issue brief

Insurance Choices: Behaviors of Firms and Their Workforces

Robert Wood Johnson Foundation

Background and Policy Context

The selection of Barack Obama as president, the change of administration, and the larger Democratic majority in Congress have once again brought health care reform to the forefront of the policy debate. There is a sense that rising health care costs are creating an intolerable burden for employers, who are increasingly unable to provide insurance for their employees while still remaining competitive. This challenge is exacerbated by the current slowdown in the economy. As a result, millions of Americans are uninsured and unable to obtain necessary health care, and that number is threatening to grow.

Building on the existing employer-based health care system, President Obama proposes to lower costs and make health insurance affordable and accessible to all Americans. The Obama-Biden health reform plan includes the following provisions:

- Requiring insurance companies to provide comprehensive coverage at fair and even premiums to all Americans, regardless of their health status or history;

- Creating a National Health Insurance Exchange offering comprehensive public or private coverage for individuals and small businesses who are seeking affordable health care options;

- Establishing tax credits for individuals and small businesses;

- Requiring large employers who do not provide health insurance coverage to contribute toward the costs of employee coverage;

- Mandating that all children have health care coverage; and

- Expanding eligibility for Medicaid and the Children's Health Insurance Program (CHIP).[1]

When trying to understand the potential impact of these and other proposed health care reforms, policymakers need to understand how firms will respond, as well as the likely responses of employees within the firms. For example, which firms or individuals will choose to purchase insurance for their employees, which will choose to contribute to an insurance pool, and what factors will influence those decisions? Data about the distribution of employees within firms are sparse. Detailed data about the wage structure and salaries within firms, as well as data on the characteristics of workers and their dependents, are not readily available. Where relevant data are available, they are often in disparate datasets with no obvious linkages between those that provide information about decisions of firms and those that examine individual choices. As a result, models estimating the impact of proposed reforms typically are based on assumptions about employer and employee responses, as well as assumptions about the behavior of the large group, small group, and individual insurance markets. As policymakers work to further develop and implement health care reform, they must identify the best data and models available to estimate the impacts of proposed interventions.

In an effort to explore the problems inherent in predicting the impact of proposed health insurance market reforms, AcademyHealth conducted a special meeting to identify the most important issues that need to be addressed under a variety of health care reform scenarios, as well as the data needed to address the questions. The meeting was supported under the Robert Wood Johnson Foundation's Changes in Health Care Financing and Organization (HCFO) initiative. In an informal, off-the-record, facilitated discussion, participants discussed:

- Key questions about the impact of health care reform proposals on firm and employee behavior;

- Data needed to answer key questions about the impact of health care reform proposals;

- The extent to which the needed data already are collected;

- The ability of existing datasets to be linked to answer the key questions;

- Data that are needed but not yet collected;

- A strategy for collecting and linking new data; and

- The ability of existing microsimulation models to address current health care financing issues.

Over the course of the meeting, participants became increasingly convinced of the need to develop specific behavioral models of employer and employee decisions under different health care reform scenarios. The questions are complex and the implications of health care reform are broad. There was consensus that it is important to consider both monetary interests and non-monetary factors that affect both employers' and individuals' decision-making. In the current environment of constrained budgets, participants expressed concern about the ability of federal agencies, insurers, and others to maintain and update existing data sources. These financial constraints also are likely to impede the collection of new data.

Participants agreed that increased efforts at linking existing data, including explorations of linkages of public and private data, are warranted and have the potential to result in powerful analytic files designed to address health care reform questions.

However, precautions will need to be taken to ensure individual and employer privacy and protect health plan proprietary information. In addition, meeting participants recognized that collaborative models, with transparent assumptions and components, represent a promising strategy for engaging more researchers in modeling efforts, allowing more comparability among analyses. However, current funding structures do not support the sharing of models, but rather encourage that they be treated as proprietary. As a first step toward promoting broad-based efforts to build data linkages and ultimately developing the complex models necessary to estimate the impact of health care reform, this issue brief highlights important questions about health care reform proposals that need to be addressed, as well as currently available data that might serve as a platform for those efforts to better understand the potential impact of health care reform efforts.

What Do We Need to Know About Health Care Reform Proposals?

