Final Report

Revised Design of a Second Generation Land Use-Transport Model for Oregon

Transportation and Land Use Model Integration Program
Task 1C

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

The University of Washington
ECONorthwest
KJS Associates, Inc.

1 February 2000

A design of the second generation model was carried out earlier this year, largely shaped by the collective experience of the study team and the consultant proposal for the second generation work. We presented the design to the Peer Review Panel in July, 1999. Their extensive comments and concerns led us to reassess the goals and objectives set for the second generation modeling work, as well as some of the theoretical and practical aspects of the proposed design. The consultant and ODOT project managers met subsequently to carry out a thorough review of the peer review panel comments, including some submitted after the meeting. A number of fundamental design decisions were made during these deliberations, which are described and expanded upon in this paper. Note that the proposed models are described in general detail. A detailed model specification will provide further information on the model design.

An earlier draft of this paper was distributed internally to the study team and the Peer Review Panel. The paper has been improved in several places because of their comments. A complete listing of the comments, as well as the response from the ODOT and consultant project managers, may be found in the Appendix.

The client's own review of the discussions and recommendations of the peer review panel led them to impose new design goals for the second generation model:

- The model should operate at a single geographic scale, roughly equivalent to traffic analysis zones in the metropolitan areas and Census tracts outside of them.
- The transportation, land use, and economic modeling components must be fully integrated. Environmental modeling components should be integrated as well, although the extent and manner of doing so is not clear at this time.
- The model should be fully dynamic.
- The model should be a hybrid equilibrium (for transportation and economic markets) and disequilibrium (for land markets and activity interactions) formulation.
- The travel model should be activity-based.
- The data required for the model must be affordable, both in terms of time and money.

As with the first generation model design, our focus turned initially to defining the scope of the model. In other words, what will the model be used for, and what questions must it be capable of addressing? Our first generation modeling design was guided by eleven key planning issues that the models need to address. Some of the issues were relevant only at certain levels of analyses (statewide, substate, or metropolitan). These issues are listed in Table 1. One quickly appreciates that a single model cannot address the wide range of issues and geographic scales at which they manifest themselves. This led us to pursue parallel tracks in model development, using TRANUS (de la Barra, 1989, 1998) at the statewide and substate levels, and the original development of UrbanSim (Waddell, 1997) at the metropolitan level. It was thought at that time that the two tracks could eventually be brought back together to form a consistent and scalable framework capable of addressing a wide variety of planning issues, including but not limited to those listed in Table 1.

We discovered in the first generation work that the cost of parallel development was unsustainable. The merits of the approach notwithstanding, we simply could not afford to fund separate efforts. Moreover, it became clear that rather than moving the two modeling streams together they are actually moving considerably further apart. To overcome these problems we postulated a very flexible modeling framework that would scale to a variety of geographic and temporal resolutions, depending on the needs of the analyst. The peer reviewers soundly rejected the validity of such an approach, suggesting that different models were appropriate at different spatial scales. Some reviewers suggested that the bi-level modeling approach should be retained, with UrbanSim or a similar model operating at a microscopic scale and a regional model operating at a macroscopic scale. This approach would permit the various social and physical systems included in the model to be represented at the scale and resolution most appropriate for each. Other advantages cited included the ability to use different modeling approaches for each subsystem, allowances for different temporal intervals for different subsystems, the ability to maximize our current investment in UrbanSim, and expected lower cost of development.

The bi-level approach makes sense for ODOT only if the anticipated applications fall into all three levels of geographic scale (statewide, substate[1], and metropolitan). We reviewed past, current, and foreseeable projects and programs that the Transportation Planning Analysis Unit (TPAU) has supported. We found that the most prevalent need was for support of major corridor studies, in which the study area typically includes several municipalities. The ongoing I-5 corridor studies, some of which cover the entire Willamette Valley, are typical applications.

We concluded that a single modeling framework and geographic scale, combining the capabilities of the statewide and substate models would best serve the needs of the TPAU. Thus our original goal of providing a flexible and scalable model has been replaced by a single model capable of addressing issues at the district level. The districts correspond to zone ensembles inside the metropolitan areas (which have operational travel forecasting models) and Census tracts outside of them. Beyond Oregon, a collar of zones will provide detail at a sufficient level to reduce boundary effects. The zones within Oregon and the collar will be considered internal zones, and external zones at larger levels of aggregation will be developed. A modeling system with approximately 3,000 zones is anticipated, most of which will lie within the Willamette Valley.

The peer reviewers also suggested adopting a fully dynamic modeling structure, as opposed to the proposed quasi-dynamic approach. The distinction between the two is somewhat artificial. Both begin with a static population of households and businesses and moves them through time. Changes in the composition of the population are modeled in each time interval, which includes the formation, death, or migration of a subset of the population. A fully dynamic model follows discrete households and businesses through time, changing their composition as they age. The quasi-dynamic approach we proposed incorporated the same changes, but did so with an aggregate representation of the population. This implied the need to regenerate the entire population, with the appropriate characteristics, in each time interval. While we believe that a robust implementation of both approaches would yield very similar final states, we have partially accepted the recommendation of the peer reviewers to adopt a dynamic approach. Our revised framework includes a fully dynamic representation of households, but for the time being retains an aggregate representation of businesses. Our proposed approach is discussed in detail below.

1. In the context of this paper, the substate area refers to the Willamette Valley, stretching from south of Eugene-Springfield to the Portland area (including Vancouver, Washington) and from the Coastal Mountain range east to the Cascade Mountain range.

Principal Modeling Components

The proposed model retains most of the same components originally suggested for the second generation model. These components and their interrelationships are illustrated in Figure 1. Macroscale economic and demographic models provide statewide activity totals, based upon national and regional trends. Information from other macroeconomic models and demographic forecasts can be included in this component of the model. These areawide total activities are allocated to households and businesses (production) in each zone. A land development module, very similar to the one in UrbanSim, is used to represent the actions and effects of the development of land and space in response to market conditions. An interactions model links up factors of production and consumption in each of the zones, including imports and exports. These activities result in the flow of people and commodities between zones. Finally, the transportation model generates trips from the demographic and economic flows and routes them through a multimodal transportation network. This framework encompasses the major urban systems that are relevant for TPAU analytical planning work, and are typical of those found in most integrated urban system models (Wegener, 1999a).

The model will step through time in one year intervals. The temporal dynamics inherent in the model are shown in Figure 2. Note that the relationships between modules shown in the Figure are somewhat at odds with classical depictions of integrated land use-transportation models, where only the output from the transportation component feeds into activity location component of the succeeding iteration. Explicit linkages between components across time (shown in different colors in Figure 2) are necessary in a dynamic modeling framework, and even more important than in a quasi-dynamic framework.

The lagged effect of the transportation costs and disutilities until the next time interval is still present in this model, although it is not clear whether the effects in one year are immediately acted upon in the next year. A multi-year lag might be a more suitable construct, and will be investigated as part of the model development process.

Economic Modeling Component

Only a cursory treatment of the economic subsystem was carried out in the first generation modeling work. An explicit model of economic and demographic activities will be included in the second generation work. The economic model will provide forecasts of statewide changes in production by industry and trade flows between the model region and external regions. In the production allocation model, increases in production are added to an “active pool” and assigned to locations where they consume space, and decreases in production are assigned to locations where they free up space.

What we need out of an economic model is, for each year:

- Production and employment by sector for the model region
- Exports by sector from the model region to each of the other region
- Imports by sector to the model region from each of the other regions

To get all of these, we need:

- An econometric forecasting model that accounts for industrial structure
- A way of getting to industry-level output predictions

– A way of getting trade flows by region pairs

The apparent solution is an integrated econometric-multiregional input-output model. Examples of such models have successfully been implemented and provide the types of information we need.

Integrating Econometric And Input/Output Models

A variety of methods have been developed for combining, in one way or another, input/output models with econometric models. Sergio Rey at San Diego State has done a lot of work recently evaluating and refining integration strategies (Rey, 1998, Rey, 1997, Rey and Dev, 1997). He divides integration strategies into three categories:

– Embedding. An embedded model is an econometric model with embedded I/O tables, which provide prior information about interindustry linkages. Embedded models are used for forecasting and their output typically is in the form of employment (and perhaps income) by highly aggregated industry.
– Linking. Linked models feed the output of one model into the other model. The links may go either way (but not both ways). Linked models where the I/O model feeds the econometric model typically are used for evaluating the impacts of one-time shocks, but tend to overestimate impacts because of double counting (the econometric equations replicate in a dynamic way what the I/O model already accounted for statically). Linked models where the econometric model feeds the I/O model typically are used to disaggregate the forecasts from the econometric model.
– Coupling. Coupled models pass information in both directions, often simultaneously. Dick Conway's WPSM model and the George Treyz, *et al*.'s REMI model are examples. Coupled models provide the greatest degree of integration and allow each model to make the most use of information from the other.

Of the available integration strategies, coupling appears best suited to our effort. An embedded model doesn't produce the disaggregate output we need and a linked model (econometric to I/O) doesn't take advantage of what we know about interindustry linkages in its forecasting equations.

Conway's WPSM model works by solving a set of simultaneous linear equations for each iteration. Some of those equations are columns from the I/O table, the rest are the equations of the econometric model (generally in log-log form). DRI's Trendlong model drives the equations for export demand.

The model also contains equations for four categories of in-region consumption demand (motor vehicles, other durables, non-durable goods, and services), three types of investment demand (residential structures, nonresidential structures, and equipment), and state and local government demand. Income and employment are driven by the output predictions form the I/O relations and feed the estimates of demand (except exports). A coefficient change submodel works on the I/O coefficients between iterations, allowing them to change in response to changes in scale and prices. A diagram of the features of WPSM, adapted from Conway (1990), appears in Figure 3.

Since we'll have a separate demographic model, we won't need to predict population in the economic model, but the demographic model should take into account changes in employment predicted by the economic model. We'll also want to use the demographic forecast in the labor supply equation of the economic model.

Multi-Regional Input/Output Models

There are several ways to incorporate more than one region into an input/output model. The two primary types are interregional input/output (IRIO) models and multi-regional input/output (MRIO) models. In an IRIO model, the coefficients in the A matrix treat each combination of industry and region as if it were a separate industry. So, looking down the column for industry 1 in region A, each coefficient is the number of dollars' worth of input from the combination of industry and region represented by that row used to produce one dollar's worth of output by industry 1 in region A. One has to know not only how much of each input is used to make a unit of output, but also what proportion of each input to each industry comes from each region.

