California Energy Commission

CONSULTANT REPORT

Energy Commission Models for Analyzing and Projecting Household Transportation Energy Demand

Evaluation, Model Improvement Options, and Recommendations

Prepared for: **California Energy Commission**

Prepared by: **David S. Bunch, Ph.D., University of California, Davis**
David Brownstone, Ph.D., University of California, Irvine

California Energy Commission

Edmund G. Brown Jr., Governor

June 2016 | CEC-200-2016-009

California Energy Commission

Primary Authors:

David S. Bunch, Ph.D., University of California, Davis
David Brownstone, Ph.D., University of California, Irvine

University of California, Davis
One Shields Avenue
Davis, CA 95616

Contract Number: 600-12-012

Prepared for:

California Energy Commission

Aniss Bahreinian
Agreement Manager

Elise Brown
Office Manager
DEMAND ANALYSIS OFFICE

Sylvia Bender
Deputy Director
ENERGY ASSESSMENTS DIVISION

Robert P. Oglesby
Executive Director

ACKNOWLEDGEMENTS

The authors acknowledge helpful conversations with Caroline Rodier, Ph.D., University of California, Davis, and helpful comments from Aniss Bahreinian, California Energy Commission's Transportation Energy Forecasting Unit, on earlier drafts of the final consultant report.

ABSTRACT

The California Energy Commission is required to regularly forecast statewide transportation fuel demand under alternative policy scenarios and, in particular, assess alternative fuel scenarios in collaboration with the California Air Resources Board. To do this, the Energy Commission relies on a modeling system called DynaSim. The recently collected *California Household Travel Survey* data offer an opportunity for updating models. In addition, other agencies have models with similar capabilities (for example, the California Statewide Travel Demand Model).

This report is the outcome of a project to review and evaluate current models and the new data sets, develop alternative options for improved models, and make recommendations. The scope is limited to fuel usage from personal vehicles in California. The review concludes that there are major flaws with the urban and intercity travel models of DynaSim, and that the Energy Commission should develop alternative modeling procedures that incorporate in some fashion the availability of the California Statewide Travel Demand Model. Two primary issues (and how they interrelate) are the quality of total vehicle miles traveled modeling and forecasting methods, and the proper allocation of vehicle miles traveled to vehicle types in the personal vehicle fleet. Both are critically important for producing fuel demand forecasts, particularly in cases involving future alternative fuels and new vehicle technologies. The report provides guidance on a future pathway of new model development with sequential testing and evaluation.

Keywords: Transportation, fuel demand, alternative fuels, vehicle miles traveled, vehicle type choice, travel demand, forecasting, DynaSim, California Statewide Demand Model, CARBITS, EMFAC, models

Please use the following citation for this report:

Bunch, David S., David Brownstone. (University of California, Davis and University of California, Irvine), 2016. ***Energy Commission Models for Analyzing and Projecting Household Transportation Energy Demand: Evaluation, Model Improvement Options, and Recommendations.*** California Energy Commission. Publication Number: CEC-200-2016-009.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

California has historically recognized the potential negative impacts associated with energy consumption and has been a leader in developing policies to address issues such as pollution, greenhouse gas emissions, and climate change. The state's transportation sector plays an important role in the state's economy and is responsible for a large share of the greenhouse gas emissions.

Senate Bill 1389 (Bowen, Chapter 568, Statutes of 2002) requires the California Energy Commission to conduct "assessments and forecasts of all aspects of energy industry supply, production, transportation, delivery and distribution, demand, and prices" to develop policies for its *Integrated Energy Policy Report*. The Energy Commission develops long-term projections of California's transportation energy demand that supports its analysis of petroleum reduction measures, introduction and commercialization of alternative fuels and technology, transportation fuel infrastructure requirements, and energy diversity and security.

In 2010, the California Department of Transportation (Caltrans) began collecting data for the *2010-2012 California Household Travel Survey*, and the Energy Commission began work on the 2013 California Vehicle Survey. The Energy Commission and Caltrans worked together to link these surveys, resulting in a very rich data set for households in California. The availability of these two new data sets motivated this project.

This report evaluates forecasting models related to household vehicles and related fuel usage in California. Other state agencies produce travel-related forecasting models that overlap functions in a variety of ways. The Energy Commission's personal vehicle choice model, which incorporates the effect of household preferences for attributes of competing vehicle technologies, is the only model of its kind. Conversely, the Energy Commission's urban and intercity travel models, which forecast vehicle miles travelled in short-distance and long-distance modes, overlap with the California Statewide Travel Demand Model. Other agency models that specifically address the California vehicle fleet and related fuel usage are the California Air Resources Board's CARBITS and the Emissions Factor database.

The proliferation of models across state agencies that use similar data as inputs, but have overlapping functionality and produce similar forecasts or projections, which leads to inefficiency, and potentially, confusion or tension, should agencies produce inconsistent forecasts or analyses for what appear to be similar policy questions. These agency models are used in conjunction, making these relationships important to consider. Duplicated effort and resources directed to similar modeling activities might be eliminated through greater coordination and perhaps even shared effort across agencies.

This review found that the California Statewide Travel Demand Model should play a role in any future model improvement pathway followed by the Energy Commission. The authors arrived at this conclusion by finding that although the California Statewide Travel Demand Model's substantial level of detail and run times restrict the volume of results, it is a theoretically superior state-of-the-art bottom-up system that uses detailed inputs to produce high-quality results. This report represents the starting point for future model improvement pathways that could be developed through a logical sequencing of activities.

CHAPTER 1: Introduction

Background

The State of California has one of the world's largest economies and is one of the largest consumers of energy. At the same time, California has historically recognized the potential negative impacts associated with energy consumption and has been a leader in developing policies to address issues such as pollution, greenhouse gas (GHG) emissions, and climate change. This viewpoint is highlighted in the California Public Resources Code, Section 25300, which discusses the importance of clean and reliable energy and establishes the essential role that state government has in addressing these issues. Part (d) specifically states that data collection and analysis activities are "…essential to serve the information and policy development needs of the Governor, the Legislature, public agencies, market participants, and the public."

To this end, the California Energy Commission is charged with meeting many of these needs, in part by producing a biennial *Integrated Energy Policy Report* (*IEPR*).[1] To provide background and context, the authors review some of the requirements of the *2013 IEPR*. The scoping order specifically states:

> "The report *makes energy policy recommendations* based on the Energy Commission's *energy assessments and forecasts* with the intent of conserving resources, protecting the environment, providing reliable energy, enhancing the state's economy, and protecting public health and safety." (Emphasis added.)

The *2013 IEPR* scoping order identified specific topics to be addressed: energy efficiency, demand response, electricity, nuclear power plants, natural gas, transportation, and climate change.

Although there are clear interactions among these topics, the primary focus of this project is the *transportation sector*, which is concerned with the movement of people and goods by vehicle, rail, airplane, and other transportation modes. Transportation is essential to California's economy and accounts for about 40 percent of energy consumption (and roughly 38 percent of the state's GHG emissions).[2] The Energy Commission's transportation-related responsibilities in the scoping order are as follows:

1 See McAllister, Andrew (2013a).

2 See McAllister , Andrew (2013b).

- Forecasts of statewide transportation fuel demand, fuel prices, and factors influencing changes in demand (such as updated forecasts of economic conditions); assessments of alternative fuel scenarios conducted in conjunction with the California Air Resources Board; and discussion of barriers to and progress toward meeting California's transportation energy goals.
- Evaluation of research, development, demonstration, and deployment activities funded under the Alternative and Renewable Fuel and Vehicle Technology Program (as required by Assembly Bill 109, Núñez, Chapter 313, Statues of 2008).

Forecasting fuel demand is a complex task that requires collection of data, development of credible forecasting models, and the judicious adoption of key input assumptions to define a baseline/reference scenario that include such factors as:

- Future population and economic growth.
- Projected advances in vehicle technology.
- Future fuel prices.
- The impact of established policies.

A thorough characterization would also address the implications of uncertainty in these assumptions. The potential impact of alternative policies would be analyzed by producing forecasts for appropriately modified scenarios and comparing the outcomes to those from the baseline/reference scenario. In the quickly changing policy landscape, it is increasingly difficult to distinguish between baseline and alternative scenarios, which include considerations related to the zero-emission-vehicle (ZEV) mandate, low-carbon fuel standard (LCFS), land-use policies in response to Senate Bill 375,[3] alternative fuel vehicle refueling/recharging infrastructure, high-speed rail, and so forth.

The Energy Commission employs a variety of models and data sources to perform these assessments, depicted at a high level in **Figure 1** (Aniss Bahreinian 2013a) and implemented using a custom-built modeling platform called DynaSim.[4]

Sources of transportation energy demand can generally be divided into two categories: "personal" and "commercial." This project specifically addresses data and models for personal travel behavior.[5] **Table 1** (see Aniss Bahreinian 2013a) summarizes the fuel demand estimates produced by the Energy Commission, where items derived from personal travel have been highlighted in bold. Personal travel behavior can be

---

3 *Sustainable Communities and Climate Protection Act of 2008*, State of California, Senate Bill No 375. 2008.

4 For identification, the term "DynaSim" will be used to refer to the Energy Commission's collection of forecasting models as implemented and operated by users.

5 Personal travel behavior is a subject with a substantial academic research literature.

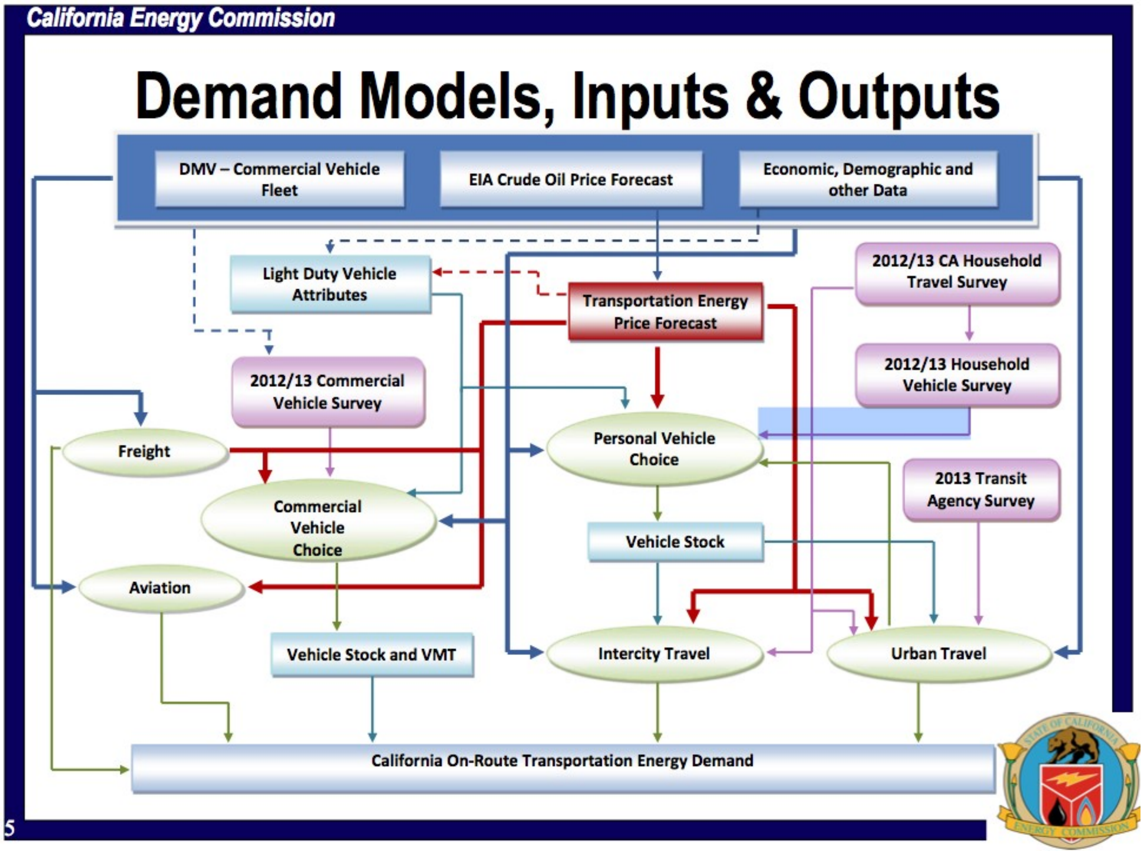
subdivided into two categories with fundamentally different behavioral motivations and implications for energy usage: *short-distance travel* and *long-distance travel*. In terms of behavioral choice, attention may be placed on decisions about individual trips and related lengths, and the mode of travel.

Short-distance trips are generally associated with routine household activities such as commuting to and from work and/or school, shopping, and short-term leisure activities. Available travel modes would typically include driving alone in a household vehicle, shared use of vehicles both within and across households, and, depending on local transportation infrastructure and land-use patterns, other local transit options (for example, bus, light rail, ferry, as well as biking and walking. Much of the state's energy usage is associated with miles traveled in personal vehicles for short-distance trips of this type. The Energy Commission's Urban Travel Model is used for these types of transportation choices, where the "urban" terminology emphasizes the richer set of local mode choices that may be available to households in urban or suburban settings.

Long-distance trips would generally be much less frequent, occurring for work-related reasons or for longer-term leisure activities (such as vacations, family-related obligations, and so forth), and be characterized by a different set of mode options. As for short-term trips, an important travel option would be using a household vehicle; however, other competing modes would include air travel and possibly train, or bus. In considering longer-term energy usage in the state, high-speed rail must be addressed in policy analysis. The Energy Commission's Intercity Travel Model is used for this purpose.

To summarize, a substantial portion of energy usage for personal travel in the state is conditional on two types of interrelated household choices: the vehicle fleet, and the way it is used. The centrality of these interrelated choices is depicted in **Figure 1**, which specifically identifies the Personal Vehicle Choice (PVC) model and illustrates the interaction of this model with the urban and intercity travel models by specifically identifying vehicle stock as an intermediate construct.

Figure 1: Current High-Level Structure of DynaSim Models, Inputs, and Outputs



Source: Aniss Bahreinian, Transportation Energy Demand Model Overview, IEPR workshop presentation, June 26, 2013

Table 1: Fuel Demand sectors in DynSim Software

Fuel Demand Sector	Vehicle Size Class and Transportation Sector (Model)
Personal	**Light-duty vehicle (LDV) fuel demand for personal travel (urban and intercity).**
	Heavy-duty vehicle (HDV) fuel demand for personal travel (urban and intercity) by transit mode: bus, rail, light rail, and others.
	Aviation fuel demand for intrastate, interstate, and international personal and business travel (**commercial passenger aviation**).
Commercial	LDV fuel demand for commercial travel.
	HDV fuel demand for movement of goods (freight) by commodity sector and by mode: rail and truck.
	HDV fuel demand for services (freight) by sector.
	Aviation fuel demand for goods movement by intrastate, interstate, and international destinations (freight aviation).

Moreover, the DynaSim software system specifically identifies LDV attributes as a required input to the PVC. The implications of this are non-trivial because they mean the system requires a forecast of future vehicle technology attributes over a relatively long planning horizon to make projections of LDV demand. Similarly, this conclusion implies that the PVC incorporates the effect of household preferences for attributes of competing vehicle technologies. This is a major distinguishing characteristic of the Energy Commission's modeling capabilities. Historically, the Energy Commission has devoted substantial attention, effort, and expertise to modeling household vehicle choice behavior. In particular, it has been unique among state agencies in pursuing data and models to evaluate longer-term scenarios involving potential new vehicle technologies that rely on alternative fuels (such as electricity, natural gas, and hydrogen), technologies, (for example, hybrid electric) and policy-related incentives (for example, purchase incentives, high-occupancy vehicle lane access, parking, and so forth).

Figure 1 also identifies other major factors for which the effects must be correctly captured to produce viable long-term forecasts of travel-related fuel consumption. Overall travel demand will change over time due to demographic changes (such as population growth) as well as general economic conditions. More specific economic factors are fuel prices, which affect (either directly or indirectly) both vehicle choice and use. Specifically, effects from these factors propagate through the model, affecting the choices of vehicles, the fuels they consume, and how much they are driven. These choices involve tradeoffs, including how many vehicles to own, which trips they are used for, and whether to shift travel to alternative modes, and others.

Another potentially important factor not specifically shown in **Figure 1** (but addressed in the modeling system) is that road congestion can affect total fuel use. Vehicles using current and past technologies continue to use fuel even when sitting in traffic. Vehicle fuel efficiencies can vary depending on the speed at which vehicles are driven.

This high-level description highlights the fit between the functions performed by DynaSim and the *2013 IEPR* fuel demand forecast for household travel. This type of modeling activity is very challenging, and it is important that models be continually updated and improved to provide the best possible policy analysis to support critical long-term decisions for California.

This project reviews personal travel demand and household vehicle choice models in DynaSim and similar models used by other state agencies, as well as recently collected household and vehicle survey data, and develops options and recommendations for possible modeling improvements at the Energy Commission (perhaps in collaboration with other state agencies). To provide greater context, the authors summarize additional background on these models in DynaSim and related models from other agencies.

History of DynaSim Personal Travel Models

This section draws from Aniss Bahreinian (2014), who reviews the historical evolution of models used at the Energy Commission for estimating fuel usage and future demand. Personal travel components of the current version of DynaSim evolved from:

- An Urban Transit Model, developed in 1982.
- An Intercity Transit Model, developed in 1982 by adapting a 1960 model.
- A Household Vehicle Demand and Utilization Model (VEDUM) developed in 1983.

