Introduction

"Focus first on those aspects of infrastructure that provide essential services, that is, those involving drinking water, wastewater, transportation, energy, and communications. … Business and population growth have already outpaced the capacity of existing systems. To meet user's expectations, planners should first determine the public's expectations with respect to the levels and resiliency of such services and the amount of money that should be spent to maintain them and then determine what alternatives exist and what actions need to be taken to meet those expectations."[1]

Toward More Sustainable Infrastructure: Better Projects and Better Programs

Modern societies depend upon vast infrastructure-based systems that support efficient transportation and communications, provide ample supplies of clean water and energy, and enable effective treatment and disposal of wastes. The performance of such systems can be measured in terms of many factors, including cost, energy consumption, resource requirements, capacity, service quality, safety, impacts on society, and impacts on the environment. Performance can also be measured in terms of sustainability, a broad concept that refers to the ability of a system to perform well over a very long period of time.

Sustainability is a particular concern for systems that rely heavily on non-renewable resources and systems that result in severe degradation of the environment. However, troubles in any aspect of performance can limit the sustainability of an infrastructure-based system. Sustainability can be enhanced by reducing costs, improving social and economic benefits, restricting the use of fossil fuels and other non-renewable resources, or reducing negative social and environmental impacts.

Many infrastructure projects and programs are aimed at improving some aspect of sustainability. Some are designed to ensure that the system continues to function properly. If infrastructure is inadequate or poorly managed, people may suffer from congestion, high costs, pollution, economic stagnation, or environmental degradation. To limit such problems, on-going investments may be required in new facilities, better materials, or new management techniques, although the nature of the infrastructure may remain about the same. Highways in 2016 may have real-time information signs, better paving materials, and synchronized traffic signals, but they still look and function much as highways did 50 years ago.

Other infrastructure projects and programs are designed to replace or upgrade systems that for some reason have become obsolete or non-sustainable. Over time, as economies develop, as societal norms change, and as certain resources become less available, the demands on infrastructure systems will change along with public perceptions of infrastructure performance. If infrastructure systems fail to evolve, they may eventually be recognized as being too costly, unsafe, disruptive to society, or overly-damaging to the natural environment. At that point, new systems are needed. For example, solar power and wind power can produce electricity that otherwise would have required additional power plants and more imported oil.

In short, infrastructure projects and programs are designed to improve some aspect of system performance. Better projects and better programs will lead to more sustainable infrastructure. The problem is how to determine which projects and which programs are better.

---

[1] One of the conclusions of *Sustainable Critical Infrastructure – A Framework for Meeting 21st Century Imperatives,* a report based upon a May 2008 workshop sponsored by the National Research Council ("Report Urges New Framework for Planning Critical Infrastructure." *Civil Engineering.* June 2009. p. 20.)

Infrastructure Projects and Programs

Infrastructure projects include large-scale, multi-dimensioned, long-term investments in transportation systems, buildings, water resources, communications, power generation, parks, schools, and other public services. Such projects always have multiple objectives, they will often be controversial, and people with many different perspectives must come together to complete the projects and make them successful. Such projects have important impacts for the public at large, because they will affect the environment, our society, and our economic prosperity. A program consists of a set of related projects, such as the Interstate Highway Program or a program designed to promote investment in wind power.

Historically, there have been numerous large-scale infrastructure projects and programs, some brilliant, some misguided, and many of them quite interesting for planners and engineers, such as the Panama Canal and skyscrapers in Manhattan. Projects like these affect the way we live, they are the backbone of much of our history, and they are the pathway to our future.

This book is about understanding where projects come from, how they are evaluated, how decisions are made to proceed with them, and what separates good projects from bad projects. This book spends considerable time on methodology, especially the methods of engineering economics that can be used to understand how projects are financed, but it also provides real-world examples and case studies that convey some of the flavor, excitement, and challenge of designing, evaluating, and implementing projects.

Figure 1
The Panama Canal

After decades of frustration, tens of thousands of deaths from tropical disease, bankruptcy and disgrace for the initial French Canal Company, the canal was finally was completed in 1914 and remains today a critical link in global transport and a highly profitable enterprise for Panama.

Implementing, operating and maintaining infrastructure requires planners and engineers to work with bankers, entrepreneurs, politicians, community leaders and the public in order to meet society's needs more effectively. Planners and engineers must learn to deal with the social, financial, and environmental issues related to infrastructure projects, and these issues will become more important over time. Engineers are likely to start out building and designing projects, and many engineers spend their entire careers concentrating on these activities. Planners and managers are likely to start out working at a low level on projects and programs that were begun years ago. However, someone, somewhere, is trying to figure out what to build next, when and where to build it, and how to convince investors and governments to pay for it. Actually, there are many such people, and some of them are destined to become famous. These people may end up proposing projects, or they may simply define problems and convince other people to begin working on them.

Engineers, planners and managers naturally expect to work on large-scale infrastructure projects. To succeed and to advance in their professions, they will need to understand the big picture - the needs of society - in order to take the lead in designing, implementing or marketing new technologies or new systems. Leaders will need a broader outlook on problems than is ordinarily conveyed in an engineering subject or a textbook on finance. They will need to understand how projects begin, how they are sold to the public, and how they become successful. They will need to combine engineering or planning skills with marketing, financial, and communications skills. Anyone who grasps this broader outlook will have a chance to become involved in projects and programs that are increasingly complex, with more possibilities for design and implementation strategy, less certainty regarding the outcomes, and greater need for imagination and leadership.

Infrastructure is usually defined in terms of public systems, and constructing and maintaining infrastructure is an on-going process and problem for local and national governments. Infrastructure refers to the physical systems that provide transportation, water, buildings, and other public facilities that are needed to meet basic societal needs. These facilities are needed by people regardless of their level of economic development. When infrastructure is not present or does not work properly, it is impossible to provide basic services such as food distribution, shelter, medical care, and safe drinking water. Maintaining infrastructure is a constant and expensive process that often is neglected in favor of more attractive political goals.

In practice, much of the civil infrastructure may be owned and maintained by private companies or individuals. Much infrastructure was originally built by private corporations with licenses or other authorization from government; private toll roads were the norm in the United States in the early 19th century, and private expressways are being built today in many parts of the world. Many railroad systems are privately owned and operated. Large office buildings or apartment buildings are mostly privately owned, and they are certainly part of the basic infrastructure of a modern city.

Project evaluation may involve assessment of proposed options for creating, maintaining, rehabilitating or decommissioning any kind of infrastructure, whether carried out by the public sector, the private sector or a public/private partnership. Thus, building a new road, adding a lane to an existing road, or paving an existing road could all be considered infrastructure projects. For administrative convenience, a large project will often be broken down into multiple smaller projects. The construction of a new road may involve construction of a dozen bridges, three major interchanges, extensive cut-and-fill operations to prepare the right-of-way, and eventually the actual paving of the road. Moreover, the road may be completed in multiple phases over a period of many years. Whether to consider each of these activities as a separate project, each phase as a separate project, or the entire road as a single project could be debated; there will certainly be a well-defined set of contracts and sub-contracts so that all of the contractors have a clear perspective on their portion of the overall project. The public, however, will likely view the whole road construction as a single, multi-phased project. The distinction is usually unimportant, although there will be times that a small segment of a road (or a small portion of some other project) will be proposed, hoping to gain approval more easily later on for an extension after "getting your toe in the door".

An infrastructure program may be established as a way to manage a series of projects or a way to simplify the design and approval process for multiple projects. A program may specify goals and criteria for measuring progress against those goals. It may also specify what kinds of projects will qualify to be included in the program and what kinds of incentives will be available to qualifying projects. For example, a state may establish a program aimed at attracting private investment in housing for low-income families. The program may provide subsidies, tax relief, or other benefits to projects that qualify according to the criteria specified in the legislation or regulations. A company may also have infrastructure programs; retailers such as Home Depot or Wal-Mart will have plans for expanding their network of stores and warehouses. A railroad may have a plan for upgrading its oldest bridges on certain high density lines; each bridge renewal would become a separate project as a part of the program. Cities and states may have programs aimed at providing housing for the elderly or for low-income residents, and they may have programs aimed at improving water supplies or sewage treatment facilities. The various interest groups and political leaders who favor or oppose a certain type of project will fight over the structure of a program, perhaps for many years, but eventually they may reach agreement about the objectives, scope, funding amounts, and funding eligibility for the program. Once a program has been established, those prolonged fights will cease, and projects can rather quickly be identified, approved, and implemented. It will be desirable from time to time to review programs to ensure that the objectives remain valid, that the funding mechanisms are adequate and fair, and that the projects as implemented under the program actually have been achieving the program's objectives.

Infrastructure projects and programs have several common and very interesting aspects:

- Infrastructure is intended to last a very long time, so it is necessary to compare what may be very large current expenses with the potential for benefits that will be gained only over a period of decades.
- Infrastructure influences and perhaps defines the location and land use of cities and regions, so the location of infrastructure will have long-term implications for local and regional land use.
- Infrastructure often involves networks of facilities that are widely dispersed, perhaps with severe consequences for the environment or for the people who live where the networks are located.
- Infrastructure benefits are frequently qualitative or difficult to measure, e.g. mobility, safety, air quality, or the availability of clean water.
- Infrastructure projects and programs will be of great concern to many different groups of people, including developers, the public, special interest groups (some of which may be public interest groups and some of which may be supporting very narrow private interests), governments (including elected officials, regulatory officials, and administrative officials), lawyers, users, abutters, construction companies, and investors.
- Infrastructure is costly to build and costly to maintain.

The long lives expected for infrastructure cannot be achieved unless funding is available for proper management, including safe operating practices, on-going inspection and maintenance, and periodic renewal and upgrades. Without such funding, infrastructure systems will deteriorate and eventually be unable to meet the societal needs they were designed to serve. Without adequate funds for renewal and expansion, it will be impossible to meet growing needs for services or to capture the benefits of new technologies.

Adequate financing must therefore be considered an essential factor in improving the sustainability of infrastructure systems, where "sustainability" refers to the ability of a system to function long into the future. Poorly managed infrastructure systems that steadily deteriorate, become congested, or become unsafe clearly are not sustainable. However, adequate financing is but one of the major factors affecting the sustainability of infrastructure.

Large-scale infrastructure, even if it appears to be adequately financed, can only be sustained over long periods of time if it is supported by society and the resources it requires are available at a reasonable cost. If infrastructure requires excessive use of non-renewable resources, if it requires too much water or energy, or if its use results in devastation of the environment, then the lack of resources, increasing costs of materials, or public outrage will force changes. If construction, maintenance, and operations continually disrupt neighborhoods, cause human suffering, or expose people to potentially catastrophic risks, then society will be reluctant to support further expansion of that kind of infrastructure.

Over time, social norms may change, the costs of resources may vary, and new technologies may emerge. What one generation viewed as highly beneficial investments may be viewed as dubious achievements or even disasters by following generations. Infrastructure systems must evolve along with society, and rising concerns about public safety, public health, climate change, pollution, environmental decline mean that society will require more sustainable infrastructure. Water shortages, highway fatalities, urban congestion, over-dependence upon fossil fuels, toxic chemicals associated with large-scale agriculture, acid rain, oil spills, and excessive amounts of solid waste are all symptoms of problems that reflect a need for more sustainable infrastructure and a more sustainable way of life. Challenges such as those posed by climate change, oil depletion, collapse of fisheries, and large numbers of endangered species combine to make stewardship of the environment and sustainable development greater concerns for society.

Achieving more sustainable infrastructure will require thought, innovation, planning, financing, regulation, and leadership. There clearly is a continuing need for large investments in infrastructure, and there will be many opportunities for evaluating projects and programs related to all types of infrastructure. Evaluating projects and programs will require methodologies for comparing current and future impacts, for considering multiple objectives, for assessing both quantitative and qualitative information, and for communicating and negotiating with diverse groups of people.

Evaluating Infrastructure Projects

The main goal of project evaluation is to help in identifying and implementing successful projects and programs. From an overall perspective, a project is successful if:

1. It was built, which proved that construction was feasible from engineering, financial, and social perspectives.
2. The benefits were indeed greater than the costs.
3. The project as built was an effective way to achieve those benefits.
4. The project was built in an efficient and effective manner:
    a. There were no clearly better options.
    b. There were no significant negative externalities.
5. Building this project did not foreclose other, even better projects.

Different participants might have far narrower definitions of success. Did the engineers design a building that was safe? Did the contractors get paid? Did clean water actually come to the neighborhoods? Did the mayor get re-elected? These different perspectives must of course be considered in evaluating projects, but it is useful for students, consultants, concerned citizens, honorable developers, and honest politicians to pay some attention to the overall issues.

Project evaluation is a qualitative process as much as it is a quantitative one. A critical step is to create a “story” for the project that can be used to explain why the project is needed, what it will do, what the benefits and costs will be, and why this is the best way to proceed. There will certainly be quantitative aspects to the process, although estimates of costs and benefits may be rather ill-defined and subject to debate.

Implementing and maintaining a project over a long period of time will require:

- Financing: sufficient income to cover expenses, whether the income comes from user fees, investors, subsidies, or contractual payments.
- Government approvals: licensing and periodic inspections to ensure compliance with safety, environmental, and other regulatory matters.
- Engineering skills: sufficient knowledge and skilled manpower to conduct the maintenance and rehabilitation necessary to perform at an acceptable level of service.
- Resources: people and materials as required for maintenance and operations and whatever additional resources are needed by users (e.g. asphalt for highway maintenance plus gasoline for drivers).
- Public support (or tolerable opposition and interference).

The financing issue is different from the economic issue. Financing provides the cash necessary to construct, operate, and maintain a project. The ability of a project to be financed depends upon the availability of money – not upon the actual economic benefits of the projects. Economic issues concern the costs and benefits associated with a project, the distribution of those costs and benefits, and whether the benefits are sufficient to justify the costs. Economic benefits may include creation of jobs, congestion relief, reduction in accidents, or improved productivity for those affected by the project. Some of these benefits may be easily described in monetary terms, and some may be very difficult to quantify in monetary or any other terms. They are economic benefits because they allow more efficient and more effective use of resources, even if the benefits do not translate directly into cash for the project or for investors.

It may be helpful for a project to have economic benefits in order to attract public or private financing. For example, governments may choose to subsidize transit operations, housing for low income or elderly residents, or agriculture. The cash provided by those subsidies can in fact attract investors, who will create commuter rail services, apartment buildings, and more productive farms. Whether or not these projects are really worth the subsidies that they receive is important for legislative bodies and elected officials to consider, but not necessarily something that will concern investors.

Government approval will be needed for any almost any project. A building permit will be needed for constructing a screen house in your back yard or for constructing a 100-story office building. Governments may establish regulations concerning land use, protection of the environment, the siting and size of buildings, construction materials and methods, the use of union or local labor, and many other factors that may affect the feasibility, cost, and ultimate success or failure of the project. Whether or not government agencies approve proposals or provide the necessary permits may depend upon legislation, regulations, the whim of administrators, and/or feedback from the public. Large projects tend to generate large criticism, so developers must always be concerned with public perceptions of their projects and they must be aware of ways to make their projects more attractive to the public.

People with the necessary skills are needed in designing projects, in constructing them, and in ensuring they continue to function. It is one thing to build a road. It is another thing to enforce weight limits to ensure that overloaded trucks do not destroy the pavement within a few years, to enforce speed limits so as to promote safe driving conditions, and to establish periodic inspections, maintenance, and rehabilitation to keep the road in safe condition.

Projects and the people who use them or depend upon them will need resources for operations and maintenance over what may be a very long lifetime. Projects may fail because the resources needed to sustain them become too costly or unavailable. Some of the most pressing issues of the 21st century relate to the continued availability of fossil fuels for transportation, electrical power generation and home heating, and the availability of water for irrigation, household consumption, and industrial use. Many projects and infrastructure choices were justified based upon usually unstated assumptions that unlimited supplies of cheap oil and water would always be available. Fossil fuels, however, will not last forever, and prices will rise as reserves of oil, coal and natural gas are used up. With cheap oil, automobiles and airlines prosper; with expensive oil, transit and rail transportation become more competitive. With abundant water supplies, crops can be grown in irrigated deserts, people can compete for the greenest lawns, and industries can use processes that consume vast amounts of water. Eventually, however, as population growth and other demands for water increase, the supply of water is no longer sufficient for all the possible uses, so the use of water will be regulated and the price of water will rise. Moreover, water supplies may diminish. Regions that are heavily dependent upon well water may find that their aquifers are drying up. In other regions, changes in climate may diminish the amounts of water that is available. Since drainage and river basins follow geographical rather than political boundaries, rival demands for the use of water have and will continue to spark political battles between neighboring states and countries. A populous region, such as the Los Angeles metropolitan area, will seek to divert water from distant regions in order to support their needs, while perhaps limiting the growth and productivity of the regions from which the water is diverted. Disputes over oil reserves have already sparked conflicts in the Middle East, and the potential for future conflict will continue as long as so much of the world's transportation, power generation, and industrial production is fueled by oil.

Public support, or at least tolerable opposition, is the final factor necessary for the long-term success of a project. The public normally does not have a direct role in decisions regarding major projects, as most decisions regarding projects are made by elected officials, appointed officials and legislative bodies. However, the public can provide input into the decision process, whether by participating in a process established to promote public involvement, by writing to newspapers or elected officials, or by organizing groups to support or oppose projects. Public opposition can prevent particular projects, it can lead to new regulations or legislation, and it can change programs and policy. In the late 1960s and early 1970s, public opposition was the major factor in halting construction of major urban portions of the Interstate Highway System, including the so-called Inner Belt and the Southwest Expressway in Boston and the Embarcadero in San Francisco. Public concerns over the safety of nuclear power plants had led to stringent regulation of the construction of such plants in the US by the 1970s; public outrage after a rather minor leakage incident at the Three-Mile Island Nuclear Power Plant effectively halted construction of such plants in the U.S. for decades.

Infrastructure, Cities, and Civilization

It can be argued that infrastructure projects are the key to urbanization, which is perhaps the chief characteristic of civilization. If people are to be able to congregate in cities, then they will need access to large amounts of clean water, and they will need to have some system for treating or isolating wastes. They will need to import food, building materials, and energy resources. They will need facilities and materials to support various kinds of manufacturing and trade. They will want to create facilities for education, sports and worship, for communications and entertainment. In short, people will have to construct the infrastructure necessary to support all of the normal functioning of a densely populated society.