As policymakers try to understand the potential impact of various proposals for reforming the health insurance market or the health care delivery system, the underlying dilemma they face is appreciating what will happen if Policy X is changed to Policy Y. Even if there are observations of behavior under Policy X, little is known about why those behaviors are observed. Therefore, it is nearly impossible to predict how those behaviors will change in response to Policy Y. Further complicating the situation, the policy changes

that will occur under health reform are likely to be transformative, and most economic theory—from which most modeling assumptions are derived—focuses on the effects of marginal changes. In addition, multiple policy changes are likely to occur at the same time, so policymakers need to understand the interactive effects of these changes, not just the effects of changing a single policy. The following are specific questions about which policymakers would value additional information.

Employer Behavior and Interaction with Employees

- **Why do employers offer health insurance? What are they trying to maximize?** Are employers attempting to actualize the economists' model of maximizing the welfare of workers and employers? Do employers offer insurance because it is part of a union agreement? Are the behaviors/responses of large firms the same as those of small firms? How do employers decide how much to contribute toward health insurance? Which employers will choose to continue/begin offering health insurance coverage and which will opt to contribute to the publicly offered/subsidized coverage under various reform scenarios?

- **What do employers know about their employees?** Does offering health insurance attract/retain employees? Does offering health insurance result in increased worker productivity? Are the behaviors/responses of workers in large firms the same as those in small firms? How do different subgroups (e.g., older, younger, immigrant, single, married, with or without dependents, etc.) of employees respond? Which employees are likely to opt not to accept an offer of health insurance and why? If health insurance were portable from one employer to another would employers/employees behave differently, given the different incentives? If employees had the option of obtaining insurance through a subsidized insurance pool, as an alternative to employer-based insurance, which option will they choose? What is the process of

actualizing the trade-off of wages and benefits that is mutually beneficial to the employer and employees?

Responses to Changes in Tax Incentives

- **How would the availability of a tax credit change the insurance offering behavior of small businesses and the purchasing behavior of families and individuals? How would this differ depending on whether the tax credit was for the individual, the employer, or both?** Will more small businesses offer insurance? Will young and healthy employees opt-out of employer coverage and purchase coverage in the non-group market? What are the non-group premiums individuals and families will face and what products will they be offered? Do individuals understand enough about health care and health insurance to make rational coverage decisions? What non-monetary factors will affect whether an individual or family chooses to purchase health insurance coverage? What do we know about the current non-group market in terms of purchasers, products offered, premiums, etc.? Will there be changes to state regulation of the non-group insurance market (e.g., community rating, cross-state purchasing, etc.)?

Reactions to a Connector/Exchange

- **What is the likely receptivity to a new government-run entity offering public coverage or subsidized public coverage for those who are uninsured or who opt-out of employer-based coverage?** What lessons can be learned from Massachusetts' health care reform that could be applied when predicting the national response to a similar intervention? How generalizable to other states is the Massachusetts experience? Who will opt to purchase from the public entity in the absence of an insurance mandate? What is the likely impact of this public entity on crowd-out (employers choosing not to offer or employees choosing not to accept pre-existing employer-based coverage)?

Responses to the Economic Downturn

- **How will employers and employees respond to health care reforms in light of the economic downturn?** Will employers drop, reduce, or change health insurance coverage to minimize costs? Will laid off employees seek coverage from the individual insurance market? Will high-deductible health plan or consumer-driven health plan uptake increase? Will the premium assistance subsidy for health coverage under COBRA—Consolidated Omnibus Budget Reconciliation Act of 1985—passed as a part of the American Recovery and Reinvestment Act of 2009 (ARRA) impact the uptake of COBRA?[2]

Data

Researchers and policymakers are in general agreement about the kinds of information and analyses that are needed to make the most appropriate policy recommendations. There is limited data available to adequately address many of the important questions identified; however, there are some datasets that can be utilized to begin to gather more information on the behavior of firms and individuals in the health care market. Examples of such datasets are listed below and in Appendix A. More detail about each survey, including information on potential data linkages and accessibility, can be found in Appendix A.

