2008 Research Report: Center for Economic Studies and Research Data Centers

Issued May 2009

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

MISSION

The Center for Economic Studies partners with stakeholders within and outside the U.S. Census Bureau to improve measures of the economy and people of the United States through research and innovative data products.

HISTORY

The Center for Economic Studies (CES) opened in 1982 to house new longitudinal business databases, develop them further, and make them available to qualified researchers—building on the foundation laid by a generation of visionaries, including Census Bureau management and outside academic researchers. Pioneering CES staff joined with qualified academic researchers who visited the Census Bureau to begin fulfilling those visions. Using the new data, they produced analyses that contributed to a revolution of empirical work in the economics of industrial organization. CES and the Census Bureau identified a strategy—research data centers (RDCs)—to expand researcher access to these important new data while ensuring the strict terms of access required by the Census Bureau's legislation and laws and regulations of other providers of data made available to RDC researchers. The first RDC opened in Boston, Massachusetts, in 1994.

ACKNOWLEDGMENTS

The contributions of individuals from within and outside the Census Bureau were essential to producing this report. **B.K. Atrostic** wrote Chapter 1; **Matthew Graham** wrote Chapter 2; **John Haltiwanger**, **Ron Jarmin**, and **Javier Miranda** coauthored Chapter 3; **Richard Burkhauser** and **Jeffrey Larrimore** of Cornell University coauthored Chapter 4; **Randy Becker** and **Cheryl Grim** coauthored Chapter 5; and **Todd Gardner** wrote Chapter 6. Thanks also are due to **Lynn Riggs** who organized Appendix 1; and **Angela Andrus**, **Cheryl Grim**, **Brian Holly**, **Sang Nguyen**, **Shigui Weng**, and **Jeremy Wu** who contributed to or worked on Appendixes 1 through 7.

Janet Sweeney and **Jamie Stark** of the Administrative and Customer Services Division, **Francis Grailand Hall**, Chief, provided publications and printing management, graphics design and composition, and editorial review for print and electronic media. General direction and production management were provided by **Claudette E. Bennett**, Assistant Division Chief, and **Wanda Cevis**, Chief, Publications Services Branch.

Finally, our RDC partners and administrators provided assistance.

DISCLAIMER

This report has not undergone the review accorded Census Bureau publications and no endorsement should be inferred. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the Census Bureau or other organizations. All results have been reviewed to ensure that no confidential information is disclosed.

2008 Research Report: Center for Economic Studies and Research Data Centers

Issued May 2009

U.S. Department of Commerce
Gary Locke,
Secretary

Vacant,
Deputy Secretary

Economics and Statistics Administration
Vacant,
Under Secretary for Economic Affairs

U.S. CENSUS BUREAU
Thomas L. Mesenbourg,
Acting Director

Suggested Citation

U.S. Census Bureau,
2008 Research Report: Center for Economic Studies and Research Data Centers,
U.S. Government Printing Office,
Washington, DC, 2009.

Economics and Statistics Administration

Vacant,
Under Secretary for Economic Affairs

U.S. CENSUS BUREAU

Thomas L. Mesenbourg,
Acting Director

Thomas L. Mesenbourg,
Deputy Director and Chief Operating Officer

C. Harvey Monk, Jr,
Associate Director for Economic Programs

C. Harvey Monk, Jr,
Acting Assistant Director for Economic Programs

Ron S. Jarmin,
Chief, Center for Economic Studies

Contents

Appendixes

A MESSAGE FROM RON S. JARMIN, PH.D.
Chief Economist and Chief of the Center for Economic Studies

The U.S. Census Bureau strives to meet the need for reliable and up-to-date information by leveraging the nation's investment in its statistical infrastructure to create new information from existing data. The Center for Economic Studies (CES) has proven to be an indispensable asset to the Census Bureau in achieving this objective. As this report demonstrates, staff and researchers working at CES and its network of Research Data Centers (RDCs) create innovative new data products and conduct insightful analyses from the rich data infrastructure housed at the Census Bureau.

Our role as a provider of innovative data products was substantially enhanced in 2008 with the move of the Longitudinal Employer-Household Dynamics (LEHD) program to CES. Through the unique federal-state Local Employment Dynamics partnership, LEHD maintains a nearly universal quarterly national job frame (quarterly observations of workers linked to employers) based on state administrative records linked to census, survey, and other administrative data housed at the Census Bureau. From this, LEHD produces several public-use products including the Quarterly Workforce Indicators (QWI) and OnTheMap, which is described in more detail in Chapter 2 of this report.

This past year also saw the introduction of the Business Dynamics Statistics (BDS), a public-use product developed from CES's confidential Longitudinal Business Database. Chapter 3 contains more information on the rich information about life cycle dynamics of U.S. businesses contained in the BDS.

The existence of these popular products owes much to perhaps the most crucial function CES plays within the U.S. federal statistical system—providing secure and authorized researcher access to confidential statistical microdata. Results obtained by researchers working at CES and the Census Bureau's RDCs, some dating as far back as the 1980s, inspired and informed the development of the QWI, OnTheMap, BDS, and other products. In fact, much of the current research on enhancing and developing these and other new products is carried out in the RDCs.

The RDCs are operated as joint partnerships between the Census Bureau and leading research institutions. Through the RDCs, researchers with approved projects can access the Census Bureau's rich data infrastructure. The RDCs are a cost-effective way for the Census Bureau to leverage internal resources. Using RDCs significantly enhances its analytical and methodological capabilities by accessing external expertise.

An excellent example of the benefits of accessing external expertise is given in Chapter 4 of this report. Chapter 4 describes methods to improve income distribution statistics generated from the Current Population Survey. Additional examples of the research conducted through the RDC network are given in this and previous CES Annual Reports and on our Web site <www.ces.census.gov>.

Our RDC partners benefit through their ability to provide their faculty and students with an unparalleled resource for applied social science research. More and more, we see RDCs used to attract talented faculty to universities with access to a RDC. I am pleased to report that the RDC network will soon be adding two nodes. First, we will be establishing a new RDC at the University of Minnesota. Second, the RDC at the University of California at Berkeley will open a branch lab at Stanford University. Branch RDCs are a new development. They are satellites of existing RDCs and share resources with the primary RDC. Branches are to be located near a significant number of users and potential users that face significant costs in commuting to the existing RDC, but close enough to the existing RDC to permit sharing resources, such as faculty advisors and RDC administrators. Branch RDCs are one way that the Census Bureau and its RDC partners strive to expand authorized access while keeping costs low.

Finally, I want to thank all those who contributed to this report. First and foremost, I want to thank B.K. Atrostic and Cheryl Grim, who compiled and edited nearly all the material in the report. I thank also the many other contributors, from both within and outside the Census Bureau, listed in the acknowledgments section of the inside cover page.

Chapter 1.
INTRODUCTION

B.K. Atrostic, Center for Economic Studies

The challenges facing our nation require high-quality information about its economy and people. A key Census Bureau objective is: "Meet the needs of policymakers, businesses, nonprofit organizations, and the public for current measures of the U.S. population, economy, and governments" (U.S. Census Bureau, 2007). The Center for Economic Studies (CES) supports this objective by producing fundamental research and innovative data products.

CES creates new research results and data by developing and applying cutting-edge techniques to existing data. Using existing data is cost-effective because it leverages the significant investments already made by the Census Bureau, other agencies whose data we use, and respondents.

Research by CES staff and researchers at its network of Research Data Centers (RDCs) creates new information that could not be produced using publicly available data.[1] In many cases, no public-use microdata even exist.

Two new innovative CES public-use data products—OnTheMap and the Business Dynamics Statistics (BDS)—fill gaps in existing public information. These products exemplify the way CES develops new data and data products in partnership with a growing array of stakeholders within and outside Census Bureau.

The Longitudinal Employer-Household Dynamics (LEHD) program created OnTheMap. LEHD, which joined CES in 2008, relies on partnerships with data providers across the country. OnTheMap could not have been

Text Box 1-1.
WHAT IS A RESEARCH DATA CENTER (RDC)?

RDCs are Census Bureau facilities staffed by a Census Bureau employee that meet all physical and computer security requirements for access to confidential data. At RDCs, researchers from academia, federal agencies, and other institutions with approved projects receive restricted access to Census Bureau data files that are not publicly available.

The Center for Economic Studies (CES) judges each proposal against five standards:

- Potential benefits to the Census Bureau.
- Scientific merit.
- Clear need for restricted data.
- Feasibility with data available in the RDC system.
- No disclosure risk.

Proposals meeting these standards are reviewed by the Census Bureau's Office of Analysis and Executive Support. Proposals approved by the Census Bureau may also require approval by the federal agency sponsoring the survey or supplying the administrative data.

Researchers must become Special Sworn Status (SSS) employees of the Census Bureau. Like career Census Bureau employees, SSS employees are sworn for life to protect the confidentiality of the data they access. Failing to protect confidentiality subjects them to significant financial and legal penalties. The RDC system and the CES proposal process are described in detail on the CES Web site <www.ces.census.gov>.

Selected confidential data from the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS) can be accessed in the RDCs. Proposals must meet the requirements of AHRQ (AHRQ, 2009) or NCHS (NCHS, 2009).

[1] RDCs are described in Text Box 1-1.

made public without state-of-the-art disclosure avoidance techniques created through partnering with RDC researchers. The BDS build on the results of years of CES and RDC research. Current research and development of OnTheMap, the BDS, and other new products is carried out at CES and in the RDCs.

Data underpin all CES products and research results. This report emphasizes new CES public-use data products and improvements in the quantity and quality of data available to researchers at CES and the RDCs. Our new data products and growing research program are recognized within and outside the Census Bureau and by organizations around the world.

2008 NEWS

LEHD Joins CES

In March 2008, the LEHD program joined CES, which comes almost a decade after LEHD was first established in late 1998. LEHD brings with it expertise in an extensive array of business and household data and in state-of-the-art disclosure protection techniques. We welcome the talent of LEHD's research and production staffs and look forward to working with our new colleagues.

The move adds LEHD's innovative data products to the CES portfolio. LEHD produces them through its relationship with the Local Employment Dynamics (LED) partnership. The LED partnership expanded in 2008 to include 50 of a potential 53 state partners (the District of Columbia, Puerto Rico, and the Virgin islands are eligible partners). We are working with LEHD and its state partners to improve and grow this important program.

Wide Recognition for CES and RDC Research

CES and RDC research was published in major journals, including the *American Economic Review*, the *Journal of Political Economy,* and *Economic Perspectives.* Our research also appeared in field journals in a range of disciplines, such as the *Journal of Agricultural Economics,* the *Journal of Ethnic and Migration Studies, Labour Economics*, and the *Journal of Urban Economics*.

CES researchers presented new findings on private equity financing at a major venue, the 2008 World Economic Forum at Davos, Switzerland, and at several conferences during the course of 2008. New research based on the same data was presented at the 2009 World Economic Forum.

Awards and Recognition for LEHD Public-Use Data Products

Two LEHD public-use data products received awards and recognition.

Older Worker Profile. The Older Worker Profile series covers the employment history and characteristics of older workers and their employers, while preserving the confidentiality of the underlying responses. Cited as "an innovative use of a unique new data source," the series will be included in the *Recognition of Excellence in Aging Research* Committee report (U.S. Senate, 2009).

OnTheMap. OnTheMap was cited in the 2008 *Economic Report of the President* as an example of the Census Bureau's efforts to improve the usefulness of the data it collects while protecting the confidentiality of respondent information (see Text Box 1-2).

OnTheMap was 1 of 4 featured papers on U.S. statistical innovations at a United Nations seminar, "Innovations in Official Statistics," held in February 2009.

Text Box 1-2.
RECOGNITION FOR ONTHEMAP

"Government agencies strive to improve the usefulness of their data products by providing greater detail while protecting the confidentiality of respondents ... The Census Bureau creates synthetic data to obscure the underlying demographic data used in its 'On the Map' feature. This feature creates maps showing commuting patterns and workforce data—where people live and work by age, earnings, and industry—for geographic areas selected by the user."

Source*: Economic Report of the President* (Executive Office of the President, 2008).

Photo by Lauren Brenner

Lynn Riggs (center left), Arnold Reznek (center), and Bill Yates (center right) receive a Bronze Medal Award from former Director Steve Murdock (far left) and Associate Director for Economic Programs Harvey Monk (far right).

Bronze Medal for New Disclosure Avoidance Module

Central to assuring the confidentiality of respondent information is the clearance of statistical output; a process known as disclosure avoidance review. CES Disclosure Officer Arnold Reznek, Lead RDC Administrator Lynn Riggs, and Special Assistant for Systems Support Bill Yates developed a new Disclosure Avoidance Review Tracking System (DARTS) module for the CES proposal management system. DARTS is a Web-based module providing an efficient way to assign and track clearance requests from RDC researchers.

The three DARTS developers received the Census Bureau Bronze Medal, the highest honor the Census Bureau can award. DARTS was cited for resulting in more effective and efficient management systems. It reduces administrative burden for RDC researchers and CES project administrators while providing transparency and accountability.

Mentorship Program

Developing an intellectual network of researchers trained in using Census Bureau data improves our research and data products. Many graduate students who worked as research assistants on RDC projects have themselves become RDC researchers after beginning their professional careers.

In 2008, CES instituted a formal mentorship program to assist graduate students actively engaged in dissertation research in economics or a related field at an RDC. CES assigns staff economists as mentors. Participants receive additional mentoring through invitations to visit CES, where they can meet CES researchers and present research.

Two Ph.D. students are participating in the new program. More information is available at the CES Web site <www.ces.census.gov>.

Strengthened Partnership for Research Using Household Data

To better support projects using household data, CES devoted significant resources over the last few years to improving documentation. Dr. John Iceland, then of the University of Maryland, developed documentation for the March supplement to the Current Population Survey (CPS) (now the Annual Social and Economic Supplement [ASEC]) and the Survey of Income and Program Participation (SIPP). In 2008, RDC researchers received improved and consistent documentation for the 1967–2005 CPS files and the 1984–1990 SIPP files.

Recognizing the benefits of RDC research to its programs, the Decennial Directorate of the Census Bureau committed to funding a CES staff position supporting projects using household data. Dr. Todd Gardner joined the CES Data Staff in 2008. Dr. Gardner, a historian, was previously with the Population Division of the Census Bureau. Dr. Gardner was responsible for bringing to fruition a long-running project to make more internal historical decennial microdata files available to

RDC researchers. The project is described in Chapter 5. CES appreciates the support of the Decennial Directorate and welcomes Dr. Gardner.

Strong and Growing Research Program

The CES research program is strong and growing. A few years ago, then-Chief Dan Weinberg noted that the RDC network was underutilized (Weinberg, 2006). For example, only 19 new projects began in 2005. Weinberg identified potential reasons why. He also laid out steps CES was taking to increase the number of research projects with potential to both create "state-of-the-art, groundbreaking, publishable, high quality academic research" and yield benefits to the Census Bureau. Those steps are bearing fruit.

In 2008, over 40 new projects started, bringing the total number of active projects to more than 100. The number of active researchers grew to more than 400. Research increasingly spans multiple RDCs and involves both CES and RDC researchers. The RDCs at CES headquarters (in Suitland, Maryland), Chicago, Michigan, and Triangle are near capacity.

More projects yield more products. More than 40 CES Discussion Papers were issued in 2008, a strong increase over the 33 papers issued in 2007. See Appendix 3.

Researchers at CES and the RDCs published more than 40 professional papers in 2008, keeping pace with recent years. CES and RDC publications and working papers are listed in Appendix 1.

Text Box 1-3.

CES PARTNERS

The Center for Economic Studies (CES) relies on networks of supporters and partners within and outside the Census Bureau. Our primary partners are listed below. All of our partners make vital contributions and we thank all of them.

Census Bureau business and household program areas. CES and the Research Data Centers (RDCs) receive ongoing help from many areas of the Census Bureau producing business and household data. That help takes many forms, including:

- Microdata—
  - Additions and expansions of data available to RDC researchers in 2008 are listed in Appendix 4.
  - Census Bureau business and household datasets that are part of the LEHD data infrastructure.
- Expert knowledge of the collection and processing methodologies underlying the microdata.
- Reviews of RDC research proposals, particularly for household data.

RDC partners. CES currently operates nine RDCs around the country in partnership with a growing roster of prominent research universities and nonprofit research organizations. Our RDC partners are recognized in Appendix 5.

LEHD partners. The LEHD program produces its public-use data products through its LED partners. Partners as of October 2008 are acknowledged in Appendix 6.

Other Census Bureau partners. Colleagues in the Economic Directorate's Administrative Staff who provide administrative support to CES are recognized in Appendix 7. CES also benefits from colleagues in several other Census Bureau divisions who support our computing infrastructures.

CES researchers helped organize and participated in a major international conference in 2008. Lucia Foster and Shawn Klimek of CES served on the Scientific Committee that selected participants to the eighth Conference on the Analysis of (micro) Economic Data (CAED). CAEDs bring together U.S. and international researchers who conduct similar research using enterprise microdata.

The May 2008 CAED was held in Budapest, Hungary. Ron Jarmin participated in the opening plenary session. CES researchers presented 11 papers, often jointly with RDC researchers.

Ron Jarmin and Shawn Klimek serve on the Executive and

Conference Planning Committees for upcoming CAED conferences planned for:

- Tokyo, October 2–4, 2009.
- San Francisco, fall 2010.
- Germany, spring 2012.

More information on past and future CAED conferences is available at the CAED Web site <www.ces.census.gov/index.php/caed/caedhome>.

PARTNERS AND SUPPORTERS

CES relies on the support of Census Bureau and external partners. Our formal partners are recognized in Text Box 1-3.

REPORT OVERVIEW

This year's report focuses on data—new data for use by CES researchers and new public-use products they developed. Creating the new data and products required strong and long-standing partnerships with other parts of the Census Bureau, RDC researchers, and academic institutions.

Two new public-use data products—OnTheMap and the BDS—help meet needs for current measures of the nation's economy and people. These products are described in Chapters 2 and 3.

Data from the LEHD program are a major addition to the stock of data available to RDC researchers. Those data, from the same complex data infrastructure underpinning the new OnTheMap product, are described in Appendix 4.

RDC research improved income measures in the public-use CPS. Users now have income measures with distributions closer to official measures produced from the internal file. That research is presented in Chapter 4.

We doubled the decennial census data available. Data for 1970 and 1980 were added to the data for 1990 and 2000 already available. The new data and their improved documentation are described in Chapter 5. All of the Census Bureau data series added or expanded in 2008 are listed in Appendix 4.

CES has identified sources for decennial census data for years before 1960 and for manufacturing data that could go as far back as 1820. Chapters 5 and 6 describe these data and the steps needed to make them usable by researchers. CES does not have the resources needed to recover and develop these data. Researchers interested in partnering with CES should contact the CES Assistant Division Chief for Research, Lucia Foster, at <Lucia.S.Foster@census.gov>.

REFERENCES

Agency for Health Care Research and Quality. 2009. Restricted Data Files Available at the Data Centers. <www.meps.ahrq.gov/mepsweb/data_stats/onsite_datacenter.jsp>.

Executive Office of the President. 2008. *Economic Report of the President*. Washington, DC: U.S. Government Printing Office.

National Center for Health Statistics. 2009. Research Data Center. <www.cdc.gov/nchs/r&d/rdc.htm>.

United Nations. 2009. Seminar on Innovations in Official Statistics. February 20. <http://unstats.un.org/unsd/statcom/statcom_09/seminars/innovation/innovations_seminar.htm>.

U.S. Census Bureau. 2007. *Strategic Plan FY2007–2012*. Strategic Goal #2 at <www.census.gov/main/www/strategicplan/strategicplan.html#4-4>.

U.S. Senate. 2008. Recognition of Excellence in Aging Research. <www.aging.senate.gov/award/award_main.cfm>.

Weinberg, Dan. 2006. "Message from the Chief, Center for Economic Studies." In *Research at the Center for Economic Studies and the Research Data Centers: 2005.* <www.ces.census.gov/index.php/ces/cmsdownloadsin>.

Chapter 2.
ONTHEMAP: AN INNOVATIVE MAPPING AND REPORTING TOOL

Matthew Graham, Center for Economic Studies

As government economic policy and private business planning increase in importance, key questions about the workforce are reiterated at local, regional, and national levels:

- Where are the jobs?
- How far do workers live from their jobs?
- What industries are employing an area's workers?
- How are these factors changing over time?

OnTheMap, a free, publicly accessible Web application from the U.S. Census Bureau, can answer those questions directly. It provides information on jobs and workers in the United States at an unprecedented level of geographical detail.

OnTheMap makes use of administrative records, censuses, and surveys to produce a vast longitudinal dataset of jobs and workers in the United States. The data system that supports OnTheMap uses cutting-edge synthetic data and noise infusion techniques to preserve analytical validity while protecting confidentiality.

Users of OnTheMap do not need any special knowledge of statistical or mapping software. Results are presented in easy-to-read maps and reports, for user needs (See Text Box 2-1).

OnTheMap was developed through the Local Employment Dynamics (LED) Partnership between the Census Bureau and its state partners. This federal-state partnership creates a new, modern, collaborative, and cost-effective statistical system.

21ST CENTURY STATISTICAL NEEDS

Understanding the dynamics of our nation's economy and society requires 21st century statistical tools. Key features of the new tools are:

- Flexible geographies.

Text Box 2-1.
ONTHEMAP VERSION 3

Version 3 of OnTheMap was released in December 2008. Current coverage includes 46 of the 53 possible states and territories (the District of Columbia, Puerto Rico, and the Virgin Islands are eligible).

Current data in OnTheMap can be separated into two types:

- *Characteristic data*—age, earnings, and the industry sector of workers.
- *Geographic data*—the spatial relationship between residence and employment locations.

Different statistics on jobs and the workforce are available based upon whether a job is in the private sector and whether it is a worker's primary job. (Only the job with the highest earnings is considered to be a worker's primary job.)

OnTheMap can perform analyses for specific segments of the workforce. The specific segments are:

- *Age*—30 or younger; 31–54; and 55 or older.
- *Earnings*—$1,200 per month or less; $1,201–$3,400 per month; and $3,401 per month or more.
- *Industry*—goods producing; trade, transportation, and utilities; and all other services.

Analyses can focus on areas of residence, areas of employment, or a combination of the two (i.e., those workers living in one town and employed in another).

OnTheMap produces maps and concise summary reports. The maps show concentrations of employment or residences as well as spatial relationships between home and work (i.e., In what areas do those employed downtown live?). The reports present socioeconomic information on selected workers or summarize their distribution by geographic category (cities, counties, ZIP codes, etc.).

- Up-to-date information.
- Easy interactive access and visualization.
- Strict protection of respondent confidentiality.

These requirements are core design elements of LED's data system and of OnTheMap.

Flexible Geography

Statistics are often reported by standard geographic boundaries such as states and counties. However, standard geography may not conform to users' specific needs.

Nor are all members of a geography set the same size. For example, the largest county in the United States is Los Angeles County, CA, with over 9 million people while the smallest is Loving County, TX, with fewer than 100 people.

A 21st century data system needs data that can be aggregated to any desired geographic boundaries. OnTheMap's data are made available at the census block level, of which there are more than 8 million in the United States. Almost any analysis area can be built by aggregating the necessary census blocks in OnTheMap.

Up-to-Date Data

Each year the entire data series is updated with new annual data and any corrections or improvements in data or methodology. This means that users of OnTheMap are viewing the most current version of the data.

Easy Interactive Access and Visualization

Large and complex datasets can be difficult to access for users without specialized computer systems or knowledge.

OnTheMap delivers a dataset with hundreds of millions of records through a simple, visual interface. Interacting with the map interface requires just a few clicks of a mouse. And only four basic steps produce an analysis. OnTheMap allows users to browse data, develop unique queries, customize data visualization, and interact with the data with flexibility.

OnTheMap also provides data in several forms that ease interpretation. Reports list detailed summary statistics, and map overlays show spatial distributions of jobs. The mapping patterns allow users to identify spatial clustering and other patterns that are not immediately obvious from tables of numbers.