In a MRIO model, the matrices of technical coefficients for each region are strung in blocks along the main diagonal of the A matrix. Another, "C" matrix of trade coefficients defines, for each sector, the proportions of inputs from that sector coming from each region. Each block in the C matrix contains nonzero values only along its main diagonal. A MRIO model can be converted into an interregional model in which the proportions of a particular commodity coming from one region to another are always the same regardless of the importing industry. The Leontief inverse in the MRIO model is $(I\text{-}CA)^{-1}$. MRIO models have the advantage over interregional models of being able to use data that is more available and, if unavailable, easier to impute. Their disadvantage is that they assume that all importing sectors have the same trade patterns for each imported good. The more aggregated the sectors, the less realistic this assumption will be.

Despite the simplification in the MRIO scheme, real data to populate the trade tables generally don't exist. Another weakness in any more-than-one region I/O model is that trade coefficients, like technical coefficients are linear and static. Trade patterns can't change as the various regions grow at different rates and can't respond to changes in trading costs (just as production functions can't change with output levels or changes in factor costs). For use in a long-run, dynamic model, this is an important weakness.

Since the coupled model strategy doesn't use the Leontief inverse, we can have a multi-regional coupled model without having to specify trade tables. That doesn't excuse us from having to predict trade flows, but it does free us to predict them in a way that allows nonlinearity and responds to changes in transport cost. It appears at this point that it would be worth exploiting that freedom.

Based on these conclusions, we propose an economic and demographic modeling framework that incorporates the following major features:

- Use a coupled integration scheme, drawing as much as we can from the experience of the WPSM model, which has been used and refined since 1979.
- Use econometric equations for each region that take into account levels in the other regions.
- Solve all the regions simultaneously.
- Make the entire model area one region and have additional regions for the areas that feed the primary entry points into the model area (I-5 from the north, I-5 from the south, and I-84 from the east) as well as a region for the rest of the US.

– Take the output of the coupled model, which will tell us how much of each commodity is going into and coming out of each region, and use a logit-based model (similar in concept to the interactions model) to figure out how much of each commodity will flow in each direction between each pair of regions (with explicit cross-hauling).
– Use employment forecasts from the economic model to inform the demographic model and use the output of the demographic model to inform the labor supply equation in the economic model.

Possible refinements of this model could allow the developer and redeveloper models to inform the investment equations, and to allow composite costs from the interactions model (which will include composite transportation costs from the transportation model) to feed back into the economic model.

Demographic Modeling Component

The region-wide demographic model will be integrated with the multi-region economic model. It is shown schematically in Figure 4. A labor force participation equation will allow the estimation of the size of the region's labor force from employment by sector and wage rates. Employment-driven in- and out-migration will be predicted from changes in labor force, using equations fit to historical data. Gross in-migration and gross out-migration will be predicted, not just net migration. A separate equation will predict non-employment-related migration, mostly of retirees and students. Migrants will be assigned to income, age, household size, and number of workers categories based on distributions appropriate to their employment sector or non-employment-related status and their mobility status.

The in- and out-migration totals from the economic-demographic model may, if desired, be overridden by exogenous specifications, e.g., from the State's forecast.

The existing population will be aged in place and relatively simple birth and death functions will be applied. Household formation (and reformation) and dissolution models also will be applied to the existing population. As the existing population's attributes are updated, their location in a zone will remain intact. These operations work on the household database that is shared with the Household Allocation model.

The Household Allocation model will then take over. It will first do a migration allocation, estimating a move out of the region probability for each household. The households will be ranked by their likelihood of moving out and out-migrants will be selected from the ranked list until the predicted out-migration has been accounted for. Households on the list that do not match the attribute combinations predicted for out-migrants will be passed over.

The space freed by out-migrants and other households that disappear (e.g, by the death of a member of a one-person household) will be added to the vacant pool.

The Household Allocation model then will assign synthetic households to the in-migrant pool (using the same routine used to create the initial population, perhaps with a seed population drawn from a larger area) and place them, along with households created or reformed from existing households (e.g., by divorce or kids leaving home and establishing new households), into the mover pool. Auto ownership levels could be assigned to households at this stage as well, but those in the mover pool would have to be assigned average accessibility measures.

Household Allocation Modeling Component

A schematic of the household allocation model is shown in Figure 5. It will work as a microsimulation of individual household actions and will provide a fully dynamic representation, where individual households either move or stay in a given time period.

A synthetic population of discrete households must be generated in the base year. The process developed by Beckman, *et al*. (1996) as part of the TRANSIMS project will be used for this purpose. This method involves fusing data from the Public Use Microsample (PUMS) with the Census enumeration of households by various categories. The Standard Tape File 3A (STF3A) is the best source of these data, providing summary tables at the Census tract level. Beckman's synthetic population generator uses an iterative proportional fitting process to generate a population of households that respects the marginal distributions in both the PUMS and STF3A for the study area. Households will be segmented by:

- Household size
- Income
- Lifecycle
- Worker A occupation category
- Worker B occupation category
- Whether a secondary residence is owned

The population generated will be located at the primary residence at the Census tract level, which corresponds to the geography represented in the STF3A. Most of the traffic analysis zones (TAZs) located outside of metropolitan areas will be either tracts or aggregations of tracts, allowing us to populate these zones with households directly. Inside the metropolitan areas, where the zone system is likely to be smaller than Census tracts, we will allocate primary residences to zones within each tract using population estimates incorporated in the metropolitan area travel forecasting model.

The household allocation process in a given year starts with the aggregate changes in total population from the previous year determined in the demographic model. This includes for each household segment:

- Total out-migration
- Total in-migration
- Shifts to other household segments (expressed as a vector of changes where each element in the vector is the number shifting to another segment) reflecting:
  - Deaths in households
  - Departures from households
  - New household formation
  - Household dissolution
  - Changes in Income
  - Changes in Lifecycle
  - Changes in Worker occupation category
  - Changes in whether a secondary residence is owned

Individual households are selected randomly for out-migration, with the probability of out-migration increasing as location utility decreases and consistent with the distribution of characteristics in the total group of out-migrating households identified in the economic and demographic model components. These individual households are then removed from the population and the residential space they occupied, including both primary residence and any secondary residence, is added to the vacant space pool in the appropriate TAZs.

Individual households are then selected randomly from the remaining population for the various shifts to other segments arising from the changes listed above, with probabilities that are functions of various location and socio-economic factors. Any specified marginal distributions for the new population totals are respected by the replacement and re-selection of households as appropriate. At the end of this process, those individual households identified for complete dissolution are removed from the population and the residential space they occupied, including both primary residence and any secondary residence, is added to the vacant space pool in the appropriate TAZs.

Each individual household in the remaining population is then designated as a "mover" or "stayer" with regard to its primary residence in a Monte Carlo simulation process that takes into account location utilities in the current and in other TAZs as well as various socio-economic factors. Those households designated to be movers are added to the primary mover pool and the primary residential space they occupied is added to the 'vacant space pool' in the appropriate TAZs.

Then each individual household in the population segments with a secondary residence is designated as a mover or stayer with regard to its secondary residence in another analogous Monte Carlo simulation process that takes into account location utilities in the current and in other TAZs as well as various socio-economic factors. Those households designated to be movers in this process are added to the secondary mover pool and the secondary residential space they occupied is added to the vacant space pool in the appropriate TAZs.

A population of discrete in-migrating households is then synthesized consistent with the expected marginal distributions for the characteristics of this group using the process developed by Beckman, *et al*. (1996), as was done in the development of the base year population. This population is added to the primary mover pool and to the secondary mover pool as appropriate.

The primary mover pool and the secondary mover pool are then merged into a single mover pool and the households in this mover pool are ordered according to income category and then primary or secondary residence being sought within each income category.

Starting with the highest income category, individual households are allocated to residential locations in a Monte Carlo simulation process that takes into account location utilities and the availability of residential space in the vacant space pools in TAZs, along with primary residential locations as appropriate.[2] When an entire income category is allocated, then the quantities of residential space in the vacant space pools is updated (with the households that are allocated using space at densities consistent with the prices in the TAZs) and the next income category is considered. When the quantity of available vacant space in a given zone is not sufficient for the total demand for a given income category, then the price of that space is increased in that zone and both

2. We realize that this is not simulating the reality of how this process works in the real world. It is proposed as a method of carrying out the allocation. We will examine other approaches in model specification and development, but feel that this provides a tractable starting point for our work.

the designation of movers and the allocation to space for that income category and all lower income categories is updated and the allocation for that income group redone.

After the allocation process is complete, the population of households by zone will be used as input to the interaction models. A parallel process will be used to transition each household into the next time period. The transitioned households will form the starting point for the next time period.

Production Allocation Modeling Component

A fully dynamic representation of production would require the microsimulation of individual businesses and firms. This is beyond our ability to carry out within the constraints of the current contract. We therefore propose to handle production at an aggregate level within zones. Each zone will be populated with production values (in dollars) for each of about 20 sectors based upon industry and white collar versus blue collar workers. The total production for each sector across the entire modeled area will be available from the economic model.

In many respects the framework for the production allocation model mimics that of the household model, although the resolution of each subsystem is considerably different. The process is shown in Figure 6. The flows represented in the production allocation model are measured in total dollar terms, and do not correspond to discrete business establishments. Like the household model, the production model begins with an estimate of migration by sector, derived from the economic model. Establishments that migrate out of the region are added to the vacant space pool for the type of land they consume (office, commercial, or industrial) and their value of production is removed from the model. Production that migrates into the modeled area is initially placed in a mover pool. Each sector currently residing in a zone will decide whether to relocate within the modeled area. A logit formulation will be used, with a composite utility based upon a vector of relative attractiveness for that zone versus other zones. The total value of production that decides to leave the zone is added to the mover pool.

The transition process is simpler than for households, as we only update total production by zone. In later generations of the modeling framework (beyond the current contract) we will move to microsimulating establishments, which will require transition functions for each sector.

Developer Component

A model of actions in the development of new space (floorspace) and the redevelopment of existing space in each TAZ in the step from each time period to the next is included in the modeling system. The operation of this developer model is shown schematically in Figure 7. It provides updated values for the quantity of space in each category in each TAZ in each time period.

An inventory is kept of the quantities of land in each TAZ, keeping track of the amounts of land in categories indicating:

- whether or not the land is developable, and if it is developable the permitted uses and permitted densities; and
- whether or not the land contains developed space, and if it does the type of space it contains (in terms of suitable use categories) and whether or not this space is occupied.

