New Expectations for Transportation Data

MARTIN WACHS
Institute of Transportation Studies
University of California at Berkeley

Winston Churchill once wrote an insightful analysis of architecture. He stated the "we shape our buildings and then our buildings shape us." He showed how we first designed buildings to accommodate our behavior and our social and cultural patterns as we understood them. But we understood them imperfectly and different architects interpreted them differently. Besides that buildings also reflected limitations presented by their sites and by the budgets the designers faced and the building materials they used, and eventually the buildings that we constructed shaped our behavior and became the determinants of our new social and cultural patterns. Some truly exciting and wonderful buildings emerge from this process; but they eventually are seen as outmoded and inadequate in at least some ways, even if they are elegant and satisfying in other ways. We are always shooting at a moving target as we design new buildings because new designs themselves create new design paradigms.

This is nearly a perfect metaphor for the relationship we've built in transportation between databases, analysis, and policy making. We formulate statements of a problem that is vexing us, and we design strategies of data collection and analysis to try to address that problem. But, the data we choose to examine are limited by questions we've previously asked, which in turn reflect the past power of our statistical tools and mathematical models. We're also limited by the costs of data collection and analysis, by concerns regarding privacy and so forth. We address the questions as best we can with the data at hand, and realizing the inadequacies in our databases we reformulate the data collection approaches. Our understanding of the phenomena we study are shaped by the data we have and the models we use and are therefore far from perfect; and information bases and analytical approaches are deeply flawed in part because they're derived from our partial and inadequate understanding of the phenomena we study. We make major breakthroughs and really do achieve dramatic advances in our understanding; and at the same time our knowledge rather quickly becomes inadequate or obsolete.

An example, perhaps, can illustrate the point. For decades, in many of our travel surveys, we defined "trips" in our databases as movements from zone to zone that involved vehicles. This framework greatly hampered our ability to analyze intra-zonal travel and travel by non-motorized modes, like walking and cycling. Yet, when some activists started to assert that pedestrian travel and bicycling were important modes in need of careful analysis and worthy of analysis in urban areas, some of us responded that these modes were not important because our data didn't show that they accounted for a very significant proportion of all trips. It was extremely difficult to squeeze useful data about walking and cycling out of our databases until we slowly added questions and changed the format of home interview studies about travel. In the process, over time, we both started to learn of a wider range of policy choices involving non-motorized travel,

and reformulated the ways in which we record information about walking and cycling.

I love to study the history of what most of us do now. I always include in my courses something about how we collected data and analyzed travel in the early days because it's much easier to appreciate the subtleties of the evolution of data analysis and policy analysis through historical examples than through immersion in current problems. People worried about urban traffic congestion for centuries, and about a 100 years ago we started to count traffic on streets in large metropolitan areas and to portray traffic patterns on maps using lines whose widths were proportional to the traffic flows. In the days before traffic counters how do you think we got the data to make these early maps? Somebody had to stand at the intersections and count the vehicles passing each point per unit of time and the way this was most commonly done in American cities as late as 1930 was by using Boy Scouts. They were trained to count and collect and record the data. They were pressed into service in late afternoons and on weekends when they didn't have to be in school. And they produced reams of data of unspecified accuracy and quality that were widely used for decades in transportation planning.

Our first understanding of traffic congestion was very much shaped by this data, however limited it was. Our professional forebears—arguing that certain streets were more impacted by traffic flow and needed to be widened, double-decked or have underpasses built at congested intersections—began to understand that streets of different width, slope and spacing between intersections could be characterized as having different capacities. Different strategies, including signs and signals, could be used to adjust and manage that capacity.

As early as 1912, a simple gravity model was used to relate traffic flows to levels of economic and social activity in different portions of a city or county, and that led to the collection of information on population and employment and retailing as part of the transportation planning effort. Much earlier than most of us would imagine—well before 1920—others in this profession came to the conclusion that looking at traffic-flow patterns on maps was incomplete and was leading them to false conclusions. These maps showed where traffic flowed, but where it flowed could be understood to be the results of two different sets of causes. One set of causes was where people wanted to go from and to, and the other was where the roads actually forced them to go by their patterns on the ground. Did heavy traffic between A and B imply that people wanted to go from A to B? Or did it imply that road capacity existed between A and B because those roads funneled traffic heading from C to D along the road between A and B where traffic was actually observed simply because the drivers had no option but to go that way? So, they invented the notion of "desire lines" as separate from traffic flow diagrams. Instead of inferring where we needed new capacity by simply observing that the streets were crowded or that volume exceeded capacity, we could ask from where to where did large numbers of people want to travel. We realized that to serve them better we might conceive of direct, diagonal transit routes or highways or high-capacity elevated or depressed facilities that could be overlaid over existing traffic flow patterns on city streets or county roads.

Origin-destination studies grew from these insights. Before World War II we had started gathering information on origins and destinations in two ways. First, cordon lines were set up and drivers intercepted on trips and asked for their origins and destinations. Later, home interviews were used and travel diaries introduced that allowed planners and

analysts to focus on trip interchanges between origins and destinations between large numbers of pairs of zones instead of only focusing on traffic flow on the networks. By the end of World War II, we were applying some of the first awkward computers to analyze this data in what were the earliest applications of computing to analysis of the performance of civil systems.

