
LONGITUDINAL EMPLOYER - HOUSEHOLD DYNAMICS

TECHNICAL PAPER NO. TP-2002-03

New Uses of Health and Pension Information The 5500 file at the Census Bureau

Date : January 2002
Prepared by : Julia I. Lane and staff from the U.S. Census Bureau
Contact : Ronald Prevost (Ronald.C.Prevost@census.gov)
U.S. Census Bureau, LEHD Program
FB 2138-3
4700 Silver Hill Rd.
Suitland, MD 20233 USA

This document reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications, and is released to inform interested parties of ongoing research and to encourage discussion of work in progress. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grant SES-9978093 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author(s) and do not represent the views of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau is preparing to support external researchers' use of these data; please contact Ronald Prevost (Ronald.C.Prevost@census.gov), U.S. Census Bureau, LEHD Project, FB 2138-3, 4700 Silver Hill Rd., Suitland, MD 20233, USA.

Introduction

The accurate measurement of employer-provided health and pension plan coverage is critically important for policy makers seeking to make state, local and national decisions that will affect the wellbeing of the current and future infirm and elderly population. However, current information could be vastly improved by improving estimates of coverage and by bettering the firm level information currently available. We propose to exploit an underutilized resource – the 5500 file – to achieve this improvement. The 5500 file, which contains data from annual returns filed annually with the Internal Revenue Service by all employers covered by ERISA, provides information on employee benefit plans, including data on different types of pension and health insurance plans.

These data could potentially help states in two ways: by providing information on how much of the workforce is covered through private health and pension benefits and by providing information on job “quality”. In the former case, the cost of providing state services to the vulnerable population is often driven by whether or not workers are eligible for private health care. Data on this coverage are scarce – the main source, CPS, is plagued with inconsistencies (see below). In addition, economic development agencies need information, not just on job creation and job destruction, but on the “quality” of jobs that are entering and leaving a state. The “quality” of a job is often measured by whether the employer offers health and pension benefits. However, state and local economic development authorities typically have no information on whether The 5500 file, in combination with LEHD data, can help individual states identify what kinds of jobs are offered, what kinds of workers are covered, and how this is changing over time.

The 5500 file can help the Census Bureau in a similar fashion. It can improve coverage estimates because it help reconcile inconsistencies in current data collection. Two major Census Bureau surveys -- the Current Population Survey (CPS) and the Survey of Income and Program Participation (SIPP) -- produce substantially different estimates of the uninsured.¹ Similar problems surround the collection of private pension information. A large part of the problem is due to that fact that respondents are typically not able to give accurate responses. For example, the health insurance information collected by the CPS is meant to represent coverage in the prior year, yet it is possible that the CPS actually provides a point-in-time estimate for some respondents. The 5500 file may help shed some insight on what the CPS actually seeks to collect. In addition, the CPS and SIPP each use a different set of questions to collect health insurance information, and the 5500 file may help to determine which set of questions produces a more accurate estimate of employer-provided health insurance. Finally, the 5500 file could be used to add information to the American Community Survey (ACS) to provide broader measures of health and pension coverage.

However, there are a number of important and difficult quality and coverage issues that need to be addressed before full use can be made of these data. The next sections provide background on employer-sponsored benefits and detail both the potential projects and the detailed data work that

¹Bennefield (1996).

could be undertaken on a pilot basis to evaluate the possible contribution of the 5500 file to the statistical data infrastructure.

Background

Employer-sponsored health insurance and pension plans play key roles in the economy. About two-thirds of non-elderly Americans received health insurance through employers in 1998 (Holahan and Kim 2000), and about four-fifths of full-time workers in medium and large firms participated in pension plans in 1997 (U.S. Bureau of Labor Statistics 1999). These pension plans provide retirement security for many workers in later life. In 1998, for example, 45 percent of Americans ages 65 and older received pension income from past employers or were married to someone who did. In addition, private pensions accounted for 26 percent of income for retirees with pensions (Social Security Administration 2000).