- *Agency for Healthcare Research and Quality (AHRQ): Medical Expenditure Panel Survey-Household Component (MEPS-HC)*—collects data from families and individuals on demographics, health status, insurance, utilization, and costs

- *AHRQ: Medical Expenditure Panel Survey-Insurance Component (MEPS-IC)*—survey of establishments conducted by the U.S. Census Bureau about the provision of health insurance benefits

- *Bureau of Labor Statistics (BLS): National Compensation Survey (NCS)*—surveys establishments on employer costs for wages, salaries, and benefits; locality occupational wages; and benefits plan incidence and provisions

- *Kaiser Family Foundation (KFF) and Health Research & Educational Trust (HRET): Employer Health Benefits Annual Survey*—collects information from firms on health insurance plans, provisions, and coverage

- *National Center for Health Statistics (NCHS): National Health Interview Survey (NHIS)*—surveys households on health status, access, employer information, health plan characteristics, and health conditions

- *Thomson Reuters™ Healthcare: Medstat MarketScan®*—has built a medical claims database of more than ten million persons with employer-sponsored health insurance with links to plan provisions

- *U.S. Census Bureau Business Surveys*—collects information from establishments on employment, total payroll, and sales. Some datasets also contain data on non-wage expenditures for employees. Also collects the MEPS-IC (see description above)

- *U.S. Census Bureau Household Surveys and Censuses [3]*

  - *American Community Survey (ACS)*–question on health insurance coverage added in the 2008 ACS. The first five-year estimates reflecting the collection of data on health insurance will be available in 2013. Data collected on demographic, social, and financial/economic characteristics

  - *Decennial Census*–short form: basic demographic data; long form: expanded demographic, socioeconomic and housing characteristics (ACS replaces the long form for the 2010 census)

  - *Current Population Survey (CPS)*–asks household respondents detailed questions for every household resident

  - *Survey of Income and Program Participation (SIPP)*–module for medical expenses and utilization of health care and module for employer-provided benefits

    – *U.S. Census Bureau Longitudinal Employer Household Dynamics (LEHD) Data*–allows construction of detailed employment histories for workers and details about employers and labor market characteristics

Data from these surveys might be used to address some questions about employer insurance offerings, employees within those firms, and the insurance market. Researchers, for example, could use Medstat data to analyze firm and health plan behavior toward high risk individuals, examining questions such as: how do employers and health plans treat high risk individuals and would employers and health plans treat high risk individuals differently if they were no longer responsible for providing high risk employees with coverage? Additionally, researchers could use the existing data to conduct longitudinal studies, potentially focused on high cost patients who are less than 65 years of age, to enhance understanding about how the individual market works.

Although the existing data can be used to answer some of the research questions posed in the previous section, the scope and sample of each survey differ and each survey may lack adequate information for answering many of the specific questions on firm and workforce behavior. The MEPS-HC, for example, provides information about whether employees are eligible for health insurance, and if firms offer insurance, information on why some employees are not eligible; however, it lacks information about insurance offers that employees turn down. Therefore, examining whether health insurance helps attract or retain employees—which may require information on employers, health benefit offerings, employee contributions, employee age, preferences, income, and tenure—would require the use of multiple datasets.

Survey samples may also impede researchers' efforts to examine the behavior of firms and their employees. The BLS National Compensation Survey, for example, allows

cross-sectional analyses of wages, salaries, and benefits across establishments, but prevents the examination of trends using current employment data, as would be involved in the analysis of staffing patterns. Moreover, some surveys sample the individual (e.g., NHIS, MEPS-HC), while other surveys sample the establishment (e.g., U.S. Census Bureau Business Surveys, MEPS-IC) or the firm (e.g., Medstat, KFF/HRET Employer Health Benefits Annual Survey), making it challenging to link the data to assess the composition and distribution of workers within firms. Furthermore, some surveys only collect data from large employers (e.g., Medstat).

Researchers are further impeded by a lack of publicly available data. Many federally-sponsored data are contained in restricted use files to ensure the privacy of survey respondents. Public use files often lack the rich data required for studying firm and workforce behavior. For example, many public use files, such as the NHIS public use files, lack geographic identifiers which are important for examining questions related to labor markets and health insurance.[4] While state identifiers are available on the restricted use files of the NHIS, researchers are required to request permission to access the restricted use files at the NCHS Data Center or the U.S. Census Bureau's Research Data Centers. Moreover, some datasets, such as the LEHD and the MEPS-IC, can only be accessed at the Census Bureau Research Data Centers with prior authorization by the sponsoring agency.