The elements of this inventory related to the first bullet point are altered to reflect changes in land use policy. The elements related to the second bullet point are altered as a result of developer actions in response to conditions in the markets for space and the availability of land for new development and vacant space for redevelopment.

In a given time period, the household allocation and production allocation components establish the prices (rents) for space and the associated occupancy densities in each of the occupied space categories in each TAZ. In the transition from one time period to the next, the quantities of vacant space of each type in each TAZ are considered for alteration in order to simulate developer actions in the re-development of existing vacant space into new vacant space of a different type. This is done for the existing vacant space of each type in each TAZ taking into account:

- the prices for each type of space in the TAZ and in neighboring TAZs as appropriate at the start of the transition;
- the allowable maximum quantity of that type of space for the TAZ consistent with permitted use and density values; and
- overall vacancy rates for that type of space.

After the existing vacant space has been considered, the quantities of vacant space in each TAZ are then increased for the transition from that time period to the next in order to simulate developer actions in the development of new space on vacant land. This is done for each type of space in each TAZ taking into account the same factors listed above concerning re-development.

In a given TAZ, the increases in space of various types are determined in order according to price, starting with the type of space with the highest price. After each increase in space is determined, the maximum allowable quantities and the overall vacancy rates are updated to reflect the changes that have occurred.

The intention is to draw on what has been learned in the development of the UrbanSim model of Eugene-Springfield in the development of the simulation of developer actions.

Activity Interactions Component

A schematic representation of the activity interactions component is shown in Figure 8. The interactions among activities in a given time period are treated in terms of flows of commodity values. These flows are determined based on the amounts of activity in each category allocated to each TAZ by the household allocation model and the production allocation model.

Commodity values are transported from production TAZs to exchange locations to consumption TAZs. An exchange location is the point in space where goods and services are transferred from producer to consumer. It is the geographic location of the relevant market, where the exchange price is determined. It is also the point where the cost (and more generally, the disutility) of transport stops being borne by the producer and starts being borne by the consumer.

For some commodities, the exchange location is the same as the production location. For others, the exchange location is the same as the consumption location. Whereas in general it is possible to have exchange locations that are separate from both production and consumption locations (possibly determined endogenously as part of the model operations); in the Oregon model all exchange locations are either production locations or exchange locations and are exclusively one or the other for a given commodity category, and all exchange locations are considered in terms of the zones containing them. In particular, the exchange location is the consumption TAZ for all primary and secondary manufacturing goods and labor; and the exchange location is the production TAZ for all services, including all wholesale and retail margins. For land — which is treated as if it is a commodity even though it is not *per se* — the production location, exchange location and consumption location are all the same TAZ, and land is said to be 'non-transportable'.

The value of the output commodities produced by a given activity in a given TAZ is allocated among exchange zones using logit formulations with 'selling interchange utilities'. These selling interchange utilities include:

- a size term;
- an inertia term, which is the quantity (proportion) of flow of commodity value from the production zone to the exchange zones in the previous year;
- the exchange price in the exchange zone; and
- the (dis)utility of transporting the commodity from the production zone to the exchange zone.

Similarly, the value of input commodities consumed by a given activity in a given TAZ is allocated to (actually allocated to be coming from) exchange zones using logit formulations with 'buying interchange utilities'. These buying interchange utilities include:

- a size term;
- an inertia term, which is the quantity (proportion) of flow of commodity value from the exchange zone to the consumption zone in the previous year;
- the exchange price in the exchange zone; and
- the (dis)utility of transporting the commodity from the exchange zone to the consumption zone.

For a given production location TAZ and commodity, the composite utility of the selling interchange utilities for the full set of exchange zones (the log-sum value) provides the selling utility for that TAZ and commodity, which is used in the determination of the production location utility for that TAZ for use in the allocation of production activity or moving households, as appropriate, in the next time period. Similarly, for a given consumption location TAZ and commodity, the composite utility of the buying interchange utilities for the full set of exchange zones provides the buying utility for that TAZ and commodity, which is also used in the allocation of production activity or moving households in the next time period.

The buying demand for a given commodity in a given exchange zone is the sum of all the commodity flows allocated to be coming from that exchange zone. The selling supply for a given commodity in a given exchange zone is the sum of all the commodity flows allocated to be going to that exchange zone. The exchange price for a given commodity in a given exchange zone is the one that equates the buying demand with the selling demand in the zone. Prices in exchange zones are adjusted until a set of exchange prices is identified. The quantities of commodities produced and consumed in TAZs are elastic with respect to selling utilities and buying utilities, respectively, which reflects the potential for an under-utilization of production capacity (including unemployment) and allows a set of exchange prices to be identified.

When the exchange location is the same as either the production or the consumption location, then the disutility of transporting commodities between such ‘same locations’ is zero. This translates into a zero disutility rather than the corresponding intrazonal disutility when locations are treated in terms of a system of geographic zones and such ‘same locations’ are ‘same zones’.

The flows of commodity values among TAZs thus determined are passed to the transportation model, where they are converted into goods and services and person movements and loaded to networks of services as appropriate.

Transportation Modeling Component

An activity-based travel demand modeling approach for households is proposed, coupled with an aggregate demand model for travel originating from establishments other than households. This hybrid approach is illustrated in Figure 9, and is necessitated by the different resolutions and techniques used to model households and economic production.

Activity-Based Travel Demand Modeling

The activity-based modeling approach builds upon the work in this area that have been carried out in Portland, although it departs from their model in some significant ways. The Portland model is perhaps best described in a recent paper by Bradley, *et al*. (1999). The framework uses a sample enumeration process to estimate travel activities for each household in the Portland region. The households are synthetically generated in a manner analogous to the previously cited work by Beckman (1996). Travel activities are generated from households with similar characteristics in activity-based travel surveys conducted in the region (the Oregon Travel Behavior Inventory [TBI][3]). Trips and trip chains are developed from the activity lists for each household. A nested logit destination choice model is used to build the origin-destination matrices, which are then allocated to mode of transport, conditioned upon several factors. The resulting flows are then assigned to a multimodal network using equilibrium assignment techniques. A variant of this approach is being developed, whereby the TRANSIMS routing and vehicular simulation models will replace the traditional macroscopic traffic assignment models.

We plan to build upon the approach developed in Portland, reusing as much of the model concepts and structure, as well as their estimation and calibration efforts and computer software developed to implement the model. That will not, unfortunately, obviate the need to estimate and calibrate the model for use in the statewide modeling context. One very important difference is that the statewide model will need to respect the constraints of the activity flows generated by the household and production allocation models, both with respect to the number and types of trips generated as well as their origin-destination patterns. In the Portland model this constraint is imposed after the trip chains are generated. In order for our model to maintain internal consistency, the constraint must be applied during tour generation.

3. The Portland portion of the TBI has been extensively analyzed and reported upon. Less well known, and very fortuitous for the TLUMIP work, is that the same survey was applied in ten Oregon counties, including six outside of the Portland region. The continuity and consistency of these surveys will enable us to build upon the work that others have already done, expanding our scope to all ten surveys.

The process for generating activity and travel lists for each household is carried out in several steps. The first step involves the synthesis of activity lists for each household, based upon the various household strata described earlier (income, size, occupational categories, and life cycle). This step will be carried out using a sample enumeration process, whereby a Monte Carlo simulation will be used to generate activity lists for each household based on the TBI. It is probable that these patterns differ across the state, and we anticipate estimating different patterns for the Portland region, the remainder of the Willamette Valley, and the remainder of the state. A slightly different scheme might be to classify the household location by settlement size: metropolitan (Portland-Vancouver region), urban, and rural.

The observed patterns will be expressed symbolically, using a letter of the alphabet to describe each activity. A vocabulary results by chaining together the activities in their order throughout the day, and can include prefixes and suffixes to denote the time period in which the activities take place, whether travel is associated with the activity, etc. Rasmussen and Barrett (1995) have applied this approach as part of the TRANSIMS development. Once these activity lists, expressed as words, are generated for each household, individual trips and tours can be derived from them using a rule-based system. Again, the relationship between activities and travel will be based upon the relationships found in the TBI, and will substantially replicate the process presently used in Portland.

At this point we will need to impose the additional constraint of respecting the activity interactions generated in the household and production allocation models. These flows, depicted in monetary terms, express the spatial interactions between households and establishments. Production requires labor, which in turn generates a demand for workers by occupational category. Households satisfy these demands, and the resulting dollar flows from the activity interaction models can be translated directly into commuter flows. Because the activity interaction model has already specified the zones from which these flows arise, the origins and destinations are known. The potential problem here is that we cannot be sure that the two principal constraints — the patterns revealed in the TBI and the flows in the activity interaction matrices — can be easily reconciled. Since we will have more confidence in the former than the latter, we will need to condition the activity interaction model to not only replicate the observed economic and demographic patterns but also those depicted in the TBI. It is acknowledged that considerable effort might be expended in this portion of the work.

The resulting tours will include intermediate stops that will also respect the flows in the activity interaction matrices. Once complete, a departure time choice model will allocate the flows to one of four time periods: a.m. peak, mid-day, p.m. peak, and off-peak. The resulting flows will be written to trip matrices by trip purpose, primary mode of transport, and period of the day.

Tour-Based Travel Demand Modeling

The microsimulation framework adopted for household-based travel cannot be used for establishments, as their activities are measured as zonal economic production and consumption rather than as discrete firms. The resulting flows are aggregate in nature. They will be represented in three-dimensional matrices, with origins and destinations in the first two dimensions and commodity in the third dimension. These flows can be translated into passenger and freight movements using factors to convert employment flows (in dollars) to number of employees (using average wage rates by sector and occupation), and to freight flows by using an average payload weight by mode of transport and commodity. This transformation is used in a number of existing packages, including TRANUS and MEPLAN, as well as many statewide and regional models.

In many ways this part of the modeling framework will not depart significantly from the approach used in TRANUS. This is due to our general satisfaction with the process, as well as a lack of data at the establishment end from which to estimate more robust models. One important advance, however, will be the introduction of trip chains. Trips are not linked in the current implementation of TRANUS. The prevalence of trip chaining, however, is too well established in the literature to ignore. The problem is that the flows handled in this part of the model — non-home-based work trips, visitor trips, and goods movement — are ones for which very little data on trip chaining and tour formation are available. We thus anticipate having to construct a fully synthetic model for tour generation. The application software will allow us to bypass the generation of tours completely by specifying an empty rule set, enabling us revert to a more traditional trip-based model implementation if necessary.

