USGS
science for a changing world

The National Map

A Cost-Benefit Analysis of *The National Map*

www.nationalmap.gov

By David Halsing, Kevin Theissen, and Richard Bernknopf

Circular 1271

U.S. Department of the Interior
U.S. Geological Survey

A Cost-Benefit Analysis of *The National Map*

By David Halsing, Kevin Theissen, and Richard Bernknopf

Circular 1271

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
Gale A. Norton, Secretary

U.S. Geological Survey
Charles G. Groat, Director

U.S. Geological Survey, Reston, Virginia: 2004

For sale by U.S. Geological Survey Information Services
Box 25286, Denver Federal Center
Denver, CO 80225

This report and any updates to it are available online at:
http://pubs.usgs.gov/circular/c1271/

Additional USGS publications can be found at:
http://geology.usgs.gov/products.html

For more information about the USGS and its products:
Telephone: 1–888–ASK–USGS (1–888–275–8747)
World Wide Web: http://www.usgs.gov/

Cataloging-in-publication data are on file with the Library of Congress (URL http://www.loc.gov/).

Produced in the Western Region, Menlo Park, California
Manuscript approved for publication, May 5, 2004
Text edited by James W. Hendley II
Layout by Susan Mayfield

Executive Summary

The Geography Discipline of the U.S. Geological Survey (USGS) has conducted this cost-benefit analysis (CBA) of *The National Map*. This analysis is an evaluation of the proposed Geography Discipline initiative to provide the Nation with a mechanism to access current and consistent digital geospatial data. This CBA is a supporting document to accompany the Exhibit 300 Capital Asset Plan and Business Case of *The National Map* Reengineering Program.

The framework for estimating the benefits is based on expected improvements in processing information to perform any of the possible applications of spatial data. This analysis does not attempt to determine the benefits and costs of performing geospatial-data applications. Rather, it estimates the change in the differences between those benefits and costs with *The National Map* and the current situation without it. The estimates of total costs and benefits of *The National Map* were based on the projected implementation time, development and maintenance costs, rates of data inclusion and integration, expected usage levels over time, and a benefits estimation model.

The National Map provides data that are current, integrated, consistent, complete, and more accessible in order to decrease the cost of implementing spatial-data applications and (or) improve the outcome of those applications. The efficiency gains in per-application improvements are greater than the cost to develop and maintain *The National Map*, meaning that the program would bring a positive net benefit to the Nation. The average improvement in the net benefit of performing a spatial data application was multiplied by a simulated number of application implementations across the country. The numbers of users, existing applications, and rates of application implementation increase over time as *The National Map* is developed and accessed by spatial data users around the country.

Results from the "most likely" estimates of model parameters and data inputs indicate that, over its 30-year projected lifespan, *The National Map* will bring a net present value (NPV) of benefits of $2.05 billion in 2001 dollars. The average time until the initial investments (the break-even period) are recovered is 14 years. Table ES-1 shows a running total of NPV in each year of the simulation model. In year 14, *The National Map* first shows a positive NPV, and so the table is highlighted in gray after that point. Figure ES-1 is a graph of the total benefit and total cost curves of a single model run over time. The curves cross in year 14, when the project breaks even. A sensitivity analysis of the input variables illustrated that these results of the NPV of *The National Map* are quite robust. Figure ES-2 plots the mean NPV results from 60 different scenarios, each consisting of fifty 30-year runs. The error bars represent a two-standard-deviation range around each mean.

The analysis that follows contains the details of the cost-benefit analysis, the framework for evaluating economic benefits, a computational simulation tool, and a sensitivity analysis of model variables and values.

Table ES-1. Yearly results for mean net present value of baseline scenario.

[The cumulative mean net of benefits and costs are presented, with negative net present value (NPV) in parentheses. These are mean values because 30 model runs were simulated to capture model variance. In year 14, the NPV becomes positive, and gains accumulate thereafter (in gray). The final mean and standard deviation NPV are included at lower right]

Model Year	Mean Net Present Value
1	$ (30,000,000)
2	$ (58,986,986)
3	$ (86,851,952)
4	$ (113,449,315)
5	$ (138,063,308)
6	$ (159,147,477)
7	$ (171,579,633)
8	$ (176,992,427)
9	$ (161,274,589)
10	$ (155,936,276)
11	$ (120,019,163)
12	$ (69,591,309)
13	$ (11,973,403)
14	$ 62,874,333
15	$ 134,560,280
16	$ 241,066,196
17	$ 362,467,550
18	$ 461,986,307
19	$ 558,838,273
20	$ 674,433,835
21	$ 806,040,161
22	$ 948,599,229
23	$ 1,070,326,997
24	$ 1,200,519,950
25	$ 1,337,230,351
26	$ 1,472,917,954
27	$ 1,591,170,175
28	$ 1,748,433,401
29	$ 1,861,787,277
30	$ 2,051,719,005
The ending Mean and Standard Deviation (in 2001 dollars) are:	
Mean NPV	StDev NPV
$ 2,051,719,005	$ 492,308,136

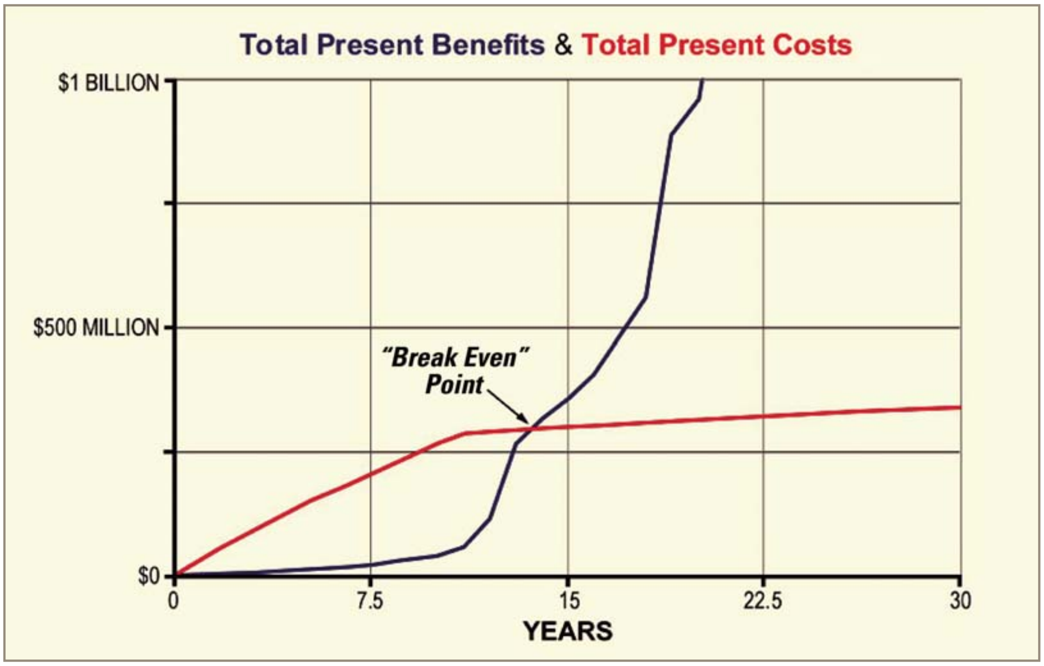

Figure ES-1. Total cost and total benefit curves of *The National Map*. A graph of the accumulating benefits (blue) and costs (red) of *The National Map* over 30 years. The break-even point is where the curves cross (year 14). The kink in the cost curve is the change from development costs to maintenance costs only. Note the slow upward climb of the benefit curve until the project nears completion, when most of the country's spatial-data users have had time to see valuable data, adopt *The National Map* as their data source for existing tasks, and innovate new applications for spatial data. The area above the cost curve and below the benefit curve between years 14 and 30 represents the net present value (NPV).

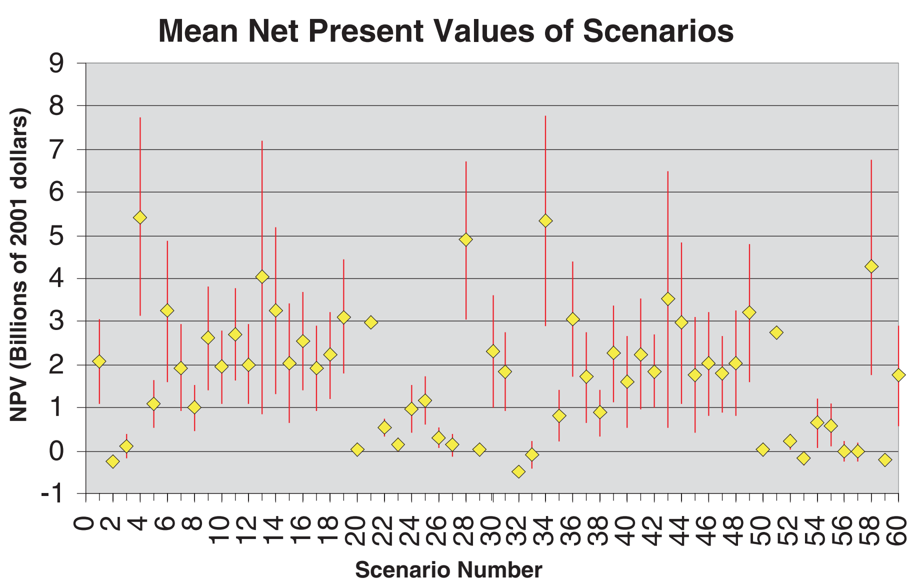

Figure ES-2. Sensitivity analysis results; mean net present value by scenario. The 60 different scenarios were each run 50 times. Results from each scenario are plotted here, as the mean net present value (NPV, yellow diamond) and the two-standard-deviation range (approximately the 95-percent confidence interval; red bars). Note that not all scenarios reach a positive value.

Acknowledgments

Several people made critical contributions to this project. Dr. Anne Wein, a volunteer operations research analyst, provided critical insights into formulating the mathematics of the theoretical framework. Amy Mathie contributed considerable time and energy to the early phases of defining and scoping this analysis. Members of the Program Management Section—including Dale Russell, Bill Coronel, John Fisher, and Janet Goodman—assisted us in arranging and conducting the interviews of GIS managers and contractors, from which we were able to derive much of our data. Dr. Stephen Gillespie generously reworked survey numbers for inclusion into our modeling. He also gave significant feedback and helpful criticism to both the theoretical and computational model development. We are deeply indebted to two technical peer reviewers, Carl Shapiro and Donald Bieniewicz, whose extremely detailed comments and insights greatly improved the final analysis and report. Finally, we appreciate the assistance given by Linda Hicklin and Joan Sziede in scoping the project.

Contents

Appendices

Tables

Figures

A Cost-Benefit Analysis of *The National Map*

By David Halsing, Kevin Theissen, and Richard Bernknopf

Introduction

History and Purpose

The Office of Management and Budget requires an Exhibit 300 Capital Asset Plan and Business Case (hereafter referred to as the "Exhibit 300") for the U.S. Geological Survey Geography Discipline's effort, *The National Map*. The Exhibit 300 requires submission of a cost-benefit analysis (CBA) of each proposed program. Western Geographic Science Center staff conducted this CBA in support of the Exhibit 300. Work on this CBA was begun in October 2002 and was completed in July 2003.

Organization of the Report

This report is organized in the following manner. Section 1 is a review of the literature involving the use and value of spatial data and Geographic Information Systems (GIS). Section 2 presents the conceptual framework that underlies the benefits-estimation model at the heart of this CBA. Section 3 explains the data gathering methods and outlines the computational model developed to simulate the spread of spatial-data use and the benefits that would ensue. Section 4 contains primary results and provides an overview of the sensitivity-analysis methods and results. Section 5 is a discussion of the major results and implications for *The National Map* program. In addition to the main report, several appendices have been compiled to provide the details of the analytical approach. These describe the economic theory (appendix A), the data gathering steps (appendix B), the model calculations and baseline parameters (appendix C), and the exact inputs for the sensitivity analysis (appendix D).

Fitting this Analysis into the Exhibit 300 Report

This analysis was initiated in support of an Exhibit 300 Report for the Office of Management and Budget that initially addressed the entire program known as *The National Map*. The Reengineering Program is essentially the information-technology (IT) component of the larger vision of *The National Map*. A more specific Exhibit 300 was created to solely address this Reengineering Program. This report accompanies that more specific Exhibit 300. Despite the narrower subject matter of the new Exhibit 300 Report, this cost-benefit analysis covers the entire program that will eventually produce *The National Map*.

There is a reason for this apparent discontinuity. The scope of the analysis was determined on the basis of the inseparability of the IT component from the other necessary components of *The National Map*, that is, the problem of joint costs. The IT portion is absolutely necessary to produce the benefits of *The National Map* but on its own is insufficient to do so. The Reengineering Program carries a clear and definable share of the cost of developing the larger *National Map*, yet it is impossible to evaluate the proportion of total net benefits that will arise from only this IT component.

A method employing a simulation approach was developed to evaluate all components of *The National Map*, instead of solely the IT component. It is suggested that the Exhibit 300 Reports for other required components of *The National Map* include results from this CBA, recognizing that, while they each bring their own definable share of the program costs, the benefits of *The National Map* arise only from their inseparable contributions.

1. Literature Review

The collection, integration, and distribution of digital spatial data at the scale and extent of the USGS effort called *The National Map* is a relatively new occurrence. As such, the literature on direct analogs of *The National Map* is comparatively rare. Studies were conducted for the British Ordnance Survey (OXERA, 1999) and the Australia-New Zealand Land Information Council (Price Waterhouse, 1995), but these are not exact comparisons. Instead, much of the available literature on spatial data focuses on the use of Geographic Information Systems (GIS) as compared to using paper maps or no spatial data at all. For example, the two studies mentioned above lean heavily on subjective valuations of GIS/no-GIS alternatives. Nevertheless, these types of studies constitute the bulk of the literature upon which this cost-benefit analysis necessarily relied.

Early in the growth of GIS, the need to measure the economic feasibility of GIS implementation was recognized and documented in the literature. Most of the economic studies take a CBA approach and were conducted internally by local, regional, or national governments and (or) by hired consultants (see for example, Center for Technology in Government, 1995; Ledbetter and others, 1997; Hardwick and Fox, 1999; Montgomery County, 1999; Price Waterhouse, 1995; OXERA, 1999). These case studies and project assessments are sometimes scorned in the literature for being anecdotal, but they serve as important "building blocks" to our understanding of the costs and benefits of *The National Map*.

The costs of implementing GIS in an organization are well understood and are not difficult to quantify by most accounts (see Korte, 1996, for an example of GIS implementation costs). Costs include data, equipment, software, and other supplies, training, operating costs, systems development, and system maintenance. These costs can be substantial (that is, millions of dollars annually) particularly the initial costs for equipment and later costs of data maintenance and systems development.

Several different types of GIS benefits are discussed in the literature. Most of these benefits fall under two broad categories termed *efficiency* and *effectiveness* benefits (Prisley and Mead, 1987; Gillespie, 1994). Efficiency benefits (often called "cost-savings" or "costs-avoided" benefits) are those that arise when the addition of a GIS reduces the costs to perform existing tasks in an organization. This is most often achieved by reducing the amount of time needed (or even the necessity) to perform a particular task. For example, once topographic maps are in digital format and can be distributed electronically there is less need for staff time to produce, maintain, and distribute paper maps. In several published studies efficiency benefits are directly measured in terms of the number of staff hours/salary that are saved by the addition of a GIS (see for example, Korte, 1996; Baltimore County, 2001). Effectiveness benefits (sometimes called "value-added" benefits) are those that arise from improvements to existing tasks or the addition of new tasks that could not be performed prior to the implementation of the GIS. For example, land use/land cover maps that are updated monthly in a GIS database could improve environmental decision-making by visualizing seasonal and interannual trends and patterns that were not apparent in a series of paper maps that were updated every 15 to 20 years.

Quantifying the full benefits of a GIS is a difficult task. Although it can be relatively straightforward to calculate some of the efficiency benefits directly, effectiveness benefits are very difficult to estimate and require a more rigorous analysis than is found in a traditional CBA (Dickinson and Calkins, 1988; Gillespie, 2000). Accordingly, some of the economic evaluations of GIS in the literature make a quantitative estimate of the efficiency benefits that a GIS can potentially bring to an organization. However, only a few studies have made quantitative estimates of any of the effectiveness benefits (see for example, Bernknopf and others, 1993; McInnis and Blundell, 1998), and we are aware of no comprehensive study of the effectiveness benefits that GIS brings to an organization. Instead, effectiveness benefits are typically treated as an important, qualitative "bonus" that can be added to the more easily quantified cost-savings resulting from the implementation of a GIS (see for example, Hardwick and Fox, 1999).

In the past it appears that this methodology has been accepted because the efficiency benefits (cost savings) alone are usually sufficient to give a positive net benefit and justify investment in the GIS. Even when this is not the case it is simply assumed that the full benefits of a GIS will outweigh the costs—and most economic studies of GIS (whether rigorous or not) conclude that positive net benefits arise from the implementation of GIS technology in an organization[1].

Estimation of the effectiveness benefits of a GIS is potentially a time-consuming and expensive process. This generally involves developing a metric for outcomes of many tasks that are, by their nature, non-market goods. One study found that a city in Florida used a computer-aided GIS to reroute their ambulance dispatch system and trimmed the response time by 1 minute per call (referenced in OXERA, 1999). To put a dollar value on that reduced response time is significantly more difficult and generally requires a separate study. For example, an analysis of the ambulance routing program developed an estimate of $150,000 in savings per year.

Without some measurement of these benefits, however, a CBA will greatly underestimate the full benefits that will arise from the implementation of a GIS. Early studies of the economic value of geologic information by Bernknopf and others (1988a,b) examined effectiveness benefits in losses avoided from landslide hazards and showed the potentially large benefit that can be derived from geospatial data. Bernknopf and others (1993) later developed an economic model to estimate the value of geologic map information. They converted a geologic map of Loudon County, Virginia, into digital format and estimated the benefits that would accrue from an improved mapping resolution for two applications—the siting of a county landfill and the selection of an appropriate transportation corridor. They found that significant benefits arise from the improved resolution of data from an updated version (1992) of the Loudon County geologic map originally made in 1963. The benefits, estimated at $1.28 to 3.50 million, were measured in avoided costs of landslides and property value loss.

Gillespie (1997, 2000) later developed a model for users to make a relatively quick estimate of the full benefits of a GIS. The authors of one study (McInnis and Blundell, 1998) used the model to estimate the benefits arising from a small number of applications, and they speculate that the model is a useful tool to be considered for more complete analyses. The main strength of the model is that it offers a method to estimate effectiveness benefits in an inexpensive manner. Yet, the authors also noted that the selection of some of the values for the variables in the model can be subjective and that the model is overly sensitive to some of these values.

In addition to efficiency and effectiveness benefits, the rates of usage and diffusion of publicly available geospatial information are crucial to an understanding of the benefits arising from *The National Map*, but these are also poorly understood quanti-

[1]An exception to this is the study of Bond (2000) who notes that whereas GIS has great potential it has primarily been used for operational tasks that have not produced a positive net benefit.

ties. The rate of diffusion of information and communication technologies is often modeled as an S-shaped curve, in which slow diffusion of information occurs prior to a period of rapid diffusion, which is in turn followed by a long, leveling-off period (Rogers, 1962; Hwang, 2002). Knowing the slope and start time of the rapid diffusion of information and communication technologies is critical to determining its benefits, but predicting these two pieces of information is difficult if not impossible. Often, the diffusion of technology turns out to be slower than models and other predictive tools might suggest. For this reason, some researchers recommend caution in the use of diffusion models as predictive tools (see for example, Lennstrand, 1998).

There is clearly a need for further theoretical and academic study of the benefits arising from geospatial information and the implementation of GIS in an organization. The development and implementation of *The National Map* provides an excellent opportunity to help address this need. In the present study, we have developed a model that can be used to estimate the benefits that arise from *The National Map*, and we report the results of our analysis.

2. Framework for Cost-Benefit Analysis of *The National Map*

Introduction

This section contains the methodology for benefits estimation. The method makes clear distinctions between several important aspects of the use of spatial data. Specifically, it:

- Distinguishes the costs and benefits of spatial data itself from those of the applications of the data,
- Compares the state of the world with *The National Map* from that without it,
- Recognizes that uses of spatial data are likely to increase over time, in part as a function of *The National Map*, and
- Accounts for variation in the baseline ability of customers to use *The National Map* data.

The purpose of this policy evaluation is to assess the costs and benefits of *The National Map* and to estimate its net present benefit to society. A positive net present benefit would mean that the costs of developing and maintaining *The National Map* are less than the benefits society reaps from its existence. These benefits come from many sources that involve the use of spatial data in a variety of public and private applications. These benefits include:

- Lower costs of developing and populating a GIS.
  - Less redundant data collection.
  - Reduced private purchase of data.
- Lower costs of carrying out tasks that require spatial data.
  - Reduced staff time because of data currency, availability, consistency, completeness, and other factors.
- Improved quality of decisions made with the data.
  - Increased quality of data (for example, current, seamless, consistent).
  - Uniform availability of descriptive metadata.
  - Instantaneous, constant, and widespread data availability.
- Increased use of spatial data.
  - Greater numbers of entities using spatial data in existing applications.
  - New applications will develop as a function of better data availability.