The current urban and intercity travel models are modified versions of the original urban and intercity transit models, extended to include travel associated with personal vehicles. They use a top-down approach that relies on aggregated travel-related statistics related to fuel usage (for example, trip counts, passenger miles, vehicle miles, fuel usage, and so forth) for an identified set of competing modes (auto, bus, rail), as well as performance-related characteristics (such as average travel times and travel costs) that would affect relative choice among modes. Important considerations for data collection and analysis are the definition of geographic regions and what constitutes short- versus long-distance trips.

A high-level description of the approach (1) computes required statistics for a specific base year using observed data from a variety of sources, (2) develops baseline projections of travel-related outcome measures for future years, and (3) performs policy analysis by estimating how projections would change relative to the baseline under alternative scenario assumptions.

It is standard practice that baseline projections are driven primarily by estimates of future population growth and income, but establishing a baseline also requires reference assumptions for other inputs. Factors include alternative assumptions on income or population growth, future vehicle technologies and infrastructure, and fuel prices. Policies that affect relative performance characteristics for competing modes would lead to shifts in travel patterns. For example, stricter fuel efficiency requirements on future autos would reduce fuel consumption, but at the same time, some additional travel could be diverted from other modes.

In contrast, VEDUM is a bottom-up model of household vehicle holdings (number and type of vehicles in a household) and usage decisions, where a “vehicle” is defined based on class (for example, compact car, minivan, small sports utility vehicle, and large pickup truck) and vintage (year of manufacture). For a household with a given set of characteristics (such as income, size, number of adults, workers, and location), it computes choice probabilities for (1) how many vehicles the household owns and (2) which vehicle class/vintage. It also projects annual vehicle miles traveled (VMT) for each vehicle. These decisions are based on household preferences for vehicle attributes, including vehicle purchase prices, vehicle types and sizes, and fuel operating costs. Estimating vehicle demand for future years requires projections of vehicle technology characteristics (such as fuel economy), purchase price and cost of fuel, as well as weights for each household "type" so that behavior can be "simulated" for each type and then aggregated to produce total results.

Kenneth Train (1986) provides a comprehensive description of VEDUM, including a detailed treatment of behavioral theory, econometric estimation procedures, and policy analysis examples. In this approach, interrelated vehicle choice and usage decisions are addressed using an internally consistent, integrated theoretical framework. The original version of the model also estimates how total VMT is allocated across multiple trip types (intracity work, intracity nonwork, non-intracity work, and non-intracity nonwork). Researchers estimated total VMT using a national-level household transportation survey and based it on revealed preference data for existing vehicles in the marketplace. Nevertheless, researchers attempted to use this model to analyze policies involving alternative fuel vehicles by constructing new vehicle classes and attributes based on researcher judgment.

As noted, VEDUM was a holdings model, based on national revealed preference data. In the mid-1990s, the Energy Commission began an ongoing program of enhancing and extending VEDUM in a variety of ways. A multiyear panel study of California households collected data on household vehicle transactions behavior, and stated preference choice experiments on alternative fuel vehicles (battery electric, methanol/ethanol, and compressed natural gas). VEDUM was updated in 1996 to become CalCars, a model of household vehicle transaction choice that could more readily address choices among alternative fuels—see Chris Kavalec (1996). In addition to a re-estimated VMT submodel, CalCars also included a "fuel choice" submodel for the case of dual-fuel vehicles. In subsequent years (2003, 2007, and 2009), the Energy Commission collected additional survey data to update CalCars. Subsequently, the VMT functionality of CalCars was dropped, and the vehicle choice model evolved into the current PVC model.

Overlap of DynaSim With Models From Other State Agencies

Other state agencies produce travel-related forecasting models that overlap DynaSim functions in a variety of ways. Two models that fall into this category are the California Statewide Travel Demand Model (CSTDM), administered by the California Department of Transportation, and CARBITS, administered by the California Air Resources Board (ARB). CSTDM functionality overlaps much of DynaSim with the exception of Personal Vehicle Choice (PVC) and Commercial Vehicle Choice (CVC). CARBITS' functionality overlaps the PVC, but each has certain features that the other does not. Specifically, CARBITS can only address currently available vehicle technology options but works at a higher level of detail (vehicle makes, models, and model years), whereas the PVC can estimate choices for vehicle technologies that do not currently exist in the market in appreciable numbers, but works only at the level of vehicle type and size class (such as small cars, midsize sport utility vehicles, and so forth). It is also sensitive to specific policy incentives that could be offered to promote purchase and use of alternative fuels (for example, high-occupancy vehicle lane access).

Yet another important model in the policy landscape that specifically addresses the California vehicle fleet and related fuel usage is the Emissions Factor database (EMFAC), also administered by ARB. Even now, these models are used in conjunction, and these relationships are potentially important to consider–for example, in the ongoing California transportation planning process (CTP 2040),[6] VMT projections from the CSTDM are being configured to be used as an input to EMFAC, and ARB provides functionality that allows CARBITS output on personal vehicle fleet projections to be used as input to EMFAC.

This proliferation of models across state agencies that use similar data as inputs, have overlapping functionality, and produce similar forecasts and projections raises two issues. One issue is the confusion and even potential political tensions or conflicts that can arise if agencies produce inconsistent forecasts and analyses for what appear to be similar policy questions. The other is efficiency, which is the question of whether duplicated effort and resources directed to similar modeling activities might be eliminated through closer coordination. In this regard, there could be an opportunity to identify areas where greater coordination and perhaps even shared effort across agencies in performing modeling activities would address these issues.

Project Motivation, Goals, and Tasks

At a high level, the main motivation for this project was the imminent availability of two new data sets: the *2010-2012 Caltrans Household Travel Survey* (CHTS) and the Energy Commission's California Vehicle Survey (CVS), both of which were completed in 2013. A major ongoing challenge for policy analysts who use models is that models can be criticized as "obsolete" and potentially less credible as time passes since they were last estimated using newly collected data. This is particularly true during periods when the economy, markets, and consumer preferences are perceived to be undergoing major structural change. There is continual pressure to collect new data and re-estimate models; however, this can be difficult due to the high cost of data collection. For this reason, data might be collected infrequently. The availability of these two new survey data sets represents a valuable opportunity.

An important additional feature of these data sets is that they were collected as part of a coordinated effort: households participating in the CVS were recruited from the sample that had already participated in the CHTS, yielding a type of single-source data that is rarely available. This provided an opportunity for considering whether DynaSim models could be improved, not only by updating them using new data, but by considering options for *improved model structures* that recognize the integrated nature

6 Reference: http://www.dot.ca.gov/hq/tpp/californiatransportationplan2040/index.shtml. (Accessed June 30, 2014).

of household travel choices. Finally, these same data are being used to update models used by other state agencies, some of which overlap DynaSim as discussed in the previous section.

With this as background, a summary of project tasks is:

1. Review DynaSim models related to personal vehicle travel.
   - Urban Travel Model
   - Intercity Travel Model
   - Congestion module
   - PVC model
2. Review data elements of the travel-vehicle integrated dataset.
   - CHTS
   - CVS
3. Review other statewide travel and vehicle demand models.
   - Caltrans' CSTDM
   - CARBITS
   - EMFAC (Note: The original plan specified MVSTAFF)
4. Develop model improvement options that address key issues.
   - Integrated vehicle choice and usage models using new travel-vehicle survey data
   - Possibilities for establishing consistency and/or possible integration between inputs and/or outputs of the CSTDM with Energy Commission models of vehicle choice and usage
   - Creating greater (or even complete) consistency between the ARB's CARBITS models and Energy Commission models

Overview of Report

A theme that emerged from the project analysis was that, given the increased availability of many types of data and the substantial recent modeling efforts across agencies, it would make sense to develop and discuss project results from the perspective of a bottom-up modeling framework in which detailed travel behavior is simulated at the level of households and individuals.

For this purpose, the recent work to update the CSTDM for use in statewide transportation planning became a focal point. The CHTS data set plays a central role in these revisions. In addition, an important recurring theme is that the CSTDM produces travel behavior results that have almost complete overlap with DynaSim, with the notable exception of household vehicle type choice. At the same time, the CSTDM is a computationally intensive model with large data input requirements; so, it is practical to only produce these results for specific “horizon years.”

One major recommendation from this project is that, in future modeling, the CSTDM should play an important role as part of the Energy Commission’s modeling efforts. The only question is exactly what the nature and extent of this role should be.

The main reason for this recommendation is that the CSTDM employs a bottom-up approach to modeling travel behavior that approaches what academics would consider to be an ideal standard. This is true because the approach directly addresses many perceived limitations and criticisms of standard modeling practices that have been in common use over the past decades. However, as noted above, there are practical tradeoffs associated with this approach.

To thoroughly address key project objectives, the approach taken in this report is to first develop supporting background material on travel behavior modeling to establish a basis for evaluating alternative modeling approaches. Chapter 2 develops a general framework for travel demand modeling from a bottom-up perspective. The framework is consistent with much of the modeling performed in transportation planning over the past few decades, including the approach used by the CSTDM. For context, the presentation incorporates an example specific to this project. Chapter 3 extends the discussion of Chapter 2 with additional material deemed necessary for the project objectives.

Chapter 4 proceeds with a review and commentary on the current DynaSim models. The primary focus is on the Urban Travel Model, which illuminates key issues related to VMT forecasting, and on the PVC, which relates to the issue of fuel usage. Alternative modeling options are developed during the discussion and evaluation of the current models, as a natural consequence of considering various limitations and features of the approaches. Chapter 5 closes by summarizing the findings and making recommendations.

CHAPTER 2:
A Framework for Modeling Personal Travel and Fuel Usage

This chapter discusses both theory- and data-related issues that are key factors for evaluating modeling alternatives. This discussion provides a basis for reviewing current models and developing alternative modeling options and evaluating tradeoffs.

Bottom-Up versus Top-Down Modeling

An important consideration when developing models is that a *market-level outcome* is the result of an accumulation of outcomes associated with a large number of choices made by decision-making units (DMUs) during a specified period.[7] When fundamental conditions defining a market do not change significantly, forecasting can rely on simple approaches, such as descriptive statistical methods and curve fitting.[8] For example, if the concern is growth in market demand and it can safely be assumed that such growth primarily arises from, for example, an increase in population size, then simple methods can be used. This is particularly true in the short run.

Major issues of concern to policy makers, however, frequently involve factors that will change the fundamental nature and structure of markets, where these changes could occur over extended periods. A fundamental principle in most academic research is that, in these cases, it is critical that models be consistent with an underlying behavioral theory of how, and on what basis, individuals make choices. Specifically, to correctly evaluate the impact of a change in market conditions, estimates of consumer response must be based on a correct understanding of what factors truly affect behavior and *how* they affect behavior. In particular, when developing forecasts or analyzing policies, future market conditions must be correctly characterized in terms of these factors. This characterization is increasingly becoming a formal requirement that government agencies impose on themselves—see, for example, UK Department for Transport (2014).

In a hypothetical ideal case, models could be developed that would literally microsimulate all DMU decisions at all points in time over a full planning horizon and produce results by aggregating the relevant outcome measures from these decisions.[9] As implied, this would require not only a model that accurately captures decision-

7 *Decision Making Units* - a collection or team of individuals who participate in a buyer decision process).

8 *Curve Fitting* - the process of constructing a curve, or mathematical function that has the best fit to a series of data points, possibly subject to constraints.

9 *Microsimulate* – computer simulation at an individual level and each is treated as an autonomous entity.

making behavior at the DMU level, but also a complete characterization of all relevant market factors over the planning horizon. This can be characterized as a bottom-up approach and represents the most extreme implementation of the principle articulated in the previous paragraph.

However, a pure and correct bottom-up approach is essentially impossible due to practical limitations of data availability and existing modeling theories and methods, particularly when addressing future scenarios that include fundamental market changes that do not correspond to current or past conditions. Nonetheless, researchers must make choices and evaluate tradeoffs with this standard in mind.

Faced with this situation, researchers rely on theory and make assumptions to produce simplified models to approximate this process. Hopefully, the assumptions can be justified. These simplifications typically yield models that produce results at a more aggregated level, with the hope that the effect of changing key factors is still adequately represented. For example, to produce a good forecast of total travel demand, there is frequently a distinction between work- and nonwork-related types of travel because they are expected to respond differently to changes in market conditions. Ignoring this distinction creates a risk of aggregation bias. Similarly, avoiding models that separately model travel choices for different household demographic segments, face similar risks because high- and low-income households may respond very differently to changes in conditions.

This issue of bottom-up versus top-down and the level of modeling detail are fundamental to this project. For example, as discussed in Chapter 1, the Urban and Intercity components of DynaSim use top-down approaches that rely on very strong simplifying assumptions and modeling tradeoffs that were made at various times in the past when data availability was more limited. Almost any improvement in the modeling system would rely on some combination of a higher level of detail and tighter integration, both of which will require moving toward an approach with fewer simplifying assumptions, that is, in the direction of a more bottom-up approach.

As a counterpoint, one major project task was to review the CSTDM. Two main features of the CSTDM are that (1) it aspires to use a bottom-up approach and (2) it relies in large measure on one of the two primary data sources motivating this project—the recently collected CHTS. In terms of functionality, the CSTDM is a direct substitute for both the Urban and Intercity Travel Models. However, it uses a highly detailed microsimulation approach that requires substantial computational time and a variety of highly detailed data inputs for alternative scenarios. The next section details the motivation and requirements for this approach.

An Idealized Bottom-Up View of Personal Travel Behavior

As just discussed, an idealized bottom-up modeling approach recognizes that total market-level demand is based on aggregating individual-level decisions. When considering personal transportation choices, each individual may be viewed as belonging to a household, within which a complex array of interrelated decisions are made. Although many adults constitute single-person households, the theory and models must take into account households with multiple members in which adults jointly make key decisions on where and how to live. Such decisions inevitably have an economic dimension in terms of how shared household income is both generated and allocated to activities. Conditional on such factors as household type (for example, single or married, and number of children), educational level, occupation, and age, these decisions include residential location, the choice of whether to work and/or attend school (and where), and a whole range of work- and nonwork-related activities.

One aspect of modeling considered here relies on principles from microeconomic analysis of consumer behavior. That is, it is assumed that shared household income provides a budget that is allocated to those expenditures required to support the household's chosen portfolio of activities. These would include regular day-to-day activities such as working and/or attending school, routine leisure or entertainment (for example, attending concerts or kids' soccer games), and maintenance activities such as grocery shopping, as well as less frequent activities such as business trips, short vacations, and so forth. To be specific, recent trends in academic research suggest that, to capture fundamental behavioral motivations, models should, as much as possible, be activity-based.

Activities of all types require some type of expenditure on goods and services, and frequently, these expenditures include the cost of transporting one or more household members to and from various locations to perform those activities. But, to be clear, moving from one location to another is also an activity that is part of the chosen portfolio. In addition to monetary expenditures, another constrained resource being allocated is time. In an economics-based paradigm, a household is assumed to choose the portfolio of activities, goods, and services that maximizes its utility, where the feasible set of portfolios is determined by budget constraints on both monetary resources and time.

Before discussing models in more detail, first consider an idealized data collection process (survey) that captures detailed information to support bottom-up models for personal travel behavior.[10] The sampling unit is the household, but data would be collected for all household members. Specifically, a large, representative sample of

---

10. Much of this discussion mirrors the type of data collection that occurs in the CHTS; however, it is important to recall that this is an idealized example and not the CHTS process.

households is identified, and detailed data on household characteristics and activities (and in particular travel-related activities) from all household members are collected for a specified set of days (perhaps distributed to cover all days in the calendar year). Assume that these data are collected from *N* households, where households are indexed by $n = 1, \ldots, N$, and a weight w_n is constructed for each household so that the entire sample can be used to represent all personal activities occurring in California for the year and specified geographical subdivisions of California.

To illustrate, consider a simplified example for a household, *H,* (with two adults and no children) that has chosen to live at a specific location in the San Francisco Bay Area and owns two vehicles; a 2008 Toyota Prius and a 2012 Ford Escape. Assume that both adults (denoted *M* and *F*) work full time. Generally speaking, the Toyota Prius is *M*'s regular vehicle, and the Ford Escape is *F*'s. *M* works at a location that cannot easily be reached by transit and usually drives the Toyota Prius when making a typical commute. In addition, *M* also makes regular trips to Sacramento for work. Depending on circumstances, he might drive the Toyota Prius, but taking the train is also a possibility and an option he sometimes uses. *F* can get to work either by taking Bay Area Rapid Transit (BART) with access by walking on both ends or by driving. *F* uses both modes, where the choice depends on the day's activities. For example, on some days, *F* combines the trip to work with grocery shopping or other errands and typically drives the Ford Escape.

In terms of leisure activities, *M* and *F* might use either vehicle for local trips (together or separately) and the occasional weekend trip to locations in other parts of California. *M* and *F* also make occasional long-distance trips outside California for work or leisure, frequently by using air travel. Finally, it is possible that *M* and *F* could make occasional trips to Los Angeles to visit *F*'s mother. These trips could be made by car, but flying is also a possibility.