> *Your imagination, your initiative, and your indignation will determine whether we build a society where progress is the servant of our needs or a society where old values and new visions are buried under unbridled growth. For, in your time, we have the opportunity to move not only toward the rich society and the powerful society but toward the Great Society. The Great Society rests on abundance and liberty for all. It demands an end to poverty and racial injustice, ... It is a place where the city of man serves not only the needs of the body and the demands of commerce but the desire for beauty and the hunger for community. ...*
>
> *Our society will never be great until our cities are great. ...*
>
> Lyndon B. Johnson, President of the United States,
> excerpts from the "Great Society Speech" delivered at the University of Michigan, May 22, 1964

The benefits of urbanization can be great for people's lifestyles and for efficient use of resources. Higher populations can support a diversity of lifestyles and greater opportunities for jobs and recreations. There can be a greater frequency of and higher quality for social events. When people no longer have to spend all of their time eking out a living, whether on a farm or in isolated rural areas, they will have sufficient time to enjoy the fruits of civilization. From a systems standpoint, having large numbers of people living in a small area allows more efficient use of resources in constructing and operating transportation networks, creating housing, supplying water and treating waste. As activities are differentiated, complementary activities can be concentrated within special districts of the city. When people are concentrated in well-situated cities with sound infrastructure, they can be protected from natural disasters, and it is possible to manage development so as to reduce the consequences of manmade disasters.

Of course, as Freud pointed out in his book *Civilization and Its Discontents,* crowding vast numbers of people into cities may not be good for everyone. The more we protect ourselves from natural disasters and the more contact that we are forced to have with each other, the more difficult it may be for us to live together. There is not only the loss of self-sufficiency that may be achievable on a farm, but there is also the possibility of extreme poverty. A city is dependent upon its infrastructure – and transportation or water resource systems may fail. If diseases break out, thousands may die, and pollution and the inability to absorb wastes may become continuous drains on health and happiness. As cities grow ever larger, congestion is likely to limit mobility, and it may become ever more difficult to limit pollution, to provide open space and to ensure adequate housing for everyone.

Whether cities evolve into safe, livable, aesthetically pleasing places or degenerate into overcrowded dens of despair depends to a very great extent upon the ability of the people of those cities to undertake the projects that will enable them to meet the needs of human life and challenges of urban life. Anticipating and responding to challenges is the driving force for successful civil and environmental projects. And there will always be new challenges.

Tomorrow's challenges may be quite different from yesterday's, but there will always be basic needs to be met and there will always be a need for evaluating and choosing the best ways to meet those needs. Even with tremendous advances in communications and computers, with automated factories and computer-controlled highway networks, with cheap transportation for freight, a global economy, and ever-improving medical care, there will still be plenty to do. After all, only about half of the world's population has access to clean running water; hundreds of thousands of people die each year in transportation accidents; earthquakes and other natural disasters cause thousands of fatalities; billions of people live in substandard housing; and nearly everyone who lives in a large city spends a large portion of their life stuck in traffic and breathing bad air.

Where Do Projects Come From?

A project begins long before the groundbreaking, long before the first contract is signed, and long before a specific plan is identified and agreed upon by people with the resources and political power to make something happen. A project begins with an idea, with a vision of what is wrong or what is needed or what is possible. Initial ideas quickly evolve into whole families of ideas and possibilities and soon different, competing options begin to emerge. Long before the time for computer analysis and project planning, strong-minded, imaginative, entrepreneurial, and political individuals are vying to promote their concepts for the future. The players might include engineers, politicians, charlatans, financiers, developers, or dreamers. There are no bounds to how they might think or talk about the project, or how they conceive the project fitting in to what is already in place or what could be put in place. Their creative processes can be slow or rapid, rational or chaotic, cooperative or acrimonious – there are no rules and there are no limits to how hard people will push.

This undisciplined, often unmannerly process eventually leads to a specific project that will be constructed to finely drawn plans with a well-defined scheme for paying for it all. At this point, and not before this point, project management skills are needed, and there will be plenty of work for those with specialized software, algorithms, and risk management techniques that can lead to more efficient designs and timely completion of the project. But those skills are not much use in the early stages of project design and evaluation.

It is these early stages where there is the greatest uncertainty, the most excitement, the widest opportunities for egregious errors, and the best chances for achieving elegance in a project. It is difficult to teach how to conduct this process for which there are no rules and few guideposts. By the time that the processes are well-enough defined to create guidebooks for planners, the damage of poorly conceived projects will be only too apparent. We built highways straight through cities for decades before stopping to think seriously about the effects on the neighborhoods and the possibility for justifying less disruptive, more effective approaches. We need to think before we leap, we need to appreciate the creative, political, and entrepreneurial efforts that are needed, and we need to avoid the pitfalls that can catch the unwary.

A Framework for Project Evaluation

Project evaluation can be broadly conceived to include five phases that cover the entire life-cycle of a project:

1. Project identification
2. Analysis of alternatives
3. Assessing and comparing alternatives
4. Implementation
5. On-going evaluation

The first three phases may require many iterations before a final project is approved, and the final phase should continue over the entire life of a project.

Project Identification. The first phase is the least well-defined and yet the most important for the ultimate success of a project or a program. Many ideas for projects arise in response to perceived problems and the needs of society. Congestion leads to ideas for new roads or new transit systems. Rising populations require new schools, housing, and drinking water. If problems and needs are understood, and if there is a process for examining possible ways to deal with them, then it should be possible to develop effective projects and programs that result in a better society. However, there will not necessarily be any process for determining and responding to societal needs. The ideas for many projects may originate when someone senses an opportunity to make some money or to create some sort of monument. Ideas for projects might well come from someone – an entrepreneur, a company or a public official - who spots an opportunity for using a new technology, for developing a particular plot of land or for expanding an existing network of facilities. It may well happen that project proponents first identify the project and then address the problems or needs that would be addressed by this project.

Nevertheless, it is useful to have a framework in which the first step examines problems or needs. For an infrastructure-based system, problems are likely to relate to cost, capacity, service quality, or safety. A problem may exist if some aspect of performance is believed to restrict the efficiency or effectiveness of the system. System operators will likely be aware of ways to improve performance, based upon their own insight into operations or based upon comparisons with similar systems in other locations. A need for better performance may be evident from user complaints, media reports, or scientific studies. Needs may be expressed in terms that are much different than the terms used to define problems. For example, transportation needs might be expressed in terms of mobility and accessibility, whereas transportation problems might be expressed in terms of travel delays and maintenance costs.

The objectives of the project need to be clear and well defined, but they can be modified based upon feedback and assessments concerning completed projects or new information related to needs and opportunities. The need for flexibility may lead to certain challenges in the overall decision-making and implementation process. Sometimes strategic objectives are too narrowly defined and remain fixed despite changing conditions and acquisition of new information. Sometimes objectives are in conflict with objectives of other programs, particularly in the public sector, so that projects can only be developed after due consideration of related programs.

The next step is to generate alternatives for addressing the problems and needs that have been identified. Problems and needs should be considered in general terms, so that different kinds of alternatives can be considered. For example, many systems must deal with potential capacity problems related to growth in population. If so, then alternatives could not only consider expanding capacity to keep pace with population growth, but also consider increasing prices in order to limit demand or increasing efficiency of operations in order to allow more effective use of existing capacity.

The project identification phase concludes with a clear statement of needs, a set of objectives and specific assessment criteria, and an initial list of alternatives for achieving the objectives. Key results from this stage of project evaluation include clear statements of needs and objectives, the establishment of criteria, and the selection of alternatives for further study.

Analysis of Alternatives. The process then enters the analysis phase, in which studies provide information that will help in assessing and comparing the various alternatives that are being evaluated. Various studies will be necessary to assess the viability of each alternative with respect to technical, financial, operational, social, economic, environmental or other objectives. Considerable discussion and thought will be devoted to identifying performance measures and evaluation criteria for each major objective. Preliminary studies may give an early indication of the viability of an alternative, along with the risk involved. The most promising alternatives will be studied in greater depth. Analysis may include market demand studies, cost-benefit analysis, environmental impact assessment, and social assessment. Very detailed analyses involving multiple groups of people with backgrounds in engineering, economics, environmental science or other disciplines may be required. Important planning decisions during this phase of project evaluation include the allocation of resources to the different types of studies and the extent to which the process allows refinement and modification of alternatives.

Assessing and Comparing Alternatives. Assessing the results of the analysis is a separate stage from analysis, because there will be many different kinds of results to be considered. During this phase, it will be necessary to compare alternatives with respect to how well they satisfy the objectives that were previously established. Assessment will involve consideration of financial, economic, environmental and social factors. To what extent does each alternative meet the needs that are being addressed? What are the costs and benefits of each alternative? Are costs and benefits measured properly? To what extent does each alternative lead to positive or negative externalities, i.e. to broader impacts on the environment or the community or the region that would result from implementing a particular alternative?

Whereas analysis requires specialists and may include many independent studies, assessment requires generalists. For public projects and for large private projects that require public approval, there will have to be opportunities for input from potential users, abutters, and the general public. Users may push for a bigger and better system. Abutters, those who live next to the construction sites, may like the concept of the project, but oppose the proposed location. This type of opposition is so common that it is known by an acronym - NIMBY – which means "Not In My Back Yard." The general public, to the extent that is informed about the issues, is likely to be more receptive to a more balanced approach that recognizes the potential benefits of the project while acknowledging the importance of externalities.

The goal at this stage is not necessarily to define the exact, best option, but to determine the general approach that is best. The outcome from this stage could be one of three broad conclusions:

- One alternative clearly is the best.
- Further study is necessary to determine which alternative is best.
- None of the alternatives is worth pursuing.

If one alternative is clearly the best, then it is possible to proceed to the next phase. If there is no alternative that is clearly the best, then more detailed analysis may be needed that focus on what are believed to be the most promising alternatives. It may also be desirable to revise some of the alternatives or to suggest new alternatives or different kinds of analysis. This phase of the evaluation process requires the consideration of multiple objectives as well as risk assessment in order to compare what could be markedly different alternatives. It also requires some mechanism for ensuring that there are no better alternatives that should have been studied, as well as a mechanism for determining that the preferred alternative in fact is a cost-effective way of meeting the needs identified at the outset. Table 1 suggests some guidelines for this phase of project evaluation.

Implementation. Project identification, analysis and assessment are iterative processes that may continue for years or decades without finding an alternative that is technically, financially, and politically feasible. Eventually, it may be possible to agree upon a particular alternative. The fine-tuning of a particular alternative may involve mitigation of environmental or social impacts, it may involve modifications aimed at reducing costs or increasing benefits (a process known as value-engineering), and it may involve modifications to incorporate recommendations resulting from public input or the various studies that were conducted. At some point, detailed engineering design can be completed, and a construction management program can be initiated. A strategy for construction must be developed. How soon should construction begin? How quickly should construction proceed? What are the possibilities for implementing the project in stages? Once these questions have been answered, a project management team will be in charge of the actual construction process, and there will be innumerable decisions related to the best construction techniques, logistics, coordination of sub-contractors, communications and cooperation with relevant public authorities and maintaining the safety and security of the site. Before construction is complete, it will be necessary to begin the transition from to operation. Eventually the construction phase ends and the project is up and operating: the bridge is open, the tenants are in the building, the water is flowing, or the park is opened to the public.

Table 1 Guidelines for Assessing Projects

---

1. Address the grand issues.
   - Economic viability – is there a clear case for supporting the project?
   - Engineering – what are the options regarding capacity, staging, and flexibility?
   - Financial feasibility – is there a way to cover investment and operating costs?
   - Environmental impacts – can the project be done with less negative impact on the environment? Can it result in improvements to the environment?
   - Political feasibility – who is likely to support or oppose the project? How can negative social impacts be mitigated?
   - Organizational structure – is the project best done as a public project, a private project, or a public/private partnership?
   - Size – would a larger or smaller project be better than what is proposed?

2. Consider comparable projects to get a quick, though rough estimate of the viability of the project.

3. Consider the possibility that the benefits are so great that there is more danger from doing too little than from doing too much.

4. Be prepared to think at all scales: local, regional and national.

5. Think about aesthetics and plan with an eye to style.

On-going Evaluation. Few projects are so well-planned and so carefully executed that everything goes perfectly on day one of the transition. There will be a period of time during which minor problems will be identified and corrected. After operations have settled down, it will be possible to compare the actual performance to what was intended. Was the project completed as planned? Was it completed on time and on budget? Most importantly, how effectively has the project addressed the original problems and needs? Answers to questions such as these will help in planning the next project and perhaps help in creating criteria for a program for constructing many similar projects.

In summary, the process of defining a project can be viewed as a logical sequence of well-defined steps beginning with identification of needs and concluding with on-going monitoring of performance. While it is useful to have a framework such as this for thinking about projects and project evaluation, it is important to recognize two fundamental aspects of the process of defining and selecting projects.

First, the process is iterative. It may begin either with identifying needs, technological opportunities, or with an idea for a specific project. Once assessment begins, new ideas may emerge or people may find serious problems with all of the proposals, so it will be necessary to reconsider the needs and the opportunities.

Second, the process may not necessarily be logical or rational. Suggestions for projects may come from those who want to build them or from those who want to operate them – whether or not the projects they propose are the best projects or the projects that respond to the most pressing needs of society. Companies that build roads and bridges want to build more roads and bridges, just as highway authorities may respond to all transportation problems by recommending construction of more highways. New technologies quickly lead to ideas for new projects, but it may be years or decades or longer before those projects can be justified. With many new technologies, the new capabilities create new needs, or at least perceived needs (continuous, instantaneous connections to the internet; high definition TV). With advertising, suppliers can create needs that drive construction of new plants and distribution facilities (bottled water is a good example – especially when the water is obtained directly from a region's public water supply). It is a mistake to expect the process to be completely rational. On the other hand, it is also a mistake not to try to impose a rational process on defining needs, identifying alternatives, and assessing, selecting, and modifying alternatives.

Essays and Case Studies

This book includes a collection of essays and cases studies that address the many phases of project evaluation. The first volume provides a framework for understanding and evaluating projects, taking into account not only the financial and economic issues, but also social and environmental factors. Examples and case studies illustrate the complexity of major projects and demonstrate the role for and the limits of analysis in clarifying and resolving issues. The second volume shows how to apply the basic methods of engineering economics in evaluating major infrastructure projects. Examples and exercises indicate how to develop and apply models for estimating the costs of resources required for such projects and how to estimate their life-cycle costs. A major goal of both volumes is to promote an approach to project evaluation that is based upon an appreciation of the needs of society, the potential for sustainable development, and a recognition of the problems that may result from poorly conceived or poorly implemented projects and programs.

Key concepts include the following:

- **Justification of large investments**: how to determine whether future benefits justify current costs.
- **Technology-based performance functions**: creating functions with sufficient detail to explore how cost, service, capacity, and safety vary with major project options related to size, design, and technology.
- **Cost-effectiveness**: how to compare options for achieving non-monetary benefits.
- **Sustainability**: environmental, financial, economic and social aspects of sustainability.
- **Evolution of systems**: understanding how systems evolve in response to changes in needs, technologies, and financial capabilities.

Analytical methodologies can be applied to each of these concepts. However, it is critical to recognize that analysis will not necessarily determine what projects are considered, what projects are proposed, which of these projects are approved or which projects are ultimately successful.

Projects may be motivated by a vision of a greater society, by an idea for addressing a specific local problem, by the prospects of making a profit while providing a needed service, or by simple greed. Some apparently excellent projects cannot be financed, while it may be easy to fund some very questionable projects. Lackluster projects may prevent outstanding projects, and highly acclaimed projects may prevent dozens of less showy, but more effective projects. Financially successful projects may be terrible in terms of their consequences for the environment, and projects sold as being good for the environment may turn out to be overly expensive or socially unacceptable.

Project evaluation is not a hard science, as there are so many factors to consider, so many unknowns, and so many different perspectives concerning what is good or bad. Nevertheless, there is a role for analysis, if only to help people to recognize and agree upon the likely magnitude of the most important costs and benefits. Past experience, a coherent framework for analysis, and a concern for sustainability will provide a sound basis for evaluating projects, whether you are the developer, the consultant, the banker, the neighbor, the user, or the politician.

This text is published in two volumes. The first volume provides an overview of project evaluation as a multi-dimensional process aimed at creating projects that meet the needs of society. This volume emphasizes the need to consider economic, environmental and social factors along with the technological and financial matters that are crucial to the success of a project. It concludes with a chapter that considers the evolution of infrastructure-based systems and the need for more sustainable infrastructure in the coming decades.

The second volume provides in-depth coverage of the engineering economic methodologies that can be used to compare cash flows or economic costs and benefits over the life of a project. That volume presents the techniques that are used by investors, bankers, and entrepreneurs in deciding whether or not to finance projects. It also shows how public policy can use taxes and other regulations to encourage projects that have public benefits. Both volumes present methodologies that are useful in developing and evaluating projects to deal with problems and opportunities.

As noted in the preface, the essays and case studies are all structured as stand-alone documents, so it is possible to pick and choose which ones to read, and it is possible to read them in any order.

Basic Economic Concepts

> *"The quality of a nation's infrastructure is a critical index of its economic vitality. Reliable transportation, clean water, and safe disposal of wastes are basic elements of a civilized society and a productive economy. Their absence or failure introduces an intolerable dimension of risk and hardship to everyday life, and a major obstacle to growth and competitiveness."*[1]

Introduction

This essay introduces various economic concepts that are useful in understanding infrastructure systems and in identifying and evaluating potential projects for improving their performance. The chapter begins with a discussion of how **equilibrium prices** result from the interaction of **supply and demand**. If prices are high, say for office space or for energy, then a great deal of investment in new buildings or oil drilling or wind power is justifiable. If too many buildings are built or if too much oil is available on the world market, then prices fall and investments based upon continuing high prices may well fail. The success or failure of any major project will depend in part upon the future interactions between supply and demand.

Costs, prices, and values are distinct concepts that should not be confused. The cost of providing a service or of manufacturing a product depends upon such things as resource requirements, capacity requirements, and unit costs associated with operations. While owners surely desire that prices be higher than costs, prices are usually determined by market forces that may have little or no relationship to cost. The value of a product or a service is something that can only be determined by potential purchasers: if they perceive the value of a product or service to be higher than the price, then they will go ahead and make the purchase. The difference between what they were willing to pay and what they actually paid is an economic benefit known as **consumer surplus**, which is in fact an economic benefit even though it does not result in any revenue to the supplier. Large infrastructure projects are often justified in part by increases in consumer surplus, so this is an important concept for evaluating such projects.

