

Institute of Education Sciences
National Center for Education Statistics

NATIONAL INSTITUTE OF STATISTICAL SCIENCES
TECHNICAL EXPERT PANEL REPORT

TAKING A LONGITUDINAL VIEW OF ADMINISTRATIVE EDUCATION DATA

TABLE OF CONTENTS

EXECUTIVE SUMMARY.....	3
PREFACE	5
I. BACKGROUND	6
II. CONTEXT	7
III. REPORTING LONGITUDINAL DATA	7
REPORTING REQUIREMENTS AND CONSTRAINTS FOR FEDERAL AGENCIES.....	7
NCES ADMINISTRATIVE DATABASES: OPPORTUNITIES.....	8
IV. VISION.....	8
OBJECTIVES.....	8
INFORMATIONAL REPORTS.....	8
RESEARCH REPORTS.....	9
V. DATA ASSETS.....	9
MANAGEMENT	9
VI. RESPONSES TO SPECIFIC QUESTIONS.....	10
OMB REPORTING REQUIREMENTS.....	10
AGGREGATE REPORTING	10
METHODOLOGY FOR LONGITUDINAL SUMMARIZATION.....	10
DIVERSE DATA STRUCTURES AND CONTENT	11
VII. VII. SUMMARY OF FINDINGS.....	11
PREMISE	11
PRIMARY RECOMMENDATIONS.....	11
SPECIFIC RECOMMENDATIONS – MANAGING RESEARCH ASSETS	11
SPECIFIC RECOMMENDATIONS – MANAGING DATA ASSETS.....	12
SPECIFIC RECOMMENDATIONS – CREATING INFRASTRUCTURE	12
VIII. APPENDIX.....	14
IX. CHARGE TO PANEL.....	15
X. AGENDA	16
XI. EXPERT PANEL MEMBERS’ BIOSKETCHES	17

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

NCES/NISS Technical Expert Panel on Administrative Data and Analysis of Panel Data

EXECUTIVE SUMMARY

The National Center for Education Statistics (NCES) conducted their first collection of administrative data on postsecondary education in 1966, almost immediately upon establishment of the Center, with continuing collections through today. The Integrated Postsecondary Education Data System (IPEDS) was initially fielded as the Higher Education General Information Survey (HEGIS). Now this data repository contains over five decades of consistently gathered administrative data on postsecondary education. Data in IPEDS effectively documents the changes over time to educational policies, practices and outcomes. These data include detailed information on institutional characteristics, student population information, financial aid information and information on outcomes.

While IPEDS data is rich and complex due both to longevity of the data and to the diversity of the population of institutions included, it also is difficult for users to navigate. Currently the value of these data is not sufficiently realized. As changes in education occur over years, the impact of some changes may only become evident in subsequent years. NCES administrative data collections comprise a unique and highly valuable resource for evaluating change and the impacts of interventions.

The NCES charged a panel of technical experts to focus on the potential of publishing longitudinal summaries of these administrative data to make this resource practically and effectively useful to decision-makers and others. In addition, the panel was asked to address two specific issues for NCES reports:

1. Issues involved in objective, policy-neutral representation of longitudinal trends present in the extensive longitudinal database of administrative information, and
2. Issues of appropriate summarization of data gathered as a census rather than as a probability sample.

The panel met at NCES on May 22-23, 2018, supplemented by teleconferences.

Primary Recommendation

Develop a vision for longitudinal analysis and reporting to meet the needs of decision-makers, policy-makers, researchers and other NCES data users.

1. **Manage Research Assets:** Immediately proceed to develop a series of publications of trends and analyses of change. Publications should be of two kinds: informational publications in a standard format for regular (annual) production and more detailed research reports on a few selected topics each year. Topics should augment the in-depth reports currently commissioned by NCES and should emphasize research that is uniquely possible via NCES resources.

Scientific integrity should be the hallmark of NCES of both kinds of publications.

2. **Manage Data Assets:** Undertake the structuring of the NCES data collections in a cloud-based structure that will facilitate multiple uses and data linkages within NCES, across the Department of Education, and with other agencies. Create a Vision for 2020 (and beyond) for the use of NCES data

and needs for access and analysis. Start now to curate NCES data to facilitate longitudinal analysis for external as well as internal use.

Details and specific recommendations for implementation are included in the final section of the full report.

Responses to Specific Questions

Objective, policy-neutral representation of longitudinal trends first requires determining which information is needed at the full population level and which is needed at the subgroup level. Note that aggregate reporting for the population may not be meaningful for subgroups.