Linkages between surveys and administrative data

There is no single dataset available to provide the scope of information necessary to study and predict firm and workforce behavior; however, linking information from different surveys or imputing synthetic data to address missing values or to prevent public disclosure of public use files may enhance the information available for answering select research questions. Existing linkages between surveys have enhanced

the richness of information available and the analytic power of the data. For example, researchers supplemented the MEPS-HC with benchmarks from the National Health Expenditure Accounts, produced by the Centers for Medicare & Medicaid Services (CMS), to develop a dataset that captures more of the nation's spending. Researchers have also linked U.S. Census Bureau Business Surveys with MEPS-IC to examine the impact of health insurance offers on worker productivity; however, comparability across industries is not necessarily straightforward. In addition to the data linkage efforts currently underway, meeting participants suggested synthetically or statistically linking the MEPS-IC with NHIS (which would combine data on employers and their employees into one dataset) and continuing efforts to link the MEPS-HC with the MEPS-IC to create actual firms.

Creating linkages between surveys and administrative data, such as Social Security, Medicare, and Medicaid claims data, may improve the breadth of information available to researchers and the accuracy of survey instruments. Meeting participants discussed the challenges in obtaining adequate income data due to low survey response rates, and suggested linking existing surveys with Social Security income data to glean necessary income information. In addition, researchers have linked the MEPS-IC to the U.S. Census Bureau's LEHD program data, which provides richer workforce data, including gender and earnings, but lacks information on marital status. Currently, federal agencies— including CMS, the Department of Health and Human Services (DHHS) Office of the Assistant Secretary for Planning and Evaluation (ASPE), and the U.S. Census Bureau—in conjunction with the University of Minnesota State Health Access Data Assistance Center (SHADAC) are linking the NHIS with Medicaid Statistical Information Statistics (MSIS) to compare Medicaid program participation reported in the survey with program estimates, allowing researchers to measure and improve the accuracy of the survey instrument.[7]

While the potential for data linkages to provide more comprehensive information is promising, there are several challenges impeding the development of such linkages. Because data linkages increase the amount of information available about individual or employer respondents, linked datasets—especially the linkage of individual data to administrative records—may inadvertently allow respondents to be identifiable. To balance privacy concerns with the need for publicly available data, federal agencies may be able to use partially synthetic data in public use files, while maintaining restricted access to original data through a research data center. Moreover, as surveys are sponsored by different federal agencies or private organizations, linking datasets from different entities requires collaboration, cooperation, and reconciliation of different privacy standards and ownership rights.[8]

Constructing synthetic workforces for firms

There is currently a dearth of public data on the composition of firms, making it challenging to predict firm and workforce behavior to health care reform. To model how different types of firms and workers would react to reform, researchers could survey a small sample of firms to obtain 'real' data. While collecting real data is ideal, it would be expensive. To address this data limitation, researchers have created synthetic firms that are populated with individuals who share similar characteristics, such as geographic location. Thomas Selden, Ph.D., and Bradley Gray, Ph.D., for example, synthetically populated establishments surveyed in MEPS-IC with individuals surveyed in MEPS-HC to estimate the value of employer-sponsored tax subsidies.[9]

The lack of data on the composition of firms is exacerbated for small firms. As such, synthetic firms may serve as an appropriate proxy for real data and allow for immediate analysis. Constructing small synthetic firms may be more challenging

than constructing large synthetic firms. Large firms are likely representative of the covered population, which would allow researchers to randomly assign individuals to a firm using data from a nationally representative dataset of large firms, like Medstat. The workforce in small firms, on the other hand, is more heterogeneous, making it challenging to accurately populate the firms or establishments. Moreover, we do not know the extent to which employees with poor health status sort themselves into the same firm. Actual data on the composition of small firms that allows researchers to examine the distribution of health risks of employees within a firm and the value of health insurance to employees, may help to understand the extent of sorting. This effort could lay the groundwork for the development of a "gold standard" for creating small synthetic firms.[10]

A national, publicly available dataset of synthetic firms that reflects the current market environment could facilitate future research and modeling efforts by a large community of researchers. This dataset could use a contributory framework to allow researchers to collaborate to ensure that they were correctly sorting workers into synthetic firms. Ideally, this dataset would contain publicly available person and firm identifiers as well as information at the state level. Such a dataset would aid in future efforts to examine employer and employee response to health reform.