From an economic perspective, a major goal of any project will be to increase **productivity**, which is defined as the ratio of system output to system input. If productivity improves, then more output can be obtained using the same or fewer resources, resulting in an overall benefit for society. If a company is able to produce more without increasing its labor force, then it may be able to afford to pay higher wages to its employees. Companies and agencies that manage infrastructure will continually be seeking ways to make more productive uses of their resources, and productivity improvement motivates many infrastructure projects and programs. In most infrastructure systems, there are **economies of scale, scope or density** that allow larger, more complex systems to offer more benefits at a lower cost.

Lower cost would seem to be a clear benefit to society, but project evaluation must consider who will capture the benefits of lower cost, the supplier or the customers? The answer to this question depends upon the extent of competition. If there are many potential suppliers, then there will be competition for customers, and prices will fall to marginal costs (**marginal cost pricing**) and customers will benefit from any productivity improvements. However, if a single supplier has no competition or very limited competition, then it will be able to charge prices that are well above marginal costs. The threat of **monopoly pricing** is therefore present whenever there are strong economies of scale. To achieve public benefits from scale economies related to essential infrastructure, it may be necessary to have public ownership or some sort of **price regulation**.

There are multiple reasons why infrastructure performance and major infrastructure projects will always be of interest to the public:

---

[1] National Council on Public Works Improvement, **Fragile Foundations: a Report on America's Public Works**, Final Report to the President and Congress, February 1988, p. 1

- First, the public uses the infrastructure, and the performance of the infrastructure affects everyone's daily life.
- Second, much of the public infrastructure is owned or regulated by public agencies, so that there is a direct public interest in managing and investing in that infrastructure.
- Third, infrastructure projects are large projects with long-lasting impacts on society and the environment, and the public has a justifiable interest in questioning whether these impacts are positive or negative and whether the costs and benefits of a major project are equitably shared.
- Fourth, investment in infrastructure projects can provide a boost to the region in terms of jobs, income, and economic growth through what is called the **multiplier effect**.

Infrastructure needs depend in part upon the economic forces that drive regional, national, and international development. Where goods are produced depends in part upon where raw materials can be found, where it is most efficient to produce the goods, where labor and other resources are cheapest, and the cost of transportation. As transportation costs decline, because of improvements in technology and expansions of transportation infrastructure, distance ceases to be an impediment to consolidation of agriculture, manufacturing, mining, and other industrial activities. Cheap transportation has enabled the rise of a global economy, and regions in one country now compete with regions in other countries for all sorts of economic activities. Two concepts that are directly relevant to understanding the global economy are **spatial price equilibrium** and **comparative advantage.** As patterns of trade and production shift, the needs for industrial facilities and transportation infrastructure also shift. In the less developed parts of the world, investments in infrastructure may be required for the economic growth. In the developed parts of the world, existing infrastructure that was designed for the economy of the 19th or 20th centuries may need to be redeveloped or replaced by infrastructure relevant to the 21st, with greater emphasis on major ports and continental distribution systems and less emphasis on access to local production facilities.

Project evaluation requires consideration of broad economic issues such as globalization, the need for regulatory policy, corporate decision-making, and the importance of regional economic impacts. However, it is also worth considering the perspective of the individuals who ultimately will be making the decisions that determine which types of infrastructure are used, how much revenue is gained, and whether or not infrastructure projects prove to be successful. Individuals decide such things as how much living space they need, whether to live in the city or a suburb, whether or not to water the lawn on a regular basis, whether to drive or take the bus to work, whether to switch from oil to natural gas for home heating, and where to go on vacation. The concept of **utility** provides a framework that can be used to understand how these decisions are made. The basic idea is that individuals are assumed to make decisions that maximize their utility based upon personal constraints related to time and money.

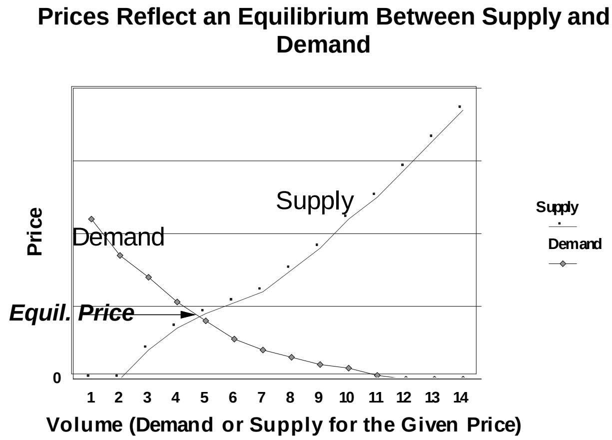
Supply, Demand, Equilibrium

<u>Overview</u>

Supply, demand, and equilibrium are central issues in economics. At the most basic level, both supply and demand are described as functions of price, and the **equilibrium price** is the price at which supply equals demand. The **supply function** shows the quantity of goods or services that will be produced for each price. Under normal circumstances, the supply of goods and services would be expected to increase as the price increases. If the price is higher, then existing suppliers will be willing to produce more, and new suppliers may be enticed to enter the market. The **demand function** shows the quantity of goods or services that will be purchased for each price. Under normal circumstances, the demand will decline as the price increases. Some people may be willing to pay a high price, but more people will be willing to pay when the price is lowered.

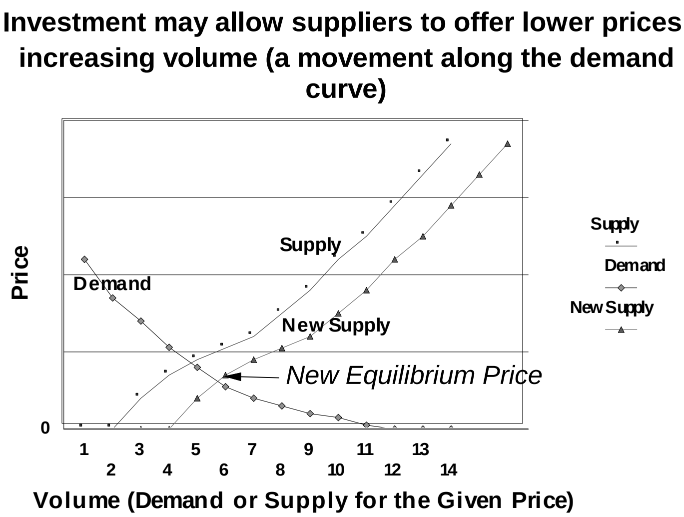
The interaction between supply and demand can conveniently be expressed in a chart as portrayed in Figure 1. Note the convention that price is shown on the y-axis, although that is assumed to be the independent variable, while the volume or quantity of supply and demand are shown on the x-axis. The point at which the supply and demand functions intersect is the equilibrium price. What is most important to understand is that this equilibrium price reflects both supply and demand: under competitive market conditions, prices will adjust to changes in supply and demand, and there will be a tendency for

Price



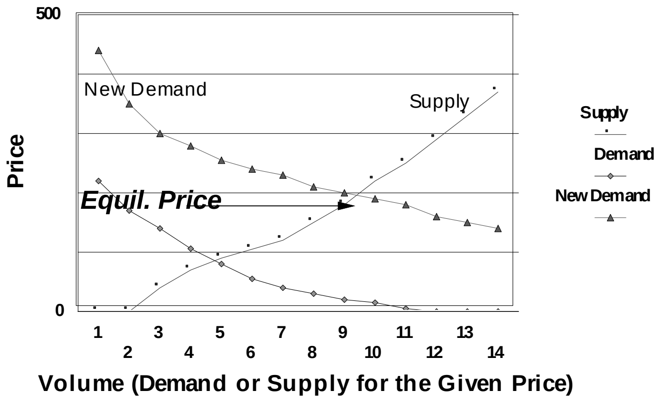
Over time, factors that affect both supply and demand are subject to change. First consider changes in supply. Investing in new technologies or in more efficient production facilities or simply adopting better management techniques may make it possible for suppliers to offer greater quantities for any given price. Graphically, this results in a shift in the supply curve to the right and leads to a new – lower – equilibrium price, as shown in Figure 2. Note that the demand curve has not changed at all: with the lower prices, people are willing to buy more, which is what is described by the demand curve.

offer
equilib



The demand curve may also change (Figure 3). For example, growth in population or increases in family income may result in an increase in cars purchased, attendance at movie theatres, or use of air transportation. These changes appear on the graph as an upward shift in the demand curve: at each price level, a greater quantity of goods and services is purchased or used.

Population Growth, Advertising, Higher Incomes, or other Factors May Cause an Increase in Demand (a shift in the demand curve)



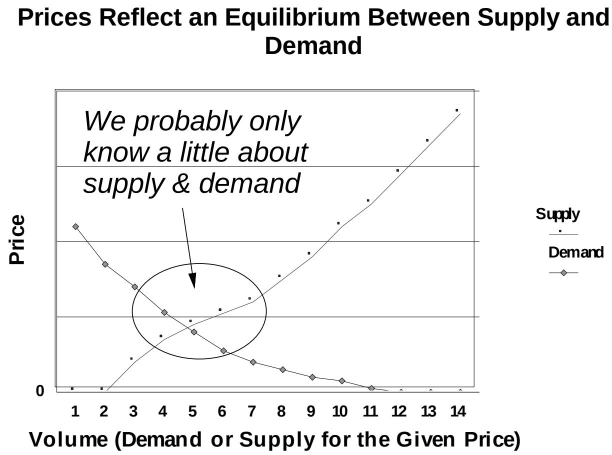
If demand increases, then prices will rise; if demand declines, then prices will fall. How much prices rise or fall will depend upon the shapes of the supply and demand curves. How quickly prices rise or fall will also depend upon the nature of the goods and services being sold. Outside a sold-out baseball stadium, the prices that scalpers charge for tickets will react within minutes to changes in demand. Achieving an equilibrium in the prices of new homes is something that may take years, as evidenced by the steady decline in home prices that began in 2007 and continued for several years thereafter because of what has been called "The Great Recession." Adjusting transportation networks to changes in oil prices or new technologies is a process that takes decades – and may never reach equilibrium, because only a small portion of the transportation network can ever be changed within just a few years.

In many circumstances, changes in demand result not in a change in price, but in poor service, congestion or long lines as too many people try to buy something or to use something at the same time. The time spent in line can be viewed as part of the price of the service that is being sold: some people will come with an intent to buy, but depart as soon as they see the line.

Over time, suppliers will react to changes in demand by adjusting their levels of production. New companies may emerge in response to increases in demand; companies may go out of business in response to decreases in demand.

It is possible to spend a lot of time trying to understand the supply and demand curves, and there are some ingenious methods for estimating these curves based upon past experience. However, it is important to retain some humility, for we probably only know a little about how supply and demand vary within a fairly small range of prices and existing conditions (Figure 4). When new projects are being considered, it is possible that the quantity or quality of services provided will be far different than what is currently available. Special studies can be undertaken to try to estimate the effects of the new projects on demand, but such studies will never be exact.

Figure 4 The int… rstood, especially if the fac… long time



<u>Consumer Surplus</u>

Given the shape of the demand curve, it is clear that the equilibrium price is lower than the price that many would be willing to pay. The difference between what someone is willing to pay and the equilibrium price is called **consumer surplus**. For each individual:

(Eq. 1) Consumer Surplus = (Willingness-to-pay) – (Equilibrium Price)

Consumer surplus is greatest for those willing to pay the highest prices. For someone willing to pay only the actual price and not a penny more, the consumer surplus is zero. The total consumer surplus in principle could be obtained by summing the surpluses for everyone using a product or service. In practice, this is infeasible, as data is collected and decisions are made based upon actual prices. Unless special studies are undertaken, little is known about how much more people would be willing to pay for things that they now buy or for services that they now use. For this reason, it is easier to focus on the changes in consumer surplus that may result from changes in equilibrium prices.

Consider the change in supply illustrated above in Figure 2. The shift in the supply curve increased consumer surplus by a) lowering the price for those who previously were willing to pay a higher price and b) allowing more people to purchase the product. The increase in consumer surplus can be estimated just by looking at prices and volumes before and after the change in supply:

(Eq. 2) Increase in Consumer Surplus = $V_0 (P_0 – P_1) + ½ (V_1 – V_0) (P_0 – P_1)$

The first term in this equation is the benefit to existing users from the reduction in price, while the second term represents the benefits gained by new users. The full decline in price is not a benefit for new users, since they were unwilling to pay the old price. If the relevant portion of the demand curve is assumed to be a straight line, then the consumer surplus for the new users will be the area of a triangle whose base B is the difference in volume and whose height H is the difference in price, and whose area is ½ BH. In effect, this assumption – sometimes called the **"rule of ½"** - provides a simple way to estimate the area under the demand curve without needing to estimate an equation for that curve.

Note that consumer surplus is an economic rather than a financial concept. Price is a financial measure, as is manufacturing cost; these are things that can be measured in dollars and cents and these are things that can and will be recorded in check books and accounting systems. Consumer surplus is not related to any such accounting, but it is still an important matter for evaluating the economic impact of projects. There is a public benefit resulting from projects that increase consumer surplus, because people will still have the money that they otherwise would have been willing to pay for the product or the service. They can save that money or use it to buy something else. Either way, there is a benefit for the individuals and for the local economy. Thus, change in consumer surplus, though not a direct concern for the private sector or for investors, is an important consideration in evaluating the public economic impacts of proposed infrastructure projects.

<u>Elasticity of Demand</u>

It is possible, but difficult, to obtain the detailed information needed to plot supply and demand, so a more abstract approach is often used. Consultants or marketing managers may use past experience in trying to answers questions such as "How much will demand change for a given change in price?" or "Will total revenue go up, down, or stay the same if the price is changed?"

These questions can be answered by using the concept of **elasticity of demand**, which is a measure of how sensitive demand is to changes in price. Elasticity of demand is defined as the negative of the derivative of the quantity demanded Q with respect to price P.

(Eq. 3) Elasticity of demand = - dQ/dP

This measure may also be referred to as **price elasticity**. The minus sign in this equation is conventionally used because the quantity demanded is expected to vary inversely with the price that is charged. Elasticity of demand can also be estimated by looking at the change in demand that occurs after a change in price:

(Eq. 4) Elasticity of demand = $-((Q_1-Q_0)/Q_0)/((P_1-P_0)/P_0) = -((Q_1-Q_0)/(P_1-P_0))(P_0/Q_0)$

In this equation, the changes in quantity and price are both normalized by dividing by their values before the price change. The equation therefore can be interpreted as the percentage change in quantity divided by the percentage change in price.

To understand the importance of price elasticity, consider the two effects of a price decrease from P_0 to P_1 on total revenue PQ. Existing customers will pay less, because the price is lowered, and there will some loss of revenue:

(Eq. 5) Reduced revenue from original customers = $Q_0 (P_0-P_1)$

However, the lower price will attract new customers, so the quantity demanded will increase from Q_0 to Q_1, providing some additional revenue:

(Eq. 6) Additional revenue from new customers = $P_1 (Q_1 – Q_0)$

If elasticity is greater is greater than one, then the percentage increase in Q will be greater than the percentage decrease in P, and the added revenue from Eq. 6 will be greater than the loss of revenue from Eq. 5. If this is the case, demand is said to be "elastic", because there is a large response to changes in price. If elasticity is less than one, then the opposite is true: total revenue will decrease if prices are lowered, as the added revenue from new customers will be insufficient to offset the loss of revenue from existing customers. When elasticity of demand equals one, there will be no change in total revenue PQ, as the effect the change in price will be exactly offset by the change in demand.

Price elasticity is an important factor in infrastructure systems, because these systems tend to have high fixed costs and low variable costs. Maximizing revenue may therefore seem to be a reasonable goal, since the greatest obstacle to making a profit is having enough revenue to cover the fixed costs of the system. During the early portion of the 21st century, tremendous investments in satellite-based communications were justified in part upon the expectation that creating a very high-capacity system with very low prices would lead to extraordinary increases in demand – which is exactly what happened as technological advances lowered the costs of email, cell phones, and wireless access to the internet.

In the short run, demand tends to be more inelastic than in the long run. For example, when the price of oil rose dramatically in 2007 and 2008, people initially had to pay the higher price and perhaps cut back on non-essential driving. Over a period of a year or two, however, people were able to adjust in part by buying more fuel efficient cars and in part by figuring out how to combine errands, share rides, and use public transportation. Over a period of a decade or longer, the automobile companies can develop cars that use alternative energy sources, allowing people to drive more while using less oil.

Elasticity of demand is an important concept to keep in mind when evaluating infrastructure projects, because demand forecasts will drive decisions related to the size and therefore the cost and capacity of infrastructure. Forecasts based upon continuation of low prices or free access will lead to extravagant statements of infrastructure needs. Such has long been the case with urban road networks: with the exception of a few toll roads, there is no charge for using highways, and there should be no surprise that these roads have become so congested. Where tolls have been introduced, as in London and Singapore, it has proved possible to reduce traffic volumes and thereby limit congestion to reasonable levels. Water supply is another area where unrealistically low prices have in many locations led to unnecessarily high rates of consumption; future "needs" for water should take into consideration the effect of more rational pricing on consumption.

Based upon the concepts of equilibrium prices and demand elasticity, it is apparent that future demand for infrastructure will depend to a greater or lesser extent upon the prices that are charged. If demand is elastic, then pricing could have a dramatic effect on demand, and raising prices could be viewed as a way to reduce or avoid investments that increase infrastructure capacity. If demand is highly inelastic, then pricing will probably not be an effective means of limiting demand, and failure to expand capacity could lead to extremely high equilibrium prices, extremely poor service, or a need for regulating use or access. Elasticity of demand is therefore an important factor both in pricing infrastructure services and in forecasting demand for infrastructure. The next section continues the discussion of pricing in the context of the degree of competition among suppliers.

Pricing

This section introduces two markedly different pricing regimes: competitive markets and monopolistic pricing. In a competitive market, there are many suppliers and many potential customers, none of whom have the power to set prices. Instead, as described above, prices reflect an equilibrium between supply and demand. Not all markets are competitive, and it is possible that geography, politics, or economic factors encourage the development of companies or agencies that have a monopoly for particular goods or services. In the absence of regulation, a monopoly can set prices and customers have little power. Monopolies are not necessarily evil, because there are many situations where a single large supplier can produce goods or services at the lowest possible cost. Moreover, the danger of monopolistic pricing can be controlled by government regulation, so that the benefits of low cost production are passed on to society and not simply captured as excessive profits by the owners of the monopoly.