Health and pension benefits also have important effects on the labor market. The availability of employer-sponsored health benefits can discourage turnover by workers (Madrian 1994). Many defined benefit pensions also provide workers with incentives to remain with their employers until they are eligible to collect pension benefits and then encourage retirement once benefits become available. More generally, employee benefits are an important and growing component of compensation and an accurate measure is necessary to gauge the returns to work.

Despite the importance of employee benefits, there is still much we do not know about why firms offer health insurance and pension coverage and how they set plan parameters. Improving our understanding of these issues is critical to evaluating proposals to change the current system. Researchers have offered possible explanations for why employers offer pensions to their workers, but there is no consensus about which factors are most important to the sponsorship decision. For example, some have argued that pensions provide an inexpensive way for firms to compensate higher-wage workers because pensions are a tax-free fringe benefit (Munnell 1982). Ippolito (1985) has suggested that firms offer pensions in order to elicit cooperative behavior from labor unions. Other economists maintain that a primary function of pensions is to bond workers to their employers (Allen, Clark, and McDermed 1993; Ippolito 1991), or that they create retirement incentives that induce older, typically high wage workers to leave the firm (Kotlikoff and Wise 1989; Mitchell and Fields 1982). The retirement incentives created by DB pension plans may be particularly important because internal labor markets and age discrimination laws may limit the ability of employers to dismiss older workers.

Pensions may also be offered because they can enable employers to elicit effort from workers when productivity is difficult to monitor (Hutchens 1987, 1989; Lazear 1979, 1983). Workers who are dismissed after being caught shirking will incur large penalties if their career compensation packages were backloaded, paying them less than their marginal products early in the career but more than their marginal products late in the career. The threat of severe penalties can deter workers with rising compensation profiles from shirking. Pension plans can also provide employers with an efficient means of deferring compensation.

Taxes obviously provide an incentive to offer compensation in the form of pensions and other nontaxed fringe benefits, but the way in which taxes affect the decision to offer pensions has been

relatively unexplored in the economic literature. As a result, there is little evidence on the impact on pension coverage of changes in tax rules, such as limits on contributions to employer pensions and the availability of tax subsidized alternatives such as individual retirement accounts. Indirect evidence on the latter point might be gleaned from the findings of Poterba, Venti, and Wise (1996) that most IRA contributions represent new saving, rather than shifting from other accounts. Under this model, one might expect the availability of IRAs to have little or no effect on participation in and thus the demand for employer pension plans. But Engen, Gale, and Scholz (1996) reach the opposite conclusion with an equally plausible model and the issue remains unresolved.

The cost savings that employer-sponsored health insurance offer to workers are often cited to explain the provision of health benefits by firms. Health benefits received from an employer are tax free, while the cost of insurance purchased in the marketplace is generally subject to income taxes. In addition, employer-sponsored insurance provides workers with access to group rates, which are generally substantially lower than the rates available in non-group plans, especially for those with health problems (Chollet and Kirk 1998).

Potential Projects

The specific projects that could be undertaken are:

Survey Improvement

1. Performing quality checks on CPS/SIPP/MEPS surveys
2. Investigating the possibility of adding employer contributed health care variables to CPS/SIPP/ACS
3. Investigating the possibility of adding employer provided pension variables to CPS/SIPP/ACS
4. Reconciling differences in health care coverage across surveys

Survey Frame Improvement

1. Adding 5500 file health and pension benefit information to the Census Business Register
2. Using selected financial asset and income information to inform the Census Business Register

State Data Enhancement

1. Adding health coverage information to the Employment Dynamics Estimates that are returned to the states
2. Adding private pension coverage information to the Employment Dynamics Estimates that are returned to the states.

Research Projects

1. Examining what kinds of firms offer different types of health insurance/pension plans – particularly by size class, industry and legal form of ownership.