Of course this circular process of redefining our policy problems based on current data and redefining our data needs based on current policy problems continues and is ongoing. It is the way our thinking and understanding advances over time, even if it is difficult to recognize that when we look at the problems of the moment and the inadequacies of our data sources.

I would like to comment on what I see to be five themes or trends that I believe will be the dominating concerns of transportation analysts and data managers over the first decade of the new century. They are, in keeping with what I have been saying, themes that at once are suggested by our existing data sources yet inadequately addressed by current data sources. They suggest ways in which our understanding of travel is changing, and ways in which transportation data collection can and should be changing over time.

My five themes are

1. Transportation in the new century will be to a far greater degree integrated with telecommunications and the flow of information; our use of data and our need for data will have to quickly change to reflect this.
2. In part because of the new relationship between telecommunications and travel, we will see new patterns emerging in the relationship between transportation and urban form and these will suggest new forms of data collection and new forms of analysis.
3. One of the major transportation problems facing policymakers in the early part of the new millennium will be goods movement. We simply must improve our understanding of patterns of goods movement and integrate it better with our analysis of people movement.
4. One of the most pressing policy issues facing transportation analysts, and one for which our data collection systems and analysis capabilities are ill prepared, is in the realm of sustainability. We will have to look more closely at transportation and global warming, and build data and analytical capabilities that will enable us to do plausible evaluations of alternative transportation policies on the basis of sustainability.
5. Equity will be one of the major themes in transportation policy for the coming decade, and we need to sharpen our tools of analysis and create data support systems for more penetrating analyses of equity in transportation policymaking.

I'd like to try to say a few words about each of these five themes in the hope that they might be relevant to the remaining sessions of this Conference and to your continuing work over the coming few years.

TELECOMMUNICATIONS AND TRAVEL

It is certainly clear to everyone here that the telecommunications revolution is affecting every dimension of our life, including travel. Several authors have recently become

sufficiently bold as to conclude that the telecommunications revolution we are now experiencing will have lasting consequences as dramatic as the industrial revolution that in an earlier century brought us factory production and widespread urbanization. The flow of information between computers, the existence of fax machines and pagers and the rapidly approaching integration of computers with television must be reflected in the way in which we think about and collect data about transportation. Let me mention only two ways in which this is critical.

First, of course, we are experiencing the emergence of what we call "intelligent transportation systems," the application of these new capabilities within the transportation system. Electronic toll collection, smart cards and integrated transit fare collection systems, global positioning systems, information on road conditions and on transit schedules in advance of our trips, and real-time automated navigation aids are all a reality, and the automated highway is only a few decades away. I have not been able to fully work out the ways in which these many capabilities can give us new forms and types of data to describe travel. Or the ways in which they can give us new policy problems that will change the ways in which we collect data or ask new questions of travel data. But I know that they will certainly do that, and that we have to be thinking more actively than we have been about this phenomenon. Won't these capabilities be able to influence and affect travel patterns in ways we are just starting to understand, and should we not reflect that in the structure of our data collection and storage methods? Won't these new tools throw off information as by products that we should incorporate into our routine methods for analyzing travel? I'm certain that they will.

A second way in which telecommunications and travel interact is that we are clearly changing our travel patterns in response to telecommunications. Certainly, we already know that telecommunications will probably increase rather than decrease travel, but our new capabilities are changing the spatial and temporal distributions of travel. The traditional morning and evening peak hours for travel are becoming less peaked but extending over longer time periods because we now have greater capacity to work at different and times and places. Service people get their assignments for the day online instead of by driving to a central dispatching point; information workers can work at home part of the time, leaving when they need to get to a face-to-face meeting. Internet purchases have enormous consequences for the temporal and spatial distribution of shopping travel and goods movement involving parcel services. Until now we have modeled travel on the basis of data on the spatial location of residences and employment because we understood those to be the principal determinants of travel? In our emerging world, should we not also be gathering information on the spatial and temporal patterns of information flows, since they are probably eclipsing land use patterns as the principal determinants of travel? Can we know when and where people will travel if we don't track when and where they communicate by wire and wireless flows of information? Could we better understand travel if we included information flows as independent variables in our travel forecasting analyses? I think so.

TRANSPORTATION AND URBAN FORM

We used to understand that the demand for travel was derived from urban form, and that

investments in transportation capacity were the principal determinants of urban form. The centralization of the city and it's enormously high densities toward the end of the 19th century, and the subsequent decentralization of urban form in the 20th century following first transit and later highway development were the result of the changing relationship between transportation and land use. The power to control land use has been jealously guarded by local governments, and we have mostly tweaked the transportation-land use relationship by investing in transportation. It has only gradually and lately become popular wisdom that we ought to and can also control traffic by deliberately manipulating land use. Urban limit lines, neotraditional development, transit villages, and smart growth are all themes that are in good currency in transportation circles these days. And places like Portland (Oregon) and San Jose (California) are taking action in accord with these principles.

