THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

APPLICATIONS OF A MATCHED FILE LINKING THE BUREAU OF THE CENSUS SURVEY OF INCOME AND PROGRAM PARTICIPATION AND ECONOMIC DATA

NO. 03

S. Haber
George Washington University

July 1985

U.S. Department of Commerce BUREAU OF THE CENSUS

ACKNOWLEDGMENTS

This paper was prepared by Sheldon E. Haber of George Washington University and the Population Division, Bureau of the Census. The author wishes to acknowledge the invaluable assistance received from the numerous Bureau of the Census personnel who gave of their time and expertise while this research was being undertaken. Clerical and editorial assistance was provided by Ellie Baugher, Clarissa Edwards, and Kathy Italiano of the Population Division, Bureau of the Census.

SUGGESTED CITATION

Haber, Sheldon E. "Applications of a Matched File Linking the Bureau of the Census Survey of Income and Program Participation and Economic Data," SIPP Working Paper Series, No. 8502. U.S. Bureau of the Census, Washington, D.C. 1985.

Table of Contents

	Page
Introduction	1
SIPP and the Economic Data	4
Survey of Income and Program Participation	4
Standard Statistical Establishment List	8
Longitudinal Establishment Data file	11
Enterprise Statistics	15
Other Economic Data Files	18
Some Applications of Micro-Worker and Firm Data	29
The Relationship Between Wage Rates and Capital	29
Labor Mobility	35
Low Wage Workers and Low Wage Firms	38
Measuring the Effects of Minimum Wage Legislation	41
Structural Unemployment	47
Identifying High Tech Workers and High Tech Firms	49
Implications of the Transition from a Goods to a Service Economy	51
Union Members and Union Firms	54
Productivity Analysis	64
Methodological Problems in Matching SIPP and the Economic Data	71
Procedures for Identifying an Individual's Employer	71
Estimating Missing Economic Data for Small Establishments and Companies	74
Concluding Remarks	77
Footnotes	82
References	93

Applications of a Matched File Linking the Bureau of the Census Survey of Income and Program Participation and Economic Data

Sheldon E. Haber

The new Survey of Income and Program Participation (SIPP) will undoubtedly become a major source of data on a wide variety of aspects of the well-being of the nation's households, families, and individuals. SIPP is designed to collect information about cash and noncash income, taxes, and assets and liabilities from which improved estimates of income, poverty, and wealth can be derived. While the principal thrust of the survey is in the area of income and program participation with the intent of assessing policy issues such as the effects of proposed changes in program eligibility rules on benefits, it will address a much wider range of policy questions and yield data for analytical studies in a variety of areas of economic inquiry. The very richness of SIPP suggests the desirability of augmenting it with administrative and statistical records. In this report, attention is focused on some potential uses of a SIPP file linked to micro-level establishment and enterprise data from the economic censuses and other data sets maintained by the Bureau of the Census.

An area where a SIPP data base enhanced by economic data is most likely to lead to significant gains in knowledge is that pertaining to the behavior of labor markets. A primary source of data for verifying established propositions relating to labor market phenomena and exploring new ones has been the Current Population Survey (CPS). The CPS is designed to provide cross-sectional data

although, because of its rotational pattern, it can be used for longitudinal studies covering a period of up to 16 months. SIPP, on the other hand, is a longitudinal survey in which each panel is queried eight times (three times a year), covering a period of 32 months. Other well-known longitudinal data sets exist, e.g., the National Longitudinal Survey maintained by Ohio State University and the Social Security Administration Longitudinal Employee-Employer Data (LEED) file; however, these do not have the same breadth of coverage as SIPP. Each of these data sets provides information almost exclusively about workers but very little or no information about the places in which they work. By bringing together information about workers and their place of employment in a single data set, denoted below as the SIPP-Economic Data (SIPP-ED) file, gaps in knowledge of how labor markets function can be filled in. Additionally, the file can be expected to add new insights into firm production and cost functions.

A list of some of the areas in which a SIPP-ED file can yield new insights includes the following:

The relationship between capital and wage rates

Labor mobility

Low wage workers and low wage firms

Measuring the effects of minimum wage legislation

Structural unemployment

Identifying high tech workers and high tech firms

Implications of the transition from a goods to a service economy

Union members and union firms

Productivity analysis

The merging of demographic and economic data will enable investigators to obtain improved estimates of the impact of economic and institutional forces which have been intensively studied but still are only partially understood. It will also enable investigators to examine aspects of labor market outcomes and production processes that have, heretofore, been difficult to study.

Besides the substantive knowledge to be gained by merging SIPP and economic data, there are externalities associated with merging these data sets. First, it will be possible to verify the accuracy of the size of firm estimates given by respondents in survey data. The economic files also permit one to more accurately identify the industry in which a worker is employed. As Mellow and Sider (1983) have shown, industry designations by survey respondents and employers differ approximately 8 and 15 percent of the time at the 1-and 2-digit SIC levels, respectively. When economists proxy monopoly power using industry concentration ratios in conjunction with CPS data, they must average the ratios for 2- or 3-digit SICs to match the census industry classification. By merging demographic and economic data, the more accurate 4-digit SIC industry concentration ratios in the source data can be utilized.

As is the case in designing and improving any data base, it is essential to have, at the outset, a clear idea of the study areas and issues to which it may be applied. The primary objective of this paper is to provide such a framework.

This is done in Section 2 where applications of the proposed SIPP-ED file are surveyed. Before examining the applications of a SIPP-ED file, attention is first given in Section 1 to the kinds of information in SIPP and the major

economic data files maintained by the Bureau of the Census, since it is the information in these files which would form the data set to be used in implementing the studies described below. Two methodological problems which need to be addressed in developing a SIPP-ED file are examined in Section 3. Concluding remarks are found in Section 4.

1. SIPP and the Economic Data

In merging demographic and economic data, it is necessary to know the information contained in the various files to be linked and how each file is constructed. In this section, we briefly describe several data sets which might be incorporated into a SIPP-ED file. As mentioned, the demographic data are contained in SIPP. The economic data are found in a number of files, the most important of which are the Standard Statistical Establishment List (SSEL), the Longitudinal Establishment Data (LED) file, and the enterprise statistics (ES). The SSEL is a comprehensive list of establishments and companies with employees and yields information on employment and payroll. The LED, as its name implies, contains longitudinal data but is restricted to manufacturing establishments. The ES, on the other hand, covers companies in the construction, mineral, manufacturing, wholesale trade, retail trade, and most service industries.[1] These and other data files maintained by the Bureau of the Census are discussed below.

A. Survey of Income and Program Participation

SIPP is an ongoing series of national panels from which income and program participation information will be derived. Prior to SIPP, the main source of

income data was the March income supplement to the CPS. SIPP is designed to obtain improved reporting of income and participation in major income security programs, as well as to expand on information needed to analyze program participation and eligibility.[2]

In the 1984 SIPP panel approximately 17,500 households will be interviewed over a two and one-half year period. Later SIPP panels will contain approximately 13,500 households. Since panels will overlap, cross-sectional estimates can be obtained for larger samples, e.g., the sample size can be doubled when two panels overlap; in some years it may be possible to combine samples from three panels. To reduce problems associated with interviewer turnover, each panel is divided into four rotation groups of about equal size. One rotation group is interviewed during the first two weeks of each month. One cycle or wave of interviewing of the four rotations requires four months; thus, each household is interviewed three times a year. The reference period for an interview is the four month period preceding the interview month, e.g., the reference period for the interview month October is June through September.[3]

SIPP consists of four parts. The first is a control card containing such information as age, sex, race, ethnic origin, marital status, educational level, veteran status, place of residence, and names of employers. The second contains a set of core questions covering labor force participation and amounts and types of income received during the reference period. In addition to wage and salary income and income from self-employment questions are asked about cash transfer payments from governmental programs such as social security, unemployment insurance, and welfare programs. Information on the receipt of noncash benefits from Food Stamps, Medicare, and Medicaid is also obtained.[4]

Of importance for linking SIPP data to the economic data for firms, persons with work experience are asked the name of the firms for which they work or worked.

The last two parts are fixed topical modules and variable topical modules. The topics covered in these modules do not require repeated measurement during the year and the reference period may be other than the four month period used for the core questions. The fixed topical modules may be repeated during a panel and over successive panels. For example, the wealth (assets and liabilities) module will be administered twice in each panel, in waves one year apart. A "round-up" module is administered after the first and second years of interviewing to obtain annual estimates of income; estimates of taxes are also obtained in this module. Additionally, respondents will be asked to provide the addresses of employers they have worked for during the year. Other fixed topical modules provide information in the following areas: work history, education history, and health and disability (denoted below as the work history module); educational enrollment; and marital history, fertility, and migration.

The variable topical modules contain supplemental questions designed for other federal agencies. Although these questions may be repeated from one panel to another, they often are of such special interest that they may be asked only one time over a span of years. Included among the variable topical modules are questions relating to reasons for not working (in which information on a worker's reservation wage will be collected); pension and retirement issues; work related expenses; child care arrangements and financing; health care utilization and financing; housing conditions and costs; energy usage; and other topics.[5]

SIPP contains retrospective and prospective labor market information which goes much beyond the scope of the CPS and other longitudinal data sets. Besides the information that is normally found in the CPS with respect to labor force attachment and employment status,[6] the SIPP core questionnaire contains information which is not routinely provided in the CPS, in particular, information on wage rates for each job held by an individual. Additional information is available in the fixed topical modules. In the work history module, for example, questions relating to the following areas are asked:

Education history

- Program of studies taken in high school
- Courses taken in high school
- Highest degree attained beyond a high school diploma
- Field of study of highest degree
- Year highest degree received

Work history

Training

- Source of latest training (e.g., apprenticeship program, training program at work, military training)
- Date and length of training program
- Who paid for the training (e.g., self or family, employer, government)

Prior work experience

- Year in which first worked at a job lasting six consecutive months or more
- Number of years worked six or more months during the year
- Typical work status during years worked (full-time, part-time)
- Number of times, duration, and reason for not working six or more consecutive months

Last job (for persons working 10 or fewer years at current job)

- Year started and ended job
- Usual hours worked
- Rate of pay at end of job
- Reason for leaving last job (e.g., layoff, discharge, other reason)
- Time between last and current job

Current job

Size of establishment and firm
Single or multi-establishment firm
Union status and coverage by a union contract
Rate of pay at start of current job
Years worked for current employer
Years worked in current occupation for current employer

Health and disability

State of health and, if disabled, length of time disabled

As indicated above, although SIPP provides extensive historical data about an individual's work experience, little information is available about the firms in which that experience is gained. The economic data described below would remedy this shortcoming.

B. Standard Statistical Establishment List

The SSEL is a comprehensive directory of single establishment and multi-establishment companies[7] with one or more employees in the private nonagricultural sector of the economy,[8] including those which are in industries that are out-of-scope with respect to the economic censuses. The SSEL links parent companies, subsidiaries, and their establishments. It contains information on approximately 4.7 million enterprises and 5.7 million establishments.

The SSEL is comprised of three files. The first is the Single Unit (EC-EI) file; it contains information for businesses with paid employees, which, therefore, must file for an employer identification number (EIN) with the Internal Revenue Service. The EC-EI file includes all businesses with an EIN provided, as noted, they have paid employees. Such businesses may consist of a single establishment

company, an establishment or subsidiary (with one or more establishments) of a multi-establishment firm which itself can be part of a larger enterprise, or an entire multi-establishment parent company. All establishments belonging to a multi-establishment parent company and the establishments of the parent company's subsidiaries are listed under the parent company's identification number, denoted as an Alpha number,[9] in a second file, the Multiple Unit (EC-MU) file.[10] The third file, known as the Master Mailing Address file, contains the mailing addresses of multi-establishment enterprises.

The SSEL also contains the address of the physical location of each establishment. The physical address of single establishment company is found in the EC-EI file; the physical address of each establishment in a multi-establishment enterprise is found in the EC-MU file.[11]

To further illustrate how the SSEL is constructed, it is useful to contrast it with the Social Security Administration LEED file. Besides differences in the data elements contained in the two files, there is another difference that should be noted. In the SSEL, the basic units of observation are single unit establishments and multi-unit establishments and their parent enterprise. In contrast, the only unit of observation in the LEED file is a firm's EIN. With only this information, it is not possible to ascertain if a firm consists of a single establishment; a subsidiary, with one or more establishments, of a larger enterprise; or the parent enterprise. If an EIN changes, as may occur when a business changes ownership, it will appear that the business has closed down. Analogously, employees of such businesses will appear to have changed jobs. In the SSEL, it is possible to identify individual establish-

ments of multi-unit enterprises through their Census File Number (CFN) even though there may have been a change in ownership.[12] A single unit company, however, cannot be linked over time when it is sold, say, by one sole proprietor to another.

While the SSEL contains a narrow range of economic data, these data impart valuable information. The establishment data contained in the SSEL are as follows:

Identification

- Name of establishment (and company)
- EIN of establishment, subsidiary, or company
- CFN of establishment
- Alpha number of enterprise
- 4-digit SIC
- Location identifiers (e.g., address of physical location of establishment)

Number of employees, 1st quarter

Payroll, 1st quarter and annual

Value of shipments or sales and receipts[13]

Legal form of organization (sole proprietorship, partnership, corporation)

Date EIN was entered into the SSEL

Reason for EIN being issued (e.g., started a new business, change in ownership)

Operational status (e.g., active, inactive)

The address of the physical location of an establishment is useful for the merging of SIPP and economic data, since it is a primary link in identifying an individual's place of work. Identification of the establishment in which a person works enables one to determine his or her industry at the 4-digit SIC level. The date an EIN was entered in the SSEL and the reason for it being issued provide information about the age of an establishment. The information

on employment yields a more accurate estimate of employer size than that which can be obtained from respondent's estimates in survey data. The employment and payroll figures also yield an estimate of average annual pay, thereby indicating whether an employer is a low or high wage employer.[14] And the ratio of sales to employment provides a proxy measure of productivity. The operational status information can be utilized to identify establishments which have become inactive.

It should be noted that the SSEL contains longitudinal information. Currently, establishment and company data are carried for three years in the SSEL.[15]

C. Longitudinal Establishment Data File

The LED is a longitudinal micro-data base containing data at the establishment level from the Annual Survey of Manufactures (ASM) and the Census of Manufactures (CM). The data begin in 1972 and currently extend through 1981; it is anticipated that data from the 1982 CM and 1983 ASM will be incorporated into the file by late 1985. The LED was developed jointly by the Bureau of the Census and Yale University under the direction of Richard and Nancy Ruggles.