The key policy questions are:

(1) Does *The National Map* increase the efficiency and (or) effectiveness of using spatial data?

(2) Does *The National Map* increase the frequency of spatial data use in existing applications?

(3) Does *The National Map* allow some who do not currently use spatial data to do so?

(4) Does *The National Map* foster the development of innovative applications for spatial data?

(5) Will these contributions of increasing the type, volume, and efficiency of spatial-data utilization justify the increased cost of the program?

Basis for Comparison

The Exhibit 300 Report for *The National Map* outlines three alternative policy proposals:

- Alternative 1—A fully implemented version of *The National Map* is built over 10 years at an annual cost of $25 million per year over 2001 funding levels.
- Alternative 2—A version of *The National Map* is built on the basis of 2001 funding levels by diverting other USGS Geography Discipline funding into the program. The description of alternative 2 states that in this case, no guarantees could be made about how long it would take to develop *The National Map*, the degree to which it would be completed, how consistent its data would be, and so on.
- Alternative 3—No attempt to build *The National Map* is made. The USGS Geography Discipline goes about “business as usual” as it was in 2001, the year before work began on *The National Map*.

The basis for this cost-benefit analysis, therefore, is to compare the differences in the states of the world that would result from achieving alternative 1 versus alternative 3. As such, it considers only the difference between the costs and the benefits of *The National Map* program, assuming that all other aspects of the Geography Discipline remain the same. We believe that this comparison adequately addresses the states of the world with and without *The National Map*, which is the main goal of a cost-benefit analysis. However, it is important to develop fuller descriptions of alternatives 2 and 3 and to include discussions of other alternatives to a fully implemented *National Map*.

Alternatives to *The National Map*

This analysis assumes that alternative 3 would see no significant changes to the net present value to society of the USGS Geography Discipline, either in terms of type, quality, distribution, or usage levels of spatial data. This constant value is represented in figure 1 by the dashed horizontal line. Conversely, the solid curved line represents alternative 1. The dip in NPV in early implementation years arises from the added cost of developing *The National Map*, but the gains thereafter cause the value to society of the Geography Discipline to break even and then increase. The dashed curved line represents alternative 2, which lies somewhere between the other two. We did not analyze alternative 2, and its shape, trajectory, and value as represented here are purely speculative.

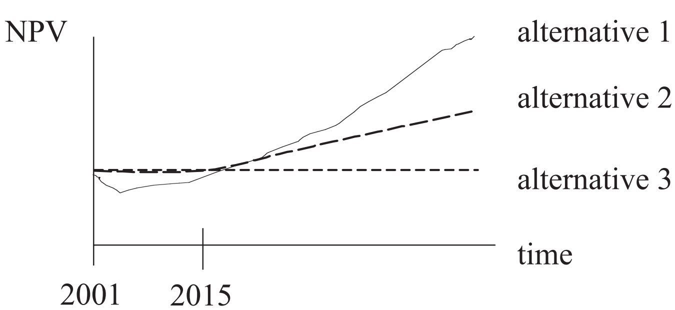

Figure 1. Values of spatial data over time. Heuristic diagram showing trajectories of value of U.S. Geological Survey Geography Discipline data over time under 3 alternatives proposed in the Exhibit 300 Report.

In reality, however, the horizontal line for the status-quo, alternative 3, world is not accurate. The value of the Geography Discipline's products would more accurately be expected to decrease over time. The Topographic Map Series, a signature USGS product, is 23 years old on average, and getting older, due to inadequate funding and staffing levels. Thus the maps become more out-of-date and less accurate each year. Further, users of spatial information are increasingly turning toward digital technologies like GIS and satellite imagery, all of which force the USGS to "go digital" or become obsolete. Finally, the Federal Geographic Data Committee (FGDC) is producing the Geospatial One-Stop (GOS), a standardized digital clearinghouse and web portal that has the goal of unifying all Federal data under one electronic banner. Failing to participate in GOS would make USGS spatial data and maps less relevant than they are today.

In figure 2 the NPV line for alternative 3 has been drawn with a downward curve to capture this decrease in value over time. In the interest of parsimony, we did not explicitly model this source of benefit from *The National Map*, this "obsolescence avoided," in our cost-benefit analysis. We conservatively assumed no loss of value in the non-*National Map* world.

Other alternatives for ameliorating the obsolescence problem do exist, however. The National Atlas is a freely available on-line database and viewer of USGS spatial data holdings (see http://www.nationalatlas.gov). Work on the National Atlas began in 1997 with the recognition of the need to transition to digital, on-line data service. It has some elements that are similar to what *The National Map* would ultimately provide, including a map viewer with user-selectable extent and layers, and some of its data is being incorporated into *The National Map*. However, like other USGS data products, there is insufficient funding to keep it current and accurate, and its decline in value is assumed to proceed along with the topographic

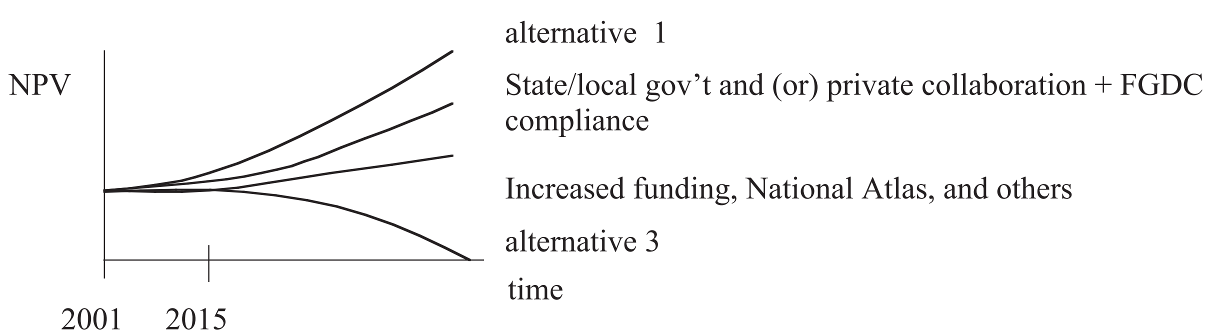

Figure 2. Possible future values of spatial data over time. Heuristic diagram illustrating trajectories of value of U.S. Geological Survey Geography Discipline data over time under various management strategies.

map series. Though certainly people are using data from the National Atlas[2], its utility is assumed to decline as it ages.

There is also the possibility of collaborations between private entities and (or) other government (including state, local, or other federal) efforts to self-organize into something like *The National Map*. This is certainly possible, and the FGDC has been working to develop the National Spatial Data Infrastructure (NDSI) to encourage this dynamic. However, we assumed that these other attempts would be likely to take longer to develop, achieve less complete coverage, and could fail to deliver the free-access that is essential to the public-good component of *The National Map*. For example, FGDC's NSDI effort began in 1990 and is not yet complete. Other examples of collaborative efforts include the Open GIS Consortium, the National States Geographic Information Council, and the National Association of Counties.

It therefore seems a valid assumption to conclude that, although substitutes for *The National Map* could eventually exist, they would not be likely to bring as much value to society as would the rapid and complete implementation of *The National Map* through alternative 1 funding.

Assumptions

Our efforts to model the institutional dynamics that address these questions require several broad simplifying assumptions. The list below states a few of these assumptions before the theoretical model is described (further discussion of these and other modeling assumptions can be found in the relevant chapters and appendices):

(1) For the purposes of this analysis, *The National Map* program consists of the fully implemented *National Map*, which is alternative 1 from the original Exhibit 300 Report. We analyze the costs and benefits of the differences between that program and what existed prior to its inception, which is alternative 3. Although the Geography Discipline of the U.S. Geological Survey has already begun to develop *The National Map* with no additional funding, as described in alternative 2, we compare the "with" and "without" cases.

(2) We assume that *The National Map* can be built and maintained as described in the Exhibit 300 Report and other vision documents. We also assume that whatever budget increase is received to produce *The National Map* will allow it to be completed over some time period that is dependent on the level of funding. Should little increased funding be available, it will delay the completion of *The National Map* but not prevent it.

(3) The sources of the benefits in this analysis come from the eight primary layers of geospatial data as described in *The National Map* vision documents. There are likely to be other layers of spatial data and products added to *The National Map* over time, but this analysis does not consider those because they are unknown at the present time.

(4) The value of digital spatial data comes from its use in a variety of "applications." The net benefit of implementing an application can be improved with better and more readily accessible spatial data. That is, having newer, better, more plentiful, or more available data allows spatially dependent applications to be implemented more efficiently, more effectively, or both. *The National Map* would provide this better data, and can be credited with some of the resulting improvements of spatial data applications. If data from *The National Map* is not used in an application, no benefit can be credited to it.

(5) The various applications of digital spatial data can be grouped into three categories, or "tiers," on the basis of how complex they are and how much technical sophistication is required to perform them:

  (a) Tier 3 applications are the simplest and are those done with paper maps and (or) simple GIS overlays.

  (b) Tier 1 applications are the most complex uses, involving spatial statistics and mathematical modeling.

  (c) Tier 2 applications are moderately complex applications and lie between tiers 1 and 3 in terms of complexity.

(6) *The National Map* will have 3 types of users/customers that are allocated to tiers of capability to utilize the data layers it provides. Each tier contains users that are described as being in one of three "states of nature." These states of nature reflect the user's current level of facility with GIS and digital spatial data:

  (a) Tier 1 users are the most sophisticated users and have access to applications from all three tiers.

  (b) Tier 2 users are moderately sophisticated users and have access to tier 2 and tier 3 applications.

  (c) Tier 3 users are the least sophisticated users and can only implement tier 3 applications.

(7) Tier 1 users have the greatest number of applications in which they can currently make use of digital spatial data, because they can make use of applications from the lower two tiers. They have the potential to receive the highest total (as opposed to marginal) benefit from using this data. Tier 3 users have the fewest number of available applications because they can only implement tier 3 applications. Tier 2 users lie between these two extremes, making use of tier 2 and tier 3 applications.

(8) The spatial unit of analysis is the county. This does not mean we examine only county government uses of spatial data. Rather, it means that we use counties as the unit of averaging and aggregating uses of digital spatial data. While recognizing that not all uses of *The National Map* are at the county scale, in order to build a feasible model, we make the simplifying assumption that they are. Applications with the spatial extent similar to a

[2] The National Atlas web site satisfies an average of 3.2 million requests for services each month (Donald J. Bieniewicz, written commun., 2003).

county are likely to use some of the base-map information *The National Map* will provide, whereas city uses might require data more specific than that. Also, the large number of counties (3078) in the United States provides a large number of units over which to integrate. Where possible, we scale other uses of data—including private as well as public—up or down to the county level. Federal uses of spatial data are handled separately in the computational model, as will be discussed more in sections 3 and 4.

Together, Assumptions 5 through 7 illustrate several key points. Tier 1 data users are able to make use of digital spatial data in many more applications than tier 3 users are. This adds more benefits to *The National Map* per tier 1 user than per tier 3 user. This is especially true of the novel applications that are likely to grow out of a data source as complete and current as *The National Map*.

A part of the USGS goal is to assist the transition of spatial data users from tier 2 and 3 into tier 1 by providing data and lowering the barriers to using digital spatial data. This is one of the important pure public good aspects of *The National Map*. Solely providing a tier 3 county with a GIS database is something close to a direct subsidy, offsetting their direct costs of developing and populating their GIS database. Benefits to society do derive from this. If, however, by providing the data layers in *The National Map*, the USGS also helps a county develop into an efficient user of spatial data in a variety of more sophisticated applications, thereby allowing innovative new applications to be developed and used, then it is an example of the Federal Government providing a public good that brings greater positive net benefits to all.

Below, we derive a quantitative framework for enumerating and estimating these benefit streams and how their rates of accrual determine the net benefit of *The National Map*. The next section explains the major steps of the model.

The Economic Model for Estimating Benefits

Values Come from Applications

The benefits of *The National Map* come from the value of information as its data is used to permit, facilitate, or improve some public or private decision or process. These benefits either can be in terms of obviating the need for users to generate, populate, or operate their own GIS database (cost savings; efficiency gains) or through making it faster or more effective to carry out tasks and projects (value added; effectiveness gains). Other value-added benefits could come from development of all-new uses of spatial data that are currently impossible or inefficient because of a lack of consistency, currency, or completeness of spatial data.

Because the benefits of *The National Map* are based on expected improvements in processing information for applications of spatial data, the willingness to pay for the program is based on its utility as a provider of that data. The value added of *The National Map* originates from improvements in projected implementation time, reduced development and maintenance costs, faster rates of data inclusion and integration, and increased expected usage levels over time for geospatial applications. This is a derived demand for the on-line distributed database of *The National Map* and assumes that data are current, integrated, consistent, complete, and more accessible to produce the desired outcome of an application. Appendix A contains a derivation of the economic foundation of a derived demand for *The National Map* and provides the theoretical support for the benefits estimation procedure.

The benefits of an application are the net present value of a partner's ability to improve a decision's effectiveness or efficiency with the use of spatial information or to use spatial data in a way that was not previously feasible. Throughout this document, "application" means a single type of use for a set of digital spatial data provided by *The National Map*. Examples of applications include:

- Creating an emergency evacuation plan.
- Designating critical habitat for an endangered species.
- Conducting property tax assessments.
- Researching land-cover change and deforestation.
- Verifying insurance claims.

Applications include any use of spatial data, whether by government or the private sector, by individuals or organizations. However, conducting the same application annually for 10 years does not create a list of 10 applications. Instead, it signifies there are 10 implementations of a single application over time. Similarly, an application conceptualized and enacted in one city, then copied by 6 others does not mean there were 7 applications. It means there were 7 implementations of a single application.

The value (or net benefits) of *The National Map* program is equal to the benefits of the information it provides minus the cost of the program, as is stated in equation (1) below. In this analysis, the program is the on-line distributed database as described in the original Exhibit 300 Report. Cost, represented by a sum of discounted cash flow of costs, C_{TNM}, is based on estimates of implementation and development costs.

The B_{TNM} encompasses the discounted benefits that come only from the parts of the larger *National Map* program that partners, customers, and the general public see on their computers, can download and use in some set of applications.

In equation (1), NB_{TNM} represents the net benefit of *The National Map*, which is synonymous with the value of information in *The National Map*, and $-C_{TNM}$ and B_{TNM} are the total costs and benefits of *The National Map.*

$$NB_{TNM} = -C_{TNM} + B_{TNM} \tag{1}$$

Here, costs are only the costs of *The National Map*, not of the applications in which it is used. Application costs are addressed directly in other portions of the model. The benefits are also written as the benefits of *The National Map*, because they refer to the benefits of *The National Map* program, not of the applications themselves. Again, these benefits of *The National Map* must come from its data being used in an application. An application implemented by a local entity using only its internal data derives no value from *The National Map*.

Deriving a Net Present Benefit

The benefit of *The National Map*, B_{TNM}, is a discounted summation of the changes in net benefits of the applications. Initially, assume that all applications for *The National Map* data already exist and are known. Taking into account the tier structure, note that some applications require more sophistication than others to implement and that users may not have the technical capacity or human expertise to implement all existing applications. In our model, tier 3 users may not use tier 1 or 2 applications, whereas tier 1 users may draw from any of the three tiers of applications. Tier 2 users may use tier 2 or 3 applications. To avoid confusion with the subscript *t* for time, in the equations, *s* refers to user tier state. The unit of time is a year and is an interval during which an application *may* occur.

Define v_{jst} as the net benefit of *The National Map* data as used in an application *j* by user *s* at time *t*. The net benefits come from a number of different sources, but generally fit into the efficiency and effectiveness benefits described in the literature review. At least some of those gains must be rightly attributed to *The National Map* if it is to have any value to society. Those improvements (which can be either gains or savings) can be written:

$$v_{jst} = NB_{jst(TNM)} - NB_{jst(SQ)} \quad (2)$$

$NB_{jst(TNM)}$ are the net benefits of an application occurring in a single place at a specific time with *The National Map* data. Similarly, $NB_{jst(SQ)}$ are the net benefits of a single application in the "status quo," without *The National Map*. The difference between them is the value of *The National Map* in that application, v_{jst}. Expanding the previous equation illustrates how this value is derived.

$$v_{jst} = (B_{jst(TNM)} - C_{jst(TNM)}) - (B_{jst(SQ)} - C_{jst(SQ)}) \quad (3)$$

Depending on the specific application and the place it is being implemented, having *The National Map* available can either reduce the costs of the application (lower $C_{jst(TNM)}$ — an efficiency gain) or increase the benefits that society, a firm, a researcher, or an individual can reap from the application (higher $B_{jst(TNM)}$ — an effectiveness gain). Either of these raises net benefits of the application ($NB_{jst(TNM)}$) by some amount. Then, by subtracting the original, status-quo net benefits [$NB_{jst(SQ)} = (B_{jst(SQ)} - C_{jst(SQ)})$] of the same application, a value can be derived for the incremental value of *The National Map* in that application as implemented by that user in that place.

Implicit in equations (2) and (3) are several assumptions. First, we assume that spatial data applications affected by *The National Map* are independent of each other. That is, the availability of *The National Map* data might increase the frequency or prevalence of application implementations, but does so in an absolute sense and not at the expense of other types of applications. Further, there is an assumption that there is no significant difference between public and private applications. This assumption means that aggregating both of these types of applications at the county scale is a fair simplification of an issue that certainly could be a concern but which is not possible to address at this level of analysis. The issue is that we are considering a social investment and that the cost-benefit analysis should be attempting to measure social benefits.

Ideally, it would be possible to make an exhaustive search for a large number of applications that had been improved with data from *The National Map* pilot projects and determine the difference in net benefits of those applications thus recognized. Those observations would then be used to extrapolate the net benefits that could be reaped across the entire country by a fully implemented *National Map*. Those benefits would be weighed against the costs of developing *The National Map* and a net benefit for the program can be estimated.

However, the current pilot projects of *The National Map* are too new to lend insight or to provide dollar values for evaluation of the on-line distribution system in a set of applications. Instead, we have had to rely on non-*National Map* data holdings and delivery systems that are partial analogs of what *The National Map* will eventually be. There are not enough of these to do a "count" of total benefits. Instead, we have developed an "average improvement in net benefit per application." Let: J_{st} = the total number of existing applications available to users in each tier, *s*, at each time, *t*. These values are expected to grow over time as innovation of new uses for spatial data proceeds. The average improvement in net benefit of an application in a tier at a given time is the summed net benefits of all appropriate applications divided by the number of applications:

$$\bar{v}_{st} = \sum_{j=1}^{J_{st}} \left[(B_{jst(TNM)} - C_{jst(TNM)}) - (B_{jst(SQ)} - C_{jst(SQ)}) \right] \Big/ J_{st} \quad (4)$$

However, within each tier, not all applications are needed or useful in every location. Also, there are different repetition intervals for each application. This means that a different subset of the set of existing applications, J_{st}, is implemented each tier-year. A different set of implementations obviously creates a different average change in net benefit of applications. Let: $\bar{v}^*_{st}$ = the average net benefit of an *implemented* application in a tier, *s*, at time *t*.

The value of $\bar{v}^*_{st}$ changes each year depending on specific implementations, accounted for in this model by ψ_{st}, discussed below. We assume that this variability is normally distributed with mean $\bar{v}_{st}$ and standard $\bar{v}_{st}/4$ deviation . In the simulation model, the exact value for this average change in net benefits is drawn randomly from a normal distribution for each model year. See section 3 for more information.

To derive the number and type of application implementations, ψ_{st}, that occur in each tier, s, during each time period, t, the following definitions are required:

K_{st} = the number of data users in each tier, s, at each time, t. These values are likely to change over time as counties advance upward through the tiers.

P_{st} = the proportion of applications available to users in each tier that are actually implemented in a given time period. This is a function of the needs of individual users—not all counties need all applications (snow removal routing, for example). It is also related to the relevant time interval between iterations of an application—some applications occur annually; others every five years. It is also a function of the cumulative diffusion of existing applications. The exact subset of specific applications varies each year and is randomly generated in the computational model as a proportion of total existing applications. The next section of this report contains more details on how this is derived.

Thus, $\psi_{st} = J_{st} * K_{st} * P_{st}$ is the tier-specific number of applications performed in a single time period.

