4.787	1.947	0.021	0.023	0.987	0.987	7.505	7.865	0.517	0.026	0.011
121	0.270	0.270	2.886	4.787	0.212	0.021	0.023	0.932	0.932	7.505	7.865	0.628	0.004	0.015
122	1.887	1.887	2.886	4.787	0.170	0.021	0.023	1.569	1.569	7.505	7.865	0.484	0.030	0.010
200	9.810		1.996	1.713	1.952			9.503	0.000	6.738	4.950	0.409	0.000	0.000
201	1.336	0.565	1.996	0.435	0.594	0.020	0.204	6.105	0.830	6.738	4.492	0.298	0.083	0.077
202	1.211	0.390	1.996	1.713	1.952	0.155	0.070	2.945	1.289	6.738	4.950	0.409	0.145	0.054
203	1.845	0.390	1.006	1.713	1.952	0.031	0.070	3.499	1.289	3.399	4.593	0.409	0.060	0.054
204	1.084	0.565	1.996	2.413	0.594	0.125	0.086	4.600	1.219	6.738	4.492	0.298	0.054	0.077

Table D-1: 1990 and 1995 average land values by zone ($\$/ft^2$) (Continued)

Zone	1990 average land values ($/SQfoot)							1995 average land values (in 1990 dollars)						
	SFD	RUR	MFD	COM	IND	AGR	FOR	SFD	RUR	MFD	COM	IND	AGR	FOR
205	1.273	0.390	2.982	1.713	1.331	0.045	0.070	3.456	1.712	6.738	9.087	0.409	0.242	0.054
206	1.929	0.390	1.996	1.713	1.952	0.068	0.074	5.485	0.913	6.738	4.950	0.409	0.032	0.054
207	2.181	0.390	1.669	0.911	1.952	0.098	0.070	4.573	1.289	4.631	4.950	0.409	0.074	0.054
208	4.528	0.390	0.947	1.713	1.952	0.066	0.070	6.434	1.289	6.738	4.950	0.409	0.106	0.054
209	1.338	0.390	1.895	1.713	1.952	0.157	0.070	3.045	1.443	3.662	4.950	0.409	0.106	0.054
210	1.296	0.390	0.368	1.713	1.952	0.066	0.070	2.719	1.000	6.738	4.950	0.409	0.074	0.002
211	1.929	0.390	1.996	1.713	1.952	0.114	0.111	2.392	1.132	6.738	4.950	0.409	0.102	0.002
300	6.894		3.467	0.949	2.137			13.435	0.000	15.394	11.605	1.753	0.000	0.000
301	0.397	0.246	3.467	0.949	2.067	0.070	0.047	3.735	0.533	1.669	2.598	1.389	0.049	0.037
302	1.421	0.246	3.467	0.949	0.923	0.010	0.047	4.556	0.533	1.669	0.055	1.389	0.130	0.037
303	1.342	0.246	3.467	0.949	1.733	0.051	0.047	6.757	0.533	2.393	2.598	1.389	0.048	0.037
304	0.425	0.246	3.467	0.949	1.733	0.062	0.047	4.463	0.533	1.669	2.598	1.389	0.071	0.037
305	0.937	0.246	3.467	0.949	0.016	0.020	0.083	5.880	0.533	1.669	0.953	1.389	0.032	0.035
306	3.025	0.246	1.955	0.949	1.733	0.005	0.002	9.049	0.533	1.669	3.891	1.389	0.048	0.037
307	1.244	0.246	3.467	0.949	1.006	0.051	0.047	6.112	0.533	1.669	8.263	1.389	0.048	0.037
308	0.581	0.246	3.467	0.949	1.090	0.037	0.083	3.677	0.533	1.669	2.771	1.389	0.114	0.350
400	0.541	0.280	1.229	1.368	0.351	0.035	0.052	3.107	1.215	5.865	10.774	0.478	0.053	0.073
401	0.227	0.102	0.434	0.404	0.022	0.008	0.015	2.831	0.798	6.870	4.447	0.478	0.016	0.041
402	0.260	0.208	0.907	0.768	0.227	0.029	0.026	5.755	0.811	6.870	11.930	0.478	0.142	0.090
500	2.101	0.193	1.480	1.831	0.393	0.015	0.009	3.225	0.296	2.271	2.811	0.603	0.023	0.013
501	7.032	0.276	4.525	7.935	0.909	0.017	0.008	11.353	0.446	7.305	12.810	1.467	0.028	0.013
502	0.404	0.193	0.000	0.779	0.244	0.030	0.002	0.582	0.277	0.652	1.121	0.351	0.043	0.003
503	0.384	0.228	0.453	0.986	0.423	0.044	0.004	0.553	0.328	0.652	1.419	0.609	0.063	0.006
504	0.441	0.160	0.000	0.808	1.142	0.021	0.004	0.635	0.231	0.652	1.163	1.644	0.030	0.005
505	4.781	0.948	1.231	2.642	1.051	0.022	0.010	6.882	1.365	1.773	3.804	1.513	0.032	0.014
506	1.515	0.322	2.310	1.934	1.449	0.018	0.011	2.283	0.485	3.481	2.914	2.184	0.027	0.017
507	1.860	0.237	1.355	1.987	0.230	0.019	0.017	3.233	0.412	2.355	3.453	0.399	0.032	0.029
508	0.149	0.246	3.467	0.949	1.733	0.005	0.024	5.520	0.533	1.669	1.378	1.389	0.043	0.019