Every year economic data are collected in the ASM from a sample of 55,000 establishments. Included in the 1984-1988 ASM sample frame as certainty cases are all plants of certainty companies, i.e., companies that in 1982 had $500 million or more in shipments of manufactured goods; establishments (in noncertainty companies) with 250 or more employees; and establishments (not falling in the preceding classes) that produced a significant proportion of the output in

their 5-digit product class.[16] The noncertainty cases are drawn from the remaining population of plants of multi-establishment companies and single establishment companies, provided that the latter employ more than five to 20 workers (depending on industry).[17] After five years, these noncertainty establishments are replaced by a new noncertainty sample.[18]

Every fifth year a Census of Manufacturers is taken. All establishments in the ASM, all establishments of multi-unit enterprises not included in the ASM, and all non-ASM single unit enterprises with 5 to 20 or more employees receive a short or long form questionnaire. Administrative records are used for the smallest firms, i.e., those with less than five to 20 employees.[19]

As noted, the LED contains data collected in the ASM and CM. Because the noncertainty portion of the ASM changes every five years, the year-to-year linkage of data in the LED is poorest for the smallest plants and improves as plant size increases. Likewise, the breadth of data improves as plant size increases. Selected data elements from the 1984-1988 ASM sample frame and 1982 CM are as follows:[20]

Identification--ASM, CM

- CFN number
- 4-digit SIC
- Location identifiers (e.g., state, SMSA, county)

Legal form of orgainzation--CM

Number of employees--ASM (pay period including March 12th), CM (annual average)

- Production workers--ASMS and CMS (pay period including March 12th); ASML and CML (average of the pay periods including the 12th of March; May, August, and November)
- All other employees (pay period including March 12th)--ASM, CM

Payroll (annual)--ASM, CM

- Production workers (annual)--ASML, CML
- All other employees (annual)--ASML, CML

- Hours worked by production workers (annual)--ASML, CML

- Cost of materials and services used--ASM, CM

  - Cost of fuels consumed for heat and power--ASML, CML
  - Cost of purchased electricity--ASML, CML
  - Quantity of purchased electricity--ASML
  - Cost of contract work done by others--ASML, CML

Inventories, beginning and end-of-year--ASM, CM

Capital expenditures--ASM, CM

- New building and structures--ASML, CM
- New machinery and equipment--ASML, CM
- Used building and structures--ASML, CM
- Used machinery and equipment--ASML, CM

Operational status (e.g., active, temporarily inactive, ceased operation)--ASM, CM

Value of shipments--ASM, CM

Percent of shipment value exported--ASMS

Value of shipments exported--ASML

Supplementary labor costs--ASML

- Legally required (includes social security tax, unemployment insurance tax, workmen's compensation tax, and state disability tax)--ASML

- Voluntary programs (includes life and medical insurance premiums, payments into pension and welfare plans, union negotiated benefits, and payments into stock purchase plans)--ASML

Gross value of depreciable assets, beginning and end-of-year--ASM

- Buildings and structures--ASML
- Machinery and equipment--ASML

Gross value of retired depreciable assets (includes assets sold, retired, scrapped, destroyed, etc.)--ASM

- Buildings and structures--ASML
- Machinery and equipment--ASML

Depreciation charges (for buildings and structures and for machinery and equipment)--ASML

Rental payments (for structures and for machinery)--ASM

Buildings and structures--ASML
Machinery and equipment--ASML

Value of shipments to other plants of the same company--ASML

It should be noted that asset and fringe benefit data are collected from the ASM long form questionnaire.[21] For these data elements universal coverage, i.e., coverage for all plants, is available only for plants with 250 or more employees. As can be seen from the figures below, 63.4 percent of the value of shipments in manufacturing originates in plants with 250 or more employees. These large plants, comprising 4.0 percent of all manufacturing plants, employed 56.6 percent of the manufacturing work force in 1977.

Percent Distribution of Manufacturing Establishments with Paid Employees by Size Class (1977 Census of Manufactures)

Size of Establishment (Employees)	Number of Establishments	Number of Employees	Value of Shipments
Less than 50	82.2	15.3	11.8
50 - 99	7.6	10.1	8.6
150 - 249	6.1	18.0	16.2
250 - 499	2.4	15.6	15.0
500 - 999	1.0	13.5	14.8
1000 or more	0.6	27.5	33.6
Percent	100.0	100.0	100.0

As indicated, the LED provides a much broader range of information about establishments than the SSEL. For each manufacturing establishment, it is possible

to estimate the average wage of production workers. For all but the smallest establishments, value added,[22] which is a measure of an establishment's contribution to gross national product (GNP), can be calculated. For establishments in the ASM, information is available on depreciable assets and rented machinery so that the capital/labor ratio can be computed. Additionally, labor compensation including voluntary supplementary labor costs, i.e., fringe benefits, can be measured. The degree to which a plant is vertically integrated with other plants can also be inferred.[23]

D. Enterprise Statistics

Like the CM, the enterprise statistics (ES) are collected every five years. The latest ES data are for 1982 and should be available by early 1986. These data cover enterprises whose primary activity is in an in-scope industry.[24] For each enterprise, the data are consolidated over all operating units. The information contained in the ES is similar to that in the CM. The following information, derived from the economic censuses, is available for all enterprises:

Identification

- Alpha number of enterprise[25]
- Census Enterprise Industry Category[26]
- Location identifiers

Legal form of organization

Single industry or multi-industry[27]

Number of owned establishments

Number of employees[28]

Payroll, annual

Sales and operating receipts (excluding the value of intra-company transfers among own establishments)

Value added (only for a company's establishments in the mineral, construction, and manufacturing industries)

Inventories, beginning and end-of-year

Capital expenditures (new buildings and structures, new machinery and equipment, and used buildings and machinery)[29]

- New computers and peripheral data processing equipment
- Other expenditures for new machinery and equipment

In addition to the above data elements, the ones shown below were collected in 1982 for large multi-establishment enterprises with 500 or more employees, using a separate enterprise questionnaire. In 1977, 93 percent of all enterprises with 500 or more workers had two or more establishments. These large multi-unit companies employed 98.5 percent of all workers in companies with 500 or more workers. Among the data elements found in the enterprise questionnaire are:

Cost of purchased advertising

Supplemental labor costs

- Legally required
- Voluntary programs

Gross value of depreciable assets, beginning and end-of-year (for buildings and structures and for machinery and equipment)

Other domestic assets for which depreciation or amortization reserves are not maintained

Net value of depreciable assets

Net foreign assets

Depreciation charges

Rental payments (for structures and for machinery)

Assets acquired through capital leases in 1982

The distribution of companies by size class is given below. As indicated, companies with 500 or more employees comprise 0.3 percent of all companies with paid employees. These large companies accounted for 47.5 percent of all employees and 48.3 percent of sales and receipts in 1977.

Percent Distribution of Enterprises with Paid Employees by Size Class
(1977 Enterprise Statistics)

Size of Enterprise (Employees)	Number of Enterprises	Number of Employees	Sales and Receipts
None[a]	10.3	-	0.8
1 - 49	86.5	32.8	31.7
50 - 249	2.7	15.0	14.8
250 - 499	0.2	4.8	4.5
500 - 999	0.2	4.2	4.4
1000 and over	0.1	43.3	43.9
Percent	100.0	100.0	100.0

[a] Companies which report annual payroll, but did not report any employees on their payroll for specified pay periods in 1977 (see footnote 28).

From the ES data, measures similar to those mentioned in discussing the LED can be derived. Additionally, a crude measure of profitability can be estimated by the ratio of an enterprise's business receipts less its variable costs (usually measured by labor costs) to its capital as measured by gross value of depreciable assets. A similar measure can also be derived from the LED, but for many problems the relevant decisions are based on profitability at the enterprise level rather than the establishment level.

Two items of interest which are contained in the 1982 ES but not in the LED are capital expenditures for computers and the cost of purchased advertising. The first of these, capital expenditures for computers, is of interest because it indicates use of a technology which underlies a number of new industries, often described as high tech industries. The second, the cost of purchased advertising, is a partial measure of the degree to which a product market is competitive.

In constructing the ES file, a number of intermediate files are created. Among these is (1) a file for single establishment enterprises, (2) a file for establishments of multi-establishment enterprises in which the in-scope records for the establishments comprising an enterprise are grouped together, and (3) a file for enterprises.[30] The first two files are based on records from the industrial and business economic censuses; the third file is based on the enterprise questionnaire described above. Information about a worker's establishment and company can be obtained by accessing the appropriate intermediate ES file. The record format for all three files is the same. Thus, a file with records for all in-scope establishments comprising an enterprise and the enterprise itself can be created by merging these files.[31]

E. Other Economic Data Files

The Bureau of the Census maintains numerous economic data sets besides those discussed above. Among those of interest for this paper are the industrial censuses, including the Census of Mineral Industries (CMI) and the Census of Construction Industries (CCI), and the business censuses, which include the

Census of Wholesale Trade (CWT), the Census of Retail Trade (CRT), and the Census of Service Industries (CSI). Additionally, the Bureau of the Census conducts special surveys and maintains other data sets that are of particular relevance for this report. These data are described below.[32]

Minerals Industries

Establishment data for the minerals industries are only available every five years from the CMI. Approximately 1.1 million workers were employed in the minerals industries in 1982. About 430,000 workers were in coal mining, metal mining, and nonmetalic mining, e.g., stone and clay products. The remaining 680,000 were in the oil and gas extraction and field services industries; the bulk of this group, about 400,000, were in oil and gas field services industries.

In the CM, an establishment is defined as a plant at a single physical location.[33] Establishments in the CMI are defined in the same manner except for the oil and gas extraction and field services industries. In the oil and gas extraction industries where workers rotate among widely dispersed sites, the data are collected only at the state level. The problem of defining an establishment is even more severe in the oil and gas field service industries where work sites are dispersed throughout the entire country; the economic data for these industries are collected on a country-wide basis. In both the oil and gas extraction and oil and gas field services industries, the economic data are allocated by county; for the latter industries this is done on the basis of receipts, for the former on the basis of employment, the quantity

of crude petroleum and natural gas shipped, and capital expenditures, depending on the data elements to be allocated.[34]

As in the case of the CM, a census questionnaire is sent to all but the smallest establishments. In coal mining, for example, only companies with five or fewer workers do not receive a questionnaire. Economic data for the smallest establishments are obtained from administrative records.

The same economic data found in the CM long form questionnaire, plus information found in the ASM, e.g., gross depreciable assets and voluntary supplementary labor costs, are also available, with few exceptions, for all establishments except small ones in the stone, sand, and gravel industry. For small establishments in this industry, the data are restricted to such basic information as employment and payroll. Also, the gross value of depreciable assets, gross value of retired depreciable assets, and depreciation charges are not collected for the oil and gas extraction industries. For these latter industries, asset and depreciation data are collected in the ES but only for large companies.

Construction Industries

In the CCI, an establishment is defined as a relatively permanent office or other place of business from which work on one or more projects (or at one or more construction sites) is managed. Where several companies share or operate out of a single office, each legal entity is considered a separate establishment.

CCI questionnaires are sent to all establishments of multi-establishment companies and to all single establishment companies with 15 to 20 or more employees, depending on industry. Questionnaires are also sent to a subsample of the smallest single establishment companies.[35] Information for establishments which do not receive a questionnaire is obtained from administrative data.

The data collected in the CCI are, with few exceptions, the same as those collected in the CMI. No information is available, however, on hours worked[36] and gross value of retired depreciable assets. The gross value of depreciable assets is available but only at the end of the census year. Quarterly employment and annual payroll information is available for construction workers, e.g., painters, plumbers, electricians, etc., and for all other employees. Instead of total value of shipments, output is measured in terms of total business receipts and receipts for construction work.

The Business Censuses

Because of the many, small establishments in wholesale trade, retail trade, and the service industries, questionnaires are only sent to the larger establishments and some small establishments; information for the remaining small establishments is obtained from administrative records.

In the CWT, a long questionnaire is sent to all establishments of multi-establishment firms; to all large single establishment firms, i.e., those

above the certainty cutoffs which vary depending on industry; and to a one-in-ten sample of single establishments below the certainty cutoffs. These wholesale establishments account for approximately 90 percent of the industry's sales and other operating receipts.

The long questionnaire is similar to the short form questionnaire in the CM; however, no information is available on hours worked and capital expenditures. Employment information is available for all employees and by type of worker (i.e., workers engaged in selling, in sales support, and in other activities) as of the pay period including March 12. Annual and 1st quarter payroll data are available, but they are restricted to all employees. Information is available on total operating expenses, including payroll, pension contributions, and overhead expenses, but no breakdown is provided of the individual expense items. Information on the cost of goods sold and materials and services used is lacking. In the CWT, output is measured in terms of sales of merchandise and other operating receipts. Of the aforementioned data elements, only sales and other operating receipts, total operating expenses, and inventories are included in the CWT short form which is sent to the 9 out of 10 single establishment firms that do not receive a long form. It should be noted that employment and payroll information is not asked for on the short form; this information is obtained from administrative records.

As in the CWT, CRT and CSI questionnaires are sent to all establishments of multi-establishment firms, the larger single establishment firms, i.e.,

those above the certainty cutoffs, which again vary by industry, and one-tenth of the single establishment firms that do not meet the certainty cutoffs. Information for the remaining small single establishment firms is obtained from administrative records.

The CRT and CSI questions are similar to those in the CWT; however, no questions are asked about operating expenses and inventories, and the type of worker question is found in some but not all of the industry questionnaires. Thus, with respect to the data elements discussed in this study, the information found in the CRT and CSI is the same as that in the SSEL. While there is a wealth of data in the CRT and CSI that are not contained in the SSEL, in particular, receipts information by merchandise line (in the CRT) and source (in the CSI), these data tend to be unique to individual industries.

Capital Expenditure Surveys

As in the ASM, annual surveys are conducted for the wholesale trade, retail trade, and selected service industries.[37] Additionally, in the years in which the business censuses are taken, special surveys, based on the same sample of firms queried in the annual surveys, are conducted for these industries. These surveys, known as the Capital Expenditures (CE) surveys, contain information on capital expenditures, the gross value of depreciable assets, and supplementary labor costs (as well as other data) that are not gathered in the business censuses.[38] The importance of the CE surveys is that they are the only source

of these data for the business sector. Asset and fringe benefit data for business sector companies are, of course, found in the ES, but these data are aggregated across all in-scope industries and are limited to companies with 500 or more employees. The CE surveys, thus, provide a means of filling an information gap on small service sector firms which is left open in major Census Bureau micro-firm data sets, i.e., the ES, the business censuses, and the SSEL.

The firms in the CE surveys are drawn from list samples and area samples. The list samples are derived from the SSEL. The area samples include new businesses with paid employees that are not found in the SSEL.[39]

All companies listed in the SSEL that meet the certainty cutoffs with respect to the relevant business industry, say, retail trade, and also have two or more establishments in the industry are sampled.[40] The data for such companies are generally aggregated over all their establishments in the industry.[41] For example, in the CE survey of retail trade establishments, the data are aggregated over all the retail establishments of a certainty company having two or more retail outlets. However, the company's establishments in industries other than retail trade, e.g., in manufacturing and in services, are excluded. Information about assets and fringe benefits would be available for the manufacturing establishments with 250 or more employees and for smaller establishments if they were included in the ASM but would not be available in the CSI for the service establishments.[42]

The data for all other companies selected off the list samples, i.e., single establishment companies in the relevant business industry that meet the

certainty cutoffs and single or multi-establishment companies that do not meet the certainty cutoffs, are collected on an EIN basis. In the business censuses, most firms identified by an EIN are single establishment firms.[43]

The data are reported by establishment for companies drawn from the area samples. Area samples are maintained only for retail trade and the service industries.

Although there is no uniform unit of observation in the CE surveys, in only a minority of the firms, estimated at approximately 12, 6, and 10 percent in wholesale trade, retail trade, and services, respectively, are the data collected at the company level (within a given business industry). Data for the remaining firms are collected at the EIN level, but, as indicated, most of these are single establishment firms. Where firms are comprised of two or more establishments, the data can be imputed to an establishment by averaging the relevant variables over the number of establishments owned by a firm.

The CE surveys includes the following data obtained from the CE questionnaires or administrative records: legal form of ownership; number of owned establishments covered by the survey; total annual payroll, legally required and also voluntary supplementary labor costs; beginning and end-of-year gross value of depreciable assets; beginning of year value of buildings and of machinery and equipment; value of assets sold, retired, scrapped or destroyed; lease and rental payments for buildings and for machinery; total capital expenditures, including separate expenditure data for new structures and for new machinery and equipment, new computers and peripheral data processing equipment, and used buildings and machinery; depreciation charges against buildings and machinery;

beginning and end-of-year inventories (retail trade only);[44] the cost of materials and services, including separate figures for office supplies, purchased electricity, fuels consumed, and purchased advertising; and sales and operating receipts.

Absent from the CE surveys is information about the number of workers employed by reporting firms. Thus, while it is possible to estimate capital formation as required for GNP accounting, an essential data element necessary for performing the studies described in the remainder of this paper is lacking. This difficulty can be remedied, however. The CE surveys can be augmented by matching them against their respective economic censuses or the SSEL. The matching process is not without problems, however, e.g., the EINs and SICs reported in the CE surveys may differ from the most current EINs and SICs in the SSEL.[45] A much simpler procedure and one which would greatly enhance the analytical potential of the CE micro-economic data, even though it would duplicate information in the economic censuses, would be to add a question on number of employees to the surveys. Despite the limitations of these data, their potential value as a micro-economic data set for studying aspects of labor market behavior and production processes is yet to be tapped.

