WORKSHOP ON THE DESIGN OF A NETWORK OF OBSERVATORIES FOR THE SOCIAL, BEHAVIORAL, AND ECONOMIC SCIENCES,

MAY 15-16, 2012

Westin Hotel, Arlington, VA

Emilio Moran

Indiana University and

University of Maryland

Moran@indiana.edu

Sandra Hofferth

University of Maryland

Hofferth@umd.edu

September 28, 2012

Funding was provided under Grant SES1052961 from the National Science Foundation, Workshop on the Future of Observatories in the Social Sciences.

**INTRODUCTION**

This workshop followed earlier meetings held in 2005, 2007, 2008, April 2009, December 2010, October 2011, and February 2012 that sought to advance the capacity of the SBE sciences in terms of theory, methods, and cyberinfrastructure for the next 20-30 years. The previous three workshops focused on defining what a national network of observatories might look like, and, in the last two, attendees began to define the requirements for collecting data at various levels that would ultimately provide a representative picture of the United States. National, regional and local data of interest to the full network of observatories could take many possible forms. At one extreme are core data extracted from existing large scale highly standardized national administrative data bases or newly collected observational data for nationally representative samples of people and places (e.g., surveys, extracts of data sources maintained at the state, county or municipal level (Tables 1 and 2). At the other extreme are specialized data sets of unique interest, applicability or availability for individual observatories.

Several challenges underlie this effort. One is the massive amount of data being generated by new sources, many of them from the Internet and social networking sites, data whose reliability, statistical validity, and generalizability are not established or known. Further, traditional data sources such as administrative, survey, and census data collected by statistical agencies and government have not been designed to facilitate interoperability and ease of use. These data issues, and the methods used to compare and integrate these data, are a matter of national importance. The turmoil of the Arab spring was not anticipated because there was an insufficient understanding of the new ways those populations were communicating with each other. How people communicate during crises such as tsunamis or hurricanes is still poorly understood and this knowledge fails to inform how best to coordinate responses to meet the urgent needs of our citizens in times of crisis. Better ways to capture social data, together with more context-specific knowledge have real and important economic and social consequences.

BACKGROUND

A key question in the design of a system of observatories is “how much standardization of core data and procedures is needed?” Too much standardization or too many restrictions on which data are acquired and how they are organized for use will stifle local creativity—making the observatories not much more than nodes in a master data collection scheme. On the other hand, if core requirements and standards for how data are collected, stored and shared are too relaxed, the system risks its ability to promote a national resource that social scientists can draw on or limits the efficiency with which one or more observatories can collaborate to develop a new line of data-based research.

The “national sample framework,” be it a national sample of Census tracts or other physical or administrative units, is simply a structure that binds the observatories’ most intensive and costly data collections into a manageable framework that when properly used will produce regional and nationally representative data. This same framework certainly can be used to organize all forms of data into a nationally representative sample (as the ACS currently does in its annual samples of small Census units). But if a data element of interest is available at the county level for all tracts within the observatories’ data capture region, does it makes sense to try to restrict the retention of the data to only a sample of small Census areas? If all of the data for larger or irregular units (counties, congressional districts, water districts, school districts) are captured, procedures will be needed to ensure that the data can be used at the observatory-level or more importantly combined across observatories in a way that supports reasonable (if not unbiased) inferences in the research settings where those data will be used. Presumably the individual observatories will know and observe the inferential limits of specialized data that they collect but how much time and effort should be devoted to developing a set of standards that facilitate inter-observatory uses of these data?

*February 2012 meeting*

In the earlier report from the February 2012 meeting we described some of the discussion and summarized what appeared to be consensus around two potential ways to determine the location and coverage of the observatories for the national sample: 1) one focused on population density and 2) one focused on place. We also summarized the discussion about the choice of census tract/block group as the basic unit. At the end of the February meeting there were three issues that still needed to be addressed:

1. Whether to focus on national vs. regional data. Core data would be collected on a set of measures. This could be produced from national administrative data or obtained through intensive local data collection. If the focus is on core data to represent the U.S.:
   1. Would this produce regionally representative information?
      1. How would additional levels of data availability be added onto tract-level data, for example, county, MSA-level data?
      2. Would there be a procedure for data augmentation whereby local observatories could augment the data by adding tracts to build up to a higher level. What is the nesting of data resources? To add new tracts, a sampling statistician would make several replicates of the sample and draw upon the additional replicates as needed – could replicate all tracts, ultimately.
      3. Could simply collect all the data in a local area - for example, the Chicago MSA.
      4. The issue is what would be the procedure for regional representation?
      5. How would cross-observatory studies be conducted?
      6. How would we accommodate data collection for places?
   2. What is the plan for observatory-specific studies? How easy would it be to conduct cross-observatory studies?
   3. How would the observatory system structure regional studies?
2. Whether to use place or population as the basis of the design
   1. This needs more discussion. These seem to be at two ends of a continuum. It seemed that the consensus fell more on the end of a population basis for a design. This would move away from an approach dividing the U.S. into regions and attempting to establish an observatory in each region.
   2. We gave up on using ecological criteria as the basis for the design. However, we could still use regional population distribution or other divisions for the design. For example, it still may make sense to think in terms of representing the 9 Census Divisions – for historical reasons. We could overlay such a regional framework on the population-based site selection.
   3. A map of population density of the U.S. is attached to show where our sample sites would lie if they are chosen proportional to population.
3. Whether priority lies with the national core sample or local control and data collection. Would the design maximize the national sample over local priorities? What fraction of resources would go to the national vs. local data structure? What would be the flexibility for local data collection?
   1. Key issue – any time an observatory is interested in collecting data, the study design must ensure a standard structure for the data collection
   2. What would the observatory system do to facilitate access to all its data?

Two tables of existing data that are currently available and an indication of the level at which they are available, to the best of our present knowledge, are found in an Appendix to this report. The intensity and cost of compiling such data are minimized if these existing data are available for the level of analysis an observatory needs: Regional observatories can conduct parallel studies using administrative and digital data at some of these levels. Some data may be available through Census Research data centers or through special confidential data contracts, whereas other data may have to be collected through surveys.

REPORT FROM THE MAY 2012 MEETING

The May 15-16, 2012 workshop focused on these issues and made considerable progress. Workshop members prepared a Vision Statement during the meeting that begins to make a case for why we urgently need to implement a network of observatories to advance SBE sciences in the service of society and to advance scientific theories and methods. This and other materials are posted on the website: http://www.popcenter.umd.edu/research/sponsored-events/nsf-workshops

*Doing Research in a Whole New Way*

The observatories seek to collect, organize, create, and disseminate data to allow research to be done in entirely new ways: observatories would be integrative of data at different levels of aggregation, use data sources that until now have not been part of the toolbox of the social sciences, and connect all these data to local context and place without losing a capacity to aggregate and serve as a national sample of people and places in the USA. In short, this national infrastructure for the country will serve as a platform for social and natural sciences to address social and environmental interactions in a changing world. It will allow researchers to answer a broad range of questions about people and places that are important both locally and nationally.

*Proposed National Design Framework*

At the meeting the participants recommended, after lengthy discussion, that the framework be a national sample of 400 census tracts (each tract of about 5,000 people), thus potentially including some 2 million people. This will provide as robust a sample as any that exists at present-- with the added value that the tracts become locations in which to examine local processes such as how persons’ and families’ proximal social contexts directly affect social, behavioral, and economic functioning and how other influences moderate these dimensions. Exogenous shocks taking place in some tracts but not all tracts will provide opportunities to conduct comparative studies of adaptation and change. By being a nationally representative sample it would also allow comparisons across the nation with regards to how local and national processes affect opportunity and mobility. In a set of the 400 census tracts, randomized trials could be conducted to understand why certain health, education, and economic interventions work in some places but not others.

Since the social observatories focus on 400 Census tracts with 5,000 people in each one, they will comprise a representative sample of 2 million people in the US from which subsamples can be drawn.   Moreover, since the goal is to obtain a wide variety of socio-demographic characteristics about these people as well as their names, telephone numbers, residential addresses, and e-mail addresses, it should be possible to identify and contact important subsets of people using multiple survey modalities This sample will make it possible to do special studies that were prohibitively expensive in the past. Moreover, since the social observatories will collect place and contextual data as well, it will be possible to understand not just individual variation, but also the impacts of spatial and contextual variation. The 400 census tracts with its 2 million residents will not serve to answer all questions scientists will have but it provides a national platform for many questions across the sciences and for policy-makers that is both flexible and comprehensive.

The advantage of having an overall national framework is enormous. First, it allows generalization across multiple contexts. The national framework permits comparison of variables and questions across multiple locations. It permits improving conceptual models that are not specific to place but can take into account such variability. The framework provides a rapid response capability because at least some of the sites are likely to experience emergencies or crises. Because of the time dimension, the design will have an improved ability to disentangle causality.