Confidentiality Protection

Whatever the potential use of a dataset, the data *cannot* be released without protecting respondents' confidentiality according to federal and state laws.

A long history of concerns about disclosure (Duncan et al., 1993; and U.S. Office of Management and Budget, 2006) partly explain the relative lack of dynamic, detailed depictions of demographic or business statistics in maps.

The actual data from the new LED system cannot be loaded directly into public-use applications like OnTheMap without adequate confidentiality protection. To make OnTheMap possible, the Census Bureau develops and implements state-of-the-art disclosure avoidance methods.

Synthetic data is one of those methods. OnTheMap is the first partially synthetic data product publicly released by the Census Bureau. Wu and Abowd (2007) explain the synthetic data protection techniques applied to OnTheMap.

INNOVATIVE MAPPING FOR THE PUBLIC

OnTheMap provides a number of innovative features:

- Customizable mapping interface.
- Many standard geographies and user-customizable selection tools.
- Characteristic maps and reports on selected jobs (by age, earnings, and industry sector).
- Maps and reports that show spatial relations between home and work.

Customizable Map Interface

Users of OnTheMap can zoom and pan the map to focus on any local or regional area. In addition, the application makes available numerous reference layers which users can display and hide along with associated labels. These include political boundaries, roads, community colleges, and career centers.

The basic map tools let users focus their analyses only on those areas of interest to them.

Standard and Custom Areas of Analysis

OnTheMap provides a number of standard geographic boundaries that can be used to aggregate the workforce data. Boundaries include cities/towns, counties, legislative districts, metropolitan/micropolitan areas, postal ZIP codes, workforce investment areas, census tracts, and tribal lands.

In addition to the standard geographic boundaries, OnTheMap offers a number of "freehand" selection tools. These tools let users interactively specify custom groupings of census blocks, such as business districts, hurricane and tornado paths, transportation corridors, and site location rings.

Characteristic Maps and Reports

OnTheMap's ability to report and map characteristic data of the workforce allows for the identification of clusters of employment or residences. The selected workforce can be grouped by age (three categories), earnings (three categories), and industry (20 North American Industry Classification System [NAICS] sectors or three industry classes).

These simple tools quickly answer questions such as:

- Where is the employment within Los Angeles County?
- What fraction of workers living in Manhattan are at least 55 years old?
- How has the workforce in New Orleans changed in age and earnings composition since Hurricane Katrina?

Linking Residence and Workplace

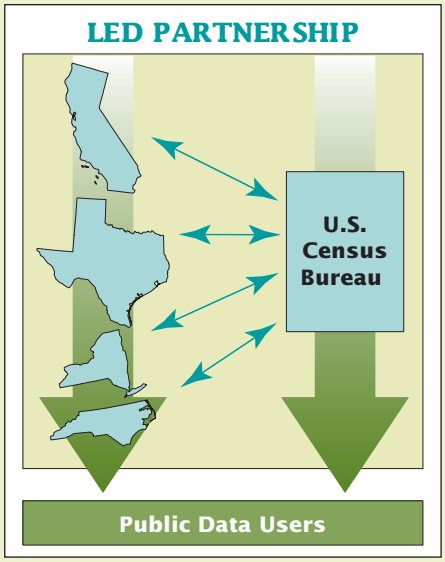
OnTheMap can access information on the spatial relationships between workers' residences

Text Box 2-2.

THE LOCAL EMPLOYMENT DYNAMICS PARTNERSHIP

The Local Employment Dynamics (LED) Partnership is a unique federal-state partnership between the Census Bureau and state labor market information (LMI) agencies.

Through the LED Partnership, the Census Bureau created a longitudinal data system covering up to 190 million jobs held by 150 million workers with over 20 million employers. This system is the combination of administrative records made available by the state partners and administrative records, censuses, and surveys that are available to the Census Bureau.



Sharing Costs and Expertise

The partnership model makes LED a very cost-effective data system because collection and processing costs are shared among the partners. The states collect much of the data through existing systems (the unemployment insurance system and the Quarterly Census of Employment and Wages) and format it for the LED data system. The Census Bureau built and maintains the national data infrastructure, performs the data integration, and releases workforce statistics back to the states and to the public.

Through the partnership, local expertise is available from the states for accurate interpretation of the statistics and for quality assurance for the raw administrative data. The Census Bureau provides national expertise in cutting-edge confidentiality protection models and data integration methods.

By sharing costs and expertise, the LED Partnership produces new, public datasets and data tools, such as OnTheMap and the Quarterly Workforce Indicators, which provide detailed and current workforce information from local to national levels.

and their workplaces. Maps display this information visually, while reports summarize it in select geographic layers.

These work-home linkages are known as "sheds."[1] Taking advantage of the shed maps and reports expands analyses beyond the selection area and lets users answer questions such as:

- How many workers live in St. Paul and work in Minneapolis?
- Which county outside of Denver contributes the most workers to the city?
- What share of St. Louis' workers do not live in Missouri?

Both maps and reports can be saved by the user to be included in papers, presentations, or further analyses.

[1] The OnTheMap application has two kinds of sheds—a commute shed shows where workers who live within a selected area are employed. A labor shed shows in which areas workers who are employed within a selected area live.

REFERENCES

Duncan, G.T., V.A. de Wolf, T.B. Jabine, and M.L. Straf. 1993. "Report of the Panel on Confidentiality and Data Access." *Journal of Official Statistics*, 9(2): 271–274.

U.S. Office of Management and Budget. 2005. "Report on Statistical Disclosure Limitation Methodology." Federal Committee on Statistical Methodology Working Paper 22. <www.fcsm.gov/working-papers/SPWP22_rev.pdf>. (accessed November 30, 2008).

Wu, Jeremy, and John Abowd. 2007. "Synthetic Data for Administrative Record Applications at LEHD." Paper presented at the Joint Statistical Meetings, Salt Lake City, UT.

Text Box 2-3.

A WORKING EXAMPLE OF ONTHEMAP

A simple example shows many of OnTheMap's capabilities: Hurricane Gustav's landfall along the Louisiana coast in 2008.

OnTheMap's tools make it easy to draw and buffer the hurricane's path. The area within 25 miles of the path is our analysis area for primary jobs in 2006. The Work Area Profile Map (Figure 2-1) shows workers' locations within the hurricane's path—jobs clustered around Lafayette and along major highways. Points of different size represent the number of primary jobs in each census block. Larger points represent more jobs.

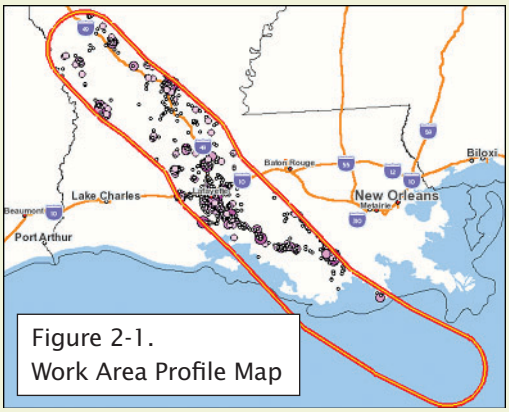


Figure 2-1.
Work Area Profile Map

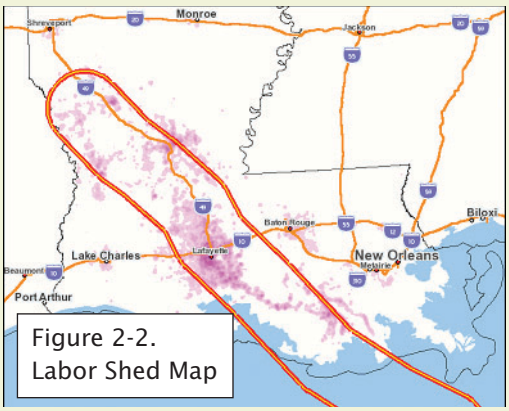


Figure 2-2.
Labor Shed Map

A companion report (not shown) tells us that 28.3 percent of the workers earn less than $1,200 per month and that the top three industry sectors for employment are health care and social assistance, (with 14.4 percent of workers), retail trade (12.5 percent), and manufacturing, (9.3 percent).

A shed analysis of the same area shows in which areas these workers live. The Labor Shed Map (Figure 2-2), shows that most workers live within the hurricane path boundary, but some live beyond it in the Baton Rouge, Lake Charles, and New Orleans areas.

The companion shed report (not shown) shows that over 11.6 percent of these workers live within the city of Lafayette. In comparison, only a combined 1.3 percent live in either Mississippi or Texas. With only a few mouse clicks, an immense amount of economic and spatial data is available.

Chapter 3.
HIGHLIGHTS FROM THE BUSINESS DYNAMICS STATISTICS

Ron Jarmin and Javier Miranda, Center for Economic Studies
John Haltiwanger, University of Maryland

U.S. businesses are very dynamic. They have high rates of business openings and closings, as well as expansions and contractions. Business dynamics are a fundamental part of innovation, productivity growth, and job creation in our economy. Much of what we know empirically about business dynamics comes from research conducted over the last 25 years at the Center for Economic Studies (CES) and the Research Data Centers (e.g., the work of Dunne, Roberts, and Samuelson [1988] and Davis, Haltiwanger, and Schuh [1996]).

Citing these and other studies, the National Research Council's Committee on National Statistics released a 2007 report recommending that federal statistical agencies, among other things, explicitly incorporate business age into their statistics and provide more statistics on young, entrepreneurial businesses. Following on these recommendations, CES released Business Dynamics Statistics (BDS) in 2008.

The BDS address several key shortcomings in official statistics. These shortcomings previously precluded data users from fully appreciating the dynamism of the U.S. economy. The BDS are the first official data product to include high quality measures of firm age. For the first time data users will be able to examine how economic outcomes differ with business age.

INNOVATIVE AND ACCESSIBLE

The richness of the BDS derives from its source, the confidential Longitudinal Business Database (LBD). CES developed the LBD by linking annual snapshots of the Census Bureau's Business Register from 1975 to the most recent year available. The LBD includes data describing the basic characteristics of every establishment with paid employees in the United States as well as of the firms that operate them.[1]

The BDS contain key economic data items, including number of establishments, establishment openings and closings, employment, job creation and destruction, and job expansions and contractions. The data can be downloaded from <www.ces.census.gov/index.php/bds>.

The LBD and BDS were developed by linking existing data and required no new data collection. For a modest investment of resources, CES was able to unlock a vast store of rich information on business dynamics, which once lay unused, and make it available to data users. The BDS represent important progress over existing products with which it shares many characteristics such as the Bureau of Labor Statistics' Business Employment Dynamics (BED) and the Census Bureau's Statistics on U.S. Businesses (SUSB).

The new elements in the BDS provide data users with unprecedented information on the life cycle of U.S. businesses. Text Boxes 3.1–3.3 provide some examples of the types of analyses possible with the BDS but examples represent only a small fraction of the possible uses of these new data. The examples in the text boxes show how the BDS can be used to track entrepreneurial activity across states, measure the employment impact of startups, and examine business dynamics by firm age.

LBD DATA PRODUCTS

The release of the BDS is part of an effort to make relevant information from the confidential LBD accessible to a wide range of users. Other efforts currently underway include creating a synthetic public-use microdata file based on the LBD. The BDS will be updated annually. Expansion to cover new areas will occur as time and resources permit.

[1] An establishment is a single physical business location. A firm is a legal entity with a controlling ownership stake in one or more establishments. Most firms operate only a single establishment. However, many establishments are operated by firms owning multiple establishments.

REFERENCES

Davis, Steven J., John Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. Cambridge, MA: MIT Press.

Dunne, Timothy, Mark J. Roberts, and Lawrence Samuelson. 1988. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries." *RAND Journal of Economics*, 19(4): 495–515.

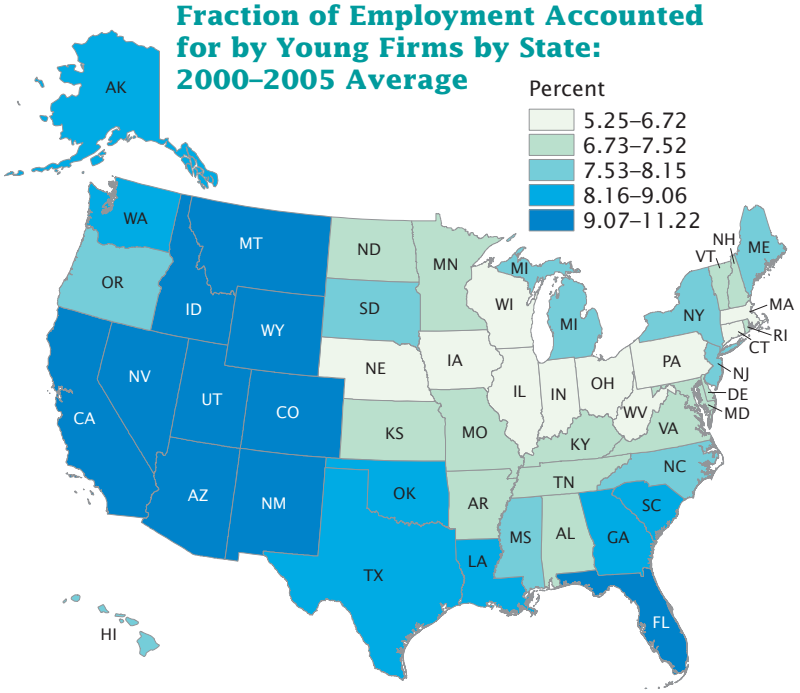
National Research Council. 2007. "Understanding Business Dynamics: An Integrated Data System for America's Future." In *Panel on Measuring Business Formation, Dynamics and Performance*, ed. J. Haltiwanger, L.M. Lynch, and C. Mackie. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC; The National Academies Press.

Text Box 3-1.

ENTREPRENEURIAL ACTIVITY ACROSS STATES

Analysts and policymakers want to understand entrepreneurial activity and the process of job creation, especially at the subnational level. Entrepreneurial activity is generally seen as crucial for innovation and growth. Currently, few reliable data sources permit comparisons of entrepreneurial activity across time and space. One possible measure of entrepreneurial activity is the fraction of private sector, nonagricultural employment accounted for by young firms.

Figure 3-1.
Fraction of Employment Accounted for by Young Firms by State: 2000–2005 Average



Source: Authors' own calculations from Business Dynamics Statistics.

Figure 3-1 depicts the average share of employment accounted for by young firms (less than 3 years old) for each of the 50 states plus the District of Columbia in the Business Dynamics Statistics (BDS). States, mostly from the West and Southwest, have as much as 12 percent of employment accounted for by young firms. In contrast, states, mostly in the East and Midwest, have about 6 percent of employment accounted for by young firms.

Text Box 3-2.

JOBS CREATED FROM BUSINESS STARTUPS IN THE UNITED STATES

Business Startups Critical for Net Job Growth

The fraction of U.S. private-sector nonfarm employment accounted for by business startups over the 1980–2005 period is about 3 percent per year. While this is a small fraction of overall employment, all employment from startups reflects new jobs. As such, 3 percent is large compared to the average annual net employment growth of the U.S. private sector for the same period (about 1.8 percent). This pattern implies that, excluding the jobs from new firms, the net employment growth rate for the United States is negative on average highlighting the importance of business startups to job creation.

Figure 3-2 shows the fraction of jobs due to business startups for all firms and for selected firm size classes: micro firms and midsize to large firms. For micro firms (firms with 1 to 4 employees), the percent of jobs in any given year accounted for by business startups is very large—about 20 percent on average (measured in the figure on the right scale). For larger firms (firms with 250 to 499 employees), the percent of jobs in any given year accounted for by business startups is substantially smaller.

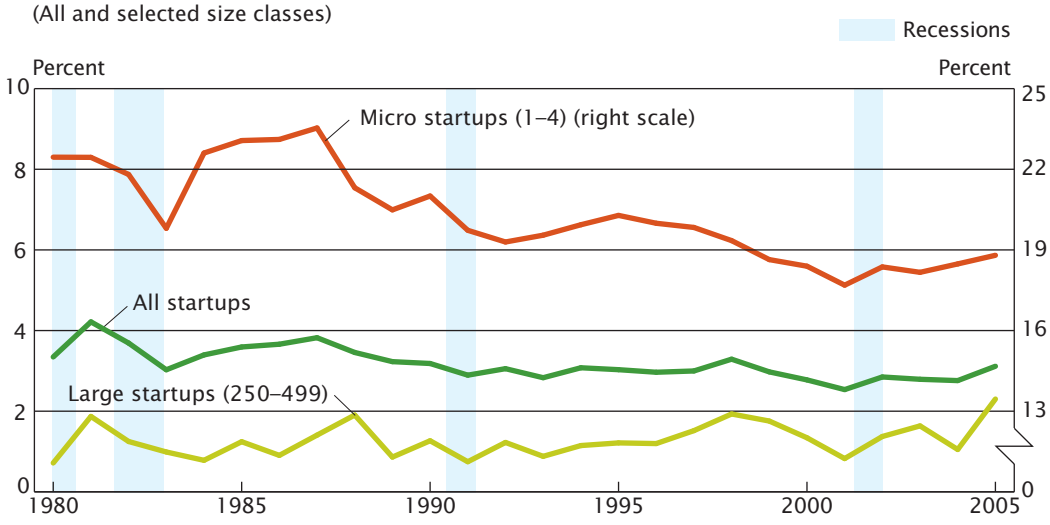
Job Creation Over the Business Cycle

The figure also shows that the role of startups over the business cycle differs in interesting ways across firm size classes. With the exception of the 2001 recession, the startup share of jobs at micro firms drops during downturns. The share of jobs accounted for by startups for all firms and for large firms are, in contrast, considerably more stable. The only notable business cycle decline in the startup rate for all firms and larger startups is in the early 1980s.

Figure 3-2.

Percent of Jobs Accounted for by New Firms

(All and selected size classes)



Source: Authors' own calculations from Business Dynamics Statistics.

Text Box 3-3.

HIGH GROWTH AND FAILURE OF YOUNG FIRMS

What is the role of young entrepreneurial firms in job creation? Until now, researchers lacked comprehensive data broken out by firm age necessary for understanding this fundamental dynamic in our economy.

Young Firms Grow Faster Conditional on Survival but Also Have Relatively High Exit Rates

Figure 3-3 shows the relationship between firm age and two measures of business dynamics: net employment growth and job destruction. The net employment growth rate—employment change at continuing (surviving) establishments and employment growth from opening new establishments—is shown in the blue bars. The job destruction rate is employment change from establishments that do not survive (known as "exits"). Omitted from the chart are establishments of new startup firms that, by definition, have an employment growth rate equal to 200 and no job destruction.

The figure shows that establishments of young firms have higher employment growth rates, if they survive, than those of older firms. For example, establishments belonging to very young firms (age 1) have a net employment growth rate of about 15 percent conditional on survival, whereas those belonging to older firms (aged 29 and over) have a net employment growth rate of about 4 percent conditional on survival. However, young firms experience much more employment loss due to establishment exit—nearly 20 percent at very young firms—than do older firms.

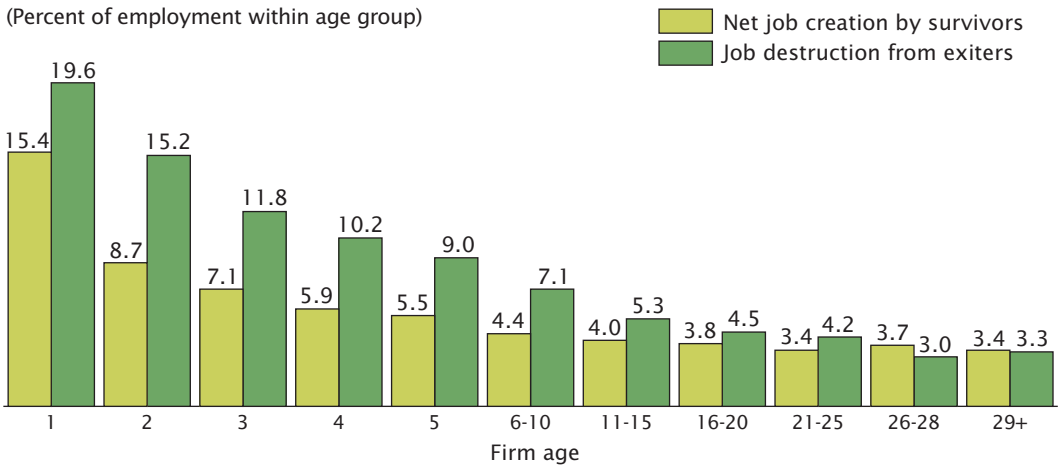
"Up or Out"

The pattern for young firms is thus one of "up or out" with very rapid net growth for survivors offset by a very high exit rate. This complex pattern highlights the importance of developing richer measures of business dynamics such as those in the BDS.

Figure 3-3.

Employment Dynamics for Survivors and Exiters by Firm Age: 1987–2005

(Percent of employment within age group)



Source: Authors' own calculations from Business Dynamics Statistics.

Chapter 4.
BETTER ESTIMATES OF INCOME AND ITS DISTRIBUTION IN THE PUBLIC-USE MARCH CURRENT POPULATION SURVEY[1]

Richard V. Burkhauser[2] and Jeff Larrimore, Cornell University

The March Current Population Survey (CPS) is the primary data source used by public policy researchers and administrators to investigate trends in U.S. income and its distribution. For confidentiality reasons, the U.S. Census Bureau topcodes each of the 24 sources of income (11 income sources prior to 1988) in the public-use CPS. However, this topcoding—the suppression of income values above some level in the public-use CPS data for confidentiality reasons—has not been consistent over time. Below we discuss a new set of papers that offer a solution to this problem using a series of created values, which, when used together with the public-use CPS data, will closely approximate income and inequality levels and their trends based on the internal CPS.

THE PROBLEM

Unsystematic topcoding in the public-use CPS data inconsistently restricts the ability of researchers to fully capture income and its distribution over time. Thus, researchers using scalar inequality measures such as the Gini coefficient with the public-use CPS will unsystematically understate both levels and trends in income inequality. In Figure 4-1, we compare the Gini coefficients using the topcoded public-use March CPS data with those obtained using the internal CPS data. The Public-Unadjusted series is the Gini calculated from the public-use March CPS data exactly as it is provided by the Census Bureau with no adjustments made for topcoding. Similarly, the Internal-Unadjusted series is the Gini calculated from the internal March CPS data with no adjustments made to the data.

Prior to 1995, the Census Bureau reported topcoded incomes as equal to the topcode threshold, which reduced the observed income of all topcoded individuals. Using the topcode threshold led to a substantial understatement in the level of inequality and distorted the trends in inequality. Starting in 1995, the Census Bureau began providing cell means—the mean of all topcoded values for each income source—which, as can be seen in Figure 4-1, allows researchers using the Public-Unadjusted CPS data to closely match results from the internal CPS data after 1995.