Summarization of data gathered as a census rather than as a probability sample does not pose a technical difficulty in reporting although probability-based standard errors or other uncertainty statements are not possible. Summarization should use sufficiently sophisticated analytic methodology to match the complexity of data and be flexible to changes.

NATIONAL INSTITUTE OF STATISTICAL SCIENCES TECHNICAL EXPERT PANEL REPORT

PREFACE

The National Center for Education Statistics charged the National Institute of Statistical Sciences (NISS) with convening a panel of technical experts to focus on the potential of publishing longitudinal summaries of administrative data collected by NCES. The panel was asked to address the issues involved in objective, policy-neutral representation of longitudinal trends present in the extensive longitudinal database of administrative information and the issues in appropriate summarization of data gathered as a census rather than as a probability sample.

On 22-23 May 2018 the panel of technical experts met in person with presentations and discussions with NCES staff followed by closed sessions for deliberation. The panel held further closed teleconferences to prepare this report.

NATIONAL INSTITUTE OF STATISTICAL SCIENCES ANALYSIS OF PANEL DATA

ADMINISTRATIVE DATA AND PANEL DATA ANALYSIS

I. BACKGROUND

Almost immediately upon the establishment of the National Center for Education Statistics in 1965 the first collection of administrative data on postsecondary education began. In 1966, pursuant to the Higher Education Act of 1965 (and subsequently as amended), NCES was charged with collecting information from institutions that participated in any Title IV student financial aid program under the auspices of the Federal Student Aid (FSA) office. The higher Education General Information Survey (HEGIS) was initiated in 1966; and participation was and still is mandated by law for all institutions that participate in FSA programs.

With revisions since 1966 and reconstitution as the Integrated Postsecondary Education Data System in the 1980s, annual collection of these data has continued without interruption. The data include detailed information on the institutional characteristics (programs offered, acceptance rates as well as fees and costs and financial and other resources), student population information, financial aid information and information on outcomes (retention, graduation, transfer, degrees/certificates).

Thus, the NCES data repository contains over five decades of consistently gathered administrative data on postsecondary education. This wealth of information chronicles and documents the many dimensions of change in education policies, practices and outcomes. The concern prompting creation of this expert panel is the desire to make both longitudinal and current views of these data practically and effectively useful to decision-makers and others.

NCES also collects other administrative data at the school district and state levels (notably the “F-33” or Survey of Local Governments: School Systems and the National Public Education Financial Survey, respectively). These universe data collections are primarily financial; F-33 data include state and federal revenues broken down by school district but limit expenditure data to salaries and benefits. NPEFS includes other expenditures (e.g., purchased services, supplies) but does not break state revenues down by district; and there are other differences as well. Decision-makers need to be able to view these data as well from a longer-term perspective than just current and preceding year.

The panel elected to focus its attention on the IPEDS data collection while acknowledging that many if not most of the recommendations in regard to these administrative data would be equally applicable to the school district and state-level finance data. In 2000 IPEDS became web-based; and in 2014 together with the rest of the collective NCES published data, IPEDS data moved to the cloud.

II. CONTEXT

IPEDS is a universe data collection and therefore is not subject to any randomness inherent in sampling. While student record-level data are not in the public domain (but may be available via restricted license), most institutional data are accessible to the public. The richness and national scale of these data make IPEDS unique as a data resource for understanding the state of post-secondary education in the US.

Consumers of IPEDS data are diverse and their purposes for consulting these data are equally wide ranging (e.g., policy-making / evaluation of long-term effects of changes / personal choice of postsecondary institution / research studies / inter-institutional comparisons). Data are used in a variety of ways (e.g., current IPEDS data / IPEDS data integrated with other federal statistical data / IPEDS data combined with state, local or institutional record data).

The richness and the complexity of the IPEDS data collection are due both to longevity of the data and to the diversity of the population of institutions defined according to Title IV eligibility and therefore included. From training programs to research universities, the breadth in terms of institutional mission, the education objectives and the clientele served is wide. But even within defined subgroups of the 7200 participating institutions the diversity is great. For example, in Pennsylvania there is a high degree of variation in the size of institutions. Universities that grant Bachelor's and postgraduate/professional degrees include the public Pennsylvania State University with 24 campuses and 100,000 students and the non-profit (Catholic and Jesuit) University of Scranton with fewer than 5400 students.