$C_{TNM(t)}$ is the cost of *The National Map* in time period t. The time horizon of interest is T. The term $(1/(1 + r)^t)$ is used to discount the benefits and costs occurring at some future point, t, by the discount rate, r (according to Federal Government standards), to bring them into present value terms. The equation that summarizes all of above discussion and represents an expression for the net present benefit of *The National Map* is:

$$NB_{TNM} = \sum_{t=1}^{T} \left[\left(\sum_{s=1}^{3} (\overline{v}_{st}^{*} * \psi_{st}) - C_{TNM(t)} \right) * \left(\frac{1}{(1 + r)^t} \right) \right] \qquad (5)$$

The increases in the available pool of applications (innovation), the number of places each one occurs (diffusion), and the change in number of users in each tier does not appear explicitly in the equations above. However, the simulation model we constructed (and which is described in a later section of this report) does have inputs for the rates of diffusion (into the calculation of P_{st}), innovation of applications, and for changes in the tier-user profile.

In general, we expect transitions from one state or tier to take place faster in those counties and other data usage areas that are active partners with USGS in *The National Map* program, but the general transition pattern should remain the same for all users. For example, after a tier 1 user begins using spatial data, *The National Map* provides a means to share J applications among K users, in part because it provides much of the data needed to implement the application. Part of the plan for *The National Map* is also to compile and distribute a “library” of applications implemented by users across the country and lists of the data needed to drive them. This is another reason why *The National Map* will increase rates of diffusion, innovation, and advancement.

On completion of this report, the actual values for many model parameters, values, and rates of change are estimates and extrapolations. This means that, in the short term, our model and results are closer to simulation than actual prediction. We have made guesses at the evolution of the program by selecting reasonable and conservative values for each of the variables in the above equations. As *The National Map* is implemented, we can substitute real values derived from observation into the cost-benefit framework, modify the simulation model, and narrow the range of uncertainties around its projections. With the current level of institutional knowledge and experience, we use the model to predict reasonable outcomes.

3. Data and Modeling Methods, Results, and Sensitivity Analysis

With the analytical framework in place, the next step was to develop a system and a method with which to quantify and account for the changes in the variables over time as *The National Map* evolves. This section first explains the analogs between the framework and the simulation model. It then provides an overview of the simulation model, outlines our data collection methods, and closes with a description of the baseline results, including a sensitivity analysis. There are additional details on these steps in a series of appendices and so they are treated briefly here.

Putting Theory Into Practice

It was first necessary to adapt the framework contained in Section 2 into a workable tool for analyzing the costs and benefits. Recall that, by providing data that are current, integrated, consistent, complete and more accessible, *The National Map* would decrease the cost of implementing spatial data applications and (or) improve the outcome of those applications. The sum of the efficiency and effectiveness gains minus the cost to develop and maintain *The National Map* would be the net benefit of *The National Map*. The initial approach to the analysis was to:

- Develop a long list of applications of spatial data.
- Sort the applications into the three tiers according to their complexity.
- Distribute each county in the United States into one of the three tiers, on the basis of the capacity they have to perform applications of varying complexity.
- Estimate the net benefits (the benefits minus costs) of performing each application of spatial data in the status quo world (that is, where no *National Map* exists).
- Estimate the net benefits of performing that same application in the case where *The National Map* exists and can provide data to assist or facilitate the application.
- The difference between the net benefits in each of these cases is the incremental value of *The National Map*

data, distribution system, and application library to that application.

- Compile these incremental values across space and time, estimating how often each application will occur and in how many places, and discount to present value.

- Take the costs to develop, implement, and maintain *The National Map* over the relevant time horizon and discount to present value.

- The difference between these totals is the net present value of *The National Map*, and represents its value to society over the time horizon of interest.

Several modifications from this initial method were required. First, because neither *The National Map* nor any completed pilot projects existed, we could not make a direct comparison of the "with" and "without" cases. Instead, we had to find examples of organizations that had made transitions similar to those *The National Map* will ultimately facilitate. This included changing from paper maps to a digital geographic information system (GIS), moving from an internal GIS to one distributed freely on the World-Wide Web, going beyond simple overlays and visual analysis to more complex spatial modeling, or increasing the data update cycle. All of these transitions will be enabled or made more accessible to local spatial data holders and users as *The National Map* is implemented. It was our assumption that empirical data from places that had made these transitions would be representative of others that could take place around the country after *The National Map* develops.

Even that modification was not enough, however, to follow the steps above. Despite significant time and effort in reviewing the literature and conducting interviews with city, county, and state GIS coordinators—as well as with *National Map* pilot project leads, Geography Discipline liaisons, and contractors who provide the USGS with data and products—getting reliable numbers for the costs and benefits of implementing spatial data applications before and after one of those transitions was extremely difficult.[3]

Many operating budgets were not set up to track costs by project. In other situations, the project costs were well known, but the portion of those costs borne by the spatial data (acquisition, manipulation, or analysis) was not known. Further, it was difficult to get credible figures on how often or widespread an application was implemented. Often, the post-transition costs could be estimated for a certain use of spatial data, but the organization had failed to keep records of the pre-transition costs.[4] Finally, as difficult as measuring the changes in the cost side of these transitions, there is a greater problem with attributing the beneficial outcomes of an application to the spatial data input. There is little consensus about measuring these "effectiveness" benefits (see the Literature Review and appendix B sections of this report).

Despite these limitations, we found enough examples of changes that addressed either the reduced costs (efficiency) or increased benefits (effectiveness) of making one or more of *The National Map*-related transitions. From these examples, we extrapolated a "mean increase in net benefits per application implemented." For brevity's sake, we often refer to it as "NB of an application," so the "improvement" and "implemented" portions are implied rather than stated. The small number of examples we found necessitated this "averaging" approach. Details of the derivation of the exact dollar amount are in appendix B, and for the reasons explained there, we used an improvement of only $1,000 gain per application implementation.

The Computational Simulation Model: NB-Sim

We built a computational simulation model to estimate how many application implementations occurred in a given year, from which a total benefit of *The National Map* was calculated. Here, we give only a brief summary of the computational model, called "NB-Sim." It is described in detail in appendix C.

NB-Sim begins with the distribution of all counties in the United States into one of three tiers of complexity, as described in section 2. Estimates of this apportionment were derived by Gillespie (2003, written commun.), who estimated that 65 percent of counties were in tier 3, 30 percent were in tier 2, and only 5 percent had the capacity to do the most complicated tier 1 applications. These numbers came from his interpretation of the Federal Geographic Data Commission Survey (Gillespie, 1999) responses. Then, once *The National Map* program is underway, NB-Sim uses an annual implementation rate to determine a rate of transition of these counties from non-*National Map*-users to *National Map* users.

Once this transition is made, an individual county can begin using data provided by *The National Map* in a range of applications. It is important to remember that "county" as used here does not mean "county government" but rather refers to all the users within that county. This includes municipal, academic, and private users, and also state government applications that are scaled to the county level. Federal Government use of data is treated separately.

Initially, there are a small number of applications available to users in each tier (less than 25 total), but as the number of users grows, the number of new applications grows as well. The model has an intrinsic rate of innovation of 2 percent per user per year. Each user in each tier can potentially utilize all the applications that exist at each point in time. However, there is a lag time as the applications need to "diffuse" around the Nation. No adoption of technology happens instantaneously or completely, so the model accounts for that lag with an internal variable for "cumulative diffusion rate." There is a second variable to account for something we termed a "need limitation." This captures the fact that not all applications are needed in every county (for example, there is no need for a hurricane

[3] Details of these interviews and literature searches are provided in appendix B and in the Literature Review section of the main body of this report.

[4] Other literature addressing ways of measuring these "efficiency" gains are discussed in the Literature Review section and appendix B of this report.

evacuation routing system in Wyoming). Nor are all applications implemented each and every year. So the need limitation adjusts the total number of implemented applications downward by randomly selecting a fractional number from a distribution ranging between zero and one but which is heavily weighted toward zero. The draws average out to about 10 to 15 percent of available applications.

One other transition that occurs is that the users of *The National Map* data can advance upward through this tier system, assisted in large part by the lower cost of data, and the ease of using data that are consistent, complete, vertically and horizontally integrated, and furnished with reliable metadata. This tier advancement shift was assumed to be slow—only 1 percent per year—but it brings substantial benefits because tier 2 and tier 1 users have access to many more applications then tier 3 users do. Each user tier can make use of the applications from the tiers below it, but none can make use of tiers that are above it. They also have faster diffusion rates, meaning that they get access to those applications sooner.

Thus far, there are a number of "application uses" for each tier of county-scale users in each year. A similar process happens in a fourth tier, representing Federal applications of spatial data. Each of these application uses is multiplied by the mean improvement in net benefits of an application, as described above. Importantly, the exact value for this variable is not $1,000 every year. Rather, $1,000 is the center of a normal distribution around which the program draws a random value. The standard deviation is set at one-fourth of whatever value is input for the mean. This allows variation in the average improvement in net benefit of an application. It is possible that, on rare occasions, this average will be a negative value. In any case, this randomness produces a different value for per-application improvement for each year in the model run.[5] These figures are multiplied by the number of implemented applications; the product represents the value-in-use of data from *The National Map* in each tier and year. These products are summed in each year, then converted to present dollars by the Office of Management and Budget's suggested rate of 3.2 percent per year (for 30-year lifespan projects).

On the cost side, the best estimates of the funding levels required to build and maintain *The National Map* over its expected 30-year lifespan are the key inputs. The exact numbers are unknown at the present time, but the Implementation Team for *The National Map* approved the selection of a $25 million annual cost for the 10-year development period and $5 million per year for annual maintenance over the entire 30 years of the project. The development funding includes the start-up costs to form the partnerships, buy or create more current data, and so on. The lower figure for maintenance can be thought of as the fixed annual costs, including salaries, regular upgrades to computer equipment, facilitating partners' efforts to update the data, and so on.

[5] The center for the normal distribution for the Federal uses of spatial data is $10,000, because they are so much broader in scope.

This analysis assumes that the major goal of the USGS in building *The National Map*—that of being the integrator and server of data, rather than the original creator—can and will be realized. We assume it will take 10 years to achieve this goal, during which time, the costs of the program to the USGS will be much greater than after it, when partners will bear the bulk of the costs (presumably a much lower total, though). Thus, the annual cost of *The National Map* in the NB-Sim model is the sum of development and maintenance costs for 10 years and then is only the maintenance cost. Each year's cost is discounted to present value as well.

In each year, therefore, there are values for the total benefits and total costs of *The National Map* in present value dollars. The model subtracts the total costs from the total benefits to determine the net benefits of *The National Map*.

4. Results and Sensitivity Analysis

Simulation Results

We ran the NB-Sim model with the best available estimates for the parameters and model inputs described above. Again, full details on the values and rates of change are available in appendix C; this is a broad overview. Each model run simulated benefits and costs for 30 years and then reported the net benefit for each year and the final total. By examining results from a run, it was possible to determine the year in which the net benefit switched from negative to positive; this would be the break-even time.

The base year for dollar comparison was chosen to be 2001, because that was the year *The National Map* was announced and work on it was begun. Even though this analysis took place in 2002 and 2003, the comparison is between the status quo, pre-*The National Map* world and the fully implemented *The National Map*. The last time there was a pre-National Map world was 2001. We are currently living in an "alternative 2 world," where there is no additional funding for *The National Map*, but work towards developing it has begun with whatever internal redirection of funds is possible. This "unfunded implementation" is not the main focus of this study and does not fit into the "with versus without" framework of a cost-benefit analysis, though we did attempt to simulate it in one of the scenarios in the sensitivity analysis discussed below.

Because of several random components in the model, any single run had the potential to generate exceedingly high or low values for a number of variables. To reduce this possibility, while maintaining variability in the system and avoiding a falsely deterministic model, the baseline parameters were run 50 times, each for 30 years. We took the mean and standard deviation of both the net present value of benefits and the break-even time. This procedure was followed for the baseline results and for the sensitivity analysis of variables.

Results from the "most likely" estimates of model parameters and data inputs indicate that, over its 30-year projected lifespan, *The National Map* could have a net present value (NPV) of benefits of $2.05 billion dollars, with a standard deviation of $490 million dollars. The average time until the initial investments (the break-even period) are recovered is 14 years. The standard deviation around the payback period is 1.9 years. Under the baseline parameters, *The National Map* was credited with a positive NPV in all 50 of the simulation runs. In other scenarios included in the sensitivity analysis—discussed below—that was not always the case.

Table 1 lists the yearly mean NPV from 50 simulation runs, as well as the ending average NPV and standard deviation around it. Note the transition from negative to positive NPV in year 14 (marked with gray background). Figure 3 presents the final NPV of each of the 50 runs, showing the variability around results from a single set of inputs. Figure 4 is a graph of the total cost (in red) and total benefit (in blue) curves over time. The sharp kink in the cost curve is caused by the switch from funding development and maintenance to only spending on maintenance. The benefit curve shows slow growth initially, and then curves up sharply as *The National Map* approaches completion. Break even follows just a few years thereafter at the point the curves cross.

The slow growth in early years may seem surprising unless the dynamics of the simulation model are examined. *The National Map* takes time to build and populate with data, meaning that in the early years, very few county-level users see data that is relevant to them and can make the transition from a non-*National Map* county to a *National Map* user. Once there, they make use of only a small fraction of the small number of applications we assume exist and are *National Map*-enabled. The innovation of new applications is dependent on how many counties have made that initial transition, and so grows slowly also. At the outset, most of the users are in tier 3, meaning that they have access to the smallest number of initial applications. Users at all tier levels have had little time to adopt and implement many of the existing applications, because the cumulative diffusion fraction is still low. All of these lead to a very small number of application implementations in the early years of the program.

As *The National Map* nears completion in year 10, many things change. Assuming adequate years of funding were provided for development, by year 10, all counties have become users of *The National Map* data and distribution system. Innovation of new applications is at its peak and continues at a high level for many years. The cumulative diffusion has reached its maximum in each tier, meaning users have access to almost all of the existing applications. Further, the tier advancement trend has proceeded, and many users have shifted upward to become tier 1 and 2 users instead of tier 3. The result of all these changes is that more applications exist, and a great many more application implementations are taking place each year. In one 50-run simulation, for example, the number of implementations rose from just 65 in year three to 1,700 in year 6 to 21,000 in year 9. This increasing rate of implementation continues through the entire 30-year run. Again, though, in any single model run, the random draw for implementation rate causes more interannual variability than what probably occurs in reality, but taking the average of many runs shows the more consistent trend that is expected.

Table 1. Mean net present values (NPV) of baseline scenario, by year.

[The change from negative to positive NPV takes place in year 14, shaded in gray. The final mean and standard deviation for NPV at the bottom of the table]

Model Year	Mean Net Present Value
1	$ (30,000,000)
2	$ (58,986,986)
3	$ (86,851,952)
4	$ (113,449,315)
5	$ (138,063,308)
6	$ (159,147,477)
7	$ (171,579,633)
8	$ (176,992,427)
9	$ (161,274,589)
10	$ (155,936,276)
11	$ (120,019,163)
12	$ (69,591,309)
13	$ (11,973,403)
14	$ 62,874,333
15	$ 134,560,280
16	$ 241,066,196
17	$ 362,467,550
18	$ 461,986,307
19	$ 558,838,273
20	$ 674,433,835
21	$ 806,040,161
22	$ 948,599,229
23	$ 1,070,326,997
24	$ 1,200,519,950
25	$ 1,337,230,351
26	$ 1,472,917,954
27	$ 1,591,170,175
28	$ 1,748,433,401
29	$ 1,861,787,277
30	$ 2,051,719,005
The ending Mean and Standard Deviation (in 2001 dollars) are:	
Mean NPV	StDev NPV
$ 2,051,719,005	$ 492,308,136

Sensitivity Analysis

The input data and rates of change used in this simulation model are a combination of numbers from literature, program implementation plans, extrapolations from interviews and empirical data points, and simplifying assumptions. All of these estimates are extremely conservative, and we believe we are, if

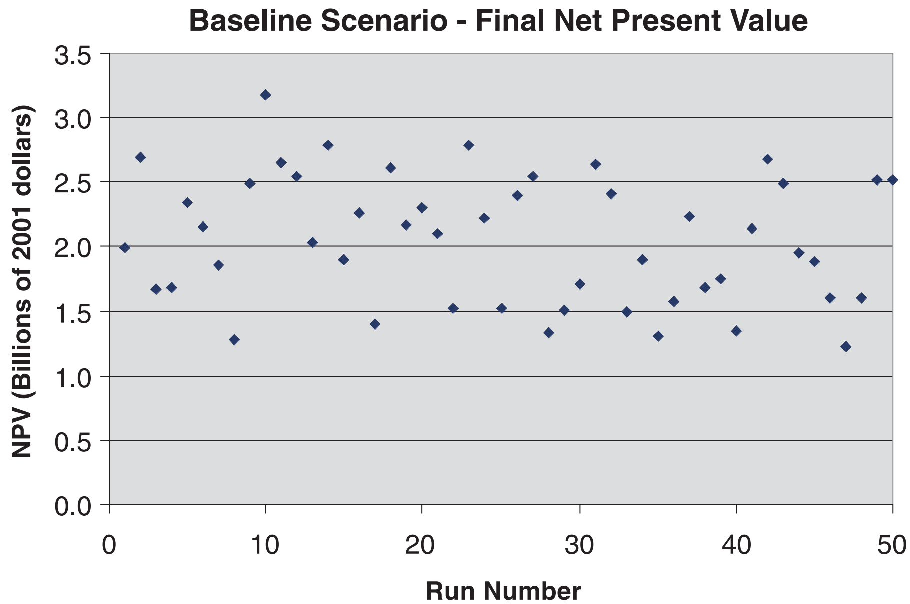

Figure 3. Variation in net present value resulting from 50 runs of the baseline scenario. The baseline scenario was run 50 times through NB-Sim, each time generating a different net present value (NPV) because of several random modeling components. Here, the NPV of each run is plotted against the run number to show the range of variation in the baseline.

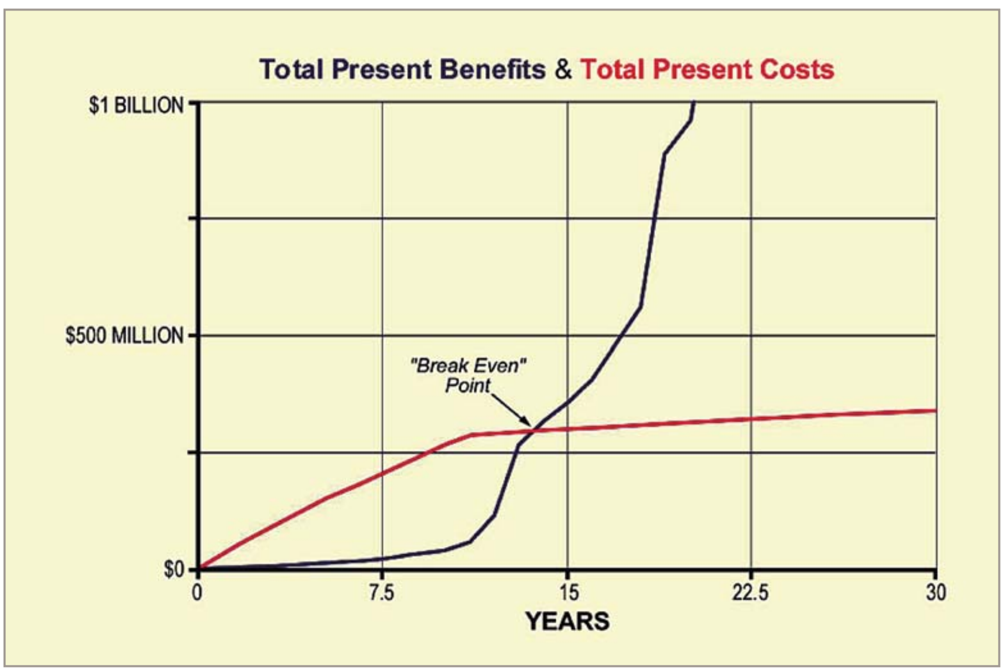

Figure 4. The total cost and total benefit curves for a single run of the baseline scenario. A graph of the accumulating benefits (blue) and costs (red) of *The National Map* over 30 years. The break-even point is where the curves cross (year 14). The kink in the cost curve is the change from development costs to maintenance costs only. Note the slow upward climb of the benefit curve until the project nears completion, when most of the country's spatial-data users have had time to see valuable data, adopt *The National Map* as their data source for existing tasks, and innovate new applications for spatial data. The area above the cost curve and below the benefit curve, between years 14 and 30 represents the net present value (NPV).

anything, underestimating the value to society of *The National Map*. However, it would be unwise to say that these results are more prediction than simulation. The NB-Sim model is an attempt to capture all the flows of changes in the streams of benefits and costs as a result of transitions *The National Map* is likely to bring and to populate them with reasonable numbers. It seemed important to go beyond merely submitting a "best guess" estimate of results and instead provide a tool to examine scenarios and input values more pessimistic and (or) more optimistic than the baseline values we used. To address this issue, we conducted a sensitivity analysis of the values and variables in the baseline scenario. A full discussion, analysis, and table of results are included in appendix D, but here, we provide an overview of the methods and results.