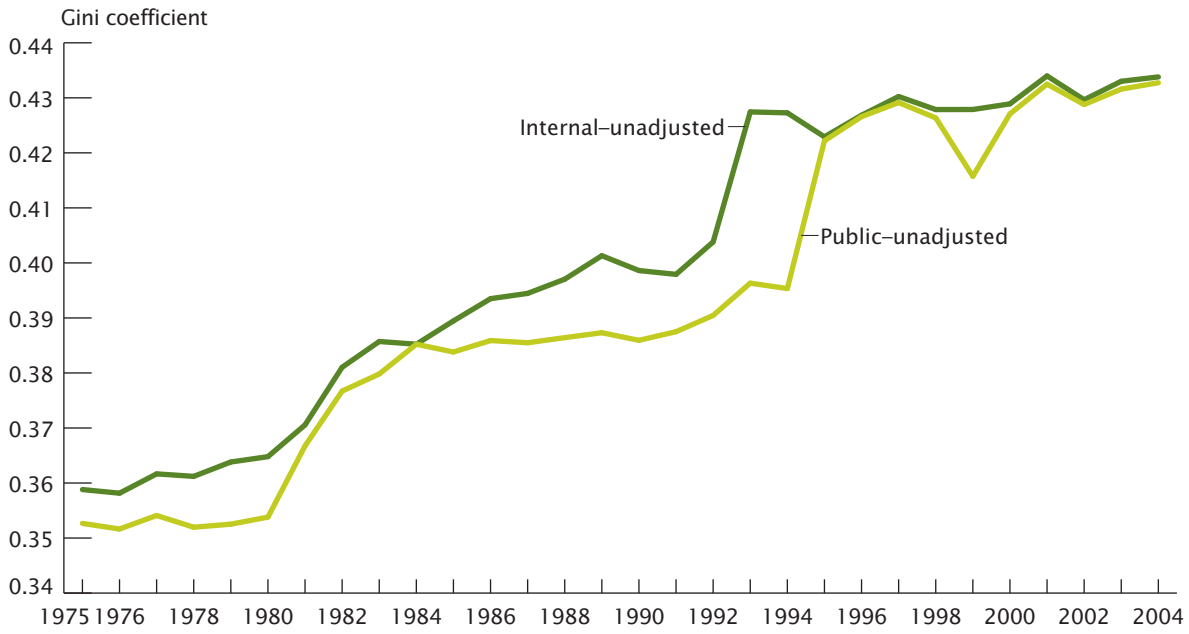
Unfortunately, researchers interested in longer-term income trends have largely ignored this valuable information because cell means were not available for prior years. Instead, many researchers have simply recoded topcoded incomes as equal to the topcode threshold or have opted to use relatively simple measures of inequality, such as the 90/10 ratio, to attempt to avoid topcoding problems. However, Burkhauser, Feng, and Jenkins (2009) show that even researchers using inequality measures, such as the 90/10 ratio, run the risk of understating income inequality. This is because even individuals whose total household size-adjusted income is relatively modest and hence below the 90th percentile may have one or more topcoded sources of household income.

Here we discuss how newly available information that we created about the means and variances of top incomes from the internal CPS data can be used in conjunction with public-use CPS data to largely remedy the problems resulting from inconsistent topcoding. Using the extended cell mean series from Larrimore et al. (2008) and the variances of topcoded incomes from Burkhauser, Feng, and Larrimore (2008), it is now possible to better capture the top part of the income distribution with public-use CPS data and closely replicate results found using the internal CPS data used by the Census Bureau for producing their official inequality statistics (U.S. Census Bureau, various years).

[1] Support for this research from the National Science Foundation (award nos. SES-0427889, SES-0322902, and SES-0339191) and the National Institute for Disability and Rehabilitation Research (H133B040013 and H133B031111) are cordially acknowledged.

[2] Burkhauser completed this paper while he was the R. I. Downing Fellow in Social Economics in the Faculty of Economics and Commerce, University of Melbourne.

Figure 4-1.
Income Inequality is Understated When Using Unadjusted Public-Use CPS Data: 1975–2004



Source: Burkhauser et al. (2008).

USING CELL MEANS AND VARIANCES TO OBTAIN BETTER ESTIMATES OF TOP INCOMES

Despite the ability of cell means to closely replicate the results from internal CPS data, the lack of cell means prior to 1995 has dissuaded researchers from using them when looking at long-term trends using public-use CPS data. Thus, in Larrimore et al. (2008) we accessed internal CPS data and used it to calculate and distribute a cell mean series to the public going back to 1975. We also show that using this series will greatly improve the ability of researchers using only the public-use CPS to capture the top part of the income distribution.

A cell mean more accurately captures the level of unobserved income for a given source of income in the public-use CPS data but does not provide information about its distribution. Therefore, in Burkhauser, Feng, and Larrimore (2008), we also calculate and distribute to the research community the variances of topcoded incomes for each source of income in the internal CPS. By using information on both mean and variance, it is possible to impute different total household income values for each topcoded individual in the public-use CPS data. While doing so, in general, will not match an individual's actual income, it will allow the resulting distribution of income to more accurately match the distribution found in the internal CPS. Below we report how we have used both our cell mean and variance series to better estimate income and its distribution using public-use CPS data.

CREATING MORE CONSISTENT MEASURES OF TRENDS IN HOUSEHOLD INCOME INEQUALITY

Burkhauser et al. (2008) analyze levels and trends in inequality, using Gini coefficients between 1975 and 2004 derived from the internal CPS, and compare them with estimates from several series derived from the public-use CPS. The series and their sources (internal or public-use data) are described in Table 4-1.

Table 4-1.
Definitions for Income Distribution Series by Source and Censoring Method

Acronym	Source	Method for Addressing Censoring Issues
Internal-Unadjusted	Internal	Uses internal data as provided in Census Bureau files, without any adjustments.
Internal-Adjusted	Internal	Topcoded observations replaced by imputations derived from multiple imputation models fitted to internal data; inequality estimates derived using multiple imputation combination methods.
Public-Unadjusted	Public-use	Uses public-use data as provided in Census Bureau files; includes Census Bureau cell mean imputations for topcoded observations from 1995 onwards.
Public-CellMean	Public-use	Uses public-use data as provided in Census Bureau files; includes cell mean imputations for topcoded observations for all years.
Public-NoMean	Public-use	Uses public-use data as provided in Census Bureau files, except that no cell mean imputations used for any year (topcoded values used "as is").
Public-Variance	Public-use	Uses public-use data as provided in Census Bureau files; topcoded observations are replaced by imputations derived from a Stoppa imputation model fitted to the mean and variance of topcoded incomes from the internal data.

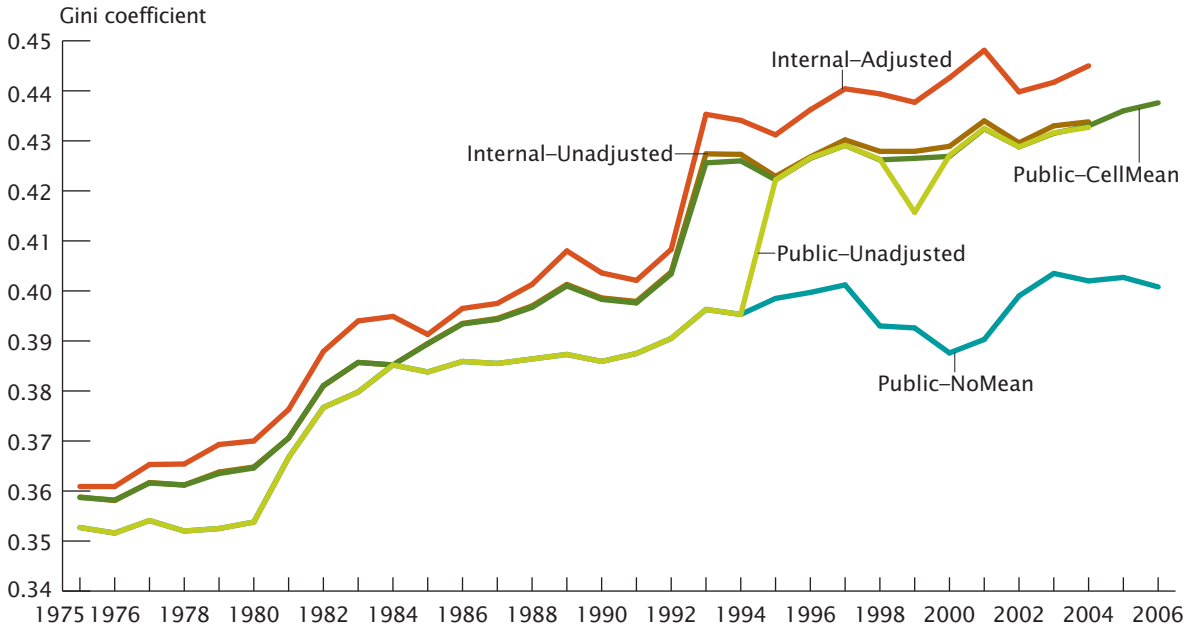
Sources: Burkhauser et al. (2008) and Burkhauser, Feng, and Larrimore (2008).

Figure 4-2, taken from Burkhauser et al. (2008), shows that those who simply use the unadjusted public-use CPS (Public-Unadjusted) will find income inequality jumps dramatically between 1994 and 1995; i.e., the Gini value increases from 0.395 to 0.422, a single year change far greater than in any prior or subsequent year. An increase in the topcode threshold as well as the use of Census Bureau derived cell mean values for all values above the topcode threshold caused this jump. Using the unadjusted internal CPS (Internal-Unadjusted), we find no such increase between 1994 and 1995. Rather, what is happening is that prior to 1995, the Public-Unadjusted CPS Gini values substantially understate income inequality because they fail to fully account for income values above the topcodes. Once the Census Bureau provided the mean value of all these topcoded values, the now more precise Public-Unadjusted CPS Gini values match the higher Internal-Unadjusted Gini values. Failure to account for this change in methodology will grossly overstate U.S. inequality increases before and after 1994–1995.

Simply ignoring Census Bureau mean values after 1994, as can be seen in the Public-NoMean Gini series, will not solve this problem. This series still inconsistently topcodes high values, and it underestimates inequality after 1994, as can be seen by the way its Gini values fall further and further below the Internal-Unadjusted Gini values. In contrast, when we use the public-use CPS together with our extended mean series (Public-Mean) in Figure 4-2, we come much closer to matching the Internal-Unadjusted Gini values in every year.

However, the internal CPS data is itself censored, albeit to a substantially smaller extent than the public-use CPS. Hence, it too has time-inconsistencies, especially in 1992–1993, as

Figure 4-2.
Measured Income Inequality Increases When Using Alternative Topcode Correction Methods Compared to the Unadjusted Public-Use CPS Data: 1975 to 2006



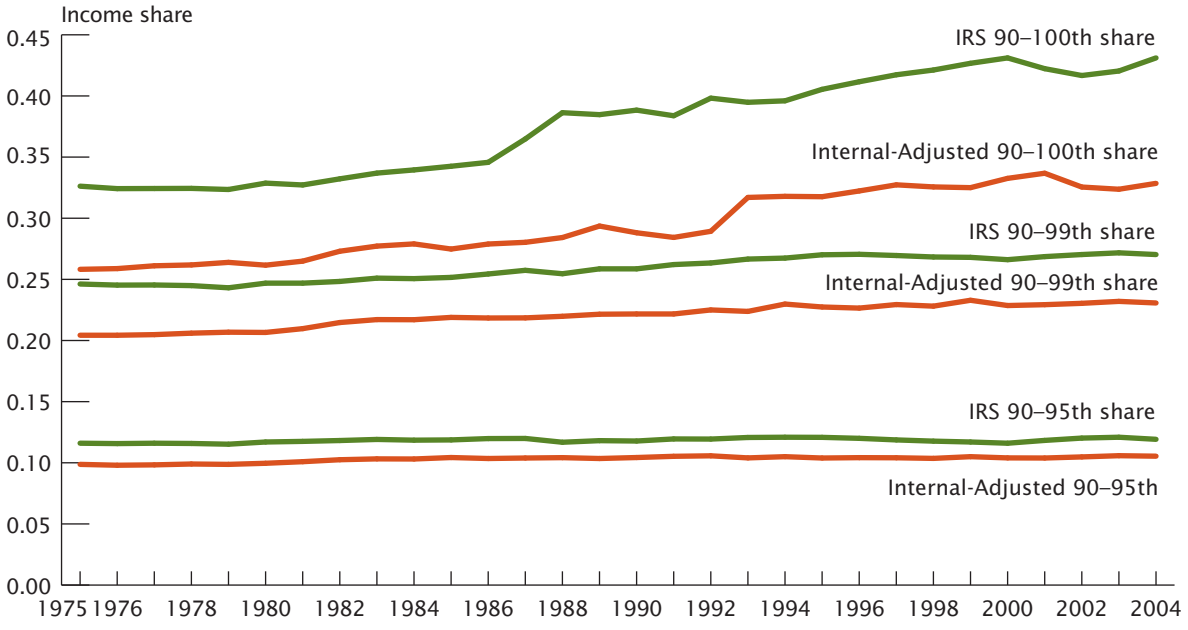
Note: Internal data were not available for years after 2005.
Source: Burkhauser et al. (2008).

can be seen by the jump in the Internal-Unadjusted Gini values between these years. To both control for inconsistent censoring and capture the missing part of the internal CPS data, we use a multiple imputation approach in which, for each year, out-of-sample values for topcoded observations in the internal CPS are imputed based on lower in-sample values that we do have. Unsurprisingly, as can also be seen in Figure 4-2, we find that compared to estimates derived from our multiple imputation approach (Internal-Adjusted), all the other series understate the level of inequality in all years.

However, just as was the case for our Public-Mean series and the Internal-Unadjusted series it replicated, the Internal-Adjusted series reveals the same trends—an increase in inequality over the entire period 1975–2004 but with a rate of increase noticeably lower after 1993 compared to before 1993. In each series, average inequality increases much more prior to 1992 than after 1993. In addition, in each series, the jump in 1992–1993 is far higher than in any other period. This is consistent with the argument that a change in the measurement of inequality rather than a real change in inequality is its cause.

To further test our Internal-Adjusted series, we compared our results to those derived by Piketty and Saez (2003) using Internal Revenue Service (IRS) administrative files. We did so by comparing our estimates of the share of income held by the wealthiest 10 percent of the population to that found by Piketty and Saez (2003). As can be seen in Figure 4-3, we find that these two estimates of the income share held by the top 90th–95th percentiles and the 95th–99th percentiles have remarkably similar levels and trends, especially given that our income units differ as we are observing household income while they are observing the adjusted gross income of tax units. It is only in the share of income held by the richest 1

Figure 4-3.
Similar Shares of Income Held by the Top Part of the Distribution—IRS vs. CPS: 1975 to 2004



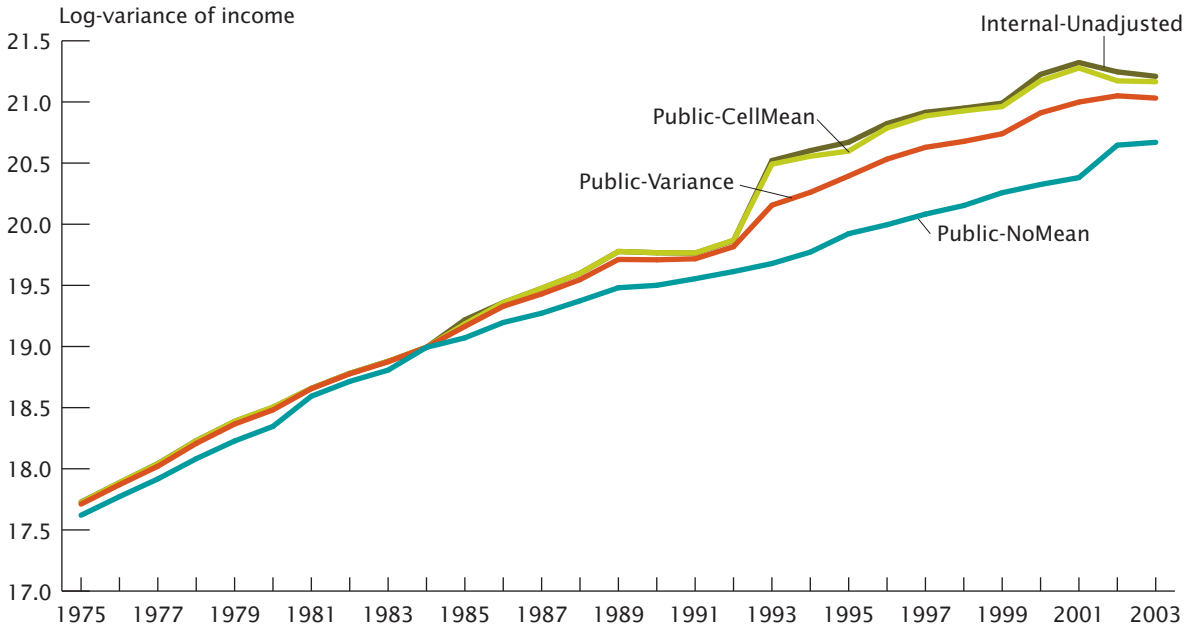
Sources: Burkhauser et al. (2008) and <http://elsa.berkeley.edu/~saez/>.

percent where our levels and trends differ. Piketty and Saez (2003) not only find that a larger share of income is held by this group but that the growth in their share has been greater over time. Differences in the measures of income we use explain differences in levels. However, it is unclear whether the differences in trends are due to an increasing inability of even the internal CPS to capture the very top part of the income distribution or to behavioral changes in the way individual tax units report their adjusted gross income in response to tax law changes in the data analyzed by Piketty and Saez (2003).

CREATING MORE CONSISTENT MEASURES OF TRENDS IN CROSS-GROUP INEQUALITY

While more sophisticated topcode correction methods can improve the accuracy of size-adjusted household income inequality calculations, as shown above, this is just one area where these newly available data and methods can improve calculations using public-use CPS data. Burkhauser and Larrimore (2008 and forthcoming) demonstrate that not correcting for topcoding in the public-use CPS can also distort the size of the earnings gap across subsets of the population. This includes the gap in mean labor earnings between individuals by sex, race, education level, and disability statuses. Researchers using the public-use CPS data without our cell mean series will understate the mean labor earnings of each of these population groups. However, because there are more males than females with topcoded earnings, even consistent topcoding will suppress a larger percentage of male earnings, thus understating the male-female earnings gap. Similarly, White, highly educated, and nondisabled individuals are topcoded at higher rates than Black, less-educated, and disabled individuals, respectively, resulting in an understatement of the race, education, and disability earnings gaps by those using the public-use CPS.

Figure 4-4.
Imputed Cell Means and Variances Better Capture the Log-Variance of Income: 1975 to 2004



Source: Burkhauser, Feng, and Larrimore (2008).

As was the case in our income inequality comparisons above, using cell means largely corrects for these topcoding-based problems in the public-use CPS. When we use the public-use CPS together with our cell means, the earnings gaps we find within sex, race, education, and disability groups are nearly identical to those we find using the internal CPS.

INCORPORATING THE VARIANCES OF TOPCODED INCOMES

Using cell means with the public-use CPS allows researchers to more closely replicate results using the internal CPS. Moreover, researchers interested in more closely depicting the upper tail of the income distribution using the internal CPS can now do so by using the variances of topcoded incomes along with cell means. Burkhauser, Feng, and Larrimore (2008), using variance information from the internal CPS, move beyond imputing the same value for all topcoded individuals. They use a multiple imputation approach to create a distribution for the public-use CPS with the same mean and variance of topcoded incomes found in the internal CPS.

Figure 4-4, taken from Burkhauser, Feng, and Larrimore (2008), compares the log-variance of the income distribution using the public-use CPS without cell means (Public-NoMean) with cell means (Public-Mean) and with cell means and variances of topcoded incomes (Public-Variance) to the log-variance of incomes using the internal CPS data (Internal-Unadjusted). Since the Public-NoMean compresses the distribution by assigning all topcoded individuals' income equal to the topcode threshold, it is not surprising that this series understates the log-variance seen in the internal data. Using the Public-Mean series results in log-variance of income increases since it more accurately captures the level of topcoded income. However, it still understates the log-variance seen in the Internal-Unadjusted series. This is the case since it incorrectly assumes that all topcoded individuals have the same level of income, thus reducing the income variance. It is only when we impute topcoded

incomes using both the cell mean and variance of topcoded incomes in the Public-Variance series that the log-variance of income from the public-use CPS closely matches the Internal-Unadjusted series.

CONCLUSION

The CPS is the primary data source used by public policy researchers and administrators to investigate trends in income and its distribution. However, failure to control for inconsistent topcoding in the public-use CPS will lead researchers to understate the levels of U.S. average income and income inequality as well as distort their trends over time. The series of papers discussed here, based on our ability to access internal CPS data, estimate and distribute the mean and variance of topcoded income values for each income source in these data.

Researchers using our cell mean and variance values together with the public-use CPS data can now closely approximate income and inequality levels and their trends based on the internal CPS. It is still important to consider internal censoring when using these data, especially when observing trends across 1992–1993. However, without making out-of-sample predictions about incomes censored in the internal data, these are the best estimates available. Moreover, they will be nearly as accurate as the official Census Bureau statistics that are published each year based on the internal CPS (U.S. Census Bureau, various years).

REFERENCES

Burkhauser, Richard V and Jeff Larrimore. Forthcoming. "Trends in the Relative Household Income of Working-Age Men With Work Limitations: Correcting the Record Using Internal Current Population Survey Data." *Journal of Disability Policy Studies.*

Burkhauser, Richard V and Jeff Larrimore. 2008. "Using Internal Current Population Survey Data to Reevaluate Trends in Labor Earnings Gaps by Gender, Race, and Education Level." Center for Economic Studies Working Paper CES-WP-08-18.

Burkhauser, Richard V, Shuaizhang Feng, and Jeff Larrimore. 2008. "Measuring Labor Earnings Inequality Using Public-Use March Current Population Survey Data: The Value of Including Variance and Cell Means When Imputing Topcoded Values." Center for Economic Studies Working Paper CES-WP-08-38.

Burkhauser, Richard V, Shuaizhang Feng, and Stephen Jenkins. 2009. "Using the P90/P10 Ratio to Measure U.S. Inequality Trends With Current Population Survey Data: A View From Inside the Census Bureau Vaults." *The Review of Income and Wealth* 55(1): 166–185.

Burkhauser, Richard V, Shuaizhang Feng, Stephen Jenkins, and Jeff Larrimore. 2008. "Estimating Trends in United States Income Inequality Using the March Current Population Survey: The Importance of Controlling for Censoring." Center for Economic Studies Working Paper CES-WP-08-25.

Larrimore, Jeff, Richard V Burkhauser, Shuaizhang Feng, and Laura Zayatz. 2008. "Consistent Cell Means for Topcoded Incomes in the Public-Use March CPS (1976–2007)." *Journal of Economic and Social Measurement* 33(2–3): 89–128.

Piketty, Thomas and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics*, 118(1): 1–39.

U.S. Census Bureau. Various years. *Income, Poverty, and Health Insurance Coverage in the United States.* (August), Current Population Reports, P60 series, Washington, DC: U.S. Government Printing Office.

Chapter 5.
EXPANDED AND ENHANCED DECENNIAL CENSUS DATA FOR RESEARCH

Todd Gardner, Center for Economic Studies

Decennial census microdata files for 1970 and 1980 are now available to researchers at the Research Data Centers (RDCs). Researchers can conduct analyses spanning the 1970–2000 period.

Formats are more consistent and documentation is improved for all years. Much of the documentation is available to the public. Researchers are now able to do significant preparatory work before entering an RDC.

The historical decennial files and documentation have been developed with cooperation from academic research institutions, particularly the Minnesota Population Center (MPC), the Inter-University Consortium for Political and Social Research (ICPSR), and the Maryland Population Research Center (MPRC). The Center for Economic Studies (CES) will continue these cooperative relationships with future data projects, such as the harmonization of the internal microdata files.

The historical decennial files potentially could be extended back to include 1940, 1950, and 1960. Recovery and preparation of the 1960 files is underway.

We need the help of the research community to transform the 1940 and 1950 files into usable research files. Transforming the files requires resources beyond those currently available at CES.

DATA

Microdata files for 1970 and 1980—both long-form (or sample) and short-form (or 100 percent or complete count) files—became available in 2008. The new files complement the existing complete count and sample files for 1990 and 2000.

Figure 5-1.
1970 Census Badge

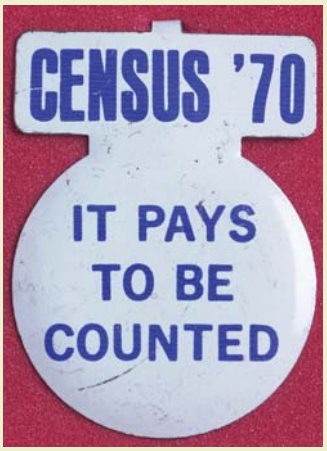


Photo Credit: U.S. Census Bureau, Public Information Office (PIO).