The complexity of the postsecondary industry is reflected in IPEDS and magnified because of changes over the five decades of collection: item inclusions/exclusions, data definitions, changing population of institutions. Even within any single year, data are complex because of differences among institutions: the data that are relevant to each; the record-keeping and accounting systems; data requirements/ definitions mandated by states.

This complexity presents a challenge in using IPEDS data from multiple years to examine change. Not all data users have the capability and/or available resources for the task, starting with assembling the information by extracting from different years' data sets and then identifying and adjusting for changes in the population and changes in definitions before proceeding to analysis.

Currently, published reporting via the NCES/IES website at present includes announcements of data availability with brief summaries of key information (*First Looks*), blog entries that present vignettes of interesting facets of the data, and a few longer articles each year that treat a particular aspect in some depth. All *First Looks* and most other publications are authored by IPEDS contractors including longer pieces that are commissioned from contractors on topics specified by NCES.

III. REPORTING LONGITUDINAL DATA

REPORTING REQUIREMENTS AND CONSTRAINTS FOR FEDERAL AGENCIES

All NCES reporting as well as data collection activities, adheres to the standards for all federal statistical agencies issued through the Office for Management and Budget (OMB) Directives. As NCES now is considering expansion of the scope of NCES reports to add informational and research reports on changes over time, NCES asked the panel to address two issues: i) proper reporting IPEDS data summaries of medium- and long-term changes that would meet the standards' requirement of objectivity and ii) reporting of statistics with associated uncertainties where appropriate.

The publication of NCES data over the past several years has been limited by imposition of a highly restrictive view of what constitutes "objectivity." For example, NCES does publish simple descriptive

statistics with their sampling variances as well as some trend lines for single variables or established indices. However statistical modeling has deliberately been left to the external research community on the general premise that selection of variables for inclusion/exclusion in a model violates the highest kind of objectivity.

The panel found this notion of “objectivity” to be excessively stringent and markedly different from the definitions at other federal statistical agencies. Agencies including the Bureau of Labor Statistics (BLS) and the Bureau of the Census, as well as others, regularly publish model-based statistics and statistical models for longitudinal data.

This in no way inhibits or obstructs the research community. Consistent, timely, regular and systematic production of standard analyses is neither the responsibility nor the interest of the research community.

NCES ADMINISTRATIVE DATABASES: OPPORTUNITIES

For decision-makers at many levels, from school district to state to Congress, longitudinal data summarization and analyses of trends and patterns are needed to provide the context for decisions and to be the basis for evaluation of the impacts of previous interventions and changes. Further, analyses are often more challenging because the impact of changes in policies or practices often is not instantaneous, but only become evident in subsequent years. For example, changes in patterns of program selections by entering students only become evident in graduation rates several years later. Single year data (even if evident for isolated events in paired current and preceding year data) cannot provide this information.

As owner of the data and with the expertise available to a federal statistical agency, NCES is *per force* the natural author of publications of longitudinal summaries, trends and models that can provide timely, valuable information to decision-makers. Many if not most decision-makers lack the time, resources and/or expertise for these longitudinal analyses. However, NCES can bring the technical expertise for appropriate analyses that place observed changes into a multi-year context.

IV. VISION

OBJECTIVES

Defining a set of objectives for reporting change is the first step in developing a cohesive set of longitudinal reports. Development and prioritization of these will require the input of user communities in order to meet their needs; but some basic longitudinal summaries can be anticipated as part of a core set.

INFORMATIONAL REPORTS

Decision-makers within the educational sphere need to be able to access information on change in both contexts, short-term (i.e., year-to-year) and long-term (i.e., multi-year trend). For a core set of variables drawn from the various types of data (financial, student, program, institution, etc.), these analyses can be predicted to have regular and wide-spread use. Once this core set has been defined and prioritized by the anticipated needs of various user communities, templates for these informational reports developed so that the reports can be institutionalized for issuing annually. For example, changes in an expenditure pattern from one year to the next may be viewed in the context of longer-term trends. Similarly, changes in the program registrations or success rates can be viewed over the longer term and/or presented in relation to the proportion of student loan funds awarded as both change over time.

The technical and statistical methodology may be challenging to develop initially but should be able to be used in subsequent years with only minor adaptation for newly introduced changes. This underscores the importance of NCES undertaking the task as many data users would not have access to the same level of expertise.