The process for synthetically generating truck tours is fairly involved, and beyond the level of detail desired in this paper. The Appendix outlines one potential technique for carrying out tour generation using available data from the truck weigh station surveys, the Commodity Flow Survey, and carrier survey data. It is a microsimulation process, whereby the aggregate flows in the activity interaction matrices are first translated into discrete shipments by commodity, and then combined into truck tours.

The resulting flows will then be allocated to periods of the day (a.m. peak, mid-day, p.m. peak, and off-peak) using the same process as the activity-based model, but the process will be calibrated using appropriate data. In the case of freight movements, for which survey data depicting temporal behavior are not available, a synthetic process will be developed based on observed hourly counts and limited carrier survey data available from Portland METRO and the Port of Portland. The resulting tours will then be collapsed into origin-destination matrices by mode of transport and trip purpose for network assignment.

Transportation Supply Modeling

The trip matrices by trip purpose, period of the day, and mode of transport will be combined together into vehicle trip matrices by period and mode for network assignment. A stochastic user equilibrium assignment process will be used to assign the flows to the multimodal network. The combined model of mode and route choice developed by de la Barra (1994) and implemented within TRANUS holds considerable theoretical and practical appeal, and is recommended for implementation within the second generation models.

Implementation Framework

The software needed to implement the model will be developed using the Java programming language. Multi-threaded applications will be developed where practical, particularly for time-consuming components such as network path-building and assignment. We will make maximum use of the existing UrbanSim model and computer programs to the extent possible, reducing the cost of development in the process. However, it is anticipated that the two models, while starting from a common foundation, will likely continue to develop along separate paths.

We plan to use a component-based architecture for the model, allowing each principal component described above to be encapsulated. JavaBeans are self-describing components that are ideal for this purpose. By encapsulating each component we hope to be able to permit a straight-forward substitution of newer (and different) components as they are developed, without having to rewrite the remainder of the programs. This will permit continued incremental development of the model at a reduced cost, and will grant considerable flexibility for future enhancements. One suggestion made earlier was the eventual development of a microsimulation of establishments. Such is beyond current capabilities, but the use of a component architecture would enable us to replace the production allocation model without rewriting the entire suite when the necessary theoretical advances and data become available.

The peer review panel has placed considerable emphasis on the presentation and visualization of the modeling process and results. We plan to place considerable emphasis on this area in the second generation, focusing on the following areas:

- Our use of geographic information systems (GIS) didn't live up to our aspirations in the first generation modeling work, for a number of reasons. Implementing the second generation models within the ArcView environment would obviously force greater integration with GIS, but it is not clear that such would be either practical or advantageous. Rather than using vendor-specified file formats (as with TRANUS), we'll focus on using ESRI's shapefile as the native format for most data, greatly facilitating our ability to use GIS to visualize data in their native format.
- We will build a process visualizer to display information about the modeling process, showing the user how the model is working and interactive displays of the processes. For example, during the production allocation model both tabular and graphical information about the state of model convergence will be shown. The user will be able to focus on price adjustments in specific zones or by commodities, and quickly switch between them. The results will be stored in XML format so that they can be directly used by other applications.
- We will include the ability to generate MPEG animations of changes in state variables over time. ArcView already provides support for building MPEG files. This will allow us to quickly and efficiently depict changes in zonal or network attributes over time in a format that is amenable to both analysts and decision-makers. An example of such a presentation might be the changes in residential land prices for zones in the Portland region from the base year (1995 or 2000) to the target year (2025 or 2030).

Some of these capabilities might be added by using commercial off-the-shelf software or Java component libraries. We will use open source and public domain software to the maximum extent possible, in order that the entire modeling platform can remain fully open for others to enhance and possibly use elsewhere.

References

Beckman, R., K. Baggerly & M. McKay. (1996) "Creating synthetic baseline populations." *Transportation Research* 30A(6), 415-435.

Bradley, M., *et al*. (1999) "A comparison of sample enumeration and stochastic microsimulation for application of tour-based and activity-based travel demand models." Paper presented at the European Transport Conference, Cambridge, U.K., September 1999.

Conway, R. (1990) “The Washington Projection and Simulation Model: A regional interindustry econometric model.” *International Regional Science Review* 13(1).

de la Barra, T. (1989) Integrated land use-transport modeling. London: Cambridge University Press.

de la Barra, T. (1994) “Dual graph representation of transport networks.” *Transportation Research* 30B(3).

de la Barra, T. (1998) “The mathematical and algorithmic structure of TRANUS.” Unpublished paper available from http://www.modelistica.com.

Donnelly, R. (1999) “A proposed framework for a bi-level urban freight model for the Portland region.” TLUMIP Working Paper 99-3, available from http://www.odot.state.or.us/tdb/planning/modeling/modeling.html.

Rasmussen, S. & C. Barrett (1995) “Elements of a theory of simulation.” Santa Fe Institute Working Paper 95-04-040.

Rey, S.J. (1997) “Integrating regional econometric and input-output models: An evaluation of embedding strategies.” *Environment and Planning A* (29).

Rey, S.J. & B. Dev (1997) “Integrating econometric and input-output models in a multiregional context.” *Growth and Change* (28).

Rey, S.J. (1998) “The performance of alternative integration strategies for combining regional econometric and input-output models.” *International Regional Science Review* (21).

Waddell, P. (1997) “Development and calibration of the prototype metropolitan land use model.” TLUMIP technical report available from http://www.odot.state.or.us/tdb/planning/modeling/modeling.html.

Wegener, M. (1999) “Land use and transport model integration: Progress and future directions.” In Oregon Department of Transportation, “Proceedings of the Oregon Symposium on Integrating Land Use and Transport Models.” Available from http://www.odot.state.or.us/ tdb/planning/modeling/modeling.html.

Table 1: Key modeling analysis requirements

Issue	State-wide	Sub-state	Urban
Effect of land supply on land use and location decisions	●	●	●
Effect of congestion on land use and location decisions		●	●
Cumulative effects of retail location choice			●
Effects of large commercial development at the UGB periphery		●	●
Effect of land supply on travel behavior	●	●	●
Effect of highway capacity increases on travel behavior	●	●	●
Effect of network connectivity on travel behavior			●
Effect of parking supply on travel behavior			●
Effect of urban form on mode choice			●
Effect of rail investment on highway use	●	●	
Effect of changes in the demographic composition of Oregon	●	●	