2. Examining how firms change coverage types
3. Examining how workers at firms offer health insurance/pensions differ from workers whose firms do not – and how this has evolved over time.
4. Examining whether firms that offer health insurance/pensions have lower turnover rates than firms that do not – and how this has evolved over time.
5. Examining differences in turnover rates for different types of workers at firms with health insurance/pension availability. For example, what are the effects of health insurance on the job mobility of a low-wage worker, relative to a high wage worker?
6. Examining the characteristics of the ‘employed uninsured’, particularly those working for firms which offer health insurance.
7. Understanding whether uninsured workers are uninsured because they don’t have the option (from the employer) of coverage, or because they choose not to take up the coverage.
8. Identifying the true effect of health insurance on labor market outcomes (wages, hours worked, etc).

Preliminary Results from the 5500 File

Preliminary work with the 5500 file has been encouraging. A “first pass” match of the 1997 5500 file with the 1997 Census Business Register revealed that 71.9% of companies on the Register with 100 or more employees matched 5500 establishment companies. These matched companies contain 88.7% of establishments, 87.8% of employment and 91.9% of annual payroll (of companies with more than 100 employees). This match exploited parent/subsidiary information which is ONLY available at the Census Bureau.

In addition, simple tabulations of the 1997 pension plan information give some idea of the richness of the data.

Number of Pension Plans and Total Participants		
Type of Plan	Number of Plans	Total Participants (thousands)
Total	720,041	94,985
Defined benefit	59,499	40,392
Defined Contribution	660,542	54,593
Profit Sharing	516,342	44,342
Stock Bonus	4,410	3,612
Target Benefit	5,860	171
Money Purchase	108,203	5,120

Adapted from Table A1: Private Pension Plan Bulletin, US Department of Labor, PWBA, Number 10, Winter 1997, Final 1997

Coverage and Quality Issues: what needs to be done

There are a number of questions that need to be researched in some detail.

1. What types of firms file the 5500?

Virtually all employers with pension plans must file the 5500 return, regardless of the tax advantages claimed under Internal Revenue Code Section 401a. All tax-qualified plans be reported on 5500 forms. Some non-qualified plans, particularly top-heavy plans, are not reported. Commercially insured as well as self-insured health plans are regulated under ERISA. States can indirectly affect benefit package and other plan provisions of commercially insured plans through their regulation of insurance companies that operate in the state. Both commercially and self-insured health plans are also reported on form 5500 file.

2. What are the gaps in reporting?

The most important gap in reporting is for unfunded or fully-insured health plans with fewer than 100 participants. There are something like 2.5 million small employers that are missed, with on the order of 20 million participants (these are SBA estimates). *Understanding size/characteristics of non-filing plans is an important research issue. We plan to begin by using the SBA methodology, and supplement this with information on the Medical Expenditure Panel Survey (MEPS).*

It is difficult to identify MEWAs: small groups of unrelated employers who buy a group policy or self-insure together. This group files an annual M-1 Form and is not required to file a 5500. However, MEWAs obtain an EIN from the IRS, specifically for filing purposes. *It will be a major research issue to get the EIN's of all participating employers to get accurate geographic location and employment information for businesses that over health and pension coverage.*

Employers may file a single form for some combination of benefits where the number of participants is not the same for the different benefits. In these cases, it is not clear what the reported number of participants refers to – the largest number of participants across the different benefits is one possibility. Thus the participant count may overstate the number of people with certain kinds of benefits. *Research using Schedule A information could address the degree to which this occurs, because this provides detailed plan level information.*