Research and Development Expenditures File

The R&D file is funded by the National Science Foundation. The first R&D surveys were conducted by BLS in the early 1950's, but since 1957 they have

been conducted by the Bureau of the Census. Tape files extend back to 1972; the most recent year for which the data are available is 1983.

Every five years a panel of companies is selected to receive annual questionnaires. In most industries these companies have R&D expenditures of 1 million dollars or more; in some industries a lower figure is used to insure coverage of 95 percent of the industry's R&D expenditures. The most recent panel includes 1,500 companies. While the sampled companies are from all the major industry categories, the vast majority are in manufacturing.[46]

In alternate years, long and short form questionnaires are used.[47] Both forms contain information on net sales and receipts, total employment (as of the pay period including March 12), employment of R&D scientists and engineers, R&D expenditures (by domestic affiliates) cross-classified by type (applied, basic, and development) and source (company, Federal agencies), R&D expenditures in the areas of energy and pollution abatement, and total R&D expenditures by foreign affiliates. Additionally, the long form provides information on basic R&D expenditures by field (e.g., physics and chemistry), applied R&D expenditures by product group (e.g., motor vehicles and equipment), total R&D expenditures by state, and R&D expenditures contracted out to other companies.[48]

Quarterly Financial Report

The Quarterly Financial Report (QFR) was initiated by the Office of Price Administration and then transferred to the Federal Trade Commission in 1971.

It has been maintained by the Bureau of the Census since December, 1982. This data set provides income and balance sheet information for the domestic operations of approximately 14,500 firms[49] whose principal activity is in manufacturing, mining, wholesale trade, and retail trade.[50,51] The certainty sample frame includes the firms in these industries with total assets of 25 million dollars or more. Although the QRF is a longitudinal data set with more than 35 years of data, in its most current format the data extend back to 1974. The QFR is used by the Commerce Department in making quarterly and annual estimates of GNP and by the Federal Reserve Board to measure flow of funds and to analyze industrial debt structure, liquidity, and profitability.

The QFR contains quarterly information on a firm's gross and net sales receipts and operating revenues, income before and after taxes, and net income retained in the business. In addition to income and loss data, the QFR contains the following asset information: cash assets and government and other securities; inventories; gross and net depreciable fixed assets, including construction in progress; land and mineral rights; and total net assets.[52] On the liabilities side of the balance sheet, the QFR contains short-term liabilities; long-term liabilities; and stockholder's equity based on historical asset costs.

Based on the aformentioned data, one can determine before and after tax profit rates on stockholder's equity, total assets, or sales receipts as well as the ratio of current assets to current liabilities and the ratio of stockholder's equity to total liabilities. These operating and balance sheet ratios are central to a firm's decision processes regarding capital expenditures, R&D expenditures, employment, and long-term growth.

With this perspective of what is available in the economic censuses for individual establishments and companies, we turn now to a more detailed discussion of the applications of a SIPP-ED file. The reader should bear in mind the limitations of the data, in particular, the lack of asset and fringe benefit data for small establishments in manufacturing, all but the largest establishments in the service sector, and small enterprises irrespective of industry. While such data are not universally available, this does not preclude the possibility that the data base can be augmented in the future, e.g., by developing analytical models for imputing assets and fringe benefits.

2. Some Applications of Micro-Worker and Firm Data

In this section, a more detailed discussion of the applications of a SIPP-ED file is provided. The main objective is to show how the uses of such a data set cut across labor market and production theory. In pursuit of this objective, we focus on a number of issues and hypotheses, some of which have been examined in the literature using data sets whose informational content is not as rich as that of a SIPP-ED; others have yet to be explored because the necessary data are unavailable.

A. The Relationship between Wage Rates and Capital

Perhaps the most immediate application of a SIPP-ED file data is in the area of wage determination. Despite the voluminous number of analyses in which investigators have sought to explain wage rates by sex, race, union status, and region, our understanding of the wage determination process still remains

incomplete. One area which warrants further investigation is the relationship between wage rates and capital. While the productivity of labor is strongly related to the amount of capital with which it is combined, we have been able to identify only a few wage rate studies that incorporate variables relating to capital use. One reason for this may be that such data are difficult to obtain. Another reason is that economic theory suggests that competitive wage rates are independent of the amount of capital utilized by a firm. Exceptions to this proposition, however, are worth noting.

In theory, variations in capital among firms have no impact on wage rates in competitive labor markets. In a competitive labor market, any firm can hire as many units of a given quality of labor as it requires at the prevailing market wage. The ability of a firm to do this does not depend on the amount of capital it utilizes in its production process (or any other attribute of the firm) as long as the amount of labor it hires is not sufficiently large relative to the amount available to influence the wage rate. Although some firms have more resources to purchase labor because they are efficient, e.g., because they are better able to incorporate capital into their production process, the fact that they are efficient means that their output will be larger than that of inefficient firms. They may, however, also purchase higher quality labor, again at the prevailing market wage for such labor, rather than pay more than the market wage for lower quality labor.[53]

Several premises underlie the competitive model which may not be met in practice, e.g., that labor quality can be precisely defined and accurately

measured so that firms are able to determine with exactness which quality of labor a particular applicant belongs to. The premise that a firm can accurately determine worker quality is cast into doubt by the large sums of money that are spent in screening applicants for employment.[54] While more productive firms can still survive even though they may overpay some workers whose quality is overestimated, less productive ones may only be able to survive if they are successful in paying workers only what their quality warrants. Thus, a positive relationship between firm productivity and the wage rate is plausible, even in competitive labor markets.

It is also clear that no matter how precisely an occupation is defined, workers within that occupation will differ in terms of their work effort, attentiveness, attendance record, and similar attributes not measured by survey instruments but for which firms are willing to pay a premium. Capital intensive firms which tend to hire high quality workers, some of whose characteristics are difficult to observe, should, therfore, exhibit higher wages than labor intensive firms, again, even when labor markets are competitive. One implication of this proposition, which is supported by findings reported by Brogan and Erickson (1975), is that the positive relationship between the wage rate in an occupation and the amount of capital utilized by a firm should be strongest among occupations directly associated with capital, i.e., the skilled occupations, in which unobserved desirable labor qualities are most likely to be found, and weakest for those occupations where this association is weakest, i.e., the unskilled and clerical occupations.[55]

Market imperfections are another reason why economic data are relevant to the wage determination process. Imperfections in the product market can result in firms securing higher than normal profits. Imperfections in the labor market can result in union members receiving higher than competitive wages. Both phenomena may be related to firm size as large firms are able to concentrate market power and pass on cost increases; they are also easier to unionize.

The relationship between wage rates and firm size has been documented in a number of studies.[56] Using industry data, Masters (1969) found that one-fifth of the variation in the hourly earnings of production workers in manufacturing was explained by plant size. The remaining variables, including the extent of unionization and concentration in an industry, raised the percentage explained by only another one-fifth. Mellow (1982b), using the May 1979 CPS and pension supplement containing information on firm size, found that compared

> "to the excluded [plant and firm] size category (less than 25 workers) ... [the] combined wage premium for an average worker in the largest plant ... and company size category ... [is] 23 percent." (p. 497)

If the combined size of plant and company effect on wage rates is as large as suggested by Mellow, its importance is as great as that of the much more heralded union effect.[57]

Size of firm captures a host of relationships that affect wage rates. Indeed, that is the problem with using a variable which captures much but reveals little by itself. To unlock the puzzle between firm size and wage rates requires information about the characteristics of individual firms. The most important

characteristic may be the amount of capital a firm utilizes. To the extent that this is so, workers of the same quality may receive a higher wage in large firms simply because they have relatively large amounts of capital to work with. In this case, the wage rate and capital/labor ratio will be positively related. Moreover, where the capital/labor ratio is positively related to other variables which themselves impact on the wage rate, the wage elasticity with respect to these other variables should decline. In particular, if there are complementarities in production between higher quality labor, e.g., better educated labor, and capital, the rate of return to education may be lower when the capital/labor ratio is included in the wage model than when it is omitted. The same will be true with respect to the wage gains of labor unions. If a labor union's ability to raise wages is enhanced in capital intensive firms, the union-nonunion differential will be overestimated if variations in the capital/labor ratio among firms are not taken into account.[58]

Besides the direct effect of the capital/labor ratio on the wage rate, interaction effects with other variables are to be expected. For example, the rate of return to education depends on whether it is combined with on-the-job training (OJT) that is specific or general. If it is combined with specific OJT, the age-earnings profile will be flatter and the rate of return less than if it is combined with OJT that is general. To the extent that the capital/labor ratio proxies the specificity of OJT,[59] the interaction effect on the wage rate of education and the capital/labor ratio is likely to be negative.

Of significance for the discussion at hand, not all large firms have the same capital/labor ratio. Large firms in retail trade have smaller capital/labor ratios than those in manufacturing. Moreover, the relationship between the capital/labor ratio and the wage rate may differ across industries, everything else held constant. It seems reasonable to assume that the kind of capital a firm invests in will also affect worker productivity. One might expect wage rates to depend on whether the most recently acquired assets are new or used, the rate at which assets are retired, and the proportion of capital expenditures invested in new computers. For these reasons, information about assets, and capital expenditures, may lead to a reduction in the percentage of the variation in wages left unexplained and, for example, attributed to labor market discrimination in human capital models of the male-female and white-black wage differential.

In addition to differences in the mix of capital and labor, firms also differ in their ability to pay high wages. It has been suggested that ability to pay is associated with the degree to which employment in an industry is concentrated in a small number of firms and firm profitability.[60] Ability to pay may also be evidenced by the age of a firm. All else being the same, young firms attempting to gain a foothold in an industry may not be able to pay as high a wage to attract labor as older firms.[61] Wage rates may also depend on changes in the level of employment; they are likely to be lower in firms that have experienced substantial contractions in their work force than in firms that are undergoing vigorous growth in employment. Variables providing these kinds of information can be derived from the economic data maintained by the Bureau of the Census.

Still another advantage to be realized from the economic data is the availability of information pertaining to supplementary labor costs for voluntary programs. These fringe benefits contribute substantially to total compensation[62] and should be taken into account in explaining labor market outcomes. Antos (1983) has found, for example, that when nonwage compensation is ignored, the union impact on employee income is seriously understated. As indicated, fringe benefit information, as well as other economic data mentioned in this section, would be contained in a SIPP-ED file.

B. Labor Mobility

The literature on labor mobility has been primarily concerned with the factors that influence workers to change employers and the effects of such mobility on individuals, employers, and the economy.

As Mincer and Jovanovic (1981) have shown, the quit rate decreases with years of experience in the labor force and with the length of time individuals work for their employer. Since specific training tends to be paid for by the employer, firms have an incentive to retain workers who receive such training. One way of doing this is by paying higher wages and/or by offering more extensive fringe benefits, e.g., pension benefits. Workers are also reluctant to leave an employer because specific training is not easily transferable among firms.

Implicit in our discussion of the determinants of the wage rate is the hypothesis that the higher the capital/labor ratio, the stronger is the relationship between specific training (which is not observable) and job tenure (which is observable).

Thus, holding job tenure, the wage rate, and other variables constant, workers in firms with a high capital/labor ratio where specific training is more likely to be offered may be expected to have a lower quit rate than similar workers in firms with a low capital/labor ratio.

Quit rates also have been found to be lower in those industries in which fringe benefits are a large proportion of total compensation (Pencavel, 1970). Since the economic data provide information by firm on voluntary supplementary labor costs, improved estimates of the elasticity of the quit rate with respect to these costs (which represents benefits to employees) will be possible.

Labor mobility is also central to the question of dual labor markets. Dual labor market theory suggests that low income is not only related to the characteristics of individuals, in particular, to inadequate or inefficient investment in human capital, but also to the characteristics of the jobs which they hold. In this view there are two kinds of labor markets. In the primary labor market, jobs are characterizied by high and increasing pay, job security, and on-the-job training. The secondary labor market, in contrast, consists of low wage jobs with few opportunities for advancement, high turnover, and little skill development.[63] Individuals who are trapped in the secondary labor market are believed to earn less than their counterparts with similar characteristics in the primary sector. Since low wage workers tend to work for low wage employers, dual labor market theory suggests that labor mobility is greater among low wage firms than between low wage and high wage firms. This essential feature of dual labor market theory can be tested using SIPP and economic data from the SSEL.

Although respondent estimates of establishment and firm size are found in some data sets, such information is generally available only for a particular moment in time. By combining SIPP with SSEL data, firm size as well as other employer characteristics can be developed for all of an individual's employers. One area where such information is crucial relates to human capital transfers between firms of different size. Schiller (1982) has suggested that by their very nature small firms expose workers to a variety of job skills and training under intensive supervision. But since small firms pay less than large ones for the same skills, workers who receive training in the former may move to the latter. Schiller estimates that "small firms are losing over 30,000 newly-trained workers (net) each year to larger firms" (p. 68). According to Schiller, although individual workers and large firms benefit from this transfer of human capital, small firms lose with the resultant outcome that they may provide less training than otherwise. Schiller's study is based on data from the Social Security Administration LEED file and, thus, suffers from the problems noted above in measuring firm size and labor mobility when EIN's are used to define a firm. These problems would, to a large extent, be absent in a SIPP-ED file.

With the availability of micro-worker data, investigators have begun to look at the return to job mobility, in particular, whether job changers experience more rapid wage growth than job stayers. For men, it appears that although job separations lead to short-run gains in wage rates, in the long-run the largest gains are registered by those who stay with an employer (Borjas, 1981). While this may be true in general, it may not be true of workers who remain

with low wage employers. Whether this is so is clearly of some interest in furthering our understanding of how labor markets function.

C. Low Wage Workers and Low Wage Firms

The relationship between poverty and low wage rates is self-evident. In 1978, the average poverty threshold for a family of four was $6,662.[64] In the same year, almost 40 percent of the wage and salary workers in families with incomes of less that $6,000 earned the minimum wage or less.[65] The characteristics of these low wage earners are the same as those of persons living in poverty, i.e., a relatively high proportion are young, black, and female.

While survey data such as the CPS provide insights into the characteristics of low wage workers, they provide no information about low wage firms. Under the plausible assumption that such firms employ low skilled labor, the price of labor relative to capital will tend also to be low. All else being the same, such firms will be labor intensive and, hence, tend to be smaller than high wage, capital intensive firms. And because recruitment and hiring costs relative to the level of wages will tend to be high, such firms will also advertise less for labor and employ fewer screening devices to measure the quality of workers; thus, their work force will be less qualified than that of high wage firms. Less qualified workers, on the other hand, e.g., younger workers and those who are less educated, will be attracted to low wage firms because their marginal product is less than that required to gain employment in high wage firms. More generally, workers with given characteristics and tastes sort themselves among firms with similar requirements for labor. The

outcome of this process is an equilibrium relationship between compensation, on the one hand, and worker and firm characteristics, on the other hand, as determined by successful job matches.[66]

Corresponding to the greater prevalence of low quality workers in low wage firms, one might expect that in these firms (vis-a-vis high wage firms) a higher proportion of capital expenditures is for used rather than new machinery and equipment; likewise, the proportion of depreciable assets retired each year is likely to be smaller in such firms. Furthermore, given that labor is of lesser quality and capital is of an older vintage, it would not be surprising if value added per worker were relatively low in low wage firms.

Other characteristics are more easily seen by focusing on high wage firms. To the extent that high wage firms are capital intensive, their need for trained workers is likely to be greater than that of low wage firms. Capital intensiveness suggests greater use of resources to monitor output; hence, a higher proportion of the work force may be needed in superivsory positions. To reduce turnover, which disrupts the production process, high wage firms are likely to offer future benefits in the form of pensions. Discontinuities in production are also reduced through vertical integration.

Information about low and high paying firms is important for another reason besides the knowledge it provides on how production is organized in these two types of firms. Since low paying firms are a source of employment for workers with relatively low productivity, it is of some interest to inquire into the extent

to which low pay among workers is attributable to their employment in such firms. In approaching the question of why some workers are paid less than others in this manner, low wage employers can be viewed as providing employment opportunities with attendant low earnings, not because they discriminate against certain groups of individuals, but because the production processes that are most efficient for their mode of operation do not require high quality labor and, furthermore, they inhibit their paying high wages.[67]

A procedure for verifying this view would be to sector firms according to whether they are low paying or high paying.[68] With this sectoring of firms, one would expect, as indicated above, that the mix of workers and capital is dissimilar between the two sectors. Assuming this is so, to what extent are differences in individual earnings in low and high paying firms due to the characteristics of the workers and capital employed in each type of firm? Also, to what extent are workers with similar characteristics renumerated in the same way in each type of firm?