Table D-1: 1990 and 1995 average land values by zone ($\$/ft^2$) (Continued)

Zone	1990 average land values ($/SQfoot)							1995 average land values (in 1990 dollars)						
	SFD	RUR	MFD	COM	IND	AGR	FOR	SFD	RUR	MFD	COM	IND	AGR	FOR
509	0.787	0.028	0.638	1.316	0.313	0.011	0.002	1.026	0.037	0.832	1.716	0.408	0.015	0.003
510	0.768	0.027	1.387	1.791	0.714	0.017	0.001	1.107	0.039	2.001	2.584	1.030	0.025	0.002
511	0.829	0.026	1.573	1.764	0.275	0.017	0.001	1.161	0.037	2.202	2.470	0.384	0.023	0.002
512	0.788	0.025	1.494	1.675	0.261	0.016	0.001	1.161	0.037	2.202	2.470	0.384	0.023	0.002
513	0.722	0.023	1.370	1.536	0.239	0.015	0.001	1.161	0.037	2.202	2.470	0.384	0.023	0.002
600	1.126	0.250	0.484	0.589	0.099	0.023	0.066	1.507	0.334	0.647	0.788	0.132	0.030	0.089
601	0.571	0.190	0.190	0.250	0.132	0.028	0.022	0.928	0.309	0.309	0.406	0.214	0.046	0.036
602	1.550	0.566	1.140	0.951	0.244	0.093	0.061	1.908	0.697	1.404	1.171	0.300	0.114	0.076
603	1.282	0.460	0.429	1.036	0.177	0.079	0.271	2.051	0.736	0.687	1.658	0.283	0.126	0.433
604	0.307	0.261	0.777	0.150	0.192	0.016	0.011	1.071	0.632	2.572	0.899	0.311	0.022	0.054
605	0.479	0.372	0.749	0.841	0.334	0.039	0.041	2.851	1.815	1.164	0.899	0.519	0.098	0.128
606	0.128	0.271	0.777	0.280	0.015	0.018	0.009	2.085	0.528	0.901	0.899	0.022	0.037	0.074
607	0.501	0.382	1.786	0.688	0.280	0.024	0.034	2.269	0.520	0.901	0.899	0.418	0.070	0.072
608	0.369	0.337	0.777	0.525	0.057	0.028	0.072	1.594	0.931	0.901	0.899	0.088	0.059	0.115
609	2.818	0.390	1.635	1.713	1.952	0.187	0.070	3.564	1.289	6.738	1.748	0.409	0.157	0.054
610	0.805	0.565	1.996	0.511	0.594	0.040	0.086	3.877	0.706	6.738	6.054	0.298	0.057	0.078
611	2.363	0.390	0.895	3.317	3.817	0.067	0.028	3.877	0.706	7.573	1.400	0.409	0.084	0.081
612	0.463	0.565	1.996	2.475	0.594	0.037	0.086	4.000	0.728	6.738	5.558	0.298	0.049	0.060
613	0.468	0.390	1.687	1.713	1.952	0.059	0.033	2.970	0.724	5.729	1.781	0.409	0.063	0.081
614	0.515	0.391	0.271	0.265	0.038	0.017	0.003	2.338	0.632	1.260	2.604	0.071	0.050	0.027
615	1.607	0.291	1.075	1.720	0.564	0.028	0.017	7.018	1.605	7.751	3.422	0.999	0.102	0.095
616	0.819	0.238	0.480	1.058	0.309	0.020	0.026	4.663	0.632	1.260	2.678	0.612	0.102	0.020
617	0.753	0.195	0.534	0.843	0.147	0.016	0.011	6.158	0.690	1.260	0.346	0.269	0.048	0.020
618	0.832	0.231	0.000	0.731	0.060	0.015	0.013	3.556	1.453	7.751	3.422	0.090	0.061	0.020
619	0.637	0.164	0.183	0.496	0.114	0.012	0.008	3.506	0.784	1.260	2.604	0.202	0.084	0.020
620	0.542	0.215	0.372	0.835	0.121	0.021	0.006	2.704	0.623	1.260	3.695	0.171	0.225	0.020
621	1.656	0.246	1.508	0.949	1.723	0.109	0.037	4.531	0.495	1.669	2.522	1.389	0.050	0.054
622	0.576	0.246	3.467	0.949	1.733	0.034	0.086	1.889	0.533	1.669	2.598	1.389	0.077	0.046