RESEARCH REPORTS

NCES is uniquely positioned to identify marked changes in education patterns, especially when these changes become apparent upon linking data either among NCES databases or with data from other statistical agencies. Understanding the patterns of change and of the time sequence of impacts observed consequent to changes in policy or practice is important to decision-makers and policy-makers in the education community. Changes may be broad trends over time (even measured in decades), or may result from advances in particular disciplines or skill sets (e.g., creation of new programs prompted by technological change), or may be policy changes that affect either students or operational issues for education institutions (e.g., student loan regulations).

NCES staff who curate the IPEDS data have knowledge of the data as these are collected and have the context for observing change. Also, requests to NCES from the public, from the institutions represented in the database, and from government (state and federal) provide other sources of topics that warrant deeper research than the informational reports.

Prioritization of research objectives for examining trends and other links between changes and their impacts is the responsibility of NCES. It would be good to start with a limited number of research projects specifically for topics in longitudinal analysis to be undertaken with research reports issued each year. By taking full advantage of the historical IPEDS data, possibly integrating other contemporaneous data, NCES staff (sometimes with collaborators) can provide a clear analytic report of the context and an in-depth longitudinal review. As researchers, NCES staff can also bring to these projects an important knowledge base with regard to alterations in the databases (definitions, population composition, etc.).

V. DATA ASSETS

MANAGEMENT

The complexity of the IPEDS database has implications for how the data need to be accessible and for how the data need to be summarized and/or analyzed. The complexity of the data, across time and across a highly heterogeneous population of institutions, presents challenges that will only expand. NCES is to be applauded for moving the NCES databases to the cloud. By moving to the cloud access to these databases has been simplified. Because this move has essentially posted the existing data structures to a new location, the process of extracting the specific data required for an analysis is fundamentally unchanged. These data structures are not well equipped to meet the need to extract data both for reporting and for linking with other federal data, while remaining cognizant of the alternative, time-limited definitions. The current file structure places obstacles to creating of a multi-year working file because data files are separate for each year so that relevant data for analysis must be extracted separately from each year's file and collated to assemble the composite working file.

Retaining this server-based data structure on the cloud fails to take advantage of valuable flexibility in data access and in data linkage and sharing and some efficiencies in incorporating new data. Summarizing and analyzing change comes with the need to extract from multiple years of the IPEDS data and often with opportunities to integrate data from other sources. A cloud-based data structure could facilitate these goals by providing the data user with simultaneous access to (extraction of) data from multiple years while taking into account changes in relevant definitions or in population base and accommodating links to other data sources (technology for accomplishing this exists and has been implemented at another large federal statistics agency).

VI. RESPONSES TO SPECIFIC QUESTIONS

OMB REPORTING REQUIREMENTS

OMB requirements do not appear to present obstacles to analyzing and reporting change. Neither is there a technical difficulty introduced in reporting on universe data. Obviously the “total error” for universe data does not include sampling error, hence a probability-based standard error or other uncertainty statement is not possible. With regard to error from other sources, data quality is handled by audit procedures to the extent that it is amenable. Apart from failed logical checks, there is no way to detect the presence or to estimate the extent of faulty data. Thus, there is no uncertainty statement to be made for this universe data.

However, in making statements about models constructed from these administrative data, measures of goodness-of-fit of the model to the data document the degree of approximation when the model is applied.

AGGREGATE REPORTING

While aggregate statistics can be reliable for the population, this summarized global information may not apply to subgroups of the population. This is a conundrum that has no single answer. The great heterogeneity of institutions in the IPEDS population (defined only by eligibility for Title IV funds) guarantees that at least some of the important population summary statistics will be inappropriate when applied to subgroups. And even within subgroups (e.g., post-graduate degree-granting institutions), the diversity is too great for averages to be meaningful at any finer granularity. Therefore, development of aggregate and subgroup reporting calls for rigorous statistical thinking. It also calls for clarity in writing.

The first decision point is to determine which information is needed on the basis of the population as a whole and which information requires reporting on a subgroup basis. When population totals or trends are needed, there is also a need to alert the reader to known or possible deviations from the overall pattern or trend when subgroups are examined, especially those known to be of primary interest. Two aspects to be considered in reporting for subgroups are the importance of the subgroups relative to the anticipated use, and the magnitude of differences of the findings for the subgroups compared to the whole population and/or among the subgroups themselves. In any case, the reporting needs to clarify the population base for each summary statistic or longitudinal model. It also needs to remind the reader that subgroup statistics – and even more so, subgroup trends or other observations over time – may show very different patterns.