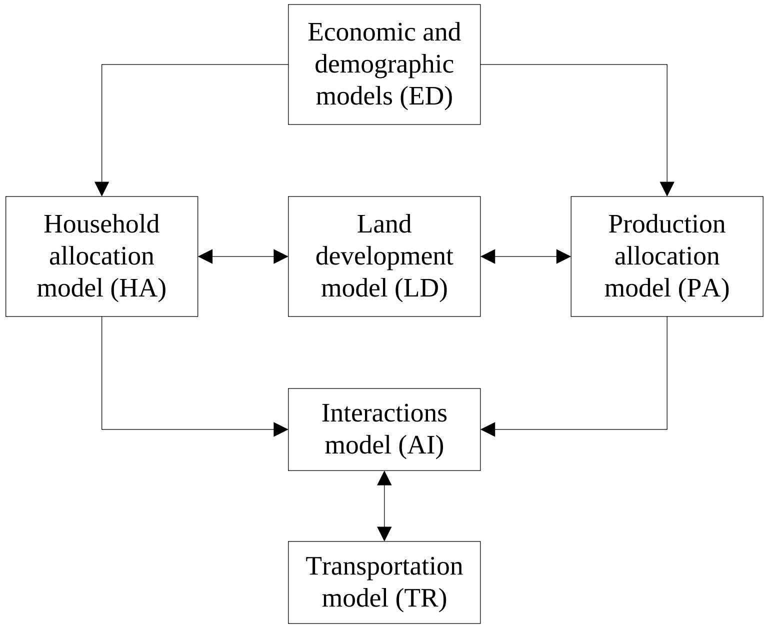

Figure 1: Principal model components

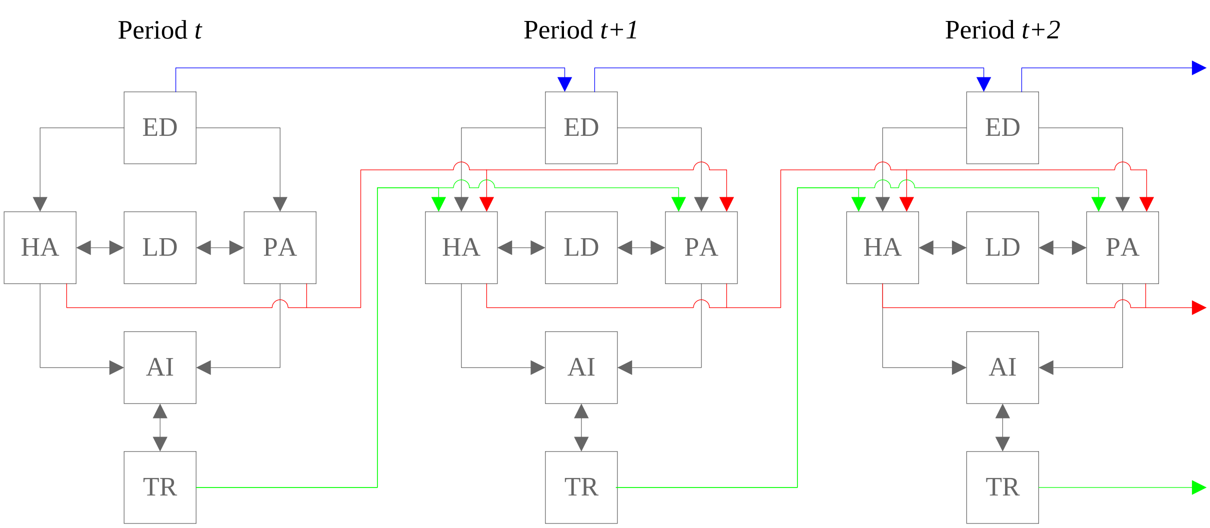

Figure 2: Temporal model dynamics

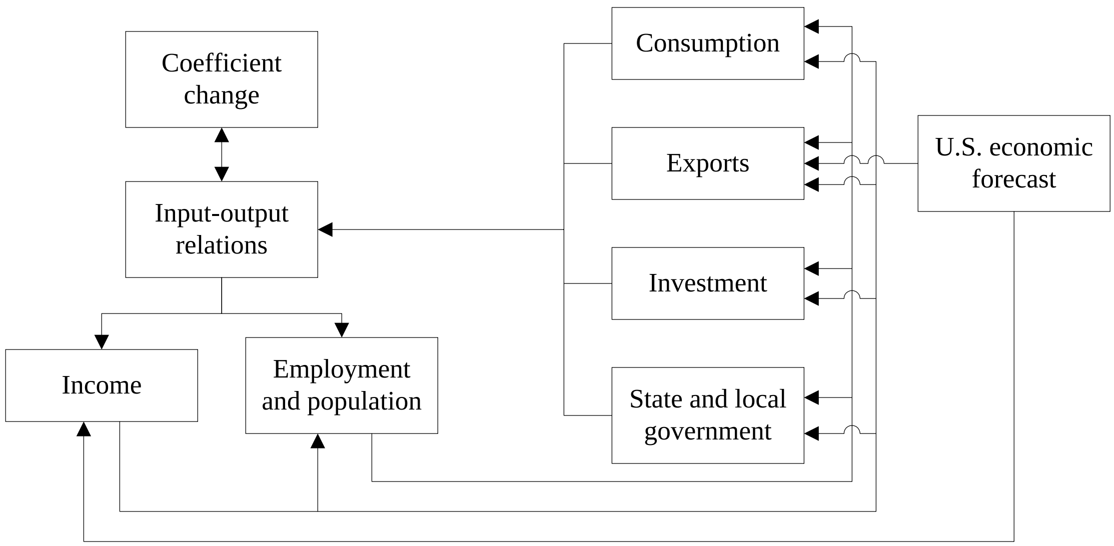

Figure 3: Schematic representation of the Washington Projection and Simulation Model