Table D-1: 1990 and 1995 average land values by zone ($\$/ft^2$) (Continued)

Zone	1990 average land values ($/SQfoot)							1995 average land values (in 1990 dollars)						
	SFD	RUR	MFD	COM	IND	AGR	FOR	SFD	RUR	MFD	COM	IND	AGR	FOR
623	0.708	0.246	3.467	0.949	1.733	0.051	0.061	4.012	0.533	1.669	2.598	1.389	0.068	0.046
624	0.741	0.246	3.467	0.949	1.239	0.033	0.037	3.162	0.197	1.669	4.461	1.389	0.020	0.028
625	1.333	0.246	3.467	0.949	2.862	0.065	0.045	4.995	0.533	1.669	2.598	1.389	0.037	0.036
626	0.934	0.246	3.467	0.949	1.698	0.034	0.028	3.200	0.533	1.669	2.598	1.389	0.042	0.022
627	0.380	0.039	0.339	0.978	0.133	0.015	0.004	0.495	0.051	0.441	1.273	0.173	0.019	0.005
628	1.225	0.051	1.058	2.172	0.462	0.024	0.003	1.550	0.064	1.340	2.750	0.585	0.030	0.004
629	0.330	0.038	0.378	0.692	0.185	0.016	0.002	0.387	0.044	0.444	0.812	0.217	0.019	0.003
630	2.536	1.204	1.230	2.108	1.013	0.098	0.079	3.445	1.635	1.670	2.863	1.376	0.133	0.108
631	0.309	0.138	0.961	0.358	0.185	0.019	0.019	2.226	0.949	6.870	3.025	0.478	0.027	0.070
632	0.494	0.167	0.277	0.434	0.403	0.022	0.030	3.600	1.109	6.870	4.447	0.478	0.100	0.056
633	0.966	0.249	0.760	1.031	0.191	0.060	0.069	1.359	0.350	1.069	1.451	0.269	0.084	0.097
634	0.360	0.161	0.326	0.559	0.304	0.021	0.023	4.571	1.277	6.870	6.856	0.478	0.140	0.036
635	0.587	0.285	1.292	1.175	0.213	0.033	0.030	8.238	1.609	6.870	11.380	0.478	0.425	0.147
700	0.202	0.072	0.070	0.133	0.540	0.006	0.003	0.254	0.090	0.088	0.168	0.679	0.007	0.004
701	0.178	0.063	0.061	0.117	0.475	0.005	0.003	0.254	0.090	0.088	0.168	0.679	0.007	0.004
702	1.344	0.080	1.579	2.047	0.446	0.018	0.009	2.048	0.122	2.406	3.118	0.679	0.028	0.013
703	0.587	0.049	0.884	0.516	0.093	0.002	0.001	0.847	0.070	1.277	0.745	0.134	0.003	0.001
704	0.386	0.075	1.020	0.336	0.065	0.005	0.000	0.483	0.094	1.277	0.421	0.081	0.006	0.001
705	0.080	0.079	0.905	0.065	0.019	0.005	0.000	0.118	0.117	1.342	0.097	0.028	0.008	0.001