There is a potential for double counting, since large firms such as GM have multiple pension plans that correspond to different groups of employees - for example, salaried, union, Puerto Rico etc. In general, double counting is a difficult issue for employers with multiple plan records. Different plans may cover different locations (particularly for health plans), or may cover the same set of participants (particularly for pension plans, but also possibility for health plans that have separate major medical plan). Pension plans are much more likely to be by business. *Research using the Schedule A information that is filed for each plan might give better information. Although these data used to be fraught with quality problems, the quality has improved over time – in addition, PWBA has an algorithm that roughly classifies plans of an employer as providing primary or supplemental benefits to a group of workers. For large employers, an appropriate research task is to begin with the summary plan description and contact the board of trustees. The Judy Diamond Group is a valuable starting point in this regard.*

3. What are the likely problems associated with linking to the Business Register?

Large firms with several EINs may file a 5500 form using only one EIN. An important component of this research is the ability of Census to use the parent/subsidiary information to identify associated EINs. If these links are used a *research question is whether all EIN's belonging to all subsidiaries offer the same health and pension plans. The answer clearly will depend on the company - and would need to be investigated. One approach would be to compare employment totals for the matched subsidiaries with the totals for the group*

Some EINs reflect Taft-Hartley plans which are plans administered by a joint board of trustees, half of whom are appointed by the employers and half appointed by the union. They file a 5500 form and list an EIN, but the EINs from the employers will be missing. However, the plan and unions are very likely to have a list of employers. *A research issue is to create these links by accessing the separate reports filed by unions - Labor Management Reports - that identify contributing employers.*

The EIN used for MEWA filings is unclear. *The research approach is to use SS4 data (the complete list of new EINs, which IRS provides to Census) to identify types of plans that do not match to Census's Business Register.*

4. What are the data quality issues?

PWBA has used pension plan data more than health plan data, so problems with that part of the data are better understood. They have done some work checking the consistency of pension plan codes with plan names -- the person who has handled that editing could probably provide some information on the accuracy of those codes.

Reasonably clean, consistent data are available for 1988-1998. Prior to 1988, IRS did not edit files, with the exception of a very clean file from 1978. There are new forms and a new processor in 1999, so that file is likely to have more problems than those from preceding years.

PWBA has a separate set of statistical files that are used for 5500 publications. These files have data on all pension plans with 100+ participants, plus a sample of smaller plans (with weights that reflect sampling). The statistical files have had additional editing of data- namely, checks to see if plan-type indicators are consistent with plan names that indicate that they are 401K or ESOP plans. These files have some records that are not on public-use files – particularly cases that failed IRS edits but for which inconsistencies were not resolved by the time the data were processed. These files do not include any health plans. *The research would involve understanding the editing process and potentially applying them to all the files.*

Proposed Approach

A team of researchers from the Census Bureau and The Urban Institute will tackle the core data linkage and quality issues with the support and guidance of Department of Labor staff and Internal Revenue staff. The initial evaluation is expected to use 1997 and 1998 plan year data, and should take one year. After that, the team could identify high priority research projects and begin the substantive research.

Census

An ad-hoc committee of Census Bureau staff from the demographic and the economics directorates has been established to investigate the potential of the 5500 file. Kristin McCue, from the Center for Economic Studies at the Census Bureau will investigate the quality of the 5500 file as matched to the Business Register and the MEPS file. She will keep Paul Hanczaryk, Bob Taylor and Mike Kornbau of the Economics Directorate apprised of her progress. Martha Stinson, from the LEHD program will match the 5500 enhanced Business Register to the SIPP and the CPS. She will keep Carole Popoff, Marc Roemer and Chuck Nelson of the Demographics Directorate apprised of her progress.

The Urban Institute

Len Burman, Richard Johnson and Julia Lane of The Urban Institute will be the principal investigators. The staff person who will be tasked to work at the Census Bureau is _____

Department of Labor

Dan Beller, of the Department of Labor, who has worked with the 5500 file since its inception has offered to advise the team on the project.

Internal Revenue Service

Nick Greenia, who was involved in developing the first reliable statistical file for the Department of Labor in plan year 1997 has also indicated his willingness to work with the team on the project.