There is also a major advantage to having the capacity for in-depth studies. It provides granularity of observation: an opportunity for focus on process, for examining micro-phenomena such as speech, behavior and community change over time. For detailed studies on more local questions, it provides a foundation and framework for augmenting the sample by adding more tracts in the area of interest (e.g. a city, a set of neighborhoods). Each observatory would propose to augment the sample to study unique research questions of local interest. This could mean additional sampled locations, increasing the intensity and nature of the data collection. Thus, this framework will allow for both breadth (national sample) and depth (greater granularity and augmented sampling for specific purposes). The observatory network can facilitate depth across locations grounded in the national sample. We are also designing this network to ensure consistency of collaboration across locations.

*Substantive Foci of the Observatories*

The major substantive foci of the observatories would be initially on questions of a) change and adaptation and b) opportunity and mobility. Two major concerns about America’s future are the adequacy of its infrastructure and the robustness of its economic structure. Many people worry that our transportation, environmental, water, sewer, educational and even governmental infrastructure are outmoded and need revitalization. They also worry that structural changes in the US economy are making it less competitive in a global world.  Because the social observatories would focus on place and context as well as people, they would provide extraordinary new information about the role of infrastructure and economic structure. They would provide evidence about how aging infrastructure and the changing structure of the US economy facilitate or impede individual opportunities and accomplishments.  They would make it possible to determine which kinds of investments in infrastructure have the most impact on well-being and how economic change can be fashioned to provide the greatest opportunities with the fewest barriers to mobility. They would provide information on how we can design "smart cities" that facilitate travel, communication, finding jobs, taking care of families, and business investment.  They would also provide information on the organizational structures of communities that may facilitate or hinder appropriate adaptation to ongoing economic, social, and environmental change. In the last decade, social scientists have increasingly recognized the importance of place and the interaction of people and place; the social observatories would provide the needed data.

*An Example of how Social Observatories might study Neighborhoods*

The social observatory will provide an unparalleled opportunity to understand neighborhoods and neighborhood change, a worthy topic in its own right but also critical to those who wish to study individuals and families in relation to neighborhood context. Current practice treats census tracts as operational definitions of neighborhood contexts and measures those aspects of neighborhood easily captured in census data (poverty or concentrated disadvantage, race/ethnic/immigrant composition, residential stability). These dimensions of neighborhood context are highly important but so are other dimensions such as social disorder, social control, cohesion, social support, and collective efficacy. Characteristics of the built environment, accessibility to health and social services, and exposure to toxins and hazards are also important. Tracts may be a relevant level at which to study Chicago’s neighborhoods in relation to attitudes and behavior, but there is evidence that points to the importance of smaller units (e.g., the “street”). Based on a study of census data from tracts and larger units, it appears that neighborhoods do not generally change very fast. Decennial census data may be adequate to describe such change. More frequent data collection may be important to describe change in smaller neighborhood units. To further complicate matters, the relevant unit varies over the life course; adolescents have a broader community to access than younger children, for example. To understand neighborhood contexts, change, and consequences, then, it is important to consider multidimensionality and multiple interrelated spatial and temporal scales.

To date, the very best neighborhood studies in sociology feature specific cases, e.g., Chicago or Los Angeles. With a social observatory network, we can contemplate the possibility of a national sample of neighborhood contexts that can be studied at multiple levels and in multiple ways. Linked with health outcomes, it will be possible to consider the effects of, for example, poverty while taking into account chemical exposures; it will also be possible to consider the effects of chemical exposures while controlling for poverty and other social characteristics of local contexts. There are virtually no studies that do this.

In order to study neighborhoods we need spatially referenced administrative data, cellphone data on the movements of individuals through their day, and survey responses if those are available. Systematic social observation could also be added to this mix because this approach captures cues in the social and physical environment that can be very important to behavioral outcomes **that are not available from other sources**. Systematic social observation is an approach that combines an observational method with a checklist and measurement at a very granular level. The observational method can be videotape (as Rob Sampson did early on in Chicago—although this is very expensive), data collectors walking or driving slowly through an area (more typically done), or Google Street View or other sensor. The unit of analysis is a block face or street segment. This approach can be adapted to fit all settings, not just cities. Although, initially, analysts simply aggregated observations up to a predefined spatial unit, more recent analysts have taken advantage of the spatial variability inherent in such a data set to describe surfaces of contextual characteristics (such as disorder) that can be used to measure contexts however they are defined and bounded.