Future Data Collection Efforts

Developing a large employer/employee linked dataset with individual and firm characteristics, including those listed below, would be ideal, but would be a huge effort and investment.

Meeting participants identified a number of alternative ways to increase the availability of data. Researchers could synthesize a new national dataset that imputes missing values from existing surveys or could create a national dataset of synthetic firms with administrative data attached to each observation. Alternatively, future data collection efforts could expand upon or update existing surveys, such as the last survey of employee benefits for small businesses, conducted in the early 1990s.[11] Expansions to existing surveys could include questions that obtain data related to the following:[12]

- Attitudes toward the value of health insurance and non-wage compensation;

- Composition of workforces within a firm;

- Distribution of workforce characteristics, including claims experience;

- Employee attitudes toward risk;

- Employee skill level;

- Health expenditures, health status, and health insurance offers for each member of employees' families;

- Immigration status;

- Income;

- New types of employment arrangements, such as leased employees;

- Premiums faced by individuals in the non-group market, including those who purchase and those who do not; and

- State microdata on employers and labor markets.

Efforts to expand and update existing datasets and to create new datasets are severely hampered by funding. Federally sponsored

Priorities for Future Data Collection Efforts

• Employer contributions • Family status • Firm size • Health insurance options and choices	• Health status • Socioeconomic status • Wage distribution within the firm • Wages

datasets are supported by soft money, and organizations can pay the marginal cost to add additional survey questions. Even if survey funding was sufficient to include all questions of interest, the increased length of the survey would likely deter individuals and employers from completing the survey.

Enhancing Health Care Reform Modeling

Models estimating the impact of proposed reforms typically are based on assumptions about employer and employee responses, as well as assumptions about the behavior of the large group, small group, and individual insurance markets. Natural experiments that evaluate employer and employee response to insurance reform, however, will likely inform future model assumptions. When predicting the effect of health care reform in the state of Massachusetts, for example, researchers used utility maximization models, assuming that employers would respond to reform in a manner that maximized their profits. By examining what happened in Massachusetts relative to the model, researchers may be able to gain a better understanding of firm and household behavior. Modeling how firms and employers respond to health care reform in light of the economic crisis, however, will be challenging as there is no historical precedent upon which to base model assumptions.

Using real-world experiences, like that of Massachusetts, to test the adequacy of existing models will likely inform future federal and state-level modeling efforts and the extent to which we can infer from state models. While the model assumptions may be similar, each state has different market and workforce characteristics. Therefore, state models may be adapted using state-level data and weights that reflect the characteristics of that particular state. Such modeling efforts may assist policymakers

in identifying potential unintended consequences that result in each state from the suggested policy levers.

Conclusion

As policymakers work to further develop and implement health care reform, they must identify the best data and models available to estimate the impacts of proposed interventions. Creative thinking about how best to use available data, create linkages among existing databases, and use limited resources strategically will permit the development of models that better predict the interactions of players in the health care market under a variety of scenarios.

Developing a baseline model upon which all modeling efforts could evolve would inform the underlying assumptions and validity of future modeling efforts. These efforts, however, would require consensus around a single theory of firm behavior, which does not exist. Therefore, the research community should acknowledge and emphasize open discussion about the different models and their underlying assumptions and triangulate toward predictions rather than focus on a single end point. Open source collaboration of models would allow researchers to access, collaborate, share, and discuss model developments. While not financially feasible, particularly for organizations whose models are proprietary, open source collaboration between private and public organizations has the potential to improve the breadth of data available to the public and expedite and improve the accuracy of analysis of these important research questions.[13]

About the Authors

Jenny Minott is an associate with the HCFO program. She can be reached at jenny.minott@academyhealth.org. Deborah L. Rogal is a director with AcademyHealth, and can be reached at deborah.rogal@academyhealth.org. Both can also be reached at 202.292.6700.

Endnotes

1   "The Agenda: Health Care," accessed at www.whitehouse.gov/agenda/health_care/. President Obama signed legislation expanding children's health insurance on February 4, 2009.

2   The American Recovery and Reinvestment Act of 2009 was signed into law February 17, 2009. See www.govtrack.us/congress/bill.xpd?bill=h111-1 for a summary of this law.

3   See www.census.gov/hhes/www/hlthins/hlthins.html for information available on health insurance from demographic data at the U.S. Census Bureau.