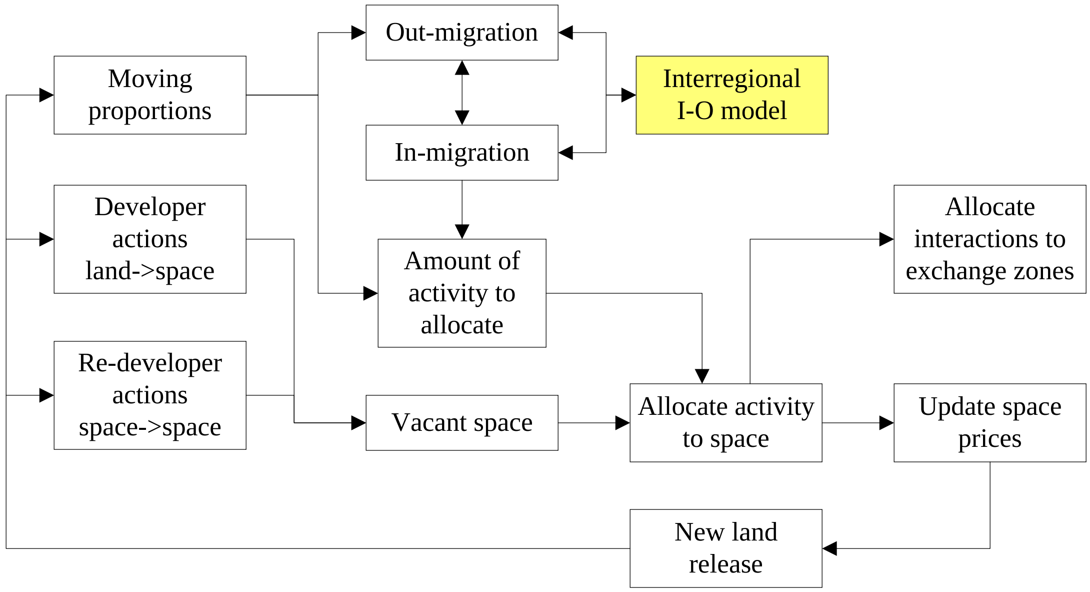

Figure 4: Demographic modeling sequence

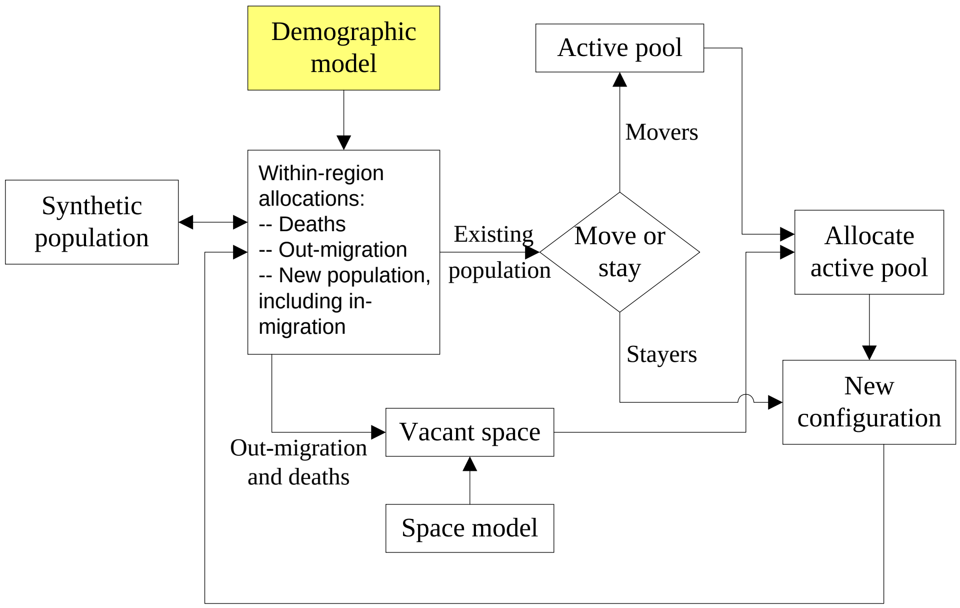

Figure 5: Household allocation model

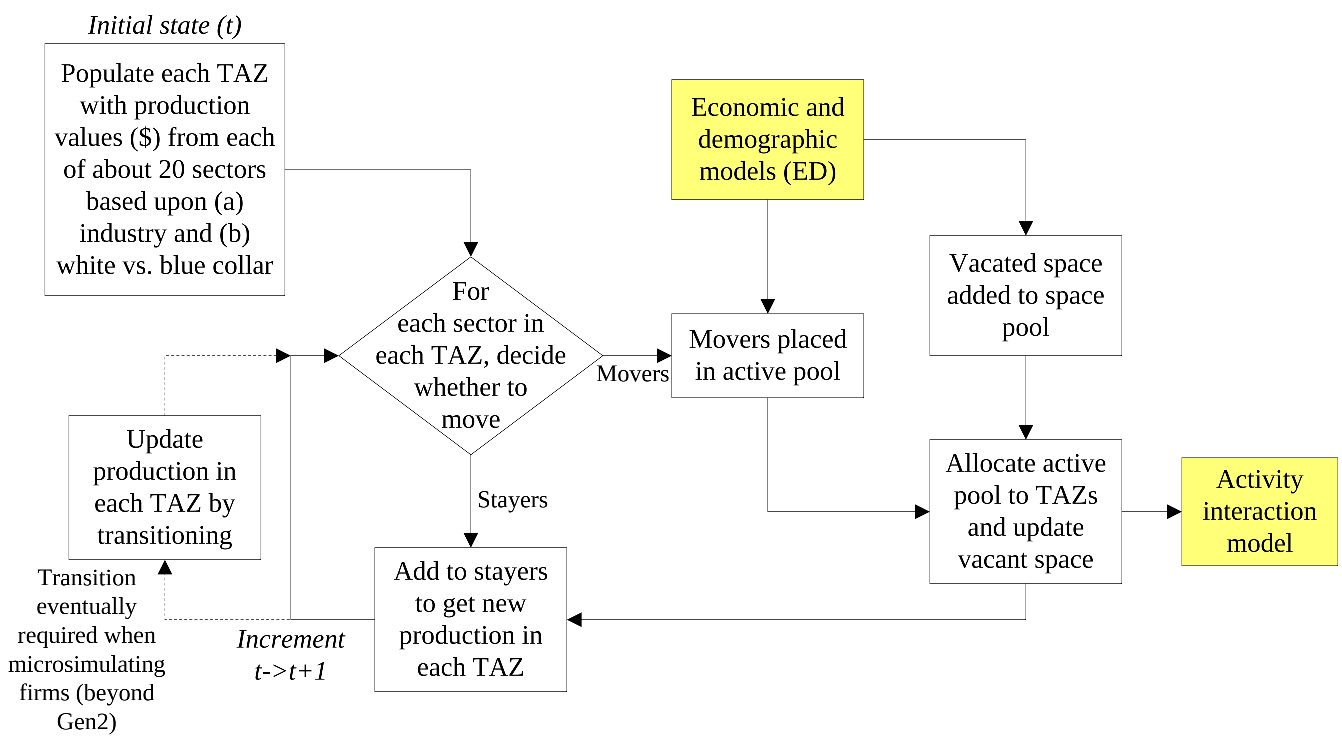

Figure 6: Production allocation model

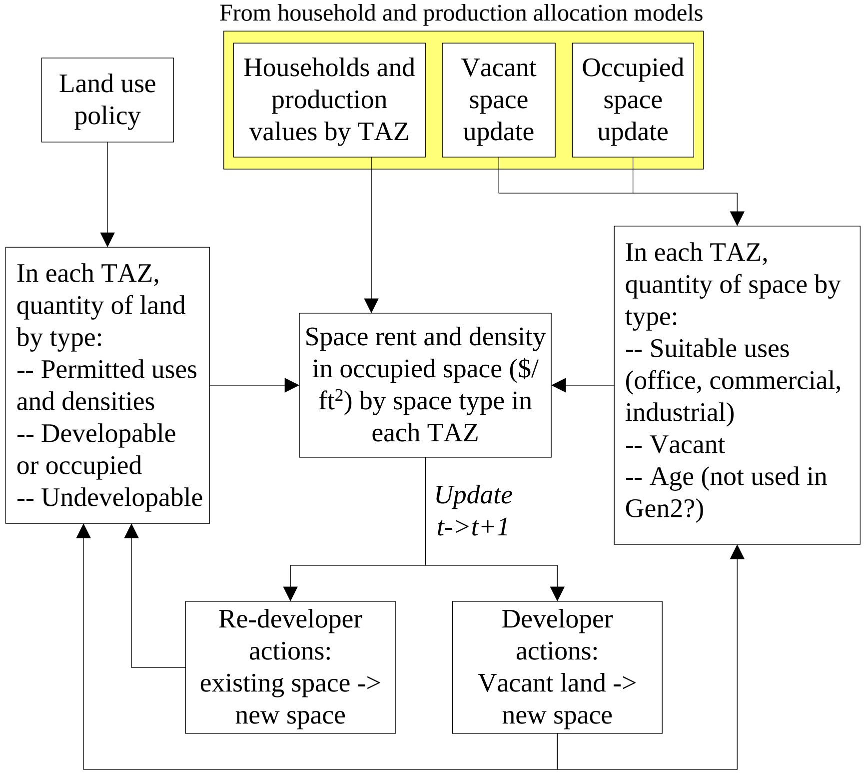

Figure 7: Land development model framework

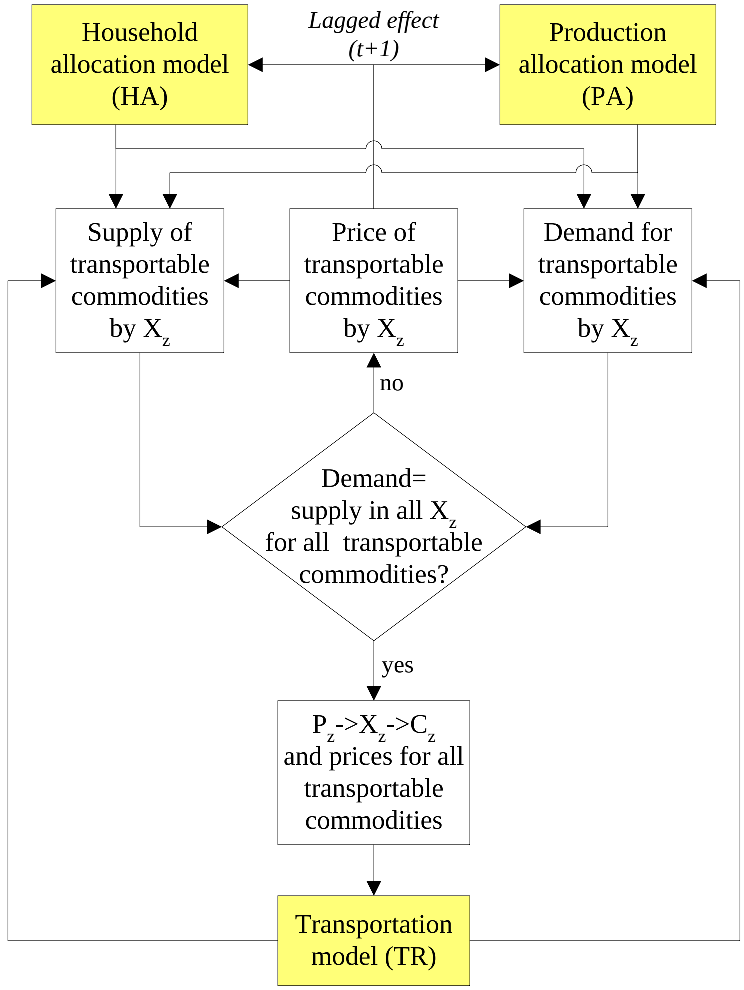

Figure 8: Activty interactions modeling framework

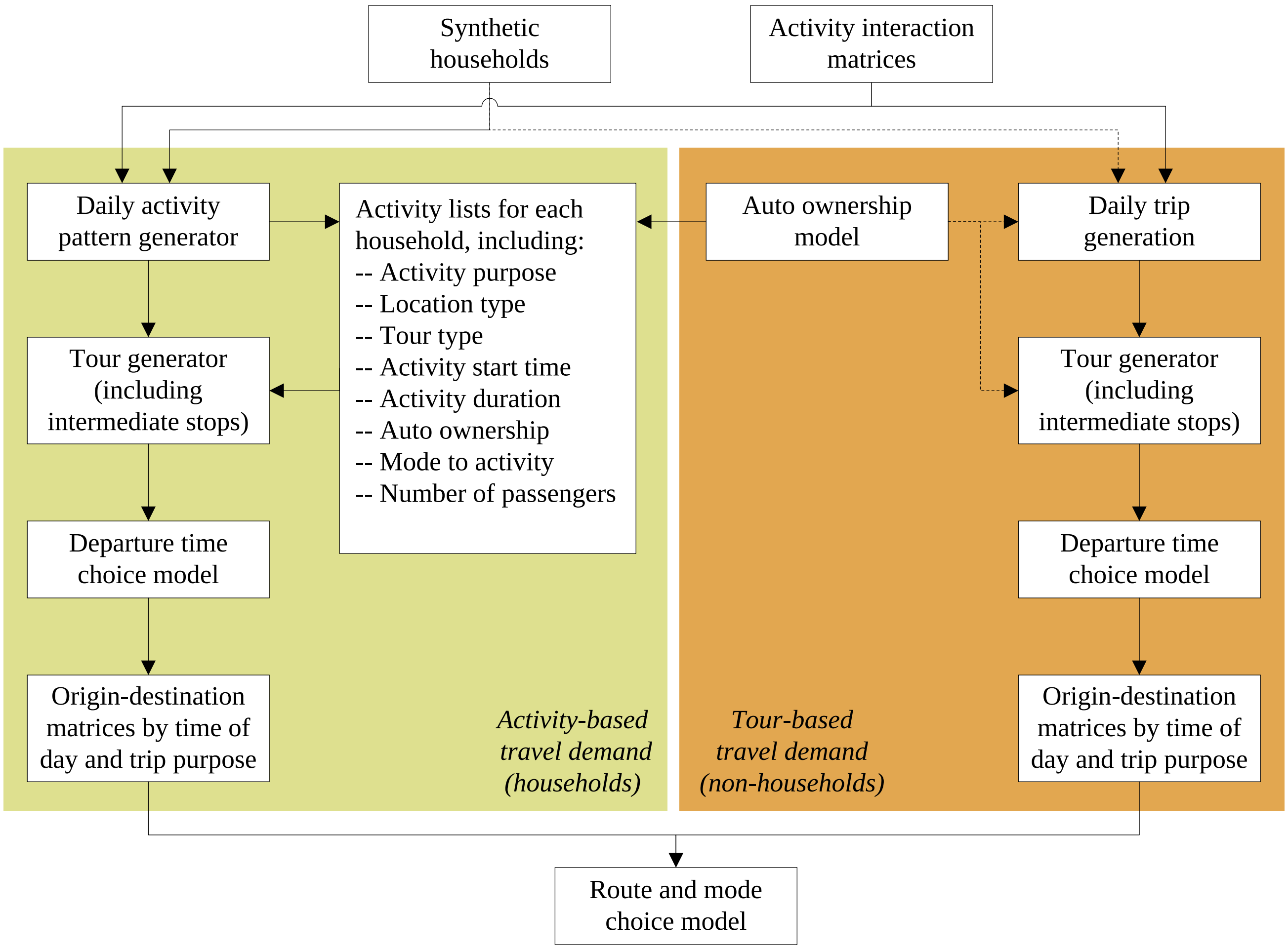

Figure 9: Transportation modeling component framework

Appendix: Reviewer Comments and Responses

The Peer Review Panel and the study team (broadly including both the ODOT staff and the consulting team) were asked to review the first draft of this paper. The written comments submitted to us are shown in italics below, and responded to as appropriate. Several sections of the manuscript have been revised to take into account the comments received.

Frank Koppelman, Northwestern University (Peer Reviewer)

The "fully dynamic" goal (p. 1) should be clarified. Dynamic could imply daily, weekly, etc. adjustments which is not what you mean or what you should be doing.

The proposed model is dynamic in the sense that once the original population of households and businesses are created, only changes to the original population are modeled in subsequent time periods. Thus, a household will change over time (and possibly die) over the course of the simulation. This is in contrast to a quasi-dynamic system, in which the entire population is regenerated in each time interval. In both cases the marginal totals are respected; the difference lies in the resolution and fidelity of the update process.

It would be useful to explain the reasons why the "two modeling streams... are... moving... apart" (p. 1).

We were unsure which way to proceed in building the first generation models, and thus we leveraged the risk by using an established package (TRANUS) for the statewide work and developing the metropolitan model (UrbanSim) as an entirely new approach. TRANUS is a fully-integrated, quasi-dynamic equilibrium model operating at an aggregate spatial scale. UrbanSim, by contrast, is a dynamic disequilibrium model operating at a fairly disaggregate scale. It incorporates business and household location and land development components, but does not include an integrated transport or macroeconomic model. UrbanSim has continued to evolve in ways that unquestionably improve its ability to model urban systems, but do little to bring the two approaches closer together. We anticipate that UrbanSim will continue to evolve separately, but its utility for statewide modeling is questionable.

I am always concerned about the use of trends (p. 2) especially if they are not adjusted to take account of differences in regional structure, etc. which are important elements of this model system.

We are assuming that the multi-regional structure of the "upper level" of the model will capture the important effects of the local economy and population. By isolating Oregon as one or more regions in that structure, we will avoid using national trends to describe Oregon activities.

I agree that a "multi-year lag" will be needed (p. 3).

I don't like the approach described in the 5th paragraph on page 6. Instead of selecting the households with greatest probability (utility) of moving out of the region, it would be better to calculate the probabilities and use Monte Carlo simulation to assign households to the stay or leave category. This approach may require some iteration to get the number to match the control totals.

I think that the outcomes will be very similar using either technique, but we will investigate your proposal more closely during model specification and development.

Also, on page 6, it might make sense to use a simple "number of cars" choice model for in-migrants, probablistically assign them to "number of cars" category and locate them based upon accessibility.

A good idea that we'll incorporate into the detailed model specification.

I am a little confused by the discussion at the top of page 7. Is this to obtain a starting point only (which makes sense to me) or repeated periodically (which is inconsistent with your dynamic structure).

The synthetically generated population is the starting point for the simulation, and as such is carried out only once.

Where will you get the marginal distributions for in-migrants referred to in paragraph 4 on page 8?

The marginal distributions will come from the demographic model.

Yes, different pattern distributions should be used for different portions of the region (paragraph 3, p. 10).

David Simmonds, David Simmonds Consultancy (Peer Reviewer)

Rather than trying to review the Revised Design from scratch, I begin my comments by considering how the revised design responds to the points I made in my note of 14 July. Those notes summarized what I said about the earlier Preliminary Draft (that of 2 July 1999) during the closing session of the last Peer Review Meeting. I then go on to a number of other points that have occurred to me in reading the Revised Design.