The 1990 decennial census data files also were updated. Record layouts and variable names are now consistent with the 1970 and 1980 files, and some minor problems with the previous version are corrected.

The complete-count files include basic demographic and housing information for all individuals and households at the time of enumeration. The sample files contain fewer records than the complete count files (1 in 5 households received the long form in 1970, while the sample density for the subsequent census years is 1 in 6), but are much richer in terms of the content of the data.

Consistent Over Time

Many of the questions on the census have been asked for decades, yielding comparable data over a broad span of time. For example, in all of the available decennial censuses, respondents were asked where they resided 5 years before the date of the census. The information on prior residence lets researchers analyze the mobility of the population over several decades.

More Observations

Public use microdata sample (PUMS) files contain only a fraction of the cases available in the internal sample file (and none from households that received only the short form). Even the largest PUMS files contain only about a third of the cases available for the long-form files. The larger sample sizes available in the internal files are sometimes necessary for studying small groups in the United States.

More Detail

The internal long-form variables have more detail than the corresponding variables in the PUMS files. For example, data items with dollar values are topcoded and rare categories are combined with other values in the PUMS files. Original values are generally available on the internal files, and while some variables are topcoded in the internal data, the threshold topcode values are generally much higher than in PUMS data.

The detail of the internal long-form files allows for matching at detailed levels of geography between long-form data and other Census Bureau products. For example, Enrico Moretti (2004) makes a block-level match between the location of workers in the decennial census and the location of their employment in the Census of Manufactures. The match uses block-level information on the individual's place of work and industry in the decennial census.

Detailed Geography

Detailed geography is one of the greatest research advantages to using the internal decennial census microdata files. To protect respondent confidentiality, PUMS files restrict geographic detail.

Internal decennial census microdata files, by contrast, have detailed political and statistical geography for census tracts, block groups, and blocks, as well as all places and unincorporated subcounty units. Such detailed geography means that individual- and household-level data can be aggregated to geographic units that are consistent over time.

For example, a city may annex territory. Using internal data, researchers can construct boundaries in each year that are consistent with the original or the expanded city.

Detailed geography also allows the user to aggregate geography into nonstandard geographic classifications, such as school districts. However, researchers must obtain the necessary geographic specifications for the classifications they need.

New Classifications

An internal project is underway to take advantage of the geographic detail in the internal microdata files to develop useful geographic classifications. The new classifications identify clusters of census tracts with high population densities and high employment densities would be available to researchers.

The current Census Bureau standard defining urban areas distinguishes only between urban and rural territory. Urban areas currently are defined at the census block level, with core areas that have population densities of 1,000 people per square mile and surrounding areas that have an overall population density of at least 500 people per square mile.

Urban areas are not homogeneous, however, as population and employment densities vary widely among these areas. The current measures do not capture this variation.

Figure 5-2 shows a proposed five-level geographic classification based on population density at the census tract level. The darkest areas are clusters of census tracts that have population densities over 10,000 people per square mile and populations of at least 50,000 inhabitants. The areas in lighter shades of blue are clusters of census tracts with progressively lower population densities, while the gray areas are outlying census tracts.

This map shows that high-density areas often extend well past the political boundaries of large cities, and that these cities often contain substantial areas of low population density.

Text Box 5-1.
DOCUMENTATION ON DECENNIAL FILES

Available to the public:

- ICPSR documentation.
- SAS zero obs files.

Available in the RDCs:

- Web-based documentation for all variables.
- SAS code with variable labels and value labels for all valid values.
- Codebooks.

DETAILED DOCUMENTATION

New detailed documentation files in a variety of formats make it easier to use the decennial census microdata files. Text Box 5-1 describes the documentation files that CES has assembled.

Some documentation is available to the public through the ICPSR

Web site <www.icpsr.umich.edu/cocoon/ICPSR/STUDY/21820.xml>. Researchers with approved projects now can access additional documentation in the RDCs.

Another publicly available resource is sets of SAS files that replicate the file structure of the internal data files but contain no data. Because they contain zero observations, they are called "zero obs" for short. Zero obs files for the 1970, 1980, 1990, and 2000 decennial census files are available to the public through the Cornell VirtualRDC <www.vrdc.cornell.edu>.[1]

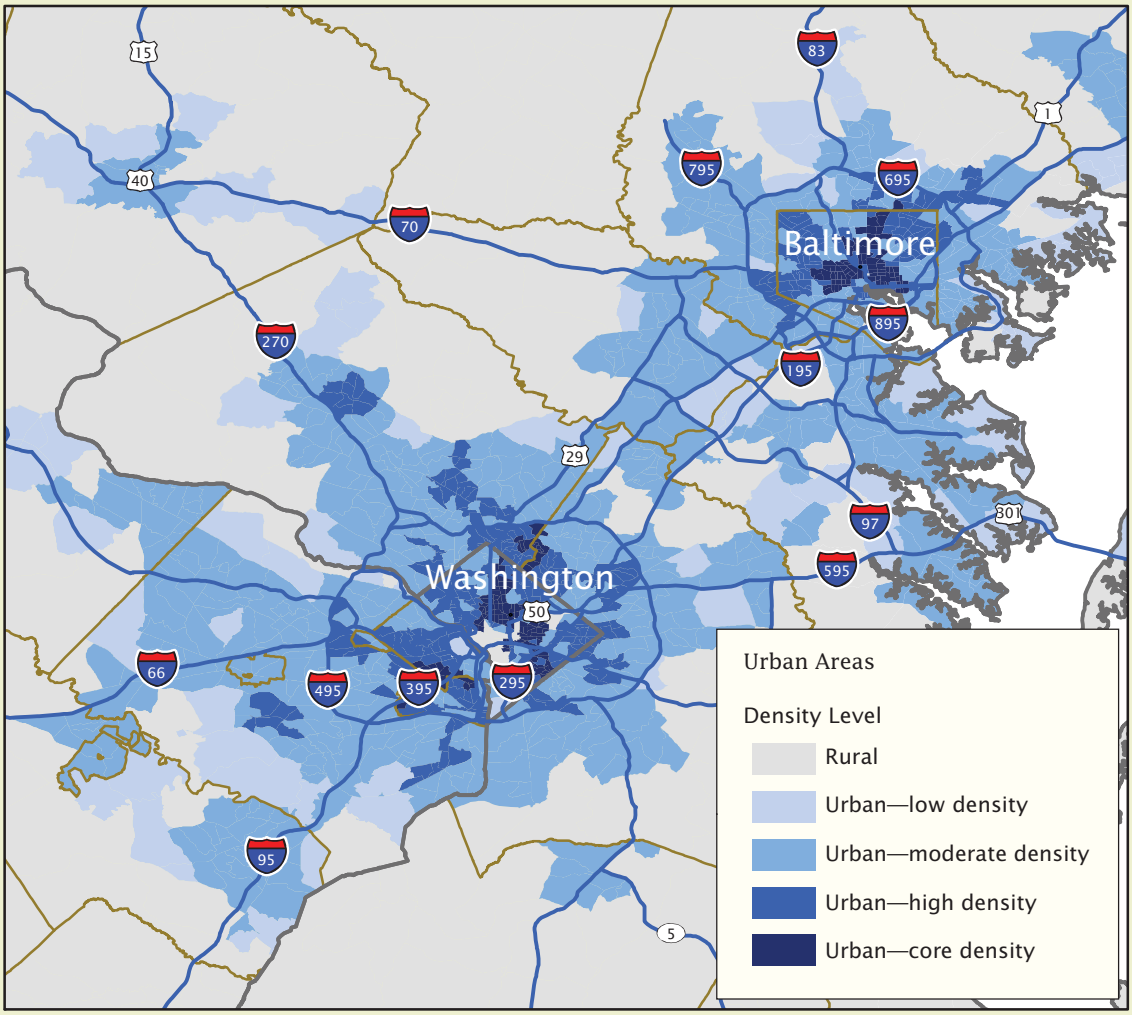
[1] The VirtualRDC lets researchers learn how to work within the constraints of the actual RDCs before starting to work at one of the physical locations, thus reducing set-up time. The VirtualRDC also provides access to synthetic data over the Internet.

EXPANDED RESEARCH OPPORTUNITIES

Many RDC research projects have used the 1990 and 2000 decennial census microdata files. The availability of additional years of decennial census microdata could provide additional historical perspective.

For example, six RDC research projects began in 2008 that

Figure 5-2.
Urban Areas in the Washington-Baltimore-Northern Virginia, Combined Statistical Area



Source: U.S. Census Bureau, Census 2000.

included data from the 1990 and 2000 decennial censuses. These projects are listed in Appendix 2. The research includes such topics as marriage and economic opportunity, income and poverty, migrant flows, the social contexts of the children of immigrants, locational attainment and residential segregation, and neighborhood economic transitions.

PROJECT HISTORY

Preparing the historical decennial census files for release in the RDCs required the foresight and commitment of many individuals and institutions. Al Nucci, formerly of CES, was instrumental in identifying these files on the Census Bureau's original Unisys computer systems and urging that they be preserved. Nucci also began the task of moving the data from the Unisys and transforming them into formats that can be used on modern computers. John Murphy and Eugene Raschlich of the Data Access and Dissemination Systems Office (DADSO) also did a considerable amount of work in this effort. Partnerships with MPRC, particularly Seth Sanders, contributed enormously to the task, and ICPSR contributed resources to create the new documentation for these files.

The Population Division of the Census Bureau contributed significant staff time to this project. Also, the retrieval and conversion of the data files and the creation of the zero obs files and other documentation could not have been accomplished without the assistance of Marie Pees of the Population Division.

ONGOING PROJECTS

Restore 1960 Data

The internal electronic version of the 1960 decennial census sample microdata file is nearly complete, but over the years, some data have been lost. A project to restore the 1960 sample data is currently underway.

The 1960 census enumeration manuscripts from affected areas are being scanned and digitized to restore the missing geographic, household, and person records to this valuable dataset. The restoration of the 1960 sample file is being done in association with MPC and the National Archives and Records Administration under a grant from the National Institutes of Health. MPC is also submitting a proposal to scan and digitize the 1960 decennial census complete-count data, as the electronic data file was not preserved after the 1960 census was tabulated.

Harmonization

Also underway is a project to harmonize all of the internal decennial census microdata. What makes the historical census files so valuable is the potential for carrying out analyses that span a long period. Unfortunately, variables from one census year may have a different set of valid values from another census year. Also, even if the sets of valid values match perfectly, variables from different censuses may have been coded differently. To overcome such compatibility problems, the data must be harmonized. That is, the data must be made as compatible as possible over time with respect to coding systems, formats, and documentation.

Recovering Additional Historical Data

As the Unisys computer is being phased out of use at the Census Bureau, we are making efforts to recover any useful data that have not yet been transferred to current platforms. These data may be lost if we do not act quickly.

Transforming these historical files into usable data products requires substantial resources beyond what is currently available at CES. In recent months, we have been able to recover the internal versions of the 1940 and 1950 PUMS files. These files contain detailed geographic information and the original alphanumeric string entries.

More processing is required to make the files a usable research resource. We need help from the research community to complete these recovery and transformation efforts.

Researchers interested in partnering with CES to recover and develop these data should contact the CES Assistant Division Chief for Research, Lucia Foster at <Lucia.S.Foster@census.gov>.

REFERENCES

Moretti, Enrico. 2004. "Workers' Education, Spillovers, and Productivity: Evidence From Plant-Level Production Functions." *American Economic Review*, 94(3): 656–90.

Chapter 6.
RECOVERING HISTORICAL MANUFACTURING MICRODATA

Randy Becker and Cheryl Grim, Center for Economic Studies

What is past is prologue.

—W. Shakespeare

This famous quote from *The Tempest* describes the motivation for much of the research conducted at the Center for Economic Studies (CES). In analyzing historical microdata, we hope to explain the present-day economy and inform debates about the future.

The oldest microdata CES currently houses are on the manufacturing sector, with plant-level information on output, employment, wages, material inputs, capital expenditures, industry, and location. Data are available for every Census of Manufactures (CM) from 1963 and for every Annual Survey of Manufactures (ASM) from 1973 to the present. These establishment-level data have been longitudinally linked.[1]

[1] These linked data were the first data CES made available to approved researchers in the early 1980s. See Atrostic (2008) for a history of CES.

A Census Bureau employee uses a card punch to transcribe data from a 1935 Business Census questionnaire to punch cards for mechanical sorting and tabulation. The use of punch cards to tabulate census data began following the 1890 census and continued into the 1960s when computers and magnetic computer tape fully replaced punch cards.

U.S. Census Bureau, Public Information Office (PIO).

Stored data files and survey forms on microfilm could be used to extend this longitudinal series back in time—by decades and perhaps by over a century.

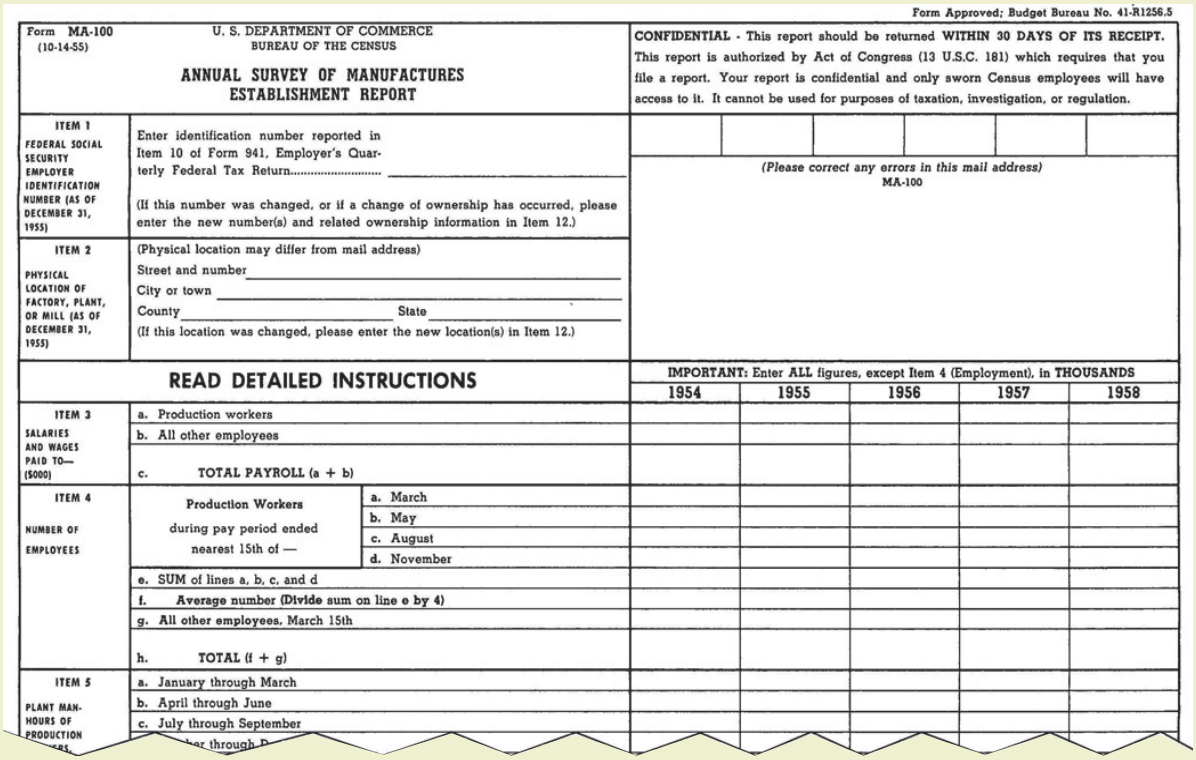
Extending the manufacturing microdata series would provide opportunities to study the evolution and behavior of plants and industries over longer time periods and more business cycles, including the Great Depression and the expansions and recessions of the immediate post-World War II period. CES and the research community recognize the value of adding earlier years to the existing longitudinal series of establishment-level manufacturing data.[2]

We need help from the research community to recover these historical data. The recovery will require substantial resources beyond those currently available at CES. Without quick action, these data may be lost.

DATA FILES STORED BY THE U.S. CENSUS BUREAU

CES has learned that vintage data files for select surveys are stored on the Census Bureau's Unisys system, an outdated and rapidly depreciating data storage system scheduled to be decommissioned in the near future. Data files for the 1955, 1956, and 1965–1971

[2] For example, recommendations from the April 19–20, 2007, meeting of the Census Advisory Committees of Professional Associations include a statement that the committee would be "very interested" in recovered data from the 1958 CM and the 1954–58 ASMs.



Form Approved; Budget Bureau No. 41-R1256.5

Form MA-100 (10-14-55)

U. S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS

ANNUAL SURVEY OF MANUFACTURES
ESTABLISHMENT REPORT

CONFIDENTIAL - This report should be returned **WITHIN 30 DAYS OF ITS RECEIPT.** This report is authorized by Act of Congress (13 U.S.C. 181) which requires that you file a report. Your report is confidential and only sworn Census employees will have access to it. It cannot be used for purposes of taxation, investigation, or regulation.

ITEM 1 — FEDERAL SOCIAL SECURITY EMPLOYER IDENTIFICATION NUMBER (AS OF DECEMBER 31, 1955)

Enter identification number reported in Item 10 of Form 941, Employer's Quarterly Federal Tax Return.....

(If this number was changed, or if a change of ownership has occurred, please enter the new number(s) and related ownership information in Item 12.)

(Please correct any errors in this mail address)
MA-100

ITEM 2 — PHYSICAL LOCATION OF FACTORY, PLANT, OR MILL (AS OF DECEMBER 31, 1955)

(Physical location may differ from mail address)
Street and number
City or town
County State
(If this location was changed, please enter the new location(s) in Item 12.)

READ DETAILED INSTRUCTIONS

IMPORTANT: Enter ALL figures, except Item 4 (Employment), in THOUSANDS

			1954	1955	1956	1957	1958
ITEM 3 — SALARIES AND WAGES PAID TO— ($000)	a. Production workers						
	b. All other employees						
	c. TOTAL PAYROLL (a + b)						
ITEM 4 — NUMBER OF EMPLOYEES	Production Workers during pay period ended nearest 15th of —	a. March					
		b. May					
		c. August					
		d. November					
	e. SUM of lines a, b, c, and d						
	f. Average number (Divide sum on line e by 4)						
	g. All other employees, March 15th						
	h. TOTAL (f + g)						
ITEM 5 — PLANT MAN-HOURS OF PRODUCTION	a. January through March						
	b. April through June						
	c. July through September						

Excerpt from the 1954–1958 ASM form, which was "shuttled" back and forth between the Census Bureau and the surveyed businesses through these years.

ASMs have recently been recovered from the Unisys.

Efforts are now underway to recover additional files that may contain historical business microdata.[3] One example is the microdata from a groundbreaking joint Census/National Science Foundation project to match the Survey of Industrial Research and Development, CMs, and other firm level information from 1957 to 1977 (Griliches, 1986).

Once recovered, the data files need to be turned into research datasets. Record layouts need to be determined, and the data need to be checked for missing records and variables. Much of this work has been done for the recently recovered early ASM data (Ladner, 2008).

We found certain industries are missing from the 1955 and 1956 ASM files, but data for these industries could potentially be filled in using the information in the 1954–1958 ASM microfilm discussed in the next section. Additional research is also required to create longitudinal links between records in these early files and the longitudinal CM and ASM data currently available at CES and the Research Data Centers (RDCs).[4]

SURVEY FORMS ON MICROFILM

CES has in its possession microfilm with images of completed survey forms from the 1958 CM and the 1954–58 ASMs. Since microfilm is vulnerable to deterioration, scanning and electronically storing the microfilm is a

[3] This is in conjunction with efforts to obtain additional historical demographic microdata (see Chapter 5 of this report). Recovering data files from the Unisys system is not a trivial exercise because many of the files are stored in an old, proprietary Census Bureau format that requires detailed knowledge to download and then convert them to ASCII format.

[4] Despite somewhat unreliable plant identification numbers in the 1950s (Kallek 1975), recent preliminary matching exercises indicate it is possible to match at least some of the plants by plant identification number. Remaining plants could potentially be matched using detailed geographic information available from the 1954–58 ASM microfilm.

necessary first step to ensure this information is preserved.

Further, to be useful for research, the scanned microfilm needs to be indexed by key variables (e.g., industry, state). Scanning and indexing alone would produce a useful resource, particularly for researchers who focus on a small set of industries and are willing to key in data from the form images. For confidentiality reasons, data must be keyed at a secure Census Bureau facility, such as an RDC. The full set of microfilm would need to be keyed at the Census Bureau's National Processing Center in Jeffersonville, Indiana.

In addition to the data on microfilm held at CES, the National Archives and Records Administration (NARA) maintains microfilm copies of completed CM survey forms from 1947 to potentially as far back as 1820 (see NARA, 1964; and NARA, 2008).[5] It may also be possible to retrieve some early nonmanufacturing microdata from the Census of Business (NARA, 1964; Nucci, 1998; and NARA, 2008).

Creating research data from the microfilm maintained by NARA first requires locating the microfilm at NARA, assessing its condition, repairing and treating it (if necessary), and obtaining a copy. Then, essentially, the same process of scanning, indexing, and keying is required to create electronic research-ready data.

The value of the microfilm at NARA is demonstrated by its use in previous studies of specific industries. Bresnahan and Raff (1991) created a panel dataset using CM microdata from 1929–35 to study intraindustry heterogeneity in the auto industry during the Great Depression. Using CM microdata from the same years, Bertin, Bresnahan, and Raff (1996) study the blast furnace industry.

[5] Completed survey forms do not appear to be available for all CM years in that span.

NEXT STEPS

CES wants to extend the longitudinal establishment-level manufacturing data back further in time. However, substantial resources beyond those currently available at CES are required. The vulnerabilities of both the microfilm and the Unisys system call for quick action.

Researchers interested in partnering with CES to recover and develop these data should contact the CES Assistant Division Chief for Research, Lucia Foster at <Lucia.S.Foster@census.gov>.

REFERENCES

Atrostic, B.K. 2008. "The Center for Economic Studies (CES) at 25: A Short History." In *2007 Research Report: Center for Economic Studies and Research Data Centers*. <www.ces.census.gov>.

Bertin, Amy L., Timothy F. Bresnahan, and Daniel M.G. Raff. 1996. "Localized Competition and the Aggregation of Plant-Level Increasing Returns: Blast Furnaces, 1929–1935." *Journal of Political Economy*, 104(2): 241–266.

Bresnahan, Timothy F. and Daniel M.G. Raff. 1991. "Intra-Industry Heterogeneity and the Great Depression: The American Motor Vehicles Industry, 1929-1935." *Journal of Economic History*, 51(2): 317–331.

Griliches, Zvi. 1986. "Productivity, R&D, and Basic Research at the Firm Level in the 1970's." *American Economic Review*, 76(1): 141–154.

Kallek, Shirley. 1975. "Potential Applications of Census Bureau Economic Series in Microdata Analysis." Papers and Proceedings of the Eighty-Seventh Annual Meeting of the American Economic Association, *American Economic Review*, 65(2): 257–162.

Ladner, Justin. 2008. "Initial Construction and Analysis of Historical Annual Survey of Manufactures Microdata, 1955–1971." Unpublished Technical Note.

National Archives and Records Administration [NARA]. 2008. *Records of the Bureau of the Census (Record Group 29)*. <www.archives.gov/research/guide-fed-records/groups/029.html> (accessed December 8, 2008).

National Archives and Records Administration [NARA]. 1964. *Preliminary Inventory of the Records of the Bureau of the Census (Record Group 29)*. Washington, DC: General Services Administration.

Nucci, Alfred R. 1998. "The Center for Economic Studies Program to Assemble Economic Census Establishment Information." *Business and Economic History*, 27(1): 249–156.