*Potential Methodological Advances*

Because the social observatories will use new methods to collect substantial socio-demographic information about a large fraction of the people (perhaps all of them) in a large number of Census tracts, they will provide an opportunity for researchers to better understand specific populations of interest. The social observatories can serve as a methodological test-bed to determine whether such new methods of data collection would produce data that are comparable, valid, and reliable compared to information from the Decennial Census, the Current Population Survey, or the American Communities Survey. Since the social observatories will collect data from Census tracts, it will be possible to compare the information gleaned from administrative data, Internet data, or other sources, with those from Census products. As a result, the social observatories can help the Census innovate in its data collection methods. The Census is a constitutional mandate aimed at improving our understanding of the lives of the American people , by advancing the ways to collect this data in less invasive ways, the observatories will strengthen the value of the Census and better serve the needs of the nation.

*Types of Data that the Observatories would Collect and Archive*

The social observatories will not only use the 400-census-tracts framework but will be expected to mine administrative data, both county and state, and to link these data to other data that they will collect such as social media data and web-scraping data. Social observatories will need to have a core of socio-demographic data on respondents. These data include the following information:

* Sex
* Age
* Race/Ethnicity
* Marital status
* Employment status
* Income
* Education
* Length of residence
* Type of housing tenure (own or rent)
* Party registration
* Household composition (other members of household)

*Survey versus Administrative Data*

One way to get these kind of data is to undertake a survey, but a survey of all members of 400 tracts would involve approximately 5000 X 400 = 2,000,000 interviews – a prohibitive number. A sampling procedure could be used that would get a small number of people in each tract, but this would sacrifice precision and it would suffer from non-response. It would get us back to simply doing surveys.

An alternative would be to get “complete” data on these people from administrative sources. This would undoubtedly require working back and forth from geo-locations, addresses, names, and other identifiers. E.g. From standard Census Tiger files you could get all streets and addresses within a census tract. Of course, some of these addresses will be commercial buildings, factories, apartment buildings, etc. They are far from a complete enumeration of all the people in the area—they just index physical locations (and they may do this incompletely, especially in rural areas. But you could then go to sources like the following to identify people within the Census tracts:

* Voting records which typically have a relatively reliable address and a person’s name although the records are sometimes not up to date
* United States Post Office Address Files which give all valid addresses (but not people)
* Information from some of the locator services that basically rely upon credit card and other information.
* Perhaps Motor Vehicle files.
* Reverse telephone directories which list people, phone numbers, and addresses.
* In cities, perhaps files from zoning agencies of buildings and their occupants. Perhaps also property tax rolls might be available.

The trick would be to work back and forth between addresses and names on files until there was a pretty good list that associated people with addresses. Services on the web (e.g, PeopleFinder) now have a lot of this information as indicated by when you try to look up someone by name and find that there are 48 Mary Smiths in a given city and it goes on to list their ages, addresses, etc. The process of adding names and addresses would probably continue as other data sets were added that had either names or addresses.

Then the following kinds of administrative data might usefully be added:

* Vital statistics – This would give recent births, parents (or just mother), and usually some address information. It would also give recent deaths. It might also give marriages. This could be used to augment and clean-up the files, but one very big problem is that vital statistics are state by state and someone might have been born or married elsewhere. This will make it hard to get complete coverage. These files also sometimes have data on race or ethnicity.
* Base wage file – This contains names, social security number (hard to get), employers, and quarterly wages. It would have to be linked to people using names and other information – this matching would be imperfect, but it would give wage data for people in the tract. This data could be compared with aggregate Census data for the tract that would be available in various categories.
* Statewide Medicaid/Welfare/Food Stamps eligibility – These are not always the same file. In some states the most basic welfare data are only available at the county level; in other states there are statewide files, but they are sometimes unreliable or only a subset of the county data. This would be very useful for figuring out something about income and perhaps ethnicity since these files sometimes have that data as well. And, of course, they would provide a wealth of data about social program participation.
* Motor Vehicle Files – These typically have names, addresses, and physical characteristics that would be very useful. They also have pictures from which some racial coding could be done.
* Political Contributions data – This just identifies a small subset of the population but it provides names and addresses and some very interesting financial data that could be used to infer lower bounds on income. (Poor people do not give $5,000 in contributions.)
* Credit Card Data – These would be a goldmine, but they might be very hard to get. Presumably they would have names and addresses, probably age, and lots of financial data that would help establish incomes.
* Tax records – These are almost impossible to get but some scholars have gotten access to these data for statistical analysis, but not the micro-data.