In the sensitivity analysis, each set of inputs was dubbed a "scenario," regardless of whether only one variable was changed from its baseline value or whether there were wholesale revisions. Each scenario was given a number and a short description, most of which only describe those inputs that are different from the baseline. As in the baseline case, each scenario consists of 50 runs with the same input values, to account for the randomness of many model parameters. The 60 scenarios were grouped into two groups of 30. The 30 scenarios in each group were identical except for the annual development cost, which changed from $25 million to $50 million per year. The other changes ranged between more conservative than the baseline and more optimistic. The conservative/ pessimistic changes were made to examine how low specific values could get and still allow *The National Map* to make economic sense. The optimistic cases were developed to see how large the benefits to society could get if the input estimates were even a little bit too low.

The bullets below highlight the most important insights from an examination of the simulation results:

- With a 10-year development time, $25 million per year to build and $5 million per year to maintain *The National Map* produce positive net benefits that are robust to large numbers of even more conservative changes.
- The following changes, enacted individually, do not dramatically reduce the NPV:
    - Eliminating tier advancement.
    - Eliminating all Federal use of spatial data.
    - Reducing the initial number of applications to 1 per tier.
    - Designating all counties as initially in tier 3.
    - Distributing the counties equally across all three tiers.
- The following changes, enacted individually, reduce NPV substantially, but still provide large positive net benefits to society:
    - Reducing cumulative diffusion to one-half of baseline at all points in time.
    - Cutting intrinsic innovation rate by one-half, down to 1 percent.
    - Lowering the increase in net benefits per application use to $500 instead of $1,000 (and $5000 for Federal applications instead of $10,000).
- The following optimistic changes, enacted individually, increase NPV:
    - Increasing innovation rate to 5 percent brings a dramatic gain to more than $5 billion.
    - Making diffusion immediate and complete, rather than gradual and partial, increases NPV to more than $3 billion; merely doubling its rate brings similar gains.
    - Doubling the per-application improvement in net benefit of an application (that is, making it $2,000 per use and $20,000 for Federal applications) roughly doubles the NPV to more than $4 billion.
    - Shifting initial tier distribution entirely to tier 1 or 2 brings substantial improvements over baseline; tier 1 more so than tier 2.
    - Increasing the value of either normal or emergency Federal applications does not bring substantial gains; there simply are not enough application implementations to make these changes stand out.
- When combining pessimistic changes, each one reduces NPV further, but even an unrealistically "worst-case" conservative scenario—no tier advancement, cutting net benefit gain per application use in half, making innovation rate 1 percent per user-year, and setting initial numbers of applications to one per tier—still brings a positive, though small, NPV of about $100 million.
- Assuming a constant total budget for development, but speeding up development to 5 years brings significant gains in NPV; conversely, slowing it to 20 years cuts NPV in half to about $1 billion.
- Failing to implement *The National Map* fully—that is, funding it for less than the time it is estimated to require—cuts the NPV severely.
- Building *The National Map* through Alternative 2—no additional funding—does bring a positive NPV, but it is very small, only $22 million over 30 years.

Importantly, the NB-Sim model is not only useful for sensitivity analysis; this tool also provides a method for evaluating *The National Map* program development as it proceeds. As better observations and empirical data are collected, the input numbers and rates of change can be improved so that they more nearly represent reality. Moreover, it can assist the Implementation Team in simulating the outcomes of various program development choices as they arise. In this use, it is as much a decision-support system as it is a cost-benefit analysis.

5. Discussion

This policy analysis was conducted in support of the USGS Geography Discipline's effort to develop *The National Map*. The research goal was to develop a system with which to estimate and analyze the costs involved in building, maintaining, and distributing *The National Map* and the various benefit streams expected from its existence. We tried to capture all the major transitions and dynamics in the amount, type, and value of spatial data uses that *The National Map* can enable and to put dollar values on the benefits that can result from those changes. Comparing those benefits and costs, we sought to estimate a net benefit to society that could result from developing *The National Map*.

We developed an analytical framework based in economic theory to model the various benefit and cost streams over time. The underlying idea was that by reducing the investment of time or money required to complete a certain project or task (an "application"), *The National Map* could reduce the cost of implementing that application. Further, by providing data that were current and consistent across the Nation, the outcome of that program or task could be improved, resulting in an improved benefit of implementing it. Finally, these gains would likely stimulate innovation of new uses for spatial data and facilitate their rate of adoption by others. Together, these effects would produce a change in net benefit each time an application was implemented. Even if those changes were small for any one project, iterating them over time and space would ultimately bring substantial benefits to society that would be directly attributable to *The National Map*.

However, finding or developing data to populate that economic model was not an easy task. The existing literature on the value of spatial data as used in real-world applications is not very extensive. Further, even that work which does exist does not analyze projects that are directly comparable to *The National Map*. Most previous work investigates the change in values when moving from paper maps or analog spatial data to a GIS or digital data. Those results had to be modified and adapted to address the question of the value of an improved data set and distribution system that will be provided by *The National Map*.

Because of these limitations, as well as the time and resources available to conduct the analysis, a full accounting of the likely costs and benefits of *The National Map* was not feasible. Instead, we developed a computational simulation model called "NB-Sim" into which the best estimates could be entered. The model simulates ranges of results that can be expected from each set of inputs. It allows rapid adjustment of the inputs so that more pessimistic or more optimistic beliefs can be simulated, and so that new results can be generated if more input data or evidence becomes available.

Using extremely conservative assumptions for initial values and rates of change, the expected NPV of benefits of a fully implemented version of *The National Map* are just over $2 billion (in 2001 dollars) in a projected product lifespan of 30 years. The standard deviation around the NPV is $490 million, meaning the net benefits are most likely to fall between $1.07 and 3.03 billion, the range covered by two standard deviations (approximately the 95-percent confidence interval). Despite significant outlays over 10 years to develop *The National Map*, the project breaks even in year 14, shortly after its completion.

The sensitivity analysis we conducted showed that these baseline results are robust to even dramatic changes in one or more of the input values. Several model parameters could be doubled, halved, or eliminated without bringing substantial changes. However, the results are quite sensitive to changes in certain other variables. The most critical variables were the:

- Average change in the net benefit of an application implementation as a result of data from *The National Map*.
- Rate of innovation of new applications.
- Amount of cumulative diffusion of data from *The National Map* and of the new and existing applications those data can inform.

This policy analysis resulted in a rigorous framework with which to evaluate the expected net benefits to society of a distributed database of digital spatial data that will be *The National Map*. Confidence in these expectations is limited by the small quantity of rigorous analytical work that puts credible numbers on the value in use of digital spatial data and the frequency and pervasiveness of applications using those data. The fields of geography, information technology, and spatial statistics would all benefit greatly from further research into these sorts of issues.

We recommend that the USGS take the lead in conducting prospective studies of the type, quantity, costs, and benefits of using digital spatial information. This would be a most opportune time to develop an empirical study in a "pre-*National Map*" situation of how data was used, what those uses cost, and what benefits they brought. Then, as *The National Map* develops and starts serving data, follow-up studies could measure changes to those cost and benefit profiles. This would allow the inclusion of more credible numbers into the simulation model, thus simultaneously providing project tracking and evaluation, decision support for future *National Map* implementation steps, and evidence of the continued importance of Federal spatial-data programs.

References Cited

Alexander, Timothy M., 2002, Measuring the value of geospatial information—a critical and doable challenge: Washington D.C., TMA Associates, [unpublished].

Alexander, Timothy M. and Schmidt, Nicholas, 2002, The one percent factor: Imaging Notes, November/December 2002.

Baltimore County Office of Information Technology, 2001, Analysis of impact of GIS to Baltimore County: [online at http://aria.arizona.edu/slg/GISCostBenefit.pdf].

Boardman, A.E., Greenberg, D.H., Vining, A.R. and Weimer, D.L., 1996, Cost-Benefit Analysis—Concepts and Practice: Prentice-Hall, Upper Saddle River, NJ.

Bernknopf, R.L., Campbell, R.H., Brookshire, D.S., Shapiro, C.D., 1988a, A probabilistic approach to landslide hazard mapping in Cincinnati, Ohio, with applications for economic evaluation: Bulletin of the Association of Engineering Geologists v. 25, p. 39-56.

Bernknopf, R.L., Brookshire, D.S., Campbell, R.H., Shapiro, C.D., 1988b, An economic and geographic appraisal of a spatial natural hazard risk—a study of landslide mitigation rules: Environment and Planning v. 20, p. 621-631.

Bernknopf, R.L., Brookshire, D.S., Soller, D.R., McKee, M.J., Sutter, J.F., Matti, J.C., Campbell, R.H., 1993, Societal value of geologic maps: U.S. Geological Survey Circular 1111.

Bond, D., 2000, GIS and spatial analysis in urban and regional research: Cities and Regions, GIS Special November 2000 [online at http://www.car.busmgt.ulst.ac. uk/papers/gis.pdf].

Center for Technology in Government, 1995, Sharing the costs, sharing the benefits—the NYS GIS cooperative project: New York State Department of Environmental Conservation Center for Technology in Government Project Report 95-4.

Croswell, Peter, Fries, J.T., Lambert, S.C., 2003, Status and potential role of state, tribal, regional, and local public entities in The National Map program: PlanGraphics, Inc. and Kentucky Governor's Office for Technology.

Dickinson, H.J., Calkins, H.W., 1989, The economic evaluation of implementing a GIS: International Journal of Geographical Information Systems, v. 2, p. 307-327.

Federal Geographic Data Committee, 2004, National Spatial Data Infrastructure Page: [online at http://www.fgdc.gov/nsdi/nsdi.html].

Fries, J.T., Metcalf, Annie, Swain, Lisa, Vincent, Andrew, 2002. Executive summary of the Ohio spatial data cost-benefit analysis: PlanGraphics, Inc., and Booz Allen Hamilton, Inc., Frankfort, Ky and McLean, Va.

Gillespie, Stephen R., 1994, Measuring the benefits of GIS use—Two transportation case studies: URISA Journal, v. 6, p. 62-67.

Gillespie, Stephen R., 1997, A model approach to estimating GIS benefits: U.S. Geological Survey [unpublished].

Gillespie, Stephen R., 2000, An empirical approach to estimating GIS benefits: URISA Journal, v. 12, p. 7-13.

Hardwick, P., Fox, B., 1999, Study of potential benefits of Geographic Information Systems for large fire incident management, prepared for USDA Forest Service Region 5, Fire and Aviation Management: Pacific Median Resources, [online at http://www.fs.fed.us/fire/planning/nist/GIS_study.pdf].

Hwang, G-H, Diffusion of information and communication of technologies and changes in skills: Sussex University Science and Technology Research (SPRU) Paper no. 48, 25 p.

Johnson, D.D., Craig, W.J., 1997, GIS in Dakota County—benefits add up: Minnesota Counties, 5 p.

Korte, G., 1996, Weighing GIS benefits with financial analysis: GIS World, v. 7, p. 48-52.

Ledbetter, M., Hunt, C., Anderson, D., 1997, Blueprints for a citywide GIS: GIS World, v. 11, p. 62-65.

Lennstrand, B., 1998, Diffusion of information and communication technology to households—How come it goes so slowly when it goes fast? Twelfth Biennial Information Technology Conference, Stockholm, Sweden, p. 1-16.

Lerner, N.B., 1994, Quantifying GIS benefits—a bottom-up approach that builds management confidence: URISA journal, p. 53-59.

McInnis, L., Blundell, S., 1998, Analysis of Geographic Information Systems (GIS) implementations in state and county governments of Montana: report prepared for The Montana Geographic Information Council [online at http://www.nysgis.state.ny.us/montana/montana.htm].

Montgomery County Council's Management and Fiscal Policy Committee, 1999, Geographic Information System cost/benefit assessment report: Montgomery County, Md [online at http://www.mc-mncppc.org/gis/gisbenefits.pdf].

National Research Council of the National Academies, 2003, Weaving a National Map: Washington, D.C., National Academies Press.

Oxford Economic Research Associates Ltd. (OXERA), 1999. The economic contribution of the ordinance survey of Great Britain: OXERA, 50 p. [online at http://www.ordsvy.gov.uk/literatu/external/index.htm]

Pistorese, T.A., 1994, Actual benefits realized through implementation of a GIS: AM/FM, 409-417.

Price-Waterhouse, 1995, Australian land and geographic data infrastructure benefits study: Price-Waterhouse Economic Studies and Strategies Unit, study conducted for the Australia-New Zealand Land Information Council [unpublished].

Prisley, S., Mead, R., 1987, Cost benefit analysis for Geographic Information Systems. Proceedings of GIS '87—

American Congress on Surveying and Mapping: American Society for Photogrammetry and Remote Sensing, Bethesda, Md, p. 29-37.

Rogers, E.M., 1962, Diffusion of innovations: The Free Press.

Sears, Garry, 2001. Geospatial Data Policy Study, prepared for GeoConnections Policy Advisory Node: KPMG Consulting, Ottawa, Ontario, Canada.

Somers, Rebecca, 1999, Framework Data Survey Preliminary Report: Federal Geographic Data Committee [online at http://www.fgdc.gov/framework/survey _results/readme.html].

Tulloch, David and Fuld, Jennifer, 2001, Exploring county-level production of framework data—analysis of the national framework data survey: URISA Journal, v. 13, no. 2.

United States Office of Management and Budget, 1992 (revised 2003), Circular A-94—Guidance and discount rates for benefit-cost analysis of Federal programs.

United States Geological Survey, 1988, Primary mapping economic analysis, phase two: Reston, Va.

Worrall, L., 1994, Justifying investment in GIS—A local government perspective: International Journal of GIS, v. 8, p. 545-565.

Appendix A. Formal Economic Theory

The National Map is an on-line database and provider of geospatial information that is a pure public good. As a public good, the intent of *The National Map* is to not exclude anyone from the benefits of its use and to supply geospatial information at virtually no cost. As shown in equation 1 in section 2, the net benefits of *The National Map* are specified as the value-in-use of the information available. The value of information is a function of the spatial and temporal applications of digital spatial data and is measured in terms of applications (that is, cost savings, value added, and (or) new uses). The willingness to pay for applications using *The National Map* is assumed to be increasing at a decreasing rate:

$$\frac{d(VOI_{TNM})}{ds} > 0 \text{ , and } \frac{d^2(VOI_{TNM})}{ds^2} < 0 \tag{A1}$$

where s is *The National Map* user-tiers, and VOI_{TNM} is the value in use of the geospatial data that is dependent on the tier of the user and the availability of *The National Map*:

$$VOI_{TNM} = f\,(s,TNM) \tag{A2}$$

where $s = g(TNM)$
where TNM refers to *The National Map program*.

Note that the order of progression of tier advancement is tier 3 → tier 2 → tier 1, and that an increase in tier sophistication does in fact result in a tier name with a lower number (for example, a tier 2 county becomes a tier 1 county, thus increasing its capability but lowering the numeral in its name).

The National Map is specified in both terms in equation A2 as (1) the increased efficiency from using *The National Map* in current applications (as a function of $s = g(TNM)$ and (2) the innovative applications that arise from being a *National Map* partner $f(TNM)$.

The willingness to pay for *The National Map* is a derived demand. To find the marginal willingness to pay for *The National Map*, totally differentiate VOI into its components:

$$d(VOI_{TNM}) = f_s ds + f_{TNM} d(TNM) \tag{A3}$$

$$\frac{d(VOI_{TNM})}{d(TNM)} = f_s \frac{ds}{d(TNM)} + f_{TNM}\frac{d(TNM)}{d(TNM)} \qquad \left[\frac{d(TNM)}{d(TNM)}\right] = 1 \tag{A4}$$

$$= \frac{\partial(VOI_{TNM})}{\partial s}\frac{ds}{d(TNM)} + \frac{\partial(VOI_{TNM})}{\partial(TNM)} \tag{A5}$$

The term $\frac{ds}{d(TNM)}$ represents the advancement rate for *The National Map* users to change tiers.

The term $\frac{\partial(VOI_{TNM})}{\partial(TNM)}$ represents the fraction of the expected applications' benefits associated with the implementation of *The National Map* and represents the derived demand for *The National Map.*

We assume that increased collaboration with the USGS—as in data-sharing partnerships—in *The National Map* increases the capability and facility with spatial data more quickly than a nonpartner data user. A partnership is defined as an acceptance of standards, data sharing, and so on.

Appendix B. Details of Data Gathering and Synthesis

Estimating Change in Net Benefit of Implementing an Application

In our analysis we have derived an average value of improved net benefit per application of spatial data from *The National Map*. This value was calculated using 40 available application (usage) values found in the literature (n = 36) and from phone surveys (n = 4). Most often, the available values were given in terms of an organization's cost savings per year to perform a given task after the implementation of GIS (efficiency gains). We designated 17 tasks as "blanket" applications; those that could be performed with *The National Map* at its minimum resolution.[6] An additional 10 tasks were designated as "quilt" applications; those that could be performed with *The National Map* at a resolution greater than the minimum. Finally, 13 applications were unspecified because it is unclear what resolution would be required to perform them. The average value per application per year of the "blanket" applications alone (n = 17) is ~$286,057. The average value per application per year of the combined "blanket" and "quilt" applications (n = 27) is ~$268,772, and the average value per application per year of all given application values (n = 40) is ~$257,536. Table B1 contains these applications and their values and calculations.

Although some of these applications may be performed only once or twice during a typical year (for example, voter redistricting, watershed analysis), others could be performed many times (for example, custom mapping). There is no clear way to determine how often an organization or an individual repeats an application. For the purposes of this analysis, it was appropriate to make a conservative estimate. Thus, we made the assumption that the average application is performed monthly. With this assumption we found a per-application value of $21,461 to $23,838 depending on which of the 36 values are used. FGDC survey data indicates that the majority (70 percent) of the geospatial data used is of the lower resolution variety, which would correspond to the "blanket" minimum resolution data offered by *The National Map* (Stephen R. Gillespie, written commun.). As shown, the decision to use application values that we have designated as "blanket" or "quilt" has a minor to moderate influence on the final value in our analysis.

In order to be even more cautious with our per application value, estimate we assume a value of $1,000 per application, or about $\frac{1}{21}$ to $\frac{1}{24}$ of the larger estimates. This is done largely because this study attempted to measure change in value, not total value. Further, the values in the literature are almost all cost savings reported by sophisticated (tier 1 or 2) GIS users with at least a moderate level of infrastructure in their organization. Less sophisticated users with little or no infrastructure may derive a much smaller cost-savings benefit per application.

For example, suppose that in terms of operating costs and infrastructure, the average tier 1 user is essentially 100 times larger than a tier 3 user (and therefore has that many more application occurrences) and 10 times larger than a tier 2 user. The per application values for tiers 1, 2, and 3 would be $21,461 to $23,838, $2,146 to $2,384, and $215 to $238, respectively. A more reliable estimate of the differences between tier usage comes from the FGDC Survey of GIS data users, producers, and distributors that indicates that tier 1 has approximately 4 times as many application occurrences as tier 2 (expenditures of $296,831 and $69,386, respectively) (Stephen R. Gillespie, 2003, written commun.). Further, assuming that tier 2 has approximately 4 times as many application occurrences as tier 3 gives per-application values of $21,461 to $23,838, $5,365 to 5,960, and $1,341 to $1,490 for tiers 1, 2, and 3 respectively.

The value of $1,000 improvement per application implementation is an extremely conservative estimate and safe starting point for several reasons. The first and most obvious is that our own analysis based on the data from the literature implies higher values (~$1,300 to $24,000 per application) than are ultimately used in this analysis. A second reason is that we are making the estimate almost entirely based on reported cost-savings (efficiency) benefits. There are few available data on effectiveness and innovation benefits, which—if available—would surely elevate the average application improvement value and might boost it considerably. As *The National Map* is implemented and data is collected on users, our understanding of the value of applications will improve.

Applications and Benefits

Some of the journal articles and technical reports we reviewed include a list of potential or actual applications that GIS is used for in particular organizations. We have compiled these lists to generate a larger, comprehensive list of GIS applications included in this report. The majority of the literature cites qualitative examples of the benefits of geospatial data. There are also rare quantitative data on the benefits of geospatial data available in the literature. The benefits are almost exclusively in the form of calculated cost savings to a particular organization. As far as our understanding of GIS applications and benefits is concerned, the most useful of the reports available in the literature is the Baltimore County Office of Information Technology (2001) cost-benefit analysis of the county's GIS. The authors list the GIS cost savings to each department within the organization, as well as the cost

[6] The National Research Council's analysis of *The National Map*, titled *Weaving a National Map* developed the "blanket vs. quilt" concept to illustrate the USGS goal of completing *The National Map* in a uniform and complete way across the entire country (the blanket), while providing more types and higher resolutions of data in certain places, where such data were available (the patchwork quilt). We examined primarily "blanket" applications in our analysis.

Table B1. Calculating the value of an application.