Yet, if I am right, we may be too late to do too much good. It is becoming necessary to think of transportation land use and telecommunications in a three-way relationship. Our modern and emerging telecommunications capabilities will change the transportation/land-use connection. I believe that in a world of ubiquitous telecommunications it may be less possible to influence travel patterns through land-use strategies, and it may be necessary to rethink these strategies given the rapid increase in telecommunications capabilities in relationship with physical mobility. Efforts to incorporate data on telecommunications into our analyses and forecasts will be absolutely necessary in order to figure out if this is right.

GOODS MOVEMENT

One of the greatest weaknesses facing transportation analysts is the prevailing absence of good data on goods movements in urban areas. This is one of the best ways of making the point with which I opened my talk. We collect most of our data on person movements and we perceive and define transportation problems in terms of person movements and we define solutions in terms of person movements. But, goods movement is growing and is a central issue in transportation policymaking, and we are ill equipped to deal with this. In several metropolitan areas there are proposals for truck-only highways; for the separation of trucks from passenger vehicles and for automated truck lanes as early deployment projects in the evolution of the automated highway system. Trucks are responsible for a substantial proportion of urban and intercity highway congestion and delay, yet many metropolitan areas continue to model truck movements by applying a multiplier to person movements. I think better information on truck movements or goods movement or freight movement, including its intermodal aspects, has been and remains one of the greatest information needs in the realm of transportation planning.

SUSTAINABILITY

It would not be an exaggeration to say that the dominating issue in transportation policy in the United States over the last quarter of the 20th century has been air quality. The provisions of the Clean Air Act and its several amendments have determined the direction of transportation planning in metropolitan America, leading many to observe that the transportation-policy dog was being wagged by the air-quality tail. Primarily because of advances in transportation technology, we have made enormous progress in meeting national ambient air quality standards; though along the way we have discovered some additional danger from sources previously not recognized as critical, such as small particulates. Interestingly, we have made great progress toward cleaner air even though our data collection and analysis tools never really got to the point where they were up to the job of adequately characterizing or forecasting key pollutants under alternative policy options.

Throughout Europe and Asia there is much more concern than there is here in the United States about the next great push in the movement to link transportation planning and policy to the broad theme of environmental quality. The term “sustainability” is more widely used. Sustainability roughly means that we should try to plan transportation systems that conserve energy, limit greenhouse gas emissions, and recycle waste materials and fluids sufficiently so that today’s mobility does not cause loss of life, injury, or depletion of needed resources tomorrow. Like many Americans I have in the past been more than a little cynical about this concept for I do believe that growth in mobility worldwide brings many social and economic and cultural advantages. But growing evidence that global warming is a credible threat has to be taken seriously, and international treaties do commit us to slowing the increase in the emission of greenhouse gasses, a substantial proportion of which have the transportation system as their source. In order to build a more “sustainable” transportation infrastructure, something I believe we are increasingly going to have to do, we need to more precisely define, to measure and to monitor the sustainability of the transportation system. I believe that will be as important a function of transportation planners during the coming decade as air quality issues have been over the last 15 or 20 years, and that there is a need to get started and to take this need seriously.

EQUITY

Transportation analysis is generally focussed on issues of effectiveness and efficiency. Our databases and tools like benefit-cost analysis and corridor studies are designed to tell us how well each alternative plan or design or course of action satisfies project or program criteria, and how efficiently they do so, per unit of capital and operating cost. One of the most pressing needs in policymaking is for more information about equity which is not well addressed by standard methods and databases. Equity analysis, of course, implies a concern with fairness and with the distribution of benefits and costs among different groups. The criteria by which we judge the equitability of different transportation policies are clearly highly subjective, but in a way that makes the question of data and modeling more urgent, more difficult and more complex. Recent disputes and

increasingly frequent lawsuits have made it clear that the distribution of project and program benefits among different spatial communities, ethnic groups, and economic classes of environmental and social and economic impacts are among the most critical dimensions of transportation projects in need of analysis and attention. Our databases and analytical models address distribution issues far less effectively than they address the effectiveness and efficiency questions, yet current policy debates almost always revolve around who benefits and who pays and how much. One of the most important areas in which I believe our tools can be strengthened in the relatively short run by careful and thoughtful adjustment and refocusing is in this realm of equity or distribution issues.

CONCLUSION

I've tried in this talk to demonstrate what I suspect everyone at this conference really already knows. First, I've tried to show that databases and information systems and analytical models interact with one another and change over time as our understanding of transportation systems and their social and economic contexts change over time. New understandings both shape and are shaped by the data and models that we use. Second, I suggested five ways in which I think planning needs and evolving transportation policy are shaping and changing our collective perceptions of data requirements in the near term future, the next 5 to 10 years.

I hope my thought processes have stimulated your own thinking about these issues, and that they may contribute in some way to the discussions that will take place at the remaining sessions of this conference. Now, if you would like to offer any comments or questions on my remarks, I would like to invite you to do so.

Thank you very much.