AMERICAN EMPLOYER SALARY SURVEYS AND LABOR ECONOMICS RESEARCH: ISSUES AND CONTRIBUTIONS

by
Erica L. Groshen

Federal Reserve Bank of New York
Research Paper No. 9604

March 1996

3/14/96
Comments Welcome

American Employer Salary Surveys and Labor Economics Research: Issues and Contributions

Erica L. Groshen
International Macroeconomics Function
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045
(212) 720-7685

Abstract

I review the uses of American employer salary surveys for labor market research. Recent computational, theoretical, and econometric advances render these surveys ripe for exploitation. I summarize theories of employer wage effects and then describe salary surveys and their preparation for analysis. Then, the surveys and the methodological issues they raise are contrasted with household data. Finally, I summarize the techniques used and contributions made in some salary survey-based studies.

The author is Research Officer and Head of the International Macroeconomics Function at the Federal Reserve Bank of New York. She thanks John Abowd for suggesting she write this paper and Richard Freeman and John Dunlop for encouraging her to use salary surveys for research. In addition, she very gratefully acknowledges the support of the Research and Personnel Departments of the Federal Reserve Bank of Cleveland during the assembly and analysis of the Federal Reserve Bank of Cleveland Community Salary Survey data set. The views expressed in the paper are those of the author and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the author.

Prepared for: Conference ADRES 1994 The Microeconometrics of Human-Resource Management: Multinational Studies of Firm Practices
December 7-9, 1994
Ministere de l'Economie, Paris, France

I. Introduction

The accessibility of large-scale household surveys, combined with the solid theoretical foundation of the human capital model, leads most US empirical studies in labor economics to focus on the impact of demographic factors (education, age, race, sex, etc.) on wages and other outcomes. From the 1960s until lately, employers' activities have been relatively neglected. This paper argues that employer salary surveys provide a valuable, underused source of information about the firm-side workings of the labor market. This resource is now ripe for exploitation, due to advances in theory, computing power, and econometric technique.

While studying and administering wage and price controls, industrial relations systems, and workplace automation during the 1940s, 50s and early 60s, a prolific generation of US economists (see Reynolds, 1951; Dunlop, 1957; Segal, 1981; Kerr, 1983) uncovered strong evidence of large, persistent wage and personnel practice variations among employers. They developed economic models to explain their findings, but lack of data and computing power limited their ability to test hypotheses econometrically. In addition, since their exploratory studies predate current methods of presenting theory, many of their hypotheses were not formalized. To the extent that the current cohort of labor economists is even aware of this work, it has been difficult to follow up on that generation's findings and conjectures.

During the 1960s and 1970s, a new generation of labor economists focused on supply-side analysis of the U.S. labor market and formalized human capital theory, aided by computational developments and the availability of household microdata (particularly the decennial Census and the monthly Current Population Survey). Hypotheses about poverty, demographic difference in earnings, and the role of education were tested with new theoretical and statistical rigor. Further refinements were possible with the addition of longitudinal data series (such as the Michigan Panel Study of Income Dynamics and National Longitudinal Surveys) and the results of several large scale social experiments (such as the Seattle and Denver Income Maintenance Experiments and Supported Work). However, the data and theories most available focused on labor supply issues.

In the mid-1980s, economists returned to investigating several strong empirical regularities that appeared linked to the demand side (in particular, industry, union, and employer-size wage effects), as predicted by the 1950s analysis.[1] The failure of even the

[1] See Groshen (1991a) for a general review of industry and establishment wage differential studies. For employer size, see Brown and Medoff (1989).

most sophisticated attempts to control for human capital to convincingly explain these employer-based wage phenomena fueled a new surge of interest in the economics of human resource management. In addition, many of the current social issues are demand-related: such as the impact of trade, immigration, technological changes and the decline of unionism. Without direct data on employers, research on the demand side of labor economics and on employer personnel policies is limited to what can be gleaned from household surveys (which report only industry, and occasionally, size and union status) and industry aggregates. Conclusions about firms' activities from such data must either be tentative, or require heroic assumptions.

Hence the interest in alternative, employer-based data. Yet such efforts face at least three challenges. First, finding appropriate data can be a challenge. Employer data can be sparse, hard to locate or access, and idiosyncratic. Unlike household data (which were developed for economic analysis), employer data is usually collected for other purposes, so it is smaller in scope and may be confidential. Second, ventures beyond the household surveys into analysis of employer data have revealed limitations in labor's standard econometric tools, which have become specialized to the analysis (and inherent problems) of household data and human capital-based issues. Thus, analysis of employer data usually requires some econometric ingenuity or innovation; each new data set or study presents the researcher with idiosyncratic problems. Third, interemployer differences require extensions of basic neoclassical theory, since in a perfectly competitive equilibrium, apparent differences among employers are either temporary or due to mismeasurement of human capital or compensation. So the growth in empirical interest in human resource activities has often required or been fueled by the development of well-grounded propositions to test and make sense of our results. Without theoretical developments in human resource economics to guide or challenge us, empirical results could wind up incomprehensible or indefensible in a simple purely competitive long-run framework. Done poorly, such exercises risk sinking into mindless data mining.