Application	Blanket apps. (n=17)	Blanket and quilt apps. (n=27)	Blanket, quilt, and undefined apps. (n=40)
Adequate public facilities	$ 13,000	$ 13,000	$ 13,000
Agriculture preservation	$ 300,000	$ 300,000	$ 300,000
Basic services mapping	$ 9,800	$ 9,800	$ 9,800
Development review and tracking	$ 156,000	$ 156,000	$ 156,000
Floodplain analysis	$ 35,000	$ 35,000	$ 35,000
Master planning	$ 353,600	$ 353,600	$ 353,600
Watershed planning/management	$ 360,000	$ 360,000	$ 360,000
GIS viewing application	$ 312,260	$ 312,260	$ 312,260
Automated mapping	$ 31,200	$ 31,200	$ 31,200
Custom cartography/spatial analysis	$ 158,000	$ 158,000	$ 158,000
Voter redistricting	$ 16,200	$ 16,200	$ 16,200
Map production	$ 135,000	$ 135,000	$ 135,000
Maps	$ 1,086,667	$ 1,086,667	$ 1,086,667
Geographic records	$ 866,667	$ 866,667	$ 866,667
Non-records related productivity improvements (crew delays, administration)	$ 793,333	$ 793,333	$ 793,333
One time savings (avoided system development, map backlog elimination)	$ 180,000	$ 180,000	$ 180,000
Garbage Truck Routing	$ 56,250	$ 56,250	$ 56,250
Average value blanket	**$ 286,057**		
Alley reconstruction		$ 15,000	$ 15,000
Curb/gutter conditions/repair/permits		$ 1,200	$ 1,200
Drainage complaint investigation		$ 6,000	$ 6,000
Legislative analysis		$ 44,200	$ 44,200
Water and sewer administration		$ 97,000	$ 97,000
Crime analysis interface		$ 20,800	$ 20,800
911 address file interface		$ 20,800	$ 20,800
Generic mail labels		$ 48,880	$ 48,880
Census bureau appeal		$ 1,800,000	$ 1,800,000
Savings from improved electrical system analysis		$ 340,000	$ 340,000
	Average value blanket + quilt	**$ 268,772**	
Building permit review			$ 6,750
Coastal Zone Management Program			$ 643,700
Lacquire			$ 20,000
Property analysis			$ 6,750
Rights-of-Way fee/maintenance			$ 11,700
Rural legacy			$ 1,220,000
Property tax assessment audit			$ 61,680
Code enforcement			$ 45,400
Zoning notifications			$ 42,400
Other			$ 806,667
Ambulance Dispatch			$ 150,000
Utility Distribution			$ 2,534
E-911 Mapping			$ 27,000
		Average value blanket +quilt+ undefined	**$ 257,536**

Key

Blanket

Quilt

Undefined

savings that the organization's GIS capability brings to particular applications.

A table summarizing the applications and, where available, data on the benefits of particular applications is provided at the end of this appendix (table B2). Those applications that are expected to be performed at the minimum resolution ("blanket" applications) guaranteed by *The National Map* are indicated in yellow along with estimated yearly cost savings. Those that can be performed at some higher resolution ("quilt" applications) are shown in blue along with estimated yearly cost savings. Other applications that *The National Map* may be frequently used for, but with no available benefits estimate are highlighted in green. Table B3 is simply a list of county-level applications collected by the Geography Discipline of the USGS.

Phone call data surveys

In order to supplement the data collected from the literature, we directly contacted several representatives of organizations that produce, distribute, and use geospatial data. Contacts included leads from pilot projects of *The National Map*, USGS State liaisons, USGS geospatial data contractors, and other sources of information recommended to us by our contacts.

In our conversations with each representative, we were particularly interested in collecting quantitative data on the following topics to aid us in our benefits estimation:

- Usage levels.
- Applications (uses) of the data.
- Efficiency and effectiveness benefits.

Usage

Respondents gave several different units of measure for usage levels of their geospatial data (for example, hits/day on a website, number of data files downloaded, number of data licenses purchased). In counties/regions with sophisticated GIS capability (such as Clark County, Nevada) websites receive a lot of traffic, often thousands of hits per day. In smaller counties with limited GIS capabilities, usage levels are, as expected, much lower.

Applications

The respondents provided several examples of geospatial-data applications. These include web interfaces that allow users to generate and print their own maps or other data plots, routing applications, environmental assessment uses, billing and addressing applications, and others. The applications listed by the respondents do not constitute an exhaustive list; rather it should be viewed as a set of examples.

Benefits

As in the literature, contacts rarely provided quantitative data regarding the benefits of geospatial data to their organization. However, many contacts implied that the use and production of geospatial data has improved their organization and were able to give anecdotal examples to illustrate the benefits. These examples ranged from several thousand to hundreds of thousands of dollars in benefits, but, again, were not based on any rigorous analysis.

Table B2. Applications and their estimated values (where available).

[Those applications that are expected to be performed at the minimum resolution ("blanket" applications) guaranteed by *The National Map* are indicated in yellow along with estimated yearly cost savings. Those that can be performed at some higher resolution ("quilt" applications) are shown in blue along with estimated yearly cost savings. Other applications that *The National Map* may be frequently used for, but with no available benefits estimate are highlighted in green]

Application	Blanket	Quilt	Est. Ann. Benefit (anecdotal)
McInnis and Blundell, 1998			
Property maps	N	Y	
Disaster and emergency planning	Y	Y	
Automated tax assessment using other data sets such as soil, topography, and climate to describe land productivity	Y	Y	
Right-of-way assessments		Y	
Internet property research	N	Y	
Growth analysis	Y	Y	
Wildlife habitat monitoring and protection	Y	Y	
Resolving areas of public/private conflict such as hunting and fishing access	Y	Y	
Tax assessment, including locating untaxed parcels	N	N	
Establish institutional controls on land use near Superfund sites	Y	Y	
Zoning/master planning	Y	Y	
Address information can be included in enhanced 911	N	Y	
Weekly ownership updates for county government		Y	
Septic permitting system	N	N	
Automated property owner notification	N	N	
Automated permit and development tracking	N	Y	
Volumetric analysis for coal mine reclamation	N	N	
Reclamation plan analysis (for example, assessing revegetation potential based on slope, and vegetation type)	Y	Y	
Database of groundwater wells in the vicinity of coal mines	N	Y	
Aesthetics (for example, view shed analysis for proposed mines)	Y	Y	
Wildlife (for example, potential fishery impacts of a proposed pipeline)	Y	Y	
One-time applications (for example, calculating the volume of material to removed from a mine tailings pile)	Y	Y	
Watershed analysis	Y	Y	
Interactive well-finder on the web	N	Y	
Drought monitoring	Y	Y	
Natural Heritage Program			
Underground Storage Tank (UST) analysis	N	N	
Road reports (updated data available at information kiosks and on-line)	Y	Y	
Road rating (condition of the road used to prioritize road repairs)	Y	Y	
Rural addressing	Y	Y	
Flood plain delineation	Y	Y	
Automated land records searches	N	Y	
Automated underground utilities information	N		
Environmental cleanup coverage	Y	Y	
Baltimore County GIS			
Accident location analysis	Y	Y	
Address matching	N	Y	
Address validation for Data entry	N	Y	

Table B2. Applications and their estimated values (where available)—Continued.

Application	Blanket	Quilt	Est. Ann. Benefit (anecdotal)
Adequate public facilities	?	Y	$13,000
Agriculture preservation	Y	Y	$300,000
Alley reconstruction	N	Y	$15,000
Approved development locations	Y	Y	
Assessor cards scanned	N	N	
Basic services mapping	Y	Y	$9,800
Bridge inventory and inspections	Y	Y	
Building permit review	N	?	$6,750
Bulk trash routing	Y	Y	
Cadastral (property) map preparation updates	N	N	
Capital project management	N		
Commercial land inventory	N	N	
Commercial properties real estate database	N	N	
Communication tower locations	Y	Y	
Complaint tracking and response	Y	Y	
Conservation master plan management	Y	Y	
County water and sewer master-plan mapping	N	N	
County-owned structures/space inventory	N		
Crime analysis	Y	Y	
Critical area analysis	Y	Y	
Curb/gutter conditions/repair/permits	N	Y	$1,200
CZMP application	?	?	$643,700
Data distribution applications	Y	Y	
Data maintenance applications			
Data quality control applications			
Data query and display application	Y	Y	
Demographic analysis	?	?	
Detour plans	Y	Y	
Development review and tracking	Y	Y	$156,000
Districting	Y	Y	
Down zoning			
Drainage complaint investigation	N	Y	$6,000
Easement mapping			
Economic development site selection	Y	Y	
Functional area/program			
Engineering design/studies	Y	Y	
Enterprise zones			
Environmental investigation review	Y	Y	
Facilities management	N	Y	
Flood control/inspections	N	Y	
Floodplain analysis	Y	Y	$35,000
Forest management plan	Y	Y	
Future water, sewer, storm drain, roads, and water tank mapping	Y	Y	
Grinder pump locations	Y	Y	
Growth management	Y	Y	
Gunpowder watershed ecological model	Y	Y	
Hazmat tracking	Y	Y	
Hydrologic Modeling	Y	Y	
Internet site posting			

Table B2. Applications and their estimated values (where available)—Continued.

Application	Blanket	Quilt	Est. Ann. Benefit (anecdotal)
Investigation of surplus property	N	N	
Lacquire	?	?	$20,000
Land acquisition databases	Y	Y	
Land use analysis Landfills and recycling facilities management	Y	Y	
Legislative analysis	?	Y	$44,200
Management of the Chesapeake Bay Program	Y	Y	
Master planning	Y	Y	$353,600
Master roads inventory/street segment integration	N	Y	
NPDES stormwater management	N	Y	
Nutrient reduction strategies			
Open space analysis	Y	Y	
Park development siting	Y	Y	
Patron analysis	N	N	
Pavement cuts permits	N	N	
Pavement marking inventory (re-stripping)	N	Y	
Preliminary alignment studies	N		
Property analysis	N	Y	$6,750
Public access	Y	Y	
Public works maintenance	Y	Y	
Repaving support	N		
Reservoir profiles	N	Y	
Rights-of-way fee/maintenance	?	Y	$11,700
Routing	Y	Y	
Rural legacy	?	?	$1,220,000
School location mapping	Y	Y	
Shoreline land use study	Y	Y	
Sidewalk inventory/repair	N	Y	
Signal inventory/design	N	Y	
Site analysis/plan development	N	Y	
Smart growth	Y	Y	
Snow removal/routing issues	Y	Y	
Solid waste collection routes	Y	Y	
Standardized map production	Y	Y	
Storm drain culvert studies	N	Y	
Street naming	Y	Y	
Street sign inventory	N	Y	
Street sweeping routing	Y	Y	
Streetscapes investigation			
Study area maps	Y	Y	
Traffic calming			
Truck traffic routing	Y	Y	
Utilities key sheet mapping	Y	Y	
Utilities maintenance programming	N	Y	
Utilization of planimetric/topographic map in lieu of surveys	Y	Y	
Vacant land analysis	Y	Y	
Water and sewer administration	N	?	$97,000
Water and sewer amendment process	N	N	
Water and sewer pumping stations	Y	Y	
Water quality monitoring	N		

Table B2. Applications and their estimated values (where available)—Continued.

Application	Blanket	Quilt	Est. Ann. Benefit (anecdotal)
Watershed planning/management	Y	Y	$360,000
Work order management	N		
Zoning-hearing case development and analysis			
Zoning layer	Y	Y	
Zoning review cases	Y	Y	
Blueprint for a citywide GIS (Scottsdale), 1997			
GIS viewing application	Y	Y	$312,260
Automated mapping	Y	Y	$31,200
Crime analysis interface	N	Y	$20,800
911 address file interface	N	Y	$20,800
Property tax assessment audit	N		$61,680
Code enforcement	N		$45,400
Zoning notifications			$42,400
Generic mail labels	N	Y	$48,880
Custom cartography/spatial analysis	Y	Y	$158,000
Census bureau appeal	N	Y	$1,800,000
Voter redistricting	Y	Y	$16,200
USFS Large Fire Incident Management, 1999			
Location of camps	Y	Y	
Division assignments			
Transportation logistics	Y	Y	
Perimeter maps every 6 hours	Y	Y	
Fire intensity information	N	Y	
BAER mapping			
Plan implementation			
Suppression damage locations	Y	Y	
Structure protection	Y	Y	
Resource determination		Y	
Line construction		Y	
Tactical design	Y	Y	
Perimeter maps every hour	Y	Y	
12 hr. expected situation report			
Strategy development	Y	Y	
Perimeter maps every 3 hours	Y	Y	
Retardant and water dumps	Y	Y	
Air operations			
Alternative creation			
Fire modeling	Y	Y	
Johnson and Craig, Dakota County GIS 1997			
Condemnations	N	Y	
New Library Siting	Y	Y	
Transit Scheduling	N	Y	
Highway Mapping	Y	Y	
Pesticide education			
Selling Tax Forfeit Properties	N	Y	
Assisted Living Planning	N		
Mosquito Control	N	Y	
Traffic Planning	Y	Y	

Table B3. Alphabetical list of county-level GIS applications.

County GIS Applications		
A through E	**E through P**	**P through Z**
Accident Location Analysis	Easement Mapping	Public Works Maintenance
Address Matching	Economic Development Site Selection	Repaving Support
Address Validation for Data Entry	Engineering Design/Studies	Reservoir Profiles
Adequate Public Facilities	Enterprise Zones	Rights-of-Way Fee/Maintenance
Agriculture Preservation	Environmental Investigation Review	Routing
Alley Reconstruction	Facilities Management	Rural Legacy
Approved Development Locations	Flood Control/Inspections	School Location Mapping
Assessor Cards Scanned	Floodplain Analysis	Shoreline Land Use Study
Basic Services Mapping	Forest Management Plan	Sidewalk Inventory/Repair
Bridge Inventory and Inspections	Future Water, Sewer, Storm Drain, Roads and Water Tank Mapping	Signal Inventory/Design
Building Permit Review	Grinder Pump Locations	Site Analysis/Plan Development
Bulk Trash Routing	Growth Management	Smart Growth
Cadastral (Property) Map Preparation Updates	Gunpowder Watershed Ecological Model	Snow Removal/Routing/Issues
Capital Project Management	Hazmat Tracking	Solid Waste Collection Routes
Commercial Land Inventory	Hydrologic Modeling	Standardized Map Production
Commercial Properties Real Estate Database	Internet Site Posting	Storm Drain Culvert Studies
Communication Tower Locations	Investigation of Surplus Property	Street Naming
Complaint Tracking and Response	Lacquire	Street Sign Inventory
Conservation Master Plan Management	Land Acquisition Databases	Street Sweeping Routing
County Water and Sewer Master Plan Mapping	Land Use Analysis	Streetscapes Investigation
County-Owned Structures/Space Inventory	Landfills and Recycling Facilities Management	Study Area Maps
Crime Analysis	Legislative Analysis	Traffic Calming
Critical Area Analysis	Management of the Chesapeake Bay Program	Truck Traffic Routing
Curb/Gutter Conditions/Repair/Permits	Master Planning	Utilities Key Sheet Mapping
CZMP Application	Master Roads Inventory/Street Segment Integration	Utilities Maintenance Programming
Data Distribution Applications	NPDES Stormwater Management	Utilization of Planimetric/Topographic Map in Lieu of Surveys
Data Maintenance Applications	Nutrient Reduction Strategies	Vacant Land Analysis
Data Quality Control Applications	Open Space Analysis	Water and Sewer Amendment Process
Data Query and Display Application	Park Development Siting	Water and Sewer Pumping Stations
Demographic Analysis	Patron Analysis	Water Quality Monitoring
Detour Plans	Pavement Cuts Permits	Watershed Planning/Management
Development Review and Tracking	Pavement Marking Inventory (Re-stripping)	Work Order Management
Districting	Preliminary Alignment Studies	Zoning Hearing Case Development and Analysis
Down Zoning	Property Analysis	Zoning Layer
Drainage Complaint Investigation	Public Access	Zoning Review Cases

Appendix C. STELLA Software and NB-Sim Model Details

The STELLA™ software produced by High Performance Systems, Inc. (HPS)[7] is a program that enables users to construct models of physical, social, or information systems and simulate changes in those systems over time. This is extremely useful software because, in addition to allowing fuzzy mental concepts to be translated into tangible systems diagrams and mathematical relationships, it can generate results of the system dynamics and allows a wide range of experiments with those systems. The differential equations and other mathematical complexities are transparent to the user, who only has to draw the simple model diagram and define a few mathematical relationships.

STELLA is based on the simple idea of stocks—those things that can accumulate—and flows—the fluxes into and (or) out of the stocks. The user draws boxes for stocks and thick, pipe-like arrows representing flows into and out of the stocks. The diameter of the pipes, and thus the rate of flow, can be adjusted by a third STELLA object, called a converter, which effectively opens or narrows the pipe and increases or decreases the rate of flow into or out of a stock. The operator sets initial conditions for stocks and flows and creates the mathematical relationships that underlie the relations between them. The inputs may be constant values, time-dependent functions, or functions that depend on values of other variables or inputs. The user also specifies the time-step interval with which to calculate the new values of model objects and the duration of the model run. Results are most commonly displayed as time-series graphs of the variables of interest but can also be shown as scatterplots, tables, or other types of data display.

We decided to use STELLA software as a way to operationalize the theoretical framework for benefits estimation. We realized that, because of the large amounts of uncertainty around the initial values and rates of change of such model inputs as the proportions of users in each tier, the number and type of applications, the costs and benefits of using spatial data in a set of applications, and the rates of application diffusion and innovation, we would not be able to produce a single, definitive answer to the net present benefit of *The National Map*. We knew that, despite a solid theoretical foundation for tracking the benefits and costs of developing and using spatial data provided through an on-line viewer and server, there were simply too many unknowns and too few empirical data sources to develop a single answer that would prove unshakeable.

The revised strategy aimed to provide a "best-guess" answer that would arise out of a rigorous analytical and mathematical simulation of the entire system, but also toward developing a simulation tool, so that readers could enter values and rates reflective of their own beliefs and observations of the system. We decided that STELLA would be the vehicle to create, run, and deliver that simulation. By giving the Office of Management and Budget a run-time version of the STELLA simulation model, into which examiners can load numbers of their own choice, we hoped to preempt concerns about bias or weaknesses in the data. We also hoped to demonstrate that, despite limitations in the data, the theoretical framework is rigorous and reliable, and can be used to track changes in the benefits and costs of *The National Map* over time, as better data and more observations become available.

Our task, then, was to convert the equations and variables in the theoretical framework into stocks, flows, converters, and equations in STELLA. In some cases, this was straightforward, but in other cases, there is not a direct one-to-one translation. The critical point, though, is to make sure that every variable and trend in equation 5 from section 2 is captured and represented somewhere in the model. Table C1 is a summary of each variable in the theoretical model's equation 5 and its analog in the STELLA model. Table C2 contains the baseline values for each variable in the model and gives a brief reference to or explanation of its source.

NB-Sim Model Steps

We named the STELLA computational simulation model "NB-Sim" for "Net Benefit Simulator." This section describes the steps NB-Sim takes in assessing the net benefits of *The National Map* system. The illustrations are screen-save images of the user interface page (fig. C1) and model wiring diagrams (figs. C2-C8). These figures should assist readers in following the processes involved in NB-Sim.

Initially, there are 3,078 counties arrayed across 3 tiers of sophistication. The majority of these counties (65 percent) are in Tier 3, whereas only 5 percent are in Tier 1 counties. The other 30 percent are in Tier 2. In the base case, none of these counties are users of data provided by *The National Map* because *The National Map* does not yet exist. The tiering structure described here relates to their existing capacity to use their own spatial data (or other non-*National Map* data), and the sophistication of the applications they do in a pre-*National Map* world. As *The National Map* gets implemented across the country, these counties may eventually become users of *The National Map* data. The applications it enables them to do can begin to diffuse across the country, new innovative uses for spatial data will be developed, and those too will diffuse, and the USGS can be fairly credited with the increase in net benefits these users derive from the improved data and distribution system.

There is also a fourth type of spatial data user, the Federal Government, which is a separate tier-like entity unto itself. As *The National Map* develops, Federal agencies are likely to experience similar increases in the amount, quality, and efficiency of their spatial data use, and will receive many of the same benefits. However, because the Federal Government represents a single entity performing tasks with data that are fundamentally different in scope than local uses, it made sense to simulate their use separately. After describing how we have

[7] STELLA™ is a product of High Performance Software, Inc. (URL http://www.hps-inc.com). Any use of trademarks in this paper is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Table C1. Variables in the theoretical and simulation models and their baseline values.