METHODOLOGY FOR LONGITUDINAL SUMMARIZATION

For the set of informational reports reporting longitudinal patterns, statistical methodology can be planned in concert with the development of the objectives and templates can be constructed. Specific methods should not be limited to simple descriptive summaries. For long-term data trends will be more complex and will have a variety of features such as jumps, breaks, outliers, etc. Longitudinal data patterns may also comprise superposed trends of different periodicities that a simple model is incapable of representing even for univariate data. For interdependent (multivariate) data, either synchronous or asynchronous with delays, the methodology must be of sufficient sophistication to match this added complexity.

Adequate methodology must also build in adjustments for changes in definitions and standard (annual) adjustments for economic data. Thus choice (possibly development) and implementation of appropriate methodology for analysis templates requires a high level of technical expertise. However, once templates are constructed, in future years analyses should be able to be conducted re-using them with each new annual data collection.

DIVERSE DATA STRUCTURES AND CONTENT

Modern cloud-based data structures allow highly heterogeneous data to be assembled without the rigidity of traditional structures of server-based data systems. As NCES data continues to expand both in scope and size, the demand to link data from multiple sources is increasing rapidly. With all NCES data now in the cloud, a data structure that takes advantage the flexibility possible in cloud-based data structure may facilitate data access and data use.

VII. VII. SUMMARY OF FINDINGS

The previous sections of this report provide details of the panel’s deliberations and the rationale for each issue identified for further consideration by NCES. Three primary findings emerge from the panel’s deliberations. To implement these broad recommendations, more specific findings follow both with regard to the content of reports and with regard to the infrastructure required for producing both informational and research reports.

PREMISE

Much important educational change transpires over years; and evidence of the impact of some changes only emerges after more than a year’s delay. Longitudinal analyses serve to illuminate these kinds of change and need to become a regular part of NCES reporting.

The NCES administrative data collections comprise a unique and highly valuable resource to a wide array of users including decision-makers, researchers, the education community and the general public as well as other federal and non-federal agencies and organizations. Currently the value of these data is not sufficiently realized. NCES as curator for these long-term data has the opportunity to expand the usefulness of this resource and to enhance its quality by publishing reports of longitudinal as well as cross-sectional (annual) summaries with descriptions of short-term and long-term change.

PRIMARY RECOMMENDATIONS

1. Develop a vision for longitudinal analysis and reporting to meet the needs of decision-makers, policy-makers, researchers and other NCES data users
2. Immediately proceed to develop a series of publications of trends and analyses of change. Publications should be of two kinds: informational publications in a standard format for regular (annual) production and more detailed research reports on a few selected topics each year. *Scientific integrity should be the hallmark of NCES of both kinds of publications.*
3. Undertake the structuring (*de novo*) of the NCES data collections – perhaps starting with IPEDS – in a cloud-based structure that will facilitate multiple uses and data linkages within NCES, in the Department of Education, and with other agencies.

SPECIFIC RECOMMENDATIONS – MANAGING RESEARCH ASSETS

Objective: To create and disseminate longitudinal reports that meet a high standard of scientific integrity.

1. Informational Publications of Longitudinal Summaries of Change
 - a) Establish a formal process for listing topics, setting priorities, developing a plan for series
 - b) Start immediately with high priority topics and plan for gradual expansion of the list
 - c) Select topics based on regularity of interest and breadth of clientele for the topic
 - d) Use respected statistical expertise to develop templates (each applicable to several topics)
2. Scientific / Research Publications
 - a) Start immediately to define topics in longitudinal analysis that augment (not replace) in-depth reports on administrative data currently commissioned by NCES

ADMINISTRATIVE DATA AND PANEL DATA ANALYSIS

- b) Select topics for research publications that are uniquely possible via NCES resources and data access using NCES data with or without links to other federal databases
 - c) Select topics that are likely to shed light on background for funding decisions, pending policies or other questions that could be anticipated to come from Congress or the federal administration
3. All Publications
- a) Protect the scientific integrity by setting up a stiff NCES technical review: at Design, Draft and Manuscript stages, with internal review followed by external review by technical experts in the scientific/research community
 - b) Engage respected statistical expertise to ensure adequate depth of knowledge in choice of methodology underpinning the analyses
 - c) Staff adequately with respect to personnel time and with respect to level of expertise; use whatever mechanisms are available (current or new NCES staff, IPA, contractor/consultant)

SPECIFIC RECOMMENDATIONS – MANAGING DATA ASSETS

Objective: To achieve longitudinal resource objectives and make longitudinal data easily accessible and usable across items and across years.