U.S. economists have begun to exploit government employer-based earnings data from the Unemployment Insurance program and from the Annual Survey and Census of Manufactures. While these sources provide a fascinating window into employer activities, they lack information on employees' skills. For that, the most readily available, largely unexploited data source is the employer wage survey, which necessarily includes detailed occupation. Challenges remain in gaining access to the surveys, in dealing with salary surveys' limitations, in learning which questions to ask, and in understanding what we find. Fortunately, the more is done, the easier it will be to do more.

The next section briefly reviews developments in neoclassical theory that point to a central role for employers in labor market outcomes. The third section describes the type of information obtained from American employer salary surveys. The fourth section covers the steps involved in preparing a particular salary survey for analysis. The fifth section contrasts salary survey data with household data, and discusses some of the methodological issues raised by the organization of the data. Section VI reviews contributions and techniques from a sample of salary survey studies on the extent, pervasiveness and permanence of variations in practices and outcomes among employers. The seventh section concludes.

II. Theories That Inspire and Direct Research into Employer Activities

While human resource management covers far more than wage-setting, labor economists focus most of their theories and empirical effort on wages. This is because wages provide an easily quantifiable measure of differences in employer practices and strategies, and understate total compensation differences (Atrostic 1983). Hence, studies of wages have provided many of the advances in the economic theory of personnel management. No doubt, awareness of employer wage differences will extend research into the many other ways that employers differ. In particular, since each theory of wage differences among employers has implications for variation in other aspects of human resources practice, wage policy acts as a springboard into the general field of interemployer differences.

So, why do some employers pay higher wages than others? This section reviews theories of employer wage differentials, beginning with a summary of the puzzle posed by interemployer wage effects, and ending with a brief discussion of the policy implications of the differences among the theories.

In neoclassical economics, wage rates and employment, like the price and quantity of any traded commodity, are determined by both supply and demand. Nevertheless, despite the simultaneous nature of the wage-setting process, until the late 1980s most empirical investigations of the determinants of wages focused primarily on factors affecting labor supply. Demand factors were relatively neglected.

In the simplest version of the labor market, (featuring perfect competition in the capital, labor, and product markets), equivalent workers at equivalent jobs will be compensated equally. Since employers in a competitive labor market face a single market wage (at which they can hire all the workers they would like) differences among employers will affect only how many workers they hire, not the wages paid. Employers

who stray from the market rate will be forced out of business by loss of employees (compensation set too low) or by the loss of capital (compensation set too high).

In the Current Population Survey (CPS), a U.S. household survey, regressions of wages on workers' narrowly-defined human capital characteristics (education and age) typically explain at most a quarter of the variation in the log of wages. Addition of occupation raises explanatory power by 15 percent or so, while race, sex, and union variables add another 4-7 percent. Adding industry indicators (broadly defined) can raise explanatory power to about 50 percent of the variation of wages.

What accounts for the half of wage variation that the equation doesn't explain? Which empirical regularities or theories are associated with the residual variation? And, are we certain that industry, unionism, race, and sex reflect differing ability on the part of workers? Despite the compelling logic of the simple human capital model, the large residual observed in cross-sectional regressions indicates that additional forces may be at work, either in individual attributes or in labor market outcomes. The empirical work summarized in Groshen (1991a) suggests that some employers pay higher wages to observationally equivalent workers, which raises two crucial questions: 1) Why do the large employers choose to pay high wages? And, 2) why don't low-wage employers lose their employees, or why aren't high-wage employers replaced by smaller low-wage competitors? In order to answer these questions, the simple model must be extended.

Most theories of the employer wage differential answer the two questions by arguing that workers in high-wage firms are more productive. However, since productivity differentials that show up in wages are usually due to differences among individuals, not to employers' characteristics, the theories need to explain the link between these characteristics and workers' productivity. In the absence of productivity differentials, imperfect competition in the product market must be present and influence wages similarly for all employees of a firm.

The proposed reasons can be grouped into five categories: 1) employers systematically sort workers by ability; 2) wages vary because of compensating differentials; 3) costly information generates or perpetuates random variations in wages; 4) the efficient wage for some employers is above the market rate; and, 5) workers inside firms exercise a claim on rents. In none of these cases does the existence of employer differentials contradict profit maximization on the part of employers. However, the last two theories predict that workers will queue for high-wage employers, while in the other three models, employer differentials clear the market.