Next, assume that *H* has participated in the survey. A completed activity diary for each person includes details on each trip taken, including starting location and time, ending location and time, mode of travel (including which vehicle was used and whether the trip was made alone or with others), and trip purpose (activity in the destination location). Assume that this idealized data collection includes full Global Positioning System (GPS) data on vehicles, uses smartphone application software to record data on other activities, and so forth. These data would be supplemented with additional data on vehicle technologies and efficiencies, local fuel prices, information on road and transit networks, and similar information on long-distance travel modes. It would be possible to ascertain exactly how many miles were traveled, the travel time, and travel cost for each trip. For example, for trips involving vehicles, information on vehicle fuel economy, fuel costs, and distance traveled could be used to compute fuel usage. Even more detailed calculations could be possible using information from GPS combined with road network information to take into account speed profiles due to congestion, as well as payment of tolls for bridges, the use of high-occupancy toll lanes, and so forth.

Similar information related to travel times and travel costs for local transit trips would also be available from these household-level data.

With these types of data available, it is worth reviewing the considerations that might go into the development of a behavioral model capable of explaining why any particular household might make the travel choices they do on the basis of various measurable factors. To narrow the scope somewhat, assume for now that the focus is on travel related to a household's daily routine, which would typically involve waking up at home each day and pursuing a pattern of normal activities that concludes with ending the day back at home. For this discussion, the authors assume that these activities produce short-distance trips, while acknowledging that many households might have daily routines that include trips that would be classified as long distance, depending on definitions.

Also, assume that this is being done for the purpose of being able to simulate behavior for a typical weekday in the fall or spring, as is the case with the CSTDM. At a very high level, explanatory factors for travel choice could fall into the following categories:

- Demographic
- Spatial
- Technological
- Economic

First, there will be many factors in this exercise that could be considered predetermined, for example, by prior longer-term decisions. For example, some basic characteristics relate to *M* and *F* themselves:

***M*, male, age 55, Ph.D., engineer, has driver's license, full-time employee**
***F*, female, age 50, M.S., health care professional, has driver's license, full-time employee**

They have no children living at home. They are both employed (which means they routinely make work-related commute trips) in occupations based on past educational and career choices, which would also imply a potential income level. (Their combined total annual income is $350,000.) These are largely demographic characteristics.

When modeling travel choices, the spatial dimension is critical. *M* and *F* live in a single-family home with a detached garage, which is within walking distance of a BART station. *M* has worked at his current job for five years at a location that is many miles away and inaccessible by transit. *F* has worked at her current job for two years, which is located closer to home and is accessible by BART. These location choices represent longer-term decisions.

Over the years, they have owned as many cars as there are drivers in the household (sometimes more), and they currently own a 2008 Toyota Prius and a 2012 Ford Escape. Compared to the previous items, these might represent decisions that are more medium-term, and have more of a technological/economic aspect.

Within this set of circumstances, *M* and *F* choose a portfolio of activities every day that necessarily include trips. It is helpful to review some terminology here. A specific trip might also be called a *trip link* or a *trip leg*, and it is defined by:

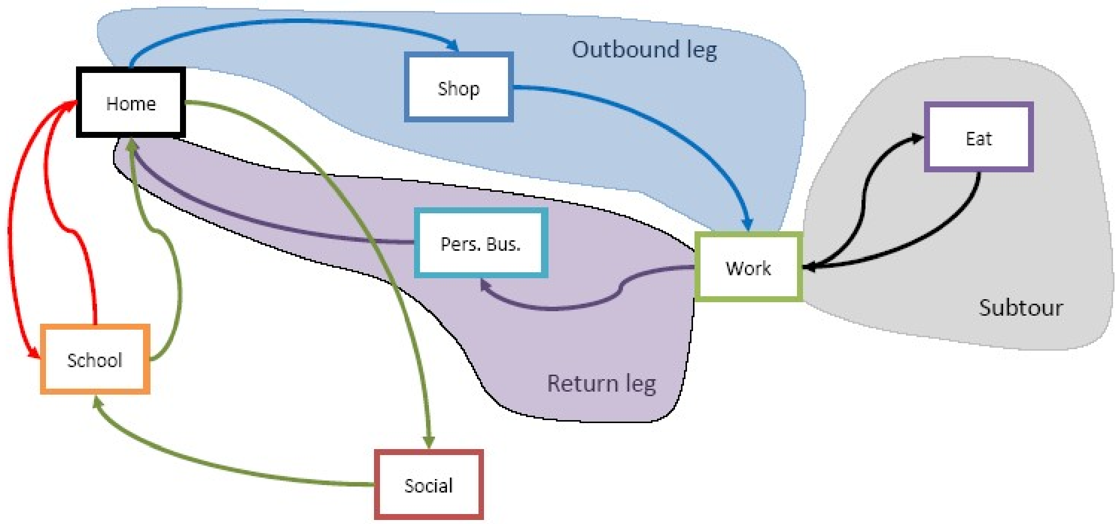
- Starting location and ending location.
- Time of day.
- Mode.
- Purpose.

The CSTDM refers to a closed or half-closed chain of trips starting and ending at home or the workplace (or, perhaps a school location) as a tour. It uses the tour as a key unit of analysis and characterizes the short-distance personal travel model (SDPTM) as a tour-based travel forecasting model. Each tour includes at least one destination and at least two successive trips. A tour is a linked sequence of trips by a person, and these are associated with generic typical day patterns. The following figure from the CSTDM documentation[11] illustrates a typical day pattern with two tours from/to home, and one subtour from/to work.

In keeping with the earlier discussion of the importance of activity-based behavioral considerations, the CSTDM assigns each member of a household to one of 103 day pattern groups. A *day pattern group* is assigned based on a combination of individual and household characteristics, where a major factor is the basic person type of which there are seven categories: preschooler, grade school student, postsecondary student, full-time employee, part-time employee, and senior. Once assigned to a group, each person is randomly assigned a specific day pattern drawn from the CHTS dataset.

11 For this and later references to CSTDM documentation, see ULTRANS and HBA Specto (2011).

Figure 2: A Typical Day Pattern With Tours (from CSTDM)



Source: Cambridge Systematics (2014), "California Statewide Travel Demand Model, Version 2.0"

For the illustrated household, *H*, both *M* and *F* would be assigned the same day pattern. They are full-time employees in a household size of two with no children, working in a non-blue-collar occupation, in the highest household income category. Based on this, they would each be expected to make 3.55 trips and engage primarily in work, maintenance, and discretionary activities (1.10, 0.65, and 0.45, respectively). Trip destinations and mode choices are simulated for each household member based on the assigned day pattern, as well as long list of relevant explanatory variables.

Although *M* and *F* have the same basic day pattern, their actual trip choices could be very different. As noted, *F* has a different work location that is accessible by both auto and BART, and such factors as travel time and travel cost for various choices are quite different from *M*'s. Once at work, access to secondary destinations for nonwork purposes could also be very different.

For the approach used in the CSTDM (for both short- and long-distance travel), trips are characterized at a relatively high level of detail. The model uses four time-of-day periods for defining trips with the noted subdivision for the off-peak period:

- An a.m. peak period (6 a.m. to 10 a.m.)
- A midday period (10 a.m. to 3 p.m.)
- A p.m. peak period (3 p.m. to 7 p.m.)
- An off-peak period (12 a.m. to 6 a.m. plus 7 p.m. to 12:00 a.m.)

Costs that affect trip choices can vary dramatically by time of day, particularly for short-distance travel involving commute trips, travel times, and travel.

In terms of the spatial dimension, starting and ending locations for a trip are defined using 5,191 (internal) traffic analysis zones, as implied earlier. Mode options are defined using a detailed specification of a roadway network (86,000 nodes, 235,000 links, and multiple modes), as well as a transit network that defines options for all air, rail, and bus travel (defining options for both long- and short-distance travel). A trip is short-distance if the centroid-to-centroid distance between zones is within 100 miles; otherwise, it is long distance.

The CSTDM uses a bottom-up microsimulation of individual/household choice behavior that contains a representative agent for every person in California for a given future scenario. Focusing again on short-distance travel, the full range of choices implied in the previous discussion is simulated. For example, work destination location choices for workers, auto ownership levels for households, daily activity patterns, tour choices for work and non-work trips, and individual trips that include time of day, destination, purpose, and mode of travel are all simulated based on behavior factors that include demographics, preferences for destination characteristics for various trip purposes, travel times, and costs for competing travel options that vary by distance between locations.

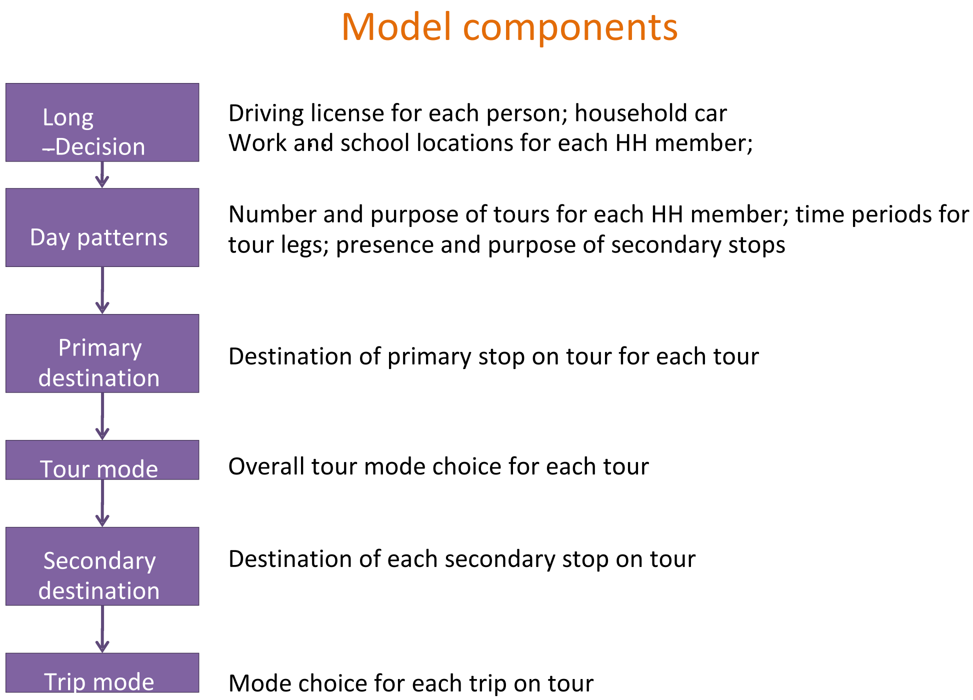
Specifically, short-distance travel choices are microsimulates using a hierarchy of no fewer than six types of choices as depicted in the figure below from CSTDM documentation.[12]

More generally, the CSTDM simulates all travel in California as part of a single integrated system that includes the following five demand models:

- A short-distance personal travel model (intra-California trips) (SDPTM)
- A long-distance personal travel model (intra-California trips) (LDPTM)
- A short-distance commercial vehicle model (intra-California trips) (SDCVM)
- A long-distance commercial vehicle model (intra-California trips) (LDCVM)
- An external vehicle trip model (for trips with an origin and/or destination outside California)

12 See ULTRANS and HBA Specto (2011).

Figure 3: Components of CSTDM Short-Distance Personal Travel Model



Credit: ULTRANS and HBA Specto (2011).

Because all travel is being simulated, the model loads all travel onto the road and transit networks, so that supply and demand interact. The demand on the road and transit system determines travel times, that is, congestion effects are simulated. However, because travel times affect travel choices, the model must be iterated to achieve equilibrium between supply and demand so that travel times and travel choices are internally consistent.

The input requirements for analyzing future scenarios demand a very large and detailed set of data, described in the following excerpts from CSTDM documentation:[13]

- Demographic data for each traffic analysis zone (population and household characteristics, employment by industry and occupation, and school enrollment)
- Other zonal attributes (area, area type, population and population density, parking costs, region)
- Travel cost data (fuel costs, public transit fares, road tolls)
- Commodity flow movements (for the LDCVM)

13 See ULTRANS and HBA Specto (2011).

- Observed vehicle flows and growth factors by type and period (for the external vehicle trip model)
- Traffic analysis zone to traffic analysis zone, travel times, and costs by mode and period (obtained using detailed network specifications)

> "The model is implemented using a software environment called CUBE, which includes packages related to transportation modeling, as well as some additional code written in Python and Java. Road network descriptions for each time period are coded in the standard CUBE format. For the road network, all freeway, expressway, and most arterial roadways are explicitly represented, with collector and local roads mostly covered through zone centroid connector links. Link distances, free flow speeds and capacities are explicitly coded…
>
> For public transit, all air and rail lines and services are explicitly coded using the standard CUBE format. For local bus transit, a simplified model is used to give level of service times and costs, based on road network speeds, land use variables, and transit operator service measures. Walk and bicycle times are derived from road network distances."

The CSTDM was designed to rely on detailed inputs for future land use patterns and transport infrastructure and uses projections of population trends and economic forecasts for simulating a synthetic population of households and individuals that then make all their choices on the basis of behavioral models. This approach is very burdensome with regard to creating input files, and the model takes a long time to run. It is applied at a few key milestone years for a relatively small number of scenarios.

As detailed as it is, the CSTDM focuses primarily on estimating trips at the level of mode choice, and completely excludes the dimension of vehicle type choice. However, the Energy Commission is being held responsible for projecting future fuel consumption, and there is a potential flaw in any method that arbitrarily separates specific vehicles on the road (in terms of type and vintage) from how they are used. The ARB faces similar issues when projecting GHG emissions.[14] At the same time, the vast majority of travel demand models take a similar approach, as discussed in the next section.

It would be possible to go into much more detail regarding the modeling approach used in the CSTDM, but the CSTDM will be discussed again later. The main purpose of this section was to provide a background on what a bottom-up modeling approach would look like.

14 Chapter 3 discusses this issue in more detail.

Travel Demand Modeling Principles and Tradeoffs

The idealized example in the previous section illustrates a bottom-up perspective on travel behavior based on collecting detailed data at the individual/household level, including a qualitative identification of various theory-based factors that would play a role in behavioral models for simulating a hierarchy of interrelated choices. Aspects of the CSTDM were incorporated into the discussion to serve multiple project-related purposes. Short of giving additional details on specific models, it is clear that implementing such an approach could potentially require many resources in terms of data, modeling, and computation, which have possible implications for the required levels of ongoing staff support. The basic tradeoff when considering alternative modeling options involves making assumptions that allow simpler versions of models to be used in place of pure microsimulation.

Over the past decades, a relatively consistent set of principles and practices has developed in response to the inherent tradeoffs discussed. In this regard, the *Handbook of Transport Modelling* edited by David A. Hensher and Kenneth J. Button (2008) has multiple relevant chapters. Chapters by Michael G. McNally (2008a) and John Bates (2008) both discuss and evaluate standard practice in the context of the "four-step model" (or "four-stage model"). The beginning of McNally, Michael G. (2008a), which focuses specifically on this model, states "The history of demand modeling for personal travel has been dominated by the four-step model." The chapter provides a concise overview and case study and emphasizes that, despite the known deficiencies of the approach, it is nevertheless widely used due to current institutional requirements and financial limitations. He concludes with the hope that future methods will be based on activity-based approaches, which are described in Michael G. McNally (2008b). Many of the features built into the CSTDM intend to address the known deficiencies of the four-step model.

McNally's two chapters build on the preceding chapter by Bates, which provides a more general perspective on transport demand modeling. It first focuses on the principles that should drive model development, then discusses the four-step model within this context. What emerges is that (1) these principles have been known and accepted for some time; (2) specific criticisms of the four-step model relate much more to the details of implementation rather than the basic structure; and (3) these are frequently attributable to modeling decisions made in response to resource limitations. For example, both McNally and Bates discuss how the four-step model has been highly criticized for being trip-based rather than activity-based, but it is clear that this is not an inherent deficiency of the four-step model.

Material from Bates is highly relevant to the purposes of this research and useful enough to warrant the following high-level summary.

Bates articulates basic principles and offers a framework for travel demand modeling that helps further clarify those key issues discussed in the previous section. Here are some excerpts (where emphasis has been added to highlight key concepts):

1. "[W]e must always remind ourselves that ***travel is a 'derived' demand***: travel is not demanded *per se*, but as a consequence of the desire to partake of ***activities*** in different ***locations.***"

2. "Since, in addition to costing ***money***, travelling between locations inevitably involves an ***expenditure of time***, it has become standard in transport economics to deal with so-called ***'generalized cost',*** which specifically recognizes both kinds of expenditure. In its simplest form, generalized cost is a linear combination of cost and time, the latter being converted to money units by means of the so-called ***'value of travel time savings.'"***

3. "***Spatial separation is the essence of travel demand***, and the majority of models aim to recognize the spatial distribution of travel explicitly, by means of ***an appropriate system of zones.*** The modeling of 'demand' then implies a procedure for predicting what travel decisions people would wish to make, given the generalized cost of all alternatives. The decisions include choice of time of travel, route, mode, destination, and frequency or trip suppression."