4   Garrett, B. and M. Chernew. "Health Insurance and Labor Markets: Concepts, Open Questions, and Data Needs," *Inquiry*, Vol. 45, No. 1, Spring 2008: pp. 30-57.

5   See www.cms.hhs.gov/NationalHealthExpendData/ for more information on CMS' National Health Expenditures Accounts.

6   Researchers linked the MEPS-HC and MEPS-IC, but were hampered by low response rates. Therefore, linked files are only available for 1996-1999 and 2001. See http://meps.ahrq.gov/mepsweb/data_stats/onsite_datacenter.jsp?prfricon=yes for more detail.

7   Cox, C. "National Center for Health Statistics Office of Analysis and Epidemiology Special Projects Branch Record Linkage Program," Presentation at National Center for Health Statistics Board of Scientific Counselors Meeting, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, April 24, 2008. www.cdc.gov/nchs/ppt/bsc/bsc_datalink_apr08.ppt#276,16,Publications & Current Projects Using Linked Medicare Data

8   ibid.

9   Selden, T.M. and B.M. Gray. "Tax Subsidies for Employment-Related Health Insurance: Estimates for 2006," *Health Affairs*, Vol. 25, No. 6, November/December 2006, p: 1568-79.

10 Garrett, B. and M. Chernew. "Health Insurance and Labor Markets: Concepts, Open Questions, and Data Needs," *Inquiry*, Vol. 45, No. 1, Spring 2008: pp. 30-57.

11 The Small Business Benefit Survey was conducted in 1992 to 1993 and surveyed over 2,000 small employers (<25 employees) in seven cities across the United States. Small employers were asked questions related to small business and employee health benefit coverage, wages, and attitudes towards proposed health care reforms. See McLaughlin, C.G., et al. "Small-Business Winners and Losers Under Health Care Reform," *Health Affairs*, Vol.13, No. 2, 1994.

12 Meeting discussion; Garrett, B. and M. Chernew. "Health Insurance and Labor Markets: Concepts, Open Questions, and Data Needs," *Inquiry*, Vol. 45, No. 1, Spring 2008: pp. 30-57.

13 Tapscott, D. and A.D. Williams. *Wikinomics: How Mass Collaboration Changes Everything.* Portfolio: New York, 2008.

Appendix A

Data	Type of Data	Focus	Sample	Information Collected	Potential Linkages	Access
American Community Survey *U.S. Census Bureau* www.census.gov/acs/www/index.html	Annual Replaces 2010 Census long form survey	Population and housing information	Geographic areas with a population of 65,000 or more; additional information forthcoming for smallest areas and groups in 2013	• Disability • Family relationships • Food stamp use • Health insurance coverage (added in 2008) • Housing costs • Income	Statistical matches with U.S. Census Bureau data	Data will be available in 2010 as public release data products, and microdata will be available at Census Bureau Research Data Centers with prior authorization.[1]
Economic Censuses www.census.gov/econ/overview/mu0000.html	Quinquennial surveys and administrative data	Industry data	Establishments surveyed in most industries; only administrative data for smallest businesses	• Core data on employment, payroll, and sales for all industries • Some data specific for particular industry		
Business Surveys *U.S. Census Bureau* www.census.gov/	**Business Register** (previously the SSEL – Standard Statistical Establishment List) www.census.gov/econ/overview/mu0600.html Annual listing constructed from administrative and survey sources	Employer data	List of all private establishments with paid employment	• Industry, employment, and payroll • Receipts or sales for single units and unallocated sales/receipts for multi-units	• MEPS-IC • LEHD • Other Census Bureau Business Data Statistical matches • MEPS-HC • CPS • SIPP • Decennial • ACS	Data are available at U.S. Census Bureau Research Data Centers with prior authorization.
	Longitudinal Business Database www.ces.census.gov/index.php/ces/researchdata?detail_key=10 Constructed from the Business Register	Employer data	List of all private establishments with paid employment	Firm-level employment and firm age constructed using Business Register measures		
Current Population Survey (CPS) – Annual Social and Economic Supplement (ASEC) *U.S. Census Bureau* http://www.census.gov/cps/	March ASEC/Supplement	Labor force characteristics	~78,000 households	• Demographic • Detailed health insurance information collected from household respondent for every household resident • Earnings • Employee benefits • Employment/unemployment • Hours of work • Income • Previous work experience • School enrollment • Self-reported health status and disability • Work schedules	Statistical matches to U.S. Census Bureau Business Surveys	Available online and researchers may request permission to use the restricted-use files at the U.S. Census Bureau's Research Data Centers.