Description or explanation	Item in theoretical model	Analogous object(s) in NB-Sim	Equation or initial value
Net present benefit of *The National Map* program	NB_{TNM}	*TNM* Net Present Benefit	Calculated as Total Present Benefits – Total Present Costs
Total program cost (calculated)	C_{TNM}	Total Present Costs	Calculated from other inputs (next 4 lines)
		Years Needed to Develop	10 years
		Years of Development Funding	10 years
		Annual Development Cost	$25,000,000 / year
		Maintenance Cost	$5,000,000 / year (*during* and *after* development)
Discount rate	r	Discount Rate	0.032/year (3.2%)
Time horizon of interest	$t = 1 \ldots T$	Run Specs…Length of Simulation	30 years
Calculation time increment	Annual	*Dt*	1 year
Tiers (or states) of users and applications	s_1, s_2, s_3	Tier1/2/3 *The National Map* Users; Tier1/2/3 Applications; Federal Applications	Definitional
Initial tier distribution of all 3,078 counties	N/A	Tier1/2/3 Initial Fraction	0.05; 0.30; 0.65
Number of applications available to each tier at each time	$j_{st} = 1 \ldots J_{st}$	Tier1/2/3 Applications; Federal Applications	Initially: 5, 8, 10; 25 applications; calculated thereafter
Average net benefit per spatial data application in each tier and time	$\overline{v}_{st}$	Tier1/2/3 Mean NB per Application; Federal Mean NB per Application	Drawn from Normal dist'n; user sets mean (default = $1,000/$10,000 per app.); SD = 1/4th mean
Total number of applications that are implemented by each tier in each time period	ψ_{st}	Tier1/2/3 AppUse; Fed'l App Use	Tier1/2/3/Fed Application *Tier1/2/3/fed Users * P_{st}
Proportion of available applications that is implemented	P_{st}	Not explicit. Calc'd as at right→	Tier1/2/3 Diffusion * RANDOM(0,1)
Number of data users in tier at each time	$k_{st} \ldots K_{st}$	Tier1 Users, Tier2 Users, Tier3 Users	Initially: (3078 counties) * (Tier1/2/3 Initial Fraction); calculated thereafter
Rate of innovation of new apps	I_{st}	Tier1/2/3 InnovationRate; Federal Innovation Rate	2% per user per year (same in all tiers)
Proportion of diffusion completed in each time step	D_{st}	Tier1/2/3 Diffusion	Graphic function of *TNM* completion percentage
Rate of one-tier User advancement in a time step	A_{st}	Tier 2 To 1 Advancement, Tier 3 To 2 Advancement	1% of tier per year 1% of tier per year
Benefit of *The National Map*	B_{TNM}	Total Present Benefits	Calculated as discounted sum of net benefits of apps. from 3 tiers

Table C2. Default NB-Sim model inputs and their explanations.

Parameter	Value	Source/Explanation
Annual development cost	$25,000,000/yr.	Exhibit 300
Annual maintenance cost	$5,000,000/yr.	Assumption
Years of development needed	10 years	Exhibit 300
Years of development funding	10 years	Exhibit 300
Annual discount rate	3.2% per year	OMB Circular A-94
Tier 1 applications	5	Assumption; based on surveys and lit review
Tier 2 applications	8	Assumption; based on surveys and lit review
Tier 3 applications	10	Assumption; based on surveys and lit review
Federal applications	25	Assumption; based on surveys and lit review
Total counties	3,078	U.S. Census Bureau
Tier 1 fraction	5%	Judgment categorization based on FGDC Survey, interviews, literature review
Tier 2 fraction	30%	As above
Tier 3 fraction	65%	As above
Tier 3 to 2 advancement	1% per year	Conservative estimate from rates of spread of other technologies and industries
Tier 2 to 1 advancement	1% per year	As above
Tier 1 mean net benefit/application	$1,000/ application	Mean of the applications we found
Tier 2 mean net benefit /application	$1,000/ application	As above
Tier 3 mean net benefit /application	$1,000/ application	As above
Federal mean net benefit application	$10,000/ application	As above
Standard deviation for normal distribution	0.25 of mean	Assumption & experimentation
Tier 1 diffusion rate	Max = 90%	Shape from literature; values are assumed
Tier 2 diffusion rate	Max = 72.5%	As above
Tier 3 diffusion rate	Max = 49%	As above
Tier 1 innovation rate	2%/yr	Assumption; based on other technologies
Tier 2 innovation rate	2%/yr	As above
Tier 3 innovation rate	2%/yr	As above
Federal innovation rate	2%/yr	As above
Federal emergency application net benefit	$1M / event	Extrapolation from history
Emergency application occurrence rate	2% of years	Assumed, random draw

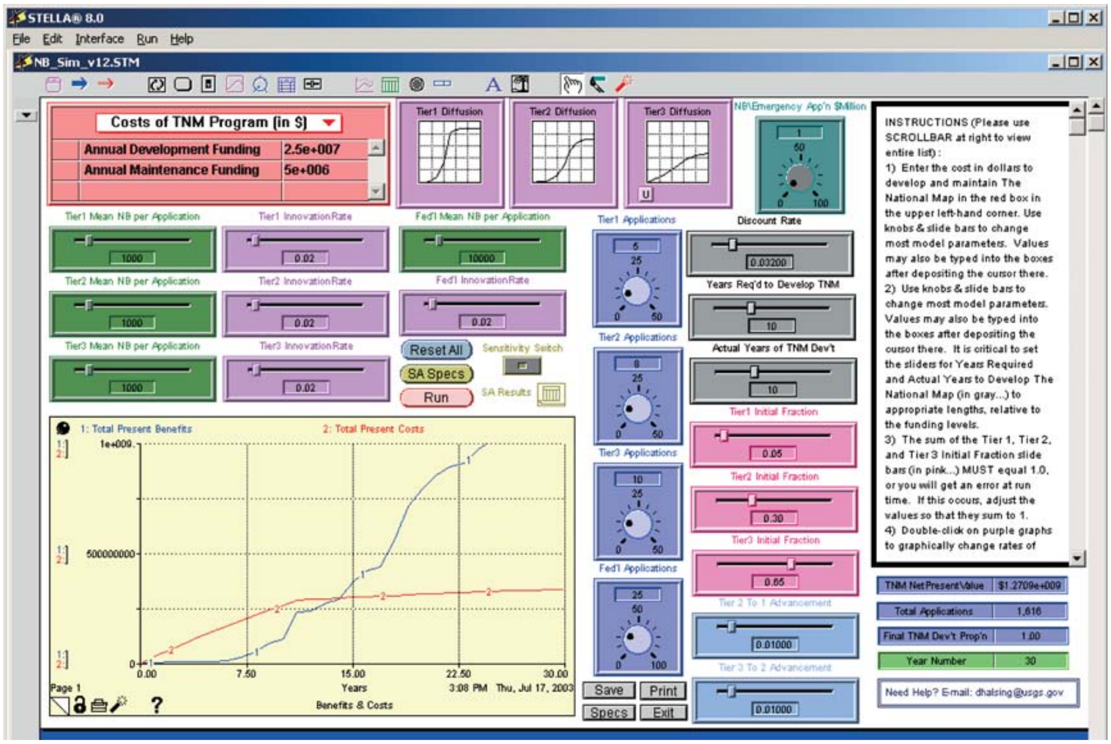

Figure C1. The user interface in the computational model, NB-Sim.

modeled what users in tiers at the county level experience, we will describe the analogous Federal transitions.

The development of *The National Map*, as well as standardizing and updating the data it contains, will take time, so the benefits alluded to above do not begin to accrue right away. Rather, there is first a process by which the "Counties" become "TNM Users."[8] As is the case throughout NB-Sim, the three tiers are handled separately so that the user may assign tier-specific initial values and rates of change (we decided to use the same values in each tier, but the model allows for different inputs than our baseline). The net benefits from each tier are summed at the end, but maintaining separation until then allows more flexibility and specificity in modeling. Note that the discussion below describes mainly tier 3, but conceptually identical processes take place in tiers 1 and 2. This transition between states is modeled as a flow between two stocks—"Counties" and "TNM Users." The flow is dependent on time elapsed and on the proportion of *The National Map* that is completed at that time step. At this stage, the choices of the simulation operator in defining *The National Map* program are critical (to avoid confusion with the term for counties that apply data from *The National Map*, called "TNM Users" or "Users," we refer to people running simulations with NB-Sim as "operators"). The operator works with the NB-Sim interface page (see fig. C1) to make simulation decisions as described below.

On the interface page, the operator inputs values for the "Years Req'd to Develop TNM" and the "Actual Years of TNM Dev't." These are not the same thing. The former is a best estimate of how long *The National Map* would take to be fully constructed given the operator-determined annual funding level; the latter is an expression of belief about how many years of funding it will actually receive. From these two settings, NB-Sim will determine a "Final Development Proportion," which is simply the ratio of years funded to years required. This is the percentage of *The National Map* that will be achieved at the end of the time period for which development funding is achieved. NB-Sim also calculates an "Annual Implementation Rate," which is the ratio of Final Development Proportion to the Years of Development Funding. This essentially takes that development which will be funded and spreads it out over the appropriate time frame. Finally, NB-Sim calculates a running total of "TNM Percent Complete" that is updated each year. This comes from iterating the Annual Implementation Rate each year and calculating how much of TNM is actually in place at the end of each time step, capping this at the Final Development Proportion.[9] Figure C2 shows these modeling steps.

The transition from a "Tier 3 County" to a "Tier 3 TNM User" is accomplished using the variables described in the

[8] Note that throughout this section, we use the acronym "TNM" for *The National Map* because STELLA allows a limited number of character spaces to name its model objects. The acronym is used only in reference to model objects, and not to the actual *The National Map*.

[9] Incidentally, the operator also sets an "Annual Development Funding" and an "Annual Maintenance Funding" to represent the number of dollars that will be spent in first the development stage and then the maintenance stage of *The National Map*. From these two inputs, combined with the Years of Development Funding and the operator-selected Discount Rate, STELLA develops an annual cost and a running total of program costs discounted to present value.

above paragraph. This flow from the County stock to the TNM User stock proceeds at a rate dependent on the number of counties in the Tier 3 Counties stock and the fraction of *The National Map* that has been completed at each point in time. Each year, *The National Map* is advanced toward its completion by including more of the country's spatial data in its database, and, gradually, more counties make the transition from Counties to TNM Users (see fig. C3). When the Final Development Proportion has been reached, this flow stops, and any counties left in the County stock stay there, never to have *The National Map* vision realized in their county. If *The National Map* is funded for as many years as the operator believes it will take to develop, then adequate time exists for all counties in all three tiers to be converted to TNM Users. Funding TNM for fewer years than necessary will not allow it to be fully developed in all places, and thus

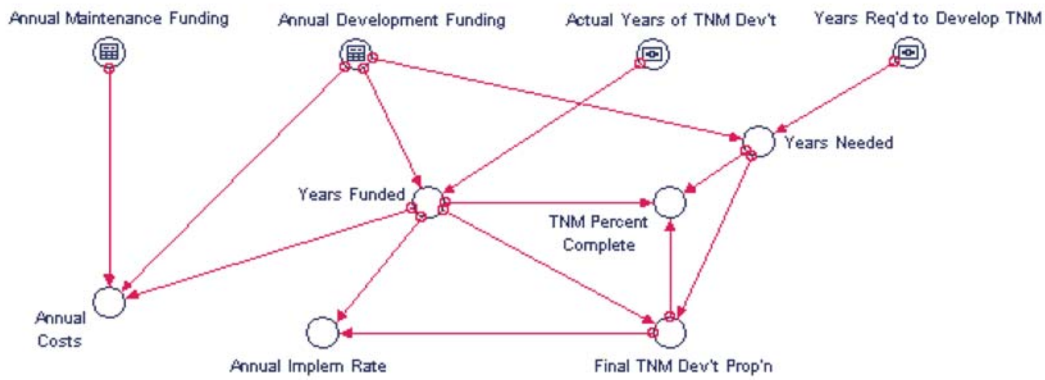

Figure C2. Wire diagram illustrating underlying dynamics of the NB-Sim model. Operator inputs for years and amounts of funding for *The National Map* (TNM) (Annual Maintenance Funding, Annual Development Funding, Actual Years of TNM Dev't, and Years Req'd to Develop TNM) allow calculation of Annual Costs, implementation rate (Annual Implem Rate), and project completion fractions (TNM Percent Complete).

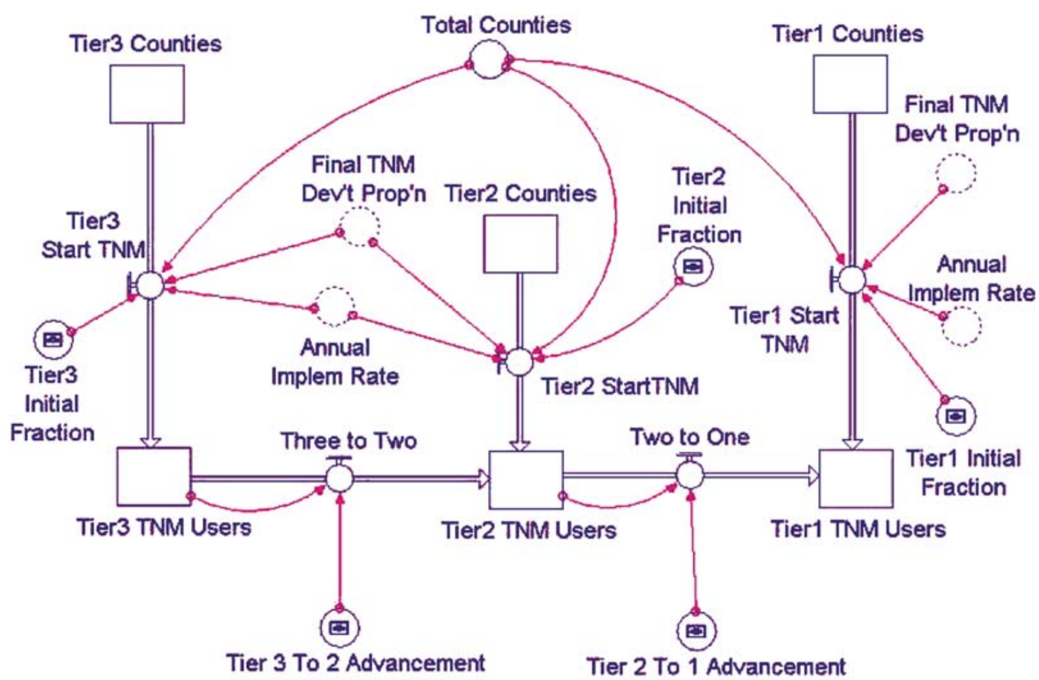

Figure C3. Wire diagram illustrating underlying dynamics of the NB-Sim model. As *The National Map* (TNM) is built, counties in each tier become TNM Users. Subsequently, some Users advance upward through the tier system. The rates of these transitions are determined by the annual implementation rate (Annual Implem Rate) and final development proportion (Final Dev't Prop'n).

fail to change all Counties to Users, and to reap many of the benefits *The National Map* could eventually provide.

The discussion thus far has centered on the rate of *The National Map* implementation and the gradual inclusion of the data from across the entire Nation into the data distribution system. This is the first step in the larger task of estimating the total benefits of *The National Map*. As soon as this process begins, several others begin and proceed simultaneously, overlapping throughout the entire process. This is properly representative of reality, as no one is likely to wait until *The National Map* is completely implemented everywhere before starting to use the data in places where it exists. The following paragraphs describe these other processes covered by NB-Sim.

Once a group of Tier 3 TNM Users exists, the members of the group may begin using *The National Map* data in a set of existing applications. There is a similar, though smaller, pool of existing Tier 2 Applications and an even smaller pool for Tier 1 Applications. The initial sizes of these pools of existing applications can be entered into the model. In each time step, some TNM Users will implement some of these Applications. It may appear that an estimate of total annual application utilization could be obtained simply by multiplying the total number of Users by the number of available application implementations within each tier, but this would be incorrect.

First of all, not every TNM User knows of every possible application. They will need time to develop both the awareness and the organizational commitment to implement certain applications. We have modeled this time lag as Diffusion. Based on concepts found in the academic and technical literature, we have modeled Diffusion as being a sigmoid curve with cumulative diffusion plotted as a function of time. The curves first increase slowly, then more rapidly, and then level off as *The National Map* achieves completion (see fig. C9). The graphs of cumulative Diffusion are not rates-per-unit-time, but can instead be thought of as a degree of "market penetration" of the set of applications. Even though applications exist and *The National Map* provides data to drive them, it takes time and a fairly complete *National Map* for these applications to see widespread use across the country. We expect that the more complete *The National Map* is at any point in time, the more quickly and completely applications will diffuse across the Nation. We also expect that Tier 1 TNM Users will adopt new applications sooner and more commonly than Tier 2 TNM Users who in turn are faster and more complete than Tier 3 TNM Users. Alternative rates of tier-specific Diffusion can be redrawn on a graph of Diffusion versus TNM Percent Complete on the operator interface page.

Note that this concept of Diffusion is separate from the transition from County to TNM User described above. That transition was about inclusion of location-specific data into *The National Map* system; whereas the diffusion process is about information awareness, organizational inertia, and the general notion that even most ready and capable data users will still take time to implement changes in their use of technology.

As alluded to above, and shown in figure C4, multiplying the numbers of TNM Users and Applications in each tier together to yield a number of implemented Applications is too simple of a solution. Instead, this product must be multiplied by the cumulative Diffusion (a fraction between 0 and 1) to represent the availability of all applications across the country. This fraction comes from graphic operator-determined inputs like that shown in figure C9.

However, even with incorporation of Diffusion, this result would be an overstatement of application use. Not every application is needed in every place. Examples of this location-specific need-limitation are that no snowplow routes need to be planned in Southern California and no hurricane evacuation routes need to be planned in North Dakota. Further, not every application needs to happen each and every year. This is a frequency-related need limitation. Given both of these limiting factors on the overall rate of Application use in a tier, it is extremely unlikely that the fraction of available applications that is actually implemented in a single year would ever approach one. Even though the necessary data and capacities may exist, the "audience" for these applications in a given place in a given year is likely to be much closer to zero than to one. Further, this fraction of available applications that are implemented is likely to vary from year to year.

To account for all of these issues, we created a process in NB-Sim to generate a random number between zero and one for each tier and each year during each model run. This random number is not drawn from the uniform distribution but from the inverse exponential distribution with an alpha parameter (α) of 6, which skews the distribution very close to zero. Most of the random draws are less than 0.25, and it is rare that more than 10 percent of the draws are greater than 0.50. On the basis of our interviews with state and county GIS managers and our review of the literature, we believe this to be a conservative estimate for need-limited use frequency of applications. This method also produces a variance in total application use from year to year as happens in reality. This process is not included in the interface page, but appears in the model diagram as "Need Limit Prop'n 1" (or 2 or 3, depending on Tier).

To arrive at a final annual application use rate specific for each tier, we multiply this random number by the product of the previous multiplication step (which was Number of Applications * Number of TNM Users * Cumulative Diffusion fraction). We will show how this tier-specific Application Use number is used shortly. First, we show how the number of Applications and Users in each tier can increase over time.

Another process we modeled was that of Innovation of new applications, also illustrated in figure C4. Recognizing that the existence of *The National Map* is likely to spur the development of new uses for digital spatial data, we built in tier-specific rates of Innovation. The intrinsic Innovation rate is a constant with a default value of 2 percent per user-year. While this per-user rate of new application development is a constant, because the number of users in each tier increases over time as *The National Map* is developed, the actual number of new applications created each year increases. Innovation increases the number of available applications available

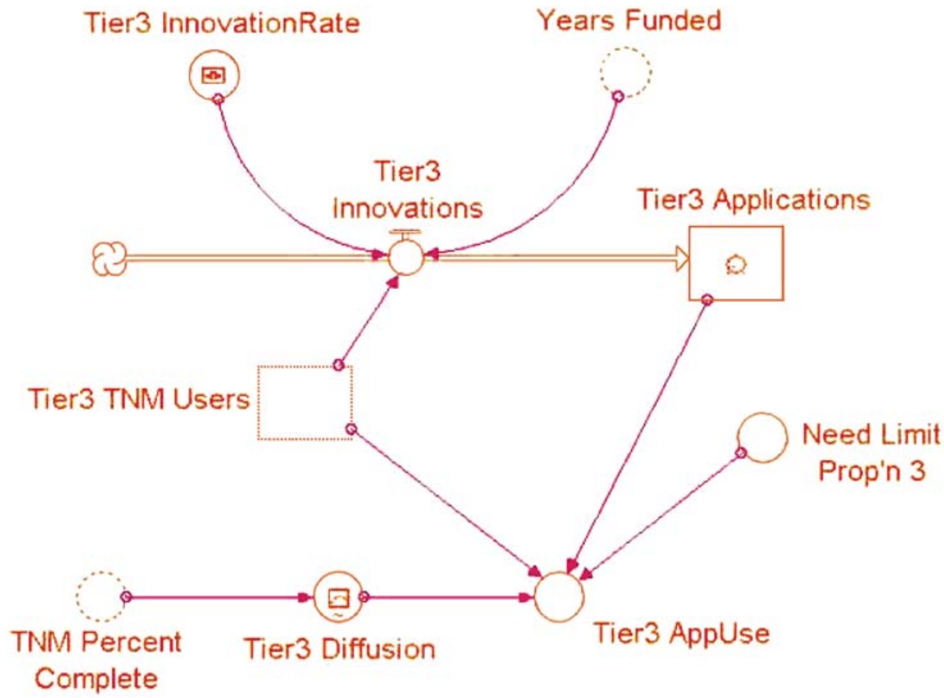

Figure C4. Wire diagram illustrating underlying dynamics of the NB-Sim model. The number of application implementations in a tier (Tier3 AppUse) is determined by the multiplication product of number of users of *The National Map* (TNM) (Tier3 TNM Users), number of applications (Tier 3 Applications), a cumulative diffusion fraction (Tier3 Diffusion), and a need limitation (Need Limit Prop'n 3). New applications are innovated at a constant rate (Tier3 Innovation Rate) multiplied by the number of users.