<u>Marginal Cost Pricing in a Competitive Environment</u>

In a competitive environment, prices will fall to marginal cost. A competitive environment is one in which many suppliers all have access to the same or similar technologies, and they are serving customers who are able to purchase goods or services from any of the suppliers. Under these conditions, a supplier who tries to raise prices above marginal costs will have a problem: another supplier will be willing to offer a slightly lower price and thereby capture the business. So long as the price is above the marginal cost, each sale will give the supplier some **contribution to overhead and profit**. Under perfect competition, no supplier has any pricing power, and prices are determined by the cost structure, the available technologies, and the level of demand.

Marginal cost pricing is efficient in the sense that prices reflect the actual cost of the product or the service. All of those who purchase the product or service are in fact willing to pay – and do pay – the marginal cost of production. While others may desire the product or service, they are unwilling to pay enough to make it worthwhile to any of the suppliers. Any supplier who can provide the product or service at a lower price is free to enter the market and make a profit by selling at or somewhat less than the prevailing price.

Situations where prices differ from marginal cost are likely to be inefficient in economic terms. If prices are too low, then demand will be too high and some users will be incurring costs that they would not be willing to pay for. If prices are too high, then many who would have been willing to pay a reasonable amount for the service will be unable to afford to make a purchase.

A major difficulty with marginal cost pricing arises in situations where marginal costs drop below average costs, as is commonly the case with infrastructure systems. In such situations, marginal cost pricing will not provide sufficient revenue to cover costs, and all suppliers will face bankruptcy. Technological advances and increasingly efficient production may help some suppliers stave off bankruptcy, but only the most efficient suppliers will be able to survive. In these situations, some kind of government regulation or subsidies may be needed to enable suppliers to remain in business. Regulation could take the form of limiting entry into the market or establishing prices at a level that allows suppliers to make a profit. Examples of governmental actions that limit entry include the following:

- Issuing taxi medallions in an attempt to limit the number of taxis to what will be efficiently utilized.
- Requiring railroads to seek regulatory approval before constructing new lines.
- Creating public utilities for communications or energy services.

Generally, when entry is limited, prices must be regulated to ensure that prices are reasonable.

Economies of Scale, Scope, and Density

A competitive market requires multiple suppliers who are free to determine whether or not to enter the market based upon the prevailing prices. If there are many companies, and it is easy to enter and exit the market, then supply and demand can quickly approach an equilibrium. However, the equilibrating process will be hampered if there are barriers to entry, such as the need to make large investments in order to be able to compete. For infrastructure-based systems, this is certainly an issue, as these systems by definition require substantial investments, and it will take time and effort to construct a competing system. Moreover, there are very likely to be economies in creating large facilities that can serve multiple purposes for many different users. Having competition among a great many – or even a few – smaller companies may be less efficient that having a single supplier. Larger systems may have three types of advantages over smaller systems: economies of scale, scope, and density.

Economies of scale exist when an increase in the size of the system results in reductions in cost. If C(Q) is the total cost of providing infrastructure adequate for usage Q, then there are economies of scale if:

(Eq. 7) $$C(Q_1 + Q_2) < C(Q_1) + C(Q_2)$$

For transportation, water resources, electric power grids, and other network-based systems, there will often be economies of scale because:

- A single management team can manage a larger system using the same basic information technology.
- The same advertising can be used for a wider audience.
- A larger network allows a company to provide single-company service to more customers, and direct service may be cheaper than service that requires cooperation among multiple suppliers.
- A larger network provides direct links between more locations, which in transportation or communication systems can be a major benefit for potential customers.
- Consolidated maintenance facilities can serve a wider area.
- The costs of energy and materials can be reduced because a larger company can negotiate lower prices from suppliers.

Economies of scope exist when it is more efficient to use facilities for two or more types of service than it is to use them for a single service. If $C(Q_i,S_i)$ is the cost of serving Q_i customers of type S_i and $C(Q_1,S_1,Q_2,S_2)$ is the cost of serving two groups of customers, then there are economies of scope if:

(Eq. 8) $$C(Q_1,S_1,Q_2,S_2) < C(Q_1,S_1) + C(Q_2,S_2)$$

A situation where there are clearly economies of scope would include highways, which serve commuters and intercity travelers moving in automobiles or buses along with local trucking and intercity trucking. Another situation would be a dam that is constructed for flood control that also can be used to generate electricity and support irrigation. If there are potential economies of scope, then there will be advantages to society from building joint facilities.

On the other hand, there are situations where it does not make sense to have a single facility for multiple services. Because of potential safety problems, pedestrians and cyclists are not allowed on high-speed, limited access highways. High-speed passenger trains cannot operate on tracks designed for freight trains, because high-speed trains cannot be safely operated on routes with sharp curves and frequent grade crossings. Swimming is not allowed in reservoirs, because of possible public health problems.

Economies of density exist when average costs decline as a result of adding more volume to an existing system:

(Eq. 9) $$C(Q_1 + Q_2) / (Q_1 + Q_2) < C(Q_1)/ Q_1$$

If applied to a single facility, economies of density would exist wherever scale economies exist. The distinction, however, between economies of scale and economies of density is very critical in transportation and other networks where there are many facilities and an extensive route structure. Costs in these networks relate to both the links and nodes of the system, and there are two major strategies for capturing more business, namely expanding the network or adding more volume to the existing network. In network systems, economies of scale refer to situations in which the network expands in proportion to the increase in demand, whereas economies of density refer to the effects of adding more traffic to existing facilities.

Even if there are no economies of scale, there could be strong economies of density. Much of the investment in transportation systems has been attempting to capture economies of density, e.g. by concentrating more cars on existing roads and more flights at existing airports. Most transport networks have strong economies of density up to the point where added traffic causes extreme congestion.

The distinction between economies of scale and economies of density can also be seen in restaurants and retail sales. Large retail outlets, such as Staples or Home Depot, are able to achieve lower fixed costs per unit of sales by having very large efficient buildings with managers and employees who can be more specialized and also more productive than they would be at smaller stores; with a larger work force, it is easier to adjust up or down for peak periods or slack periods. Big box retail stores therefore capture economies of density.

Fast food outlets such as McDonalds and Burger King, which have thousands of restaurants all over the country and around the world, are able to achieve scale economies. While they have some facilities that are larger than others, they have vast numbers of similar restaurants that benefit from brand recognition, common procurement, common design, and standard management. These restaurants have lower unit costs than the individual restaurants and smaller chains that they compete with, and they use extensive marketing to convince us that their food is not just cheaper, but just as tasty. People know what to expect when they walk into one of these restaurants, and therefore people are likely to go to these restaurants not just when they are near home, but also when traveling or vacationing in another state or another country. These companies clearly profit from scale economies.

Monopoly Pricing

Where there are possibilities for economies of scale, scope, and density, there can be strong forces leading to supply-side consolidation. The motivation initially is to save costs or to expand markets, but if competition is reduced, then it could be that a single company achieves monopoly pricing power. If demand is inelastic, this can lead to extremely high prices, not to mention extremely high profits.

If there are economies related to size, a larger company can always underprice smaller companies and still make a profit. Hence, they can drive competitors out of business, then, when no one else is left, they may have the opportunity to raise prices so as to maximize their profits. Naturally, the public and public officials are against monopolistic pricing, but they also are likely to favor lower prices. Therefore, in situations where there are strong economies of scale, public agencies will often allow one or a few companies to exist, but regulate their prices and perhaps their services. Examples of publicly owned or regulated monopolies include most transit systems in the United States, most agencies that provide water and sewage treatment, and most public utilities.

Productivity

Productivity, a classical consideration in economics, is defined as output divided by input. Productivity can be increased either by increasing the outputs obtained from the same level of inputs or by reducing the inputs required to obtain the same level of output. Improving productivity allows a company, an industry or a society to produce more and/or to consume fewer resources. Improving productivity therefore is generally viewed by most everyone as an important goal. Officials in the private sector believe that productivity improvements will lead to higher profits, while those in the public sector believe that higher productivity will lead to higher income for workers, lower prices for consumers, and better opportunities for growth in the economy. Achieving higher productivity motivates many infrastructure projects.

Measuring productivity is complicated by the fact that there are usually multiple types of outputs and inputs. Thus, to get the ratio of outputs to inputs it is necessary to have some kind of weighting system for measuring both outputs and inputs. For example, in looking at the productivity of the air transportation system, output cannot simply be measured as the number of passengers or the number of passenger-miles; it is necessary to consider the differences among short- and long-distance flights for business and non-business travelers. With water supply systems, it is necessary to distinguish among water supplies for residential, industrial, and agricultural users. With electricity companies, it is necessary to consider peak loads vs. off-peak loads. In all of these systems, inputs will include many factors that can be summarized under broad headings of labor, capital, land, materials, and energy.

One simplifying approach is just to consider a single measure of output and a single factor of production, leading to measures such as the total number of air passenger trips per airline employee or the amount of electricity generated per unit of investment. These single-factor measures tend to be somewhat arbitrary, as production really does require multiple factors, but such measures may be reasonable for measuring productivity changes for a single, fairly stable operation.

Another possible approach is to weigh outputs by their prices and to weigh inputs by their costs, perhaps using information from a base year for both prices and cost. If this is done, then the aggregate measure of output will be something close to "total revenue" and the aggregate measure of input will be something close to "total cost", and the ratio of output to input will be the ratio of revenue to cost. A variant of this approach is to assume that prices reflect marginal cost, in which case price can be assumed to be a measure of cost. Since price information is more readily available than cost information, this can be a useful assumption. If there is a meaningful measure of output, then the inverse of the price per unit may be a reasonable measure of productivity. Over time, changes in the price per unit can therefore be viewed as an indicator of changes in productivity.

For example, tremendous gains in productivity have been achieved in freight transportation over the past 200 years. Table 1 shows three factors that contribute to freight productivity: the cost per mile for construction, the tons carried per vehicle, and the miles traveled per day. The most commonly used measure of output in freight transportation is the ton-mile, which is one ton carried one mile. Typical costs and prices are shown in this table for two periods, the early 19th century and the early 21st century. The costs are current costs, unadjusted for inflation.

Before 1800, there were only two types of freight transportation that were generally available: wagons moving over bad roads and somewhat larger wagons moving over improved roads, which were usually limited to a few turnpikes radiating out from major cities. At that time, a turnpike was often just a dirt road that was maintained to allow slightly heavier vehicles to travel the straightest possible path between two towns. When using the turnpike, it was possible to use a larger wagon and to go a little faster simply because the road was smoother and a bit wider. Even with the turnpikes, transport was slow and expensive, and typical prices exceeded 15 cents per ton-mile.

Table 1 Increasing the Productivity of Freight Transportation

Technology	Cost per Mile to Construct	Tons/vehicle	Miles/day	Ton-miles per vehicle day	Typical Prices (cents/ ton-mile
Early 19th Century					
Rough road	$1-2,000	1 per wagon	12	12	20 to 40
Turnpike	$5-10,000	1.5 per wagon	18	27	15 to 20
Canal	>$20,000	10 to 100 per canal boat	20 to 30	200 to 3,000	5 to 10
Railroad	$15-50,000	500 per train	200	100,000 per train	3 to 5
Early 21st Century					
Arterial roads	$1-5 million	10 per truck	100	1000	10 to 50
Interstate Highway	$5-100 million	20 per trailer	500	10,000	10 to 15
Heavy-haul railroad	$1-5 million	5-15,000 per train	500	5 million per train	2
Inland waterway	Highly variable	1500 per barge; up to 40 barges per tow	50-200	6 million per tow	1

Rivers and canals allowed larger loads and longer daily hauls, as it is much easier to pull a canal boat along a river than to drive a horse and wagon up and down the hills. In the early 19th century, canals were built for upwards of $20,000 per mile; the expense was justified by the increased productivity for the freight carriers, and typical freight rates fell below ten cents per ton-mile. Canals were limited by geography, so rail technology had a great advantage as soon as it became available. Even with only 500 tons per train, a railroad allowed much more productive freight operations than was possible with small canal boats, and typical freight prices dropped below five cents per ton-mile. Technological improvements continued throughout the 19th and 20th centuries, so that today, the prices for freight transportation are actually lower than they were 200 years ago. The lowest prices are achieved for fully loaded vehicles traveling at the maximum speed on the main routes – tractor/trailer combination trucks on the Interstate Highway, heavy coal trains on high density, well-maintained rail lines (Figure 5), and tows of 40 barges moving along the major rivers.

This example used the ton-mile as a simple measure of output for freight transportation, even though the costs and benefits of transporting different commodities different distances can vary widely. For example, it is easier to move coal in single shipments of 10,000 tons than it is to move 10,000 tons of general merchandise as 200 separate shipments. Nevertheless, even though the ton-mile is far from a perfect measure of rail output, the cost/ton-mile remains useful in highlighting the dramatic productivity improvements that have been achieved in freight transportation.

Perceived productivity problems often suggest the types of projects that need to be undertaken:

- Peak demands may cause delays at bottlenecks in transportation or systems (so consider investing to relieve bottlenecks).
- Engineering constraints, such as weight limits on bridges or band-width limitations in communications networks, may restrict the usage of the system (so consider investing to increase the ability of the infrastructure to handle larger or heavier loads or higher volumes of usage).
- Lack of communication and control may inhibit efficient use of resources (so consider investing in communications and control systems).
- Facilities that were designed and built many years ago may no longer match what is needed today or in the future (so consider rehabilitating or expanding or redesigning facilities or networks).

Figure 5: Coal trains carrying up to 15,000 tons of coal operate on narrow rights of way through difficult territory like this canyon in Colorado. A double-track route like this can handle more than 100 million tons of freight per year – plus a pair of 79 mph passenger trains. The interstate highway is visible on the other side of the river. To minimize its footprint, it was double-decked for a portion of the route. There is also a bike path beneath the highway.

Photo: S.J. Martland 2000.

Measuring and Improving the Economy

Infrastructure projects have impacts that go well beyond the financial affairs of owners and users. Infrastructure allows and supports other economic activity, and the greatest benefits of investment in infrastructure may be the new opportunities made available to society. In evaluating large infrastructure projects, two types of economic impacts are commonly considered:

- The short-term boost to the local economy resulting from the planning and construction of the project.
- The long-term impact of the project on the region's economy once the project is completed, including the benefits to users of the new infrastructure as well as the permanent jobs directly linked to the project.

Constructing a light-rail line to the airport in order to relieve highway congestion and improve access does much more than provide jobs for operating the trains and maintaining the tracks; it also saves time for air travelers, airport employees, commuters who use the new line, and highway commuters who experience less congestion, as well as creating opportunities for developing real estate near the light-rail stations.

Short Term Economic Impacts: the Multiplier Effect of New Investment

The design and construction of a major project boosts the economy, because of what is known as the multiplier effect. Consider the construction of a new building in a city. Much of the investment cost will be made up of wages and salaries paid to local construction workers and payments to local merchants for materials and services. These workers may save some of their wages, but they are likely to spend most of it; likewise, the local suppliers will spend most of what they receive. The proportion of the new income that they consume is called the **marginal propensity to consume**.

Let's say that the construction of the building resulted in wage payments of $1 million to local workers and companies. This $1 million will in itself be an addition to the regional economy, but that is just part of the story. If the marginal propensity to consume is 0.5, then workers and companies will spend another $0.5 million – which will be another addition to the local economy. And that $0.5 million will go to other workers and companies who will save some and spend some. If they also save half and spend the rest, then there will be another $0.25 million added to the regional economy. And some of that money will also be spent. If half of the money is saved at each step, then the total addition to the regional economy will be $1 million (1 + 0.5 +0.25 + 0.125 + …) which will converge to $2 million dollars. In this case, the multiplier is 2, as each dollar invested leads to an increase of $2 in the regional economy.

In general, the total addition to the economy can be expressed as a function of the marginal propensity to consume MPC:

(Eq. 10) Addition to economy = Investment in Region $(1 + \text{MPC} + \text{MPC}^2 + \text{MPC}^3 + \ldots)$

So long as MPC is less than 1, this sequence converges to 1/(1-MPC). The factor (1-MPC) is the marginal propensity to save, so the multiplier effect increases inversely with the marginal propensity to save:

(Eq. 11) Multiplier Effect = 1/(1-MPC)

For example, if the marginal propensity to consume increases from 0.5 to 0.75, then the marginal propensity to save drops to 0.25. If so, then more money goes into the economy. The total addition to the regional economy would be:

(Eq. 12) $1 million (1+.75+.75(.75)+(.75)(.75)(.75) …) = $1million/(.25) = $4 million

With less money going into savings, the multiplier effect jumps from 2 to 4.

The multiplier effect would apply both to the construction phase and to the operations phase of a project. For infrastructure projects, however, since investment costs are so much higher than continuing costs, the greatest interest is in the multiplier effect from the investment. Multipliers are typically found to be between 2 and 3 for transportation and other infrastructure projects. Note that the multiplier effect relates only to the money spent within the region, so that a project that imported costly materials and used highly automated equipment would have a much lower regional impact than a more labor intensive project that used local labor and materials.

The presence of a multiplier effect motivates governments to initiate stimulus programs during a recession. In the short-run, the stimulus will be most effective in reviving the economy if it is directed toward projects and programs that direct money toward people who will be likely to spend most of what they receive. The long-run economic benefit will depend upon the success of the project in providing permanent jobs, making society more productive, or enabling other economic benefits to society.

Long-Term Economic Impacts: Gross Regional Product

The most common economic measure used to monitor the health of the economy is the **gross domestic product (GDP)**, which equals the sum of private consumption C, investment I, and government expenditures on goods and services G plus exports E minus imports M:

(Eq. 13) $$GDP = C + I + G + E - M$$

The growth in the economy is measured as the change in GDP, and growth in GDP is generally viewed as a critical objective for a nation. A growing economy provides opportunities for more jobs, higher wages, provision of more goods and services, and higher profits for companies. If GDP declines for two successive quarters, then the economy is said to be in a recession. Unemployment rises during recessions, wages may fall, and company profits decline. Thus, maintaining GDP is an important economic objective for a nation.