1. Create a Vision for 2020 (and beyond) for use of NCES data and needs for access and analysis to realize the potential of these databases for cross-sectional information and longitudinal perspectives and for linking with other federal databases
2. Move to a modern infrastructure built de novo based on the Vision for 2020 (and beyond)
 - a) Design infrastructure to enable efficient linkage to other data resources both internal to the Department of Education and external to other federal statistical agencies
 - b) Employ a contractor with specialized expertise in cloud-based databases and successful experience with federal statistical agencies
3. Curate NCES (administrative) data – past, present and future – to facilitate longitudinal analysis
 - a) Document curated data for external as well as internal use; provide public users with Recommendations for use
 - b) Start now to curate and document – don't wait
4. With a modern infrastructure in place, continue the NCES practice of developing data descriptors and analysis tools for public use that provide correct analyses and accurate summaries that respect the underlying survey design

SPECIFIC RECOMMENDATIONS – CREATING INFRASTRUCTURE

Objective: To organize resources and to initiate response to the overarching recommendations

1. Build an NCES team to create the Vision for longitudinal reporting
 - a) Set goals for data access (e.g., freely connect across time and items to identify patterns of change)
 - b) Set expectations for depth of analysis and sophistication of methodology
 - c) Set research objectives in terms of contextualizing change
 - d) Continue to act as a stimulus to report consistently, but with sufficient flexibility to reflect the dynamic educational landscape
- 2) Build a NCES team to set a Research agenda
 - a) Define the informational report series
 - b) Set selection criteria and vet topics for research reports

ADMINISTRATIVE DATA AND PANEL DATA ANALYSIS

- 3) Structure a review process that covers all stages from design manuscript submission or dissemination
- 4) Commit the Resources required – use internal plus contract effort and expertise
 - a) Take advantage of what is possible with freer cloud infrastructures
 - b) Contract to technical expert in cloud data structure design
 - c) Take advantage of technical expertise – use external (panel, contractor, consultant,...) to extend internal deep statistical expertise
 - d) Take advantage of external resources in the education community to identify user needs and to set research directions
- 5) Set goals
 - a) Initiate regular production of a core set of informational longitudinal reports
 - b) Take ownership within the federal statistics community of “education issues” - database item specifications, reporting formats and definitions of terms
 - c) Increase NCES visibility as a resource, in particular for longitudinal studies of administrative education data

VIII. APPENDIX

1. CHARGE TO PANEL
2. AGENDA
3. EXPERT PANEL BIOSKETCHES

IX. CHARGE TO PANEL

In addition to surveys and assessments, NCES collects extensive administrative data at all levels of education. IPEDS (Integrated Post-secondary Education Data System) is a very large, continuing data collection that is “your primary source for information U.S. colleges, universities and technical and vocational institutions” according to the NCES website. These data are effectively a census as the Higher Education Act of 1965, Title IV, requires that all institutions that participate in any federal student aid program report their data. Annual reporting began in 1966 and covers eight specific areas: institutional characteristics; institutional prices; admissions; enrollment; student financial aid; degrees and certificates conferred; student persistence and success; and academic libraries, institutional, and human and fiscal resources.

The potential for IPEDS as a resource to policymakers, as an information source for education administrators and professionals and as source material for education researchers is immense. There are extraordinary opportunities to capture a snapshot or a detailed picture of many aspects of post-secondary education and its impact. Also, with over forty years of data, IPEDS data can be queried to examine changes over time and to investigate relationships among the specific areas that are not contemporaneous but systematically lag in impact from the first area to the second. In addition, IPEDS can be integrated with other federal agency data when specific objectives go beyond the eight specific areas listed.

Proper analysis of IPEDS data is not without challenges – and these challenges are the focus for this Expert Panel. First, by OMB requirement, federal statistics are to be quoted with their associated uncertainties. As IPEDS is a census, it is not free of measurement and reporting error but sampling error is not the source of IPEDS data uncertainty. Second, the population of over 7200 post-secondary institutions is extremely diverse, even within broad classification as college, university, technical or vocational institution. If size is measured by students – or even by full-time students – universities vary by an order of magnitude; if size is measured by endowment or other financial measure, then colleges vary by several orders of magnitude. Simple summary statistics become almost meaningless with this diversity. Third, methodology for investigating time-relationships for panel data (universe data) have advanced with increased computing power; so, the limitations on analyses may come from the structure of the data rather than the capability of computation. Fourth, there are differences in institutional structure that affect data reporting in ways that are often relevant to the analysis (i.e. comparing public or private institutions or institutions with different missions and serving different populations). Understanding and respecting the limits of legitimate analysis – as opposed to hypothesis seeking explorations that can also be useful but are different – is both difficult and important.