706	0.281	0.122	0.847	0.395	0.108	0.005	0.001	0.445	0.193	1.342	0.626	0.170	0.008	0.001
707	0.240	0.132	1.143	0.308	0.084	0.007	0.000	0.281	0.155	1.342	0.361	0.099	0.008	0.001
708	0.796	0.027	1.487	2.420	0.924	0.018	0.001	1.093	0.037	2.043	3.325	1.269	0.024	0.002
709	0.796	0.158	0.971	0.819	0.380	0.032	0.006	1.340	0.265	1.634	1.379	0.640	0.053	0.010
710	0.475	0.143	1.223	1.403	0.758	0.014	0.007	0.873	0.264	2.248	2.578	1.393	0.025	0.013
711	0.244	0.097	0.667	0.774	0.023	0.011	0.004	0.585	0.231	1.598	1.855	0.056	0.027	0.010
712	0.290	0.061	0.431	0.237	0.101	0.009	0.004	0.297	0.063	0.441	0.243	0.104	0.010	0.004
713	0.314	0.023	0.659	0.812	0.804	0.006	0.001	0.380	0.028	0.799	0.984	0.975	0.008	0.002
714	1.342	0.428	0.771	1.226	0.769	0.035	0.002	1.458	0.465	0.838	1.332	0.836	0.038	0.002

Table D-1: 1990 and 1995 average land values by zone ($\$/ft^2$) (Continued)

Zone	1990 average land values ($/SQfoot)							1995 average land values (in 1990 dollars)						
	SFD	RUR	MFD	COM	IND	AGR	FOR	SFD	RUR	MFD	COM	IND	AGR	FOR
715	0.370	0.179	0.551	1.306	0.804	0.025	0.002	0.887	0.430	1.320	3.130	1.926	0.059	0.005
800	0.047	0.142	0.002	0.013	0.034	0.105	0.002	0.038	0.115	0.002	0.010	0.028	0.085	0.002
801	0.638	0.189	0.792	1.018	0.198	0.011	0.001	0.781	0.231	0.969	1.247	0.242	0.014	0.002
802	0.691	0.219	0.984	0.738	0.206	0.012	0.001	0.839	0.266	1.195	0.896	0.251	0.015	0.002
803	0.711	0.593	0.852	1.069	0.370	0.073	0.001	0.997	0.832	1.195	1.499	0.518	0.103	0.002
804	0.701	0.419	0.913	0.914	0.294	0.045	0.001	0.918	0.549	1.195	1.197	0.384	0.059	0.002
805	0.934	0.147	1.566	2.042	0.735	0.019	0.002	0.966	0.152	1.619	2.110	0.760	0.020	0.002
806	0.821	0.411	1.177	1.510	0.535	0.051	0.001	0.981	0.492	1.407	1.805	0.639	0.061	0.002
807	2.372	0.843	0.820	1.571	0.652	0.063	0.002	2.536	0.901	0.877	1.680	0.697	0.068	0.002
808	1.483	0.587	0.963	1.469	0.563	0.054	0.001	1.759	0.696	1.142	1.742	0.668	0.064	0.002