The first two explanations relax the assumption of inter-firm uniformity in employees or working conditions. Then, hourly wages may mismeasure either the

workers' units of work (because workers in high-wage employers produce more per hour) or their compensation (because wage omits nonpecuniary returns to employment). Recent extensions of the theory of human capital and ability sorting give a role for sorting by employer (examples include Salop's 1979 model of adverse selection, Akerlof's 1981 model of jobs as dam sites, or Kremer and Maskin's 1994 model of quality segregation). The theory of implicit contracts also suggests that some compensating differentials may be employer-wide. Abowd, et al. (1994) suggest that a major portion of firm wage effects may be the result of firms' sorting of workers by qualities unobserved by the economist. In the case of compensating differentials, many aspects of working conditions are common among a firm's employees, including possible implicit contracts.

By contrast, errors, the efficiency wage and rent-sharing models assume imperfections or lack of competition in product or labor markets because they imply the existence of job rationing or queues for high-wage employers. The latter two classes of models (efficiency wages and bargaining) are largely employer-based and the recent subjects of substantial theoretical developments.

This attention has been fueled by accumulated empirical stylized facts which suggest that employer wage differences may be important components of observed earnings inequalities among demographic groups. Each possible explanation of employer differentials suggests that a different type of policy for reducing these inequalities would most effective and efficiency-enhancing.

If employer differentials are the result of errors, the efficiency of the labor market may be enhanced by their elimination, perhaps through government subsidies for information gathering and dissemination. On the other hand, efficiency wage differentials or implicit profit-sharing wages may be appropriate second-best solutions to monitoring or agency problems endemic to the labor market. However, their existence implies that jobs are rationed to workers on a queue, introducing implications for other policies (such as trade or antidiscrimination policy--see Bulow and Summers 1986--or macroeconomic policy--Weitzman 1986). Finally, if employer wage differences reflect compensating differences or returns to unmeasured human capital, they are first-best and efficient. Apparent inequities in the market arise from inequality of access to human capital development, rather than from workplace discrimination.

III. Characteristics of American Employer Salary Surveys

A. Who Conducts Salary Surveys in the U.S.

Salary surveys constitute most large U.S. employers' primary source of information about their employees' opportunity wages (Freedman, 1976). A wide variety of organizations sponsor salary surveys, including the following examples: the federal government, most of the regional Federal Reserve Banks (private employers), Hay Associates, Inc. (a consulting firm), the American Hospital Association (an employers' association), the National Association of Business Economists (a professional society), and the American Association of University Professors (a union).

The sponsor usually considers the survey to be a service it offers to participants, and often uses the data for its own purposes as well (e.g., for internal wage-setting, monitoring, promoting local economic development or lobbying). Employers participate in return for entitlement to receive or purchase the results. Purchase fees, when charged, can be large--either as direct fees, or as part of membership in the surveying organization. If the latter, access to the salary survey is often a primary benefit or membership, along with access to group health insurance (small employers), lobbying, and conferences.

B. The Contents of Salary Survey Data

Each year, salary surveys record the wage of every person holding positions in certain well-defined occupations with participating employers, along with some information on the employer itself, such as industry, number of employees, date of last and next company-wide pay adjustment, or fringe benefit policies. Generally, no information is collected about workers in nonsurveyed occupations. Normally, instructions specify that all cash payments be included in the wage reported, exclusive of overtime and shift premia. Thus, bonuses, incentives, and piece-rate earnings are included, but most nonwage benefits are not. Increasingly, this information is requested and reported in machine-readable form, to speed processing and improve reliability of the data. The data are often available for study within six weeks of collection.

Salary surveys are organized by the industry, geographic area, or professional group, with coverage limited accordingly. If a survey is geographically based, then the occupations covered will be those commonly found and most comparable across industry: usually clerical, administrative, maintenance, and managerial positions. The blue-collar occupations range from janitors, drivers, and laborers, to painters, plumbers and electricians. Clerical jobs include clerks of various types, secretaries, receptionists. Computer jobs, personnel staff, payroll clerks, bookkeepers, even attorneys, accountants,

economists, and industrial nurses will often be included. These surveys have the advantages over industry and professional wage surveys that they allow control for regional wage differences, they include many different industries, and they are longitudinal in establishments. While they do not cover all occupations, they do cover a broad mix: white- and blue-collar, professional, skilled, and unskilled.

Because the surveyed occupations are found in many industries and firms, their labor markets may be more competitive than the markets for more industry-specific or firm-specific occupations. Workers should be more mobile when their skills are readily transferable among many different employers. Thus, we would expect the wages of workers in these occupations to be more standard across employers than would the wages of workers in less common occupations. However, because they are common to most firms, these occupations often work outside the major productive activities of their establishments and represent a minority of employment in their establishments.

Industry-based surveys cover production occupations for the industry in their purview. They may be limited geographically, or national in scope. However, these surveys are the most likely to cover a high proportion of workers in the surveyed establishments--usually about 60% in U.S. Bureau of Labor Statistics Industry Wage Survey samples. The jobs included tend to be particularly narrowly-defined and particular to the industry. In manufacturing, these are primarily blue-collar occupations and their supervisors: various machine operators, inspectors, packers, skilled trades. The amount and type of training and the actual machines tended and tools used are often specified. In service industries, these jobs are more likely to be white-collar jobs (such as tellers or data entry clerks) or service occupations (such as food service workers and sales personnel).