Appendix 1.
CENTER FOR ECONOMIC STUDIES (CES) STAFF AND RESEARCH DATA CENTER (RDC) PUBLICATIONS AND WORKING PAPERS

[Term inside brackets indicates work by CES staff or the main RDC involved]

PUBLICATIONS

Andersson, Fredrik, Matthew Freedman, John Haltiwanger, Julia Lane, and Kathryn Shaw. Forthcoming. "Reaching for the Stars: Who Pays for Talent in Innovative Industries?" *Economic Journal*. [CES]

Arzaghi, Mohammad and Vernon Henderson. 2008. "Networking Off Madison Avenue." *Review of Economic Studies*, 75(4): 1011–1038. [Boston]

Atrostic, B.K. 2008. "Measuring U.S. Innovative Activity: Business Data at the U.S. Census Bureau." *Journal of Technology Transfer,* 33(2): 153–171. [CES]

Bayer, Patrick, Stephen Ross, and Giorgio Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy*, 116(6): 1150–1196. [Boston]

Becker, Randy A. and Ronald J. Shadbegian. Forthcoming. "Environmental Products Manufacturing: A Look Inside the Green Industry." *Topics in Economic Analysis & Policy*. [CES]

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. 2008. "Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. That Trade Goods." In *Producer Dynamics: New Evidence From Micro Data*, ed. Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts. Chicago, IL: University of Chicago Press. [Boston]

Boyd, Gale. 2008. "Estimating Plant-Level Manufacturing Energy Efficiency With Stochastic Frontier Regression." *Energy Journal*, 29(2): 23–44. [Triangle]

Boyd, Gale, Elizabeth Dutrow, and Walt Tunnessen. 2008. "The Evolution of the Energy Star Energy Performance Indicator for Benchmarking Industrial Plant Manufacturing Energy Use." *Journal of Cleaner Production*, 16(6): 709–715. [Triangle]

Burgess, Simon, Julia Lane, and Kevin McKinney. Forthcoming. "Matching, Reallocation, and Changes in Earnings Dispersion." *Oxford Bulletin of Economics and Statistics*. [CES]

Cappelli, Peter and Monika Hamori. 2008. "Are Franchises Bad Employers?" *Industrial and Labor Relations Review*, 61(2): 147–162. [CES RDC]

Chen, Susan and H. Wilbert van der Klaauw. 2008. "The Work Disincentive Effects of the Disability Insurance Program in the 1990s." *Journal of Econometrics*, 142(2): 757–784. [Chicago]

Culter, David, Edward Glaeser, and Jacob Vigdor. 2008. "Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation." *Review of Economics and Statistics*, 90(3): 389–405. [Triangle]

Davis, James C. and J. Vernon Henderson. 2008. "The Agglomeration of Headquarters." *Regional Science and Urban Economics*, 38(5): 445–460. [CES]

Davis, Steven J., John Haltiwanger, Ron Jarmin, C.J. Krizan, Javier Miranda, Al Nucci, and Kristin Sandusky. 2008. "Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes." In *Producer Dynamics: New Evidence From Micro Data*, ed. Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts. Chicago, IL: University of Chicago Press. [CES]

Eisfeldt, Andrea and Adriano Rampini. Forthcoming. "Leasing, Ability to Repossess, and Debt Capacity." *Review of Financial Studies*. [Chicago]

Fairman, Kristin, Lucia Foster, C.J. Krizan, and Ian Rucker. 2008. "An Analysis of Key Differences in Micro Data: Results From the Business List Comparison Project." In *2008 Proceedings of the American*

Statistical Associations, Business and Economics Statistics Section [CD-ROM], Alexandria, VA: American Statistical Association. [CES]

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1): 394–425. [CES]

Glaeser, Edward and William Kerr. Forthcoming. "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" *Journal of Economics and Management Strategy*. [Boston]

Guo, Xuguang (Steve). 2008. "The Viability of the Employer-Provided Health Insurance System." Ph.D. diss. Rutgers University. [New York]

Hellerstein, Judith, David Neumark, and Melissa McInerney. 2008. "Changes in Workplace Segregation in the United States Between 1990 and 2000: Evidence From Matched Employer-Employee Data." In *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches*, ed. Stefan Bender, Julia Lane, Kathryn Shaw, Fredrik Andersson, and Till von Wachter. Chicago, IL: University of Chicago Press. [CES]

Hellerstein, Judith, David Neumark, and Melissa McInerney. 2008. "Spatial Mismatch or Racial Mismatch?" *Journal of Urban Economics*, 64(2): 464–479. [CES]

Henderson, Vernon and Yukako Ono. 2008. "Where Do Manufacturing Firms Locate Their Headquarters?" *Journal of Urban Economics*, 63(2): 431–450. [Boston]

Ioannides, Yannis M. and Jeffery E. Zabel. 2008. "Interactions, Neighborhood Selection, and Housing Demand." *Journal of Urban Economics*, 63(1): 229–252. [Boston]

Kerr, William and Ramana Nanda. Forthcoming. "Democratizing Entry: Banking Deregulations, Financing Constraints, and Entrepreneurship." *Journal of Financial Economics*. [Boston]

Kerr, William and Shihe Fu. 2008. "The Survey of Industrial R&D-Patent Database Link Project." *Journal of Technology Transfer*, 33(2): 173–186. [Boston]

Linn, Josh. Forthcoming. "Why Do Oil Shocks Matter? The Role of Inter-Industry Linkages." *Economic Inquiry*. [Chicago]

Linn, Josh. 2008. "Energy Prices and the Adoption of Energy Saving Technology." *Economic Journal,* 118(533): 1986–2012. [Chicago]

Lombardi, Britton and Yukako Ono. 2008. "Professional Employer Organizations: What Are They, Who Uses Them, and Why Should We Care?" *Economic Perspectives*, 32: 2–14. [Chicago]

Melzer, Brian. 2008. "The Real Costs of Credit Access: Evidence From the Payday Lending Market." Ph.D. diss. University of Chicago. [Chicago]

Nguyen, Sang V. and Mary L. Streitwieser. 2008. "Energy-Capital Substitution Revisited: Evidence From Micro Data." *Journal of Economic and Social Measurement*, 33(2-3): 129–153. [CES]

Nguyen, Sang V. and Michael Ollinger. Forthcoming. "Mergers and Acquisitions, Employment, Wages, and Plant Closures in the U.S. Meat Product Industries." *Agribusiness*. [CES]

Ollinger, Michael and Danna Moore. Forthcoming. "The Direct and Indirect Costs of Food Safety Regulation." *Review of Agricultural Economics*. [CES RDC]

Ollinger, Michael and Danna Moore. 2008. "The Economic Forces Driving Food Safety Quality in Meat and Poultry." *Review of Agricultural Economics*, 30(2): 289–310. [CES RDC]

Steffenson, Kristina. 2008. "Productivity Leadership and Strategic Investments in Innovation: The Adoption of E-Business Capabilities." Best Paper, Academy of Management Proceedings. [Chicago]

Valdez, Zulema. 2008. "The Effect of Social Capital on White, Korean, Mexican, and Black Business Owners' Earnings in the U.S." *Journal of Ethnic and Migration Studies*, 34(6): 955–973. [Michigan]

Woodcock, Simon. 2008. "Wage Differentials in the Presence of Unobserved Worker, Firm, and Match Heterogeneity." *Labour Economics*, 15(4): 771–793. [Cornell]

WORKING PAPERS

Angrist, Joshua, Stacey Chen, and Brigham Frandsen. 2008. "The Effects of Vietnam-Era Conscription on Mortality and Disability: Draft-Lottery Estimates From the 2000 Census." Unpublished. [Boston]

Arzaghi, Mohammad, Ernst Berndt, James Davis, and Alvin Silk. 2008. "Economic Factors Underlying the Unbundling of Advertising Agency Services." National Bureau of Economic Research Working Papers 14345. [Boston]

Atrostic, B.K., Kazuyuki Motohashi, and Sang V. Nguyen. 2008. "Computer Network Use and Firms' Productivity Performance: The United States vs. Japan." Center for Economic Studies Discussion Paper CES-08-30. [CES]

Basker, Emek, Shawn D. Klimek, and Pham Hoang Van. 2008. "Supersize It: The Growth of Retail Chains and the Rise of the 'Big Box' Retail Format." University of Missouri-Columbia Department of Economics Working Paper 08-09. [CES]

Basker, Emek, Shawn D. Klimek, and Pham Hoang Van. 2008. "Supersize It: The Growth of Retail Chains and the Rise of the 'Big Box' Retail Format." Center for Economic Studies Discussion Paper CES-08-23. [CES]

Bates, Timothy and Alicia Robb. 2008. "Analysis of Young Neighborhood Firms Serving Minority Clients." Center for Economic Studies Discussion Paper CES-08-11. [Berkeley]

Becker, Randy A. and Ronald J. Shadbegian. 2008. "The Green Industry: An Examination of Environmental Products Manufacturing." Center for Economic Studies Discussion Paper CES-08-34. [CES]

Becker, Randy A. and Ronald J. Shadbegian. 2008. "The Green Industry: An Examination of Environmental Products Manufacturing." National Center for Environmental Economics Working Paper 08-10. [CES]

Benedetto, Gary and Simon Woodcock. 2008. "Partial Synthesis of the Longitudinal Employer-Household Dynamics (LEHD) Database." Unpublished. [Cornell]

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. 2008. "Transfer Pricing by U.S.-Based Multinational Firms." Center for Economic Studies Discussion Paper CES-08-29. [Boston]

Bernard, Andrew B., J. Bradford Jensen, Stephen Redding, and Peter K. Schott. 2008. "Intra-Firm Trade and Product Contractibility." Unpublished. [Boston]

Bernard, Andrew B., J. Bradford Jensen, Stephen Redding, and Peter K. Schott. 2008. "The Margins of U.S. Trade." Unpublished. [Boston]

Bernard, Andrew B., Stephen Redding, and Peter K. Schott. 2008. "Multi-Product Firms and Product Switching." Center for Economic Studies Discussion Paper CES-08-24. [Boston]

Bernard, Andrew B., Stephen Redding, and Peter K. Schott. 2008. "Products and Productivity." Center for Economic Studies Discussion Paper CES-08-22. [Boston]

Bernard, Andrew B., Stephen Redding, and Peter K. Schott. 2008. "Multi-Product Firms and Trade Liberalization." Unpublished. [Boston]

Bjelland, Melissa, Bruce Fallick, John Haltiwanger, and Erika McEntarfer. 2008. "Employer-to-Employer Flows in the United States: Estimates Using Linked Employer-Employee Data." National Bureau of Economic Research Working Paper 13867. [CES]

Burkhauser, Richard and Jeff Larrimore. 2008. "Trends in the Relative Household Income of Working-Age Men With Work Limitations: Correcting the Record Using Internal Current Population Survey Data." Center for Economic Studies Discussion Paper CES-08-05. [Cornell]

Burkhauser, Richard and Jeff Larrimore. 2008. "Using Internal Current Population Survey Data to Reevaluate Trends in Labor Earnings Gaps by Gender, Race, and Education Level." Center for

Economic Studies Discussion Paper CES-08-18. [Cornell]

Burkhauser, Richard, Jeff Larrimore, and Shuaizhang Feng. 2008. "Measuring Labor Earnings Inequality Using Public-Use March Current Population Survey Data: The Value of Including Variances and Cell Means When Imputing Topcoded Values." Center for Economic Studies Discussion Paper CES-08-38. [Cornell]

Burkhauser, Richard, Shuaizhang Feng, Stephen Jenkins, and Jeff Larrimore. 2008. "Estimating Trends in U.S. Income Inequality Using the Current Population Survey: The Importance of Controlling for Censoring." Center for Economic Studies Discussion Paper CES-08-25. [Cornell]

Celikkol Geylani, Pinar and Spiro Stefanou. 2008. "Linking Investment Spikes and Productivity Growth: U.S. Food Manufacturing Industry." Center for Economic Studies Discussion Paper CES-08-36. [Michigan]

Chemmanur, Thomas, Karthik Krishnan, and Debarshi Nandy. 2008. "How Does Venture Capital Financing Improve Efficiency in Private Firms? A Look Beneath the Surface." Center for Economic Studies Discussion Paper CES-08-16. [Boston]

Chemmanur, Thomas, Shan He, and Debarshi Nandy. 2008. "The Going Public Decision and the Product Market." Center for Economic Studies Discussion Paper CES-08-20. [Boston]

Chen, Susan. 2008. "The Effect of Disability Insurance on the Labor Supply Decisions of Married Couples." Unpublished. [Chicago]

Davis, Lucas. 2008. "The Effect of Power Plants on Local Housing Values and Rents: Evidence From Restricted Census Microdata." Center for Economic Studies Discussion Paper CES-08-19. [Michigan]

Davis, Steven J., Cheryl Grim, and John Haltiwanger. 2008. "Productivity Dispersion and Input Prices: The Case of Electricity." Center for Economic Studies Discussion Paper CES-08-33. [CES]

Davis, Steven J., Cheryl Grim, John Haltiwanger, and Mary L. Streitwieser. 2008. "Electricity Pricing to U.S. Manufacturing Plants: 1963–2000." National Bureau of Economic Research Working Paper 13778. [CES]

Davis, Steven J., Jason Faberman, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2008. "Business Volatility, Job Destruction, and Unemployment." Center for Economic Studies Discussion Paper CES-08-26. [CES]

Davis, Steven J., Jason Faberman, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2008. "Business Volatility, Job Destruction, and Unemployment." National Bureau of Economic Research Working Paper 14300. [CES]

Davis, Steven J., John Haltiwanger, and Ron Jarmin. 2008. "Turmoil and Growth: Young Businesses, Economic Churning, and Productivity Gains." Ewing Marion Kauffman Foundation Research Report. [CES]

Davis, Steven J., John Haltiwanger, Ron Jarmin, Josh Lerner, and Javier Miranda. 2008. "Private Equity and Employment." Center for Economic Studies Discussion Paper CES-08-07. [CES]

Eisenberg, Daniel and Brian Rowe. 2008. "The Effects of Smoking in Young Adulthood on Smoking and Health Later in Life: Evidence Based on the Vietnam Era Draft Lottery." CES Working Paper CES-08-35. [Michigan]

Fairman, Kristin, Lucia Foster, C.J. Krizan, and Ian Rucker. 2008. "An Analysis of Key Differences in Micro Data: Results From the Business List Comparison Project." Center for Economic Studies Discussion Paper CES-08-28. [CES]

Ferreira, Fernando. 2008. "You Can Take It With You: Proposition 13 Tax Benefits, Residential Mobility, and Marginal Willingness to Pay for Housing Amenities." Center for Economic Studies Discussion Paper CES-08-15. [Berkeley]

Fitzpatrick, Maria. 2008. "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Pre-Kindergarten." Center for Economic Studies Discussion Paper CES-08-04. [Triangle]

Frandsen, Brigham. 2008. "Long Term Effects of Military Service on the Distribution of Earnings." Unpublished. [Boston]

Frey, William and Kao-Lee Liaw. 2008. "Choices of Metropolitan Destinations by the 1995–2000 New Immigrants Born in Mexico and India: Characterization and Multivariate Explanation." Center for Economic Studies Discussion Paper CES-08-27. [Michigan]

Glaeser, Edward and William Kerr. 2008. "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" Center for Economic Studies Discussion Paper CES-08-37. [Boston]

Glaeser, Edward and William Kerr. 2008. "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" Harvard Business School Working Paper 09-055. [Boston]

Glaeser, Edward and William Kerr. 2008. "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" National Bureau of Economic Research Working Paper 14407. [Boston]

Gurak, Douglas, Min-Ah Lee, and Mary Kritz. 2008. "Internal Migration of New Immigrant Groups to New Destinations." Unpublished. [Cornell]

Hellerstein, Judith, Melissa McInerney, and David Neumark. 2008. "Neighbors and Co-Workers: The Importance of Residential Labor Market Networks." National Bureau of Economic Research Working Paper 14201. [CES]

Hellerstein, Judith, Melissa McInerney, and David Neumark. 2008. "Spatial Influences on the Employment of U.S. Hispanics: Spatial Mismatch, Discrimination, or Immigrant Networks?" Unpublished. [CES]

Hortacsu, Ali and Chad Syverson. 2008. "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." Center for Economic Studies Discussion Paper CES-08-41. [Chicago]

Hortacsu, Ali and Chad Syverson. 2008. "Why Do Firms Own Production Chains?" Unpublished. [Chicago]

Hyson, Rosemary and Alice Zawacki. 2008. "Health-Related Research Using Confidential U.S. Census Bureau Data." Center for Economic Studies Discussion Paper CES-08-21. [CES]

Larrimore, Jeff, Richard Burkhauser, Shuaizhang Feng, and Laura Zayatz. 2008. "Consistent Cell Means for Topcoded Incomes in the Public Use March CPS (1976–2007)." Center for Economic Studies Discussion Paper 08-06. [Cornell]

Lee, Yoonsoo and Toshihiko Mukoyama. 2008. "Entry, Exit, and Plant-Level Dynamics Over the Business Cycle." Center for Economic Studies Discussion Paper CES-08-17. [Michigan]

McCue, Kristin. 2008. "A Comparison of Employee Benefits Data From the MEPS-IC and Form 5500." Center for Economic Studies Discussion Paper CES-08-32. [CES]

McKinnish, Terra, Randall Walsh, and Kirk White. 2008. "Who Gentrifies Low Income Neighborhoods?" National Bureau of Economic Research Working Paper 14036. [Triangle]

Morgenstern, Richard D., William A. Pizer, and Jhih-Shyang Shih. 2008. "Evaluting Voluntary Climate Programs in the United States." Resources for the Future (RFF) Working Paper RFF DP 08-13. [CES RDC]

Ono, Yukako and Daniel Sullivan. 2008. "Manufacturing Plants' Use of Temporary Workers: An Analysis Using Census Micro Data." Center for Economic Studies Discussion Paper CES-08-40. [Chicago]

Osili, Una Okonkwo and Anna Paulson. 2008. "Bank Crises and Investor Confidence." Federal Reserve Bank of Chicago Working Paper 2008-17. [Chicago]

Patterson, Rhiannon, Alicia Robb, and Timothy Bates. 2008. "Neighborhood Effects on High-School Drop-Out Rates and Teenage Childbearing: Tests for Non-Linearities, Race-Specific Effects, Interactions With

Family Characteristics, and Endogenous Causation Using Geocoded California Census Microdata." Center for Economic Studies Discussion Paper CES-08-12. [Berkeley]

Pierce, Justin. 2008. "Plant-Level Responses to Antidumping Duties: Evidence From U.S. Manufacturers." Unpublished. [Boston]

Pierce, Justin and Peter K. Schott. 2008. "Concording U.S. Harmonized System Categories Over Time." Unpublished. [Boston]

Pierce, Justin and Peter K. Schott. 2008. "Concording U.S. Harmonized System Categories to U.S. SIC and NAICS Industries." Unpublished. [Boston]

Puri, Manju and Rebecca Zarutskie. 2008. "On the Lifecycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms." Center for Economic Studies Discussion Paper CES-08-13. [Triangle]

Saldanha, Terence. 2008. "The Impact of Electronic Integration on Labor Productivity of U.S. Manufacturing Plants: Differential Impacts of Technology and Information Sharing." Unpublished. [Michigan]

Schultz, Jennifer and David Doorn. 2008. "Rising Health Care Costs: Effects on Labor Demand and Retiree Health Insurance Benefits." Unpublished. [Chicago]

Silk, Alvin and Charles King III. 2008. "Concentration Levels in the U.S. Advertising and Marketing Services Industry: Myth vs. Reality." Harvard Business School Working Paper 09-044. [Boston]

Steffenson, Kristina. 2008. "Delegation in Multi-Divisional Firms: The Organizational Structure of IT Systems Purchasing Authority." Unpublished. [Chicago]

Steffenson, Kristina. 2008. "The Prevalence of Discretion: Impacts of Firm-Level Covariates and Aggregation on the Organizational Structure of IT Systems Purchasing Authority." Unpublished. [Chicago]

Steffenson, Kristina. 2008. "Productivity Leadership and Strategic Investment in Innovation: The Adoption of E-Business Capabilities." Unpublished. [Chicago]

Tang, John P. 2008. "Financial Intermediation and Late Development: The Case of Meiji Japan, 1868–1912." Center for Economic Studies Discussion Paper 08-01. [CES]

Appendix 2.
ABSTRACTS OF PROJECTS STARTED IN 2008

THE EFFECTS OF THE TAX TREATMENT OF HEALTH INSURANCE AND OF HEALTH POLICY ON EXPENDITURES AND SERVICE USE

Sharon Glied—Columbia University

This project examines the impact of the tax treatment of health insurance on health coverage and spending patterns. The second component is an assessment of the effect of health insurance coverage, and policies that affect health insurance coverage, on measures of the quality of care based on the Commonwealth Fund state scorecard.

AN EXAMINATION OF ADOPTION AND UTILIZATION OF ELECTRONIC COMMERCE TECHNOLOGIES WITHIN U.S. FIRMS

Sathasivam Mathiyalakan—Winston-Salem State University
Raja Velu—Syracuse University

This project examines the impact of the tax treatment of health insurance on health coverage and spending patterns. The second component is an assessment of the effect of health insurance coverage, and policies that affect health insurance coverage, on measures of the quality of care based on the Commonwealth Fund state scorecard.

The Internet and its underlying technology have enabled businesses to redesign their processes to take advantage of the capabilities of the Internet as well as to create new ways of communicating and coordinating diverse activities. In addition, both the intranet and extranet have offered ample means for an enterprise to create (or add) value. This has resulted in new and unique challenges for data collection. This project investigates measurement issues related to electronic commerce and produces estimates of the effects of e-commerce on productivity. The project will also examine the impact of computer networks on business processes and on various elements of the supply chain. The study will examine the types of enterprises that use Electronic Commerce Technology (ECT), examine how ECT affects e-supply chain and transforms production, determine if and how the use of ECT results in value addition (or creation), and determine the adequacy of the measures used by the U.S. Census Bureau.

The research will generate estimates of the effect of e-commerce on productivity in manufacturing plants. The investigation will determine whether plants that invested in e-commerce technology received returns in the form of added productivity gains. The researchers' investigation using Census Bureau data can lead to a detailed examination of whether the data are consistent with one of these hypotheses and identify sources of potential e-commerce measurement problems. They will investigate in an e-commerce framework the applicability and robustness of traditional production function estimation, issues related to incorporation of knowledge, manufacturing processes, and quality improvement into the production function, and issues related to industry type and geographic concentration of industries.

IMPORTS AND PRODUCTIVITY IN U.S. MANUFACTURING

James Harrigan—Federal Reserve Bank of New York
Victor Shlychkov—Columbia University

The measurement of foreign trade flows—imports and exports—is an important component of the information the Census Bureau collects on the overall U.S. economy. The research will bring data from these two programs together to provide more extensive characterization of foreign trade activity in U.S. manufacturing firms and variation in firm productivity. The project will prepare estimates of the relationship between productivity and how firms organize production with a particular focus on the role of trade and foreign sourcing in the supply chain. This project will also produce estimates of the relationship between exports, imports, and growth in productivity and firm size. The project will contain analyses that serve to assess the quality of data collected on the supply chain in the 2002 Economic Census.

The project will address whether productivity level determines the organizational form of the firm and whether importing leads to productivity growth in U.S. manufacturing or if higher productivity leads to imports. By estimating models of productivity growth that take the impact of both import and export behavior into account, this research will provide new, enhanced estimates of the relationship between exports, imports, and productivity and how such relationships might vary with firm characteristics and industry. Estimates will also be produced of the relationship between firm growth (employment and revenues), imports, and exports. Accurate estimation of such relationships is an important tool to be used in the consideration of how trade liberalization will impact businesses and the economy. The results will also provide empirical evidence regarding the hypothesis that imports play a relatively insignificant role in productivity growth.