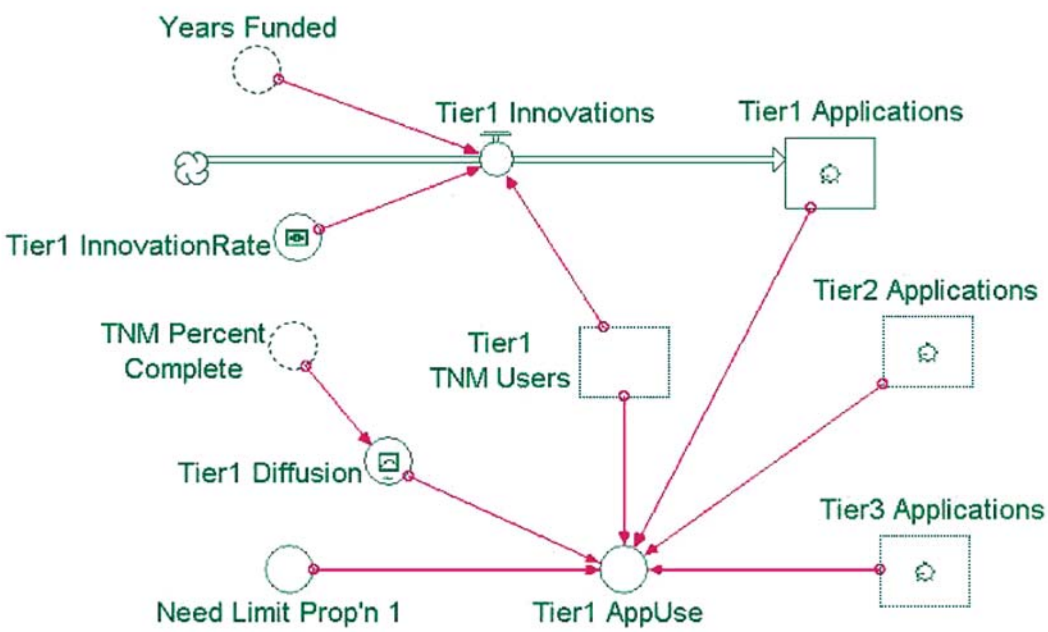

Figure C5. Wire diagram illustrating underlying dynamics of the NB-Sim model. The application innovation, diffusion, and implementation processes for Tier 1 users of *The National Map* (TNM). Note the availability of applications from lower tiers (Tier2 Applications and Tier3 Applications).

to each tier, but does not, on its own, increase the use rate of those applications. That is achieved through the Diffusion process and the transition from Counties to TNM Users.

In theory, some time after development and spread of *The National Map* is complete, innovation could again decline as all practical and useful applications are developed and implemented. Conversely, innovation could proceed indefinitely (though likely at a slow rate). We did not find any literature that addressed the “back-end” of this innovation curve, nor did we claim to know a maximum number of applications that could possibly exist at each tier. Any argument we would make based on analogous technologies would be suspect, at best. Further, whatever the magnitude of this effect, it is likely to have its influence well after *The National Map* has achieved a positive net benefit. Therefore, to maintain a conservative approach to this simulation, we decided that innovation could continue for 20 years after development funding stopped. After that time, the flow of new applications would go to zero.

Importantly, not all Applications are available to all Users. Tier 3 TNM Users may only utilize Tier 3 Applications. Tier 2 TNM Users may utilize Tier 2 and Tier 3 Applications, and Tier 1 TNM Users have the technical expertise and data to make use of Applications from all three tiers. Figure C5 illustrates the Application Use portion of the diagram. Comparing it to Tier 3’s Application Use diagram (fig. C4) reveals two more stocks, representing the additional availability of tier 2 and 3 applications. A Tier 1 User thus has a greater number of potential applications to implement.

One process related to this idea is that of Advancement through the tier system, represented in figure C3. As *The National Map* provides low-cost data, a library of spatial data applications and (or) technical support and collaboration, it is likely that many TNM Users at tiers 2 and 3 will develop greater expertise with applications of digital spatial data. As they do, they will be likely to embark on increasingly sophisticated and complex applications. This could upgrade them from tier 3 to tier 2 or from tier 2 to tier 1. Thus, the number of Tier 1 Users is expected to grow over the life of *The National Map*; the number of Tier 3 Users is expected to shrink; and Tier 2 Users are expected to increase then decrease over time.

Both of these stages of Advancement increase the number of potential applications available to TNM Users. This Advancement process of TNM Users moving through the tiers is separate and distinct from that of Counties becoming TNM Users. On the user interface, there are operator-selected rates of TNM Users advancing in each time step. These are labeled “Tier 2 to 1 Advancement” and “Tier 3 to 2 Advancement.”

By way of summarizing the discussion thus far, in each tier and in each time step, we have:

- A process of Counties becoming TNM Users.
- TNM Users adopting some of a pool of existing applications.
- The breadth of use of this pool of applications increases through the Diffusion process.
- The numbers of available applications increases through the Innovation process.
- The Innovation, Diffusion and County-to-TNM User processes are all dependent on the degree of development of *The National Map*, which itself is dependent on time and the relative Actual and Needed funding for development.
- Many TNM Users advance upward through the tier structure, thus increasing the number of applications available to them.
- The number of TNM Users, the number of Applications, the amount of cumulative Diffusion, and a random number between zero and one are multiplied together to produce a total number of Application Uses occurring in each tier and time period.

As discussed in the analytical framework section, we decided it would be a fair assumption to multiply an estimated number of Application Uses in each tier and time by an average net benefit of *National Map* data as they would be used in those applications. The alternative would be to determine an exact dollar value for each application and summing those net benefits over the number of applications we could find. This did not seem realistic, so our averaging approach seemed the best option.

Because the bundle of implemented Applications is likely to be different from year to year, it seemed inappropriate to use a constant value for net benefits. Drawing random values from a normal distribution centered on an accurate mean seemed more rigorous. The software features a built-in command to draw a random number from a normal distribution, wherein the programmer can select the mean and standard deviation. We developed an average value to use as the default mean for these random draws on the basis of survey results, interviews, and our literature review (see appendix B for more details on this choice). The standard deviation is set at one-fourth of the randomly drawn value. We felt this produced sufficient variation without causing the values for net benefits to skyrocket and plummet in consecutive years. Additionally, it does allow the potential for very small or, on rare occasions, negative average changes to per-application net benefits. The interface page allows the operator to alter the value for the mean net benefits by simply moving a slide bar on the screen. This value will not be used by the simulator every year, but will form the mean for a distribution from which the random draw will occur in each model simulation year.

In each tier, the number for Application Uses is multiplied by the randomly generated dollar value for Mean Net Benefit Per Application to produce a tier-specific annual dollar value for benefits of *The National Map*. These are summed across tiers in each year and discounted to present value by the user-adjustable Discount Rate. All of this creates a flow of Discounted Annual Benefits into a stock of Total Benefits in Present Value.

Earlier, the inclusion of a “fourth tier,” the Federal use of spatial data, was mentioned (see fig. C6). This simulation of Federal spatial data use is similar to what happens at the county level, with some notable exceptions. First, there is only one User (in our model, all Federal agencies fit into one "User" group), so no Diffusion is necessary. Second, there is no County-to-User transformation, so Federal users begin reaping benefits from *The National Map* right away. Also, the average change in net benefit of an application of spatial data is likely to be much higher at the Federal level than at the county level. The model uses a random draw from a normal distribution, and the operator has the ability to set the mean of that distribution, but the default is $10,000 per application implementation rather than $1,000.

There is one other important difference. Recognizing that there are occasional national disasters and emergencies, the responses to which would involve widespread use of digital spatial and could be greatly improved by *The National Map*, we attempted to capture this dynamic in the model. The September 11, 2001, attack on the World Trade Center is an example of the sort of emergency use of spatial data of interest here. When such events occur, spatial data will be used in a way that is beyond the normal purview of any Federal agency. We wanted to capture these random non-annual benefits.

To do this, NB-Sim generates a random number between 1 and 100 for each model year. If the number is greater than 98, an added net benefit of data use is added to the normal annual net benefits. That is, the default of this added benefit is $1 million dollars each year the emergency occurs (which should be in no more than 2 percent—or once every 50 years—on average), but the operator has the ability to change this value as well. It would of course be possible to annualize these benefits, but averaging out these improvements does not capture the stochastic nature of this benefit stream.

The number of Federal "Application Uses" is multiplied by the randomly drawn "Average Net Benefit per Application Use" to yield a total "Federal Net Benefits" in each year. To this product, the "Net Benefit per Emergency Application" is sometimes added. The annual Federal Net Benefits are added to the total of Net Benefits from the three other tiers to produce an annual "Sum of Applications Net Benefits" in each year. Figure C7 illustrates this process.

While this is happening on the benefits side of the equation, a much simpler process is happening on the cost side (see fig. C8). The values for Annual Development Funding and Annual Maintenance Funding (set on the interface page…) are adjusted to present value by the same Discount Rate[10] and flow into a growing stock of total costs in present value. For the initial years of *The National Map*, the model uses the figure for Development Funding plus the figure for Maintenance Funding, after which it switches to Maintenance Funding only, a much smaller figure. The switch between these two funding

[10] By default, the discount rate was set at 3.2 percent per year based on OMB's guidance for 30-year projects in real dollars. The user can adjust this rate with a slide bar on the user interface page.

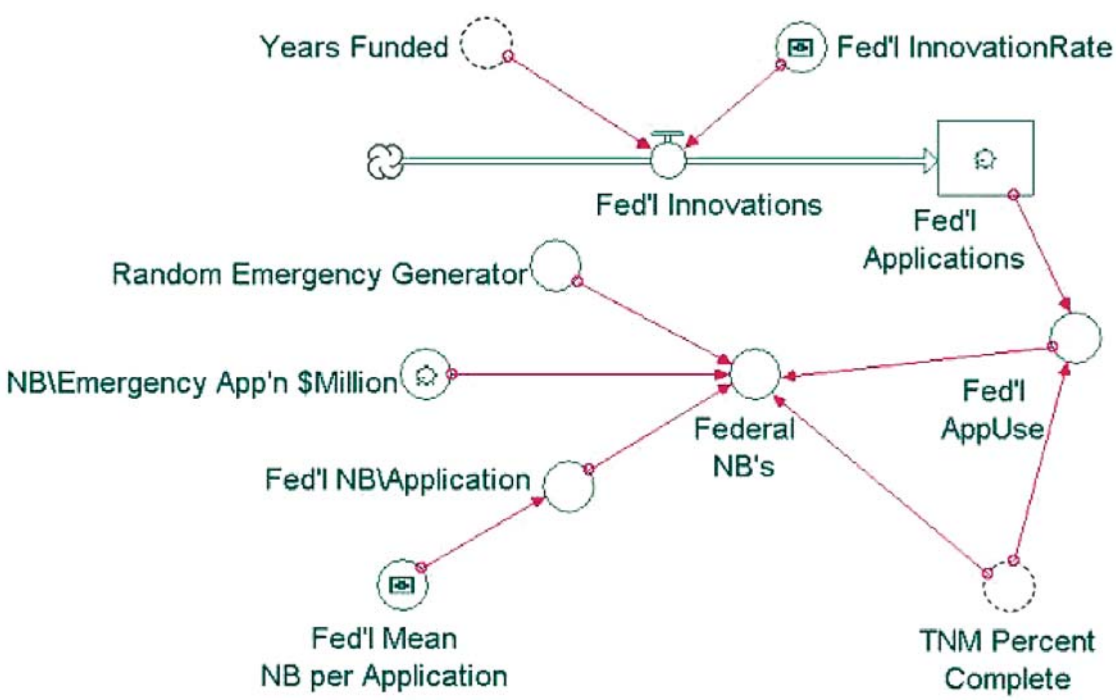

Figure C6. Wire diagram illustrating underlying dynamics of the NB-Sim model. The innovation and implementation of Federal applications (Fed'l Applications), including occasional "emergency" applications of spatial data (Random Emergency Generator).

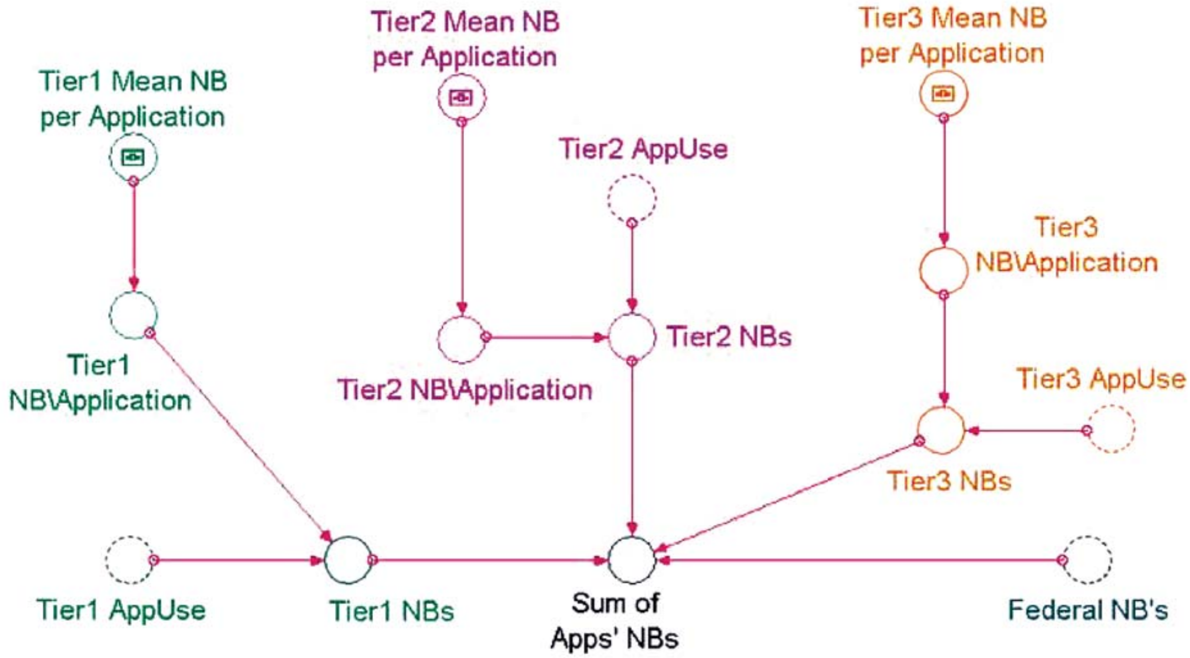

Figure C7. Wire diagram illustrating underlying dynamics of the NB-Sim model. Numbers of implementations (Tier*n* AppUse) in each tier are multiplied by a randomly drawn value (Tier*n* NB\Application) from a normal distribution centered on an operator-determined average (Tier*n* Mean NB per Application) to yield tier-specific net benefits (Tier*n* NBs). The three tiers and the Federal data uses are then summed to yield annual benefits (Sum of Apps' NBs).

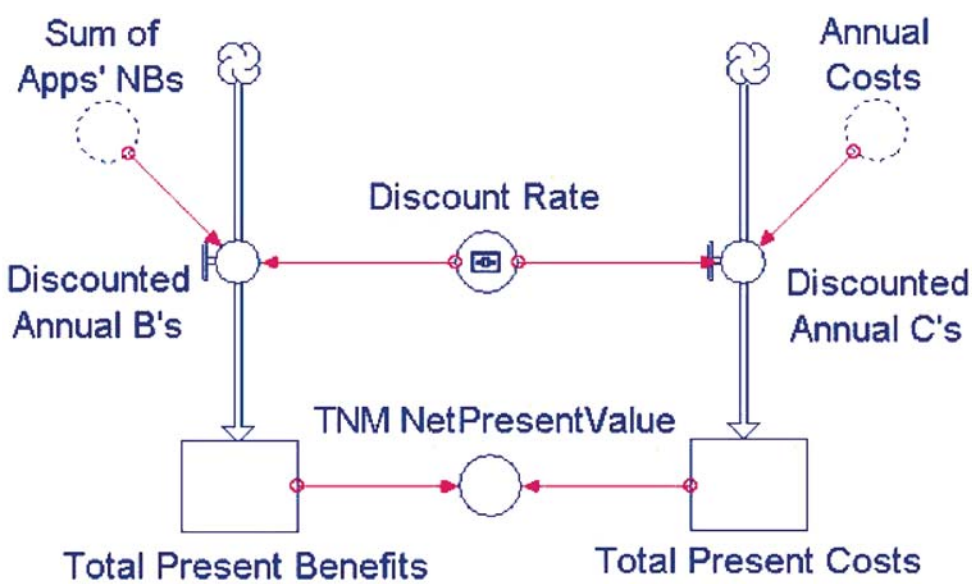

Figure C8. Wire diagram illustrating underlying dynamics of the NB-Sim model. Each year, the total benefits (Sum of Apps' NBs) and annual program funding (Annual Costs) are discounted and flow into accumulating stocks of Total Present Benefits and Total Present Costs. The difference between them is the net present value of *The National Map* (TNM NetPresentValue).

levels is set at 10 years as a default, but is adjustable by the operator on the interface page.

The difference between the two stocks, Total Present Benefits and Total Present Costs yields the Net Present Benefit of *The National Map*, also shown in figure C8. The graph on the interface page is set up to show these benefit and cost streams as they change over time. Where the curves cross, the net present benefit equals $0. The year in which this occurs can be read out at the bottom of the graph. Continuing the simulation beyond this point allows further surplus net benefits to accrue to society, the USGS, and *The National Map*.

There are other graphs available for viewing. The operator must simply click the small "dog-ears" in the lower left-hand corner of the yellow graph to flip through the other graphs. These other graphs can reveal a curve of net benefits over time, the number and distribution of TNM Users by tier, the number of Applications in each tier, the net benefits of applications by tier, and the degree of completion of *The National Map*.

It is also possible to conduct systematic sensitivity analyses quickly and easily. That topic is covered in appendix D.

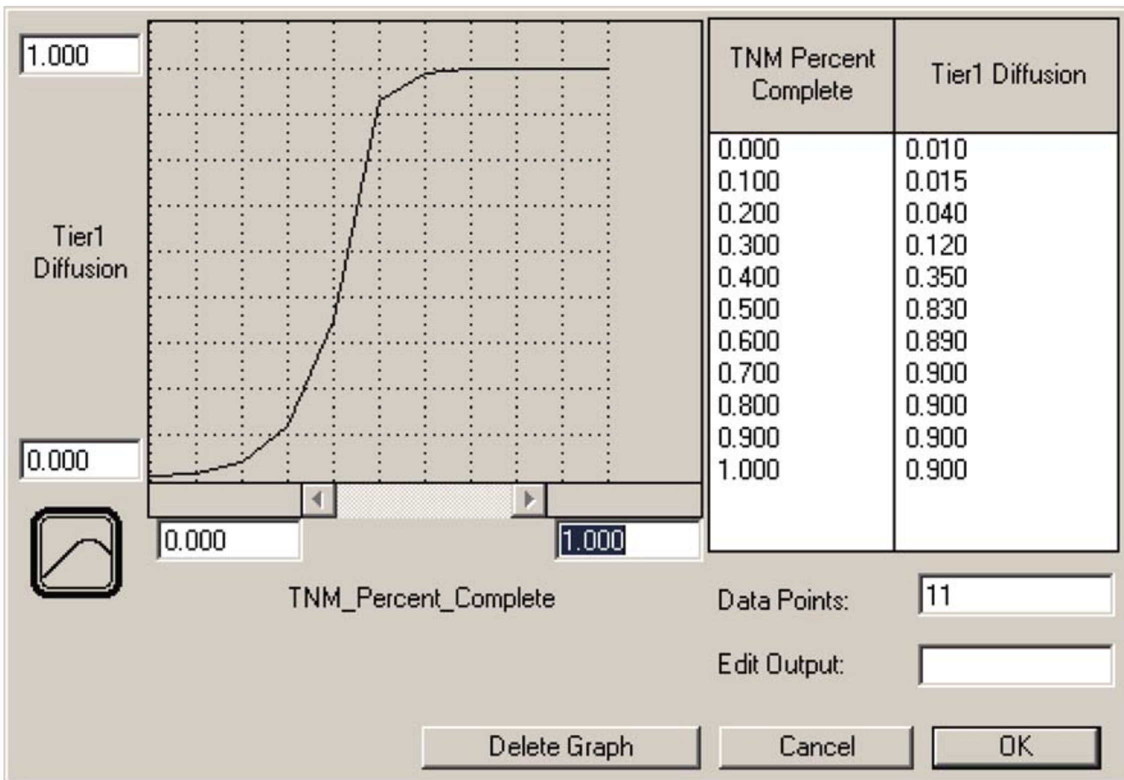

Figure C9. The input window for graphically entering non-constant rates of change. In this case, Tier1 Diffusion is not a constant rate, but is dependent on how complete *The National Map* is (TNM_Percent_Complete).