By contrast, professional salary surveys typically cover only a few closely-related variants of the same occupation, and are usually national and interindustrial in scope. These tend to be conducted for professions which have a national job market. Thus, the employees are highly skilled, and easily identified, and often hold only a small share of their employer's jobs. Examples include physicians, corporate librarians, attorneys, business economists, and college professors. Some of these surveys also collect demographic data on the employees (such as degree, years of experience or tenure, etc.) because such information is more uniformly applicable within a narrowly-defined occupational group. One interesting variant are the corporate executive surveys, which have garnered a lot of recent attention in the academic, popular, and business press.

The typical salary survey specifies a detailed job description for each occupation included. The specificity of the occupation definitions, which are actually job classifications, exceeds the detail in four-digit Dictionary of Occupational Titles or Census

codes. Job descriptions are typically two to three paragraphs long, and specify the responsibilities, training requirements, how the job is done (including the precise machinery or tools worked with, if applicable), what is produced, position in the corporate hierarchy, the occupation of direct supervisor, and number of supervisees. This is apparently the level of detail required by employers for the results to be useful.

Where applicable, occupations are divided by grade and/or listed as belonging to "job families" -- members of a well-defined career path -- through which new entrants are expected to move as they gain experience and seniority. For example, secretaries are divided into at least four occupation classes, by their responsibilities, and distinguished from other clerical occupations such as stenographers, typists, and file clerks. For brevity in the discussion that follows, the term occupation will be used instead of job classification, the more accurate term. Generally, the survey sponsor provides the name and phone number of a person to consult if the responding employer has questions about whether a certain employee's job is consistent with a particular description.

Normal household-survey demographic data (age, sex, race, marital status, and education) are not included in the typical salary survey. The Equal Pay Act and Equal Employment Opportunity regulations dictate that these distinctions are an illegal basis for setting wages. Accordingly, employers no longer report wages separately by race and sex in private salary survey results, although it was once standard practice. For historical reasons, the BLS Industry and Area Wage Surveys continued to collect separate data by gender. But, in general, U.S. employers are reluctant to release demographic and wage information, for fear of exposing themselves to bad publicity or lawsuits.

Sponsors report aggregated results of the surveys in tabular "result books"--with the identity of each reporting institution disguised. The heart of survey results are the salary statistics by occupation. Some result books list the full set of incumbent salaries for each occupation for each employer. In such cases, the sponsor preserves the anonymity of individual responses by assigning a code to each institution and revealing to each respondent only its own code. In result books where the full distribution is not displayed, employers learn only summary statistics such as the number of responding institutions, the number of incumbents and the minimum, maximum, and mean salary.

Table 1 compares the features of salary surveys with employer data found in two other major sources: the Longitudinal Research Database -- which is constructed from the U.S. Census and Surveys of Manufactures -- and the Unemployment Insurance employer and worker files. Although these alternative sources provide broader samples, only salary surveys provide firm-level occupational information in a level of detail that can control for

human capital as productively used. In the other sources, human capital must be merged in from industry aggregates.

C. How Salary Surveys are Used: Wage-Setting in Large U.S. Firms

Salaries set by firms are bounded on the high end by workers' marginal products and from below by their employees' outside opportunities. However, these constraints determine a range more often than a unique wage, because both parties have limited current information on individuals' productivity and labor market prospects. In fact, Levine's (1993) survey of 139 compensation executives shows that wage change decisions rarely reflect unemployment rates, quit rates, and corporate returns on assets.

As a substitute for the unobserved supply and demand functions, firms develop compensation policies to attract and retain employees appropriate for the firms' needs. While these policies vary across firms, large employers' practices share a number of common features, in particular: a job evaluation system to rate jobs; salary grades or a wage line which specify earnings according the job evaluation system; and a merit- or seniority-based system for wage growth within salary grades.[2] Based on experiences assisting an employer's personnel department, discussions with other personnel executives, compensation textbook descriptions of the process, and the responses of compensation managers in Levine's survey, Groshen and Schweitzer (1996) advance the following institutional description of wage adjustments, which is summarized in figure 1.

Overall annual budgets for compensation, and therefore average pay increases, are determined by top management, typically the chief executive officer or some senior officer at corporate headquarters.[3] Proposals for the adjustment originate with, and are defended by, the local personnel department, primarily on the basis of salary survey comparisons. Once the budget is approved (two to six months in advance of the actual salary adjustments), the total "pie" for wage increases is split up among and within departments in accordance to perceived merit and labor market conditions for particular workers (again, often using salary surveys to justify arguments). Although the degree of decentralization varies among companies, the basic mechanism usually takes the form described above.