GDP is not a perfect measure, in part because there is more to life than economics. Even in the realm of economics, however, there is a major problem with GDP, because it fails to account for the losses associated with the depreciation of the capital stock of the country. Machines wear out, buildings age, infrastructure deteriorates, and these losses from depreciation will not be captured until and unless repairs are made or facilities are replaced. The net domestic product is calculated by subtracting total depreciation from GDP. Net domestic product is less commonly used because it is difficult to estimate depreciation of assets, while it is relatively easy to monitor consumer purchases, investments, government expenditures and foreign trade. Since the two measures will usually rise and fall in tandem, the GDP figure is what is most frequently used.

GDP is an aggregate measure that will not reflect conditions for particular regions, groups of people or sectors of the economy. However, similar measures can be estimated for each region of a country. The **gross regional product (GRP)** would be defined in the same manner, with the various factors defined so as to apply to the region, not to the nation. As with the national economy, growth in GRP will be a major economic objective for any region.

Jobs and average income are other important aspects of the regional economy. Adding jobs to the regional economy is always viewed positively, but especially so during a period of high unemployment. Higher-paying jobs are preferred, and local governments may provide tax breaks and other incentives to attract or to retain companies that have such jobs.

Economic models can be constructed to predict the impact of infrastructure investment on the regional economy. Such models may be able to show that transportation investments will make the region more attractive to new businesses or that investments in dams and irrigation will make local agriculture more profitable, leading to growth in all activities related to agriculture. Analysis may also show that investment in infrastructure is expected to have a measurable impact on congestion, public safety, or public health. Savings in time, reductions in risk, and improvements in health can be translated into economic benefits by using the average value of time for commuters, the expected savings in accident costs, and the expected reduction in health care costs. While such benefits do not result in cash flows that help pay for infrastructure investments, they are quantifiable factors that can help justify (and gain public approval for) public investments.

Trade

A great deal of infrastructure investment is based upon projections for population growth and growth in regional economies. Over the long-term, both types of growth depend to a large extent upon forces that act on a national or international scale, such as technological change and trade. **Technological change** results in new products, new materials, new development opportunities, and new processes for manufacturing and distribution. Over time, there can be marked changes in what types of things are produced, how and where they are produced, and how they are distributed around the world. These changes influence and respond to changes in **economic geography**, i.e. the location of people and economic activities throughout a region, a nation, or the world. And it is these changes in economic geography that require and motivate many investments in infrastructure.

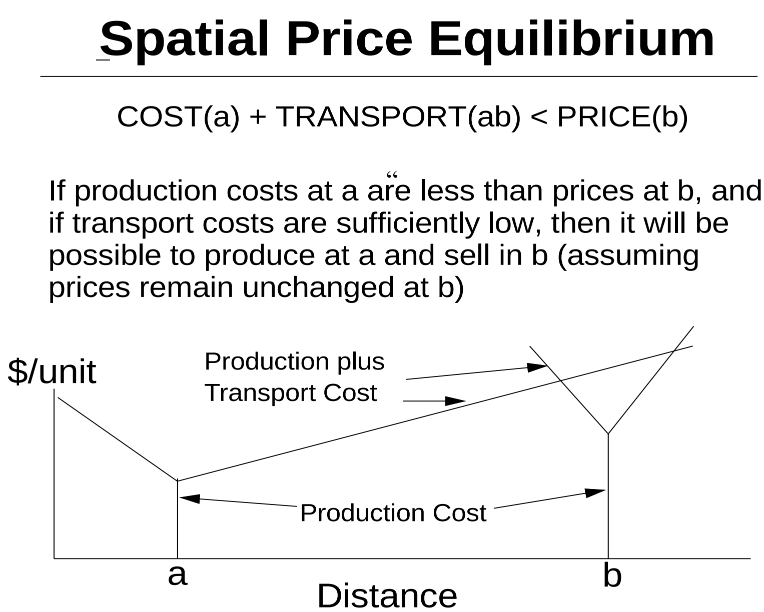
Trade is the exchange of goods among regions or countries. Transportation makes trade possible. Differences in regional resources, economies of scale in production, and differences in costs and capabilities make trade desirable. The ability to exchange currencies of different countries and to transfer monies between countries makes trade financially feasible. The ability of wealthy countries to buy vast amounts of goods and materials makes trade flourish.

Two key concepts are helpful in understanding how trade works and why trade is important. First, if it is possible to produce something for a lower cost in one region than in another, then there is an opportunity for trade, but only if transport costs are sufficiently low. **Spatial price equilibrium** is the process by which transport costs and manufacturing costs together determine prices for products that can be produced in one area and sold in another. Second, it makes sense economically for different regions to exploit their **comparative advantage**, i.e. to concentrate on what they each do best. Understanding these concepts is essential to understanding the global economy, and recognition that there is a global economy is essential for understanding the kinds of infrastructure investment that will be needed to support industrial production, trade, and population growth.

Spatial Price Equilibrium

To begin, consider a product that can be made in two locations. At location “a”, production costs are lower than at location “b”. T... ough costs are lower at locati... st of all, costs are not prices. ... ey will share whatever mark... costs must be added to the m...

Figure 6 illustr... s of this chart represents dist... ”. The y-axis represents the ... indicated by the vertical line...



Assume that transport costs T_{ax} from Manufacturer A to any potential location X are proportional to the distance D_{ax} from the manufacturer:

(Eq. 14) $$T_{ax} = k_{ax}D_{ax}$$

Thus, if Manufacturer A decides to market its product at location X, then the total cost will be:

(Eq. 15) $$C_a = M_a + T_{ax} = M_a + k_{ax}D_{ax}$$

A similar equation would apply for Manufacturer B:

(Eq. 16) $C_b = M_b + k_{bx}D_b$

The total costs C_a and C_b are both plotted in Figure 6. The costs for each manufacturer are Y-shaped, as total costs rise linearly by moving in either direction in this two-dimensional figure. The slope of the total cost line is the cost per mile for transporting the product, which is k_{ax} for Manufacturer A and k_{bx} for Manufacturer B. The slope is steeper for Manufacturer B, indicating that transportation costs are more expensive for that manufacturer. Manufacturer A will presumably capture the market for all locations where its total costs are less than those of Manufacturer B. At points where they are equal, the manufacturers will share the market. In the two-dimensional world depicted above in Figure 6, the total cost lines intersect in two places, one just to the left of Manufacturer b's location on the graph and one a little further to the right of that location. This indicates that Manufacturer B will only be serving its local market, while Manufacturer A will capture all of the rest.

In general, if two manufacturers of identical products are competitive in a market, then they must be charging the same price at that location. For a manufacture to compete and make a profit, the sum of their production cost plus transportation cost must be less than or equal to the prevailing price at that location. Prices will vary from one location to another, reflecting the differences in production and transportation cost. Over time, if transportation and production technologies are stable, a spatial price equilibrium will be achieved, and the markets served by each manufacturer will be well-defined.

Of course, neither production nor transportation technologies remain stable for very long. As we have already seen in this chapter, transportation costs have declined dramatically over the past 200 years, enabling today's manufacturers to compete globally. With low cost global distribution feasible, it makes sense a) to have large manufacturing facilities that take advantage of whatever economies of scale can be found in manufacturing and b) to have those manufacturers located in regions where costs are lowest, whether because of local labor rates or the local energy costs or the geographical position relative to sources of inputs and major markets. Because of the tremendous improvements in freight transportation productivity, it is now feasible to manufacturer many consumer goods in Asia, where labor and production costs are very low, and to ship those goods on large container ships to major ports for distribution throughout the rest of the world.

Comparative Advantage

The unequal distribution of resources, including capital and skilled labor as well as natural resources, is another force promoting trade. Because of accidents of location or history, one region may be able to produce certain products at a lower cost or higher quality than other regions. If this region makes an excess amount of such products, then it can sell them to other regions and use the proceeds to purchase other types of goods from those regions. For example, one country may be very good at making automobiles, while another country is very good at agriculture. Opportunities for trade would seem pretty clear: trade automobiles for food, and allow the country to concentrate on the products where they are the best.

The potential benefits of specialization were first highlighted nearly two hundred years ago by David Ricardo, who developed the theory of **comparative advantage** in 1817. Samuelson summarized this theory as follows:

> "Whether or not one of two regions is absolutely more efficient in the production of every good than is the other, if each specializes in the products in which it has a *comparative advantage* (greatest *relative* efficiency), trade will be mutually profitable to both regions. Real wages of productive workers will rise in both places."[2]

[2] Paul Samuelson, **Economics: An Introductory Analysis**, McGraw-Hill Book Company, New York, 1964, 6th edition, p. 665.

This theory is the basis for reducing tariffs and other barriers to trade. Tariffs are taxes that are charged on inputs as a means of protecting local manufacturing. However, the theory of comparative advantage indicates that it is better to allow imports, so that local workers and local capital can be put to work more productively in areas where the region enjoys its greatest efficiency relative to other regions.

Currency Exchange and International Banking

International trade depends upon a banking system that is able to do two things. First, there must be a common medium of exchange, so that the money used in one country can be used or converted into an amount of another country's money that has the same value. Second, it is necessary to have some system of credit so that a person in one country can borrow the funds needed to buy goods that will be exported from a second country and perhaps transported and sold in a third country.

Today, when it is easy to stick a credit card into an ATM to get local currency when traveling abroad, it is difficult to imagine how important – and how difficult - it was for traders to have access to a banking system in order to carry out their business without personally carrying vast sums wherever they went.

The exchange rates between currencies can be based upon market forces or regulatory forces. Major newspapers provide daily reports on the exchange rates for the major currencies. In early 2009, for example, $1.00 was worth about 0.77 euros or, to put it the other way around, one euro was worth about $1.30.

At times, exchange rates will be quite volatile. In September 2008, when I traveled to Ireland, the euro was worth $1.60. While I was in Europe, there was a worldwide credit crisis, major banks appeared in imminent danger of collapse, and the value of the euro dropped to about $1.35 by the end of my vacation. Although the prices of our hotels and our meals remained unchanged – in euros – it appeared to us as though everything was now 15% off! This change made our trip a little bit cheaper, but this same change affected every transaction between anyone in the U.S. doing business with anyone in Europe. Suddenly, everything priced in euros was 15% cheaper for anyone who had dollars to spend – and everything priced in dollars was 15% more expensive for anyone who had euros to spend. A change of this magnitude is equivalent to putting a 15% tariff on everything exported from the US into Europe and having a 15% sale on everything imported to the US from Europe. Changes of this magnitude have broad repercussions on international trade and travel, even without a credit crisis.

A credit crisis can be devastating to trade and economic growth. Without credit, it is hard for businesses to get the loans they need to expand production and it is hard for consumers to borrow money to buy houses, cars and other items. The credit crisis in late 2008 and early 2009 resulted in stock markets plunging, the auto industry teetering upon bankruptcy, and many banks and investment banks collapsing. Without credit, trade declined abruptly and the world economy slipped into a serious recession.

Making Decisions: Utility and Sunk Costs

Economists use the concept of **utility** as a way to understand how individuals make decisions. It is assumed that people act so as to maximize their utility, subject to budgets for both time and money. Utility is a useful concept, even though few of us will be able to say why we do or do not do something or why we prefer one product over another. It is possible to study utility by documenting the choices that people make or by conducting surveys, i.e. by considering what are called **observed preferences** or **stated preferences**. Analysis of actual choices is likely to provide better insight into behavior, but it is much easier to obtain detailed information by using surveys and documenting stated preferences. For example, a survey could ask people to say whether or not they would buy a particular product for various prices. More complex surveys could be devised to explore quality of service, timing, and other factors that might be important in addition to price.

For example, consider the journey to work. A person may have the choice of driving alone, riding with a friend, or taking the bus. Direct observation may show that this person drives alone 60% of the time, rides with a friend 20%

of the time, and takes the bus 20% of the time. An interviewer can probe further, seeking to understand how cost, travel time, work schedules, weather, errands, and shopping factor into the decision. Utility models may be constructed based upon the actual decisions or the stated preferences. These models will typically include variables that reflect cost, service quality and convenience. The result of a utility analysis will be something like "people seem to value the time spent commuting at something close to their average wage". Such an estimate of the value of time could be used in estimating the benefits of a transportation project that saved time for commuters or other travelers.

Economic decisions concern future costs and benefits. Money spent in the past should not affect what we decide to do today, and such costs can be viewed as **sunk costs**. For instance, if you are about to buy a new car, and you plan to trade in your old car, it does not matter – to you or to the car dealer – what you paid for that car. What matters is what the car is worth today, which will depend upon the condition of the car and the demand for used cars. On a larger scale, when trying to decide whether or not to buy or sell a building, it does not matter what that building cost to build. It is only the market value of the building that will affect the price. Of course, if you have yourself put a great deal of money into buying a car or your house, you may well perceive that the car or house is worth a lot more than anyone else does, but that is only a factor in deciding whether you are willing to part with it. The current market value is what should enter your economic analysis.

Summary

This chapter has introduced supply, demand, equilibrium, competitive and monopolistic pricing, productivity, utility and other economic concepts that are relevant to project evaluation. These concepts provide a framework for thinking about needs, projects, and project evaluation.

Supply and Demand

The **supply function** describes the amount of output that will be provided as a function of the price per unit that is sold. The **demand function** describes the amount of output that will be purchased as a function of the price per unit. The **equilibrium price** is the price for which supply will equal demand. In most complex systems, there will be continual changes in both supply and demand, and it is more realistic to think about systems moving toward equilibrium rather than always being in equilibrium, especially for systems where it is costly and time-consuming to adjust capacity. Congestion, delays and poor quality are likely when demand exceeds supply, while underutilization of equipment and reductions in the work force of suppliers will be common when supply exceeds demand.

The **elasticity of demand** with respect to price can be estimated by observing the effects of price changes on demand. If demand is elastic, then demand will be more responsive to price changes, and an increase in price will lead to a decrease in total revenue. Demand tends to be more elastic in the long-run than in the short-run, as people and businesses will generally find ways to reduce their dependence on higher-priced goods and service.

Productivity

Productivity is defined as the ratio of output to input. Improving productivity is an important objective, because productivity improvements make it possible to produce more goods and services using fewer people and resources. Many projects eliminate productivity problems related to bottlenecks, constraints on usage, inadequate control, or outmoded facilities. Productivity may also be improved by changing the structure, design and size of networks or facilities so as to achieve economies of scale, scope, or density. **Scale economies** exist when expanding the size of the system leads to reductions in average cost. **Scope economies** exist when it is cheaper to use facilities for multiple uses. **Density economies** occur in a network when more volume is concentrated on each route.

Pricing

In a **competitive environment**, no individual supplier has the power to set prices, and prices will fall to marginal cost. In infrastructure-based systems, **marginal cost pricing** will generally be well below average cost as long as the system is operating below its design capacity. Thus there may be a need for price regulation or subsidies to ensure that revenues are sufficient to cover total costs of operation. If demand approaches capacity, then marginal costs for both users and operators will rise as delays and high utilization levels make it difficult to use and maintain the system.

If a supplier has a **monopoly**, it can set prices well above marginal costs so as to maximize profits. Monopolies may also be slow in adopting new technology or expanding capacity to meet demand, and they may display little concern for service quality. However, for many infrastructure-based systems, there are tremendous economies of scale and density, so that the cost of service can be greatly reduced by limiting competition. Thus, many transportation companies and public utilities are allowed to operate as monopolies in order to achieve cost savings, while being subjected to regulation in order to ensure reasonable prices and service.

Measuring and Improving the Economy

Public officials and the general public are naturally interested in expanding economic output, which is commonly measured as the **gross domestic product (GDP)**. GDP is the sum of all private consumption, private investment, government expenditures and net exports. Other measures of the economy include **total jobs**, **unemployment** levels, and **average income** per person or per family. All of these measures can be developed for a region as well as for a country.

A major project will have both a direct and an indirect effect on the regional economy. The direct effect will be related to the jobs created and the expenditures required to complete and subsequently to operate and maintain the project. In addition, there will be a **multiplier effect**, because the people who work on the project and the companies that sell materials to the project will spend much of what they earn, whether on food or cars or housing, and other people and other companies will enjoy some added income.

Trade

Trade allows regions to specialize in economic activities where they have a **comparative advantage** relative to other regions. By producing more of what they need for some types of products, they are able to trade for other things that they need or desire. The ability to trade is dependent upon the ability to transport goods efficiently between regions, because of **spatial price equilibrium**. In order for trade to make sense, the cost of producing something in one location plus the cost of transporting the product to another location must be less than the price that can be charged in that location. Investments in transportation systems have produced dramatic reductions in transport costs, thereby enabling the shift of manufacturing, mining, agricultural production, and other activities to the regions of the world with the lowest costs. The global economy reflects low transportation costs and the fact that there are generally high economies of scale in production.

Banking and currency exchange are another essential aspect of global trade. Exchange rates between currencies of different countries may be determined by market forces or by regulations. Over time, **exchange rates** may vary substantially, which will tend to change the patterns of trade by making some countries relatively cheaper and other relatively more expensive. Growth in trade and changes in trade routes are important considerations in many infrastructure projects.

Making Decisions

Despite the fact that we know ourselves often to be less than rational in our decisions, economists assume that individuals will generally make decisions so as to maximize their own **utility**. Utility is a rather vague – but thoroughly useful – concept that can incorporate disparate factors such as cost, convenience, reliability, safety or aesthetics that might affect our choice of a new car, a new house, or where to eat dinner. By observing what decisions people make (**revealed preferences**) or asking people about hypothetical choices (**stated preferences**), it is possible to infer what factors they consider in making choices. Those who plan projects must, at some level, consider how many people will use the completed project (road, water system, park or office building) and how much they will be willing to pay for their use of it.

When evaluating proposals, only future costs and benefits need to be considered. Money spent in the past is a **sunk cost.** Whether that money was well-spent or wasted does not and should not affect decisions concerning what to do in the future.

Manchester, New Hampshire

Water power enabled Manchester to become a dominant manufacturing center during the 19th century. When the mills closed, the old buildings were transferred into office space museums, restaurants, and small businesses, and walkways were constructed along the Merrimac River in order to attract visitors.

Public Perspectives: Economic, Environmental and Social Concerns

Happiness lies not in the mere possession of money; it lies in the joy of achievement, in the thrill of creative effort. The joy and stimulation of work no longer must be forgotten in the mad chase of evanescent profits. These dark days will be worth all they cost us if they teach us that our true destiny is not to be ministered unto but to minister to ourselves and to our fellowmen. ... Our greatest primary task is to put people to work. This is no unsolvable problem if we face it wisely and courageously. It can be accomplished in part by direct recruiting by the government itself, treating the task as we would treat the emergence of a war, but, at the same time, through this employment, accomplishing greatly needed projects to stimulate and reorganize the use of our natural resources.