4. "The preferred approach nowadays is to set up ***'choice hierarchies'*** making use of ***discrete choice theory.*** This allows the 'lower level' choices to be made conditional on higher choices (e.g., mode choice might be assumed to be conditional on destination choice) in a theoretically consistent way…"

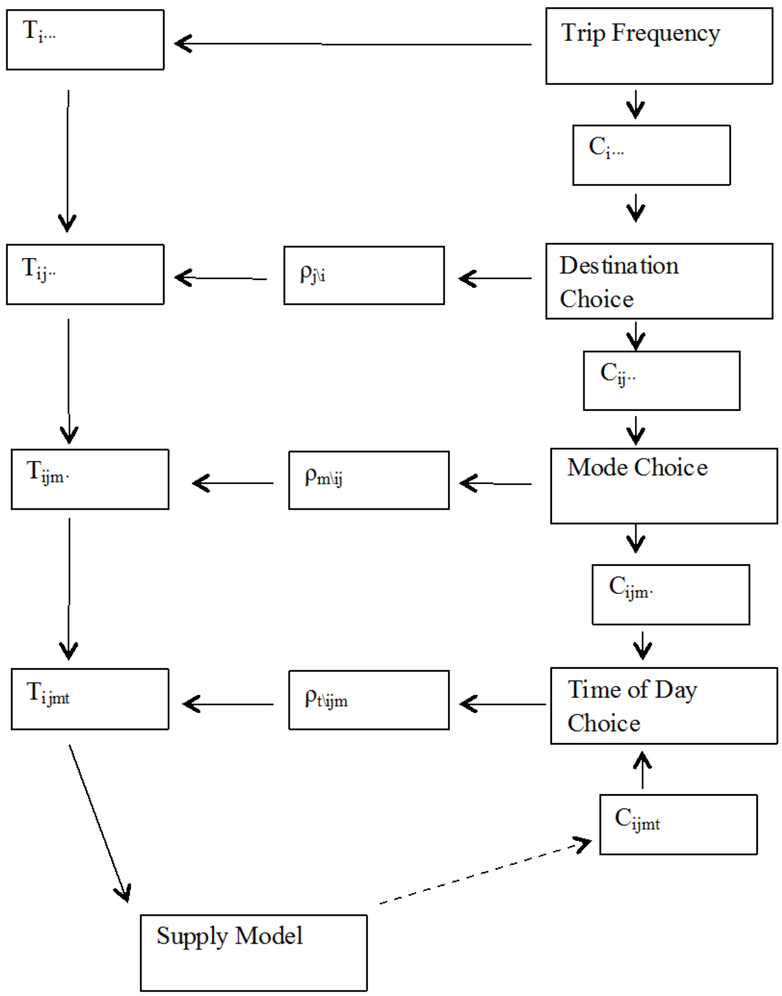
5. "The overall level of demand clearly depends not only on the costs that are directly related to the transport system but also on those factors that relate to the ***demographic composition of the population***, together with other 'external' changes (e.g., effects due to ***land use, income***)."

6. "In particular, it is well established that the level of ***car ownership is a key determinant of demand.***"

7. "***[D]ifferent persons have different basic demands for travel.*** For example, employed persons need to get to work, children need to get to school, retired people have more free time, etc. …[therefore] ***it is sensible to take reasonable account of this variation by person type between areas, or "zones."***

To illustrate some of these principles, Bates provides an example of a hierarchical demand model that, at a high level, can be used as the basis for developing and evaluating modeling options. (See **Figure 4**.)

Figure 4: Example of Hierarchical Demand Model From John Bates (2008)



Source: Bates, John (2008), "History of Demand Modeling"

In this example, the main output is a *trip matrix* T_{ijmt} that provides trip count totals for a typical day, where: i = origin zone, j = destination zone, m = mode, t = time of day. This illustrates excerpts 1 and 3. In terms of behavioral models, it is possible to think of these trip counts as occurring at varying levels of aggregation, all the way down to an individual.

The final output can be considered as arising from interrelated choices that can be structured as conditional choice probabilities (excerpts 3 and 4). For example, the probability of choosing which *time of day* to make a trip can be modeled as being conditional on other choices that have already been made for mode and destination

($p_{t|ijm}$).[15] The trip total for T_{ijmt} is given by the number of trips estimated for *(i, j, m)*, i.e., T_{ijm*}, multiplied by $p_{t|ijm}$, and so on.

The factor determining choice from among competing time-of-day options is characterized as a generalized cost C_{ijmt} (an overall measure of disutility), as discussed in excerpt 2. At the next higher level, the generalized costs appropriate for evaluating competing *modes* (on an overall basis) must adequately address the fact that, in this choice structure, the competing modes have multiple generalized costs for different times of the day. In other words, one must compute an average or representative generalized cost C_{ijm*}. In the context of discrete choice models, this is known as an *inclusive value*, and it represents a way of implementing the "feedback" arrow in **Figure 4** to capture dependencies and relationships among choice dimensions.

At the top level of **Figure 4**, the total number of trips generated from a particular origin (T_{i***}) is conditional on factors related to the origin (*i*), plus some overall level of the attractiveness for making trips, conditional on characteristics of the entire rest of the system (destinations, modes, and so forth). Excerpts 1, 5, 6, and 7 discuss the various factors that affect these measures. Although they are not always explicitly identified using this notation, they are readily incorporated within this framework when developing models.

In addition to the feedback effects for the generalized costs that have already been discussed, a key feature is the feedback effect from the *supply model* that shows how the trip matrix *T* will have an impact on the generalized cost at the bottom level (and therefore at all of the upper levels as well). This feature allows the model to be iterated to an equilibrium solution with internally consistent generalized costs and travel choices.

With this as a standard of comparison, Bates proceeds to discuss the four-step model. The four stages relate to:

1. Trip generation (and attraction)
2. Trip distribution
3. Modal split
4. Assignment

Bates notes that, in relation to **Figure 4**, "…[T]he top three models on the right-hand side correspond with the stages of trip generation, trip distribution, and modal split." The "time-of-day choice" model can be seen as a "fifth stage" that was initially ignored, but has more recently been considered important to include. He also notes that, because

15 The depicted structure does not actually mean that choices are literally made in a sequential fashion. Rather, this is one possible mathematical structure that captures the possible behavioral interdependencies among the choices.

car ownership levels play such an important role in this process, a model for this is frequently included.

For completeness, Bates gives additional details on the assignment stage, which relates the trip matrix (determined in steps 1-3) to the transportation network. Because many of the applications of the four-step model are evaluating large infrastructure projects, this has been a major focus. Trips are loaded onto the appropriate networks (for example, highway network for car trips, plus various competing transit modes where appropriate) to determine network performance in terms of link flows. For cars, this can involve “route choice” or “route assignment,” and various algorithms are available to produce travel times under the assumption that equilibrium between “demand” (represented by the trip matrix) and “supply” (the transportation network) occurs.[16]

Once the assignment step has been performed, it is possible to generate an impedance matrix that associates an average travel time for each *(i, j, m, t)* trip combination. This, of course, is one component of the generalized cost that is a required *input* for determining *T*! However, this is routinely ignored in practical applications of the four-step model for reasons of practicality, and *omitting feedback effects* is one of the major criticisms of the four-step model. On this specific point, Bates indicates that, if feedback effects are properly included, this would address much of the criticism leveled at the four-step model.

However, Bates offers an additional comment that is potentially *very important for this project*:

> In addition, the earliest versions of this model were applied at an extremely aggregated level, without taking account of variations in purpose, person type, etc. Although a defining characteristic of a typical four-stage model is its fairly detailed network representation, with consequent implications for the number of zones, which may limit the amount of traveler-type disaggregation that is feasible in practice, more recent versions of the model do include a reasonable amount of travel variation.

With these issues in mind, options for how to model and forecast a trip matrix *T* were explored. But first, to explicitly review an earlier point, in a bottom-up approach, it is straightforward to produce *T* by aggregating together all of the simulated trips made by individuals that fall into each category *(i, j, m, t)*. Trips by individuals are simulated as arising from a choice hierarchy of the type discussed previously, using models that address the principles articulated above. Moreover, in the CSTDM approach, the total

---

16 This is the type of modeling and analysis that the CUBE system (as well as other commercially available systems) is designed to address.

modeling system is iterated to achieve equilibrium with internally consistent results, which explicitly addresses the feedback issue highlighted by Bates.

However, this does not represent the usual state of practice. In what follows, additional useful material is summarized, drawing in part from *Modeling Transport* by Juan de Dios Ortúzar and Luis G. Willumsen (2001).

To review, Bates' perspective is that models should include:

- Sufficient spatial detail (*zones* and *networks*).
- Sufficient detail on *person types* and *trip purposes* to capture travel variation.
- *Feedback effects* to capture *interrelationships among choices*.

With regard to the first two points, additional dimensions of disaggregation were considered by adding two new indices to the framework: n = person type, and p = purpose. A more granular model would then involve generating trip counts of the form T_{ijmt}^{np}. In this approach, a wide range of detail is possible depending on the number of person types and the number of zones, but this still would not represent a microsimulation approach. A helpful extension is to include a dimension h = household type and information on household membership (see below). Household types could be defined by attributes such as income level, car ownership level, household size/structure, and so forth.

The first two steps of the four-step model are trip generation and trip distribution to produce a trip matrix T_{ij} before any consideration of mode choice. Because any zone can be both an origin and a destination, the main distinction is the direction of flow. From a behavioral perspective, researchers consider the concepts of *production* and *attraction* more useful. For example, household members typically start their day in their zone of residence, and activity patterns that require trips arise with some frequency. The simplest example is the need to commute to work. In the early part of a typical day, this would generate a *home-based work* (HBW) trip (perhaps during a period of peak travel) from i to j (the destination zone of the work location), and later in the day there would typically be a return non-home-based trip from j to i. Other types of trips occur for nonwork purposes and exhibit more variability than patterns related to work and attending school.

Details on specific trips are collected in household surveys. However, the goal is to develop models that can adequately capture total trip counts for future scenarios. In such models, researchers generally believe that trip production is better understood as a function of household characteristics, whereas attractiveness of destinations for alternative purposes is less understood, and characterized by, for example, land-use-related measures. So, trip generation is framed as production of trip frequencies from an origin zone, followed by a destination choice. In this regard, household survey data are helpful for estimating trip production models, but destination choice models require other data related to zonal characteristics. At the same time, vehicle traffic count data, transit ridership, and other variables can give more accurate information on actual network use. So, model development can involve a combination of modeling approaches and data sources, where behavioral models are estimated using household survey data, but then the overall modeling system requires calibration to match known performance of actual networks.

To proceed with a specific example, let $w_i(h)$ be the number of households of type *h* in zone *i*, and let *H(n)* be the set of households of type *h* containing persons of type *n*. Then, the estimated production of trips with purpose *p* by persons of type *n* in zone *i* can be expressed as:

$$O_i^{np} = \sum_{h \in H(n)} w_i(h) t^{np}(h)$$

where $t^{np}(h)$ is the expected number of trips with purpose *p* for person type *n* in household type *h*. The estimated total trips for purpose *p* by person type *n* would be given by:

$$T_{ij**}^{np} = O_i^{np} D_j^{np} \exp(-\lambda C_{ij**}^{n})$$

where D_j^{np} is a measure of attractiveness of destination *j* for *np*, and C_{ij**}^{n} is a generalized cost (or impedance) for travel between zones *i* and *j* for person type *n*. Additional complexity is possible by incorporating measures of accessibility into the *O* and *D* terms, respectively. Even more complexity arises if it is not possible to specify the *O* and *D* terms separately, that is, the approach of modeling production of trip origins and attraction of destinations is an oversimplification of the behavioral factors that give rise to final trip patterns.

This model takes a form similar to a variation of the well-known gravity model:

$$T_{ij} = \frac{O_i D_j}{d_{ij}^2}$$

where d_{ij} is the distance between zones *i* and *j*, the inverse square of distance captures impedance, and the distinction between origin and destination has been maintained with regard to causes of attractiveness. Although the focus here is on short-term travel choices, these same general forms can be applied to long-distance/intercity travel choices as well.

Allocating the trips to competing modes would be a matter of specifying an appropriate discrete choice model based on generalized costs and performing the multiplication discussed previously. The model could be a simple logit model, but more recently researchers have used nested logit to capture similarities among groups of models that are likely to be closer substitutes.

One important point that can be discussed using this framework involves how such models (and the associated data) can be used for forecasting. In a future forecast year, many important factors may have changed. For example, based on population and employment projections, it would be possible to update the $w_i(h)$ factors that specify how many households of type *h* are living in zone *i*. Appropriate projections would take into account both population size and changes in income distribution.[17] Any changes to the transportation infrastructure would need to be taken into account. For example, adding a rail stop in a zone could effectively add another mode option that did not exist before. Improved level of service or widened highways would affect travel time estimates and, therefore, generalized costs. Changes in gasoline prices or transit fares would also affect generalized costs.

Aside from the concern about whether it would be possible to ensure equilibrium between travel choices and transport network performance, the obvious way to produce a forecast would be to use models that have been estimated using observed trip matrices from a base year. In general, however, estimated models do not exactly reproduce the dependent variable data used to estimate them: there are residual differences that reflect random effects but also possibly systematic factors that are not captured by the explanatory variables of the model.

In these cases, it is common practice to use *pivot point* methods to preserve all of the original information in the base year observed data. In this case, forecasted values for the trip matrix are given by:

$$T_{ij**}^{np}(f) = T_{ij**}^{np}(b)\frac{O_i^{np}(f)D_j^{np}(f)\exp(-\lambda C_{ij**}^{n}(f))}{O_i^{np}(b)D_j^{np}(b)\exp(-\lambda C_{ij**}^{n}(b))}$$

where *"f"* denotes the forecast year, and *"b"* denotes the base year. In this case, $T_{ij**}^{np}(b)$ represents the actual observed counts from a base year data set that may have been used to perform model estimation. (They could also be adjusted or calibrated to take into account observed traffic data.)

In later chapters, this material will be useful for identifying and characterizing alternative modeling options. However, there is an additional area that requires exploration.

---

17 More complex forms of demographic projection could address other dimensions such as household size, number and ages of adults and children, and so forth.

CHAPTER 3: Car Ownership, Vehicle Type Choice, Vehicle Miles Traveled, and Fuel Usage

The domain of travel modeling covered by Chapter 2 of this report arguably represents a large percentage of modeling performed by professional transportation planners, which historically has been largely oriented toward managing local or regional transportation infrastructure within a state context. A major motivation for providing this material is that this domain is highly consistent with the currently specified functionality of both the CSTDM and DynaSim (with the notable exception of the PVC). In fact, a review of the detailed guidelines for regional transportation planning in California reveals that they are in large measure a more detailed version of the material in Chapter 2—see California Transportation Commission (2010).

To review, characteristics of the Chapter 2 modeling framework include the following:

- Key outputs of behavioral models are person-trips.
- Personal vehicle usage is measured in terms of mode choice for trips.
- Household vehicle choice focuses on car ownership levels.
- An important concern is how travel choices and the performance of the transportation network interact with one another.

Despite the overlap, there are additional considerations that must be addressed to fully support a discussion and evaluation of modeling options that meet the Energy Commission's requirements. By way of review, these include:

- Producing forecasts of future transportation fuel demand.
- Examining implications of factors that influence changes in demand.
- Assessing the future role of alternative fuels, including barriers to and progress toward meeting California's transportation energy goals.

Formally, the Energy Commission also produces fuel price forecasts, which would be another obvious requirement of forecasting fuel demand.

Aside from any concerns about the quality of the models themselves, additional discussion is warranted on the following questions:

- How is VMT determined?
- How is fuel usage computed?
- How are these addressed for alternative fuels?

Before proceeding, note the close correspondence between computing fuel usage and greenhouse gas emissions. In particular, ARB's EMFAC model (see the next section) includes factors that allow calculation of fuel usage based on fuel carbon content.

First, with regard to the VMT question, Bates provides a succinct description that holds true for both DynaSim and the CSTDM: "Private trips, estimated on a *person* basis, are converted into vehicle trips by adjusting for average occupancy…" The final step is to convert vehicle trips to VMT by using information on trip distances.

To compute fuel usage, this conversion would normally be considered a simple computation involving vehicle efficiency and VMT if there were full data available for vehicles and how they are used. A potential issue is the wide variation in vehicle fuel efficiency across vehicle types and vintages in the marketplace, but recall information from the model framework of Chapter 2 is limited to *mode choice*.

For illustration, consider an expanded model using the notation of Chapter 2, where the set of modes (indexed by m) is replaced by a more detailed set of options that incorporates information on a household's vehicle holdings. Specifically, recall the trip matrix notation T_{ijmt}^{np}, where i = origin zone, j = destination zone, m = mode, t = time of day, n = person type, p = trip purpose, and, if needed, h = household type (not shown).

Using this notation, trips made by auto ("a") could be represented by $T_{ij[m=a]t}^{np}$. However, the notation could be expanded to include more detailed information on household vehicle types and the assignment to trips, e.g., $T_{ij[m=vfy]t}^{np}$ where v = vehicle type, f = fuel type, and y = model year. In this case, total fuel usage F could be computed as:

$$F = \sum_{vfyt} \left[e_{vfyt} \sum_{n,p,i,j} T_{ij[m=vfy]t}^{np} d_{ij[m=a]t} \right]$$

where $d_{ij[m=a]t}$ denotes the distance traveled (miles) between i and j when the trip is made during time period t, and e_{vfyt} denotes the fuel consumption rate (gallons per mile) for vehicle "vfy" driven during time period t.[18]

However, this level of detail (i, j, m,t,n, n,p,h,v,f) does not exist in the Chapter 2 framework, and in particular, it does not exist in the CSTDM. Similarly, in DynaSim software, it is the Urban and Intercity Travel Demand Models that provide the estimate of total VMT for all personal LDVs in the system.