One way of answering the first of these questions is to separately estimate wage rate equations for workers in low and high paying firms. The variables in each equation would reference the quality of labor and the quantity and quality of capital, and would control for occupation, industry, geographical location, union status, etc. Earnings differences between each type of firm due to factors other than labor quality can then be estimated by holding constant the quality of workers in low and high paying firms. Additionally, earnings differences between each type of firm due to factors other than the quantity and quality of capital can be estimated by holding constant the quantity and

quality of capital in low and high paying firms. Subtracting the sum of these differences from the total differential in earnings of workers in high and low paying firms yields an estimate of the earnings discrepancy which is due to the differential rate of return to labor and capital in both sectors.

The answer to the second question posed above is obtained by determining which coefficients of the variables referencing labor and capital are significantly different from zero (and have the right sign) in each sector, and where both coefficients for a given variable are significant, whether they are significantly different from each other. For example, it may turn out that being female or black has no effect on earnings in low paying firms but both groups earn less than their white male counterparts in high paying firms. Were this outcome observed, one could then go on to estimate the amount by which the earnings of these groups would rise in high paying firms if they had been paid at the same rate as white males in high paying firms.

The primary point to be emphasized by this discussion is that information about firms, in particular, whether they are low or high paying, adds an important additional dimension in assessing how earnings outcomes are determined in the labor market.

D. Measuring the Effects of Minimum Wage Legislation

Economists have long been interested in government policies which are believed to impose restrictions on the free operation of labor markets. Minimum wage legislation falls into this category. As a result of amendments to the Fair Labor

Standards Act (FLSA) in 1977, the federal minimum wage rose from $2.65 per hour in 1978 to $3.35 per hour in 1981; the same minimum wage prevails today. As of September 1982, 85.5 percent of all private sector nonsupervisory employees were covered under the 1977 amendments; the corresponding figure for all employed wage and salary workers was 63.5 percent.[69]

By raising the wage above that which would prevail in a competitive labor market, workers whose marginal revenue product is less than the minimum wage are subject to disemployment. With respect to the firms in which such workers are found, the disemployment effect may be so severe that they will cease operation. In general, all else being the same, the disemployment effect will be greater the lower a firm's average wage rate. However, the adverse impact of the minimum wage may be greater for both low wage and high wage workers in low wage firms than in high wage firms, since the minimum wage impacts on the firm, and only indirectly affects individual workers through changes in firm behavior. As a result of the minimum wage, some high wage workers in low wage firms may become disemployed if total firm employment shrinks. On the other hand, if the proportion of low wage workers is small in a high wage firm, there may be no or only a small disemployment effect.

The earliest studies of the effects of minimum wage legislation focused on low wage industries, since such effects are most reliably detected when a significant proportion of an industry's work force is comprised of low wage workers. These studies attempted to isolate the impact of the minimum wage by comparing changes in employment before and after imposition of (or an increase in) the minimum wage between a test group and control group of firms, i.e., between

covered and noncovered firms.[70] While this approach is the most direct one in getting at the impact of minimum wage legislation, the data that have been collected suffer from several deficiencies. Among these deficiencies are 1) retrospective data were utilized so that firms that closed down between the time legislation was initiated and a survey taken were omitted from the studies, thereby understating the adverse impacts of the minimum wage, 2) the studies typically measured employment only in terms of numbers of workers with no correction for hours worked, again, possibly understating adverse impacts, 3) inclusion of workers, such as professional workers and managers, who are normally exempt from the FLSA makes interpretation of the data difficult, and 4) it has not been possible to control for prior employment trends.

The need to control for prior employment trends is due to the implicit assumption that in the absence of minimum wage legislation, the test and control groups would grow at the same rate.[71] For example, assume that it is found that employment in retail industries characterized by low wage firms grew more rapidly than retail industries characterized by high wage firms, despite extension of the FLSA to the former group. In this case, if employment in the former group had been growing even faster than in the latter group prior to the extension of coverage, it would be incorrect to infer that no disemployment had occurred. Because of the aforementioned difficulties, industry studies of the disemployment effect of the minimum wage have fallen out of vogue. Were a SIPP-ED file in place, each of the problems just noted could be resolved.

In recent years, investigators concerned with minimum wage issues have turned their attention to groups in the population with specific demographic

characteristics. The group receiving the most study is that of young people; most of the time-series studies have been confined to this group. More recently, cross-sectional data have been utilized to study the effects of minimum wage legislation, holding constant factors which cannot be controlled for in the time-series data.

The study of the effect of the minimum wage on the substitutability of adult labor for that of youths is particularly insightful as to how a SIPP-ED file can be used in policy analysis, since, from time to time, consideration is given to lowering the minimum wage for young people. The argument for a youth differential is based on the supposition that the minimum wage has led employers to substitute adult workers for younger ones, and a lower minimum wage for youths would result in more jobs being made available to them. To test this proposition, Cotterill and Wadycki (1976) estimated the percentage of workers employed in establishments with annual sales of $300,000 or more in eight retail trade industries in 31 Standard Metropolitan Statistical Areas (SMSAs).[72] This percentage was used to proxy the extent to which workers in each industry and SMSA were covered by the Federal minimum wage. The need for this proxy stems from the absence of information in their data set which would permit identification of individuals employed in covered establishments.

Cotterill and Wadycki tested two hypotheses: 1) the wage rate of retail trade employees is higher in SMSAs where minimum wage coverage is greater and 2) firms in SMSAs where minimum wage coverage is greater compensate for higher wage rates by employing more adults and fewer youngsters, i.e., higher quality labor.[73] The first hypothesis is tested using a wage model which excludes

personal characteristics variables; not surprisingly, the hypothesis is consistent with the data. The second hypothesis is tested by adding personal characteristics variables to the original wage model. If employers in high coverage SMSAs substitute adult workers for younger ones, the higher wage in these areas would be "picked up" by the personal characteristics variables, thereby reducing the coefficient of the coverage effect variable in the modified model. No reduction in the coverage effect variable was observed, suggesting that young people had not been replaced by older workers when the minimum wage was extended to retail trade.

It should be clear from this brief review of Cotterill and Wadycki's study that a much simpler and more direct test of their hypotheses would have been possible if a SIPP-ED file had been available to them. The SSEL portion of this file would have permitted them to identify individual firms and workers covered (and not covered) by the federal minimum wage law, thereby vitiating the need to develop a proxy variable based on geographical variations in coverage. Moreover, it would have eliminated the need for an indirect test of the substitution effect of the minimum wage. By grouping workers in covered and in noncovered firms in various industries, and controlling for other factors that govern the ratio of young to adult workers, e.g., the geographical and occupational distribution of an industry's work force, variations in the youth/adult ratio attributable to employment in firms covered by the federal minimum wage law could be directly ascertained.

It should be noted that in studying the impact of the extension of Federal minimum wage coverage on employment in say, retail trade, minimum wage effects

should be distinguished from size of firm effects. For example, assume that after the extension of the Federal minimum wage in retail trade it was found that the youth/adult worker ratio was lower in SMSAs where the coverage was high than in SMSAs where it was low (or lower in covered than noncovered firms were such data available). One could not infer from these data alone that as a result of the extension of coverage adult workers were substituted for younger ones, since covered firms are larger in size than noncovered firms and would tend to have a lower youth/adult worker ratio even in the absence of the minimum wage. This problem is amenable to analysis using a SIPP-ED file since it would contain information on size of firm; hence, one could control for variations in this variable.

Cross-sectional data have also been used by Leighton and Mincer (1981) in assessing the impact of the minimum wage on OJT. Since OJT is part of the total compensation package, any exogenous increase in the wage rate, in this case, due to the minimum wage, should restrict the amount of OJT that low wage firms can provide. Given that the Federal minimum wage is uniform across states, the authors estimate the level of wages in each state for workers with the same characteristics. By identifying low wage and high wage states in this manner, it can be assumed that, on average, firms in low wage states pay lower wages than firms in high wage states. Holding the proportion of workers covered by Federal minimum wage legislation in a state constant, the lower the value of the state wage, the greater the negative impact of the minimum wage[74] and the less likely are workers to have participated in OJT. Leighton and Mincer also measure OJT directly from responses indicating whether a

person received training in his or her current job and indirectly from the change in an individual's wage and from their length of job tenure. The empirical data suggest that, indeed, the minimum wage tends to discourage OJT.[75]

As before, availability of a SIPP-ED file would provide a means of getting directly at the effect of the minimum wage on OJT, since the average wage paid by a firm is a datum in this file; hence, there would be no need to estimate state wage proxies. Additionally, one can control for the characteristics of the firms themselves to explore how OJT is related to the production process.

Besides the possibility that the minimum wage may affect the amount of OJT that a firm offers, it may also affect a firm's ability to provide health insurance and retirement fringe benefits, since a wage floor limits the trade-off between wages and other forms of compensation. Here again, a SIPP-ED file could be utilized to assess still another aspect of the economic consequences of minimum wage legislation.

E. Structural Unemployment

An important issue in maintaining full employment is the proper mix of policies to meet the challange of structural change. The direct effects of structural change, whether arising from the introduction of new technologies, the substitution of foreign for domestic output, or long-term shifts in consumption patterns, are typically localized to firms producing a particular product or to specific areas. In either instance, the structural disequilibria are such as to result in relatively large changes in firm employment, often of such

magnitude that firms are forced to close their doors. One recent estimate places the annual number of jobs lost due to major plant closings between 1978-1982 at 900,000.[76]

An issue of long standing is what happens to workers who are displaced from their job as a result of structural disequilibria. How long do they remain unemployed vis-a-vis other workers who separate from an employer? What sources of income, including cash and noncash government transfers, do they draw on when they are unable to find work? When they find a job, how do earnings in the new job compare to earnings in the old one? If there is an earnings loss, how much of this loss is recouped, say, after one year?

A major problem in answering these questions is that workers do not know if they are structurally unemployed.[77] One way of identifying such workers is to look at employment changes in the firms in which they were last employed. If the firm has undergone a substantial decline in employment or has closed, one may presume that it has undergone a shock which is typical of the shocks experienced by firms subject to structural disequilibria. It also can be presumed that the employees of these firms experience the aftereffects of such shocks. For some, the aftereffect is loss of a job. For others who are able to retain their job, the aftereffect may be reduced earnings or diminished pay raises instead of actual wage cuts. Just how poorly workers affected by structural disequilibria fare, relative to job changers and job stayers in firms where demand conditions are stable, is an unresolved question.

As indicated previously, the SSEL contains longitudinal data on employment and operational status. A SIPP-ED file would enable one to determine the

extent to which firms are subject to severe, long-term shocks as evidenced by plant closures and substantial reductions in employment, and how such shocks affect their work force.

F. Identifying High Tech Workers and High Tech Firms

Despite the importance of new technologies for improving productivity, regaining our competitive advantage in international markets, and maintaining our defense posture, there is no widely accepted definition of a high tech industry. Using three different definitions, Riche, Hecker, and Burgan (1983) estimate that from 2.8 to 13.4 percent of all wage and salary workers were employed in high tech industries in 1982.[78] The first figure is based on a definition which includes industries with an R&D to net sales ratio of at least twice the average for all industries. The second is based on a definition which includes industries with a ratio of technology oriented workers[79] to all workers of at least 1.5 times the industry-wide average.[80, 81]

High tech industries have been cited as having a large group of high and low wage workers whereas other industries are comprised of workers who are concentrated in the middle of the earnings distribution. It is useful to know how workers in high tech and other industries differ and the differential growth of employment in the two kinds of industries; however, it is equally important to know the characteristics which differentiate high tech from other <u>firms</u> and the differential in the rate of growth of the two types of firms.

Not all firms in high tech industries utilize the latest technology, and new techniques of production are utilized by firms in industries besides those

labeled as high tech. One approach to distinguishing between the two types of firms would be to compare the characteristics of the industries denoted on a priori grounds as high tech with other industries and then to use this information to identify high tech firms.[82] To illustrate this approach, assume that the a priori criterion used to denote high tech industries is one of the definitions noted above, e.g., that the ratio of high tech to all workers in a given industry to the similar ratio for all industries is higher than some minimum value. Assume also that the high tech industries exhibit high values of the following ratios: capital expenditures for new computers to all capital expenditures, capital expenditures to asset value, and capital to labor. Given a set of characteristics which permit the bifurcation of industries, the multivariate technique of cluster analysis can then be applied to identify high tech firms within both high tech and other industries.

Cluster analysis is a way of analyzing multivariate data.[83] It is particularly useful in creating a classification system in that it enables one to group observations, in this case, individual firms, into homogeneous classes or clusters without imposing a priori specifications on the data other than the choice of variables to be used in the analysis. Using cluster analysis, each observation can be compared with every other observation and a measure of distance can be computed for every matched pair. Grouping the two observations with the smallest distance between them, the number of observations can be reduced by one and the process repeated. The end result is two or more clusters, determined by the data, where each cluster of firms represents a homogeneous set of observations.

The outcome of the cluster analysis is a partitioning of firms into categories, in this case, high tech and nonhigh tech firms. An advantage of applying the

aforementioned two-stage procedure using a SIPP-ED file is that it provides an independent test of how well the procedure works. For if the approach is successful, the proportion of workers who are technology oriented among the firms classified as high tech will be higher than the similar proportion for firms classified as nonhigh tech, and the difference in proportions will be greater than the corresponding difference when industries are classified as high tech and nonhigh tech. An additional advantage of the SIPP-ED file is that industries can be disaggregated to the 4-digit SIC level. At this level of detail, a better determination can be made of the variables to be included in the cluster analysis than at the more aggregated census industry level.

Having identified high tech firms, in contrast to high tech industries, insights can then be obtained as to how production processes in these firms differ from their nonhigh tech counterparts. At the same time, it will enable one to better define high tech occupations and how workers in these and other occupations in high tech firms differ from similar workers in nonhigh tech firms.

G. Implications of the Transition from a Goods to a Service Economy

One of the most striking changes in the U.S. economy since World War II has been the growth of service sector employment. Between 1950-1983, the annual rate of growth in goods producing industries has averaged 1.0 percent. In the service industries, however, the rate of growth has averaged 3.1 percent per year over the same period. This rapid growth, coupled with the fact that 1 out of 2 wage and salary workers in the service sector is employed in enterprises with less

than 100 workers, compared with 1 out of 4 in the goods sector, suggests that small service firms are becoming a more common feature of the economic landscape.[84]

Small firms employ a larger fraction of younger and older, female, and part-time workers among their work force than large firms (Barth, Cordes, and Haber, 1984). And as indicated above, small firms pay less than large ones, holding worker characteristics constant. The incidence of fringe benefits, particularly private pension benefits, is also lower among small firms (Mellow, 1982a).

Besides the characteristics of their work force, goods and service firms differ, in terms of their capital/labor ratio. For firms of the same size, the capital/labor ratio is lower in the service sector than in the goods sector. Given the changing distribution of firms between the two sectors, this could imply a reduced demand for capital vis-a-via the level that would otherwise prevail.[85] Additionally, because of the household orientation of the service sector, it is likely that the productivity of a worker in this sector will be less than that of a similar worker in the goods sector working with the same amount of capital. To the extent that this is so, profit margins may be lower in the service sector, and this may also explain why employee compensation is less in the service sector than in the goods sector.

It has been suggested by some that shifts in employment to high tech industries within the manufacturing sector and from the goods to the service sector have resulted in a decline in the middle class (Thurow, 1984). The high tech

industries, it is said, are composed of high paid professional workers and low paid assembly workers in contrast to the "smokestack" industries in which high paid skilled workers and almost as highly paid operatives predominate. Likewise, high and low paying jobs are believed to be more characteristic of the service sector than the goods sector.