Franklin D. Roosevelt, president of the United States, First Inaugural Address, March 4, 1933
(*Record,* 73 Congress, Special Session of the Senate, pp. 5-6)

Overview

The public sector, which is responsible for many kinds of infrastructure systems, takes an entirely different perspective than the private sector when identifying needs and evaluating potential projects. The primary motivation for the private sector projects will be the financial returns to the owners, not the broader effects on the economy or society. For the public sector, the motivation is not to earn money, but to satisfy public needs or to promote growth in the economy. Financial issues are important, but not necessarily dominant, and every major project will have multiple purposes and multiple measures of effectiveness (Table 1). Social and environmental impacts are central to the public evaluation process, and equity in the distribution of costs and benefits will be critical. In dealing with non-monetary objectives, cost effectiveness will be a more relevant concept than return on investment: which of the proposed alternatives is the best way to achieve the desired objectives?

Table 1 Examples of Public Infrastructure: Multiple Purposes and Multiple Measures

Type of Infrastructure	Purpose	Measures
Transportation	Mobility Accessibility Regional competitiveness	Service levels (travel time, congestion) Cost of transportation Fuel consumption Safety Emissions
Dams	Flood control Irrigation Hydropower Recreation (boating, swimming, camping, picnic sites)	Risks associated with floods Volume of water available for irrigation Land area to be irrigated Electricity production (cost & revenue) Impact on wildlife
Water & sewage	Clean water for consumption Water for industry & irrigation	Volume of water available for each type of use Cleanliness (risk of disease) Cost per unit
Public Housing	Housing for elderly Housing for low income residents Housing for homeless	Number of units Size and quality of buildings Cost per unit (construction & operation) Safety & Security Aesthetics
Parks & recreation	Open space for residents Protect environment Aesthetics	Open space as % of total space Visitors per year Diversity of wildlife Safety

Another difference is that the time frame of the analysis will be much longer for the public than for the private sector, as the public entity is presumed to endure indefinitely. The long time frame requires the consideration of sustainability – will projects or programs be sustainable over long periods of time taking into account economic, financial, social, and environmental factors?

For many kinds of public infrastructure projects, the tolls, fees, and other direct revenues from the project will be insufficient to cover the costs of the investment. However, the non-monetary benefits could be considerable. A public transportation project may relieve congestion, improve air quality, and promote mobility for those without access to an automobile. While these benefits can at times be quantified by using concepts like consumer surplus or the multiplier effect of new investment, such benefits do not produce cash flows that cover the interest on bonds or the operating costs of the transit agency. If the benefits are clear, and if the public generally believes that these benefits are worthwhile, then public agencies may be able to use tax revenues to supplement the direct cash flows from the project. If taxes are used to finance a project, then that project will be competing not just against similar projects, but against all of the other projects that might be undertaken by that city, state or country. Transportation projects compete with housing for the elderly, and water projects compete with health care projects. Decisions for or against projects will be political decisions, and the relative importance of various kinds of costs and benefits will be subject to considerable debate.

To complicate the situation further, a project will also be evaluated in terms of its impact on the community:

- Economic impacts, including employment, regional economic growth, regional competitiveness
- Environmental impacts, including air quality, water quality, noise, loss of wetlands, and impact on ecosystems.
- Equity, including the distribution of costs and benefits across regions and groups of the population and the relative impact on current and future generations.
- Aesthetics, including the appearance of the new infrastructure, its effect on neighboring areas, and its effect on long-term changes in land use.
- Other social impacts, including such things as impacts on communities during construction, displacement of residents, and long-term changes in population distribution

Multiple objectives and multiple measures mean that these projects are inherently complex, with many conflicts possible among different objectives. The decision-makers ultimately will include the public, who may have a chance to vote for or against the funding sources proposed for a project, and the politicians or appointed officials who must justify their decisions to the public in order to be re-elected or to retain their jobs. Large projects will be politically sensitive, and it will be necessary to consider and to balance all of the conflicts. There will be real and apparent conflicts of interest among those who are supposed to be proposing, evaluating, and approving projects. It will not be possible to satisfy everybody, and there will likely be determined opposition to almost any major project. People commonly do not want anything built too close to them, even if they are going to be major beneficiaries of the project. This phenomenon, which can lead to intense community opposition, is known as the NIMBY response: "not in my backyard".

In summary, a major public project will be evaluated by many different groups of people, from many perspectives, with varying concerns for the relative importance of various features of the projects, and with potential disputes about how to measure or estimate costs and benefits.

Benefit/Cost Analysis

Public projects require an evaluation process that includes, but is much broader than financial analysis. A simple dictum is mandated both by law and by common sense: for any public project, the total benefits should exceed the total costs. This does not mean that every project with more benefits than costs is a good project; it simply means that projects whose costs exceed their benefits are bad projects that should not be funded by the public. This may seem to be a rather obvious principle, but it surely is necessary, as there are many instances of projects being built, at taxpayer expense, whose costs far exceeded their benefits. There are even names for such projects: "gold plated" projects that could have been constructed for far less money; "pork barrel" projects that were approved in order to get a crucial politician's support for some larger political scheme; and "white elephants" that are constructed at great expense, but that afford no greater benefits than ordinary elephants! The political process can provide a means to fund many different projects, and it is possible that many projects will be "earmarked" (i.e. specifically authorized in the legislation) rather than subjected to a rigorous examination of their costs and benefits. A requirement that the benefits of every project should exceed the costs is therefore a step toward a more rational allocation of public funds and a defense against mismanagement, stupidity, and corruption.

Measurement will be a major problem in determining whether or not costs exceed benefits: how can different types of non-cash costs and benefits be converted to monetary terms? How can important benefits such as savings in travel time or reductions in risk of accidents be converted into monetary terms? What about aesthetics? In some cases, the monetarization is straightforward; in other cases it is convoluted and controversial; and in still others it is essentially impossible.

For example, consider a proposal to construct a new highway that is intended to provide a safer, more attractive route around the congested core of a city. The basic question is whether the savings in travel time, the expected reduction in fatalities, and the prettier route justify the cost and the environmental impacts of constructing the highway through the surrounding region.

<u>Travel time</u>: Traffic engineers are able to model how commuters, truckers, and others will use the new facility, and they will be able to predict traffic flows on the new facility along with the changes in traffic flows on other facilities. Based upon the changes in traffic flows, they will be able to predict travel times on the new road and changes in average travel time on each segment of the existing network. The overall effect can be summarized as a reduction in travel time measured as vehicle-hours per day or per year, with details for commuters, local delivery trucks, long-distance trucks traveling through the region, and any other group of interest. The value of these time savings is commonly estimated by making a series of assumptions. For instance, the time saved by commuters could be valued by using the average hourly wage for workers in the region, and the value of time saved by truckers could be valued by using the average hourly wage for truck drivers, the hourly cost of truck ownership, and the hourly value of the contents of the truck. Some might argue that something less than the average hourly wage should be used, and others might challenge the methodology used to estimate the hourly cost of truck ownership, but these estimates of the value of time are commonly accepted, and the benefits are clear and verifiable.

<u>Safety</u>: estimating the value of the safety benefits will be trickier. Traffic engineers will be able to predict the number and severity of accidents based upon traffic flows and highway geometry, and safety analysts can use past history to quantify the expected damage to vehicles and the highway. However, no one can readily place an economic value on the most important safety benefits, namely a reduction in the expected number of injuries or fatalities. Instead, departments of transportation in some countries will consider the value to society of reducing fatalities and serious injuries resulting from automobile and other transportation accidents. In the United States, the U.S. Department of Transportation uses a value of approximately $2.5 million in its risk analysis. This amount represents the benefit to society of eliminating a single, future fatality. It does not represent the value of a human life, for it is impossible to say who would have been hurt or killed. In effect, the $2.5 million can be viewed as an aggregate benefit to all users of the system of a slight reduction in the probability of a fatal accident. Every user benefits because the probability of an accident is reduced.

<u>Aesthetics</u>: now we are close to the "impossible" in trying to quantify the benefits of the new highway. Whether or not aesthetics is viewed as an important component of the decision will depend upon the political situation in the region. It could be viewed as an afterthought, which might mean planting some flowers along the right-of-way, or it could be a major design consideration, as in the construction of roads in national parks or the construction of parkways into major cities. The argument may well boil down to someone showing artist's renditions of the options (or photos of similar projects elsewhere) and saying something like "isn't it worth spending an extra $5 million to get a nice

facility?" If people agree that "$5 million" is a small amount, then they will choose the more aesthetic option. If people note that $5 million" is equivalent to the park budget for 20 years, then they will probably be vocal in their opposition!

Economic Impacts: Measures Related to the Regional or National Economy

Governments and public agencies will be concerned with the effects of projects on the local, regional, or national economy. Primary measures will include gross regional or national product, jobs created or lost, average income, and personal and industrial productivity. These economic benefits could come from several types of benefits:

- Construction jobs and income
- The multiplier effect of construction
- Jobs and income related to the eventual operation of the new project
- The multiplier effect of operation of the new project
- Continuing productivity benefits resulting to citizens, users, industries or public agencies as a result of the project
- Growth in the economy related to the productivity benefits provided by the project

For example, consider the construction of a toll road. The initial construction may take two years, provide hundreds of jobs, and increase sales of construction materials within the region. The direct expenditure of several hundred million dollars would have a multiplier effect that would more than double the economic benefit to the region during the period of construction. Once the toll road is opened, there could be long-term jobs for toll-collectors (or for those who maintain any electronic toll collection devices) and for highway maintenance forces, providing both a direct and a multiplier effect to the regional economy. The toll road presumably offers benefits to the public in terms of higher capacity for rush hour traffic, reduced risks of accidents, and perhaps reduced travel time. With less congested highways, the region may be able to continue to attract new businesses and to absorb additional population growth. Land near interchanges is likely to increase in value and attract hotels, restaurants, trucking terminals, warehouses, and other businesses that depend on highway access or serve highway users.

These benefits could be offset by the impacts of both the construction and the continued operation of the highway. Disruption of normal activities can be a major economic cost of a highway project. Although construction of a new highway interchange will ultimately relieve congestion, it may cause increased delays for a year or two. Once the highway is built, it may act as a barrier that limits access between different parts of the region. Over time, land use will adjust to the existence of the highway, which could result in rapid growth in some areas and equally rapid declines in other areas.

Environmental impacts

Any project will alter the complex relationships between what might be thought of as the natural world and the manmade world. Construction activities convert more space from the natural to the manmade world. Projects require construction materials such as wood, steel, and concrete which ultimately depend upon activities such as forestry, mining, and manufacturing that certainly disrupt and may at times destroy the environment. Continued operation and maintenance of infrastructure require energy and other materials that ultimately come from the natural world. Normal operations, accidents and decay may release toxic substances that can affect air quality, water quality, soil composition and limit or destroy the ability of plants and animals to survive near project sites. Constructed facilities will cast shadows, they may be noisy, and they might just be ugly or interfere with people's day-to-day lives. Whether or not the benefits of the project are worth the environmental costs will always be a relevant question, especially when those receiving the benefits are not those who bear the costs. The extent to which this question is considered will depend upon the social, cultural and political institutions.

In many countries, developers must prepare an environmental impact statement (EIS) that at least states the goals of a project, presents major alternatives for achieving these goals, identifies the major environmental impacts, and suggests ways to mitigate the most negative impacts. Preparing an EIS ensures that information is made available to the public and to public officials who must approve a project; the extent to which environmental considerations affect decisions about a project may well depend upon legal and political battles.

Courts and legislative bodies are well-structured for dealing with controversial trade-offs between environmental and economic issues and the extent to which developers must deal with environmental concerns. Legislation has limited the development of wetlands, promoted soil and water conservation, required more fuel-efficient automobiles, and limited land use via zoning and other restrictive matters. However, courts and legislative bodies are not well-structured for dealing with the underlying science, as evidenced by the controversies related to the extent of, the causes of, and the possibilities of responding to global warming and climate change.

Mankind has certainly transformed the world. Over a period of many thousands of years, humans have converted vast portions of the earth's land area to agriculture, drained innumerable wetlands, developed much of the land near the oceans, seas, and major rivers, and cut down vast areas of forest. These activities have changed the chemistry of the atmosphere, altered the natural flows of fresh water, and restricted the natural habitats crucial for many species of plants and animals. These activities have also allowed humans to prosper by helping to ensure adequate food supply, clean water, housing, abundant energy sources, and protection against floods and other natural disasters. In the future, we will still eat, drink, use energy and improve the way we live – but we will have to pay more attention to our impact on the environment.

There will be a whole of range of environmental issues that must be addressed in evaluating any major project, and there will be many major projects whose primary objective will be to improve the environment. Environmental issues will range from very local debates as to what gets built in whose backyard to regional and national questions related to the use of resources to international questions concerning the future of the planet. Since we can't expect to answer all of these questions every time we want to build a new hotel or new segment of a highway, we need to provide a reasonable structure for addressing these issues within the project evaluation process.

Environmental Concerns

Let's start by considering some basic environmental concerns. Soil, water, sunlight, and temperature are among the factors that determine what plants can grow in any location. Plants that are well-adapted to local conditions will prosper, those that are poorly adapted will struggle, and those that cannot survive the extremes of temperature or hydrological conditions will never gain more than a short-term foothold. Insects, birds and small mammals are necessary to the propagation of many plant species, and worms, amphibians, and insects make soil into a complex, living community. Animals may feed on plants or other animals, and they prosper in locations where there is an abundance of food along with sufficient cover for their own safety and appropriate places to raise a family, whether in trees, burrows, rotten logs, stream banks, or wetlands.

Left undisturbed, any location will eventually develop with a characteristic set of plants and animals that can survive or flourish within the constraints posed by soil conditions, sunlight, and climate. Biologists have identified distinct **ecosystems** that can be characterized by the kinds of plants and animals that will be found there. Within any healthy ecosystem, there will be a diversity of species, each of which is somehow related to the health of the overall system. Pileated woodpeckers make holes in dead trees as they look for insects, and these holes are later used as nesting sites for chickadees. When the dead tree finally collapses, it will provide cover for mice and other small rodents, as well as an ideal place for fungi to grow or for grouse and hares to hide.

In many cases, there are species that will only be found in certain ecosystems, so that they can be considered to be **indicator species** that are useful in documenting the existence of unusual ecosystems. For instance, wood frogs lay their eggs in vernal pools, which are small pools that are formed in rainy seasons or in spring as the snow melts. Vernal pools dry up for part of the year, so they cannot support fish, which means that eggs deposited in a vernal pool will be safe from predation from fish. Wood frogs are an indicator species for vernal pools. A single female wood frog lays hundreds of eggs in the early spring, so if the pool retains water long enough for the eggs to turn into tadpoles and for tadpoles to grow into tiny frogs, then the wood frogs will prosper. If the vernal pools are filled as part of the process of building a parking lot or a suburban sub-division, then the wood frogs will die off.

In most regions, a few types of ecosystem will dominate the landscape, while a dozen or more other types will be commonly scattered throughout, and some will be found only in a few locations. Preserving the rare ecosystems may be essential for preserving bio-diversity, as there will be plants and animals that are only to be found in those locations. Preserving a good distribution of the more common ecosystems will prevent populations of plants and animals from becoming too isolated. Preserving large tracts of the dominant ecosystems will ensure healthy conditions for all of the region's most common species.

Ecosystems can be harmed in several ways. **Pollution** – the introduction of foreign elements into the air, the water, or the soil – may lead to the death of certain plants or insects and of the animals that depend upon eating them. Pollution could be in the form of toxic chemicals that are poisonous to certain species, but it could also just be the introduction of sediment into a pristine stream, thereby making the water quality unsuitable for certain types of fish. Pollution can also refer to the heated water that is discharged from a nuclear power plant, as the heated water will be lethal to the some species, while attracting others that may be alien to the previously existing ecosystem.

Disruptions to the flow and retention of water can have devastating effects on ecosystems. Draining wetlands to increase the land available for highways, housing or agriculture will lower the water table and make the remaining wetlands more susceptible to drought and fire. Extensive development in Florida, for example, has changed the flow of water through the Everglades, threatening the future of what was once the seemingly endless wetlands of southern Florida. More rapid runoff of water means that both floods and droughts are more likely, which means that certain species of plants and animals will have greater difficulty surviving.

Fragmentation of an ecosystem will eventually create areas that are too small to support the wildlife that formerly flourished there. A black bear requires a range of 10 to 100 square miles; if a region that formerly supported large populations of black bears is crisscrossed by roads and disrupted by housing developments and malls, then the habitat will no longer be large enough for the bears to survive.

For species requiring less extensive ranges, it is not so much the fragmentation of the habitat as the total **loss of habitat** that will be decisive. As agricultural land is turned into housing developments or malls, the birds that used to feed on the insects and seeds will have to go somewhere else, and the deer that used to feed on the leftover corn cobs will be hit by cars as they try to feed on the shrubs and gardens of the new developments. Colonies of butterflies and dragonflies will be lost, along with vast numbers of mice, voles, and moles and the hawks, owls, and weasels that feed on them. For migratory birds, the loss of habitat is especially problematical, as they need places to feed and to breed, perhaps on two continents, and they need extensive areas for resting and feeding along their migration routes.

A final threat to ecosystems comes from the introduction of **alien species**. In a well-functioning ecosystem, everything is in balance. Insects or other animals eat some but not all of the seeds, none of the animals eat all of any of the species of plants, and none of the plants grows so rapidly that it crowds out all of the other plants: it is a complex system of natural checks and balances. An alien species is one that originated in a distant ecosystem where it had adapted to competition with the other plants and animals that comprised that ecosystem. It undoubtedly served to control some of the other species, and other species controlled it. However, when introduced as an alien into a new ecosystem, there may be no controls and balances, and an alien species may prosper to the extent that it out-competes and eventually crowds out the native species. Purple loosestrife is a tall, tough wildflower that has a large woody ball of roots; it has numerous flowers on a spike, and it grows profusely in wetlands. When introduced to wetlands in the United States, it faces only modest competition from less aggressive plants, and it has no natural insect competitors. As a result, it can, within a few years, fill the wetlands, creating what is a beautiful purple covering but what is also a barren wetland. As there are no native insects that eat the stalks or the flowers, there are no native birds that are attracted to the plant, and there are no hawks circling to catch any of those birds off-guard. The weeds grow so close together that it is difficult or impossible for muskrats or beavers to keep their channels open, and there is too little space between plants to support families of ducks.