---

[2] General compensation policy references include: Frederick S. Hills, *Compensation Decision Making*, Milton L. Rock, *Handbook of Wage and Salary Administration*, and Robert E. Sibson, *Compensation*. Firm participation rates are over 90% in these activities, even with a sample including small businesses, according to Crandall (1979), *Journal of Management*.

[3] See Freedman (1976).

Crucial to this process is some explicit recognition of the employer's desired position in the wage distribution. Personnel managers will all characterize, justify, and describe their efforts to maintain their corporate wage-level policy.

For unionized employees, negotiations on more detailed terms and conditions of pay increases (or reductions) take place further in advance since contracts typically last about three years. Nevertheless a prospective compensation budget, similar to nonunion budgeting, is typically completed by the firm prior to negotiations to establish the acceptable range. Salary surveys are also sometimes used in union contract negotiations.

One of the interesting features of this method of wage-setting is that firm-wide wage effects are determined largely independently of occupational wage effects, particularly in a large, diverse companies. Firm effects and adjustments are centralized, but occupational and individual adjustments are decentralized--although possibly limited by the "pool" determined by upper level management.

IV. Issues in Preparing a Salary Survey for Analysis

A. Gaining Access

Since salary surveys are decentralized and include information that respondents consider sensitive, economists' access to them depends on the relationship they forge with the survey sponsors. Such relationships depend on the economist's finding out about the survey, respecting the confidentiality concerns of the sponsor and participants, demonstrating an ability and propensity to do research the sponsor considers relevant, and not imposing high costs on the sponsor.

Use of salary surveys for research necessitates direct contact with human resources practitioners. To find out about such surveys, one needs to ask compensation experts which surveys they use, and if they sponsor one themselves. Possible introductions to compensation administrators may be gained through Industrial Relations Research Association meetings, contacts with the American Compensation Association, one's own employer, employers for whom other consulting is being done, social contacts, or gainfully-employed former students. Then the survey sponsor must be persuaded to release the data for research.

The benefits to survey sponsors of allowing research analysis are the economist's informal or formal ongoing consultation on the results or conduct of the survey, plus whatever they learn from the actual research. Economists' statistical expertise and familiarity with non-salary survey data can prove very useful to human resource

practitioners, particularly if coming from someone already familiar with the survey itself. Academic-style studies of salary surveys can yield results readily translatable into the terminology of personnel officers, since the data are collected for their purposes. Such involvement can pay off for the researchers, too, increasing their awareness of current issues and of uses of the data.

Survey sponsors may, however, incur two kinds of costs from releasing their data for research. First, possible breaches of confidentiality might expose them to liability, or endanger future participation from their respondents. Second, the demands of locating the data and its documentation, delivering it or accommodating its in-house use, and responding to questions about it may unduly burden the sponsor's staff. The first kind of costs are highest for current data, the second for historical data.

In order to cooperate, sponsors must be convinced that the benefits will outweigh the costs, or that the costs will be trivial. Possible solutions to confidentiality concerns include pledging that no published paper will include the names of participants, using the data only in-house, and purging data of recognizable employer information (but retaining unique identification numbers). Unfortunately, the last solution eliminates the option of merging in firm data from other sources.

B. Steps in Processing

This section describes the steps in preparing a salary survey for analysis. To make the issues concrete, this section uses the example of the 38-year Federal Reserve Bank of Cleveland Community Salary Survey (CSS), which was converted for analysis during my tenure with the Bank. After receiving assurances that the data would remain in the Bank and that results would be reported only in aggregate form (with no identification of participants), the Personnel Department agreed to cooperate in return for consulting services and because sanctions for noncompliance are enforceable on employees.

The first step was physical acquisition of the data, in hard copy (result books) and machine-readable form (individual responses--for the later years). Collection required searches of personnel offices in three cities and of 30 archive boxes, for both result books and documentation. Most (but not all) books were found. The longer the time span, the more likely are gaps when no survey was done or the results cannot be located.

Then the data from the result books were keyed in, and the machine-readable files were converted to a uniform format over the years. The result books usually documented occupational descriptions and codes, and employer names and codes--allowing creation of uniform, consistent occupation and employer codes across years. The difficulty of this step increased with the age of the data, and with each change of survey administrator. For

the years in which documentation was unavailable, information from the adjacent years was used to assign occupation and employer codes. This procedure was straight-forward in the case of occupation, because we had the name of each surveyed occupation, and occupational descriptions did not change much from year to year. When employer names were missing, we used the pattern of occupations employed and wages paid from adjoining years to determine respondents' identities.

To facilitate merging other data, we also matched the surveyed occupations with Dictionary of Occupational Titles codes and employers with firm identifier numbers used for publicly-traded firms. Once the investment was made in processing the historical data and in setting up procedures for current data, updating and expanding the file each year became relatively easy. Ongoing involvement in analysis of the survey and in consulting for the Personnel Department led to improvements of the data from the research standpoint. At my suggestion, and because participants deemed them relevant, several questions (such as size of unit, information on corporate parent, etc.) were added to the employer portion of the survey. I have also spoken with the personnel administrators of many of the respondents, and have conducted at least one supplementary survey of participants to obtain answers to specialized questions.