Appendix A

Data	Type of Data	Focus	Sample	Information Collected	Potential Linkages	Access
Decennial *U.S. Census Bureau* http://factfinder.census.gov/	Point-in-time data; conducted every 10 years from 1940–2000	Demographic, housing, and socioeconomic characteristics	Sample of 1 in 6 households receive long form; everyone receives short form	Short form: basic demographic data Long form: expanded demographic, disability status, housing, and socioeconomic housing Long-form data includes: • Census block geographic data for place of work and place of residence (also available in the short-form data) • Disability status, including separate measures for vision/hearing impairment, limited mobility, difficulty performing mental tasks, difficulty with self-care and movement at home, difficulty with mobility to go outside home alone, difficulty in ability to work at a job or business • Immigrant status, language ability • Income • Marital status • Occupation and three-digit industry	Statistical matches to U.S. Census Bureau business/ employer surveys	1970 onward: data are available at Census Bureau Research Data Centers with prior authorization.
Employer Health Benefits Annual Survey *Kaiser Family Foundation (KFF)/Health Research and Educational Trust (HRET)* ehbs.kff.org/?page=abstract&id=1	Nationally representative survey with random sample	Employer-sponsored health benefits	Public and private large and small employers	• Cost of health insurance • Covered benefits, including prescription drugs • Demographics • Dependent eligibility • Employee contributions • Employer cost-sharing • Employer opinions • Employer preference for wages versus benefits • Enrollment patterns • Insurance eligibility • Offer rates • Premium increases • Reasons for not offering health insurance • Retiree health benefits • Types of plans • Wellness programs	N/A	Public use tape
Longitudinal Employer-Household Dynamics (LEHD) *U.S. Census Bureau* lehd.did.census.gov/led/	Longitudinal quarterly administrative data; unit of analysis is job or person/employer matching	Workforce dynamics, human capital, labor market characteristics	Universe of workers	• Worker: can construct detailed employment history (e.g., if working multiple jobs, tenure, etc.) and demographic characteristics of the co-worker pool • Employer: information on industry, location, human capital measures, churning, and jobs created/ destroyed by gender, and age	• Census Bureau Business Surveys • MEPS-IC	Data are available at Census Bureau Research Data Centers with prior authorization.

Appendix A

Data	Type of Data	Focus	Sample	Information Collected	Potential Linkages	Access
Medical Expenditure Panel Survey (MEPS)-Household Component (HC) *Agency for Healthcare Research and Quality* www.meps.ahrq.gov/mepsweb/survey_comp/household.jsp	Annual, nationally representative survey; panel survey for two years	Health insurance and health conditions	Families and individuals	• Access to care • Demographics • Employment • Health conditions • Health insurance costs • Health insurance coverage • Health status • Income • Medical charges • Medical utilization • Satisfaction with care • Socioeconomic characteristics • Source of payment	• NHIS Statistical/synthetic matches: • MEPS-IC • U.S. Census Bureau Business Surveys	Public files are available online, on CD-ROM, and at the AHRQ Data Center. Researchers may request permission to use the restricted use files at the AHRQ Data Center or U.S. Census Bureau's Research Data Centers.
Medical Expenditure Panel Survey (MEPS)-Insurance Component (IC) *Agency for Healthcare Research and Quality* www.meps.ahrq.gov/mepsweb/survey_comp/Insurance.jsp	Annual, nationally representative survey; cross-sectional sample	Employer-sponsored health insurance	Private and public sector establishments	• Access to employer-sponsored insurance • Benefits included in plans • Enrollment • Employer characteristics • Employer/employee contributions • Non-health fringe benefits • Number of health plans offered • Out-of-pocket costs • Premiums • Provider arrangements • Retiree health benefits • Type of health plan offered	• LEHD • Census Bureau Business Surveys Statistical/synthetic matches • MEPS-HC • Census Bureau household data	Data is not available for public use. Researchers may request permission to use data at U.S. Census Bureau Research Data Centers.
Medstat MarketScan®2 *Thomson Reuters™* www.medstatmarketscan.com/	Convenience sample; sample is large enough to comprise a nationally representative sample of people with employer-sponsored insurance	Privately insured individuals	Large employers	• Claims data for the full continuum of care • Diagnosis and procedure codes • Employer provided data • Outpatient prescriptions and patient copayments • Payment/charge information • Prescription fills	N/A	Proprietary