Appendix D. Sensitivity Analysis

Methods

Because of the large ranges of uncertainty around many of the input variables to this analysis, it is important to measure how sensitive the results are to changes in the "best guess" estimates. To conduct this sensitivity analysis, we created "scenarios," which are combinations of input variables that are different from the baseline scenario. In some scenarios, only one variable was changed relative to the baseline; in others, many were altered. We ran optimistic and pessimistic versions of each single-variable change and of many multiple-variable modifications.

As in the main body of the report, each model simulation runs for 30 years, because that is the most relevant time horizon of interest. Conservatively estimating that *The National Map* will take 10 years to develop and populate with data, many of the scenarios do not break even until around year 12 or 16. After that, substantial positive net present values (NPV) are generally recognized, but the realm of information technology progresses so rapidly that assuming knowledge of benefits streams 40 or 50 years into the future seems unrealistic. The 30-year analysis horizon seemed a fair compromise between these extremes.

Because the NB-Sim model contains several random components, a single run was insufficient to provide a result in which complete confidence could be placed. It was important to run a number of simulations and use NPV of benefits from each run to calculate the mean and standard deviation of NPV for that scenario. Similar mean and standard deviation calculations were included for the break-even time. Each scenario was run 50 times, to reduce the likelihood that a particularly high or low draw from any of the random model components would disproportionately affect the end results. The means and standard deviations are taken from these 50 runs.

Table D1 contains a list of the scenarios we have included in the sensitivity analysis. The table is organized along scenario number. Following each number, there is a description of the changes—again, relative to the baseline values—that were made. If no variable is listed in the description, it was no different than the base case. After the descriptions, columns list the means and standard deviations around the net present value and the break-even time (in years) of each scenario. There is also a column for the break-even rate, which is the percentage of runs in which a positive net benefit was achieved within the 30-year time horizon. If not all of the runs in a scenario show a positive net benefit, then no average break-even year can be calculated, which is why those columns are labeled with are labeled with "NA" in those scenarios. Figure D1 shows the mean NPV of each scenario with error bars representing the two-standard deviation range (roughly the 95 percent confidence interval) around those means.

In total, there were 60 scenarios simulated. These represent two sets of 30 scenarios that are identical except for the value entered for Annual Development Cost. Group 1 consists of 30 scenarios with the Annual Development Funding set at $25,000,000 per year. Scenario 1 is the baseline or "best-guess" scenario. The other 29 are described below. Group 2 consists of the same 30 scenarios—with Scenario 31 being a "baseline" for this group—except funded for development at $50,000,000 per year. These two different values were chosen because, even though the original Exhibit 300 for the entire *The National Map* program asked for $25,000,000 per year for 5 years, no one knows the exact cost of building a complete version of it. We wanted to compare many different input scenarios against cost figures that covered a wide range of possibilities. The two values we used were regarded as reasonable as any by the Implementation Team for *The National Map*.

Sensitivity Results Discussion

Table D1 and figure D1 reveal several important insights. First, most scenarios show a positive NPV, even those that are far more pessimistic than even the highly conservative baselines. Further, most scenarios in group 1 take between 10 and 15 years to break even. Scenarios in group 2 take longer; generally 15 to 20 years. Looking more closely at the scenarios in group 1, with regard to the impact of specific variables or combinations, the changes relative to the baseline of $2.05 billion and a 14-year break-even time are as follows:

- The Innovation rate matters quite a bit. Scenario 2 assumes there is no flow of new applications into any of the tiers, and its NPV is negative in every run. If the Innovation rate is 5 percent instead of its baseline of 2 percent (scenario 4), the NPV more than doubles to more than $5 billion, and the break-even time moves in by 3 years.

- Cumulative Diffusion also heavily affects the results, but not as much. Eliminating Diffusion (scenario 3) so that Users will only implement those applications they already do or that they create for themselves, drops the NPV to only $100 million. Conversely, making Cumulative Diffusion immediate and complete—so that all Users have immediate access to every application as soon as it is created—increases NPV by more than $1 billion, and moves the break-even date in by 2 years (scenario 19). Keeping the sigmoid shape of Diffusion, but either halving (scenario 5) or doubling (scenario 6) its rate at each point, decreases or increases the NPV by about $1 billion.

- Scenarios 8 and 9 assume a constant total budget for development funding but either spread that amount out

Table D1. Sensitivity analysis scenarios and results. The 60 different scenarios are described and the mean and standard deviation of both the net present value (NPV) and the break-even time are reported, along with the percentage of each scenario's model runs that achieved a positive NPV. Because of NB-Sim's randomness, even a scenario with a positive NPV may not break even in every run.

	Scenario Number	Description	Mean Net Present Value (NPV) *2001$*	Std. Deviation NPV *2001$*	Breakeven rate *%*	Mean Breakeven *years*	Std. Deviation *years*
Group 1: Annual Development Funding = $25,000,000 per year	1	Baseline	$ 2,051,719,005	$ 492,308,136	100%	14.0	1.9
	2	No Innovation	$ (271,515,685)	$ 11,248,142	0%	NA	NA
	3	No Diffusion	$ 103,925,239	$ 140,143,712	78%	NA	NA
	4	Innovation Rate = 5%	$ 5,425,730,132	$ 1,151,585,837	100%	11.2	1.3
	5	Diffusion = 1/2 of Baseline	$ 1,083,330,665	$ 279,926,803	100%	16.7	1.7
	6	Diffusion = 2 x Baseline	$ 3,234,312,771	$ 810,634,383	100%	12.0	1.6
	7	No Federal application use at all	$ 1,924,846,325	$ 500,557,643	100%	13.8	1.7
	8	Fixed total budget; 20-yr implementation	$ 992,682,531	$ 268,293,606	100%	19.2	1.8
	9	Fixed total budget; 5-yr implementation	$ 2,612,823,196	$ 600,438,861	100%	11.1	1.8
	10	No tier advancement	$ 1,934,175,893	$ 432,172,567	100%	14.0	1.9
	11	Tier advancement = 5% per year	$ 2,706,643,787	$ 530,867,804	100%	13.5	1.3
	12	Emergency fed'l application adds $50M	$ 2,003,886,405	$ 459,564,003	100%	14.1	1.9
	13	All counties are tier 1 initially	$ 4,028,577,651	$ 1,585,187,501	100%	12.2	2.6
	14	All counties are tier 2 initially	$ 3,248,466,604	$ 969,116,128	100%	13.0	2.2
	15	All counties are tier 3 initially	$ 2,018,828,229	$ 687,961,169	100%	14.4	2.6
	16	Initial tier distribution = 33% per tier	$ 2,543,773,440	$ 572,399,896	100%	13.1	1.8
	17	Initial number of applications = 1 per tier	$ 1,908,503,573	$ 493,567,616	100%	14.2	1.8
	18	Initial number of applications per tier = 4 x Baseline	$ 2,211,806,340	$ 500,067,700	100%	13.2	1.6
	19	Diffusion is immediate and 100% in all tiers	$ 3,107,374,056	$ 661,311,553	100%	11.8	1.5
	20	Alternative 2; Funding increase = $0	$ 22,698,680	$ 6,164,110	100%	2.0	0.0
	21	No randomness; Deterministic model	$ 2,967,654,414	NA	100%	12.0	NA
	22	10 years of development needed; get 5 yrs. Funding	$ 536,209,842	$ 103,616,250	100%	17.7	2.0
	23	20 years of development needed; get 10 yrs. funding	$ 143,593,366	$ 87,362,539	84%	NA	NA
	24	Avg. increase in net benefit per app'n = 1/2 baseline	$ 961,845,496	$ 272,647,663	100%	17.6	2.2
	25	No tier advancement; halved net benefits	$ 1,169,501,998	$ 276,976,548	100%	17.9	2.5
	26	No tier advance; halved net benefits; Innov. = 1%	$ 305,667,340	$ 112,447,397	100%	23.2	2.7
	27	No tier advance; halved net benefits; Inn.=1%; 1 init'l app	$ 122,181,021	$ 120,273,966	98%	NA	NA
	28	Net benefits per application = 2 x Baseline	$ 4,888,612,720	$ 917,964,995	100%	11.5	1.3
	29	No Innovation; maximum initial number of applications	$ 13,052,376	$ 62,143,573	56%	NA	NA
	30	2x net benefit gain/app; 20 yrs of dev't needed/received	$ 2,299,879,300	$ 649,018,101	100%	18.2	1.9
Group 2: Annual Development Funding = $50,000,000 per year	31	Baseline	$ 1,832,342,316	$ 461,444,159	100%	17.0	2.0
	32	No Innovation	$ (505,285,518)	$ 12,591,476	0%	NA	NA
	33	No Diffusion	$ (101,843,260)	$ 161,642,645	24%	NA	NA
	34	Innovation Rate = 5%	$ 5,333,869,822	$ 1,221,505,981	100%	12.8	1.3
	35	Diffusion = 1/2 of Baseline	$ 827,118,933	$ 295,357,078	100%	21.0	2.1
	36	Diffusion = 2 x Baseline	$ 3,045,901,436	$ 671,060,362	100%	14.6	1.7
	37	No Federal application use at all	$ 1,697,109,734	$ 525,170,848	100%	16.5	1.8
	38	Fixed total budget; 20-yr implementation	$ 873,732,652	$ 268,733,735	100%	21.8	1.7
	39	Fixed total budget; 5-yr implementation	$ 2,253,219,392	$ 554,344,427	100%	14.5	1.8
	40	No tier advancement	$ 1,609,427,876	$ 534,497,607	100%	16.9	2.0
	41	Tier advancement = 5% per year	$ 2,241,980,048	$ 634,952,851	100%	16.0	1.5
	42	Emergency fed'l application adds $50M	$ 1,853,781,868	$ 424,843,586	100%	16.5	2.1
	43	All counties are tier 1 initially	$ 3,523,315,329	$ 1,485,605,442	100%	14.7	3.1
	44	All counties are tier 2 initially	$ 2,967,702,450	$ 938,851,506	100%	14.8	2.5
	45	All counties are tier 3 initially	$ 1,750,167,695	$ 674,530,148	100%	17.3	3.0
	46	Initial tier distribution = 33% per tier	$ 2,012,047,371	$ 601,729,889	100%	16.6	2.1
	47	Initial number of applications = 1 per tier	$ 1,790,328,106	$ 443,785,544	100%	17.0	1.9
	48	Initial number of applications per tier = 4 x Baseline	$ 2,026,415,237	$ 612,295,267	100%	16.2	2.3
	49	Diffusion is immediate and 100% in all tiers	$ 3,203,586,772	$ 799,894,542	100%	13.9	1.8
	50	Alternative 2; Funding increase = $0	$ 22,698,680	$ 6,164,110	100%	2.0	0.0
	51	No randomness; Deterministic model	$ 2,731,559,572	NA	100%	14.0	NA
	52	10 years of development needed; get 5 yrs. Funding	$ 213,876,597	$ 91,349,262	100%	23.6	2.4
	53	20 years of development needed; get 10 yrs. funding	$ (169,518,000)	$ 76,754,160	0%	NA	NA
	54	Avg. increase in net benefit per app'n = 1/2 baseline	$ 636,962,952	$ 291,247,432	98%	NA	NA
	55	No tier advancement; halved net benefits	$ 591,947,009	$ 250,657,412	100%	21.7	2.1
	56	No tier advance; halved net benefits; Innov. = 1%	$ (21,969,199)	$ 116,557,903	38%	NA	NA
	57	No tier advance; halved net benefits; Inn.=1%; 1 init'l app	$ (34,102,146)	$ 109,251,344	36%	NA	NA
	58	Net benefits per application = 2 x Baseline	$ 4,259,828,228	$ 1,249,327,730	0%	NA	NA

over 20 years or compress it into 5 years. Essentially, these scenarios are a slower and a faster implementation, and it is clear that the faster *The National Map* is developed, the sooner and larger its benefits will accrue. Scenario 8 drops NPV to less than $1 billion, whereas the faster implementation in scenario 9 increases it to $2.6 billion.

- The Initial Tier Distribution and Tier Advancement Rates do affect the results, but less dramatically than might be expected. Eliminating advancement (scenario 10) reduces NPV by less than $100 million. Increasing the advancement rates to 5 percent per year (scenario 11) brings an extra $700 million, but only moves the break-even date in by 6 months. Changing the Initial Tier Distribution to one-third of all counties in each tier (scenario 16) makes only a $500 million improvement. Putting all counties into tier 3 (scenario 15) causes a reduction of only a few million dollars; whereas starting all Counties in tier 1 (scenario 13)—an admittedly impossible starting point—doubles the NPV to more than $4 billion. Starting with all counties in tier 2 (scenario 14) brings an NPV of more than $3 billion.
- The Initial Number of Applications also fails to produce dramatic changes in the final results. Dropping the starting number to just 1 application per tier (scenario 17) brings a decrease of $50 million, whereas quadrupling the initial number in each tier (scenario 18) increases NPV by about $200 million.
- The model treats Federal use of spatial data separately from county-scale use. To see if this modeling decision made a significant difference, all Federal applications net benefits per application use and innovation were eliminated (scenario 7). This brought a decrease of about $75 million, which did not indicate high sensitivity. Similarly, setting the additional benefit of a randomly generated emergency application to $50 million, instead of its default of $1 million (scenario 12), actually brings a lower NPV than in the baseline. It is about $50 million less, but the standard deviations are so much wider that these results are essentially equal. This apparent contradiction is explained by recalling that the need for the emergency application is randomly generated with a likelihood of 2 percent per year. Each scenario consists of fifty 30-year runs, and even a difference of one or two "emergencies" could explain this difference. The important insight is that this is a variable for which the exact value is not significant relative to other inputs.
- A very sensitive variable is the Increase in Average Net Benefit of an Application. In the baseline, each tier sees an additional $1,000 in net benefits each time an application is implemented, and the Federal

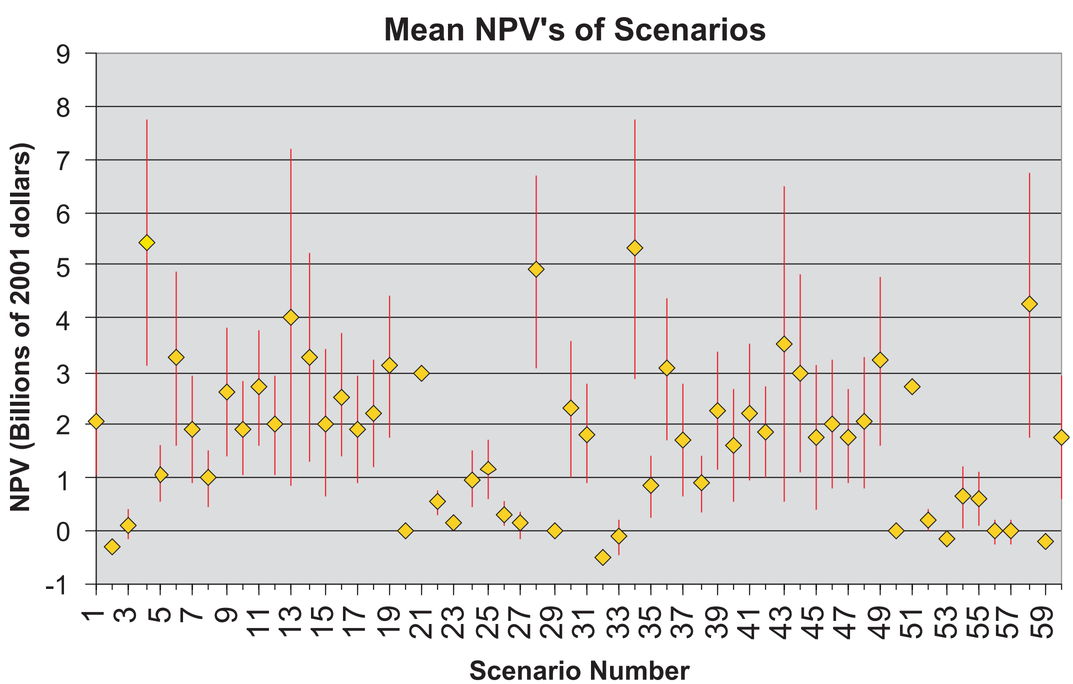

Figure D1. Sensitivity analysis results; mean net present value by scenario. The 60 different scenarios were each run 50 times. Results from each scenario are plotted here, as the mean net present value (NPV, yellow diamond) and the two-standard-deviation range (approximately the 95-percent confidence interval; red bars). Note that not all scenarios reach a positive value.

Government receives an improvement of $10,000 per application. If those values are halved (scenario 24) or doubled (scenario 28), the NPV drops to less than $1 billion or increases to almost $5 billion.

- Two scenarios addressed the effects of partial development of *The National Map*. These examined what would happen if the program would need 10 years to develop at a certain funding level, but only receives 5 years of funding (scenario 22), or if it needed 20 years of funding but received only 10 years (scenario 23). Both of these scenarios broke even on average, but scenario 23 only achieved a positive NPV in 84 percent of the model runs. Also, the NPV dropped to $500 million and $100 million in the two scenarios, underscoring the importance of funding *The National Map* to completion.

- We considered combinations of pessimistic changes in variables. Scenario 25 cuts the net benefits per application in half and eliminated tier advancement. This dropped NPV to $1.1 billion. Combining those changes with an Innovation rate of 1 percent (scenario 26) brings the NPV down to $300 million. Those changes and lowering Initial Number of Applications in each tier (scenario 27) to 1 lowers the NPV to just more than $100 million. These results show that the positive NPV of *The National Map* is robust to even a combination of extremely pessimistic assumptions.

- We also combined pessimistic and optimistic changes in the same scenarios. Scenario 29 eliminated Innovation, but set the Initial Number of Applications to 50 in each tier and 100 at the Federal level. The increased applications were not truly sufficient to replace the lack of new applications, as NPV dropped to only $13 million. In scenario 30, *The National Map* could be built much more slowly—taking 20 years—but if the net benefit per application doubles to $2,000 (or $20,000 for Federal applications), the NPV is almost $2.3 billion. Compared to Scenario 8—the 20-year implementation version of baseline—this brings an NPV that is more than twice as large.

- Finally, we investigated two particularly odd scenarios. In scenario 22, we attempted to simulate alternative 2 from the original Exhibit 300. That alternative would begin to implement *The National Map* with no additional funding, but only through an internal shift of funding to that task. We assumed that would take 50 years and adjusted the model accordingly. With all other values equal to baseline, the NPV decreased to $22 million. This is a very small number, but, because there was no initial investment made, the break-even year is year 1, when NPV's are zero dollars because no additional money has been spent.

- In scenario 21, we removed all randomness from the model, made Diffusion immediate and complete, and removed the random emergency application completely. The Need Limitation Proportions were set at 15 percent for each tier, meaning that in each year 15 percent of all diffused and available applications were implemented. That number was not based on any empirical belief and is perhaps overly optimistic. The net effect of these changes was to make every model run identical and the total benefits curve smooth, rather than kinked as it is in all other scenarios. The NPV jumped to $2.9 billion. Although it would be a mistake to put much faith in this result, though. If the Need Limit Proportion is 25 percent, then the NPV is more than $5 billion; if it is 5 percent, then NPV is only $700 million. In all other scenarios, this number is randomly drawn from a declining exponential function; here, that proportion is fixed. This non-random model is fundamentally different from the baseline or any other version of the model, and comparing them would probably not be fair or accurate.

In the group 2 scenarios, the general directions and the magnitudes of changes seem to be largely the same as in the group 1 scenarios, so that discussion will not be repeated here. It is important to note, however, that in some cases (scenarios 33, 53, 56, 57, and 59), a scenario that produced a small, but positive, NPV in group 1 will bring a negative NPV in group 2 because of the added development cost. In general, as noted above, the projected NPV was smaller and the break-even period came several years later.

The most important insight to be gained from this sensitivity analysis is that the exact values used in the default case are less important than recreating a fair representation of their interactions. Even with extremely conservative assumptions of initial values and rates of change, *The National Map* can be expected to bring about $2 billion in net benefits to users of the data it will provide. This exact value cannot be expected to be fixed, but seems relatively robust to major changes to all but the most sensitive variables.