V. Methodological Issues in the Use of Employer Salary Surveys

A. How Salary Surveys Differ From Household Survey Data

Salary surveys differ from household surveys in their coverage and in their use of institutionally relevant, rather than neoclassical theoretical concepts. While both differences entail advantages for research, they also pose methodological challenges.

Figure 2 compares the coverage of household survey data to that of employer surveys. Consider a small labor market, consisting of only jobs 1 through 10 and employers A through I. All possible jobs in that market can be classified within the grid, but companies vary in their size and use of occupations. Actual workers are designated by filled or empty circles. Ideally, labor economists would have access to full data on all workers, and be able to use any detail meaningful to their purpose. In fact, such samples are not available, so American labor economists mainly use household surveys to answer our questions about the labor market.

A household survey amounts to a random sample of the workers in the universe, represented in figure 2 by filled circles. From such a survey, one learns about the characteristics and activities of the work force, including those not employed or holding

multiple positions. Note however, that few workers are observed for any employer or detailed occupation.

Thus, even if the data included full detail on each worker, little could be learned about the particular occupations or employers, or about variation between versus within occupations or employers. Nor can their effects be fully controlled for or linked together. Of course, cost considerations (primarily in maintaining data quality), and confidentiality restrictions rule out collecting at that level of detail over the entire country anyway. Instead, surveys collapse occupations into large occupational classes and companies into industries. Guided by theory, policy interests, and awareness of institutions and stylized facts, the surveys also collect alternative measures of human capital differences (such as education) and employer differences (such as size).

By contrast, salary survey data provide a census of all workers within a class of occupations (preserving most of the relevant occupational detail) in a set of employers. The detail is possible only because the scope of the survey is limited by geography, industry, and/or occupation. A national sample with this level of detail would be prohibitively expensive. Such a sample allows study of the impact of employer and narrow occupational differences. Although the data may not cover the full range of employer or occupational differences, an understanding of the sample, and replication in other samples will indicate the relevance of the findings for the rest of the universe.

To illustrate the effect that these differences in sample design have on economic analysis, compare the basic household survey approach to wage regressions with an employer salary survey approach. In research using household surveys, economists estimate variants of the following regression for the wages of person i:

(1) $\text{Ln Wage}_i = (\text{Return to human capital}) * (\text{Amount of human capital used on job})_i + \text{Observed unexplained factors}_i + \text{Unknown factors}_i$

Starting from the theory of human capital, we look for appropriate measures of its use on the job--primarily by the job holders' years of education (or educational attainment), the years of experience (or usually, age), and broad occupational classification. For pragmatic reasons, economists often include other variables--either because of posited relationships with unobserved human capital or compensating differentials, or to control in an agnostic way for unexplained empirical regularities. Examples of variables included for these reasons are industry, race, sex, marital status, and union representation. Finally, the unknown factors term picks up the remaining wage variation, including unmeasured

human capital or compensating differentials, short idiosyncratic variations in productivity, and errors.

By contrast, the employer salary survey approach begins with the following relationship for person k, holding occupation i, working for employer j:

(2) Ln Wage_{ijk} = Job classification i (occupation) wage effect
+ Employer j wage effect
+ Internal labor market (job-cell) effect_{ij}
+ Individual k factors.

Here the constructs onto which economists project wages are easily observed labor market institutions, for which we have plausible theoretical interpretations. The occupation effect captures the current market valuation of the skills and working conditions represented by occupation i. The employer effects capture most demand-side influences on wages, such as those introduced in section 2. An internal labor market effect arises when a job-cell mean wage deviates from the sum of its occupation and employer effects. It represents a higher or lower establishment effect for a particular occupation, or the cell's deviation from the average tenure or merit. Similarly, the individual factor presumably arises from individual ability (as rewarded by tenure- or merit-based wage increases) or taste differences.

In order to link occupation to human capital or compensating differentials, attributes of occupations can be merged in so that occupation effects can be decomposed by job and jobholder attributes. Otherwise, detailed occupation can be used as a very careful control for human capital and compensating differentials. Likewise, facets of interemployer wage differences can be examined by decomposing them according to their relevant characteristics, or employer identity can be used as a very complete control for demand-side differences.

Unlike the residuals in household data, the individual effects in regressions such as equation (2) are very small (1-5% of total variation), providing a very interesting and complementary way to view labor economics phenomena previously only addressed with household data. Occupation (including sex and incentive) and employer often explain more than 95 percent of wage variation. From a policy standpoint, this means that other characteristics of the individual (for example, tenure, marital status, or race) must operate through job classification or through employer in order for them to have a large effect on wages. That is, since a large improvement in earnings can be attained only through a promotion or a change of employer, barriers to entry into highly remunerative occupations

or establishments can have a devastating impact on the earnings of otherwise-qualified workers.