In summary, the current capabilities of these models preclude the possibility of computing fuel usage using the above equation.

---

18 At the risk of creating confusion, this example includes the notion that fuel consumption rates could vary by time of day for reasons that will be more apparent later. The assumption is that average speed of travel could vary by time of day, affecting fuel consumption rates. A different version using more detailed information on trip speeds for specific origin-destination pairs was another possibility.

That being said, both models in the current forms are being used to produce forecasts that should, in theory, use the above equation. Specifically, the purpose of DynaSim is to compute fuel usage forecasts, and CSTDM VMT output is being used to compute GHG emissions using EMFAC. Because the above equation cannot be used, both models rely on approximations and assumptions that are potentially suspect.

This observation relates to one of the major issues being addressed in this project: the lack of integration between households' choices of which vehicles to buy, and how they are used. On this specific issue, a major question is whether the assumptions and approximations being used to produce fuel/greenhouse forecasts are generally valid, or whether they could lead to misleading results. One important factor is the type of policy scenario being evaluated, particularly if it involves future alternative fuel vehicles.

The next two sections provide additional detail on how the current versions of the CSTDM and DynaSim are being used to compute important forecast measures.

California Statewide Travel Demand Model and EMFAC in the California Transportation Plan 2040

As a starting point, aspects of ARB's EMFAC model were reviewed. This is important in its own right and is relevant because producing estimates of GHG emissions is functionally equivalent to fuel usage. The following discussion is based on documentation of the most recent version EMFAC 2011—see California Environmental Protection Agency-Air Resources Board (2011a, 2011b).

EMFAC characterizes vehicles according a particular classification system, which, for the purposes of this report, is limited to LDVs of the type listed in **Table 2**.

Table 2: Excerpt of EMFAC Light-Duty Vehicle Classes

Vehicle Class	Fuel Type	Code	Description	Weight Class (lbs)	Abbr.
1	All*	PC	Passenger Cars	All	LDA
2	All*	T1	Light-Duty Trucks	0-3750	LDT1
3	Gas, Diesel	T2	Light-Duty Trucks	3751-5750	LDT2
4	Gas, Diesel	T3	Medium-Duty Trucks	5751-8500	MDV
5	Gas, Diesel	T4	Light-Heavy-Duty Trucks	8501-10000	LHD1
6	Gas, Diesel	T5	Light-Heavy-Duty Trucks	10001-14000	LHD2

*gas, diesel, and electric
Source: California Energy Commission staff

The current version employs a detailed analysis of California Department of Motor Vehicle (DMV) and smog check program data from 2009. Vehicle identification number (VIN) decoding is used to classify the existing vehicle population into the above classes for 45 age groups in 69 subregions of California. Vehicle classification using VIN decoding has a variety of issues when it comes to correct assignment of vehicle technologies (for example, VIN-based information is frequently missing information on transmissions and drive trains), and ARB staff developed an improved methodology that involved matching DMV records to smog check compliance data, which fills in missing gaps on technology. These data also provide information on odometer readings that give information on VMT patterns as a function of vehicle type and age.

Emission rate factors (including factors that can be used to estimate fuel usage based on carbon content) are also established. Generally speaking, vehicle emissions can vary depending on the speed at which the vehicle is driven, so EMFAC determines emissions rates for different speed categories. Denote these emission rate factors by e_{vafsy}, where v = vehicle class, f = fuel type, a = (vehicle) age, and s = speed. The index y = (calendar) year has been included because EMFAC also provides *projected* estimates of these factors for future years. This is necessary because emissions are a function of both vehicle model year and age. These factors are representative values for a more diverse underlying set of specific vehicle makes and models, not only for the base year, but for projected years. In other words, EMFAC essentially employs a hard-wired set of assumptions regarding vehicle type choices in California by projecting base year vehicle fleet characteristics out into the future.

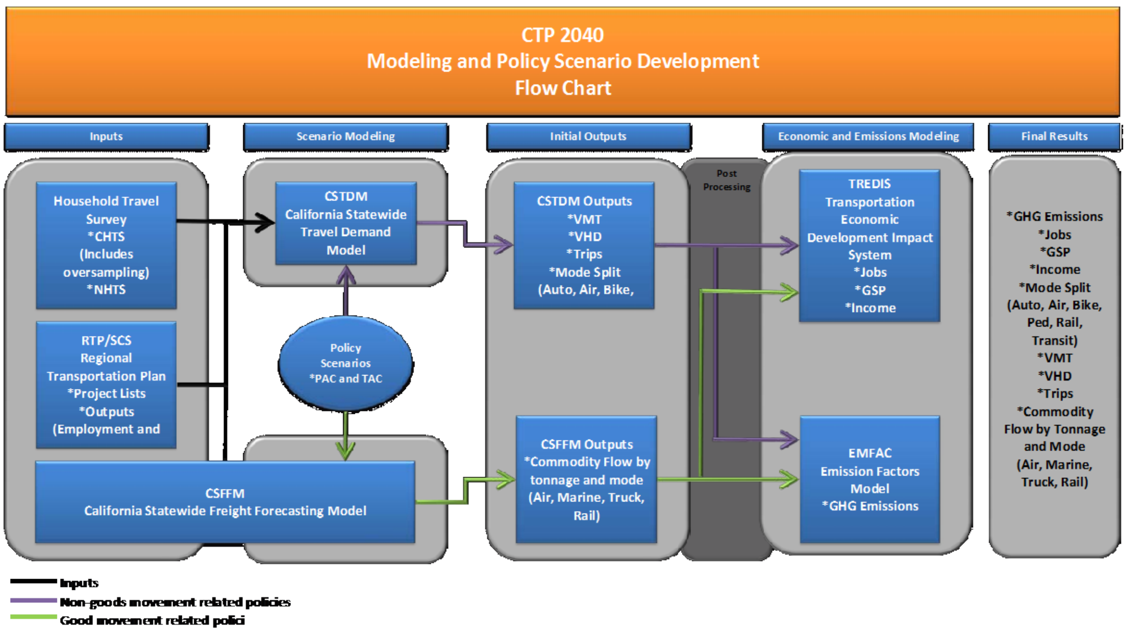
For EMFAC to have practical value, it also maintains data that allows it to be used as a tool by various stakeholders for policy analysis. For example, because emissions vary by speed, it is necessary to have information on VMT speed distributions that, in general, could be difficult to come by otherwise. According to the documentation:

> For air quality and transportation planning purposes, EMFAC2011-LDV uses the VMT provided by regional transportation planning agencies (RTPA). For EMFAC 2011, ARB received VMT and speed submittals from the Southern California Association of Governments (SCAG), Bay Area Metropolitan Transportation Commission (MTC), San Diego Association of Governments (SANDAG), and San Joaquin Valley Councils of Government. In the absence of recent RTPA data, the model contains default speed distributions and estimated VMT as a function of vehicle population (from DMV) and mileage accrual rates (from the Bureau of Automotive Repair Smog Check program).

As noted, EMFAC includes a full set of default forecasts on the future vehicle fleet population defined according to vehicle class, including information on how VMT varies by age, and the aforementioned VMT and speed estimates, for future calendar years. Based on these considerations, an example of the most accurate fuel usage computation obtainable from EMFAC would involve running the model using base-year values for a specific region where the quality of data submitted to ARB is considered to be very high.

For policy analysis, EMFAC has a variety of modes whereby users can submit their own input in varying levels of detail. In this regard, the authors discuss the modeling and scenario planning process adopted in the most recent statewide transportation planning process (see **Figure 5**). It shows that CSTDM VMT outputs are used to produce GHG emissions estimates using EMFAC.

Figure 5: Relationship between CSTDM and EMFAC in CTP 2040



How would this calculation be performed, and how accurate would it be? For discussion purposes, assume that one uses CSTDM output for the entire fleet of light-duty vehicles for a *"typical week day"* in a specified year. What would be required is a program that aggregates trip counts into a trip matrix, where trip travel times and distances are used to assign trips to speed bins in addition to time of day. Following the notation in Chapter 2, this would take the form T_{ijmts}, where the aggregation has been performed over all trip purposes and person types, and the notation has been augmented by adding *s* (speed). In addition, the program would produce the matrix d_{ijm} (distance between *i* and *j* for mode *m*).

Recall that the CSTDM generates trips only at the level of car-based modes, not vehicle types and vintages. For ease of discussion, assume the data have been collapsed into one mode for car, denoted *c*. Based on the EMFAC documentation, there are several options that could be used, where various options implicitly allow the user to rely on built-in EMFAC assumptions regarding the vehicle fleet. This would absolutely be required in this case, because CSTDM includes no information on vehicle types or vintages.

In summary:

- The CSTDM can produce estimates of VMT totals and has the capability of subdividing VMT estimates by speed category. In this regard, it is capable of evaluating congestion issues associated with alternative scenarios because the CSTDM contains an equilibration feature for the transportation network.
- There is no vehicle type choice model, and it cannot capture any possible relationships between vehicle characteristics and trip choices (include trip lengths). It relies on EMFAC assumptions for all these relationships, which are hard-wired and based on simple extrapolation of historical data. On the positive side, this means that these calculations would appear to avoid an even simpler assumption, that is, that all vehicles are driven exactly the same distances. However, again, this behavioral assumption is built into EMFAC and is not a behavior that is explicitly modeled by the CSTDM.
- The combined CSTDM–EMFAC approach has little or no ability to address issues related to alternative fuels. Adding this capability would require new modeling features of some type to be introduced into the process.[19]

Remarks About CARBITS

In addition to EMFAC, ARB also uses the CARBITS model. This is a stand-alone vehicle choice model that could be used to produce alternative estimates of future vehicle population distributions. In fact, ARB's website includes a tool that can be used to perform CARBITS runs and then import the updated vehicle distributions into EMFAC.

CARBITS has no capability to address future alternative fuel technologies. One notable feature of CARBITS is that it operates at a very high level of detail with regard to vehicle choices, in other words, at the level of year/make/model. But the results from a CARBITS run must be aggregated to the level of EMFAC classes, and to the authors' knowledge, the emissions rates are still those that are based on EMFAC's original analysis of vehicle distributions in the base year. In other words, the full vehicle distributions from CARBITS cannot be used to produce consistent emissions factors. Finally, CARBITS addresses only vehicle choice and not usage. Similar to the emissions factors, employing CARBITS results in a process of estimating fuel usage/greenhouse gas emissions would rely on EMFAC or some other model for this information.

19 It is possible that someone with a deeper technical knowledge of these models (particularly EMFAC) might disagree with this assessment. However, this represents the authors' best understanding and judgment, and it seems important to highlight this point.

DynaSim Fuel Calculations

As discussed, DynaSim determines total personal VMT using the Urban and Intercity Travel Models for short and long distances, respectively.[20] Similar to the CSTDM, the VMT is obtained for a generic auto mode with no additional detail on vehicle type or vintage.

One obvious observation is that VMT estimates could be used as inputs to EMFAC to estimate fuel usage. However, these VMTs are more limited than those of the CSTDM in at least two ways: (1) they would be limited to daily totals, that is, there is no time-of-day dimension and (2) allocating VMT to different speed bins would be problematic.

In particular, DynaSim does not really model spatial effects (see Chapter 4). Specifically, the notation developed in Chapter 2 for trip counts involving origins and destinations (and therefore trip distances) does not exist in DynaSim. And, even if some approximation were available based on, for example, past data containing speed distributions, it is clear that DynaSim would not have the ability to equilibrate travel times on a travel network as does the CSTDM (because there is no travel network).

But, in contrast to the CSTDM, DynaSim does have a detailed vehicle choice model that can produce estimates of vehicle fuel and technology types and vintages for households. Moreover, the choice model has the capability of including alternative fuel vehicle types that do not exist in the marketplace.

However, the current versions of travel demand models in DynaSim make a very strong assumption: *All vehicles, regardless of type, vintage, technology, and so forth are all driven the same number of miles.*

DynaSim therefore computes fuel usage based on the following approach:

1. Obtain estimates of total VMT from the Urban and Intercity Travel Models.
2. Allocate VMT to all vehicles equally, that is, proportionally to all vehicle types and vintages based on vehicle counts.
3. Use nominal fuel efficiencies from a database of vehicle attributes to compute fuel usage.

Based on this discussion, it seems clear that this calculation could be improved by adding some type of information that would capture differences in VMT patterns by vehicle type and vintage, similar to EMFAC. The Energy Commission regularly performs its own analysis of DMV data with regard to the vehicle fleet population for its own system of vehicle classes and vintages; however, this analysis does not address the issue of VMT. It might be possible to develop a procedure similar to ARB's that would take

20 This is discussed in more detail in the next chapter.

into account odometer readings from smog check data, or there could be other data sources available for this purpose. (See the next section.)

Another possibility is that this could represent an opportunity to use a VMT model of the type that was previously dropped from the personal vehicle choice (PVC) model in DynaSim. Specifically, the previous light-duty vehicle VMT model that was estimated in conjunction with the vehicle choice model was dropped. The model could be reframed to be used for allocating VMT across personal light-duty vehicles. This will be discussed later in the "Personal Vehicle Choice Model" section in Chapter 4.

Additional Remarks on Vehicle Type Choice and Vehicle Miles Traveled

The information in the previous sections was important to document for multiple reasons. When considering these issues, it is important to keep in mind that vehicle choice and VMT forecasting are inherently challenging, and multiple modeling approaches can be considered for this purpose.[21]

The perspective that influences the current approaches used by the CSTDM and DynaSim is discussed in Chapter 2. Bates provides some opinions that shed light on this perspective. In discussing car ownership levels, he reviews that key determinants are household demographics (such as income, household size, number of adults, number of workers) and emphasizes the importance of income. However, he also opines that attempts to include other, more detailed effects have not been established as empirically important. These effects include the effect of prices (capital and operating costs) on car ownership levels and a more detailed treatment of vehicle type choice. In his view, although including vehicle type choice "can be significant in predicting environmental effects, it contributes little to the overall question of travel demand, where there is little reason to discriminate between various types of car." In other words, the main thing to model is total travel demand, and it is reasonable to assume that there is a weak interaction between demand and the type of car.

This view succinctly captures a key issue for this project: for many purposes, it may be sufficient to capture overall VMT demand; however, how this VMT is distributed is relevant when it comes to issues of fuel usage and greenhouse gas production.

Finally, to highlight the point about alternative modeling approaches for analyzing VMT-related policies, one widely used source of VMT forecasts is the U.S. Energy Information Administration's (EIA) National Energy Modeling System model.

Kenneth A. Small (2012) summarizes the details of this model, which includes many features typical of models that focus primarily on long-term aggregate forecasting of vehicle choice and usage with much less emphasis on infrastructure.

---

21 In fact, even VMT measurement is considered difficult.

> …consumers as a group make several choices, modeled as aggregate demand functions (EIA, 2008, pp. 51–77). First, they choose the shares of cars and light trucks according to a logit-like formula that predicts the change in market share from the previous year as a function of changes in variables including income, fuel price, and new-vehicle fuel efficiency. Second, they choose among the six size classes available for each of the cars and light trucks according to an aggregate model, again predicting change in market share from the previous year as a function of changes in several variables such as fuel price, vehicle price, and income (EIA, 2008, pp. 10 and 41). Third, consumers choose market shares of various fuel types through a three-level aggregate nested logit model whose variables describe vehicle price, fuel cost, range, acceleration, and other factors (EIA, 2008, pp. 51–61). EIA has calibrated the coefficients of these aggregate choice models to match known market shares in recent years, and has added some projected variation in them over time representing judgments about the likely evolution of tastes and marketing practices.
>
> Finally, the stock of LDVs on the road is determined by combining new-vehicle sales, as described above, with exogenous vehicle survival rates (EIA, 2008, pp. 78–84). Total VMT are modeled as a consumer choice determined by a lagged adjustment process following a log-linear regression with two variables: income and fuel cost per mile (EIA, 2008, pp. 84–85). These VMT are apportioned exogenously by vintage, a key part of determining total energy consumption.

The consumer choice referred to in the penultimate sentence relates to VMT per licensed driver, so that final overall VMT demand is determined by using population projections.

This approach is consistent with Bates' emphasis on the importance of population growth and income levels, as well as the generalized cost of traveling, in determining total VMT demand. At the same time, it addresses the need for allocating VMT across vehicle types and vintages for correctly determining energy consumption. This topic is addressed in more detail in section titled "Personal Vehicle Choice Model" in Chapter 4.