Appendix A

Data	Type of Data	Focus	Sample	Information Collected	Potential Linkages	Access
National Compensation Survey *U.S. Bureau of Labor Statistics* www.bls.gov/ncs	Nationally representative survey with probability sample	Employer costs for wages and benefits; occupational wages; and incidence and provision of selected employee benefit plans	Private industry and state and local governments; establishments; occupation is the unit of observation	• Access to health insurance benefits: take-up rates, participation rates, who is offered plans • Employer costs per hour worked for employee benefits • Employer costs per hour worked for employee wages and salaries • Health insurance provision: summary of insurance plan, type of plan, limitations on coverage, distribution, vision, dental, mental/substance abuse • Information available for a variety of characteristics: all workers, major occupational and industry groups, size of the establishment, union/non-union and full-time/part-time status, hourly wage percentiles • Occupational pay • Premiums: single and family premiums, employer/employee cost sharing arrangements, average employer health care costs in dollars and as a percent of total compensation, employer cost sharing	OES	Published data are available online and in printed news releases and reports. Unpublished data are available onsite at the Bureau of Labor Statistics. Researchers must submit a proposal for review before receiving permission to use the data. (www.bls.gov/bls/blsresda.htm#availabledata)
National Health Interview Survey (NHIS) *National Center for Health Statistics* www.cdc.gov/nchs/nhis.htm	Nationally representative, cross-sectional survey with multistage area probability design	Health status, utilization, and insurance	Households	• Assets • Health care behavior • Health conditions • Health insurance • Health services, utilization, and access • Health status • Income • Measures of uninsurance • Socioeconomics and demographics • Some employer-related information • Some health plan characteristics	• Mortality data • Medicare enrollment and claims data • Social Security Benefit Data • MEPS-HC • Medicaid	Some public use files available online; others available at data research centers. Researchers may request permission to use restricted-use files at the NCHS Data Center or U.S. Census Bureau's Research Data Centers.
Occupational Employment Statistics (OES) *U.S. Bureau of Labor Statistics* www.bls.gov/OES/	Panel survey; sample data collected every six months	Occupational employment and wage estimates; based on 6 panels (3 years)	1.2 million establishments	• Data produced: mean hourly wage, percentile wage, and annual wage • Data useful for analysis of staffing patterns • Number of employees paid within specific wage intervals	• National Compensation Survey • QCEW	Available online

Appendix A

Data	Type of Data	Focus	Sample	Information Collected	Potential Linkages	Access
Quarterly Census of Employment and Wages *U.S. Bureau of Labor Statistics* www.bls.gov/cew/cewover.htm	Reported quarterly	Employment and wages	Not a sample: universe count of employment covered by unemployment insurance laws 9 million establishments	• Employment • Number of establishments • Quarterly wages	OES	Available online
Survey of Income and Program Participation (SIPP) *U.S. Census Bureau* www.census.gov/sipp/overview.html	Longitudinal panel survey lasting 2.5 to 4 years	Income and public program eligibility	Households	• Cash and non-cash income • Demographics • Disability • Health care costs • Health insurance coverage including employer-provided health benefits with data on who is covered, premium cost, plan details, employer, and industry • Health status • Labor force information • Long-term care • Medical expenses • Public program coverage • Public program eligibility • Utilization	Potential statistical matches to other U.S. Census Bureau data	Available online and researchers may request permission to use the restricted-use files at the U.S. Census Bureau's Research Data Centers

Footnotes:
1 The following link provides information to researchers interested in accessing confidential data at the nine Census Bureau Research Data Centers located around the country. www.ces.census.gov/index.php/ces/researchprogram
2 Adamson, D.M., et al. "Health Research Data for the Real World: The MarketScan® Databases," *White Paper*, Thomson Reuters™, July 2008.

Additional sources: Hyson, R. and A. Zawacki. "Health-Related Research," 2007 Research Report: Center for Economic Studies and Research Data Centers, U.S. Census Bureau, October 2008.