B. Methodological Issues in the Use of Employer Salary Surveys

Although each new survey and use of a survey presents unique challenges, some common themes emerge. This section introduces the reasoning behind three of these concerns.

A major issue is the lack of standard human capital variables, or demographic variables. The solutions lie in three directions, depending on the study: merge in mean demographic or occupational training requirement data; use occupation indicators as controls for human capital; or construct other variables, such as seniority, from the job families included in the job descriptions.

The second frequent issue is the representativeness of the sample. The selection issues are basically two-fold: what are the sponsor's biases in its choice of participants, and on what basis do employers agree to participate. In general, for the latter, it is safe to assume that are firms who participate have wage-setting mechanism is fairly similar to that described above. And, since submitting the data and paying a possible fee are not costless, entities that participate must feel that the information they receive is useful, presumably because it reflects other possible employers for their staff.[4].

Respondents tend to be large firms: They have the best-defined jobs, and the personnel, resources, and record-keeping infrastructure necessary to complete such a report. Larger employers' negotiations with workers are less bilateral. They tend to have bureaucratic, systematic wage-setting procedures, so they rely more on salary survey results. Parent companies usually require reference to salary surveys in salary budget proposals. Survey sponsors are more likely to know about the large companies and ask them to participate, and they get many more matches per participant (lower average cost per observation) from a large company.

Small companies' wage setting is more idiosyncratic for a variety of reasons. For example, their pay structure may be more driven by matching outside offers and by direct, individual rewards for performance. Thus, jobs in small companies may evolve rapidly to reflect the skills, interests, and abilities of the incumbent, rather than being circumscribed by a two-paragraph description.

---

[4] Some exceptions happen, for example, several governmental agencies participate in the CSS. they explain their participation as being good citizens. The results may also be useful for lobbying purposes

Since private salary surveys are not scientifically-drawn random samples, comparisons to randomly sampled published data -- such as the Current Population Survey or the government's Industry or Area Wage Surveys-- for characteristics, trends, and outcomes is the best recourse. If peculiarities of the sample emerge, the careful researcher confines claims of generality to the sector analyzed, noting how important it is. Similarly, when the data are differenced or compared across years, sample drift (the non random entry and exit of firms and occupations from the sample) appears. The papers summarized have methods for dealing with this issue.

The importance of salary surveys in wage determination and the trading of confidential data also introduce concerns that the surveys may be the vehicle for the enforcement of monopsonistic labor market practices. If the survey is being used as an instrument to constrain the wages of workers--as feared by some observers--this could affect the pattern of data observed. Then comparison to the Area Wage Surveys, which is clearly not a monopsony vehicle (since it is too broad, large, and late at publication date to be of any use for this purpose) might prove useful. Asking respondents if they use it as their primary source of wage information might also be informative.

One final concern is the peculiarities in respondent error in salary surveys. Because salary survey data are collected from the administrative records of firms who hope to use the results in setting wages or classifying occupations, there is reason to believe that they contain fewer errors from loss of memory or intentional misinformation than the responses of individuals or their family members in household data. Nevertheless, the single source of the data for many people introduces concern about systematic error. To my knowledge, this issue has not yet been directly addressed in any studies of salary surveys to date.

VI. Examples of the Research Uses of Employer Salary Surveys

This section summarizes a selection of research done by this author using employer salary surveys. Two types of surveys are featured here, the federally-sponsored Industry and Area Wage Surveys (IWS and AWS) and the private Federal Reserve Bank of Cleveland Community Salary Survey (CSS). The AWS are ongoing geographically based program of the Department of Labor. The IWS are currently -- and possibly permanently -- suspended. The IWS and AWS programs replicate the design and content of private

salary surveys, although they draw from a larger, almost random sample.[5] The discussion of each paper focuses on the unique contribution made possible by salary survey data, and the methodological issues encountered and how they were handled.

A. Sources of Intra-Industry Wage Dispersion: How Much Do Employers Matter? (Groshen 1991b)

Findings: Recent studies document the existence and correlates of industry wage differentials and challenge the assumption that interindustry wage variation reflects only human capital differences, but remain inconclusive as to the source of the differentials (Dickens and Katz 1987). Indirectly, they also suggest that employers' wages may vary within industry, because sources of interindustry variation may also operate within industry, and because interindustry studies may miss wage variation associated with factors that do not vary much among industries. Other empirical studies suggest the existence of establishment wage differentials, but focus only on a single aspect (such as plant size (Brown and Medoff 1989) or proportion female (Blau 1977), or on particular occupations (Ward 1980) or cities (Rees and Schultz 1970), rather than taking a general approach to estimating employer wage differences.