It has also been observed that the distribution of earned income, i.e., wages and salaries plus income from self-employment, has become more unequal over time for men but has remained almost constant for women (Henle and Ryscavage, 1980). The sectoral shifts just noted are consistent with these trends. Given the posited distribution of earnings in high tech and service industries, a relative shift of employment among males to these industries would result in a more unequal distribution of male earnings. On the other hand, women have been employed in the service sector for decades, so that the expansion of jobs in this sector (and low paying jobs in high tech industries) would have little effect on the distribution of their earnings.

As indicated, small (large) firms are characteristic of the service (goods) industries. To the extent that both small and large firms pay less in the service sector than in the goods sector, when men shift from the latter to the former sector there is a tendency for them to "slide down" the income distribution, even when they find employment in the same size firm. This same tendency could be less pronounced for women if the wage differential between small firms in the goods and service sector is less than the analogous differential between large firms.

A SIPP-ED file would provide a basis for measuring the effect on the earnings distribution of shifts in employment between the goods and service sectors and

different types of firms, for both men and women. The advantage of using this data set is that it enables one to examine the process by which the earned income distribution changes over time. One way in which this can be done is to partition the economy into small and large firms within both the the goods and service sectors, and then to further stratify each group into high tech and nonhigh firms.[86] For each of these groups, one can compute a measure of inequality of earned income and obtain a weighted inequality measure over all groups taken as a whole.[87] By comparing the measure of inequality and its components for successive SIPP panels, the effects of diverse structural changes impacting simltaneously on the economy can be separately determined. With a decomposable measure of earnings inequality, one can also partition each subsector by age, sex, race, and marital status of the head of household, and compare the effects of changes in demographic and economic variables. In this case, changes in the distribution of earned income would occur within an approximately closed population, since the composition of any given panel, once determined, is essentially fixed.[88]

H. Union Members and Union Firms

Besides government policies, institutional arrangements, in particular, labor unions, affect outcomes in the labor market. Among the outcomes of special interest is the impact of unions on wage rates, worker productivity, and employment. Even though the proportion of workers who belong to unions has declined, from 33.4 percent of nonagricultural employment in 1960 to 21.9 percent in 1980,[89] the role of unions as economic agents continues to be important and to draw the attention of economists.

A basic difficulty in estimating the extent to which unions raise their members' wage rate is that one cannot observe what they would earn in the absence of unions. Simply comparing union and nonunion wage rates e.g., for establishments in the same local area, is not sufficient, since wage rates in nonunion establishments reflect both union threat and spillover effects. To the extent nonunion employers in an industry raise wages in response to increases in the union rate, the union effect on wages will be underestimated. On the other hand, the union effect on wages may be overestimated as workers displaced in the union sector, because of rising wage rates, seek employment in the nonunion sector, thereby depressing wages in that sector. Additionally, in measuring the union wage gap, it is necessary to adjust for as many differences as possible between union and nonunion workers, apart from union status and the presence of unions, which might affect their pay. For example, where unions are successful in raising wages, employers are likely to select higher quality labor than they could attract at lower nonunion wages, necessitating that worker characteristics be accounted for in estimating an adjusted union wage gap.

Despite these and other difficulties, progress has been made in understanding the diverse relationships between unions and wage rates. In the earlier literature the adjusted wage gap was estimated by regressing the average wage rate in an industry on the percentage of the industry's work force which was unionized, holding other variables constant. With the advent of micro-worker files, cross-sectional data have been utilized to determine the relationship between an individual's wage rate and his or her union status. Longitudinal data have

also been utilized to estimate how an individual's wage rate changes as one moves from jobs covered by a union contract to noncovered jobs and vice versa. By using longitudinal data, work habits and traits affecting worker productivity, which cannot be inferred from cross-sectional data, can be presumed to be constant over time. Thus, it is not surprising that longitudinal data yield lower estimates of the adjusted union wage gap then cross-sectional data. For example, Mellow (1981) reports a union wage premium of 19 percent based on CPS cross-sectional data for May 1974 and May 1975. CPS longitudinal data for May-May matches for the years 1974-75 and 1977-78 indicate, however, that workers who leave a union job realize a wage premium over the subsequent year that is less than half of that for workers who remain on a union job. Workers who join a union within a 12-month period earn 3 percent more than non-union members. A similar relationship between cross-sectional and longitudinal estimates of the adjusted union wage gap is reported by Mincer (1981).

Much of the recent literature on the labor market effects of unions has utilized demographic survey data. A small number of studies have utilized establishment data developed by BLS. Using these data, Freeman and Medoff (1981) have examined the global impact of unions by comparing the wage rates of union workers in industries that are highly unionized with the wage rates of union workers in less unionized industries. Their results indicate that in manufacturing there is a strong positive relationship between the wage rate of union workers and the degree of unionization in an industry. They also estimated the extent to which threat, spillover, and demand effects combined raise or lower the wages of nonunion workers in industries that are highly

unionized vis-a-vis industries that are less unionized; among nonunion workers, there is at most only a weak positive association between wage rates and the degree of unionization in manufacturing industries.

In another study using BLS establishment data, Freeman (1981) analyzed the impact of unions on fringe benefits paid to production and nonproduction workers. His estimates indicate that not only are fringe benefits, both in absolute and relative terms, higher among unionized production workers than nonunionized ones, but also the union fringe benefit effect exceeds, in percentage terms, the union wage effect. Additionally, it appears that fringe benefits paid to nonproduction workers are higher when production workers within an establishment are unionized than when they are not.

In both of these studies, 1968, 1970, and 1972 establishment data were used from the BLS Employer Expenditures for Employee Compensation (EEEC) surveys. These surveys contain establishment information on employment, hours worked, payroll, and total compensation (including legally required and voluntary supplementary labor costs) for both nonoffice and office employees. Additionally, they indicate in which establishments a majority of the nonoffice employees and office employees are covered by a union-management collective bargaining agreement. Establishments in which a majority of the nonoffice workers are covered by a collective bargaining agreement can, thus, be identified as union establishments.[90] Because the EEEC is a micro-economic data file, union and nonunion establishments can be studied separately without recourse to indirect estimation procedures based on regression interaction effects. Moreover, union effects relating to production and nonproduction worker wages and fringe benefits can be readily modeled.

Except for information on union status, the EEEC and SIPP-ED file would be structured in a similar manner. A SIPP-ED file, however, would contain important economic data not found in the EEEC, e.g., asset and output data -- both of which are necessary for estimating the elasticity of substitution between different qualities of labor and between labor and capital. It should be noted, too, that because of confidentiality requirements, information for some large establishments, representing 15 percent of employment in manufacturing,[91] were deleted from the EEEC public use file. The EEEC also lacks demographic data. Although such data can be merged into the EEEC from CPS files, individuals cannot be matched to their employer, thereby reducing the variability of the data. While the EEEC surveys still retain their utility, the last one was conducted in 1977.[92]

One of the distinctive features of the EEEC was a question on whether 50 percent or more of nonoffice and office workers were covered by a collective bargaining agreement. A similar question is not asked on any questionnaire in the economic censuses. A SIPP-ED file could serve to partially fill this gap. This can be accomplished by defining a union establishment or company as one in which 50 percent or more of the relevant group of SIPP respondents is covered by a collective bargaining agreement. In the manufacturing and minerals industries, the relevant group of respondents would be production workers; in construction, it would be the construction trades. For wholesale and retail trade, the relevant group might be sales and clerical workers; in the service industries, service and clerical workers; and in the remaining industries, e.g., real estate, clerical workers. In the most likely case where only a single worker

is sampled from a firm's relevant work force, the union status of the firm would be the same as the union status of the worker.[93]

Although the convention suggested above enables one to identify a unionized firm when the sampled workers in SIPP are in a relevant group, it cannot be used when the sampled workers fall outside the group. In 1980, for example, 26.0 percent of the workers in manufacturing were in occupations other than those associated with production. If the ratio of production to all workers is independent of establishment size, the union status of approximately one-quarter of the establishments in manufacturing would not be ascertainable. Nonetheless, it is still possible to check the accuracy of the proposed convention. Considering manufacturing as an example, if for any given size class of establishments, the ratio of production to all workers is unrelated to the establishment's union status, the estimated number of production workers in establishments identified as unionized should equal the population weighted number of workers in SIPP who are union members times the percentage of manufacturing workers engaged in production in that size class. If so, the establishments identified as union and nonunion would be a representative sample of all union and nonunion establishments. Hence, the inability to classify all manufacturing establishments in a SIPP-ED file in terms of union status should not inhibit analyses of union labor market effects that are not discernible from other currently maintained data sets.

It should be noted that it is a relatively simple matter to modify the SIPP questionnaire so that the union status of all establishments can be determined. Respondents might be asked if any employees are covered by a collective

bargaining agreement at the location where they work. Workers answering in the affirmative could then be asked to estimate the percentage of nonoffice and office workers covered by such an agreement. Inclusion of questions in the economic censuses similar to those in the EEEC would, of course, obviate the need of inferring the union status of an establishment from SIPP data.

In assessing the union wage effect, consideration needs also to be given to the union productivity effect. Union make-work rules can reduce productivity, but unions also contribute to productivity by improving worker morale and motivation, expanding formal and informal on-the-job training while at the same time protecting more experienced workers under seniority rules, and providing a mechanism whereby worker dissatisfaction is alleviated through labor-management dialogue rather than labor turnover.

The effect of unionization on worker productivity in manufacturing has been estimated by Brown and Medoff (1978) using a Cobb-Douglas production function which relates output, measured by value added, to the amount of capital and labor inputs utilized by a firm. Brown and Medoff utilized 1972 CM data classified by 2-digit SIC and state area. In the absence of information that would enable them to distinguish between the union and nonunion sector of manufacturing industries, they modified the production function so that output depended on the fraction of labor unionized in each industry-state area (derived independently from CPS data) and the capital/labor ratio. Based on their analysis, they concluded that other factors held constant, including labor quality,[94] union establishments were 22 percent more productive than nonunion ones, and that the differential in productivity approximates the union wage effect.[95] A

similar approach using the Brown and Medoff estimating procedure, and adjusting for differences in housing construction costs among geographical areas, was applied by Allen (1984) to 1972 CCI data. A substantial positive union productivity effect, between 17 and 22 percent, was also found by Allen for the construction industry, but this was about half the estimated union wage effect in this industry.

These studies break new ground in understanding union effects on the labor market. They also illustrate areas of inquiry where a SIPP-ED file can be applied. With such a file, the union productivity effect could be estimated with greater accuracy. Because Brown and Medoff, and Allen, were unable to separately estimate production functions for union and nonunion establishments, their analyses rest on the assumption that the elasticity of output with respect to capital is the same in both.[96] To the extent that the productivity of capital is greater in the union sector, the Brown and Medoff model overestimates union productivity. With a SIPP-ED file, production functions could be estimated for union and nonunion establishments within each industry-state area, and capital productivity differences between union and nonunion establishments could be controlled for within the model.

Another advantage of a SIPP-ED file stems from its being a micro-level file. The published data used by Brown an Medoff are aggregated to the 2-digit SIC level. At this level of aggregation, it is not clear that an industry's

technology is invariant among state areas; for example, SIC 37--transportation equipment--includes the motor vehicle, aircraft, shipbuilding, railroad equipment, bicycle, and guided missile and space vehicle industries. It is probable that the distribution of employment among these industries differs among state areas. A related problem is that when union and nonunion establishments sell in separate markets within a 2-digit SIC, it is possible that the higher paying union establishments sell their output at a higher price than nonunion establishments. Thus, the union productivity effect may reflect a union price effect. These difficulties can be mitigated with a SIPP-ED file by grouping establishments at the 3-digit SIC level where the assumptions that firms face the same technology and sell their output at the same price can be maintained with greater confidence.

Another area where unions may have an important impact is on employment. Whereas much attention has been focused on the employment effects of minimum wage legislation, comparatively few studies have been made of union employment effects.[97]

Economic theory suggests that in competitive labor markets, union wage increases are achieved at the expense of reduced employment or hours worked by union members. But the extent to which this is true is difficult to ascertain because of the sensitivity of union employment effect estimates to small changes in the specification of estimating equations applied to longtitudinal data (Pencavel and Hartsog, 1984) and the lack of cross-sectional data pertaining to output and capital intensiveness of union and nonunion establishments. Since, in cross-sectional studies the dependent variable is the level of employment,

another measure of firm size besides employment must be found to control for differences in production levels among establishments. Absent information on output, it must be assumed that production levels in union and nonunion establishments are the same. Likewise, absent information on capital, it must be assumed that union and nonunion establishments utilize the same amount of capital. Even where both output and capital data are available, a finding that union establishments employ fewer workers than nonunion ones may be explainable by differences in productivity between union and nonunion labor. As indicated, differences in productivity between union and nonunion workers could be estimated from SIPP-ED data. By taking account of productivity differences among industries, and also variations in output and capital utilization between union and nonunion establishments, understanding of the relationship between union wage and employment effects can be significantly enhanced.

In manufacturing, union establishments are more than twice as large, in terms of employment, as nonunion ones.[98] Union and nonunion establishments in manufacturing are different in other ways: in union establishments value added is substantially higher, the ratio of capital to production labor hours is also much higher, and the ratio of nonproduction worker hours to production worker hours in moderately lower.[99] As noted in the discussion of the union wage effect, union establishments in manufacturing employ higher quality labor and a higher proportion of their labor bill is comprised of voluntary supplemental labor costs. Not so clear but of some interest is the relationship

between a manufacturing firm's union status, its capital expenditures per worker for new machinery, and the proportion of its capital expenditures spent on computer hardware. More generally, to what extent are relationships between union and nonunion firms found in manufacturing representative of other industries? A SIPP-ED file offers promise of providing at least a partial answer to these questions.

I. Productivity Analysis

One of the more self-evident applications of a SIPP-ED file, and the one where the lack of micro-worker data has been keenly felt, is in the area of productivity analysis. Central to productivity analysis and, indeed, to the theory of the firm, is the production function which maps inputs into outputs.[100] Traditionally, inputs into the production process have been classified into three broad classes, i.e., land, labor, and capital. For the economy as a whole and, for simplicity, for industries and firms, land has been treated as if it were a fixed factor of production, i.e., incapable of being changed by significant amounts. In this circumstance, the level of output depends only on the different quantities of labor and capital entering into the production process. To produce final products, however, other factors, such as materials and fuels are also needed. But if production is viewed simply as the transformation of intermediate goods into final products, output can still be thought of soley as the end result of combining the cooperative efforts of only two agents, labor and capital. Viewed in this manner, raw materials and fuels play only a passive, i.e., parametric, role in the production process; the greater the amount of materials and fuel, the more output that can be produced by

any given combination of labor and capital. In this conceptualization of the production process, the output generated by labor and capital is net output or value added as measured in the GNP accounts.[101]

For the economy as a whole, intermediate goods cancel as both inputs and outputs, but they do not cancel within a firm or industry. At the latter levels of aggregation, materials and fuels and similar intermediate goods should be included, if possible, in the production function. This is particularly important where the technology of production is such that capital and/or labor can be substituted for materials and fuels as, for example, in the generation of electrical power. Where this is the case, the correct measure of output is gross output as measured by total value of shipments or operating and sales receipts.[102]

In the well-known work by Douglas,[103] data for manufacturing were used to estimate the elasticity of net output with respect to labor and capital. These elasticities yield estimates of the labor and capital share of the value of output which are consistent with those based on national income data. While most studies of productivity employ data aggregated by industry, some studies have been based on plant data.[104] Plant data, however, are relatively uncommon because of the difficulty of collecting the requisite information from establishments. The development of the LED is a major effort to overcome this shortcoming.

The advantage of micro-firm data in productivity studies is clear: they portray substantially more complex variations in output and in inputs then do industry data and, hence, permit more detailed testing of alternative specifications of

production functions. Moreover, micro-firm data provide information on capital, material, and fuels so that substitution effects between these variables and labor can be examined at the plant level, as well as substitution effects between different kinds of labor and different kinds of capital, holding other factor inputs constant. The new insights to be gained from analyzing micro-firm data will surely be as large as those realized when micro-worker data became readily available in the 1970's. Matching micro-worker and micro-firm information will not only further increase the variance of the data, but will enable investigators to examine questions and issues not heretofore amenable to analysis.