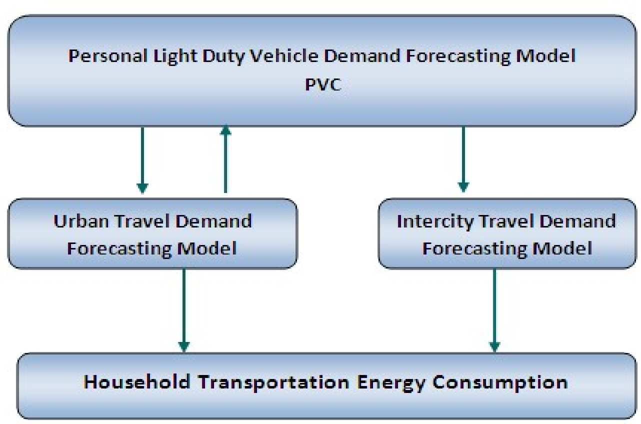
CHAPTER 4: Review of DynaSim Personal Travel Behavior Components

The current model structure of DynaSim was shown in **Figure 1**. The following is a list of DynaSim modeling modules sorted by history (specific modules addressed by this project are highlighted in bold)—see Stanfield Systems (2012):

- **Urban Travel Model**
- **Intercity Travel Model**
- Urban Transit Finance Model
- Other Bus Model
- **PVC Model**
- Commercial Vehicle Choice and Travel Model
- Freight Model
- **Congestion Model**
- Aviation Model
- Petroleum Reduction and Benefits Model

In what follows, the terms Urban Travel Model, Intercity Travel Model and PVC Model will generally be used. The scope of this project is limited to personal travel specifically addressed in a report by Aniss Bahreinian (2013c) that includes the following figure that shows the relationship between the highlighted models above.

Figure 6: Household Vehicle and Travel Demand Forecasting Models



Source: California Energy Commission

One recommendation of this report is that future improved modeling approaches and processes at the Energy Commission should draw on the existence of both the CSTDM and the statewide transportation planning process cited in Chapter 3. Two main reasons for this are:

1. The CSTDM is based on a theoretically superior modeling approach that is capable of producing better long-term VMT forecasts. Many of the features of the CSTDM directly address long-standing criticisms of similar models used for producing forecasts (all of which apply to DynaSim).
2. The CSTDM is being used as part of the statewide transportation planning process. Substantial resources and effort are being expended to produce the large input data sets required for key horizon years, including highly detailed specifications of future planned transportation infrastructure in California. Taken as a whole, all efforts expended for these horizon years represent a potential benefit to modeling efforts at the Energy Commission.

These two points should be kept in mind when considering various aspects of the DynaSim review. As noted, the Urban and Intercity Travel Models of DynaSim generally correspond to the STPDM an LTPDM, respectively, of the CSTDM. In this regard, a primary issue relates how forecasts of total VMT for personal travel are produced. Both DynaSim and CSTDM produce VMT forecasts as an outcome of mode choice decisions.

However, there is ultimately a requirement to produce fuel forecasts in a world where there is substantial diversity in vehicle characteristics (particularly fuel efficiency) and, as discussed in Chapter 3, both approaches rely on very strong assumptions of independence between vehicle choices and travel choices to perform these calculations. In this regard, the situation can be characterized as a lack of integration between vehicle choice and travel choice.

In what follows, a detailed discussion is reserved for the Urban Travel Model. The primary focus will be on issues related to quality of total VMT forecasting. This discussion is also applicable to the Intercity Travel Model. Alternative modeling approaches and their tradeoffs will be explored. The other major discussion involves vehicle choice and how to integrate these choices more effectively with other travel choices.

Urban Travel Model

The Urban Travel Model deserves special scrutiny for the following reason. It produces what is arguably the most important behavioral measure in DynaSim: VMT forecasts for using personal vehicles in California for short-distance trips.

As noted in the introduction, until recently the VMT model of VEDUM/CalCars was used to produce the total VMT forecast. Recently, this was discontinued, and a decision was made to use combined VMT projections from the Urban and Intercity Travel Models.

Although it is unknown how this actually occurred, some of the factors are readily understood. These are stand-alone models that produce redundant versions of the same required outcome measure but use different approaches.[22] Without any obvious way to combine them, the question becomes, “Which set of results should we use?” In addition, the idea of first producing a total VMT forecast and then allocating VMT to vehicles has been discussed multiple times.

The Behavioral Modeling Approach Used by the Urban Travel Model

A high-level verbal description of the model employs concepts discussed in Chapter 2. Travel behavior is defined and modeled primarily in terms of passenger trips, with conversion to vehicle miles required at some stage. Here is some basic background:

- Individual (passenger) trip behavior involves consideration of both *trip links* and *round trips* (aka “tours”). The term *trip* is interpreted to mean “round trip,” and the term *link* appears when necessary to avoid confusion.
- While avoiding specifics, at some level of aggregation the following relationship might be used:

  $$UrbanTrips = UrbanTripLinks \,/\, LinksPerUrbanRoundTrip$$

- Trip generation behavior is modeled in terms of *trip links*; however,
- The other key behavior being modeled is *mode choice* for *(round) trips*. Urban travel modes in DynaSim include both auto and transit modes (but not walking or biking) that can vary by region.[23]
- The model is defined using three “dimensions”: *mode*, *period*, and *time*. *Period* refers to “time of day”, and *time* is the time dimension along which forecasting occurs (in this case the scale is “years”, with the initial time being the “base year”).[24]

Chapter 2 identifies an important aspect of travel behavior modeling that is notably absent from the formal specification of the Urban Travel Model: a *spatial* or *geographical* dimension. This must somehow be addressed as part of using the model for analysis, but the dimension does not formally exist in the DynaSim documentation.

---

22 The fact that these models are “stand alone” and not part of a single, integrated modeling system (that would produce one consistent outcome measure) was one of the main motivations for this project.

23 One key difference between DynaSim and CSTDM is the definition of auto mode. DynaSim has a single auto mode, whereas CSTDM includes multiple auto-related modes, based on occupancy level of the vehicle: Single, 2, or 3+ household members. In addition, CSTDM models walking and biking as modes.

24 This dimension is included for reasons already discussed in Chapter 2: characteristics of competing modes that affect choice can differ by time of day in important ways. Therefore, the time of day dimension is apparently not used.

The authors' best interpretation is that the absence of a spatial dimension may be related to the design philosophy underlying DynaSim. A design philosophy reflects a viewpoint on how to define models, how they function, and what their relationship to data should be. The goal in DynaSim is clearly to create a flexible and generic modeling platform. As implied above, it assumes that models can be defined on the basis of fundamental dimensions, which is theoretically sensible. However, there is an implicit view that issues of scale, geography, spatial relationships, and aggregation are largely matters that can be implicitly addressed on the basis of details contained in data sets that are used as the modeling inputs. At some level this view may be valid, but in its current form, this creates structural problems that were identified in Aniss Bahreinian (2013b).

For example, the Urban Travel Model is viewed as an entity that can be run at any arbitrary level of geographical detail (and whose definition can flexibly change). The current implementation is linked to data sets at the following levels of aggregation: transit agency, county, region, or the entire state. (See Aniss Bahreinian [2013b]). For this discussion, it is understood that the model is being used for an arbitrarily defined geographical region.

Continuing with the description of the model, it follows these basic steps:

1. Input data for the base year are totals of the following:

    a. *UrbanTripLinks(m,b)*

    b. *UrbanTripLinkMiles(m,b)*

    c. *UrbanPassengerMiles(m,b)*

    where m = mode, b denotes base year, and in this treatment, p = period is dropped. In addition, base year data are used to establish the following constants: *LinksPerUrbanRoundTrip(m,b)*, and *UrbanPassengersPerVehicle(m,b)*.

2. Compute *UrbanTravelModeShares(m,b)*, that is, the share of each mode in the base year as measured by (round) trips.

3. For each time t, compute an (initial) trip forecast denoted as *UrbanGeneratedTrips(m,t)*. Although there are additional complicating details, for clarity of presentation, the primary assumption is that future trips can be obtained by a simple extrapolation based on population projections:

$$GeneratedUrbanTrips(m,t) = \frac{UrbanTripLinks(m,b)}{LinksPerUrbanRoundTrip(m,b)} \square \frac{Population_t}{Population_b}$$

    where $Population_t$ is a forecast of population for time t, and $Population_b$ is the base year population.

4. Compute the final forecast, *UrbanTrips(m,t),* by adjusting *GeneratedUrbanTrips(m,t)* to account for any changes in mode share. The generated trips reflect what would happen if there were no change from mode shares computed for the base year. However, if mode shares were to change, the number of trips for each mode would be adjusted accordingly:

$$UrbanTrips(m,t) = GeneratedUrbanTrips(m,t)\frac{UrbanTravelModeShare(m,t)}{UrbanTravelModeShare(m,b)}$$

   Changes in mode share are computed using a pivot-point calculation relative to the base year shares. For this discussion, the details of the equations are not required. The idea is that mode shares would be the same, unless there are *changes* in any of the following factors related to the *generalized cost* of using each mode (versus base year levels):

   a. *InVehicleTime*
   b. *OutOfVehicleTime*
   c. *OutOfPocketCost*
   d. *PersonalIncome*

   The first three of these are attributes for the competing modes, and the fourth is used to compute an adjustment factor to reflect changes in sensitivity to costs associated with changes in income levels.

   Remarks: Changes in the above variables could come about due to a variety of effects. *InVehicleTime* could be reduced for autos by assuming a change in the average speed traveled by vehicles on the road network. One potential source of change would be additional congestion on the road network due to population growth. For transit, this could be improved by increasing the frequency of transit services. *OutOfVehicleTime* relates specifically to transit: it is the amount of time spent waiting for a vehicle to arrive, which could also be affected by changes in transit vehicle stock. (This is assumed to be zero for the auto mode.) *OutOfPocketCost* for autos could be affected by changes in fuel prices and/or average vehicle efficiencies. For transit, this could change due to changes in fare levels.

Based on this description, it is important to emphasize that *this model relies on extremely strong assumptions that specifically exclude the most important factors that were identified in Chapter 2.*

For example, although data on "trips" are compiled for this model at the level of county, transit district, and so forth, and these are referred to in some discussions as "zones," *trip links are not defined by zone-to-zone movements in this model* in the manner discussed in Chapter 2. If anything, trips are defined as behaviors that are being "counted up" and assigned exclusively to an origin zone. There is, quite literally, no spatial aspect to trip definition in terms of trip link end points (destinations) in this model. As noted, trips are defined using counts associated with what would otherwise be known as "origin zones"; moreover, these "zones" really are more akin to "regions" in terms of their level of detail.

One of the most important aspects of mode choice modeling is a proper definition of the choice situation faced by the decision maker. As discussed in Chapter 2, the preferred level of detail includes an origin, a destination, a specification of the set of available mode options, and generalized costs associated with all options. This level of detail is required because it is critical to accurately model trip choices before aggregation. At the level of a region, households will be distributed across very different choice situations with regard to mode accessibility and other factors. Forecasting changes in mode shares based on high level averages for trips over an entire region would be unlikely to capture the proper behavioral response. These issues arise even more directly in the Intercity Travel Model. In this case, structures were added to DynaSim to allow defining trip links in terms of both origins and destinations.

The Current Urban Travel Model in Relation to the California Statewide Travel Demand Model

Before pursuing spatial issues any further, as a first step, the authors review in more detail some features of the current model. The approach will be to use the CSTDM (as incorporated into the statewide transportation planning process) as a point of comparison, from the following perspective:

1. Annual VMT for personal vehicles is a critical output needed by the Energy Commission
2. There are ways in which the CSTDM could be productively exploited even if no changes were made to the current model.

The most basic idea to explore first is that the CSTDM could be used as a data source for DynaSim at specific milestone years.

The first obvious requirement is that base-year data would need to be consistent for the two models. A related requirement would be the development of set of common assumptions for defining a reference scenario.

On the first point, both models draw very heavily on the CHTS for key baseline values related to trips, VMT, and the like. They have a slightly different definition for short- versus long-distance trips, but this should be a minor issue. Geographical location definitions play a key role in some of these processes. However, the CHTS has multiple geographic location variables (for example, "county FID code 2010," "census tract ID 2010"), and all the model-related documents the authors have reviewed seem to use very similar definitions for geographical hierarchies. (This is true for DynaSim, CSTDM, and EMFAC.) So, this is not a major issue.

With regard to reference scenario inputs, some of the key factors are population, employment, and factors relating to fuel costs of operating vehicles.

It is clear from the review of EMFAC, California Transportation Commission (2010), and other sources that developing common reference scenarios for population and employment projections are possible. The number of sources and options for producing these assumptions are relatively limited, and there is already substantial overlap in current practice. With some increased coordination, reaching agreement on a consistent set of assumptions for a reference scenario seems possible.

Issues related to fuel price forecasts could be slightly more complicated but solvable. The Energy Commission is the official entity in California for producing fuel price forecasts. Some modelers rely on forecasts from other sources, for example, EIA. However, the EIA and Energy Commission forecasts are closely related, and reaching an agreement is at least possible, as ARB adopts Energy Commission prices for its analyses. Some additional coordination could be required because for choice modeling, the estimates of, for example, auto operating costs frequently include additional figures for maintenance costs, registration fees, insurance, and the like. Another potential complication is that it is unclear on what constitutes a reference scenario for the Energy Commission when it comes to future alternative fuel vehicles, and how this relates to scenario development for the CSTDM. As noted, the only personal vehicle attribute used as an input to the CSTDM is the average operating cost of a vehicle. It may be that this value is relatively insensitive to these assumptions.

A major motivation for a common reference scenario is the amount of time and effort being expended on building detailed data sets that represent planned future transportation infrastructure. These are factors that would be nearly impossible to replicate in DynaSim. An important aspect of this is that, at each horizon year, CSTDM performs a network equilibrium that creates information on travel times, trip distances, and the like, that could represent a vast improvement over what could be obtained from DynaSim.

In terms of progress on these fronts, Cambridge Systematics (2014) reported on the status of CSTDM as of June 17, 2014.

- All models have been fully completed.
    - » Base year validation and sensitivity testing are complete.
- Future year forecasts are nearly complete.
    - » Results and growth statistics will be available within two weeks.
    - » Horizon years:
        - 2015, 2020 (complete)
        - 2035, 2040, 2050 (in progress)
- Documentation is in progress.

  - » Expected to be fully completed within one month.

With regard to the base year, they report on a validation exercise to compare network traffic counts (for so-called "screen lines") produced by the CSTDM versus actual traffic data for the year 2010, and results appear reasonable.

In what follows, the working assumption that CSTDM runs for these horizon years have been obtained for an agreed-upon reference scenario. Under that assumption, any output from these runs can be treated as data sets analogous to simulated CHTS survey data at specified horizon years.

Deficiencies of the Current Urban Travel Model

With this as background, the authors turn to an evaluation of the Urban Travel Model as a source of VMT estimates (for short-distance travel). The review reveals that the current Urban Travel Model has these deficiencies:

1. Trip generation (and therefore total VMT) is affected only by projections of future population. In other words, there are no income effects and no effects due to vehicle operating costs.
2. The only way in which household income and vehicle operating costs are taken into account by the Urban Travel Model is through their impact on mode shifts between auto and transit alternatives.

These should be major concerns for the Energy Commission, given their responsibilities under the *IEPR*. Aniss Bahreinian (2014) provides a helpful review of the type of scenario analysis requirements the Energy Commission addressed for the *2013 IEPR*. The primary factors are:

- Population.
- Income.
- Energy prices for liquid fuels.
- Energy prices for electricity/natural gas/hydrogen.

The approach is to establish a reference scenario (R) for all these factors and then define alternative scenarios through various combinations of low (L) and high (H) levels on these factors.

The fact that transit modes in the Urban Travel Model are appropriately responsive to only one of these four factors (population) is a major deficiency that must be addressed. There are similar issues with the Intercity Travel Model.

A Potential Role for the California Statewide Travel Demand Model

As a point of comparison, the CSTDM takes into account the effect of population, income, and projected mode operating costs on both total trip generation and mode shares.

This is because the behavioral models in the CSTDM specifically model a critical factor (car ownership level) as a function of both household demographic variables (which include income), as well as destination accessibility measures, which are affected by operating costs. Moreover, accessibility measures include feedback loops of the type discussed in Chapter 2, whereby the overall generalized costs of the various travel modes are represented at higher levels of choice (including car ownership levels). For a more detailed discussion of latent and induced demand in the SDPTM of CSTDM, see HBA Specto (2013).

An additional feature is that the CSTDM also includes an equilibration of the transportation network to produce appropriate travel times.

As has been noted, the CSTDM taken as a whole could be regarded as a substitute for all of the DynaSim modeling modules with the exception of the PVC model. The simplest incremental option that is worth considering is to use the CSTDM to produce mode-level VMT forecasts to replace forecasts from DynaSim. This would be done by (1) producing forecasts for specific milestone years, and then (2) interpolating for intermediate years. Although an initial reaction to this idea might be negative, this is clearly a superior approach to what is being done now.

As mentioned, all the output from any particular CSTDM run would be available to be treated as a data set. This opens a variety of other options for how CSTDM could be used. However, before considering these, further exploration of practical issues is warranted because of concerns about the long run times and large data requirements for the CSTDM.

The stated working assumption was that an agreed-upon reference scenario had been established. This means that CSTDM results for the horizon years 2015, 2020, 2035, 2040, 2050 would be automatically available and no additional computations would be required.

However, suppose that there were some disagreements and inconsistencies about the reference scenario. This would mean that the CSTDM might need to be run again. It is argued that, with regard to the reference case, (1) any such disagreements would not be major, (2) differences in results would be small, and (3) rerunning the model with slight changes in these assumptions should be much faster than starting from scratch. Moreover, the initial step would involve only a couple of runs just to evaluate the situation. The bigger question would involve the alternative scenarios required by *IEPR* work.