Alien species often get their start when ground is disturbed for some sort of construction project. If these species are not dealt with – which often requires people who search for the first aliens and then pull them out by hand – then they can rapidly spread and destroy many acres of land. The key point to remember is these alien species overflow their niche, eliminating the chance for native species to prosper, and also eliminating the niche that was occupied by insects and animals that depended upon the native plants and animals. Alien species may be beautiful, but they tend to limit biodiversity.

Maintaining the health of ecosystems requires local, regional, and national strategies. One useful concept is **Green Infrastructure**, which refers to the network of natural areas that is necessary to support the diverse populations of native plants and animals that live within a region. This term does not refer to man-made infrastructure that is constructed in an environmentally friendly manner. Rather, it refers to the connected natural system of open spaces, forests, waterways, and wetlands that allows plant and animal species to prosper. Green infrastructure includes the following kinds of components:

- Very large areas of undeveloped land that are able to support and protect habitat for the widest-ranging animals and ensure the continued existence of diverse ecosystems
- Small areas of undeveloped land that protect uncommon or rare ecosystems
- Numerous small or medium-sized natural areas that are large enough and close enough together to avoid isolation of plant and animal species
- Connecting corridors of open spaces that can be used by animals to move between the larger open areas

By acknowledging the existence of and the need for green infrastructure, it is possible for government agencies and conservation groups to develop plans that preserve and protect suitable green infrastructure. National parks, state parks, public conservation lands (e.g. national forests or wildlife management areas) can provide the critical large areas. Smaller parks, wetlands, and private land-holdings can protect enough smaller areas to ensure diversity and density of ecosystems. The hardest part is ensuring that wide enough corridors are maintained between and among all of the open spaces so that wildlife can in fact move throughout the region. The corridors need to be wide enough to be perceived as safe routes for animals to travel. For the largest mammals, 100 to 200 foot-side corridors will be needed. For smaller mammals and amphibians, narrower corridors will suffice. Land adjacent to waterways and wetlands is ideal for use as connecting corridors, as is land next to railroads, power lines, or other infrastructure networks.

Pollution can be controlled by limiting emissions, by confining emissions, or by cleaning up emissions. The cheapest control strategy is to prevent emissions, but that may or may not be feasible depending upon the nature of the process that causes the pollution. Some pollutants are extremely toxic, and even a small release can be hazardous to anyone living close to where the release occurs. Hence, special consideration is necessary in dealing with the most toxic chemicals and spent nuclear fuels or other radioactive substances. Finding a safe means of sequestering nuclear waste is one of the main challenges facing the nuclear power industry.

Climate change caused by excessive emissions of carbon dioxide, methane, and other so-called greenhouse gases is a major challenge for the world in the 21st century. Scientists believe that increasing concentrations of these gases in the atmosphere will trap heat, thereby leading to warmer temperatures. With warmer temperatures, there will be more energy available to power hurricanes, tornadoes, tsunamis and other extreme weather conditions. Warmer temperatures will also accelerate the melting of glaciers and the ice caps, which will raise the level of the oceans and threaten flooding of the many cities and developed regions along the coasts. Changes in climate could also include regional changes in precipitation, which could have major implications for agriculture and for the natural environment.

Environmental Impacts of Projects

Proposals for major infrastructure projects can raise many different kinds of environmental issues. Abutters, politicians, environmental organizations and others are likely to express concerns about some or all of the following:

a. Use of materials in construction and operation
b. Pollution: impacts on air quality, water quality, and soil toxicity
c. Loss of habitat and disruption of ecosystems: Impacts on plants and wildlife
d. Impacts on the local environment (noise, shade, aesthetics)
e. Sustainability or the lack thereof

It will not be possible or necessary to consider all possible levels of impacts for every project that is considered. Regulations can be developed that govern the use of materials and that establish acceptable limits for pollution. Regional plans can help identify the necessary green infrastructure, and zoning can be used to direct development away from the most critical natural areas. Local impacts on noise, aesthetics, and land use will of course be a concern for nearly any project, and some sort of community involvement can be helpful in anticipating and responding to potential problems.

In the United States, a process has been created to ensure that environmental impacts are considered along with the economic and social impacts of any major project or program involving federal funding or approvals. This process emphasizes the need for determining and disclosing environmental impacts in what is called an **environmental impact statement (EIS)**, and it requires developers to consider how to mitigate negative impacts, but it does not indicate what can or cannot be done. That determination is left to the legislatures and the courts. An EIS is required for any major federal legislation or action "significantly affecting the quality of the human environment."[1] The federal agency proposing the changes must prepare the EIS, which must include "a detailed statement of these environmental effects."

> *"The National Environmental Policy Act of 1969 (NEPA), as amended, (42 U.S.C. 4321 et seq., Public Law 91-190, 83 Stat. 852), requires that all Federal agencies proposing legislation and other major actions significantly affecting the quality of the human environment consult with other agencies having jurisdiction by law or special expertise over such environmental considerations, and thereafter prepare a detailed statement of these environmental effects. The Council on Environmental Quality (CEQ) has published regulations and associated guidance to implement NEPA (40 CFR Parts 1500-1508)."*[2]

The Environmental Protection Agency (EPA) is responsible for reviewing the draft EIS and rating it according to two criteria. First, the EPA must decide whether or not the EIS is acceptable in terms of the depth of its analysis and the completeness of its findings. Second, EPA rates the environmental impact according to one of four categories:

- Lack of objections (LO): "The review has not identified any potential environmental impacts requiring substantive changes to the preferred alternative. The review may have disclosed opportunities for application of mitigation measures that could be accomplished with no more than minor changes to the proposed action."
- Environmental Concerns (EC): "The review has identified environmental impacts that should be avoided in order to fully protect the environment. Corrective measures may require changes to the preferred alternative or application of mitigation measures that can reduce the environmental impact."
- Environmental Objections (EO): "The review has identified significant environmental impacts that should be avoided in order to adequately protect the environment. Corrective measures may require substantial changes to the preferred alternative or consideration of some other project alternative (including the no action alternative or a new alternative)."
- Environmentally Unsatisfactory (EU). "The review has identified adverse environmental impacts that are of sufficient magnitude that EPA believes the proposed action must not proceed as proposed."

[1] The material in this sub-section is based upon EPA's "Policy and Procedures for the Review of Federal Actions Impacting the Environment" October 3, 1984, pp. 4, 19-20. The document is available from EPA's web site (www.epa.gov)
[2] Ibid. p. 4

The EPA review is a major hurdle for any project involving the federal government or requiring federal approval. After a draft EIS is published, time is allowed for public comments concerning what the draft includes or fails to include. The draft EIS and all of the comments and procedural rulings are available to the public on-line. If the EPA finds environmental concerns, it may require substantial changes in the proposed actions or prevent the project from proceeding as proposed. Moreover, if EPA finds the EIS to be inadequate, EPA may require it to be revised or redone, an action that could delay a project for a year or more.

The conditions that would allow the EPA to raise environmental objections are specified by government regulations. Objections can be raised in five situations:

1. *"Where an action might violate or be inconsistent with achievement or maintenance of a national environmental standard;*
2. *"Where the Federal agency violates its own substantive environmental requirements that relate to EPA's areas of jurisdiction or expertise;*
3. *"Where there is a violation of an EPA policy declaration;*
4. *"Where there are no applicable standards or where applicable standards will not be violated but there is potential for significant environmental degradation that could be corrected by project modification or other feasible alternatives; or*
5. *"Where proceeding with the proposed action would set a precedent for future actions that collectively could result in significant environmental impacts."*

In other words, EPA must have a clear reason for raising objections, and other guidelines and policies will be used to determine whether proposed actions are acceptable or not. More stringent guidelines are in place for finding a proposal with environmental objections to be environmentally unsatisfactory:

1. "*The potential violation of or inconsistency with a national environmental standard is substantive and/or will occur on a long-term basis;*
2. *"There are no applicable standards but the severity, duration, or geographical scope of the impacts associated with the proposed action warrant special attention; or*
3. *"The potential environmental impacts resulting from the proposed action are of national importance because of the threat to national environmental resources or to environmental policies."*

Thus, the environmental review process places the onus on the proposing agency to identify the potential impacts, while establishing an agency with the necessary skills and responsibility to review and interpret the EIS. The criteria cited above could be quite qualitative, leaving approval up to the judgment of the EPA. Since the whole process is open to the public, it is possible for groups opposed to any action to make their objections to EPA.

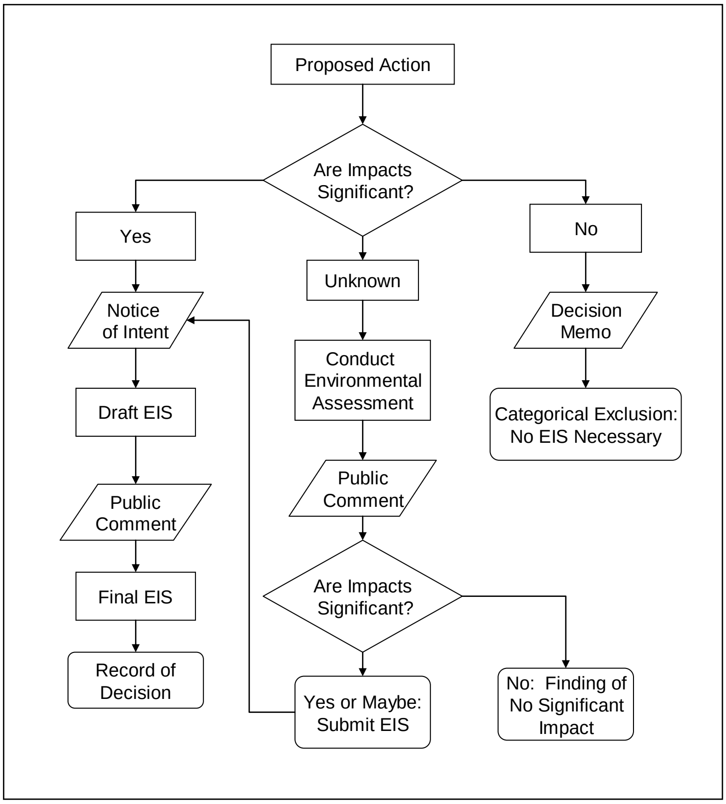
Figure 1 summarizes the review process used by EPA. If a major project or action is proposed, the first step is to determine whether or not there will be significant impact on the environment. If EPA believes that there will be no such impact, then it can allow the project to proceed without an EIS. If EPA finds that there will be significant impact on the environment, then the proponents of the project must prepare a draft EIS, which will be available for public comment and review by EPA prior to submission of the final EIS. EPA will then make its decision, as described above. If EPA finds that the environmental impact is unknown, then an initial environmental assessment can be required, which could lead to a finding of no significant impact or to the preparation, review, and revision of an EIS.

A great deal of judgment is involved in preparing and reviewing environmental impact statements. EPA has developed two sets of checklists of questions that might be asked in order to guide reviewers as they evaluate an EIS.[3] The first set of checklists address general areas of environmental concern that might apply to any proposed project: energy management, habitat preservation, landscaping, water use, and pest management. In each area, EPA has summarized

[3] Science Applications International Corporation, "Pollution Prevention – Environmental Impact Reduction Checklists for NEPA/309 Reviewers", Final Report, EPA Contract No. 68-W2-0026, January 1995

the relevant scientific factors and posed a series of questions that could and probably should be asked when reviewing any project. The second checklist provides questions that address the problems likely to be encountered in different types of projects. In each special area, such as a transportation or waterway project, the guidebook includes a brief summary of the likely problems and then provides several dozen questions that could be asked to determine how well an EIS has addressed issues that might be anticipated.

Figure 1 Summary of NEPA Decision Process



Social Impacts

Almost any project will have social impacts that may be related to the users of the project, people who live near the project, people who are displaced or competitively disadvantaged because of the project, or people who are hurt or whose lives are hindered as a result of the construction or operation of the project. Social impacts could be positive as well as negative, but it is the negative impacts that must be considered most carefully. Positive social impacts will help make a project more attractive, whereas negative impacts may be sufficient to arouse intense public opposition that prevents or markedly restricts a project. Anticipating negative impacts is therefore something that should be done early in the evaluation process, so that there will be an opportunity to adjust plans so as to reduce the negative impacts or to provide means for mitigating them.

Major projects may have far-reaching consequences that are difficult or impossible to quantify or comprehend. In some cases, projects that appear at first to be wholly desirable turn out to have unexpected consequences that are viewed very unfavorably by some people. In their famous study of Middletown, Robert and Helen Lynd reported that some residents recognized the social benefits of having automobiles for commuting, but were outrage by the dreadful social impacts of auto ownership:

"No one questions the use of the auto for transporting groceries, getting to one's place of work or to the golf course, or in place of the porch for "cooling off after supper" on a hot summer evening; however much the activities concerned with getting a living may be altered by the fact that a factory can draw from workmen within a radius of forty-five miles, or however much old labor union men resent the intrusion of this new alternate way of spending an evening, these things are hardly major issues. But when auto riding tends to replace the traditional call in the family parlor as a way of approach between the unmarried, "the home is endangered", and all-day Sunday motor trips are a "threat against the church;" it is in the activities concerned with the home and religion that the automobile occasions the greatest emotional conflicts."[4]

In the United States, social impacts must be considered as part of the process required for environmental impact assessment. A set of principles and guidelines for social impact assessment was developed by the Interorganizational Committee on Principles and Guidelines for Social Impact Assessment (ICPGSIA), a group of social scientists who sought to help public agencies and private organizations in carrying out responsible social impact assessment (SIA). Their motivation was that "SIAs help the affected community or communities and the agencies plan for social change resulting from a proposed action or bring forward information leading to the reasons not to carry out the proposal."[5] Like the environmental impact assessment, a major purpose for the SIA is to provide a mechanism for understanding and responding to the potential negative impacts of proposed policies, programs or projects.

> *By social impacts, we mean the consequences to human populations of any public or private actions that alter the ways in which people live, work, play, relate to one another, organize to meet their needs and generally cope as members of society. The term also includes cultural impacts involving changes to the norms, values, and beliefs that guide and rationalize their cognition of themselves and their society.*
>
> The Interorganizational Committee on Principles and Guidelines for Social Impact Assessment

This group defined social impacts and identified six principles for social impact assessment. The first principle calls for identifying the people who will be affected by the proposed action and collecting information about their social conditions so as to establish a base line for evaluating changes to those conditions. The second principle is that the analysis should be focused on the most important social and cultural issues that are likely to be affected. The SIA need not address every possible social or cultural impact that might be imagined. The third principle emphasizes the need for using proper methods and input from the public in identifying and quantifying problems that might be encountered. In other words, social scientists know how to do this kind of analysis, and they should be involved early in the design and evaluation process for major projects. The fourth principle establishes the role of the SIA as providing information to be used by decision makers and the public; the SIA and the people conducting the SIA are not the ones who ultimately make decisions about whether or not to go ahead with the project.

The fifth principle deals with environmental justice, which refers to the sometimes commonly used approach to locating or structuring projects: "locate them in the poorest neighborhoods and don't worry about how the disadvantaged will be hurt by the project." Who benefits and who pays are important considerations in SIA and in project evaluation in general. The final principal indicates that SIA doesn't end when the project or program or policy is implemented. It is necessary to monitor what happens to ensure that mitigation measures are actually implemented and to ensure that unforeseen social impacts will be recognized.

---

[4] Robert S. and Helen Merrell Lynd, *Middletown: A study in Contemporary American Culture*, New York, 1929 [the authors spent more than a year in Muncie, Indiana, interviewing residents about all facets of life in that small city].

[5] The Interorganizational Committee on Principles and Guidelines for Social Impact Assessment, "Principles and guidelines for social impact assessment in the USA", **Impact Assessment and Project Appraisal,** Vol. 21 No. 3, Beech Tree Publishing, Guilford, Surrey, UK, September 2003

Predicted social impacts may be temporary or long-lived, and there may be minor impacts that affect a lot of people or intense impacts for a few people. To understand the importance of social impacts, it will help to consider the kinds of social impacts that might be encountered in typical infrastructure projects. Table 2 lists some of the notable types of impacts.

Table 2 Examples of Negative Social Impacts of Projects

Type of Impact	Examples
Relocation of people	• Entire villages displaced for the construction of a dam. • Hundreds of people and small businesses relocated to allow the construction of a highway through a city.
Deaths and injury during construction	• Deaths of more than 20,000 from tropical disease in the various efforts that eventually led to the Panama Canal. • Deaths resulting from workers falling off bridges or buildings in situations where safety nets were not installed.
Deaths, injury or illnesses resulting during normal operation of infrastructure	• Millions of people severely injured or killed in highway accidents • Bridges and tall buildings serving as jump-off points for suicides • Asthma and other illnesses resulting from air pollution caused by emissions from power plants, automobiles, or home heating • Tens of thousands of people injured or killed annually worldwide in grade crossing accidents between highway vehicles and trains
Deaths and injuries resulting from infrastructure failure	• Thousands of deaths and destruction of cities resulting from dam failures. • Loss of life from buildings and structures that collapse in earthquakes
Disruption of neighborhoods	• Limited access highways serving as barriers when they are constructed so as to divide urban neighborhoods. • Loss of property values following construction of large, noisy, or ugly buildings or infrastructure. • Creation of suburbs and decline of central cities following construction of better highways and policies that encouraged home ownership.
Loss of livelihood caused by negative environmental aspects of a project	• Destruction of fishing and shell-fishing areas following construction of bridges, port facilities, or oil spills • Decline in use of informal taxis and buses following opening of new subway lines in large cities in Latin America and Asia.
Loss of livelihood related to projects that help competitors	• Bankruptcy of canal companies following construction of railroads. • Bankruptcy of railroads following construction of highways and invention of cars, trucks and airplanes. • Decline in newspapers following widespread use of the internet.
Loss of privacy	• Disruption of the life of native peoples following construction of roads or railroads through their previously remote homelands.
Reduced quality of life	• Noise and dust resulting from construction of a highway • Shade resulting from construction of tall buildings

It might be even more helpful to consider some projects where the social impacts, whether foreseen or unforeseen, turned out to be devastating or reprehensible. If the leaders of the French company that set out to build the Panama Canal knew that tens of thousands would die in their failed attempt, would they have ever begun the project? If city officials in New Orleans had long ago understood the risks posed by hurricanes, would they have allowed housing to be built in the lowest-lying areas of the city? If automobile manufacturers, highway engineers, and government officials truly understood the dangers of automobiles (hundreds of thousands killed worldwide each year), would we have the system that we have today? These questions are worth some discussion. Hindsight may suggest that we would have done things differently if we had only known – but maybe the drive for a route to the Pacific, the need for more housing, and our great love affair with the automobile would have led us exactly to where we ended up.