Irrespective of the level of aggregation of output and inputs, a major problem in understanding the determinants of productivity is measuring the quality of labor. One approach to this problem is to measure the quality of labor in terms of its marginal productivity as reflected by wage rates. Underlying this approach is the assumption that wage rates are determined soley by marginal productivity and not at all by employee credentials or employer perceptions of the worth of jobs. A second assumption is that economic models can adequately explain wage rate variations. Granting these assumptions, an index of skill level, i.e., quality, can be calculated by combining experience and education where the weights are the coefficients derived from a wage rate regression equation whose arguments are experience and education, holding constant age, sex, race, occupation, and other variables. Multiplying hours worked by the index of quality yields a measure of labor adjusted for quality.[105] This approach, based on the work by Gollop and Jorgenson (1980), is currently

being implemented by BLS to refine its multifactor measure of productivity (Bureau of Labor Statistics, 1983).

Although the Gollop and Jorgenson approach provides an index of worker quality as a by-product of adjusting for the experience and education embodied in a given work force, it introduces an extra step which can lead to misestimation of the relationship between output and investment in human capital. To avoid this difficulty, human capital variables should be entered explicity into the production function.[106]

For many labor market studies, and also for productivity analysis, it is important to know an individual's work experience. Since almost all men work upon leaving school, their work experience is typically estimated by their age minus years of schooling minus six (assuming that the normal school entry age is six years). Work experience for women cannot be estimated in this manner as many women leave the labor force one or more times. Information on the work experience of women is lacking in the CPS and decennial censuses of population -- the data sets used by investigators to adjust labor inputs for differences in quality.[107] In SIPP, the work experience of women, as well as men, can be determined from questions which ask how many years a person has worked for six or more months; additionally, SIPP provides information on up to four periods of six months or more in which an individual did not work. Another measure of experience, which may be more closely related to worker productivity than the total number of years spent in the labor force,

is tenure on one's current job (Jusenius, 1977). Job tenure information is collected in SIPP.

Another kind of human capital formation, i.e., investment in formal education, has been given particular attention in the literature as a factor contributing to labor productivity.[108] Besides the conventional questions relating to years of school completed, SIPP provides information on highest degree attained, year highest degree attained, and major field of study. Inclusion of these variables in the production function could yield more accurate estimates of the contribution of education to economic growth. Moreover, other variables found in SIPP pertaining to frequency of turnover, vocational training, and health and disability, measure aspects of investment in human capital besides education.

The effect of including a wide range of variables available in SIPP into the production function could be large. Some indirect evidence of this can be inferred from Norsworthy and Zabala's (1985) study of the relationship between worker behavior and productivity and costs in the automobile industry. Their translog cost model with an embedded worker behavior function is found to have "significantly greater explanatory power than the standard model" (p. 19).[109] Of the four variables they use to derive an objective measure of worker behavior, i.e., grievances filed, grievances unresolved, unauthorized strikes, and quits, only the last is included in SIPP; however, information on total weeks lost due to strikes is included in SIPP, and this may correlate with grievances and unauthorized strikes.

By matching SIPP and economic data, individual workers can be paired with their employer. The particular advantage of matching worker and firm records, rather than simply merging the data and aggregating up to the industry level, is that the affect on output of differences in the composition of the work force among industries, i.e, the marginal productivity of alternative qualities of labor, can be determined, holding constant firm attributes.

Of special significance, estimates of the marginal productivity of various qualities of labor could then be compared with their corresponding wage rates, providing a consistency test of marginal productivity theory at the micro-level comparable to the one utilized by Douglas at the macro-level. This contrasts with the approach taken in wage rate studies where the quality of labor is taken as a proxy measure of worker productivity, and the focus is on ascertaining how well wage rates are explained by labor quality.

The relationship between labor quality and output is largely an unexplored area within the larger framework of productivity analysis, but another fundamental issue is how to measure output itself, i.e., in net or gross terms. The recent study by Norsworthy and Malmquist (1983) of long-term productivity growth in the United States and Japan is illustrative of this problem. Their main finding is that in manufacturing a Japanese worker has substantially more capital to work with than an American worker, even after adjusting for differences in wage rates. An equally important conclusion from a theoretical viewpoint is that the value added measure appears to be inappropriate for productivity

studies based on aggregated industry data, in particular, data for all manufacturing industries.[110] Further verification of this conclusion based on micro-worker and micro-firm data, which would allow one to control for the quality of labor, can provide a more definitive answer to this critical issue.

It should be mentioned that since SIPP panels are followed only for two and one-half years, the primary use of a SIPP-ED file in the area of productivity analysis would be in cross-sectional studies. However, it may be possible to link successive panels of workers and their employers to form a longitudinal file at the same level of industry detail that would be used in cross-sectional studies.

It should also be recalled that capital information is lacking for small establishments in manufacturing, all but the largest establishments in the service sector, and all small enterprises regardless of industry. This limits, at least for the near future, the range of firms for which micro-level productivity studies can be performed. For large enterprises, however, there are data sets which if brought together would greatly enrich the materials available for studying productivity and other related issues. A firm's decision to invest in an asset depends on a number of determinants, a principal one being its profit position. Likewise, the decision to undertake research and development projects is related to the availability of funds to finance them. Both the acquisition of new capital and research and development lead to higher productivity.[111] More generally, firm profits, investment, R&D expenditures, and employment are functionally related. While capital investment

and employment growth are positively related, productivity and employment growth can be positively or negatrively related. Whether the latter relationship is positive or negative and the circumstances under which it is one or the other have significant policy implications. By linking the ES, R&D, and QFR files, it should be possible to track these complex processes. Since a substantial proportion of total output is produced by the large companies contained in these files, a better understanding could also be achieved of the relationship between micro-level firm decisions and macro-level outcomes.

3. Methodological Problems in Matching SIPP and Economic Data

In this section, attention is focused on two methodological problems. One is central to the development of a SIPP-ED file; the other is peripheral but places a constraint on the way in which the file can be applied. The first problem deals with procedures for matching workers to their establishment and company. The second relates to the estimation of data which are not universally available in the economic censuses.

A. Procedures for Identifying an Individual's Employer

Essential to the creation of a SIPP-ED file is the ability to determine the establishment and/or company in which a person is employed. The most promising and least expensive way of doing this is to match on firm name and physical address of an individual's place of work. Questions relating to both pieces of information are being asked of SIPP respondents. Although the physical address is not necessary for identification of an individual's work place, its availability greatly facilitates the matching process.

For employers with only one establishment in an area, the firm name and employee's address will typically be sufficient to determine where a person is employed. For companies with more than one establishment in an area, the firm name and employer's address should be sufficient to identify the place of work. If an employer has more than one establishment in an area and the place of work cannot be determined using the employer's physical address, other information in SIPP can be utilized. Thus, in cases where more than one activity is located at a given physical address, e.g., a firm's manufacturing and sales activity may be located in the same building, the kind of work an individual does can be used to determine in which activity he or she works. When no address is available, the kind of business in which an employer is engaged and a respondent's estimate of size of establishment can be used to identify the person's work place. For example, a firm manufacturing bottles may have several establishments but only one large plant in a local area. In the event that an individual's work place cannot be identified, as may happen in the case of person working for a chain of fast-food stores, the person can be randomly assigned to one of the stores.

Another aid in identifying an individual's work place is the EIN. While a company may have a number of establishments in a local area, its subsidiaries, when identified by their own EIN, may have only one establishment in the area. Thus, the EIN of the employer for whom an individual works can be sufficient to uniquely determine the establishment in which that person is employed. In the annual roundup fixed topical module, EINs are being recorded from W-2 forms when the latter are referenced by respondents.[112]

A third aid in identifying an individual's work place needs also to be mentioned. In developing the journey to work statistics from the 1980 Census of Population, a Major Empoyer List (MEL) was prepared in which are given the exact physical address and zip code of large establishments located in SMSAs extant as of the date the decennial census was initiated.[113] An advantage of this list is that it cross-classifies both firm and company names, facillitating entry into the SSEL. Thus, if a respondent reports he or she worked for firm A and firm A is part of company C, in searching the SSEL it is a good deal easier to determine which of company C's establishments the respondent worked in if company C is known. This is particularly so when all of company C's establishments are covered by a single EIN. Additionally, where only a partial address is given by the respondent, the complete address, including the zip code, can be obtained from the MEL. Knowing the zip code can materially reduce the search time needed to identify an individual's work site, particularly if the person works in a labor market whose boundary crosses several states.

One problem not noted but which is of practical importance is the time lag between data obtained from SIPP and the posting of establishment data in the SSEL. Because of this lag, an imputation process must be used to link employees and the firms for which they work when the latter are newly established. In these cases, actual employers can be identified but only after the SSEL is updated to include new firms extant during the period a SIPP wave was interviewed.

B. Estimating Missing Economic Data for Small Establishments and Companies

Some economic data, e.g., the gross value of depreciable assets and supplemental labor costs, are not universally available for all establishments and enterprises. Given that the primary objective of the Bureau of the Census in collecting economic data is the measurement of industry output, and that the bulk of output in any industry is produced by large establishments and companies, the quantity and quality of the data that is collected for small businesses need not be complete or as accurate as that for their larger counterparts. This is particularly so for accounting data, such as asset information, since accounting practices may vary from firm to firm. Even if substantial improvements were made in estimating missing economic data for small establishments and companies, these would translate into only small improvements in estimating industry aggregates. Hence, while there is interest in improving the economic data for firms of all sizes, to do so by collecting additional data from small ones would require a disproportionately large outlay of resources.

In the CM, asset values for small establishments are imputed by multiplying their value of shipments by the average capital/value of shipments ratio for larger establishments in the same 4-digit SIC class (as calculated from the ASM).[114] Despite the fact that asset information is not collected for small plants in manufacturing, unless they are in the ASM, it may be possible to obtain reasonably accurate estimates of capital for these small plants using an economic model.

Economic theory suggests a number of relationships which influence the amount of capital that a firm employs in its production process. In particular, since the capital/labor ratio varies inversely with establishment size, it seems reasonable to assume that information about the number of employees in establishments in the same or closely related industries can be used to derive estimates of their capital assets. Additionally, holding establishment size and other factors constant, low wage establishments will substitute labor for capital in order to economize on the use of the relatively expensive factor, i.e., capital; thus, low wage establishments will tend to have a lower capital/labor ratio than high wage establishments.

Even among establishments of the same size whose average wage rate is also the same, one would expect a higher capital/labor ratio the smaller the ratio of production workers to all workers, since high capital/labor ratios are implied by high monitoring costs. Additional relationships between assets and other variables may exist. For example, it may be that newer establishments in an industry are more capital intensive than older ones; likewise, regional variations in entrepreneurial ability may give rise to corresponding variations in capital intensity.

Besides economic relationships, engineering relationships also may be useful in estimating capital intensity. For example, it is plausible that in manufacturing an establishment's capital/labor ratio is positively related to purchased electricity per employee; while the former is known only for establishments in the ASM, the latter is available for all but the smallest establishments in the CM.

It will be noted that the dependent variable in this economic model is the capital/labor ratio rather than the capital/value of shipments ratio. When firms change their level of output, they do so by adjusting overtime hours as well as the number of workers in their employ. To the extent that overtime hours is chosen to adjust labor inputs, variations in employment over the business cycle will be smaller than the corresponding variations in value of shipments. For this reason, where the focus of inquiry is specifically on micro-estimates of firm capital, use of the capital/labor ratio as a multiplicative factor should yield more stable estimates and, hence, is preferred to the capital/value of shipments ratio.[115]

An assessment of the utility of an economic model to estimate the assets of an establishment can be made using the ASM. Given the large size of this sample, an economic model can be fitted to establishments in closely related industries, omitting a subset of observations which could be used as a "live" test of how well the model performs against the current procedure. Such a test should be restricted to establishments with information as originally reported by respondents to avoid cases where reported values have been computer or analyst corrected to conform to Bureau of the Census editing procedures.

Finally, an economic model can also be utilized to estimate employer contributions for fringe benefits where this information is missing in the economic data. It is plausible to assume that a firm's fringe benefits are related to its size, average wage level, legal form of organization, industry, and region where it is located. Given a firm's fringe benefit expenditures, whether obtained directly from the economic data or estimated from an economic

model, the (average) market value of the fringe benefits received by each of its employees can be determined. From a SIPP-ED file, one could then estimate the population weighted market value of fringe benefits received by various demographic groups, e.g., men and women, whites and blacks, etc. Although it should be evident from the discussion of this paper, this last illustration is indicative of the ways in which a SIPP-ED file can be applied.

4. Concluding Remarks

The principal objective of this study has been to assess the availability, sources, coverage, and content of various economic data files maintained by the Bureau of the Census and to examine study areas and issues to which a data set combining micro-worker and micro-firm data could be applied by investigators. In the course of the study, specific demographic and economic variables have been identified which should be incorporated into such a data set, which we have referred to as SIPP-ED file. Additionally, it was anticipated that methodological problems inherent in this undertaking would be revealed; indeed, this has been the case.

In looking at the problem of how to identify an individual's work place, it became apparent that an important piece of information for doing this is the employer's physical address. This information is now being collected in SIPP in the "annual roundup" waves. Given the individual's work address and the name of the employer, it should be possible to determine the establishment in which he or she works from the Standard Statistical Establishment List

(SSEL). However, respondents other than the person for whom information is being sought may not know the person's work place address. Even when the employer's physical address is obtained, it may not be correct or complete, e.g., the zip code may be missing or a building name may be given rather than the street address. An aid in locating an employer's physical address has also been identified, namely, the Major Employer List (MEL), which was compiled in developing journey to work statistics from the 1980 Census of Population. The MEL can be used to determine the physical address of large establishments in metropolitan areas. Other sources of information can also be utilized to determine a person's place of work, e.g., Dun and Bradstreet reference books, Standard and Poor directories, and telephone books.

Once an employee and employer have been matched in the SSEL, identifiers can be extracted from this file which enable one to access the economic data files maintained by the Bureau of the Census. Some economic data can be extracted directly from the SSEL; indeed, the SSEL is the only source of economic data maintained by the Bureau of the Census for firms that are out-of-scope with respect to the economic censuses. Other files where more extensive economic data can be obtained include the Enterprise Statistics (ES) file, the Longitudinal Establishment Data file, and the economic census files themselves.

The ES file is particularly suitable for linking workers and their firm because data from all the economic censuses are brought together in this file along with data from a special questionnaire sent to large enterprises. Thus, it

will be possible to retrieve economic data at feasible cost for the establishment and enterprise in which an individual works.

A few illustrations of the study areas and issues which can be examined with a SIPP-ED file should suffice to indicate the returns to linking the micro-worker and micro-firm data sets maintained by the Bureau of the Census. As a by-product of identifying an individual's employer, studies of employment and unemployment by place of work (versus place of residence) can be undertaken; intercensal data on journey to work by the labor force characteristics of workers could also be derived. While much is known about the determinants of wage rates, little is known about the relationship between low pay among workers and their employment in low wage firms. For example, it may be that low wage employers provide employment opportunities with attendant low earnings, not because they discriminate against certain groups of individuals, but because the production processes that are most efficient for their mode of operation do not require high quality labor, and, furthermore, inhibit their paying high wages. Investigators have long been stymied in measuring the employment effects of minimum wage legislation because of the difficulty of identifying low and high wage firms. Likewise, they have had great difficulty in quantifying union employment effects. Verification of the dual labor market theory has been hampered by the lack of data on mobility of workers between employers who comprise the primary labor market and employers who

comprise the secondary labor market. In studying structural unemployment, attention has focused almost exclusively on supply side aspects because of the lack of information on changes in plant employment. And little is known about the income distribution effects of the shift in employment from the goods to the service sector where small firms are prevalent. In the area of productivity analysis, it is not possible to obtain extensive, refined measures of labor quality. If one attempts to determine whether union workers are more productive than nonunion workers, one is faced with the problem that one cannot separately estimate production functions for unionized and nonunionized firms. These and similar issues could be resolved by linking SIPP with Census economic data sets.