Based on Bahreinian (2014), reference values for population and income are used in only one scenario. Because the CSTDM requires the use of its population synthesizer, one specific question would be how to produce two additional population/income-related projections (denoted L-L, and H-H in Bahreinian [2014]). Once these are available, there would be anywhere from four to seven additional scenarios required for varying levels of the key factors.[25] Producing a full set of CSTDM results for the horizon years could therefore require anywhere from 20 to 35 additional runs.

Based on current discussions, this could be time-consuming. Focusing again on the case where the main purpose is to produce total VMT estimates, a logical starting point would be to run the CSTDM using reference values for fuel-related factors at the two additional levels (L-L, H-H) population/income scenarios and to do so in a sensible order of horizon years. These would presumably have the largest effect on VMT demand. After that, selected runs could be performed that relate to the effect of overall price levels on total VMT demand.

Based on this discussion, here is a summary of potential key questions:

1. What would be required to reach an agreement on a common reference scenario that could be adopted by all agencies? This would presumably involve Caltrans and ARB, but there is also a policy advisory committee for the statewide transportation planning process that could perhaps address this.
2. Is there any shared interest on the part of the other agencies in running alternative cases that correspond to the *IEPR* cases?
3. What is involved in running only the population synthesizer of CSTDM?

Possible (Minor) Modifications to Current Urban Travel Model

The next question is, working under the assumption that the CSTDM could generally be used to produce specific sets of results at horizon years to flesh out major population/income/fuel scenarios, what role could be played by the current version of the Urban model?

To address this carefully, the first step is to review additional details about the information available from the CSTDM for the horizon years. In addition to total VMT projections, the CSTDM would be able to provide summary statistics analogous to base year statistics obtained from the CHTS, but for all the horizon years. (See the list above.) In particular, this would include mode shares for all horizon years based on a detailed characterization of future infrastructure, including equilibrated travel times.

25 One issue is that high-speed rail scenarios would play a role, and the CSTDM is an integrated model that includes all travel in California.

Perhaps the most important aspect of the current version of the Urban Travel Model would be its functionality for computing other statistics required by the Energy Commission for non-auto modes. One area the authors are less familiar with what the potential contribution of CSTDM would be with regard to infrastructure data. For example, it probably does not maintain the type of data on transit stock, fuel types, efficiencies, etc., that the Energy Commission needs for computing fuel usage.

The bottom line is, with regard to VMT demand estimation, the capabilities of the current version of the Urban Travel Model are so inherently limited that it is unclear what the value of the model might be, versus judicious use of interpolated scenario statistics that can be obtained from the CSTDM. With regard to run times based on experience, there would probably be a variety of ways to produce the needed results more quickly, once a full set of horizon year results is available and the stakeholders using the model gain more familiarity. As noted, the current Urban Travel Model would probably perform the necessary energy-related calculations once the demand-related quantities are appropriately provided to the system as inputs. This assessment would generally apply to other DynaSim models as well.

Improved Versions of the Urban and Intercity Travel Models

At the outset and during the course of this project, various stakeholders have expressed hope that some new modeling options might emerge that would represent an improvement over the current version of the models in DynaSim while also being simpler than the CSTDM. This is based on an accurate perception of tradeoffs associated with the CSTDM that have already been discussed. In particular, the authors interpret the adjective *simpler* to include at the very least an avoidance of microsimulation and perhaps the notion that it should be possible to devise acceptable models that rely on less detail.

Having developed sufficient background, it is now possible to explore in more detail additional options and related tradeoffs. Returning to the Urban Travel Model, the intended functionality of the model fits squarely in the domain of travel demand modeling addressed in Chapter 2. At the same time, it has been determined that the current version of the Urban Travel Model lacks even the minimal desirable features identified in Chapter 2.

Having said this, in addressing concerns identified in the first paragraph of this section, it is important to recognize that, even if the Urban Travel Model was improved along the lines of models described in Chapter 2, it would still suffer from the same criticisms as standard four-step models. In addition, the CSTDM represents the type of approach that is required to overcome these criticisms. With this caveat, the remainder of this section is devoted to exploring possible approaches that lie somewhere between the current version of the travel demand models in DynaSim and a version of the CSTDM that has been enhanced with a vehicle choice model.

To review, basic structural features that can create modeling deficiencies are:

1. Treating interrelated choices as independent by severing the feedback loops identified in Figure 2.
2. Not taking into account the interaction of supply and demand in the transportation network that determines performance characteristics, such as travel time.

With regard to the second feature, addressing this requires detailed network specifications of transportation infrastructure, and capabilities similar to those employed by the larger metropolitan planning organizations and the CSTDM using Geographic Information System (GIS)/spatial systems like CUBE, TransCad, and the like.

Of course, developing these capabilities might not be out of the question. As already discussed, the authors recommend that the CSTDM play at least some role in future Energy Commission modeling, and the authors understand that the Energy Commission is developing in-house CSTDM capabilities. The CUBE software would, therefore, already be available. Moreover, the DynaSim system documentation includes the following remark: "The long term goals for DynaSim include the integration of geographic information systems (GIS), land-use, and regional economic evaluation capabilities…"

Still, running a pre-existing model like the CSTDM is a very different approach than developing an entirely new modeling system that includes a full-blown transportation network model. For this reason, options with this level of detail were excluded from consideration. Some obvious improvements of the Urban Travel Model would include:

1. A zone-to-zone structure with an appropriate level of detail.
2. A synthetic population in each zone of representative household types, along with assigned weights designed to provide an appropriate demographic representation of each zone.
3. Indices to capture attractiveness of zones as destinations for multiple trip purposes.
4. Infrastructure performance measures for zone-to-zone mode choices.

Depending on the specific details, this structure includes the possibility of integrating vehicle choice with other household travel-related choices.

As discussed, it would be possible to exploit the CSTDM for creating such a framework by using it as a data source to populate the system at horizon years.[26]

26 The earlier question about what would be involved in using the CSTDM Population Synthesizer in a "stand alone" mode might be relevant.

The primary advantages of this approach are the following:

1. A zone-to-zone structure would allow a much more realistic model of mode choice than the current version.
2. An appropriately defined structure could also be used for long distance/intercity travel choices.
3. Explicit inclusion of demographics allows the model to directly embed a module for car ownership level and vehicle choice.
4. The use of weights and representative households avoids the requirement for pure microsimulation.
5. The combination of spatial and demographic features could allow inclusion of period and trip purpose dimensions.
6. In the absence of vehicle choice, the basic approach described here is one that is reasonably well covered by current state of practice.

The disadvantages of this option include the following:

1. Although this approach includes more detailed zones, it would not include details of transportation networks, supply and demand equilibration, and so forth.
2. As such, the model would have difficulty addressing congestion issues.
3. Implementing such a modeling approach could require a substantial amount of effort, since the entire model is being constructed from scratch. The amount of effort would depend on some of the finer details of zone and household definitions, etc. It is, therefore, difficult anticipate what the resource requirements might be.
4. It is unclear whether the current DynaSim software architecture would be the best platform for this approach (although disagreement from DynaSim software designers on this point would be expected). Even though this proposed structure does not include detailed network specifications, it has additional spatial features that place it squarely in the state-of-practice modeling approaches discussed in Chapter 2. These are part and parcel of various transportation-oriented commercial packages (including the aforementioned CUBE and TransCad systems).

The approach suggested here actually creates the possibility of improvement in two areas. The first is an improvement in total VMT and mode choice forecasting (for both the Urban and Intercity models), and the second is the possibility of integrating vehicle choice.

With regard to the first area, there are at least two suboptions that should be considered. First, it is recommended that the system structure for this model be built using seed data from the reference scenario CSTDM horizon year results. This immediately and automatically ensures an improved set of results for total VMT and mode shares, because they correspond to CSTDM results for the reference scenario. But, what are the implications for analyzing alternative scenarios?

In describing the above framework, it is not indicated how the model would be adjusted for alternative population, income, and fuel price forecasts. However, there would be reasonable ways to do this and, in this regard, the model would be considered an improvement. But it would still be subject to all the criticisms similar to four-step models. Specifically, there would be no feedback loop to correct supply and demand effects on travel times. Again, this would still be a substantial improvement over the current version.

Another alternative to consider would be to reseed the model using CSTDM results, particularly for alternative population and income scenarios. Running the model for alternative fuel price scenarios would be subject to similar criticisms as those above.

Finally, it would be possible to reseed the model for all population, income, and fuel price scenarios. This recommendation would represent a complete replacement of the DynaSim with the exception of the PVC. It would, in fact, remove criticisms related to the network equilibration issue. The major difference is that the approach in this section represents a possible way of creating integration between vehicle choices and travel choices. Only the details are different and they relate to the specific way that CSTDM results are extracted.

This represents a complete discussion of the possible options for the Urban and Intercity Travel Models. The only remaining discussion of modeling options relates to vehicle choice, which is summarized in Chapter 5. The next two sections provide additional details specific to the Intercity and Congestion models, respectively.

Intercity Travel Model

As discussed, the Urban Travel Model generates VMT forecasts for short-distance trips using personal vehicles. The Intercity Travel Model provides the remaining VMT estimates for the long-distance personal trips. In the current version of DynaSim, one-way trips greater than 50 miles are long-distance trips.

The high-level structure of the Intercity Travel Model is similar to that of the Urban Travel Model. The first step is to predict total intercity travel, and then a logit model is used to allocate this travel across modes. Of course, the modes are defined differently than in the Urban Travel Model, since air and long-distance train are not options for urban travel. Unlike the Urban Travel Model, the total number of intercity trips is sensitive to income and trip costs, in a manner consistent with theory: total number of trips is an increasing function of income and a decreasing function of trip cost.[27] Moreover, the mode choice utility specification makes mode choice less sensitive to trip costs as income increases. Presumably the total number of intercity trips is also adjusted for increases in population as in the Urban Travel Model, but this is not clear from the documentation provided.

Another key difference relative to the Urban Travel Model is that the intercity model is run separately for a set of representative actual trips, which are then reweighted to match the forecast total intercity travel. Bahreinian (2013b) lists 10 representative trips that were actually used as input. These trips were between cities using the location of the home district office of California State Senators for the cities, but two of the trips (numbers 22 and 39) show distances of fewer than 50 miles. Even though these trips are set arbitrarily (indeed the model has previously been run using only one intercity trip to represent intercity travel for the entire State), the Intercity Travel Model does not allow for substituting origin or destination as costs change. For example, if fuel costs increase, the Intercity Travel Model would still predict the same relative number of trips for each of the representative trips, but the predicted mode splits would change based on changes in relative costs.

The previous two paragraphs describe how the Intercity Travel Model forecasts passenger miles by mode. In order to convert passenger miles into vehicle miles and fuel usage, exogenous occupancy (or load) factors are used and are fixed over the forecast period. These baseline occupancy factors come from CHTS trips for auto and other administrative data sources for air and rail. The baseline occupancy factors are calculated for the entire state, and this implies that auto occupancy does not vary by cost or distance of trip.

It is also not clear how the Intercity Travel Model deals with external trips to out-of-state destinations like Las Vegas. These trips are captured by the Caltrans screen line counts, then the representative trips used by the Intercity Travel Model are scaled up to include these in the total travel. This is probably not too bad an approximation unless scenarios that lead to big price disparities between California and neighboring states are considered. For example, if the price of gasoline is 50 cents per gallon cheaper in Nevada, then that cheaper price is the relevant price for auto trips to Las Vegas.

It is useful to contrast the Intercity Travel Model with the CSTDM to show how a bottom-up approach differs. The long-distance travel model in the CSTDM considers all

---

27 However, it is not known exactly what type of income projection was being used to drive the increase in overall trips. The original 1983 SYDEC model specifically uses per-capita income to capture income elasticity effects but separately includes a factor for population growth. But, even the SYDEC documentation is unclear on where the population growth factor is introduced. The authors cannot find any specific population factor mentioned in DynaSim documentation. One possibility is that this is somehow combined with income, but this would be inconsistent with the intended model specification.

trips among the approximately 5,000 zones more than 100 miles in length. Each household is assigned 0, 1, or 2 daily long-distance trips for a particular purpose (business, commute, recreation, and other), and each trip is assigned a party size, destination zone, main mode choice (car, air, conventional rail, or high-speed rail), and access/egress mode. Each of the underlying models is linked through accessibility and destination log-sum variables that are sensitive to travel times, costs, and desirability of the destination zones. Therefore, an increase in fuel prices will lead to households taking fewer trips, with shorter distances and higher occupancies for each trip. The differential models for trip purpose allow business travelers to be less sensitive to cost and more sensitive to time than other travelers, and these differential models allow for higher auto occupancy for recreation and other purpose travelers.

The CSTDM long-distance model handles external travel by defining a number of external zones (including Southern Nevada). This approach allows for different travel costs and is clearly more accurate for the key Interstate 5, 10, 80, 15, 10, and 40 routes.

All of the analysis in previous sections of this chapter would also apply to the question of how the CSTDM could be used to improve the Intercity Travel Model. In particular, the CSTDM could be used to provide network time and cost data at specific horizon years that could be used to provide a structure for running the Intercity Travel Model on a much larger set of representative trips. All the comments on the merits and disadvantages of this approach discussed in Chapter 4 are relevant here as well. The CSTDM will also provide forecasts of vehicle occupancy for each origin/destination pair in the network but does not provide occupancy forecasts for air or rail. The current DynaSim approach of using external data from Amtrak and airline surveys on load factors and equipment used will still need to be followed to convert passenger miles into vehicle miles, fuel use, and emissions.

Finally, the CSTDM has a separate mode for high-speed rail but does not consider intercity bus as a possible mode. The CHTS includes intercity bus in its data collection (accounting for 0.4 percent of long-distance trips as compared to 1.1 percent for train), so it would, in principle, be possible to expand the CHTS long-distance model to include intercity bus. However, the lack of the intercity bus mode is a weakness of the CSTDM relative to DynaSim. This implies that intercity bus travel times and costs will need to be collected separately for each origin/destination zone pair used in DynaSim.

Congestion Model

The Congestion Model captures the impact of congestion on auto and freight travel time, and it models the impact of congestion on auto fuel economy. The model is based on work from the Texas Transportation Institute, which is also the source of most of the baseline data for most major cities in California. The key inputs are daily VMT and lane miles on freeways and arterials (congestion on collector and local roads is ignored due to lack of data). The outputs are indices that can be used to adjust the travel time and fuel efficiency inputs to other DynaSim models or to adjust final fuel usage.

Baseline data on freeway and arterial daily VMT come from the Texas Transportation Institute Mobility report, and baseline data on lane miles come from Caltrans Public Road Data. These figures need to be adjusted manually to correspond to counties, which may introduce errors. The congestion model then uses exogenous forecasts of lane miles plus VMT output from the Commercial Vehicle Model, the Urban Vehicle Model, and the Freight Vehicle model to modify the Road Congestion Index (RCI). The updated Travel Time Index (TTI) is then predicted using a simple linear regression from RCI. This updated TTI can then be used to modify travel times in the other models.

The key output of the Congestion Model is a Fuel Efficiency Index that can be used to modify average fuel efficiencies input into other DynaSim modules as well as final fuel usage calculation. This is done by applying a regression relationship between fuel economy and speed developed by Raus[28] in 1981. Given the large changes in engine control systems as well as the penetration of hybrid electric vehicles, it would be very desirable to update the Raus study with more recent data on a representative sample of modern vehicles.

Although the Congestion Model is clearly an aggregate approximation to a very complex problem, it very much complicates running the DynaSim system since it creates feedback loops with other models. These feedback loops should be iterated until outputs stabilize. While it is likely that they will eventually stabilize, the increased computational time to run DynaSim could be substantial.

The microsimulation approach used in the CSTDM is probably the only way to accurately model the impacts of congestion. The CSTDM simulates all personal and commercial road travel and loads this travel onto calibrated road networks to predict congestion endogenously. Since congestion changes network travel times and therefore the accessibility measures that impact all vehicle use decisions, the CSTDM needs to be iterated until these accessibility measures stabilize.

The Energy Commission and ARB's main concern with congestion is the impact on fuel use and emissions. The equilibrated CSTDM networks can give traffic volumes and speeds on each network link, and these can be used as input to the detailed EMFAC calculations of fuel and emissions as discussed in the section "California Statewide Travel Demand Model and EMFAC" in the California Transportation Plan 2040. In this case, the same issues discussed in Chapter 3 with regard to fuel and emissions calculations would apply, which relates to how a vehicle choice model would be integrated to produce these results.

On the more general question of how DynaSim could be improved with regard to its treatment of congestion and infrastructure issues, much of the discussion in the first

---

28 *Raus, J. A Method for Estimating Fuel Consumption and Vehicle Emissions on Urban Arterials and Networks, Report No. FHWA-TS-81-210, April 1981*

section of Chapter 4 applies. For example, it is unclear how road capacities are being forecast for DynaSim. If the base year data obtained from TTI are being used, then accepting CSTDM data on infrastructure scenarios as a reference case might be worth consideration as an improvement. As discussed in the section "Urban Travel Model," it is unclear how, given the current lack of spatial structure in DynaSim, using these indices to correctly capture the impact of travel choice on travel times would be implemented. As noted previously, the CSTDM was specifically designed to address this difficult modeling challenge.