The collective expertise on the panel brings together the technical perspectives of statistics, econometrics, and sociology with the perspectives of education and data to consider how NCES might address these four challenges.

X. AGENDA

Tuesday, May 22

10th Floor Auditorium

8:30 am	Arrival and Building Security
9:00 am - 11:00 am	Welcome Introductions Selected NCES Staff Presentations of NCES Databases and Requirements for Panel Data Analysis
11:00 am - 12:00 pm	Discussion and Questions from the Panel
12:00 pm - 1:00 pm	Lunch (on your own)
1:00 pm - 4:30 pm	Panel Executive Session
4:30 pm - 5:00 pm	Clarification Requests of NCES from the Panel
5:00 pm	Adjourn

Wednesday, May 23

Room 6008

8:30 am	Arrival and Building Security
9:00 am - 11:00 am	Panel Executive Working Session
11:00 am - 12:00 pm	If useful: NCES Staff Responses to Panel Requests

Wednesday, May 23

Room 5083 (after Lunch)

12:00 pm - 1:00 pm	Panel Executive Discussion and Working Lunch
1:00 pm - 3:30 pm	Panel Executive Working Session
3:30 pm - 4:30 pm	Panel Feedback to NCES
4:30 pm	Adjourn

XI. EXPERT PANEL MEMBERS' BIOSKETCHES

John M. Abowd, Ph.D.

Title: Associate Director for Research and Methodology and Chief Scientist, Census Bureau

Dr. John M. Abowd is Associate Director for Research and Methodology and Chief Scientist at the Census Bureau and the Edmund Ezra Day Professor of Economics, Professor of Statistics and Information Science at Cornell University. He is also Research Associate at the National Bureau of Economic Research (on leave while serving in the federal government), Research Affiliate at the Centre de Recherche en Economie et Statistique (CREST, Paris, France), Research Fellow at the Institute for Labor Economics (IZA, Bonn, Germany), and Research Fellow at IAB (Institut für Arbeitsmarkt-und Berufsforschung, Nürnberg, Germany). He is the past President (2014-2015) and Fellow of the Society of Labor Economists; past Chair (2013) of the Business and Economic Statistics Section and Fellow of the American Statistical Association; elected member of the International Statistical Institute; and a fellow of the Econometric Society. He has served as Distinguished Senior Research Fellow at the United States Census Bureau (1998-2016) and on the National Academies' Committee on National Statistics (2010-2016). He currently serves on the American Economic Association's Committee on Economic Statistics (2013-2018). He was the Director of the Cornell Institute for Social and Economic Research (CISER) from 1999 to 2007. His current research focuses on the creation, dissemination, privacy protection, and use of linked, longitudinal data on employees and employers.

Sandy Baum, Ph.D.

Title: Senior Fellow, Urban Institute

Dr. Sandy Baum is a fellow at the Urban Institute and professor emerita of economics at Skidmore College. Dr. Baum earned her B.A. in sociology at Bryn Mawr College, where she is currently a member of the Board of Trustees, and her Ph.D. in economics at Columbia University. She has written and spoken extensively on issues relating to college access, college pricing, student aid policy, student debt, affordability, and other aspects of higher education finance.

Dr. Baum has co-authored the College Board's annual publications Trends in Student Aid and Trends in College Pricing since 2002. Through the College Board and the Brookings Institution, she has chaired major study groups that released proposals for reforming federal and state student aid. She has published numerous articles on higher education finance in professional journals, books, and the trade press. She was the principle researcher on the Urban Institute's new website on college affordability and her recent work includes Urban Institute briefs on graduate student enrollments and financing. She is the author of Student Debt: Rhetoric and Realities of Higher Education Financing (Palgrave Macmillan 2016) and co-author with Harry Holzer of Making College Work: Pathways to Success for Disadvantaged Students (Brookings Institution Press 2017). She is a member of the Board of the National Student Clearinghouse.

Mike Brick, Ph.D.