Economic data, e.g., employment and payroll data, are available from the SSEL for all private nonagricultural establishments and companies. A larger set of economic data, including information on the gross value of depreciable assets and fringe benefits, is available for large establishments and large enterprises in in-scope industries. Besides these data, the Bureau of the Census maintains special files with R&D expenditure and income and balance sheet information for large companies. Thus, for large companies a wide range of micro-level data can be explored and related to analogous macro-level data.

While there is a plethora of economic data, important gaps exist. In particular, universal coverage for asset and fringe benefit data is available only for large establishments and enterprises in in-scope industries. These gaps

can be filled however, by developing economic models to estimate assets and fringe benefits from sample data in the Annual Survey of Manufactures and the Capital Expenditure surveys.

A perennial problem with linked data is their confidentiality. One way of maintaining the confidentiality of a SIPP-ED file is by inviting scholars to visit the Census Bureau through the American Statistical Association/Census Research Fellowship and Associate Program. An attractive proposal would be to establish joint working agreements between the Census Bureau and governmental agencies; nonprofit organizations; and individual academic researchers with university and foundation support. As this paper attempts to indicate, a SIPP-ED file holds great promise for those concerned with public policy.

FOOTNOTES

1. Out-of-scope industries with respect to the ES and the economic censuses include transportation, communications, and public utilities; finance, insurance, and real estate; and some service industries. The ES data for 1982 will include agriculture.

2. Nelson, McMillen, and Kasprzyk (1983).

3. Ibid.

4. Ibid.

5. Ibid.

6. For a discussion of the measurement of the labor force, employment, and unemployment based on SIPP and CPS data, see Ryscavage (1984).

7. For ease of exposition, the terms company and enterprise are used interchangeably. Both terms reference a parent company, comprised of one or more establishments and/or subsidiaries, which is a completely independent business organization.

8. Agricultural firms owned by multi-establishment enterprises in both the agricultural and nonagricultural sectors are also included in the SSEL, but single establishment enterprises in agriculture are excluded.

9. The Alpha number identifies a parent company in the ES and, hence, is the link for integrating the economic data in the SSEL for establishments with similar data in the ES for the parent company.

10. To further illustrate the structure of the SSEL, consider the following cases:

Company	Establishment or Subsidiary Number	Number of EINS	Alpha Number
A	1	1	No
B	1 2	2	Yes
C	1 2	1	Yes

Company A has only a single establishment; no Alpha number is found in the EC-EI file because it is not a multi-establishment firm. Company B has two subsidiaries, each with an EIN, since Company B maintains a decentralized reporting system. Only the names of these subsidiaries appear in the EC-EI file; each subsidiary has the same Alpha number which identifies Company B. Company C also has two subsidiaries but has only one EIN, since it maintains a centralized reporting system. Only Company C's name and its Alpha number appear in the EC-EI file; the names of its subsidiaries are not included in the file because

they have no EIN. Both Company B and Company C with their corresponding Alpha number are in the EI-MU file. Additionally, the two subsidiaries of both companies, as well as all of the establishments owned by each subsidiary, are included in the EI-MU file under their respective parent company's Alpha number.

11. In the illustration of footnote 10, the physical address of Company A appears in the EC-EI file; the physical address of each establisment comprising companies B and C are in the EC-MU file. If either parent company has a physical location, the address of that location is also in the EC-MU file. The mailing address of each parent company, irrespective of whether or not it has a physical location, is contained in the Master Mailing Address file. For single establishment companies such as Company A, the mailing address, as well as the physical address, is in the EC-EI; however, for some single establishment companies only a mailing address, e.g., a Post Office Box number, is available.

12. The CFN is composed of 10 digits. For single establishment enterprises, the CFN is the company's EIN preceded by a zero. For establishments of enterprises with two or more establishments, the CFN is the parent company's 6 digit Alpha number followed by a four digit establishment number. When establishments in the latter category change ownership, both the old and new CFNs are retained in the SSEL.

13. Economic census years only and only for businesses identified by an EIN in the EC-EI file.

14. A more accurate estimate of the average wage can be obtained by correcting the pay data to take account of hours worked per employee in the 1st quarter in the industry in which the employee is employed. Hours worked per employee can be derived from SIPP.

15. Additional information about the SSEL is found in Bureau of the Census (1979).

16. Although the ASM (and CM) certainty criteria have changed over time, continuous longitudinal data are available in the LED for large establishments with 250 or more employees.

17. In 1984, the ASM sample size was approximately 57,000. Of this number, about 19,000 were establishments of certainty companies with $500 million or more in shipments of manufactured goods. Another 6,700 and 5,500 met the establishment certainty and product certainty criteria, respectively. The noncertainty establishments numbered 26,000.

18. Short and long form questionnaires are sent to the establishments in the ASM. In the non-census years 1984-86 and 1988, a long form is being sent to all establishments of multi-unit enterprises (approximately 39,000 places) and to single establishment enterprises with 18 to 35 or more employees, depending on industry (approximately 13,500 plants). The remaining establishments, accounting for about 4,700 of the 57,000 in

the ASM, receive a short form. Additionally, economic data are available from administrative records of other agencies for plants not included in the ASM sample, but these data are maintained in a separate file.

19. Of the 340,000 manufacturing establishments in the 1982 CM, 70,000 establishments of multi-unit companies (34,500 of which were in the ASM) and 20,000 single unit companies (all of which were in the ASM) received a long form. Of the remaining 250,000 single plant non-ASM companies, 65,000 and 50,000 received a long and short form, respectively, The plants receiving a short form were in selected 4-digit SIC industries where they comprised a small fraction of all establishments or accounted for a small fraction of the total value of shipments. Administrative records are found in the CM file for the 135,000 small, single unit enterprises that did not receive either a long or short form.

20. In the list of data elements given in the text, ASMS, ASML, and ASM denote information contained only in the short form, only in the long form, and in both the short and long form ASM questionnaires. Similarly, CMS, CML, and CM denote information contained only in the short form, only in the long form, and in both the short and long form CM questionnaires.

21. In economic census years, all ASM plants receive an ASM long form which also serves as their CM long form.

22. Value added is the difference between the total value of shipments and the cost of materials and services plus the net change between beginning and end-of-year inventories.

23. Additional information about the LED is found in Monahan (1983).

24. The ES data also include in-scope eatablishments of companies primarily engaged in out-of-scope industries, e.g., retail appliance stores of a public utility are included in the ES.

25. In the ES file, single establishment enterprises are accessed by using the establishment's CFN; otherwise, as noted, the Alpha number is used.

26. Census Enterprise Industry Categories are based on the Enterprise Standard Industrial Classification system which, in turn, translates into 2- to 4-digit SIC codes depending on industry (see Office of Management and Budget, 1974 and Bureau of the Census, 1982).

27. Single establishment companies and multi-establishment companies all of whose establishments are engaged in a single 4-digit SIC industry are classified as single-industry enterprises. All other companies are denoted as multi-industry companies. From payroll data, the 4-digit SIC of companies in the latter category is obtained by first ascertaining their largest census industry division (e.g., manufacturing, retail trade) and then determining their primary 4-digit SIC industry within that division.

28. For mineral and manufacturing industries, employment is given by the average of paid production workers for the pay periods including the 12th of March, May, August, and November plus all other paid employees in the pay period including March 12; for construction, employment is given by the average of all paid workers in the pay periods including the 12th of March, May, August, and November; for all other industries, employment represents the number of employees on the payroll in the pay period including March 12.

29. For companies with fewer than 500 employees classified in the mineral, construction, and manufacturing industries, capital expenditures are only cumulated for their operating establishments in these three industries.

30. Also available at the time the ES files are processed are administrative records for firms with and without employees.

31. Where it is useful to include additional data elements in the ES besides those noted above, say, from the CM, the establishment data can be aggregated by enterprise.

32. Besides the data described in this section, the Bureau of the Census also maintains a specialized file containing concentration ratios for manufacturing industries at the 4-digit SIC level.

33. In cases where significant amounts of products classified in different SICs are produced and shipped from the same location, the different activities may be identified as separate establishments in the CM. Most often, however, different products are made on the same production line by the same employees and cannot be distinguished in a firm's records; in these instances, establishment and plant are synonomous.

34. In the SSEL, each state in which an oil and gas extraction and field services company operates is treated as if were a separate establishment. Additionally, off-shore installations in specified local areas are grouped together as a single establishment.

35. In the 1977 CCI, approximately 180,000 establishments were sampled from a universe of about 550,000 establishments with paid employees.

36. As noted, hours worked can be derived from SIPP. This is also the case for other industries where hours worked are absent in the micro-firm data.

37. Monthly surveys are also conducted for wholesale and retail trade. The data from these surveys are utilized by the Bureau of Economic Analysis in constructing GNP estimates.

38. In 1977, the CE surveys covered approximately 3,500 firms in wholesale trade, 29,500 in retail trade, and 11,500 in selected service industries.

39. The 1977 CE surveys for retail trade and services included firms without paid employees. The 1982 CE surveys exclude such firms.

40. The data for determining whether a company meets the certainty cutoff for a particular SIC are obtained from the economic census prior to a CE survey. The level of industry aggregation, i.e., 2-, 3-, or 4-digit SIC level, varies to conform to specified sampling error constraints. For purposes of estimating GNP, the CE data are usually allocated to 4-digit SIC industries on the basis of sales and other variables.

41. In some instances, multi-establishment companies provide disaggregated data. To simplify the exposition, it is assumed that the data are aggregated for all of a company's units in an industry.

42. Data for the service establishments might be available from the CE survey for service industries but only if they were included in that survey.

43. In 1977, 82.1 percent of the 1.9 million establishments in retail trade and 95.1 percent of the 1.8 million establishments in in-scope service industries were operated by single establishment companies. The comparable figure for the 380 thousand wholesale trade establishments was 69.6 percent.

44. Information on inventories is collected in the CWT. Because inventories are minimal in the service industries, inventory data are not collected for this industry.

45. As the SICs of reporting firms in the CE surveys are derived from the business census preceding the one in which the surveys are conducted, the survey SICs can differ from the ones developed from the most current business census. The SICs in the CE surveys have, heretofore, not been updated.

46. As in the case of the ES, companies can be identified by their CFN or Alpha number. Unlike the ES, industries are classified by SIC code.

47. Since 1978, the short questionaire has been used in even numbered years, the long one in odd numbered years. Only a long form was utilized prior to 1978.

48. R&D information is solicited from a second panel of approximately 1,600 companies which also rotates every five years, but this additional panel has in the past received only one questionnaire during the rotation period. This questionnaire is limited to information on net sales and receipts, total employment, employment of R&D scientists and engineers, and total R&D expenditures by domestic and by foreign affiliates.

49. The coding system for identifying firms differs from that of the SSEL. In the QFR, firms are identified by their EIN. These firms can be linked to their parent enterprise through the SSEL.

50. In determining which major division a firm is in, a plurality rule is used based on gross sales receipts for in-scope industries. A plurality rule again based on gross receipts, is also used to determine the detailed industry of a firm. In the QFR, a modified version of the Enterprise Standard Industrial classification system is used to classify firms to a 2-digit level of detail (see Brannen, 1982).

51. The QFR is a quarterly report. Of the almost 4,600 certainty firms, as of the second quarter, 1984, about 60 percent were in manufacturing.

52. In the CM and other economic censuses, information is only available on the gross value of depreciable assets.

53. One reason for more efficient, larger firms to acquire hire higher quality labor is that potential losses in output, when production is disrupted due to worker absence, negligence, or error, are greater than for less efficient, smaller firms. Hence, it pays the former to incur monitoring costs or to attempt to defray them by hiring higher quality labor (Oi, 1983).

54. Indeed, the difficulty of screening applicants has given rise to the theories of signalling (Spence, 1974) and of statistical discrimination (Phelps, 1972).

55. Of interest, when Brogan and Erickson included industry concentration and establishment size variables in their wage equation along with a capital per establishment variable (based on aggregated industry data), they found that while the last variable remained highly significant neither the concentration or size variables were statistically significant.

56. Kwoka (1983) summarizes some of the major wage rate studies in which industry information on plant size is utilized. In the seven studies cited in which plant size is a variable entering the analysis, it has a positive and significant effect on wages in every one.

57. In the Mellow study, size of plant and size of firm are assumed to be independent variables. The assumption is not met, however, for single establishment firms where, by definition, plant and firm size are the same.

58. Some evidence for this latter proposition, based on a 1968 BLS survey of 1,149 establishments, is found in Bailey and Schwenk (1971).

59. Of practical import is the problem of determining whether a given kind of OJT is specific or general. The capital/labor ratio provides a way of proxying the specificity of OJT. It is likely that in establishments where the capital/labor ratio is high, the capital is sophisticated and performs much of the work. The role of the employee is to insure that the equipment is maintained in good order and its capabilities are fully utilized, but to succeed in these tasks specific training is often required. On the other hand, specific OJT should be less prevalent in establishments where the capital/labor ratio is low.

60. Kwoka (1983).

61. On the other hand, young firms may have more modern equipment which would contribute to higher productivity and earnings.

62. In 1977, fringe benefits accounted for 15.5 percent of total compensation (Bureau of Labor Statistics, 1980).

63. Doeringer and Piore (1971).

64. Bureau of the Census (1984).

65. Minimum Wage Study Commission (1981), Chapter 1.

66. Salop and Salop (1976).

67. The rapid growth in labor force participation among women and their possible absorption by low paying firms would be consistent with the puzzling phenomenon of why the female/white male wage rate failed to decline during the decade of the 1970's (see Green, 1984).

68. As an example, a low paying firm might be defined as one with an average wage rate which is less than one-half the median of all firms; this criterion is similar to one proposed by Fuchs (1967) in defining the poverty level of income. A desirable feature of this criterion is that it focuses attention on the distribution of firms in terms of the wage they pay their employees. The more conventional approach would be to classify firms in terms of their average wage level. We assume that firms are categorized in this manner; for simplicity, two categories are used and are denoted as "low" and "high" paying firms.

69. Department of Labor (1983).

70. See, for example, Department of Labor (1959).

71. Brown, Gilroy, and Kohen (1982).

72. At the time of their study, retail establishments were covered by the Federal minimum wage if they had annual sales of $250,000 or more and were part of an enterprise with annual sales of $500,000 or more.

73. Cotterill and Wadycki consider state as well as Federal minimum wage affects. For simplicity, and with no loss in generality, the discussion in the text is restricted to the latter.

74. See also Welch (1974) where this approach to measuring the impact of the minimum wage is developed in greater detail.

75. Using a different method to estimate a minimum wage measure, Hashimoto (1982) reaches the same conclusion.

76. Bluestone, Harrison, and Gorham (1984).

77. A recent study that attempts to identify such workers is described in Bureau of Labor Statistics (1984).

78. For other approaches to defining high tech industries, see Department of Commerce (1983).

79. Defined as engineers, life and physical scientists, mathematical scientists, engineering and science technicians, and computer specialists.

80. Of the approximately 12.3 million wage and salary workers in high tech industries under the second definition, 7.5 million were employed in manufacturing.

81. Under a third definition, which is a composite of the two noted in the text, 6.2 percent of all wage and salary workers were employed in high tech industries in 1982. See also Lawson (1982) where a composite measure based on R&D expenditures and R&D scientists engineers is used to identify high tech industries.

82. To simplify the problem, the universe of industries might be restricted at first to those in manufacturing.

83. For a discussion of clustering techniques, see Hartigan (1975).

84. For a discussion of the growth of small firms and the relationship between such growth and that of total employment, see Birch (1979).

85. Whether this is the case is not clear, since the demand for capital depends on the number of firms as well as the capital requirement of the typical firm in each sector. Moreover, the number of service sector firms is growing relative to the number of firms in the goods sector.

86. In the absence of being able to identify high tech and nonhigh tech firms, high tech and nonhigh tech industries can be distinguished. To simplify the analysis even further, only four sectors might be used, i.e., small and large firms within the goods and service sectors.

87. To do this, it is necessary to use a measure of income inequality which is decomposable; hence, a measure other than the Gini coefficient is required, e.g., Thiel's entropy measure (Theil, 1967).