CAUSES AND CONSEQUENCES OF NEIGHBORHOOD ECONOMIC TRANSITIONS

Ingrid Gould Ellen—New York University
Keren Mertens—New York University
Katherine O'Regan—New York University

This project aims at improving understanding of neighborhood economic change by studying household residential choices and examining the circumstances under which households are willing to make moves into neighborhoods with incomes lower than their own. It will study whether and how these households differ from households making other types of residential choices and whether certain aspects of neighborhoods or their residents make such pioneering moves more attractive. The project will also investigate household exit decisions and examine whether mobility rates are higher in economically gaining neighborhoods. We will consider renters and low-income renters separately. The research will study other changes taking place in neighborhoods experiencing gains in income—such as whether housing costs increase and whether there are compensating changes in quality of the neighborhood, like increased satisfaction with neighborhood safety, schools, local transit, and availability of local businesses. The research will compare mobility rates and the prevalence of certain types of household turnover across different types of neighborhoods. A series of regression analyses will model pioneering moves, household exit, and various measures of household satisfaction by using the national and five metropolitan area versions of the American Housing Survey.

EMPLOYMENT AND TANF OUTCOMES FOR LOW-INCOME FAMILIES RECEIVING CHILD CARE SUBSIDIES IN ILLINOIS, MARYLAND, AND MINNESOTA: PHASES II AND III

Robert George—Chapin Hall Center for Children
Allison Harris—Chapin Hall Center for Children

This project and another currently underway analyze the child care subsidy (CCS) take-up decision and a range of employment and welfare outcomes among all low-income families in several states. The research seeks to improve the Census Bureau's understanding of who uses the child-care subsidy and how the subsidy aids different groups of low-income families in their quest for economic independence. The groups under study are those who currently receive cash assistance through the Temporary Assistance for Needy Families (TANF) program, those who have recently left TANF, and those who have had no recent contact with the TANF program. This project uses data from Illinois, Maryland, and Minnesota. The primary dataset to be used for this analysis, the Social Services Analysis File (SSAF), is an output of an internal Census Bureau project. The primary research questions are: (1) What are the factors related to child-care subsidy use among low-income working families who are eligible for the CCS through employment or training, and (2) what is the relationship between subsidy use and employment and welfare outcomes.

EMPLOYMENT, AGGLOMERATION, AND THE SPATIAL SORTING OF HOUSEHOLDS AND FIRMS

Patrick Bayer—Duke University
Stephen Ross—University of Connecticut
Giorgio Topa—Federal Reserve Bank of New York

This research will estimate an empirical model that describes the decisions made by both households and firms within large, urban labor markets. The projects' three main tasks are to estimate a model that describes household residential location choice and the employment status of household members, estimate a model describing the establishment location and employment decisions of firms, and conduct general equilibrium simulations using the estimated parameters from the two models. Analysis of the household problem should provide new insights into the effect of residential location on employment outcomes. The analysis of the firm problem is intended to examine the extent and nature of agglomeration economies within metropolitan areas, and the simulations that use the estimates from both models allow for an assessment of the general equilibrium effect of changes in economic conditions on the patterns of individual and firm choices.

MEASURING INCOME AND POVERTY FROM A MULTIYEAR PERSPECTIVE

Jeffrey B. Liebman—Harvard University

This project is a next step in research on the value of supplementing official Census Bureau measures of poverty, income, and the income distribution with measures based on multiple years of income, potentially up to an individual's entire lifetime. The aim of this research is to produce a comprehensive analysis of the impact of government tax and transfer programs on the lifetime income distribution, incorporating not only the components studied in earlier work but also additional components, such as the Temporary Assistance for Needy Families program and supplemental security income. The project will analyze how a multiyear approach alters measures of poverty among the elderly. The data for this project consist of various Survey of Income and Program Participation (SIPP) panels matched to administrative data from the Social Security Administration. The primary benefit to the Census Bureau

of this project is the additional information it will provide on the value of incorporating multiyear measures of income or poverty into its data collection and data analysis strategy. The second benefit is that, by combining administrative and survey measures of similar concepts, it will help to evaluate the quality of data collected by the SIPP. The third benefit to the Census Bureau involves improving the quality of imputation procedures for item nonresponse in the SIPP. A final benefit to the Census Bureau is that it will provide information helpful in analyzing methods for adding a "potential Primary Insurance Amount (PIA)" variable to the public-use version of the SIPP.

THE EFFECTS OF MATERNAL LABOR SUPPLY ON CHILD HEALTH

Melinda Morrill—North Carolina State University

This project uses a restricted version of the National Health Interview Survey (NHIS) from the National Center for Health Statistics. Over the past several decades, women have been steadily increasing their labor supply. The economic impact of this trend cannot be completely characterized without an understanding of the externalities involved. Children's health is considered important not just because of the contemporaneous costs and burden, but also because child health is strongly linked to adult health and well-being. Maternal labor supply is generally thought of as a consequence of, rather than the cause of, child health. It has been shown that mothers reduce their labor supply when their children are unhealthy or disabled, which induces a positive correlation between maternal labor supply and child health. By using variation in a mother's youngest child's eligibility for kindergarten, this research will be able to provide a causal estimate of the effect of maternal labor supply on child health. The success of this research relies on access to the restricted NHIS, which contains unique data allowing a child's kindergarten eligibility to be precisely calculated.

MIGRANT LIFE HISTORY PROJECT ANALYSIS AND REPORT

Maritsa V. Poros—City College of New York, CUNY

This research is designed to evaluate and improve the quality of existing nativity questions on Census Bureau surveys. External data provide new information on the characteristics and patterns of migration that the Census Bureau's migration data are not able to capture at present. The research design involves an analysis of qualitative data on the implications for producing intercensal demographic estimates of the population. The data consist of approximately 300 unstructured interviews with adult immigrants (aged 18 or over) who were born in any of 12 sending countries and who have lived in the United States for at least 3 months. The sending countries represent top source countries of recent migration (1995–2000) and/or very diverse types of migration flows and experiences. The primary purpose of the project was to collect detailed data on what migrant flows look like and to examine the demographic and other characteristics and experiences associated with different types of flows. Data include the socioeconomic background of migrants and their education, migration, work, and health histories. These data will be used to address the importance of social networks for international migration, occupational attainment, and residence (including internal migration and changes in household composition). These analyses will identify the limitations of and gaps in existing data that are currently used for intercensal demographic estimates by providing the first systematic evaluation of migration questions since they were introduced. Second, they will provide a basis for proposing revisions to survey content, which can improve those estimates, and, in general, improve the quality of census survey data on the foreign born. Third, the results will address issues regarding the economic, political, and social impact of migrants on American society.

THE SOCIAL CONTEXTS OF THE CHILDREN OF IMMIGRANTS IN THE U.S.

Richard Alba—University at Albany, SUNY

This project examines the social contexts in which children of immigrant parents are being reared in the United States and identifies the metropolitan, neighborhood, family/household, and parental characteristics associated with these contexts. Theoretically, it is guided by the conceptualization in the literature of three distinct patterns of incorporation: conventional assimilation, downward assimilation (or racialized incorporation), and the pluralist alternative. Since no one of these patterns is adequate for a full understanding of the incorporation of contemporary immigrant groups, the major goal is to identify more precisely the circumstances under which each pattern comes into play. The project addresses in particular the factors that lead the children of immigrants to grow up in socially advantaged or disadvantaged households and neighborhoods. Since limited economic resources and residence in poverty-stricken neighborhoods have implications for access to health care and exposure to health risks, the findings will provide improved understanding of important determinants of health outcomes. Multilevel models of household and neighborhood outcomes will be estimated for 36 different immigrant groups, as well as for the children in native families of all the major race groups and Hispanics. These models will include parental, household, and metropolitan characteristics as independent variables and will allow the examination of effects specific to kinds of metropolitan regions through interactions between metropolitan-level variables and lower-level ones.

THE IMPACT OF SPIN-OUT GENERATION ON PARENT AND PROGENY FIRMS

Rajshree Agarwal—University of Illinois at Urbana-Champaign
Benjamin A. Campbell—The Ohio State University
April Franco—University of Iowa

Spin-outs are distinct entrepreneurial ventures where founding team members are ex-employees of an incumbent firm. In many industries, spin-outs (progeny firms) compete directly with the incumbents (parent firms) and are an important source of innovation and growth. Previous research has emphasized the beneficial performance implications of the knowledge diffusion from the parent to the progeny and have found that the inherited knowledge provides spin-out firms both superior capabilities and performance vis-à-vis other entrants. However, little is known about the competitive effects of these knowledge flows on the subsequent capabilities and performance of parent firms. This project will investigate the processes that generate firm births and deaths in general but also those that generate spin-off births in particular. The research will identify new firms, spin-off firms, and the parents of these new spin-off firms in the legal services industry. It will analyze the characteristics of these new firms as well as the parents of spin-off firms. This process is especially interesting in the services sector where barriers to entry are generally low and firm birth and death rates are very high. As a result of this high turnover, it has traditionally been difficult to accurately measure new firm generation, the characteristics of new firms, and the performance of new firms in the services sector. However, with a custom Longitudinal Employer-Household Dynamics (LEHD) extract in conjunction with the external Martindale-Hubbell directory, this project will be able to accurately portray characteristics of new and dying firms as well as to estimate birth and death rates in this sector. Moreover, this project will develop and test new methodologies of spin-off generation and knowledge as well as examine the quality of the LEHD data in comparison to the Martindale-Hubbell data.

THE IMPORTANCE OF OBJECTIVE HEALTH MEASURES IN PREDICTING EXIT FROM THE LABOR FORCE VIA EARLY OA, DI, AND SSI PROGRAMS: THE CASE OF FATNESS

Richard V. Burkhauser—Cornell University
John C. Cawley—Cornall University
Maximilian D. Schmeiser—Cornell University

Theoretical models argue that poor health will contribute to early exit from the labor market and the decision to take early social security retirement benefits (Old-Age or OA), Social Security Disability Insurance (SSDI), or Supplemental Security Income-adult disability (SSI) benefits. However, most estimates of the importance of health on the decision to take such benefits have been forced to rely on subjective measures of health, such as self-rated work limitations or self-rated health. Using a special data extract from the National Center for Health Statistics, this project contributes to the empirical literature by using a more objective measure of health, fatness, to predict application for and receipt of such benefits. The research will also determine which of several measures of fatness—body mass index, total body fat, percent body fat, waist circumference, and waist-to-hip ratio is the best predictor of participation in OA, SSDI, or SSI.

MULTIPLE IMPUTATION AND ESTIMATING AGGREGATE PRODUCTIVITY GROWTH IN MANUFACTURING

Jerome Reiter—Duke University
Amil Petrin—The University of Chicago
Kirk White—U.S. Department of Agriculture, Economic Research Service

Economists have noted that the quality of imputed data is a problem for researchers using plant-level Annual Survey of Manufactures (ASM) and Census of Manufactures data. Using the detailed item impute flags in the later years of the ASM, this project proposes to develop a model of missing data by applying methods of multiple imputation to improve imputations for nonresponse in these data. The research will measure improvement in data quality by analyzing how both aggregate and plant-level productivity, as well as other measures, are different with multiply imputed data versus the methods currently in use by the Census Bureau.

PROSPERITY AND THE CLUSTER COMPOSITION OF REGIONS: LINKAGES BETWEEN RURAL AND URBAN AREAS

Richard Bryden—Harvard Business School
Mercedes Delgado—Harvard Business School
Christian Ketels—Harvard Business School
Michael E. Porter—Harvard Business School
Scott Stern—Northwestern University

This research project employs data that will enable a better understanding of the impact a region's economic composition has on its economic performance, the impact of differences in the geographic profile of regions on this relationship, and the impact of the cluster composition and performance of neighboring regions on a region's economic performance. Previous work on the relationship between cluster composition and the economic performance of regions and their clusters has revealed data limitations, in particular the significant level of data suppression of industry-level data in smaller, largely rural regions. Using the economic census and the Longitudinal Business Database, this project will investigate the ability of a cluster concept of colocated industries (those linked through their supply chain and other

agglomeration forces) to provide an opportunity to publish regional aggregations with less need for disclosure suppressions. The project will further investigate the definitions of clusters and their development over time. The research will generate estimates of the economic performance of a region and its clusters. It aims at a more sophisticated understanding of linkages across clusters within a region and of linkages between regions, and to develop typologies of regional economies by cluster profiles and other dimensions.

INTERFIRM JOB MOBILITY, LOCAL LABOR MARKETS, AND ORGANIZATIONAL DYNAMICS IN RETAIL

Troy Blanchard—Louisiana State University
Michael Irwin—Duquesne University
Charles Tolbert—Baylor University

American public opinion reflects a growing concern about employment stability and a weakening of the bond between employer and employee. Less stability means more frequent interfirm job changes that can result in interruptions in company benefits (especially health insurance), decreases or stagnation in wages, and involuntary entry into contingent or temporary employment relations. This project examines interfirm mobility by introducing locality into an analysis of the retail industry. Job mobility is best gauged with measures of the dynamic characteristics of firms as they function in their local context. At the organizational level, workers in locally oriented retail business establishments exhibit different mobility patterns than their counterparts employed in non-locally oriented firms. This research seeks to analyze the influence of local organizational and labor market dynamics on interfirm mobility of workers, examine the mobility of workers in locally oriented and non-locally oriented firms with special attention to times of economic expansion and recession, and explore the direction of interfirm mobility in terms of earnings stability, appreciation, or depreciation over time.

MARRIAGE AND ECONOMIC OPPORTUNITY IN THE UNITED STATES

Catherine Fitch—University of Minnesota
J. Michael Oakes—University of Minnesota
Steven Ruggles—University of Minnesota

This research will assess the impact of changes in male and female economic opportunity on marriage formation in the United States since 1960. The working hypothesis is that rising female opportunity discouraged marriage, but that effect diminished over time; after 1970 the decline in the supply of young men with good jobs was the key factor behind late marriage. Previous efforts to address these questions were stymied by inadequate geographic precision in available public-use microdata and by insufficient sample sizes, especially for racial minorities. The decennial long-form census microdata present an opportunity to deepen understanding of this extraordinary demographic transition.

MEASUREMENT ERROR IN EXPECTED AND ACTUAL SSI BENEFITS: CAUSES AND CONSEQUENCES

Todd Elder—Michigan State University
Elizabeth Powers—University of Illinois

This research will measure the extent of elderly poverty by using information contained in the Survey of Income and Program Participation (SIPP) data linked to Supplemental Security Record files. The project uses supplemental security income (SSI) rules and administrative data on benefits and SSI eligibility to construct measures of earnings and eligibility for SIPP

sample members and contrast these measures to those obtained from public-use data. Preliminary estimates suggest that measurement error in income and assets in the public-use SIPP is substantial among the low-resource elderly. Moreover, this measurement error translates directly into errors in estimates of the size of the SSI-eligible population and the rate at which eligible individuals enroll in SSI. The potential benefits of this study are to increase awareness of the quality of SIPP data, increase understanding of the sources of measurement errors in financial resources, and prepare estimates of the characteristics of the population that are likely to be more accurate than those obtained from previous research using public-use data.

THE EFFECTS OF VIETNAM DRAFT LOTTERY STATUS ON LATER LIFE OUTCOMES

Dalton Conley—New York University
Jennifer Heerwig—New York University

The goal of our proposed project is to examine important economic, family, health, and residential outcomes in the Vietnam-era service cohort through an instrumental variable estimation. By using this statistical technique, we hope to provide the Census Bureau with estimates of the effect of military service on the veteran population purged of selection bias. These estimates will demonstrate how Vietnam-era males, chosen randomly through the draft lottery, compare with the nonservice population on variables such as income, wealth accumulation, marital stability, and residential mobility. The analysis will provide important information about the later life characteristics of the male Vietnam-era population while highlighting the needs of aging Vietnam veterans.

AN EMPIRICAL STUDY OF ENTRY AND EXIT WITH MULTIPLE FIRM TYPES

Timothy Dunne—Federal Reserve Bank of Cleveland
Shawn Klimek—U.S. Census Bureau
Mark Roberts—Pennsylvania State University
Yi Xu—New York University

Until recently, progress in developing dynamic empirical models of entry, exit, and firm investment and growth has been slowed by methodological difficulties and a lack of coordination between research teams. Now that theoretical research has provided a number of alternative estimation approaches, coordinated access to longitudinal microdata on firms or plants for a wide range of industries is needed to make progress on estimating the underlying structural parameters across a variety of business sectors and geographic markets. The relationship between the size and the competitiveness of markets is an area of longstanding interest in industrial organization. Recent empirical studies have used the relationship between the size of a geographic market, often measured as market population, and both the number of firms in the market and the average sales of the firms to indirectly draw inferences about the degree of competition in the market. A second line of inquiry has relied on dynamic models that make predictions about how impediments to new firm entry, such as the magnitude of sunk entry costs, affect patterns of firm turnover in order to infer the extent of competitive pressure from potential entrants. This project will apply a class of dynamic models to Census Bureau microdata for a range of industries and geographic markets. The goal is to estimate underlying variables and parameters of interest (e.g., average profits in a market and the parameters describing the distribution of sunk entry costs and scrap values) in fully dynamic models with imperfectly competitive markets.

THE REAL COSTS OF CREDIT ACCESS: EVIDENCE FROM THE PAYDAY LENDING MARKET

Marianne Bertrand—The University of Chicago
Brian Melzer—The University of Chicago

This project uses nonpublic data from the Survey of Income and Program Participation (SIPP) to study the effect of access to short-term personal loans on household welfare, particularly among low-income populations. Specifically, it will estimate the impact of loan access on the following outcomes: difficulty paying rent and utilities, eviction, termination of utilities due to default, delay of needed health care, and household debt. The measure of loan access, geographic proximity to payday loan stores, depends upon fairly detailed information on household location (census tract and county), necessitating access to nonpublic geographic identifiers in the SIPP. By investigating a determinant of economic well-being among low-income individuals, this project fulfills one of the Census Bureau's goals in conducting the SIPP: to provide improved measures of economic well-being. If short-term borrowing does influence welfare, then this would suggest a useful revision to the SIPP, whereby the Census Bureau could offer a clearer picture of the financial situation of low-income individuals by inquiring about payday loan usage among SIPP respondents.

POLLUTION ABATEMENT EXPENDITURES AND OUTCOMES: INTERNAL AND EXTERNAL DETERMINANTS

Wayne B. Gray—Clark University
Ronald J. Shadbegian—U.S. Environmental Protection Agency

This project seeks to improve the understanding and the quality of the plant-level data on environmental spending collected in the Census Bureau's Pollution Abatement Costs and Expenditures (PACE) survey. The project combines PACE data with other Census Bureau datasets and with external data to model the impact of pollution abatement spending on economic factors, such as the plant's production costs and productivity, as well as its pollution emissions. The research will test accuracy of reported abatement expenditures by modeling their impact on total factor productivity levels, which should decrease productivity on a one-for-one basis if abatement activities contribute nothing to production. It also models the plant's production function, testing whether the productivity of specific types of inputs are more seriously affected by pollution abatement activities. Analyses include tests for differences across plants in abatement costs and in their impact on productivity, based on plant size, age, and other observed characteristics. The project will also model the impact of reported abatement costs on a variety of business decisions, including shifts in economic activity and investment, providing an indirect test for the reality of abatement costs. The project will benefit the Census Bureau in several ways. The PACE survey has been recently resumed after an extended hiatus, so information about its data quality and comparisons to data from earlier versions of the survey are valuable. Our models of the impact of reported PACE spending on productivity and emissions test their validity in two ways: are they costs, and do they abate pollution? Our external datasets provide information on production technology and material use that will be used to assess the quality of comparable Census Bureau-collected information. Finally, the external datasets (particularly the Environmental Protection Agency's Facility Registry System) provide carefully maintained name-address, latitude-longitude, and plant ownership data that will provide information about the quality of the comparable Census Bureau information and how quickly that information is updated.

SEMIPARAMETRIC ESTIMATION OF ADVERSE SELECTION AND MORAL HAZARD IN EMPLOYER-SPONSORED HEALTH INSURANCE

Matthew S. Rutledge—University of Michigan

The estimation of the effects that adverse selection, inefficient provision of insurance to certain risk types, and moral hazard (the overconsumption of medical care by insured patients who do not face the true cost) has on the health insurance market has been of great interest to economists in recent years. Key elements of the estimation of these information asymmetries are the menu of insurance plans available to the consumer, the degree to which consumers have to share in the cost of their medical care (i.e., the coinsurance schedule), and the relative health of the consumer, which is usually unknown to the insurer due to legal limitations on experience rating. Most attempts to estimate adverse selection and moral hazard, however, lack data that explicitly provide these key elements, and researchers are forced to make additional assumptions and ancillary estimations that add uncertainty to their results. This project uses data that provide the menu of plans offered by employers to their employees, including explicit coinsurance schedules for each plan, which are then linked to data on employees' medical expenditure and health conditions. These data have been collected from the employers of respondents to the Medical Expenditure Panel Survey (MEPS) between 1996 and 1999.

MICROFOUNDATIONS OF AGGREGATE LABOR DEMAND FOR SKILL

Lawrence Katz—Harvard University
Joshua W. Mitchell—Harvard University

This project uses the Quarterly Workforce Indicators (QWI) and human capital files to better understand how workforce and production aggregation choices affect estimates of labor demand for skill and influence our understanding of recent changes in wage inequality. Decompositions of changes over time in employment and wage shares of skill and demographic groups, formal modeling and estimation of labor demand parameters, and correlations between demand and indicators of technology, capital, and trade from public-use datasets will be performed. These estimates will be generated at distinct levels of aggregated producer data: establishment, detailed industry, industry group, and economy-wide. The results will be used to reconcile microeconomic and macroeconomic trends in the wage structure and help evaluate competing explanations for the evolution of wage inequality.

TRADE, GLOBALIZATION, AND THE ENTERPRISE-ESTABLISHMENT RELATIONSHIPS OF MULTINATIONAL COMPANIES

Maria Borga—U.S. Bureau of Economic Analysis
Bryan Goudie—U.S. Bureau of Economic Analysis
Marilyn Ibarra—U.S. Bureau of Economic Analysis
Jennifer Koncz—U.S. Bureau of Economic Analysis
Raymond Mataloni, Jr.—U.S. Bureau of Economic Analysis
Robert Yukavage—U.S. Bureau of Economic Analysis
Steven Zemanek—U.S. Bureau of Economic Analysis

This project examines the impact of trade and globalization on U.S. businesses. The project has three objectives. It will study the regional impacts of rapidly growing U.S. foreign exports of services by developing estimates of exports of services at the state level. It will examine the impact of globalization by U.S. multinational companies through direct investment, trade, and the offshoring of services on the U.S. economy. It will examine the

relationship between an enterprise and its establishments and explore methods of allocating enterprise-level data to establishments in various industries. All three components of this project will involve supplementing data collected at the Census Bureau on U.S. business activities with data collected at the U.S. Bureau of Economic Analysis (BEA) on the operations of multinational firms and on trade in services. Most Census Bureau business data are collected at the establishment level, while BEA's surveys are conducted on a consolidated business enterprise basis. Linking these datasets greatly increases their analytic value by allowing questions about decisions made at the enterprise level—such as where to locate production to serve foreign customers—to be examined using the establishment-level data best suited to answering them. This project will build on previous projects that have linked Census data with BEA enterprise data. These previous linking projects have provided benefits to the Census Bureau, and this project will expand on these benefits. For example, it will enable the Census Bureau to improve its sample frames, verify and improve the accuracy of data reported on its surveys, and improve the industry classifications of enterprises and establishments. In addition, it will greatly increase the utility of the Census Bureau datasets for examining the impact that globalization—in the form of direct investment and trade in services—is having on the U.S. economy.

FIRM PERFORMANCE EFFECTS OF BROAD-BASED EMPLOYEE STOCK OWNERSHIP

E. Han Kim—University of Michigan
Paige Ouimet—University of North Carolina

This project examines whether broad-based equity-based incentives are effective at aligning incentives and the potential costs associated with providing these incentives. If employee ownership provides appropriate incentives, worker productivity should be higher and employee turnover should be lower. However, employee ownership also gives employees voting rights, which can be used to extract employee benefits at the expense of other stakeholders in the firm. For example, workers with voting rights may be more successful at obtaining above-market wages or in delaying or preventing layoffs or plant closures. Examining the effects of employee ownership on firm and establishment performance measures—such as productivity and wages—which in turn affect firm value, will provide important insights into the benefits and limitations of equity-based compensation.