Title: Vice President, Westat

Dr. Brick is a Vice President of Westat, where he is co-director of the Survey Methods Unit and the Statistical Staff. With more than 40 years of experience, he has special expertise in sample design and estimation for surveys, the theory and practice of surveys conducting using telephone, address-based sampling, the Internet and mixed mode surveys. He has developed and implement studies on nonresponse and different sources of bias in surveys. He is also a Research professor in the Joint Program in Survey Methodology at the University of Maryland. His Ph.D. is in Statistics from American University. He has published numerous articles on survey sampling, nonresponse rates and bias, and modes of data collection.

Edward W. (Jed) Frees, Ph.D.

Title: Emeritus Professor, University of Wisconsin-Madison

Dr. Edward W. (Jed) Frees is an emeritus professor affiliated with the University of Wisconsin-Madison where he served as the Hickman Larson Chair of Actuarial Science. He received his Ph.D. in mathematical statistics from the University of North Carolina at Chapel Hill and is a Fellow of both the Society of Actuaries (SoA) and the American Statistical Association (the only Fellow of both organizations). Professor Frees has provided extensive service to the profession, including serving as the founding chairperson of the SoA Education and Research Section, a member of the SoA Board of Directors, a Trustee of the Actuarial Foundation, the Editor of the North American Actuarial Journal, and as an actuarial representative to the Social Security Advisory Board's Technical Panel on Methods and Assumptions. He has written three books, edited a two-volume series on Predictive Modeling Applications in Actuarial Science, and is editing an online, open source book Loss Data Analytics. Regarding his research, Professor Frees has published extensively and won several awards for his work. He has won the Society of Actuaries' Annual Prize for best paper published by the Society, the SoA's Ed Lew Award for research in modeling, the Casualty Actuarial Society's Hachmeister award, and the Halmstad Prize for best paper published in the actuarial literature (four times).

Jamienne S. Studley

Title: President, WASC Senior College and University Commission (WSCUC)

Jamienne S. Studley became the sixth president of the WASC Senior College and University Commission (WSCUC) on January 16, 2018. Ms. Studley was deputy undersecretary of the U.S. Department of Education from 2013 to 2016. She also acted during vacancies in the positions of undersecretary and assistant secretary for postsecondary education. Ms. Studley served on the federal advisory committee on accreditation, the National Advisory Committee on Institutional Quality and Integrity (NACIQI) from 2008 to 2013 (chair, 2011-13). Earlier Ms. Studley was the department's deputy and then acting general counsel from 1993 to 1999.

Ms. Studley has served in a number of other higher education roles, notably as the first female president of Skidmore College and as Associate Dean & Lecturer in Law at Yale Law School. She has also been professor of practice (public policy), Mills College; adjunct faculty, UC Berkeley and Stanford Law Schools; board member, Association of American Colleges & Universities; and Visiting Committee, Harvard Law School. A graduate of Barnard College (magna cum laude and Phi Beta Kappa) and Harvard Law School, Ms. Studley also served as National Policy Advisor for Beyond 12 and an independent consultant on institutional effectiveness, accreditation, and leadership. Her nonprofit leadership experience includes CEO and now President Emerita of Public Advocates Inc. and executive director of the National Association for Law Placement. She serves on the boards of KQED and the Foundation for Student Success.

Daniell Toth, Ph.D.

Title: Senior Research Mathematical Statistician, Office of Survey Methods Research, Bureau of Labor Statistics

Dr. Daniell Toth is the Senior Research Mathematical Statistician in the Office of Survey Methods Research at the U. S. Bureau of Labor Statistics. He has a Ph.D. from the Department of Mathematics at Indiana University and is currently an associate editor for The American Statistician and for Survey Methodology. He has published research in the areas of survey methodology especially in developing tree-based methods for complex-sample designs, their application to nonresponse analysis, and disclosure limitation methodology, is the author of the R-package rpms, which allows users to build design consistent regression trees using survey data, and an active member and fellow of the American Statistical Association.

Panel convened by National Institute of Statistical Sciences

Nell Sedransk, Ph.D.

Title: Director, National Institute of Statistical Sciences-DC

Dr. Nell Sedransk is the Director of the National Institute of Statistical Sciences. She is an Elected Member of the International Statistical Institute, also Elected Fellow of the American Statistical Association. She is coauthor of three technical books; and her research in both statistical theory and application appears in more than 60 scientific papers in refereed journals. The areas of her technical expertise include: design of complex experiments, Bayesian inference, spatial statistics and topological foundations for statistical theory. She has applied her expertise in statistical design and analysis of complex experiments and observational studies to a wide range of applications from physiology and medicine to engineering and sensors to social science applications in multi-observer scoring to ethical designs for clinical trials.