Personal Vehicle Choice Model

This area of modeling has historically been an area of strength for the Energy Commission. Over the years, Energy Commission staff has pursued an agenda of improving its models through collection of new survey data and updating of model estimates. This is important to ensure that model parameters are current and capture any fundamental shifts in preferences or behavior of households. However, they have also collected data (using stated preference approaches) that support the extension of modeling capabilities to address potential new technologies and policy choices for which real-world market data either do not exist or are extremely limited.

The following review of the PVC Model is based on all the various DynaSim-related documentation and historical documentation on CalCars (Chris Kavalec [1996]). Although it would be possible to summarize the technical aspects of this model in some detail, at this stage it should be clear that the primary issue for this project is how the model might be better integrated with other models. Another potential issue relates to the level of detail to use in defining vehicle choices.

The current version of the PVC forecasting model defines vehicles using 15 vehicle classes and 7 fuel technology types (see **Table 3**). The model includes used vehicles as well as new vehicles, and so it has a vintage dimension.

In contrast to the Urban and Intercity Travel Models, the PVC functions more like a bottom-up, disaggregated model (although it does not produce results using pure microsimulation, as does the CSTDM). It takes into account details of household characteristics, preferences for vehicle attributes, and the variation of preferences by region in California. Preferences for alternative fuel vehicles that do not yet exist in appreciable numbers in the market have been estimated using stated choice experiment data. While this approach has detractors, it is one of the few ways to perform policy analysis on this subject. Although some of the behavioral models are based on the notion of transaction choice, the bottom line output yields estimated probabilities for how many vehicles, and which vehicle types, are held for any particular household type, which are then converted to vehicle counts by the appropriate application of weights. The result is a representation of the personal vehicle population by vehicle class, fuel type, and vintage.

A major concern is the complete lack of integration between vehicle choice and vehicle usage. To produce fuel usage estimates, DynaSim assumes that all vehicles are driven the same distance, regardless of vehicle features such as type, size, fuel efficiency, and vintage.

This is clearly a major oversimplification, and first and third sections of Chapter 3 demonstrate approaches that other models have taken to address this issue. The second section of Chapter 3 suggests one possible improvement that could be made with minor changes to the current modeling approach.

However, this section considers modeling options for closer integration between vehicle choices and other travel-related choices. Recall that Chapter 4 laid some of the groundwork for this. But as a first step, a specific option in keeping with our earlier approach of first considering the "theoretically ideal case" was identified. For reference, this is called the fully integrated bottom-up model (FIBM) option.

> It would be possible to estimate a household vehicle holdings choice model using the combined CHTS/CVS data that would be incorporated into the CSTDM.

What this would exactly entail requires additional discussion. Such a bottom-up model would explicitly capture the complete interaction between household vehicle and trip choices, and do so in an environment that incorporates supply and demand equilibrium of the transportation network.

Table 3: PVC Vehicle Classes and Fuel Technology Types

Vehicle Class	Fuel/Technology
Subcompact car	Gasoline
Compact car	E-85
Midsize car	Plug-in Hybrid Electric Vehicle (PHEV)
Large car	Compressed Natural Gas/CNG
Sports car	Diesel
Small cross utility (car)	Hybrid Electric
Small cross utility (unibody SUV)	Battery Electric Vehicle (BEV)
Midsize cross utility	Fuel Cell Electric Vehicle (FCEV)
Compact Sport Utility Vehicle (SUV)	
Midsize SUV	
Large SUV	
Compact van	
Large van	
Compact truck	
Standard truck	

Source: California Energy Commission staff

The CSTDM only models car ownership levels. This is used to determine whether a household has more cars than drivers, the same number of cars as drivers, or more cars than drivers, and this is used as an explanatory variable in mode choice models for individuals making trip choices.

An expanded choice model would model not only the choice of how many vehicles the household would own, but also which vehicles it would own, and how each household vehicle would be used for making trips. In this regard, there are many possible interactions among choices that would occur.

To illustrate, consider the example in the section "An Idealized Bottom-Up View of Personalized Travel Behavior" on page 16. In this case, imagine that various demographic and locational details for household *H* would be established first. Based on factors such as household income, working status, job type, and others, the number of vehicles owned (2) and the vehicle types (2008 compact hybrid, and 2012 midsize SUV) would be simulated. With two full-time workers, one approach would be to assign each vehicle to an adult to be his or her vehicle, which then becomes the default auto mode vehicle for his or her work tour mode choice problem. In this approach, specific vehicle characteristics and other factors would impact the interrelated choices. For example, *M* has a longer distance regular work commute, which would increase the probability that the 2008 compact hybrid would be assigned to him for regular work trips. The 2012 SUV would be assigned to *F*, and the features of this vehicle would be related to the frequency with which it is used for work- and nonwork-related trips.

The FIBM is not necessarily being recommended as the preferred option. It is, however, the best option from a theoretical perspective, all else equal. But, as is typically the case, the best option is rarely the least costly option. In this regard, the FIBM establishes a point of comparison with other options and invites an examination of tradeoffs.

The most obvious cost of FIBM is associated with the bottom-up aspect, where the primary concern is still the substantial run times and data requirements of the CSTDM. Perhaps a more challenging aspect is that it would require direct modification of the CSTDM, and this is not an Energy Commission model.

Perhaps the most important point here is that the details of developing an integrated choice model are largely independent of whether it would be embedded in the CSTDM or would be used in the framework described in the section "Improved Versions of the Urban and Intercity Travel Models" on page 49. In this regard, it would be possible to consider alternative model improvement pathways for moving forward, rather than specific options.

It is perhaps worth revisiting what the fundamental purpose is for creating a more integrated model. Recall from Chapter 3 that the main concern is whether ignoring particular interactions would have an adverse effect on computing fuel usage (and greenhouse gas emissions). Current practice suggests that it might be sufficient to rely on general patterns obtained from a data set to then apply exogenously to vehicle choices.

Suppose this were true. Then the primary concern might be judiciously using the CSTDM to ensure improvements in total VMT forecasts as discussed in the section "Urban Travel Model" on page 40, and then separately consider improved ways to allocate VMT to vehicle choices, where the vehicle choices are produced relatively independently. However, an organized research plan could involve a pathway, whereby testing could occur to evaluate the degree to which interaction effects might matter.

For example, an initial step would be to pursue estimation of a more integrated model specification before committing to how it would specifically be used (or if it would be used) for forecasting. Testing could be performed to compare what the implications would be for performing fuel usage calculations using a simple approach, versus a more complex approach. However, a careful approach would be required. The idea is that, even if such interactions were not important in current and past markets, they could become so in the future. There is likely to be a much stronger interaction between the choice of future vehicle technologies and household trip patterns, availability of refueling infrastructure, and so forth. Additional issues, such as the level of detail required for characterizing vehicles and others, are also matters for future investigation.

CHAPTER 5: Summary and Recommendations

This project sought to develop and evaluate improved modeling options related to household travel, and to make recommendations, based on a review of current Energy Commission models, related models of other state agencies, and recently collected CHTS/CVS data. The approach in this report was to first establish a framework and standards for how this type of modeling should be conducted (Chapters 2 and 3), taking into account the specific statutory requirements of the Energy Commission to produce fuel forecasts under alternative policy planning scenarios. Chapter 4 combines a review of the current DynaSim models with a development of possible approaches for model improvement, taking into account a range of concerns.

A general recommendation is that the CSTDM should play a role in any future pathway for model improvement followed by the Energy Commission.

The only question is what that role would be. The overall situation can be summarized as follows:

1. The CSTDM generally performs all of the same functions as DynaSim, with the exception of vehicle choice modeling. They both draw heavily on the CHTS for base year modeling data, and consistency with regard to geographical definitions, trip definitions, and other factors is easily feasible.
2. The CSTDM is a state-of-the-art bottom-up system that uses a theoretically superior approach to modeling, including supply and demand equilibration of transport networks. But it requires highly detailed inputs and substantial run times. It can therefore produce high-quality results, but only for a relatively small number of scenarios.
3. Substantial effort is being invested in updating the CSTDM to provide statewide modeling results as part of the CTP 2040 for horizon years 2015, 2020, 2035, 2040, 2050. This requires highly detailed input data sets compiled from metropolitan planning organizations and RTPAs, including for planned transportation infrastructure and land use patterns.
4. Through some process, a valid reference case (which relates to inputs) should be established for these horizon years, so that CSTDM results for these years can be treated as correct. CSTDM output can then be used as inputs for other modeling purposes.

All of the above points are assumed true for all model improvement pathways under consideration.

During the analysis, it was recognized that it is possible to evaluate modeling options based on two basic issues:

1. Quality of total VMT forecasts
2. Allocation of VMT to the vehicle fleet to perform fuel usage calculations.

In DynaSim, the Urban and Intercity Travel Models produce total VMT forecasts for personal vehicles, and PVC output is used to allocate VMT for purposes of fuel usage calculations. It was found that on both of these issues the travel demand models in DynaSim have major flaws.

1. Total VMT demand should respond to population, household income, and fuel costs. The Urban Travel Model responds only to population change (via simple linear interpolation). This is troubling, given the requirements of *IEPR* cases.
2. VMT is allocated across vehicles under the assumption that all vehicles are driven the same distance.

The issue of total VMT projection is potentially very important, and the area where the CSTDM could immediately contribute under the previous working assumptions. VMT forecasts for the reference scenario could be obtained by:

1. Using CSTDM numbers for horizon years.
2. Use interpolation (or perhaps a statistical regression) to fill in values for other years.

Then, the only remaining question would be how to improve the VMT allocation. One simple way would be to reintroduce the old VMT model and use it to create an improved VMT allocation.

This represents the starting point for a future pathway of model improvement that could be pursued by developing a logical sequence of activities to evaluate tradeoffs for the two basic issues: total VMT forecasting and VMT allocation.

For example, consider the initial suggestion above. The first task would be to obtain CSTDM numbers for horizon years and devise an interpolation. It would immediately be possible to compare DynaSim results to CSTDM results. (This would presumably be done for specific regional definitions). Although the focus is on total VMT, the other related output involves mode choice results. These could also be compared. It might be possible to consider other tests.

On VMT allocation, the CSTDM would have nothing to contribute. However, the first study that could be performed would be to allocate VMT using the current approach and to allocate VMT using the proposed approach. Another alternative would be to use an exogenous source of information on how VMT declines by age. Finally, one specific study could be performed using the original CHTS data. Fuel usage could be computed using fully available information, and compared to other approaches.

Continuing with the VMT allocation issue, this relates directly to the necessity for a more highly integrated choice model. A research study could be sponsored that focuses specifically on developing an integrated vehicle and travel choice model that targets the issues identified in this project. A priority in producing the model would be evaluating the impact the model would have on fuel usage calculations compared to simplified approaches that assume more independence across these choices.

Either way, at the conclusion of the study, a model would be available. A next step (or perhaps part of the same study) would be to run a test of some type with the model embedded in the CSTDM. Alternatively, if it is determined that a choice model is required and embedding it in the CSTDM is not practical, the main additional requirement would be to use CSTDM output to produce more detailed scenario structures rather than just VMT forecasts.

Returning to the issue of total VMT projection, an initial question would be whether it would be feasible to simply run the CSTDM in some fashion to fill in additional values for alternative scenarios and, as before, use extrapolation to fill in values for other years. Specifically, this would involve considering requirements for alternative population/income assumptions (called L-L and H-H in the Energy Commission's *IEPR* nomenclature). This would involve carefully deciding which runs to perform, and in what order. How this would proceed would depend on what can be learned as events evolve. Depending on the outcome of these efforts, it could be learned if it is practical to simply use CSTDM runs to replace DynaSim VMT demand. If this is not feasible, other alternatives could be pursued. If it were determined that an integrated choice model is required, the only change to this discussion would be to use more of the CSTDM output to produce the necessary inputs for the choice model.

One final issue that is not completely clear concerns other requirements for computing fuel usage associated with non-auto modes. It may be that some version of the current DynaSim models would continue to exist in order to perform these calculations, where the demand-related portions are replaced using the procedures discussed here.

In conclusion, the best recommendation is in the form of a sequential process that leverages the insights developed in this project, rather than in the form of a specific model specification. The discussion in this concluding chapter provides a starting point and a framework for proceeding, but any specific plan should incorporate the more detailed information contained in the earlier chapters.

References and Links

California Environmental Protection Agency – Air Resources Board (2011a), *EMFAC 2011-LDV User's Guide*, released September 19, 2011, updated January 2013. Accessed June 30, 2014. http://www.arb.ca.gov/msei/modeling.htm.

California Environmental Protection Agency – Air Resources Board (2011b), *EMFAC 2011 Technical Documentation*, released September 19, 2011, updated January 2013. Accessed June 30, 2014. http://www.arb.ca.gov/msei/modeling.htm.

California Transportation Commission (2010), *2010 California Regional Transportation Plan Guidelines*. Accessed June 30, 2014. http://www.catc.ca.gov/programs/rtp/2010_RTP_Guidelines.pdf.

Bahreinian, Aniss (2013a), "Transportation Energy Demand Model Overview: Inputs and Methods for the 2013 Transportation Energy Demand Forecast," June 26, 2013 presentation, link: http://www.energy.ca.gov/2013_energypolicy/documents/2013-06-26_workshop/presentations/02_Model_Overview-Aniss_Final_RAS_21Jun2013.pdf.

Bahreinian, Aniss (2013b), "Personal Travel Demand Model Data," document provided by Bahreinian, Aniss (Microsoft Word document, "Personal Travel Demand Model Data V2.doc," created November 25, 2013).

Bahreinian, Aniss (2013c), "Attachment A: Behavioral Vehicle and Travel Equations in DynaSim," document provided by Bahreinian, Aniss (Microsoft Word document, file name "Behavioral Vehicle and Travel Equationsv2.doc," created March 1, 2013).

Bahreinian, Aniss (2014), "Vehicle Choice and Utilization Model Integration: The Motivation," PowerPoint presentation (ref. filename "Model Integration Whys.ppt", California Energy Commission, June 30, 2014.

Bates, John, (2008), "History of Demand Modeling," Chapter 2 in *Handbook of Transport Modelling*, 2nd edition, David A. Hensher and Kenneth J. Button ed., Elsevier, Amsterdam.

Cambridge Systematics (2014), "California Statewide Travel Demand Model", PowerPoint presentation to California Statewide Travel Demand Model Update, June 17, 2014. Accessed June 30, 2014. https://www.google.com/search?client=safari&rls=en&q=www.dot.ca.gov/hq/tpp/offices/osp/ctp2040/ctp2040_pac/jun_17_2014/CSTDM_Model.pptx&ie=UTF-8&oe=UTF-8

HBA Specto (2013), "Induced and Latent Demand in the SDPTM," technical memo, October 19, 2013. Accessed June 30, 2014.

http://www.dot.ca.gov/hq/tpp/offices/osp/ctp2040/ctp2040_tac/nov_13_2013_tac_mtg/Tech_Memo-SDPTM_induced_demand.pdf

Kavalec, Chris (1996), *CALCARS: The California Conventional and Alternative Fuel Response Simulator,* California Energy Commission document, April 1996.

McAllister, Andrew (2013a), "2013 Integrated Energy Policy Report Scoping Order," March 7, 2013, (http://www.energy.ca.gov/2013_energypolicy/documents/2013-03-07_scoping_order_2013_IEPR.pdf).

McAllister, Andrew (2013b), *2013 Integrated Energy Policy Report*, Draft Lead Commissioner Report, California Energy Commission, October 2013, CEC-100-2013-001-LCD).

McNally, Michael G., Michael G. (2008a), "The Four-Step Model," Chapter 3 in *Handbook of Transport Modelling*, 2nd edition, David A. Hensher and Kenneth J. Button ed., Elsevier, Amsterdam.

McNally, Michael G., Michael G. (2008b), "The Activity-Based Approach," Chapter 4 in *Handbook of Transport Modelling*, 2nd edition, David A. Hensher and Kenneth J. Button ed., Elsevier, Amsterdam.

Ortuzar, Juan de Dios and Luis G. Willumsen (2001), *Modelling Transport*, 3rd edition, Wiley, Chichester.

Small, Kenneth A. (2012), "Energy Policies for Passenger Motor Vehicles," *Transportation Research Part A*, 46, 874-889.

Stanfield Systems (2012) "Dynamic Simulation Transportation Energy Model (DynaSim) SOFTWARE DESIGN DOCUMENT, Version 5.03, April 17, 2012," provided by Bahreinian, Aniss (two versions of Microsoft Word documents: "DynaSim SDD V5.03.docx", created March 1, 2013; and "DynaSim SDD V5.03_2.docx", created March 22, 2013).

Train, Kenneth (1986), *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*, MIT Press, Cambridge.

UK Department for Transport (2014), "TAG Unit M1: Principles of Modelling and Forecasting," *Transport Analysis Guidance* (TAG), January 2014. Accessed June 30, 2014. https://www.gov.uk/government/publications/webtag-tag-unit-m1-1-principles-of-modelling-and-forecasting.

ULTRANS and HBA Specto (2011), "CSTDM-09 Statewide Travel Demand Model: Final System Documentation Technical Notes," (18 volumes). Accessed June 30, 2014. http://ultrans.its.ucdavis.edu/doc/cstdm-documentation