88. Even within a SIPP panel, some individuals will enter or re-enter the labor force while others leave or retire from the labora force. To obtain a completely closed population, only individuals who were employed in both the beginning and ending reference periods should be considered.

89. Between 1960 and 1980 the percentage of workers unionized in manufacturing declined from 51.2 to 32.3 percent. The decline was even more dramatic in construction where the percentage fell from 77.6 to 31.6 percent (Newmann and Rismann, 1984).

90. Freeman (1981) states that in private discussions with BLS personnel he was informed that "it is reasonable to assume that all workers...are covered when 50+ percent of them are covered and that none are covered when fewer than 50 percent are covered" (p. 495). In an earlier paper, Freeman and Medoff (1979) also refer to Douty (1960) as having concluded that "[in] the aggregate, situations involving minority coverage are believed to affect comparatively few workers" (p. 345).

91. Freeman and Medoff (1981), p, 568.

92. Freeman (1981) notes that the "abandonment of the [EEEC surveys]... represents a serious loss of information on compensation" (p. 495). As indicated in the text, this loss can be offset by a SIPP-ED file.

93. In classifying a firm as union or nonunion on the basis of a single observation, there is an obvious risk of misclassification. On the other hand, much information can be lost on how labor markets function when the union status of a firm is unknown.

94. Also derived independently from CPS data.

95. Freeman and Medoff (1984) note that unionized establishments use more capital than predicted by the production model; correcting for the productivity of capital under alternative assumptions, they estimate that the "[Brown and Medoff] results are consistent with a 10 to 15 percent [union] productivity effect" (p. 167).

96. Additionally, because hours worked data are unavailable in the CCI, it is assumed that union and nonunion construction workers work the same number of hours per week.

97. Much of what is known about union employment effects is found in Lewis (1963).

98. Freeman and Medoff (1982).

99. Ibid.

100. The use of cost functions, which relate output to total cost (which in turn reflects the marginal productivity relationships implicit in the production function), has become increasingly widespread in productivity analyses because they impose fewer constraints than production functions. For example, the CES production function constrains the elasticity of substitution between different pairs of factors to the same constant. This

constraint is absent when the translog cost function is utilized. For ease of exposition, the discussion in this section focuses on the production function rather than the cost function.

101. Specifically, net output is defined by $Q= Q[V(K, L), M, F]$ where K, L, M, and F, respectively, represent capital, labor, materials, and fuels. M and F are parameters affecting the level of output, but the marginal rate of substitution between K and L is independent of M and F (Arrow, 1974).

102. When output is measured in gross terms, the production function is defined by $Q= Q(K, L, M, F)$. Comparison of this function with that in the previous footnote indicates that the value added measure of output assumes the production function is weakly separable.

103. Douglas (1934).

104. See, for example, Klotz, Madoo, and Hansen (1980).

105. To take account of labor quality, L in the equation in footnote 102 is replaced by $L^*= Lq$ where $q = \sum_i \sum_j (aE_{ij} + bH_{ij})$ is an index of labor quality. E_{ij} and H_{ij} measure units of experience and education, respectively, of the ith individual in the jth experience-education class, and a and b are the coefficients of experience and education variables in a wage rate regression equation. L^* is thus an adjusted value of L, reflecting the composition of labor with respect to experience and education among different industries at a specified point of time or at various points of time for a given industry.

106. Instead of $Q(K, Lq, M, F)$, the production function becomes $Q(K, L, q, M, F)$ where q is a vector of variables related to investment in human capital, e.g., age, sex, race, occupation, education, experience, etc.

107. It should be noted that SIPP, the CPS, and the Census of Population do not distinguish between hours paid (which is what is reported in these data sets) and hours worked. This distinction is important for productivity studies, particularly since there are systematic differences in the ratio of hours worked to hours paid by firm size (Kunze, 1984). Information on paid vacation and paid sick leave, the primary components of hours paid but not worked, could be obtained from SIPP respondents and enhance its usefulness for productivity analysis.

108. Denison (1974).

109. The duality relationship between a firm's production function and cost function insures that a change in one direction in the total unit cost of production associated with worker behavior results in a change in the opposite direction in total factor productivity. Hence, Norsworthy and Zabala's results imply that inclusion of a worker behavior variable should lead to an improvement in the specification of a firm's production function as well as its cost function.

110. A similar conclusion is reached by Norsworthy and Zabala (1983) for the automobile industry.

111. Mansfield (1980).

112. In the 1979 Income Survey Development Program about 50 percent of the respondents used their W-2 forms to provide income information.

113. In the MEL, large establishments are defined as having 50 to 500 employees, depending on the size of the SMSA in which they are located; the larger the SMSA, the larger is the minimum establishment size required for inclusion in the MEL. It should be noted that since the MEL was compiled, some of the establishments listed in it may have moved or failed; nonetheless, it should prove to be a useful tool in linking individuals to their place of work.

114. This approach is used because in some 4-digit SIC classes the number of small establishments and their response rate is low.

115. A similar argument is made by Klotz, Madoo, and Hansen (1980) for defining establishment size in terms of assets rather than value of shipments. On the other hand, because some small high tech establishments can account for a large share of an industry's output, it may be desirable in some cases to define establishment size in terms of value of shipments.

REFERENCES

Allen, Steven, "Unionized Construction Workers Are More Productive," Quarterly Journal of Economics, May 1984, pp. 251-274.

Antos, Joseph, "Union Effects on White-Collar Compensation," Industrial and Labor Relations Review, April 1983, pp. 461-479.

Arrow, Kenneth, "The Measurement of Real Value Added," in David Paul and Melvin Reder, eds., Nations and Households in Economic Growth, Academic Press, 1974, pp. 3-19.

Bailey, William and Schwenk, Albert, "Wage Differences Among Manufacturing Establishments," Monthly Labor Review, May 1971, pp. 16-19.

Bailey, William and Schwenk, Albert, "Wage Rate Variation by Size of Establishment," Industrial Relations, Spring 1980, pp. 192-198.

Barth, James; Cordes, Joseph; and Haber, Sheldon, Worker Characteristics and Size of Business: Are there Systematic Empirical Relationships? U.S. Small Business Administration, Office of Advocacy, 1984.

Birch, David, The Job Generation Process, M.I.T. Program in Neighborhood and Regional Change, Cambridge, Mass., 1979.

Bluestone, Barry; Harrison, Bennett; and Gorham, Lucy, Storm Clouds on the Horizon: Labor Market Crisis and Industrial Policy, Economic Education Project, Brookline, Mass., May 1984.

Borjas, George, "Job Mobility and Earnings over the Life Cycle," Industrial and Labor Relations Review, April 1981, pp. 365-376.

Brannen, Patrick, Explanation of Sample Codes and Terminology, Appendix D, Industry Codes, Bureau of the Census, September 16, 1982.

Brogan, Robert and Erickson, Edward, "Capital-Skill Complementarity and Labor Earnings," Southern Economic Journal, July 1975, pp. 83-88.

Brown, Charles; Gilroy, Curtis; and Kohen, Andrew, "The Effect of Minimum Wage Legislation on Employment and Unemployment," Journal of Economic Literature, June 1982, pp. 487-528.

Brown, Charles and Medoff, James, "Trade Unions in the Production Process," Journal of Political Economy, June 1978, pp.355-379.

Bureau of the Census, The Standard Statistical Establishment List Program, Technical Paper 44, January 1979.

Bureau of the Census, Economic and Agricultural Censuses Manual, Attachment B, 82EACES-E, 1-A-7, Supplement No. 1, 1982.

Bureau of the Census, Characteristics of the Population Below the Poverty Level: 1982, Current Population Reports, Series P-60, No. 144, February 1984.

Bureau of Labor Statistics, Employee Compensation in the Private Nonfarm Economy, 1977, Summary 80-5, April 1980.

Bureau of Labor Statistics, Trends in Multifactor Productivity, 1948-81, Bulletin 2178, September 1983.

Bureau of Labor Statistics, BLS Reports on Displaced Workers, News Release, November 30, 1984.

Cotterill, Philip and Wadycki, Walter, "Teenagers and the Minimum Wage in Retail Trade " Journal of Human Resources, Winter 1976, pp. 69-85.

Denison, Edward, Accounting for United States Economic Growth, 1929-1969, Brookings Institution, Washington, D.C., 1974.

Department of Commerce, An Assessment of U.S. Competitiveness in High Technology Industries, International Trade Administration, February 1983.

Department of Labor, Studies of the $1 Minimum Wage, Wage and Hour and Public Contracts Division, 1959.

Department of Labor, Minimum Wage and Maximum Hours Standards under the Fair Labor Standards Act - 1983 Report, Employment Standards Administration, 1983.

Doeringer, Peter and Piore, Michael, Internal Labor Markets and Manpower Analysis, D.C. Heath and Co. 1971.

Douglas, Paul, The Theory of Wages, The MacMillan Company, 1934.

Douty, H. M., "Collective Bargaining Coverage in Factory Employment, 1958," Monthly Labor Review, April 1960, pp. 345-349.

Freeman, Richard, "The Effect of Unionism on Fringe Benefits," Industrial and Labor Relations Review, July 1981, pp. 489-509.

Freeman, Richard and Medoff, James, "New Estimates of Private Sector Unionism in the United States," Industrial and Labor Relations Review, January 1979, pp. 143-174.

Freeman, Richard and Medoff, James, "The Impact of the Percentage Organized on Union and Nonunion Wages," Review of Economics and Statistics, November 1981, pp. 561-572.

Freeman, Richard and Medoff, James, "Substitution between Production Labor and Other Inputs in Unionized and Nonunionized Manufacturing," Review of Economics and Statistics, May 1982, pp. 220-233.

Freeman, Richard and Medoff, James, What Do Union Do?, Basic Books, 1984.

Fuchs, Victor, "Redefining Poverty and Redistributing Income," The Public Interest, Summer 1967, pp. 89-95.

Gollop, Frank and Jorgenson, Dale, "U.S. Productivity Growth by Industry, 1947-73", in John Kendrick and Beatrice Vaccara, eds., New Developments in Productivity Measurement and Analysis, National Bureau of Economic Research, University of Chicago Press, 1980 pp. 17-136.

Green, Gordon, Wage Differentials for Job Entrants, by Race and Sex, The George Washington University, Ph.D. Dissertation, 1984.

Hartigan, John, Clustering Algorithms, John Wiley and Sons, 1975.

Hashimoto, Masanori, "Minimum Wage Effects on Training on the Job," American Economic Review, December 1982, pp. 1070-1087.

Henle, Peter and Ryscavage, Paul, "The Distribution of Earnings Among Men and Women," Monthly Labor Review, April 1980, pp. 3-10.

Jusenius, Carol, "The Influence of Work Experience, Skill Requirements, and Occupational Segregation on Women's Earnings," Journal of Economics and Business, Winter 1977, pp. 107-114.

Klotz, Benjamin; Madoo, Rey; and Hansen, Reed "A Study of High and Low Labor Productivity Establishments in U.S. Manufacturing," in John Kendrick and Beatrice Vaccara, eds., New Developments in Productivity Management and Analysis, National Bureau of Economic Research, University of Chicago Press, 1980, pp. 239-286.

Kunze, Kent, "New BLS Survey Measures Ratio of Hours Worked to Hours Paid," Monthly Labor Review, June 1984, pp. 3-7.

Kwoka, John, "Monopoly, Plant, and Union Efffects on Worker Wages," Industrial and Labor Relations Review, January 1983, pp. 251-257.

Lawson, Ann, "Technological Growth and High Technology In U.S. Industries," Industrial Economics Review, Spring 1982, pp. 11-18.

Leighton, Linda and Mincer, Jacob, "The Effects of Minimum Wages on Human Capital Formation," in Simon Rottenberg, ed., The Economics of Legal Minimum Wages, American Enterprise Institute for Public Policy Research, Washington, D.C., 1981, pp. 155-173.

Lewis, H. Gregg, Unionism and Relative Wages in the United States: An Empirical Inquiry, University of Chicago Press, 1963.

Mansfield, Edwin, "Basic Research and Productivity Increase in Manufacturing," American Economic Review, December 1980, pp. 863-873.

Masters, Stanley, "Wages and Plant Size: An Interindustry Analysis," Review of Economics and Statistics, August 1969, pp. 341-345.

Mellow, Wesley, "Unionism and Wages: a Longitudinal Analysis," Review of Economics and Statistics, February 1981, pp. 43-52.

Mellow, Wesley, "Determinants of Health Insurance and Pension Coverage," Monthly Labor Review, May 1982, pp. 30-31.

Mellow, Wesley, "Employer Size and Wages," Review of Economics and Statistics, August 1982, pp. 495-501.

Mellow, Wesley and Sider, Hal, "Accuracy of Response in Labor Market Surveys: Evidence and Implications," Journal of Labor Economics, October 1983, pp. 331-344.

Mincer, Jacob, Union Effects: Wages, Turnover and Job Training, Working Paper No. 808, National Bureau of Economic Research, Cambridge, Mass., November 1981.

Mincer, Jacob and Jovanovic, Boyan, "Labor Mobility and Wages," in Sherwin Rosen, ed., Studies in Labor Markets, National Bureau of Economic Research, University of Chicago Press, 1981, pp. 21-63.

Minimum Wage Study Commission, "A Demographic Profile of Minimum Wage Workers," Report of the Minimum Wage Study Commission, May 1981.

Monahan, James, Procedures for Using the Longitudinal Establishment Data File, Bureau of the Census, April 1983.

Nelson, Dawn, McMillen, David, and Kasprzyk, Daniel, An Overview of the Survey of Income and Program Participation, Bureau of the Census, December 1983.

Newmann, George and Rissman, Ellen, "Where Have All the Union Members Gone?" Journal of Labor Economics, April 1984, pp. 175-192.

Norsworthy, John and Malmquist, David, "Input Measurement and Productivity Growth in Japanese and U.S. Manufacturing," American Economic Review, December 1983, pp. 947-967.

Norsworthy, John and Zabala, Craig, Output Measurement and Productivity Growth in the U S. Auto Industry, Bureau of the Census, April 1983.

Norsworthy, John and Zabala, Craig, "Worker Attitude, Worker Behavior, and Productivity " Industrial and Labor Relations Review, (forthcoming, 1985).

Office of Management and Budget, Enterprise Standard Industrial Classification Manual 1974.

Oi, Walter, "Heterogeneous Firms and the Organization of Production," Economic Inquiry, April 1983, pp. 147-171.

Pencavel, John, An Analysis of the Quit Rate in American Manufacturing Industry, Industrial Relations Section, Princeton University, 1970.

Pencavel, John and Hartsog, Catherine, "A Reconsideration of the Effects of Unionism on Relative Wages and Employment in the United States, 1920-1980," Journal of Labor Economics, April 1984, pp. 193-232.

Phelps, Edmund, "The Statistical Theory of Racism and Sexism," American Economic Review, September 1972, pp. 659-661.

Riche, Richard; Hecker, Daniel; and Burgan, John, "High Technology Today and Tomorrow: A Small Slice of the Employment Pie," Monthly Labor Review, November 1983, pp. 50-58.

Ryscavage, Paul, SIPP and CPS Labor Force Concepts, Bureau of the Census, August 1984.

Salop, Joanne and Salop, Steven, "Self-Selection and Turnover in the Labor Market," The Quarterly Journal of Economics, November 1976, pp. 619-628.

Schiller, Bradley, Human Capital Transfers from Small to Large Businesses, U.S. Small Business Administration, Office of Advocacy, July 1982.

Spence, Michael, Market Signalling: Informational Transfer in Hiring and Related Screening Processes, Harvard Univeristy Press, 1974.

Theil, Henri, Economics and Information Theory, North-Holland Publishing Company, 1967.

Thurow, Lester, "The Disappearance of the Middle Class," The New York Times, February 5, 1984, p. 21.

Weiss, Leonard, "Concentration and Labor Earnings," American Economic Review, March 1966, pp. 96-117.

Welch, Finis, Minimum Wages: Issues and Evidence, American Enterprise Institute for Public Policy Research, Washington, D.C., 1978.