Various researchers and agencies have used different categories of social impacts, and it is possible to construct quite elaborate topologies of social impacts. Whatever the categories, the main concern for the SIA will be to determine who is going to be hurt by the project or program, when will problems arise, and what can be done about them. As with environmental impact assessment, it is believed that there is much to be gained simply requiring these questions to be asked, with the answers and supporting information made public. If the problem is understood, then it may be possible to take action.

There broad categories of actions can be identified

- Adjust the design so as to avoid or reduce the social impacts
- Require mitigation as a condition for approval of the project
- Compensate those who are hurt by the project.

Whether or not any or all of these are necessary is something that will ultimately be decided by those who are threatened, local governments, developers, other stakeholders, and the courts.

Safety and Security

Safety and security are particularly important and emotional social concerns. Public reactions to projects seldom derive from a calm, rational assessment of the costs and benefits. Sometimes the public response is driven by fears and emotions, whether the fears relate to the potential for disaster or for national security. Proponents are likely to downplay the potential problems, while opponents are likely to stir up people's emotions. The classic case is nuclear power. While there have been only a very few serious accidents involving nuclear power plants in the U.S., the public has been very fearful of such accidents and very leery of proposals for sequestering nuclear waste. If nuclear power plants are built to modern safety standards, and if radioactive waste is properly sequestered, then they would seem to provide an efficient, clean alternative to the use of fossil fuels for generating electricity. However, public fears have forced extraordinary measures to be taken to limit the risks of such plants, and some countries have banned such plants altogether.

In the United States, a rather inconsequential incident at the Three-Mile Island Power Plant in 1979 was "the most serious in U.S. commercial nuclear power plant operating history, even though it led to no deaths or injuries to plant workers or members of the nearby community."[6] The accident led to very little off-site release of radiation. The average dose (1 millirem) of radiation to the population of 2 million closest to the site was only one-sixth of the dose received during a full set of chest x-rays. In addition, the NRC reports that studies have determined that there were no more than negligible effects on the environment. However, the accident at Three Mile Island led to sweeping changes in the regulation of the nuclear industry, including the addition of many costly safety procedures. The negative media attention that it received also created a terrible public image for the nuclear power industry as a whole, and no new plants were built in the US for 30 years. Acceptance of nuclear power varied widely in other countries. In Germany, public opposition led to a political decision in 2003 to phase-out of nuclear power plants, but in France, nuclear power had by then become the dominant source of electrical power.[7]

[6] Fact Sheet on the Accident at Three Mile Island. 1 Mar. 2004. United States Nuclear Regulatory Commission. 10 Nov. 2004. < http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html>
[7] Stephen Graham, "Germany begins nuclear phaseout," The Boston Globe, November 15, 2003

Summary and Discussion

The public sector's perspective regarding projects differs in several important ways from that of the private sector. Entrepreneurs view projects as a means of making their fortune; investors view projects as a way to earn a return on their investment; public officials and charitable organizations view projects as a way to achieve goals related to society's needs or desires. Many of the differences can be captured by the distinctions between financial analysis and economic analysis. Financial analysis addresses cash flows, whereas economic analysis also considers safety and security, growth in the economy, and provision of education, infrastructure, and other public services.

Another difference between the public and private perspective is that environmental and social concerns are more likely to motivate public projects, whereas such concerns are apt to be constraints on private projects. This is not an absolute difference. In particular, many infrastructure projects may have very negative consequences on the environment or upon people who live near project or who must move to accommodate the projects. Construction of roads through cities, creation of dams that displace thousands of people, irrigation projects that require flooding of fertile valleys, and use of public land for forestry and mining are examples of public efforts that can lead to serious controversy among agencies and among those who favor or oppose the projects. When controversies arise about public projects, a change in government, legal battles, or the passage of new legislation may be needed to resolve the issues. When controversies arise with respect to the social or environmental impacts of private projects, there could be legal and legislative battles to determine what kinds of activities are allowed and what kinds are prohibited.

A principle underlying public projects is that the benefits should exceed the costs. The project itself does not have to produce the cash necessary to cover the investment or even the operating costs, as the government has the power to raise taxes to pay for projects. However, the project should provide measurable economic benefits that are at least as great as all of the economic costs of the project. Just as the benefits can be more than the cash flow directly related to the projects, the costs can include social, economic and environmental impacts that are directly or indirectly related to the project. The logic underlying benefit/cost analysis can easily be misinterpreted. The proper interpretation of this kind of analysis is that the government should not pursue projects if the expected benefits are less than the expected costs. This does not mean that projects where benefits exceed costs should be approved, it only means that they deserve further consideration. Many controversies about the need for projects boil down to controversies as to what counts as a benefit, what counts as a cost, and how values should be put on factors such as safety, air quality, or job creation.

The extent to which concerns about environmental impacts, social impacts, or safety should affect the design or implementation of a project is something that will ultimately be decided by governments, developers, other stakeholders, and the courts. Over time, as the validity and magnitude of such concerns become clearer, it may be desirable or necessary to require improvements to existing infrastructure and develop more stringent regulations for locating, designing, and constructing new projects.

This essay has addressed the key factors that must be considered in enhancing the sustainability of infrastructure systems: financial feasibility, economic impact, environmental impact, and social impacts including safety and security. Finances are important, because cash will be needed to construct, maintain and operate infrastructure. Economic impacts are important, because they include many types of short- and long-term impacts that can help to justify a good project or to prevent a bad project. Environmental impacts are important, because an infrastructure-system that is too destructive to the environment or that requires excessive use of limited natural resources will not long endure. Social impacts are important, because it is ultimately society that decides whether or not to proceed with infrastructure projects and that bears the brunt of failures to consider hidden costs of projects.

It is clear that there are many diverse factors that will influence what needs are addressed, what projects are considered, and how projects will be evaluated. There will be no easy methods for determining the best projects, and no simple ways to gain public support for a particular project, although various methods and concepts can be used in reaching a consensus about what is needed and what should be done.

Comparing Strategies for Improving System Performance

"The basic idea is to define alternatives to just sufficient a level of detail to allow different stakeholders to at least rank order them in terms of desirability according to each identified criterion."

Theodor J. Stewart, "Thirsting for Consensus"

Introduction

Any major project is likely to have multiple objectives and measures of performance, especially but not only when public agencies are involved. Comparing strategies for improving system performance may at times be a straightforward matter of comparing financial costs and benefits, but it will be much more likely that important factors will be difficult or impossible to express in monetary terms. That is why project evaluation often focuses on cost effectiveness and why weighting schemes are commonly used when evaluating competing options.

The benefit of stating all costs and benefits in monetary terms is obvious: the methods used to compare financial costs and benefits can then be used to compare all of the economic costs and benefits. The danger of trying to state all costs and benefits in monetary terms is that the evaluation may simply disregard important measures that cannot be converted into monetary measures. The fact that it is difficult to put a price tag on beauty or equity is no reason to ignore aesthetics or to forget about being fair.

Discounting and Net Present Values

For some projects, it is possible to translate costs and benefits into monetary terms. At times, the translation can be straightforward, as when measuring changes in congestion by using the average values of time for the automobiles and trucks that are stuck in traffic. Other factors, such as safety, may require potentially controversial assumptions, e.g. the value to society of avoiding or incurring injuries or fatalities. Still others, such as aesthetics, can only with great difficulty be given a monetary value. Economists and other researchers have been very creative in developing methods of estimating the value of changes in measures that at first glance appear to be purely qualitative. For example, it is possible to link air quality to health and also to more mundane matters such as the need for window-washing and periodic cleaning of buildings.

Even if we can state the most important types of costs and benefits in monetary terms, we still have a problem in comparing options, since costs and benefits of an infrastructure project are likely to be spread across a period of many years or decades. The basic question is whether or not the expected future benefits will be sufficient to justify the initial investment. To answer this question, it is necessary to compare current and future costs and benefits. For any cost or benefit that can be monetarized, a process known as **discounting** can be used. Discounting provides a means of reducing future costs or benefits so that they can be expressed as equivalent **present values** that will be directly comparable to current costs and benefits. Discounting is used in both the public and the private sectors, with the public sector considering broad categories of costs and benefits and the private sector concentrating on cash flows. .

Discounting is necessary because money in the future is worth less than it is today or, to say the same thing in different terms, money today can be invested and grow into a larger sum in the future. The basic relationship between present and future value is a function of the discount rate r and the time period t:

(Eq. 1) $$\text{Present value} = \text{Future Value}/(1+r)^t$$

By use of this equation, future costs or benefits can be discounted to be equivalent to a current cost or benefit. For example, if an expected benefit of $1 million in year 4 were discounted at a rate of 5% per year, it would be worth only $822,000 today. The higher the discount rate, the lower the present value of a future sum of money. The same $1 million in year 4 would be worth less than $700 thousand today if it were discounted at 10%.

The **net present value (NPV)** of a project is the sum of the discounted values of all the benefits and costs associated with the project. If the NPV is positive, then the benefits outweigh the costs; if the NPV is negative, then the costs outweigh the benefits. Maximizing NPV may be viewed as the primary objective of a project, particularly in the private sector. In the public sector, this objective corresponds to maximizing the overall net benefits to society, taking into account any and all costs and benefits that can be expressed in terms of dollars and cents. If only financial costs are considered, as would commonly be the case in the private sector, then this objective would be stated as "maximize the net present value of cash flows".

The choice of a discount rate is an important factor in project evaluation. A very high discount rate will emphasize initial costs and make it more difficult to justify projects that require large investments; a very low discount rate will favor projects that have net benefits that continue over a very long time horizon. Private companies typically use a discount rate of 10% or greater in evaluating their projects; higher rates are used if projects are perceived to be riskier. Public agencies, which do not have to pay taxes, will typically use a lower discount rate. Discount rates will rise with inflation in both the private and public sectors.

This section has only introduced the most basic concepts of discounting and NPV, but even these basic concepts are very powerful. If we are given costs and benefits that occur over any stretch of time, and if we have a discount rate, we now know how to calculate the NPV of these benefits.

Measuring Cost effectiveness

It is always possible to calculate the cost effectiveness of any investment with respect to any non-financial, but quantifiable objective. Cost effectiveness is the ratio of the investment cost to the improvement in the measurement of interest, i.e. the $ per unit improvement. If different options are compared, the most cost effective option will have the lowest cost per unit of improvement.

Consider the case shown in Table 1, where there are three alternatives for reducing the impact of highway noise on the residents of two new apartment buildings that will be built adjacent to a heavily traveled interstate highway. The first option is to erect a quarter-mile long sound barrier along the edge of the property that abuts the highway. The second option is to require special sound-proofing for all of the windows and walls that face the highway. The third option is to design floor plans to minimize the windows on the walls facing the highway, e.g. by placing bathrooms, closets, and stairways along that wall. The effectiveness of each measure is estimated in terms of the average reduction in noise levels that would be experienced by residents living in the structure. In this hypothetical example, the three strategies result in a similar reduction in noise levels, and redesigning the floor space is the most cost-effective way to gain peace and quiet.

Table 1 Cost Effectiveness of Three Options for Reducing Noise Levels

Option	Cost	Reduction in Noise Levels	Cost-Effectiveness ($/% improvement)
Noise barrier	$500,000	20%	$25 thousand
Sound-proofing	$200,000	22%	$9.1 thousand
Redesign floor space	$100,000	21%	$4.8 thousand

Using Weighting Schemes in Multi-Criteria Decision-Making

Any major project, but especially public projects, will be trying to satisfy multiple objectives, many of which will be non-financial. There are likely to be many competing designs, some of which are markedly different in approach. Each of these projects can be evaluated in terms of each of the objectives, producing a (very large) matrix showing the predicted impacts of each option on each of the objectives. It is very unlikely that any option will be rated best on all measures; instead, one design will be the cheapest, another will have the highest capacity, another will have the least impact on the environment, and another will be the easiest to construct. Which alternative is the best is dependent upon how much weight is placed on each objective.

Choosing the best project will therefore ultimately require making judgments concerning the relative importance of the various objectives and the validity and uncertainty of the evaluation process. Unless there is a single individual with authority to make all design decisions, selecting the best project will inherently be a political process: people with different perspectives and agendas will have to work out a process to determine the best way to proceed.

There are many methods that can be used to help structure the political process. Weighting schemes can be developed and applied to each of the criteria. However, weighting schemes may appear to be more objective than they really are. This is not an argument against using weights, and it is certainly not an argument against having multiple criteria. However, it is a caution to avoid thinking that it will be easy to agree upon criteria or weights or to think that the public will agree with whatever criteria or weights are proposed by any of the parties.

Choosing weights is simply another way of making judgments. In situations where there is general agreement concerning a) the options to be considered, b) the relevant criteria, and c) the relative importance of the criteria, a structured weighting scheme can be helpful in ranking the options. However, weighting schemes will not be of much use if there is strong disagreement about which alternatives should be considered, which criteria are most important, and how impacts should be measured. Presenting a weighting scheme as a way to obtain an "objective ranking of the options" will, in such cases, be impossible, as the different groups will simply push for weights that favor their own preferred options. Difficulties in reaching consensus will be exacerbated if many of the criteria are highly qualitative.

The role of the analyst is clear: provide the best information possible within the available time and budget; identify what is certain, what is likely, and what is possible. Let people know how much faith you put in each part of the analysis. Explain the assumptions, and indicate whether or not you or your expert colleagues believe such assumptions to be reasonable. Where have you had to make guesses, say where those guesses are most uncertain. You also want to make sure that the range of options is wide enough to cover the major strategies that might be considered. Finally, in public meetings or in private meetings with stakeholders, you can try to ensure that the discussion deals with the actual data and credible options, and you can try to provide insight as to the cost-effectiveness of various proposals with respect to various criteria.

Seeking Public Input

Input from the public and key stakeholders can be helpful in identifying how to look at a complex problem. Input will be most helpful – and the exercise is likely to be most productive – at a point in time when there is general awareness that something needs to be done about a problem, but no one knows (or thinks they know) what is best to do. The main reasons for seeking input from the public and from stakeholders are to clarify the nature of the problem, to identify potential alternatives or strategies for dealing with the problem, and to discuss the relevant criteria for selecting and evaluating specific alternatives. Preliminary discussions can also be very useful in identifying where there is consensus about needs and opportunities, where additional information is needed, and where potential controversy will be most likely. This feedback will be helpful in allocating research and planning resources and in determining how best to structure the process.

For example, many projects are aimed at promoting development while at the same time improving living conditions for people living in a region and protecting the environment. Different groups of people would likely place much different weights on the costs and benefits related to such things as regional development, median income for residents, pollution, and loss of open space. Possible types of projects might include investments in new infrastructure, upgrading technology, providing regulatory incentives, increasing awareness of best practices, or changing governmental structure or policies. Input from citizens could help prioritize the problems and place weights on various costs and measures, thereby providing a framework for seeking the most effective strategies. Consultants could conduct workshops on the issues and facilitate discussions among people representing a wide range of interests and expertise. An informed examination of the needs, problems, and strategies for dealing with them may lead to approaches and projects that are less controversial, more responsive to the needs of the region, and better for the environment.

Summary

Large projects will always be trying to meet multiple, sometimes conflicting objectives. Sometimes there will be obvious ways to define economic or financial criteria that capture the essence of a project. More often, it will be necessary to make assumptions concerning the monetary value that might be associated with some of the impacts of a project. And there may be times when the problems and potential solutions are primarily concerned with factors such as aesthetics or equity that are inherently difficult to quantify.

Expressing Costs and Benefits in Monetary Terms

If the most important objectives can be expressed in economic terms, then it will be possible to create a single monetary estimate of the costs and benefits associated with a project. Many private sector projects are primarily concerned with financial matters: if this project proceeds, will the revenues be sufficient to cover the investment and operating costs? Many public projects will be concerned with broader economic benefits, such as economic growth, job creation, and average income. While these economic benefits are not necessarily readily tied to the project or to specific individuals, they can certainly be expressed in monetary terms.

Comparing Present and Future Costs and Benefits

When evaluating costs and benefits over the life of a project, it is necessary to discount future costs and benefits for comparison with present costs and benefits. By using a discount rate, any future cost or benefit can be reduced to an equivalent present value. By summing all of the discounted costs and benefits, it is possible to obtain the net present value (NPV) of a project. If the NPV is positive, then the project provides net benefits and may be worth pursuing.

Cost Effectiveness

Cost-effectiveness is a very useful concept when dealing with criteria that are quantifiable but difficult to monetarize. Cost-effectiveness is the ratio of the cost of a proposed project to the change in performance. So long as it is possible to quantify a key aspect of performance, such as risk or capacity or service quality, it should be possible to evaluate alternatives by considering their cost-effectiveness in improving this aspect of performance. Cost-effectiveness will be less relevant in projects where there are multiple objectives, so that it is impossible to focus on just one critical aspect of performance.

Weighting Schemes for Projects with Multiple Criteria

When there are multiple objectives, some of which cannot be monetarized, then the evaluation cannot focus on a single metric. Instead, some kind of weighting scheme – whether objective or subjective - will have to be used to compare alternatives. If a decision is made by an individual or by the vote of a committee or by a referendum, then each person involved can make their own subjective judgment in determining their preferred option. Often there will be a structured process for making the decision that requires participants to consider multiple options, to consider impacts upon various objectives, and to follow a specific procedure for ranking the options. If so, then some sort of a weighting scheme may be helpful. With a weighting scheme, it is possible to collapse any set of multiple measures into a single measure of effectiveness.

It is critical to remember that there is no objective way to determine the "correct" weighting scheme or even the selection of the "proper" criteria. The choices of what criteria to consider and how much weight to give to each one could cause intense debate among those trying to address a problem or those trying to promote a particular project.

Using Public Input in Identifying and Evaluating Projects

Input from the public and/or stakeholders can be helpful in identifying and evaluating projects. Public input can help in clarifying the nature of the needs or problem, as well as in identifying measures or criteria that can be used to define the problem and to evaluate potential solutions. In some cases, very rough measures of potential impacts may be useful in reaching consensus about what to do, how to do it, and when to begin.

Resource: Project Evaluation: Essays and Case Studies
Carl D. Martland