As before, I have concentrated on issues relating to "land-use" and land-use/transport interaction.

<u>Sequence of activities</u>

I accept the points the Rick made in reply to my note, primarily that
- *I missed earlier discussion of what was emerging from work on the Gen1 models.*
- *Further testing of the Gen1 models may be worthwhile, but has to be justified in terms of their relevance to Gen2.*

In addition there is some discussion of the results from UrbanSim, of the kind that I was recommending, in the revised (27 August) report on the Longitudinal Calibration.

<u>A single, scalable model</u>

If I follow the revised design correctly, there are now two different levels of operation in the proposed model:

- *The detailed level of some 3000 zones, covering the whole of Oregon and the surrounding "collar" but most detailed within the Willamette Valley (as defined on page 2). All of the zones in Oregon and "collar" will be "internal"; there will be "external" zones beyond these.*
- *A regional level, with one zone for the entire model area (i.e. Oregon and the collar), and other zones for the corridors north, east and south and for the rest of the US.*

A linked economic-demographic modelling system will operate at the regional level, and will predict:
- *Total employment for the model area by sector*
- *Total trade (by value?) between the regions (critically imports to and exports from Oregon+collar)*
- *Natural demographic change (is this actually to be done at the detailed zonal level? the description on p6 says "aged in place...")*
- *Gross in- and out-migration.*

The processes working at the detailed zonal level are
- *Household allocation (to vacant residential space)*
- *Production allocation (to vacant office, commercial or industrial space)*

both constrained to model area totals output from the regional level.

This is clearly a two-level model, with different processes operating at the different levels, and as such represents a large step in the direction I was suggesting. However, it would still appear (p8) that all households moving within the model area are added to and allocated from a single mover pool. This would appear to mean
- *That there is no difference in process between (for example) moves within Eugene/Springfield and moves between Eugene/Springfield and Portland*
- *That there is no influence of distance on households' movements within the model area.*

Unless there is evidence to suggest that local and inter-urban moves should all be treated as one choice set, I think there is still a need either to separate these two processes or to use distance-related variables as an influence on how far households move.

The space pool will be implemented at the zonal level, not as a scalar value for the entire system. Thus I believe that your concerns will be adequately addressed.

<u>Modelling processes of change</u>

It seems to me that the present design is much more satisfactory - it is defined in terms of a number of processes which at least broadly represent different kinds of decisions made by different types of actors. The emphasis is now on these processes interacting with each with time lags (one or more years), whereas the previous draft emphasized reaching a short-run equilibrium (with potentially problematical inertia terms) in each year. It may that some of this is a presentational change (page 2 says that "The proposed model retains most of the same components originally suggested for the second generation model.") but even so I think the change of emphasis is valuable in working towards a model which is defensible both to specialists in particular disciplines and to a wider audience.

I think it is advantageous that the design has reverted to representing households as such, rather than trying to imbue individuals with characteristics describing household behavior and choices. I say this partly because in another context I have (for reasons of data availability) been considering the possibility of modelling individuals whilst retaining a reasonable representation of household behavior, and I have come to the conclusion that it is more likely to create problems than to remove them.

Elaboration of the model

The new design comes across as a new approach which looks for existing models to help to fill in particular components (such as the regional economic model and the household allocation model). It is an entirely subjective judgement, but I feel much happier about this than about the previous draft, which seemed to be trying to add additional features and characteristics to an existing spatial-economic model structure.

Conclusion in relation to previous comments

With one exception - the lack of a spatial level for migration between the different urban areas within Oregon - I feel that the revised design is much more satisfactory, and that the criticisms I made before are no longer needed.

It should be acknowledged that the new design is distinctly more similar to the models I have developing (as the DELTA package) over the last five years[4]. There are two interconnected reasons for this:

- *The arguments I have been putting forward in TLUMIP are the same as those that I have been using in my own work (give or take a little more experience)*
- *Both TLUMIP and DELTA are influenced in different ways by Michael Wegener's views about urban and regional modelling, and his application of those ideas in the IRPUD model.*

The remainder of the comments relate to the present draft of the paper.

Comparison with other models

Page 3 seems to suggest (in the paragraph about the one year time step) that in "classical depictions of integrated land-use/transportation models" the output of the transportation model is the only time-lagged input into the next period's activity allocation. This seems an exaggeration. It is true that the most common form of integrated land-use/transportation model over the last 20 years has been that in which interactions are used to predict locations, and that in these lagged transportation outputs are an extremely important input to activity allocation. Nevertheless, most if not all of these models incorporate some other lagged terms or relationships, notably in relation to the development process.

Relationship of economic model and production allocation

4. *For information: a paper describing the approach at the urban level has been published since the last Peer Review Group meeting as Simmonds, D C (1999): The design of the DELTA land-use modelling package. Environment & Planning B, vol 26 pp 665-684. 2 Since writing this I have found Figure 8, but I am not sure I am much wiser.*

Some further thought is needed about the relationship between the economic model and the production allocation model. The plan outlined on page 3 is that

- *Decreases in model area total production by sector, output from the economic model, will be added to a "negative pool" and "delocated" in the allocation model so as to free up space, whilst*
- *Increases will be added to the "pool" to be located into available space.*

In addition, there will be a proportion of existing-and-continuing production will be allocated to the movers' "pool" and relocated. The problem here is that the economic model will only produce the net changes over time in each sector, i.e., each sector in each year will show either an increase or a decrease. As currently expressed, the model will not consider the turnover of activity within each sector, with jobs being created in some firms or establishments and lost in others - the net change reflecting the difference between these. This can perhaps be dealt with (in the necessarily simple approach which is rightly being taken in the production allocation component) simply by defining "movers" more broadly - not just the number of jobs which may relocate during the year (which I understand it to be at present) but also including the gross number of jobs created and lost each year even within a static total.

A good idea that we will consider during the detailed model specification (Task 3A).

<u>Household allocation process</u>

The description of the household allocation process on page 8 proposes

- *A sequential allocation process, with households of the highest income group having first choice of location, then the second highest, etc.*
- *Prices of space being updated within that process, but with the revised prices applying only to the households under consideration when space is exhausted and to those in lower income groups.*

This seems to be a compromise between two common approaches:

- *Those where prices are held fixed within each period (my understanding of IRPUD and UrbanSim) and some kind of rule (possibly the same sequential one) is used to control allocation*
- *Those where price are variable within each period and the allocation process of all household categories is reiterated each time prices are adjusted (as in MEPLAN, TRANUS and DELTA).*

The effect of the proposed approach would seem to be that demand from higher income households can affect the prices paid by lower income households, but that the cumulative demand from lower income households has no influence at all on the prices paid by higher income households. This doesn't seem realistic (certainly not from European experience). It would also leave different prices being paid by different household groups for the same kind of space in the same place, and hence an apparent ambiguity in defining which price should be adjusted to get the starting price for the following year.

It seems consistent with reality that upper income households do have a disproportionally large effect on housing prices paid by all income groups. The upper incomes households can more easily adjust to rapidly escalating land prices than lower income households, and can in and of themselves bid up the price of tight supplies of land so as to drive out or exclude lower bidders. The ability of the lower income households to moderate or dampen this influence seems almost non-existent. So within a given time interval where the demand exceeds the supply of housing, the prices will adjust upwards until this fundamental imbalance is resolved. It will mean that some households will be forced to relocate to lower cost land in the process (which happens in real life). But the price will be unambiguous at this point; it will simply be the price at which demand and supply come back in line. This price will then find its way into the next time interval, where it will be adjusted upwards or downwards according to simulated market conditions. Notwithstanding, we will consider some alternative allocation processes as model development proceeds.

Interaction models

I am unclear how the interaction models produce the money flows once the household and production allocations have been determined. The input of allocation results to interaction calculations is mentioned at the foot of page 8, and the use of the interactions to constrain the activity-based travel modelling is mentioned near the foot of page 10, but I can't find the step in between. Does the comment at the top of page 11, about needing to develop a way of conditioning the interaction matrices to be consistent with TBI data as well as with economic and demographic data, mean that this is as yet an unfilled black box? In passing I wonder whether if

- *The production allocation works entirely in terms of the value of production, but*
- *The household allocation model works entirely in terms of numbers of households (and hence of potential workers),*

such that it might be more appropriate to build the interaction matrices representing labour in terms of workers, so as not to have to convert from wages to workers in constraining the travel patterns?

A section describing the treatment of interactions has been added to the design paper. It indicates that the use of a conceptual framework where commodities move from production location to exchange location to consumption location has been retained.

Accessibility and locational utility

The paper doesn't seem to specify what variables contribute to the locational utilities for the household and production allocation models. I assume that these will be similar to those in UrbanSim (with values per unit of production replacing the UrbanSim values per establishment). The section on transportation modelling should include a definition of how accessibilities, or generalized costs from which to calculate disutilities, are obtained for use in the allocation models. This will presumably be done at the point where the nested logit destination choice model is run (see foot of page 9).

I think that we'll be able to specify candidate variables for the location models in the detailed model specification, but we will obviously have to wait until model estimation is undertaken to evaluate their explanatory power. Most certainly the work that Paul Waddell has done with UrbanSim will provide a starting point for model specification. We'll likewise look closely at Paul's definition of accessibility, but we will not commit to a specific approach until we complete the detailed model specification.

It is important to note that a nested logit destination choice model is only part of the present Portland implementation of their activity-based models. There is no analogue to it in the proposed statewide model. As noted at the top of page 10, destination choice — expressed in terms of the activity interaction matrices — will be a constraint imposed upon the transportation model rather than something modeled within it.

Conclusion

There are clearly some important issues still to resolve, or at least to clarify, within what seems a very promising approach.

Paul Waddell, University of Washington and Urban Analytics

The one-level model statewide is an appealing simplification in one way (it's one model!), but it does force a compromise on geography. The level you've suggested seems to be slightly aggregated from the metropolitan model scale, but not much. Running the entire state at this level, with households and production choosing between zones at far corners of the state, makes me a bit nervous. But if we get the underlying locational behavior specified well at a disaggregate level, and can add effects that explain the anomalies - or at least K-factors, then it should work.

We expect that the correct specification of transport disutilities would prevent households and businesses at opposite ends of the state from interacting in an inappropriate or unrealistic manner.