It might be possible to get a fairly reasonable enumeration of everyone in a tract—especially those over 18 and it might be possible to learn a lot about their demographics. Perhaps the most elusive demographic is education. Very few datasets record it, and when they do, they do not do so very accurately. For those under 18, one way to get it would be to have school records but there are lots of problems here: School records are notoriously hard to get. Not everyone goes to public schools so that private school data must be obtained as well. And, of course, even if the under 18 data can be obtained, it is very hard to get the over 18 data.

One of the first things that the social observatories might try to do would be to get an enumeration of people within their tracts and to get as much socio-demographic information as possible. Then a validation procedure might take place where surveys would be done to cross-check the data developed in this way.

In short, it would be possible to get basic socio-demographic data from these other, non-census sources. Indeed, the prospect could be of interest to the Census Bureau and may lead to get them to match the social observatories files with Census data through a Census Research Data Center. The next steps would be to attach e-mail addresses to people and to see to what extent social media and internet data could be attached to them.

Needed are baseline environmental data on the places where people live, work and play. Everything should be geographically tagged or referenced geospatially —houses, neighborhoods, townships. We need weather data that can then be linked to economic outcomes, mobility patterns, and physical structures’ condition. We need to cross physical sciences’ information with social sciences information and personal information to address issues of the energy system that sustains us, how information is used, and what kind of transportation systems we need to better serve people.

*Preferences*

An important aspect of individual variation — individuals’ underlying preferences, such as risk tolerance, patience, altruism, trust, and cooperation — is not captured by survey-based demographic measures. These underlying preferences are important determinants of individual and group decision-making and of the economic success of individuals, households and neighborhoods. Individuals who are more risk tolerant and patient are more likely to engage in the kinds of behaviors, such as investment in human capital, that lead to economic success. Groups with higher levels of trust and cooperation are more successful at organizing to provide public goods. These preferences are not measured – or measured inadequately – in current social surveys. Considerable research has shown the superiority of behavioral measures of preferences. In these behavioral measures, subjects engage in decision-making that has real consequences. The decisions they make determine monetary payoffs in the games. While these protocols were developed for use in lab experiments, they have begun to disperse into the field. The studies to date show that individuals vary in their preferences, and that these variations are consequential in terms of economic well-being. In addition, preferences appear to vary systematically at the household or neighborhood level, and there is some evidence that they are ‘contagious’ within social networks. Social observatories would allow deep data collection and thereby provide more complete explanations for the economic success and failure of individuals in our society.

*Social Media Data Sources and Approaches*

Large amounts of data are currently available from sources such as Google, Twitter, and social media websites. The location of twitter accounts is known and twitter messages can be organized geographically as well as classified according to topic. Organizing such messages by geographic area may provide useful information on the concerns of local residents. This is also a way to identify networks and to locate emergencies in time and space.

*Smartphone Enabled Data Collection*

In addition, we will pursue another novel data-collection (and attrition-reduction) strategy. We propose to collect extensive data on adolescents’ use of space and time using the capabilities of smartphones distributed to the study subjects. The smartphones will allow the collection of multiple types of data. Real-time GPS data will allow us to construct geographical descriptions of how the neighborhood environment is used by adolescents, including the full extent of geographical space covered in a typical day or week, the proportion of time spent at home, the school or other supervised settings, or in public, unsupervised spaces. In addition, interactive text messaged questions will provide dynamic time use data that can supplement and validate the locational data from the GPS. Below we describe some of the measures we will glean from this technology:

Geographic extent: The geographic space in which respondents travel over the course of a week or a year. Geographic extent will be constructed by identifying locations in which subjects spend at least five percent, for example, of their daytime hours over the week (~5 or more). These criteria are used to avoid bias from isolated trips outside of the subject’s typical environment. Using GPS data supplied over a sampled period and matching it with spatial data identifying key locations (the respondent’s home, school/workplace, local parks and “third” spaces) we will construct aggregate measures of time spent in each of these key locations in an subject’s life over the course of the year.

Social Network: In order to assess the heterogeneity of social networks, we will upload their “address book” feature from their smartphones as well as the call and text logs and then use reverse lookup directories in order to determine—to the extent possible—the address of the numbers called (and received). We then will use these locational data as proxy for SES by comparing with means for census block groups. (We will ask respondents to share geographic information about those we are unable to find online.) Finally, we will assess duration of ties.