RESETTLEMENT AND WELL-BEING OF NEW ORLEANS RESIDENTS AFTER HURRICANE KATRINA

William Frey—University of Michigan
Michael Rendall—RAND Corporation
Narayan Sastry—University of Michigan

This project employs data from the Census Bureau's American Community Survey (ACS) to examine the current location and well-being of residents of New Orleans in the year after Hurricane Katrina struck the city. The aims of this project are to describe the return or resettlement in the year following Hurricane Katrina of people who resided in New Orleans before the storm and to examine the well-being of the pre-Katrina New Orleans population in the year after the hurricane, compared to a matched population from the prior year. The ACS data provide a unique opportunity to examine the geographic dispersion of New Orleans residents throughout the United States in the aftermath of Hurricane Katrina and to assess several important dimensions of well-being. This project addresses a number of unanswered research questions about the effects of Hurricane Katrina on the New Orleans' population. It also explores the strengths and weaknesses of the ACS data for examining

the effects of Hurricane Katrina and for future studies of the effects of large-scale natural and man-made disasters. The researchers will evaluate the suitability of propensity-score reweighting techniques for studying these topics using the ACS by, for example, examining which population groups are underrepresented and overrepresented between the two ACS cross-sections.

PRIVATE EQUITY DEALS: OPERATIONAL AND FINANCIAL PERFORMANCE OF U.S. BUYOUTS

Fernando Chaddad—University of North Carolina
David Ravenscraft—University of North Carolina

The present study is based on prior work that studied the financial performance of a sample of leveraged buyouts (LBO) and going-private transactions occurring between 1978 and 1989. That project identified shortcomings of census data collection programs by comparing Census Bureau data to other overlapping data by conducting quality checks of Quarterly Financial Report (QFR) data against other sources such as Standard and Poor's COMPUSTAT®. The present study will extend the methods developed for the 1993 Center for Economic Studies (CES) study over many more years of data to examine the recent phenomenon of private equity buyouts using the QFR data.

The researchers will directly compare the LBO of the 1980s with present-era private equity deals. This will improve the Census Bureau's understanding of firm structure and ownership change, as well as the impact of these on firm performance.

This project will also increase the Census Bureau's knowledge base regarding the financial and operational performance of private equity buyouts in the United States from 1990 to 2007 and, hence, improve the Census Bureau's understanding of firm structure and ownership change. The main question to be answered is how do present-era private equity buyouts perform financially and operationally. Moreover, this project will determine whether the short-term (1 year) and the longer-term performance (3 years) of private equity buyouts increases, decreases, or remains unchanged for the sample of buyout firms, relative to their industry peers. It will also determine whether the sources of improvements were purely financial (debt) or operational in nature. This can lead to serious implications with regards to firm and national competitiveness. This study will also address the question of how similar today's private buyouts are vis-à-vis the 1980s LBO. As the recent problems in subprime lending have shown, problems in one debt class can have large repercussions on the economy as a whole.

TAXES AND THE STRATEGIC TIMING OF CHILDBIRTH

James Sallee—University of Michigan

The federal tax system in the United States offers tax benefits worth several thousand dollars per child to the average family. Existing research has found evidence that parents manipulate the timing of childbirth in order to maximize tax benefits. Using birth record data from the Vital Statistics system at the National Center for Health Statistics, this project examines whether or not the mechanism (elective induced labor and elective caesareans) could explain this phenomenon. This study investigates whether or not this mechanism appears to be responsive to changes in the tax code. The findings of this research bear on larger questions about the degree of responsiveness of fertility to government incentives and to public health concerns about rising rates of induced labor and caesarean births in the United States.

LONG-TERM HEALTH OUTCOMES AND SERVICE IN VIETNAM: ESTIMATING CAUSALITY VIA THE VIETNAM DRAFT LOTTERY

Daniel Eisenberg—University of Michigan

Estimating the causal effects of military service on health outcomes is complicated by selection bias. People who serve in the military may be different, on average, in characteristics, such as tolerance for risk, that are difficult to measure. Such differences can confound conventional estimates of health consequences of service that depend on comparing health outcomes for people who served to people with similar characteristics who did not serve. This research design addresses this issue by making use of the fact that from 1970 to 1973, an individual's risk of induction into the military was primarily a function of a randomly assigned number based on date of birth. This Vietnam draft lottery is used to construct a natural experiment in which the comparison groups are essentially those who had low draft numbers (and therefore higher probability of service) and those who had high draft numbers (and therefore lower probability of service). Due to the random nature of the lottery, these two groups are virtually identical on average, and estimates of causality are therefore unbiased by the prior differences in the groups.

THE DEMAND FOR MULTIPLE PRIVATE INSURANCE POLICIES WITH PREMIUM SUBSIDY AND CHANGES IN MEDICAL CARE DEMAND

Hyo Jung Tak—The University of Chicago

This study investigates how a high premium subsidy in employer-sponsored health insurance induces changes in people's behavior both in insurance take-up decision and choice when multiple choices are available. It is expected that a high premium subsidy makes people choose a more generous plan when multiple choices are available and also makes a family purchase multiple employer-sponsored health insurance when there are two working spouses. This research, employing a special extract of the Medical Expenditure Panel Survey, investigates whether the high premium subsidy changes the insurance structure, such as deductible and copayment/coinsurance that people choose, and analyzes whether it increases medical care demand.

DO STABLE NURSING HOME STAFF IMPROVE RESIDENT OUTCOMES?

Laura D'Arcy—University of North Carolina
Sally Stearns—University of North Carolina

Many nursing homes suffer from high rates of turnover, especially among staff providing direct care to residents. High levels of staff turnover can have deleterious effects on resident outcomes, ranging from mental status to clinical measures such as rates of bedsores and pain management. This study has two empirical objectives: (1) to examine the relationship between resident outcomes and staff turnover (including registered nurses, licensed practical nurses, and certified nursing assistants); and (2) to test two strategies (fixed effects and instrumental variables) to correct for endogeneity of turnover, providing an improved estimate of the effect of turnover on resident outcomes. The project employs data from the National Center for Health Statistics' National Nursing Home Survey.

HARMONIZATION OF HISTORICAL DECENNIAL CENSUS DATA

Joseph Alexander—University of Minnesota
Todd Gardner—U.S. Census Bureau
Matthew Schroeder—University of Minnesota

The goal of this project is to harmonize the Census Bureau's internal decennial census microdata files from the 1960, 1970, 1980, 1990, and 2000 decennial censuses. This effort is part of the National Historical Census Files Project (NHCFP), the ultimate goal of which is to construct a time series of easy-to-use historical decennial census files and make these files available to researchers within the Census Bureau and to qualified researchers through the Census Bureau's Research Data Centers. In the first phase of NHCFP, all of the available microdata from the 1960, 1970, and 1980 decennial censuses were converted to standard formats. The next step is to merge the historical decennial census microdata files, along with the microdata from the 1990 and 2000 censuses, into a single database with consistent coding and integrated documentation. Harmonization of the Census Bureau's internal historical microdata files will follow the model established by the Minnesota Population Center (MPC) for the Integrated Public Use Microdata Series (IPUMS), which is the harmonized database merging all of the original Public Use Microdata Samples (PUMS). MPC is sharing its software and expertise to produce the internal counterpart to the IPUMS, which will be known as the Integrated Confidential Use Microdata Series, or ICUMS.

AGGLOMERATION EFFECTS: THE ROLE OF SELECTION

Jason Faberman—Federal Reserve Bank of Philadelphia
Matthew Freedman—Cornell University

This research investigates the extent to which firm learning and selection account for observed geographic agglomeration effects. A vast literature documents positive relationships between the wages and productivity of firms and various measures of agglomeration, effects that persist even after controlling for a broad array of worker, firm, and other local area characteristics. In many cases, researchers attribute the positive observed effects of spatial agglomeration on different outcome variables as evidence of agglomeration economies, suggesting that there may be knowledge spillovers or externalities associated with the geographic clustering of economic activity. This project explores the role of selection in explaining observed agglomeration effects using the Census Bureau's Longitudinal Business Database (LBD). The LBD provides detailed geographic classifications that permit construction of measures of agglomeration at fine geographic levels. This research will shed light on the sources of observed agglomeration effects, the drivers behind differential rates of firm turnover and economic growth across geographic areas, and the potential ramifications of policies aimed at encouraging or discouraging clustered business activity. This work will also yield a number of benefits to the Census Bureau by identifying shortcomings of existing data and methodologies, as well by improving the quality and utility of the data through longitudinal editing and the preparation of new population estimates.

THE EFFECT OF HEALTH CARE COSTS ON THE GROWTH AND SURVIVAL OF SMALL BUSINESS

Adela Luque—The Urban Institute

Whether or not small businesses offer health insurance to their employees is a critical factor in the health care coverage of many Americans, yet little is known about the relationship between health care costs and business growth. The project links data from the annual Medical Expenditure Panel Survey—Insurance Component (MEPS-IC) data files to each other as well as to the Longitudinal Business Database (LBD). The study would use MEPS-IC data from 1996 to the present, longitudinally linked using the LBD, to examine the relationship between offering health insurance and firm outcomes. The primary emphasis is on the effect of insurance offering on firm performance, rather than the effects of firm performance on the decision to offer insurance. The project uses differences over time for establishments that are observed more than once to estimate these effects, using instrumental variables to deal with endogeneity. The project also will develop a methodology for estimating nonresponse to the MEPS-IC.

THE ROLE OF INDUSTRY CLASSIFICATION AND FIRM STRUCTURE IN THE ESTIMATION OF RESEARCH AND DEVELOPMENT EXPENDITURES

G. Andrew Bernat, Jr.—U.S. Bureau of Economic Analysis
Bryan Goudie—U.S. Bureau of Economic Analysis
Leonard Loebach—U.S. Bureau of Economic Analysis
Gabriel Medeiros—U.S. Bureau of Economic Analysis
Leo Sveikauskas—U.S. Bureau of Labor Statistics

This project will reclassify the estimates of research and development (R&D) performance in the Survey of Industrial Research and Development (SIRD) from firm-based industries to establishment-based industries using a consistent industry classification (NAICS) for the entire time series of the SIRD. The research will convey benefits to the Census Bureau along three dimensions by producing estimates of R&D using establishment-based industry codes; by conducting data quality assessments of the SIRD data, which exploit the longitudinal nature of the data; and by analyzing firm-establishment relationships. The Census Bureau is currently in the process of a major redesign of the SIRD, so the insights gained from this project will be particularly timely in contributing to the improvement of the future versions of the SIRD. The project will make an important contribution to the Census Bureau's efforts to classify economic information on an establishment basis by identifying the establishments within each firm that are most likely to perform the firm's R&D. Identifying these establishments within each firm will also substantially improve the Census Bureau's efforts to accurately measure economic activity within states because many R&D-performing companies have establishments in more than one state.

COMMUNITY-LEVEL INFLUENCES ON YOUNG MEN'S SEXUAL AND REPRODUCTIVE HEALTH BEHAVIORS: 1988 TO 2002

Laura Lindberg—Columbia University
Mark Orr—Columbia University

This study examines changing disparities in young men's sexual and reproductive health (SRH) from 1988–2002, focusing on differences by level of community disadvantage. Data from unmarried males aged 15–19 are pooled from two nationally representative household surveys, the 1988 National Survey of Adolescent Males and the 2002 National Survey of

Family Growth. From the pooled samples, researchers estimate a series of multivariate models across a range of relevant SRH outcome variables. In the first stage, each model includes controls for individual and family influences, as well as a variable for year of survey (1988/2002). At the second stage, added to the models are measures of community disadvantage based on characteristics of the neighborhood in which the respondent resides at the time of the survey. Finally, interactions between year and the community-level measure test the extent to which disparities by community disadvantage in young men's SRH have changed over time.

LOCATIONAL ATTAINMENT AND RESIDENTIAL SEGREGATION IN U.S. METROPOLITAN AREAS

John Iceland—University of Maryland
Melissa Scopilliti—University of Maryland

Immigration of Asians and Hispanics has fueled recent growth in the non-White population in the United States. As of 2000, 31 percent of the population was of a group other than non-Hispanic White, up from nearly 25 percent a decade earlier. The influx of immigrants, particularly to metropolitan areas, is changing the demographics of America's neighborhoods. The first part of this project examines the relationship between individual race/ethnicity, nativity, and human capital characteristics with levels of neighborhood economic advantage and racial diversity. Often termed residential or locational attainment, this research investigates the effectiveness of spatial assimilation and place stratification theories for understanding racial and ethnic stratification across neighborhoods. The second portion of the analysis explores the relationship between metropolitan context and locational attainment. Metropolitan-level residential segregation indexes by race and ethnicity will be developed, and the relationship between metropolitan characteristics (segregation and ecological factors) and locational attainment will be examined.

RENEWABLE ELECTRIC POWER GENERATION

Allan Collard-Wexler—New York University
Joan Farre-Mensa—New York University

The goal of this project is to understand the decision to build renewable electric power generators and how policy could encourage the build out of wind turbines over the next 20 years. Datasets to be used include the Longitudinal Business Database and the Census of Transportation, Communication, and Utilities. Measures to reduce emissions, such as the introduction of a Carbon Tax or subsidies for wind power, can only be evaluated using a model of how utilities decide to build and maintain wind turbines or solar generators and when they will choose to shut down coal fired plants. Wind power is increasingly seen as an alternative to coal, oil, and gas generators for producing electricity. For instance, Denmark now produces 19 percent of its electricity using wind power, while in the United States wind power accounts for 1.1 percent of total electric power. This project will describe generators of wind power and other renewable power sources in the United States over the last 20 years. Second, it will construct and estimate dynamic models of entry and investment that will be used to simulate the effect of policies, such as tax subsidies, that would lower the cost of building new wind power generators in the future.

FEMALE LABOR FORCE PARTICIPATION AND ADULT HEALTH

Melinda Morrill—North Carolina State University

This project investigates the impact of married women's labor force participation on own and husband's health. It focuses on measures of body weight, high blood pressure, physical activity levels, and self-reported levels of stress. The opportunity cost of a woman working is substantially lowered when her youngest child becomes eligible for public school, leading to an increase in maternal labor supply at that time. This study uses state-level and year-specific variation in kindergarten eligibility laws to precisely measure the youngest child's eligibility for kindergarten. Information on state of residence is only available on the restricted-use version of the National Health Interview Survey. Preliminary results find that female employment has a positive impact on body mass index (BMI) for married men with less than a high school education. This finding is consistent with men facing an increase in the cost of home cooking with a positive impact on body weight. Women face an offsetting rise in the level of physical activity, and households whose husbands have higher income can afford less calorie-intensive prepared food. The magnitude of these findings is larger than found elsewhere in the literature. The results also show that for married men with higher levels of education, female demand shocks produce an increase in the levels of physical activity. This last element, plus the positive impact of female demand shocks on BMI and obesity rate, suggests that the channel through which female labor force participation raises a man's weight must be through a higher consumption of calories, which is consistent with a lower consumption of fruit and vegetables that are associated with female demand shocks for this group.

SCHOOL ENTRY LAWS AND ADHD DIAGNOSIS

Melinda Morrill—North Carolina State University

Attention-deficit/hyperactivity disorder (ADHD) is a neurobehavioral disorder of childhood that is difficult to diagnose and has no known cause. There is some concern that diagnoses for ADHD are not for underlying behavioral problems but rather represent age-appropriate differences in behavior. This project seeks to determine whether a child that is "young for their grade" and is hence relatively immature, is more likely to have been diagnosed with ADHD. The research exploits the sharp discontinuity in school enrollment generated by kindergarten eligibility laws, which are determined at the state-level. Children born just before the cutoffs will, on average, be young for their age; while children born just after the cutoff must, on average, wait a year to enter school and therefore will be the oldest children in their class. ADHD is an underlying neurological problem where incidence rates should not change dramatically from one birth date to the next. If diagnosis rates do shift appreciably based on small changes in birth dates, then the diagnosis is not based on underlying conditions but rather behavior relative to peers in class. This project uses data from the National Center for Health Statistics from 1985 through 2004 to demonstrate the relationship between eligibility date and grade for age. These data are used to test whether children born just prior to the kindergarten eligibility cut-off date have higher rates of ADHD diagnosis than children born just after those dates.

THE RELATIONSHIP BETWEEN MEDICAL EXPENDITURE AND QUALITY OF LIFE FOR GENERAL POPULATION IN UNITED STATES

Meiying Han—Stony Brook University

This project seeks to add to the existing health literature by studying how medical expenditure is associated with quality of life for the general population in the United States. It employs a two-stage least squares analysis, which uses peoples' exposure to local health service intensity, including number of health service professions, facilities, and local population characteristics. The research also examines a reduced form method to determine how regional resources that increase medical expenditure affected their quality of life. This study uses data from individual interviews collected from the Medical Expenditure Panel Survey regarding general health status and medical expenditure variables.

Appendix 3.
CENTER FOR ECONOMIC STUDIES (CES) DISCUSSION PAPERS 2008

CES Discussion Papers are available at <www.ces.census.gov>.

08-41 "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices," by Ali Hortaçsu and Chad Syverson, 12/08.

08-40 "Manufacturing Plants' Use of Temporary Workers: An Analysis Using Census Micro Data," by Yukako Ono and Daniel Sullivan, 12/08.

08-39 "Gender Differences in Business Performance: Evidence From the Characteristics of Business Owners Survey," by Robert W. Fairlie and Alicia M. Robb, 12/08.

08-38 "Measuring Labor Earnings Inequality Using Public-Use March Current Population Survey Data: The Value of Including Variances and Cell Means When Imputing Topcoded Values," by Richard V. Burkhauser, Shuaizhang Feng, and Jeff Larrimore, 11/08.

08-37 "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?," by Edward L. Glaeser and William R. Kerr, 10/08.

08-36 "Linking Investment Spikes and Productivity Growth: U.S. Food Manufacturing Industry," by Pinar Celikkol Geylani and Spiro E. Stefanou, 10/08.

08-35 "The Effects of Smoking in Young Adulthood on Smoking and Health Later in Life: Evidence Based on the Vietnam Era Draft Lottery," by Daniel Eisenberg and Brian Rowe, 9/08.

08-34 "The Green Industry: An Examination of Environmental Products Manufacturing," by Randy A. Becker and Ronald J. Shadbegian, 9/08.

08-33 "Productivity Dispersion and Input Prices: The Case of Electricity," by Steven J. Davis, Cheryl Grim, and John Haltiwanger, 9/08.

08-32 "A Comparison of Employee Benefits Data From the MEPS-IC and Form 5500," by Kristin McCue, 9/08.

08-31 "The Direct and Indirect Costs of Food Safety Regulation," by Michael Ollinger, 9/08.

08-30 "Computer Network Use and Firms' Productivity Performance: The United States vs. Japan," by B.K. Atrostic, Kazuyuki Motohashi, and Sang V. Nguyen, 9/08.

08-29 "Transfer Pricing by U.S.-Based Multinational Firms," by Andrew B. Bernard, J. Bradford Jensen, and Peter K. Schott, 9/08.

08-28 "An Analysis of Key Differences in Micro Data: Results From the Business List Comparison Project," by Kristin Fairman, Lucia Foster, C.J. Krizan, and Ian Rucker, 9/08.

08-27 "Choices of Metropolitan Destinations by the 1995–2000 New Immigrants Born in Mexico and India: Characterization and Multivariate Explanation," by Kao-Lee Liaw and William H. Frey, 9/08.

08-26 "Business Volatility, Job Destruction, and Unemployment," by Steven J. Davis, R. Jason Faberman, John Haltiwanger, Ron Jarmin, and Javier Miranda, 8/08.

08-25 "Estimating Trends in U.S. Income Inequality Using the Current Population Survey: The Importance of Controlling for Censoring," by Richard V. Burkhauser, Shuaizhang Feng, Stephen P. Jenkins, and Jeff Larrimore, 8/08.

08-24 "Multi-Product Firms and Product Switching," by Andrew B. Bernard, Stephen J. Redding, and Peter K. Schott, 8/08.

08-23 "Supersize it: The Growth of Retail Chains and the Rise of the 'Big Box' Retail Format," by Emek Basker, Shawn Klimek, and Pham Hoang Van, 8/08.

08-22 "Products and Productivity," by Andrew B. Bernard, Stephen J. Redding, and Peter K. Schott, 8/08.

08-21 "Health-Related Research Using Confidential U.S. Census Bureau Data," by Rosemary Hyson and Alice Zawacki, 8/08.

08-20 "The Going Public Decision and the Product Market," by Thomas Chemmanur, Shan He, and Debarshi Nandy, 7/08.

08-19 "The Effect of Power Plants on Local Housing Values and Rents: Evidence From Restricted Census Microdata," by Lucas W. Davis, 7/08.

08-18 "Using Internal Current Population Survey Data to Reevaluate Trends in Labor Earning Gaps by Gender, Race, and Education Level," by Richard V. Burkhauser and Jeff Larrimore, 7/08.

08-17 "Entry, Exit, and Plant-Level Dynamics Over the Business Cycle," by Yoonsoo Lee and Toshihiko Mukoyama, 6/08.

08-16 "How Does Venture Capital Financing Improve Efficiency in Private Firms? A Look Beneath the Surface," by Thomas Chemmanur, Karthik Krishnan, and Debarshi Nandy, 6/08.

08-15 "*You Can Take it With You*: Proposition 13 Tax Benefits, Residential Mobility, and Willingness to Pay for Housing Amenities," by Fernando Ferreira, 6/08.

08-14 "Market Forces, Plant Technology, and the Food Safety Technology Use," by Michael Ollinger and Danna Moore, 6/08.

08-13 On the Lifecycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms," by Manju Puri and Rebecca Zarutskie, 05/08.

08-12 "Neighborhood Effects on High-School Drop-Out Rates and Teenage Childbearing: Tests for Non-Linearities, Race-Specific Effects, Interactions With Family Characteristics, and Endogenous Causation Using Geocoded California Census Microdata," by Rhiannon Patterson, 5/08.

08-11 "Analysis of Young Neighborhood Firms Serving Urban Minority Clients," by Timothy Bates and Alicia Robb, 5/08.

08-10 "Horizontal Diversification and Vertical Contracting: Firm Scope and Asset Ownership in Taxi Fleets," by Evan Rawley and Timothy S. Simcoe, 5/08.

08-09 "A Warm Embrace or the Cold Shoulder: Wage and Employment Outcomes in Ethnic Enclaves," by Roberto Pedace and Stephanie Rohn, 4/08.

08-08 "The Efficiency of Internal Capital Markets: Evidence From the Annual Capital Expenditure Survey," by Sumit Agarwal, Victor Souphom, and Guy M. Yamashiro, 4/07.

08-07 "Private Equity and Employment," by Steven J. Davis, John Haltiwanger, Ron Jarmin, Josh Lerner, and Javier Miranda, 3/08.

08-06 "Consistent Cell Means for Topcoded Incomes in the Public Use March CPS (1976–2007)," by Jeff Larrimore, Richard V. Burkhauser, Shuaizhang Feng, and Laura Zayatz, 3/08.

08-05 "Trends in the Relative Household Income of Working-Age Men With Work Limitations: Correcting the Record Using Internal Current Population Survey Data," by Richard V. Burkhauser and Jeff Larrimore, 3/08.

08-04 "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Pre-Kindergarten," by Maria D. Fitzpatrick, 3/08.

08-03 "What Happens When Firms Patent? New Evidence From U.S. Economic Census Data," by Natarajan Balasubramanian and Jagadeesh Sivadasan, 1/08.

08-02 "Who Gentrifies Low Income Neighborhoods?" by Terra McKinnish, Randall Walsh, and T. Kirk White, 1/08.

08-01 "Financial Intermediation and Late Development: The Case of Meiji Japan, 1868 to 1912," by John P. Tang, 1/08.