For geography, it appears you've dropped the grids. I've had good luck using grids in the Honolulu and Salt Lake City model applications, and it seems to really help clean up data noise (I've not done any parcel cleanup for those applications). Also pretty easy to use data synthesis to fill gaps in the data. But the main advantage is the ability to generate efficient spatial queries. I've been using a 150 meter grid (too detailed for a statewide model, but easy to change) to measure neighborhood effects like density, land use mix, proximity to highways and arterials, and employment clustering - and these seem to be very strong variables. Anyway, it was disappointing to see this aspect dropped, especially now that we have the software architecture completed to support it.

The grid system admittedly has many attractive features, and a lot of thought went into our proposal to not implement one. We were particularly impressed with the range of environmental analyses that can be carried out in SPARTACUS using a grid system. The downside of grid systems is that it appears not all that efficient for a statewide modeling application. If we use a uniform grid system, then cell sizes appropriate for the Willamette Valley seems way too fine for the remainder of the state (especially Southeastern Oregon). The way around that appears to be variable size, irregular, or nested grids, which appear to be many orders of magnitude more complex and difficult to implement than uniform grid systems. Having said all that, however, we will reconsider the idea of using a grid system during detailed model specification.

The macro economic and demographic process recommendations seem sensible, but still quite challenging. There's more microsimulation and more new development.

The additional work required to implement these components is acknowledged.

I can’t quite tell what the Interactions Model is, and don’t see much discussion of it. What does it look like, and does it attempt to resolve simultaneously the location of production and households through iteration? Or is it a price adjustment that influences choices in the next time period? Can’t make much comment without more information, and this seems to be a core component of the proposal.

The document has been revised to include a description of the Interactions Model.

Household location seems to be a bit unnecessarily simplified and probably biased, in the sense of sequentially allocating high to low income. Why not use a logit model of location choice in which price enters the specification, and assign households in random order? Higher prices should lower the utility for lower income households. Also, I don’t recall seeing anything before on the 2nd home issue. How critical is this? It seems to add unnecessary complexity and may be quite difficult to estimate.

Household location choice will be done using a logit model of location choice with price entering the logit utility functions as an attribute. The expectation is that a variety of ordering strategies will be considered as development proceeds. Note that proximity to previous location may also be an attribute in the utility functions, which would allow this issue to be addressed without necessarily having to split the movers pool.

Second homes have an important influence on land markets in certain recreational areas and assist in the explanation of the congested conditions that arise with recreational travel to and from the coast in particular. We have gone back and forth on this issue for several years with the statewide model. The ODOT staff believe that there are a significant number of them in Northwest Oregon, and that their characteristics are markedly different from first homes. If we fail to capture these dynamics in the model then it seems likely that the model will not be able to handle them in a competent manner.

If this specification has moved away from a Spatial I/O framework (as it appears to have done), then why not use employment as the unit of analysis for location rather than dollars? It is clearer, there is better data, it is easier to verify, etc. You can convert from employment to tonnage of goods, or trip production, etc. by industry sector as easily as these conversions can be made from dollars.

The specification retains elements of spatial I/O in the determination of interactions. As such, it is more consistent to retain to representation of production activity in units of value (dollars). Employment (labor) is one of the elements that is consumed in production, so the model will also determine employment (as labor consumed) as part of its treatment of consumption arising with production, and therefore provide employment numbers that can be used in verification as desired. Variations in labour productivity in both time and space can also be represented explicitly.

Implementation framework: I can’t tell from the discussion whether you intend to build on the modeling and visualization architecture we’ve got built, or to start over. Seems a waste to do the latter, and I’d like a chance to discuss the flexibility of the architecture to implement the model specifications you’ve put out. I know last year my suggestions along these lines didn’t go over well, but I hope we’re not at the same place now.

This issue will be addressed when we carry out the detailed Platform Specification (Task 3C). We are most certainly amenable to reusing components rather than recreating them.

Michael Wegener, University of Dortmund (Peer Reviewer)

The revised design for a second generation land-use transport model for Oregon seems to be a major step towards an integrated, policy-relevant and feasible model design.

The nice things first:

1. *The design goals set up for the second generation seem to be reasonable and realistic: The decision to concentrate on a single geographic scale for the model geared to corridor studies covering several municipalities is good. The principle of zones increasing in size from the areas inside municipalities to the areas outside of them is reasonable. The remaining design principles (full integration of land use and transport, full dynamics, blend of equilibrium and disequilibrium, activity-based travel model) are excellent.*

2. *It is a great advantage that the new model design takes advantage of the rich modelling experience available at Oregon DOT and Portland Metro, in particular the experience with both TRANUS and UrbanSim and with the Portland TRANSIMS application.*

3. *The overall model structure as presented in Figure 1 and the temporal model dynamics as presented in Figure 2 are convincing. It is particularly interesting how the economic-demographic model (ED) interacts with the household and production location models (HA and PA) and the activity interaction and transport models (AI and TR). This preserves the best of TRANUS and UrbanSim yet gives the freedom of introducing some of the exciting new elements of TRANSIMS.*

4. *The plan to model household formation and location as a microsimulation is an innovation going beyond TRANSIMS. This submodel would place the second generation model among the very few efforts world-wide to apply microsimulation to other urban processes than travel. The decision to keep the production allocation model aggregate for the time being seems to be reasonable and leaves future options open.*

5. *The land development submodel as presented in Figure 7 (not explained in my copy) seems to pick up some of the best features of UrbanSim.*

The description of the land development and activity interaction components (Figures 7 and 8 in the paper) were inadvertently left out of the copy that was sent to you. This has been corrected in the current version of the document.

6. *The activity interactions submodel as presented in Figure 8 (not explained in my copy) seems to take advantage of the equivalent feature in TRANUS.*

7. *The travel model promises to be an interesting blend of activity-based travel demand modelling and aggregate tour/trip assignment, which solves the problem of combining individual person trips with aggregate freight flows and avoids the excessive detail of the routing and vehicular simulation in TRANSIMS.*

8. *The ideas about programming principles, GIS integration, visualization, animation and open-software policy sound exciting.*

Some problems:

9. *The volume of effort likely to be spent on the economic submodel (ED) seems very large. To develop a full-scale 'integrated econometric-multiregional input-output model' as depicted in Figure 3 is a very large effort by itself and may distract too much work capacity from the land-use transport model. After all, the economic submodel is only needed to provide the regional control totals in terms of production and employment for the land-use transport model - could this not be taken from other, e.g. national forecasts or simply be specified in the form of scenarios?*

Your point is well taken, particularly with respect to the resources the work might distract from the land use and transport components. We will monitor that and provide periodic reports at the peer review panel meetings. Some of the specified functionality could indeed be provided with a flexible scenario manager. We wrestled with this topic, but in the end concluded that a robust endogenous economic model was required. Part of the impetus for such was the requirement to test several economic and demographic development scenarios, which the proposed model would be ideally suited to. While some of the information could be specified exogenously, the feedback interactions (which you allude to in your comment 3, above) would not be possible. Much of the pioneering research work on combined econometric-multiregional input-output models has already been done; we plan to implement the scheme suggested by Conway (1990) rather than pave new ground here. Finally, we feel that intercity freight modeling is more heavily dependent upon a good economic model at its base. The proposed framework should provide much better data for freight modeling than the current TRANUS and UrbanSim frameworks.

10. *The separation of (activity-based) travel demand modelling and (aggregate) assignment poses the problem of feedback between the two leading to adjustments of destination, route and departure time in the tour-generation step. This issue, which goes beyond meeting the marginal constraints of origins and destinations, is not addressed in the paper. The solution implemented in TRANSIMS, several iterations of both tour generation and assignment, is largely responsible for the excessive computing time requirements of TRANSIMS. The obvious solution would be to integrate both steps without iteration (as in reality). I discussed this recently with Kai Nagel, who is now in Zurich. Which solution is envisaged here?*

Integrating the demand and supply sides of the transport model would obviously be an elegant solution to the problem you've described. The proposed model moves in the direction suggested, in that it uses a combined model of route and mode choice. Our motivation for leaving the remainder of the steps separated stems from our desire to re-use as much of the existing framework developed by Portland Metro. The development of a true simultaneous model of activity-based transport demand and supply in this phase of the TLUMIP work is beyond our means. We are hoping that using the network state from period t as the starting point for the transport model in t+1 years will provide enough feedback such that multiple passes through the transport supply model will not be required within a given year. This is an area of on-going research that will be reported upon at the next peer review panel meeting.

11. *Despite the shortcuts of separating tour generation and assignment and of using aggregate assignment, the model is likely to run into severe computing time problems, in particular if the design goal of full integration of land-use and transport is to be met, i.e. if feedback between land use and transport is to be implemented in each simulation period. Long computing times, however, will collide with the planned interactive visualization and animation features. Has this aspect been considered? Are there realistic forecasts of the likely computing times to be expected?*

At the present time we plan to separate the process visualizer from the simulation engine. The former will run on the user's workstation, allowing them to configure scenarios, run the model, interactively review the progress of the model, and to retrieve data that will permit the visualization of results. The simulation engine will likely be implemented on a quad-processor server running the Linux operating system. We plan to rely heavily on multi-threaded code to take full advantage of multiple processors, especially for those components involving microsimulation and network modeling. We feel that this will provide an acceptable balance between cost and performance, although we obviously will need to continually assess this throughout the project. Our best estimate at the present time, based upon experience with other models and limited testing of certain ideas to date, suggests that runtimes on the order of about five to six hours should be possible.

<u>Two small suggestions:</u>

12. *The demographic submodel is said to be at the region-wide level, yet the existing population "will be aged in place, i.e. "their location in a zone will remain intact". That is a contradiction. It would probably be better to clearly distinguish between region-wide and zonal calculations, i.e. to move ageing of the zonal population to the household allocation submodel.*

The regional model will specify the total level of change (i.e., total in-migration, out-migration, etc.) that the household allocation model will add or subtract from the household database.

13. *Figures 1 and 2 give the impression that household allocation and production allocation are only driven by the economic-demographic submodel, i.e. that in each period the whole population is allocated to zones. This is in contradiction with what really happens: in both models the 'stayers' are aged and retained from the previous period, and only the 'movers' are newly allocated. This should be expressed in Figure 2 by additional arrows leading from HA in period t to HA in period t+1, and from PA in period t to PA in period t+1.*

The Figures have been changed as suggested. A single red arrow shows the dynamic linkage between the PA and HA in one time period and the next. They are combined into a single arrow from one period to the next to make the Figure more readable, but in reality will operate in the manner that you describe.