Self-reported time use: Using regularly scheduled text messages over the sampled period, we will ask the adolescent to provide a description of where he/she is at the given moment, whom he/she is with, and what activity he/she is taking part in. These messages will only be sent in non-school hours, and will be used on their own and also to verify the accuracy of the data from the GPS readings. In the Unites States, differences in structured versus unstructured activity have been linked to differential educational outcomes for low- and high-SES children (see Lareau 2006; Lareau, Weingartner and Conley 2009). For example, children in the top income quintile spend about 50 percent more of their non-school time in formalized activities. Whether these differences mediate differential outcomes is a matter of some debate; for instance, the theory of “summer setback”—in which SES and race differences emerge during the summer months when the activities diverge the most by class—has recently gained currency in the educational literature (see, e.g., Heyns 2002). The extent to which such SES-based activity differences are actually a product of local environments and peer influences remains to be seen.

Physical Movement: In combination with GPS location, smartphone technology allows for the assessment of physical movements through built-in accelerometers.

### *Sensors*

### A variety of new tools are available to passively obtain information about neighborhoods.

### Air quality: A series of new semi-conductor sensors have been developed and calibrated for distributed usage such as proposed here. One particular sensor measures the following airborne toxicants in the specified ranges: Carbon monoxide (10-1000ppm); Hydrocarbons/Volatile Organic Compounds (50-5000ppb) and Nitrogen Dioxide (50-5000ppb). This is the MICS-4514 sensor produced by Ev2 Gas Sensors, a Swiss-based semi-conductor and gas sensor company. These sensitivity levels work for our purposes since recommended safety threshold levels fall within their ranges. For example, the National Institute for Occupational Safety and Health (NIOSH) has established a recommended exposure limit (REL) for nitrous oxide of 25 parts per million (ppm) parts of air as a time-weighted average (TWA) for the duration of the exposure assuming exposure for a forty hour work week (NIOSH 1992). Likewise, homes with poorly installed gas stoves, for example, average carbon monoxide levels of 30 ppm (though no official standards have been set by the government [EPA 2009]). Nor has the EPA established guidelines for volatile organic compounds; however, HUD has set an upper limit of formaldehyde (one of the more common VOCs) levels of .4 ppm while the EPA “suggest” mitigating exposure to more than .1 ppm (EPA 2009a).

### Noise Levels: Noise can be assessed in two ways. First, levels can be recorded for two seconds each hour and transmitted to the data server. This provides a mean level of noise (linked to GPS location and time of day) both in daytime and night hours. Second, anytime the ambient decibel level exceeds 85 decibels (the level at which damage to hearing begins) during day light hours, the duration and level of the noise can be recorded automatically if the phone is on (Goines and Hagler 2007). Likewise, the threshold during night time hours can be set at 45 decibels, which is the level at which the average person cannot sleep (ibid).

SUMMARY AND CONCLUSION

The workshop participants felt that a network of social observatories must be capable of fully representing the people and the places where people live, both in the aggregate and in fine detail, in order to address local and regional differences that characterize this diverse nation. As such, it will provide a representative picture of the nation’s population as well as a placed-based sample capable of addressing the differences that characterize our population and the diverse settlements where they live and work. Unlike most other existing research platforms, this place-based capability will ensure that we understand not only the high-density urban places where many of the population lives, but also the important medium and low density exurban and rural places that represent a vast majority of the land area in the nation. To do so, the observatories will follow approximately 400 census tracts sampled to represent the U.S. population while also fully capturing the diversity that characterizes local places. These census tracts will be systematically studied over time and space from some 20 observatories spread across the country. Unlike the Census, taken every ten years, the observatories will provide a near continuous snapshot of the nation and address the dynamics of change in our society, quickly be able to see how national policies affect local places, an all important task because what happens at the local and state level matters in a federal system such as that of the US.

These observatories will transform how the SBE sciences go about their work by encouraging the integration of the SBE sciences, rather than reinforcing the fragmentation that we have experienced since the 1960s. The last four decades were a necessary phase to achieve greater specialization but have had over the years the effect of making it ever more difficult for SBE scientists to share methods and approaches to address issues of national importance. The observatory network will explicitly promote what is now a broad call to address issues of importance with the best tools available without regard for disciplinary origins. Having places across the nation with the explicit charge of ensuring that teams of scientists are working together around questions of national interest will serve to integrate the sciences, and also serve the nation better by providing diagnostic and policy-relevant solutions at a variety of scales from local to state to national and international issues. There are issues of privacy to be addressed in this sort of geocoded world, and the observatories will be a place where these concerns can be addressed systematically and lead to the creation of standards for ensuring privacy of sensitive information.