Appendix 4.
NEW DATA AVAILABLE THROUGH RESEARCH DATA CENTERS (RDCS) IN 2008[1]

Table A-4.1.
BUSINESS DATA

Data product	Description	New or updated years
Annual Survey of Manufactures	The Annual Survey of Manufactures provides statistics on employment, payroll, workers' hours, payroll supplements, cost of materials, value added by manufacturing, capital expenditures, inventories, and energy consumption. It also provides estimates of value of shipments for over 1,800 classes of manufactured products.	2003–2006
Auxiliary Establishment Survey	The Auxiliary Establishment Survey (AUX) covers auxiliary establishments of multiestablishsment firms, reporting separately. The primary function of auxiliary units is to manage, administer, service, or support the activities of the other establishments of the company. Information on the auxiliary unit, the owning company, and the type of management or support service function provided is collected. The AUX is sometimes referred to as the Company Auxiliary Establishments. The Center for Economic Studies (CES) still maintains AUX after it was discontinued in 1997. The 2002 AUX is constructed from 2002 Economic Census data.	2002
Business Expenses Survey	The Business Expenses Survey, formerly known as the Assets and Expenditures Survey, supplements basic economic statistics produced by the economic census for wholesale trade, retail trade, and service industries with estimates of operating expenses. Data collected on operating expenses include payroll and fringe benefits, contract labor costs, taxes and license fees, depreciation and amortization charges, software and other computer expenses, office supplies, repair and maintenance expenses, lease and rental payments, utilities, advertising, accounting, and legal services.	2002
Foreign Trade Transaction Import Data	Foreign Trade Transaction Import Data contain information on U.S. imports of merchandise compiled primarily from automated data submitted through the U.S. Customs' Automated Commercial System. The data are also compiled from import entry summary forms, warehouse withdrawal forms, and Foreign Trade Zone documents required by law to be filed with the U.S. Customs and Border Protection. Data on imports of electricity and natural gas from Canada are obtained from Canadian sources.	1992–2005

[1] These tables do not include custom extract data made available to approved projects from the Longitudinal Employer-Household Dynamics program, National Center for Health Statistics and Agency of Healthcare Research and Quality.

Data product	Description	New or updated years
Foreign Trade Transaction Export Data	Foreign Trade Transaction Export Data contain information on U.S. exports of merchandise from the United States to all countries, except Canada. The data are compiled from copies of Shipper's Export Declarations (SEDs) from qualified exporters, forwarders, and carriers. Each SED represents a shipment of one or more kinds of merchandise from one exporter to one foreign importer on a single carrier.	1992–2005
Medical Expenditure Panel Survey—Insurance Component	The Medical Expenditure Panel—Insurance Component collects data on health insurance plans obtained through employers. Data collected include the number and type of insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, eligibility requirements, and employer characteristics.	2003–2006
Quarterly Financial Report	The Quarterly Financial Report collects information on the financial results and position of U.S. corporations.	1992–2008 (first quarter only for 2008)
Survey of Industrial Research and Development	The Survey of Industrial Research and Development (SIRD) is the primary source of information on research and development (R&D) performed by industry within the United States. The survey is a nationally representative sample of all for-profit companies, publicly or privately held and with five or more employees that performed R&D within the United States. The SIRD collects detailed information on the cost of R&D performed and the employment of scientists and engineers.	2001–2005
Survey of Business Owners	The Survey of Business (SBO) includes data from all nonfarm businesses filing tax forms as individual proprietorships, partnerships, or any type of corporation and with receipts of $1,000 or more. The SBO covers both firms with paid employees and firms with no paid employees. The SBO is conducted on a company or firm basis rather than an establishment basis.	2002
Standard Statistical Establishment Listing	Standard Statistical Establishment Listing files maintained at CES are created from the old Standard Statistical Establishment List (prior to 2002) and the new Business Register (2002 and forward).	2005, 2006

Data product	Description	New or updated years
Survey of Plant Capacity Utilization	The Survey of Plant Capacity Utilization (PCU) provides current data on the rates of capacity utilization in U.S. manufacturing plants. Data collected are for the fourth quarter and include number of days and hours worked, estimated value of production at full production capability, and estimated value of production achievable under national emergency conditions. Additional items include reasons why the plant may operate at less than full production, reasons why the estimate of full production capability changed from the prior year, and how quickly the plant could reach full production and national emergency levels of production. The PCU was discontinued after 2006. The Quarterly Survey of Plant Capacity Utilization replaces the PCU beginning in 2007.	1975–2006

Table A-4.2.
HOUSEHOLD DATA[1]

Data product	Description	New or updated years
Decennial Census Long Form	The decennial census long form is a 1 in 6 sample of all households in the United States. Long form questions include age, race, sex, educational attainment, income, place of work, occupation, household relationships, and housing unit characteristics.	1970, 1980, 1990
Decennial Census Short Form	The decennial census short form questionnaire includes population and housing questions. The questions on the short form are also asked on the long form, along with additional questions.	1970, 1980, 1990
Survey of Income and Program Participation Panels	The Survey of Income and Program Participation collects information on source and amount of income, labor force information, program participation and eligibility data, and general demographic characteristics to measure the effectiveness of existing federal, state, and local programs; to estimate future costs and coverage for government programs, such as food stamps; and to provide improved statistics on the distribution of income in the country.	2001

[1] These demographic or decennial files maintained at CES and for the RDCs are the internal versions, and they provide researchers with variables and detailed information that are not available in the corresponding public-use files.

Table A-4.3.
LONGITUDINAL EMPLOYER-HOUSEHOLD DYNAMICS (LEHD) DATA[2]

Data product	Description	New or updated years
Business Register Bridge	The Business Register Bridge (BRB) is a link file between LEHD employer microdata and Business Register (BR) firm and establishment microdata. Since the concepts of "firm" and "establishment" differ between the LEHD employer microdata and the BR, the BRB provides a crosswalk at various levels of business-unit aggregation. The most detailed crosswalk is at the level of Employer Identification Number (EIN)–State–four-digit Standard Industry Classification (SIC) industry–county. The bridge includes the full list of establishments in the LEHD data and in the BR that are associated with the business units (e.g., EIN-four-digit SIC-State-County) in the crosswalk and measures of activity (e.g., employment, sales).	1990–2001
Employer Characteristics File	The Employer Characteristics File (ECF) consolidates most firm-level information (size, location, industry, etc.) into two easily accessible files. The firm-level file has one record for every year and quarter in which a firm is present in either the covered Employment and Wages (ES-202) program data or the unemployment insurance system (UI) wage records. Firms are identified by the LEHD State Employer Identification Number (SEIN). The data in the firm-level file is aggregated from the core establishment-level file, where establishments are identified by reporting unit number within SEIN, called SEINUNIT.	1990–2004
Employment History File	The Employment History File (EHF) provides a full time series of earnings at all within-state jobs for all quarters covered by the LEHD system and provided by the state. It also provides activity calendars at a job, firm, and subfirm reporting unit level. It can be linked to other Census Bureau files through the Protected Identity Key (PIK) and the LEHD SEIN.	1985–2004
Geocoded Address List	The Geocoded Address List (GAL) is a dataset containing unique commercial and residential addresses in a state geocoded to the census block and latitude/longitude coordinates. It consists of the GAL and a crosswalk for each processed file-year. The GAL contains each unique address, a GAL identifier, its geocodes, a flag for each file-year in which it appears, data quality indicators, and data processing information. The GAL Crosswalk contains the GAL identifier.	1990–2004 (Years covered vary by state)
Individual Characteristics File	The Individual Characteristics File (ICF) for each state contains one record for every person who is ever employed in that state over the time period spanned by the state's unemployment insurance records. It consolidates information from multiple input sources on gender, age, citizenship, point-in-time residence, and education. Information on gender, education, and age is imputed ten times when missing.	1985–2004

[2] Detailed information on how the LEHD files were constructed, documentation of variables, and some computer programs can be found in the technical documentation at <lehd.did.census.gov>.

Quarterly Workforce Indicator	The Quarterly Workforce Indicators (QWI) establishment file contains quarterly measures of workforce composition and worker turnover at the establishment level for selected states. The LEHD establishment-level measures are created from longitudinally integrated person and establishment-level data. Establishment-level measures include: (i) Worker and Job Flows—accessions, separations, job creation, job destruction by age and gender of workforce; (ii) Worker composition by gender and age; (iii) Worker compensation for stocks and flows by gender and age; (iv) Dynamic worker compensation summary statistics for stocks and flows by gender and age. The LEHD-QWI may be used in combination with the LEHD BRB to match to other Census Bureau micro business databases and can be matched by firm-establishment identifiers to other LEHD infrastructure files.	1990–2004 (Years covered vary by state)
Unit-to-Worker	The unemployment insurance records underlying the LEHD infrastructure files provide neither establishment identifiers (except for Minnesota) nor industry or geographic detail of the establishment—only a firm identifier. Between 60 and 70 percent of state-level employment is in single-unit employers (employers with only one establishment) for which a link through the firm identifier is sufficient to provide such detail. For the remaining 30 to 40 percent of employment, such links have to be imputed. The Unit-to-Worker Impute (U2W) file contains ten imputed establishments for each employee of a multiunit employer. The file can be linked to other Census Bureau datasets through the PIK and the LEHD SEIN-SEINUNIT.	1990–2004

Appendix 5.
RESEARCH DATA CENTER (RDC) PARTNERS

Boston Census Bureau RDC
Wayne Gray, Executive Director

National Bureau of Economic Research

California Census Bureau RDC (Berkeley)
Jon Stiles, Executive Director

University of California, Berkeley

California Census Bureau RDC (UCLA)
David Rigby, Executive Director

University of California, Los Angeles

Census Bureau Headquarters RDC (CES)
Lucia Foster, Assistant Division Chief for Research, Center for Economic Studies

Administration for Healthcare Research and Quality
Bureau of Economic Analysis
Federal Reserve Board of Governors
National Center for Health Statistics

Chicago Census Bureau RDC
Bhash Mazumder, Executive Director

Federal Reserve Bank of Chicago
Northwestern University
University of Chicago
University of Illinois
University of Notre Dame

Michigan Census Bureau RDC
Margaret Levenstein, Executive Director

University of Michigan
Michigan State University

New York Census Bureau RDC (Baruch)
Diane Gibson, Executive Director[1]; Sanders Korenman, Executive Director[2]

Baruch College, City University of New York City University of New York
Columbia University
Cornell University (RDC administered by CISER, William Block; Director)
Federal Reserve Bank of New York
Fordham University
National Bureau of Economic Research
New York University
Princeton University
Russell Sage Foundation
Rutgers University
Stony Brook University, State University of New York
University at Albany, State University of New York
Yale University

New York Census Bureau RDC (Cornell)
Sanders Korenman, Research Director[3]; Warren Brown, Research Director[4]

Baruch College, City University of New York City University of New York
Columbia University
Cornell University (RDC administered by CISER; William Block, Director)
Federal Reserve Bank of New York
Fordham University
National Bureau of Economic Research
New York University
Princeton University
Russell Sage Foundation
Rutgers University
Stony Brook University, State University of New York
University at Albany, State University of New York
Yale University

[1] September 2008–present.
[2] January–August 2008.
[3] October 2008–present.
[4] January–September 2008.

Triangle Census Bureau RDC

Gale Boyd, Executive Director

Appalachian State University
Duke University
East Carolina University
Elizabeth City State University
Fayetteville State University
North Carolina Agricultural & Technical State University
North Carolina Central University
North Carolina State University
University of North Carolina at Asheville
University of North Carolina at Chapel Hill
University of North Carolina at Charlotte
University of North Carolina at Greensboro
University of North Carolina at Pembroke
University of North Carolina Wilmington
University of North Carolina School of the Arts
Western Carolina University
Winston-Salem State University

Appendix 6.
LONGITUDINAL EMPLOYER-HOUSEHOLD DYNAMICS (LEHD) PARTNERS

LOCAL EMPLOYMENT DYNAMICS (LED) STEERING COMMITTEE

New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont)
John Dorrer, Director
Center for Workforce Research and Information
Maine Department of Labor

New York/New Jersey
Leonard Preston
Labor Market Information
New Jersey Department of Labor and Workforce Development

Mid-Atlantic (District of Columbia, Delaware, Maryland, Pennsylvania, Virginia, West Virginia)
Tim Kestner
Economic Information Services
Virginia Employment Commission

Midwest (Illinois, Indiana, Iowa, Michigan, Minnesota, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin)
Rick Waclawek
Labor Market Information and Strategic Initiatives
Department of Labor and Economic Growth

Mountain-Plains (Colorado, Kansas, Missouri, Montana, Utah, Wyoming)
Alexandra E. Hall
Colorado Department of Labor and Employment

Southeast (Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee)
Robert Brown, Jr., Director
Labor Market Information
South Carolina Employment Security Commission

Southwest (Arkansas, Louisiana, New Mexico, Oklahoma, Texas)
Richard Froeschle
Labor Market and Career Information
Texas Workforce Commission

Western Region (Alaska, Arizona, California, Hawaii, Idaho, Nevada, Oregon, Washington)
Greg Weeks, Ph.D., Director (Cochair LED Steering Committee)
Labor Market and Economic Analysis
Washington Employment Security Department

FEDERAL PARTNERS

U.S. Department of Labor, Employment and Training Administration

U.S. Department of Health and Human Services, National Institute on Aging

STATE PARTNERS

State partner information as of October 2008.

Alabama
Jim Henry, Chief
Labor Market Information
Alabama Department of Industrial Relations

Alaska
Brynn Keith, Chief
Research and Analysis Section
Alaska Department of Labor and Workforce Development

Arizona
Dennis Doby, Senior Director
Research Administration
Arizona Department of Commerce

Arkansas
Vacant, Division Chief
Labor Market Information
Arkansas Department of Workforce Service

California
Steve Saxton, Chief
Labor Market Information Division

Colorado
Alexandra Hall, Labor Market Information Director
Labor Market Information
Colorado Department of Labor and Employment

Delaware
Thomas MacPherson, Director
Division of Unemployment Insurance
Delaware Department of Labor

District of Columbia
John Kangethe, Ph.D., Acting Director
Office of Labor Market Research and Information
District of Columbia Department of Employment Services

Florida
Rebecca Rust, Director
Labor Market Statistics Center
Florida Agency for Workforce Innovation

Georgia
Rosa Hayes, Director
Workforce Information and Analysis
Georgia Department of Labor

Hawaii
Naomi Harada, Chief
Research and Statistics Office
Hawaii Department of Labor and Industrial Relations

Idaho
Bob Uhlenkott, Bureau Chief
Research and Analysis
Idaho Department of Labor

Illinois
Evelina Tainer Loescher, Ph.D., Division Manager
Economic Information and Analysis
Illinois Department of Employment Security

Indiana
Hope Clark, Ph.D., Director
Research and Analysis
Indiana Department of Workforce Development

Iowa
Jude Igbokwe, Ph.D., Director
Workforce Data and Business Development Bureau
Iowa Workforce Development

Kansas
Dorothy Stites, Director
Labor Market Information Services
Kansas Department of Labor

Kentucky
Lelia K. Todd, Manager
Research and Statistics Branch
Kentucky Office of Employment and Training

Louisiana
Michael "Dino" DeMarte, Director
Research and Statistics Division
Louisiana Workforce Commission

Maine
John Dorrer, Director
Center for Workforce Research and Information
Maine Department of Labor

Maryland
Carolyn J. Mitchell, Director
Office of Workforce Information and Performance
Maryland Department of Labor, Licensing and Regulation

Michigan
Richard Waclawek, Director
Office of Labor Market Information and Strategic Initiatives
Michigan Department of Labor and Economic Growth

Minnesota
Steve Hine, Ph.D., Research Director
Minnesota Department of Employment and Economic Development

Mississippi
Mary Willoughby, Bureau Director
Mississippi Department of Employment Security

Missouri
William C. Niblack, Labor Market Information Manager
Missouri Economic Research and Information Center
Missouri Department of Economic Development

Montana
Todd Younkin, Chief
Research and Analysis Bureau
Montana Department of Labor and Industry

Nebraska
Phil Baker, Administrator
Nebraska Workforce Development—Labor Market Information

Nevada
Bill Anderson, Chief Economist
Research and Analysis Bureau
Nevada Department of Employment, Training, and Rehabilitation

New Jersey
Yustina Saleh, Director
Labor Market and Demographic Research
New Jersey Department of Labor and Workforce Development

New Mexico
Arthur Martinez, Chief
Economic Research and Analysis Bureau
New Mexico Department of Workforce Solutions

New York
Peter Neenan, Ph.D., Director
Research and Statistics Division
New York State Department of Labor

North Carolina
Elizabeth (Betty) McGrath, Ph.D., Director
Labor Market Information Division
Employment Security Commission of North Carolina

North Dakota
Duane Broschat, Manager
Labor Market Information Center
Job Service North Dakota

Ohio
Keith Ewald, Ph.D., Bureau Chief
Labor Market Information Bureau
Department of Job and Family Services

Oklahoma
Lynn Gray, Director
Economic Research and Analysis
Oklahoma Employment Security Commission

Oregon
Graham Slater, Administrator for Research
Oregon Department of Employment

Pennsylvania
Deep Gupta, Director
Center for Workforce Information and Analysis
Pennsylvania Department of Labor and Industry

Puerto Rico
Clarissa Muniz, Director
Labor Market Information/ Bureau of Labor Statistics
Puerto Rico Department of Labor and Human Resources

Rhode Island
Robert Langlais, Assistant Director
Labor Market Information
Rhode Island Department of Labor and Training

South Carolina
Robert O. Brown, Jr., Director
Labor Market Information
South Carolina Employment Security Commission
Rebecca Gunnlaugsson, Director
Division of Research
South Carolina Department of Commerce

South Dakota
Bernie Moran, Director
Labor Market Information Center
South Dakota Department of Labor

Tennessee
Joe Cummings, Director
Research and Statistics Division
Tennessee Department of Labor and Workforce Development

Texas
Mark Hughes, Director
Labor Market Information
Texas Workforce Commission

U.S. Virgin Islands
Gary Halyard, Director of Survey and Systems
Bureau of Labor Statistics
U.S. Virgin Islands Department of Labor

Utah
Stacey Floyd, Director
Workforce Information
Utah Department of Workforce Services

Vermont
Andrew Condon, Chief
Economic and Labor Market Information
Vermont Department of Employment and Training

Virginia
Donald P. Lillywhite, Director
Economic Information Services
Virginia Employment Commission

Washington
Greg Weeks, Ph.D., Director
Labor Market and Economic Analysis
Washington Employment Security Department

West Virginia
Ben Parker, Director
Research, Information and Analysis Division
Workforce West Virginia

Wisconsin
Linda L. Schultz, Labor Market Information Director
Division of Employment and Training
Department of Workforce Development

Wyoming
Tom Gallagher, Manager
Research and Planning
Wyoming Department of Employment

Appendix 7.
CENTER FOR ECONOMIC STUDIES (CES) STAFF LISTING 2008

CES Staff as of December 2008 in **bold.**

* Indicates contractor.

Name	Position
CES Senior Staff	
Jarmin, Ron	Chief Economist and Chief, CES
Atrostic, B.K.	Senior Economist and Special Assistant to the Division Chief
Foster, Lucia	Assistant Division Chief for Research
Holly, Brian	Project Review Coordinator
Kydd, Walter	Chief, LEHD Production and Development
Mildorf, Mark	Assistant Division Chief for Research Support
Riggs, T. Lynn	Lead RDC Administrator
Walker, George (Chip)	Communications and Marketing Consultant
Weng, Shigui	Chief, Data Staff
Wu, Jeremy	Assistant Division Chief for LEHD
Senior Research Fellows	
Abowd, John	Senior Research Fellow
Haltiwanger, John	Senior Research Fellow
CES Research Staff	
Akinyooye, Larry	Survey Statistician (detailed from Service Sector Statistics Division)
Bailey, Paul	Graduate Research Assistant
Becker, Randy	Senior Economist
Davis, Ronald	Research Assistant
Fort, Teresa	Graduate Research Assistant
Grim, Cheryl	Economist
Handley, Kyle	Graduate Research Assistant
Hayes, Natalie	Research Assistant
Herritz, Joshua	Research Assistant
Klimek, Shawn	Senior Economist
Krizan, C.J.	Senior Economist
McCue, Kristin	Economist
McInerney, Melissa	Graduate Research Assistant
Miranda, Javier	Economist
Nguyen, Sang	Senior Economist
Reznek, Arnold	Disclosure Officer
Tang, John	Economist
Zawacki, Alice	Economist

CES Research Support Staff

Chancellor, Jason	Data Staff, Survey Statistician
Gardner, Todd	Data Staff, Survey Statistician
Goodloe, Mike	Data Staff, Survey Statistician
Iceland, John*	Sociologist
Ryan, David	Data Staff, IT Specialist (Data Management)
Singal, Anurag	Data Staff, IT Specialist (Data Management)
Tsai, Ya-Jiun	Data Staff, IT Specialist (Data Management)
Wu, Xiaoyu*	Graduate Research Assistant
Yates, Michele	Data Staff, Survey Statistician
Yates, William	Special Assistant to the Assistant Division Chief for Research Support

Longitudinal Employer-Household Dynamics (LEHD) Program

Becker, Patti	Assistant to the Assistant Division Chief for LEHD

LEHD Quality and Research Staff

Andersson, Frederik*	Senior Research Associate
Chesser, David*	Economist
Ciccarella, Steve	Graduate Research Assistant
Garcia-Perez, Monica*	Graduate Research Assistant
Goetz, Chris	Graduate Research Assistant
Graham, Matthew	Geographer
Haywood, Heath	Geographer
Isenberg, Emily*	Economist
McEntarfer, Erika	Economist
McKinney, Kevin	Economist
Moody, Erin	Graduate Research Assistant
Pitts, Robert*	GIS Analyst Project Manager
Rosenblum, David*	Visiting Economist
Sandusky, Kristin	Economist
Sousa, Liliana*	Research Assistant
Strain, Michael	Graduate Research Assistant
Tibbets, Stephen*	Economist
Vilhuber, Lars	Economist

LEHD Production and Development Staff

Black, Bruce*	Task Manager
Li, Tao*	SAS Programmer
Ma, Cindy*	SAS Programmer
Mulato, Bong*	System Engineer
Norwood, Camille*	SAS Programmer
Perez, Claudia	Office Automation Technician
Zheng, Chaoling*	Webmaster

Research Data Center (RDC) Administrators

Andrus, Angela	Berkeley
Carter, J. Clint	Ann Arbor (Michigan)
Davis, James	Boston
Dragoset, Lisa	New York (Ithaca)
Grider, Bert	Research Triangle (North Carolina)
Hyson, Rosemary	New York (Baruch)
Limehouse, Frank	Chicago
McKinney, Kevin	Los Angeles (UCLA)
Riggs, T. Lynn	Lead/Washington, DC (CES Headquarters)
Schmutte, Ian	New York (Ithaca)
White, T. Kirk	Research Triangle (North Carolina)

Administrative Staff

Anderson, Dawn	Division Chief Secretary
Brown, Holly	Secretary to the Assistant Division Chief for LEHD
Burchinal, Jen	Secretary to LEHD Production and Development Staff
Cross, Henry	Student Intern
Schatzer, Ann	Secretary to the Project Review Coordinator
Turner, Rebecca	Secretary to the Assistant Division Chief for Research
Wright, Deborah	Secretary to the Assistant Division Chief for Research Support

Economic Directorate Administrative Office

Conley, Anita	Administrative Assistant
Dennison, Marilyn	Lead Financial Assistant
Estep, Tasha	Reimbursable Team Lead
Farri, Vicki	Reimbursable Team Lead
Kiatta, Cheryl	Administrative Officer
Oliver, Deborah	Lead Financial Assistant

Computer Services Division

Bishop, Eileen	Customer Relationship Manager, Headquarters
Johnson, Kyle	Customer Relationship Manager, RDCs
Lessard, James	Customer Relationship Manager, Headquarters
Razewski, David	Customer Relationship Manager, RDCs