This paper first presents the basic facts about establishment wage differentials, and then briefly evaluates the plausibility of competing explanations for their existence, clarifying the nature of the puzzle posed by the results. The paper demonstrates clearly that intra-industry establishment wage differentials measured in BLS Industry Wage Surveys are a large portion of wage variation that merit continuing research. Controlling for detailed occupation, wages vary almost as much among establishments within industry as they do among industries: corresponding to a standard deviation in wages of 14 percent of the mean (a major portion of the economy-wide standard deviation of about 50 percent) and accounting for 20-70 percent of intra-industry wage variation. Measured characteristics of establishments (such as size and product) can account for about one-half of wage variation among them. Internal wage structure variations generate under 10 percent of wage variance in these samples.

Investigations in the second part suggest that wage differences among employers are not random variations or returns to usual measures of human capital. Thus, further investigation is needed into sorting by unmeasured worker quality, compensating variations, efficiency wages, and rent-sharing.

---

[5] The samples are stratified by size, participation in the survey is voluntary, and the population from which they are drawn depends on the currency of information maintained by the local Bureau of Labor Statistics office, hence the samples may have some biases.

Job classification and establishment alone can explain over 90 percent of wage variation. This implies that unless one has an explicit incentive component to one's compensation, only a promotion or a change of employer can raise wages significantly. Indeed, attributes of individuals (for example, race, education, marital status) must operate through occupation, employer, or job-cell in order for them to affect wages. Thus, economic research into labor market outcomes needs to cover the impact of the major activities of employers, such as recruitment, promotion, supervision, forced separation, and general wage-level policies.

Methodological issues: Since this study sought to test for significant, but unmodeled, employer differences, the challenge was to adequately and appropriately characterize variance components without imposing structure. Of particular interest is employer wage differences' relative contribution to total wage variation, as a measure of the economic significance. The paper partitions the variance of wages into four portions associated with the particular effects shown in equation (2) using analysis of variance (ANOVA) techniques to compare the dispersion of the following wage components:

1) occupation--a measure of the importance of external labor markets;
2) establishment--a measure of the importance of demand-side differences;
3) job-cell--a measure of the importance of independent wage structures; and
4) within job-cell--a measure of the importance of individual differences.

One option might have been to rate the importance of establishment wage differences through statistical significance -- as measured by F-tests for employer indicators in wage equations. The alternative of concentrating attention on proportions of variance is preferred for two reasons. First, the large sample sizes found in typical salary surveys can yield strongly significant F-statistics (the critical value is 1 in most cases) -- even when economic significance is slight. Second, establishment identity is presumably an inefficient measure of the economically relevant differences between establishments. By construction, it captures all differences and thus identifies the maximum amount of variation that understanding of employer wage policy could explain. However, since establishment may be an unnecessarily fine measure of employer differences, the F-statistic can mislead because it averages out the impact of all estimated levels. While the additional variation explained by unnecessary levels is negligible, inclusion of irrelevant levels could potentially wash out the significance of the relevant ones. To demonstrate this point, note that the F-statistic of a factor X is defined as follows:

$$(3) \qquad F_x = [(RRSS-URSS)/k]/[URSS/(n-k)],$$

where RRSS and URSS are, respectively, the restricted and unrestricted residual sum of squares, k is the number of restrictions or levels in parameter x, and n is the degrees of freedom in unrestricted equation.

In salary surveys, n is large relative to k. If k is the number of correctly specified levels of the factor X, then let α = measure of irrelevant fineness in another measure, say Y. That is, suppose instead of using k levels, we use the αk levels of Y, where $\alpha > 1$. Then, as long as the levels of X are a linear combination of the levels of Y, and n is large relative to αk, the URSS of the equation will be almost the same, the RRSS will be the same, so the ratio of F_y (the F-statistic of the inefficient parameter Y) to F_X is as follows:

$$(4) \qquad F_y / F_x \cong (n - \alpha k)\, k \,/\, [(n-k)\,\alpha] = [(n/\alpha) - k\,] \,/\, (n - k) \cong 1/\alpha$$

The maximum of the ratio is one (where X=Y, so α=1); otherwise it decreases monotonically with increasing α, and approaches $1/\alpha$ for n large and k small. So the F-statistic depends not only on the economic relevance of the parameter X, but also on the efficiency with which it is measured. Since the purpose of this work is to identify the potential explanatory power of variables based on establishment, attention should focus primarily on the percentage sum of squares explained by factors rather than on F-statistics.

The essential complication to variance-component decomposition of salary survey data is that the data are unbalanced. An unbalanced design produces collinearity between the vectors of employer and occupation indicator variables in equation (2), preventing a simple separation of their impacts. If an establishment employs a relatively large number of workers in a skilled occupation, we cannot distinguish whether a differential paid to those workers is due to their employer or to their occupation.

Techniques for estimation of variance components of a model of unbalanced design are detailed in Searle (1971) and Henderson (1953). Restricted maximum likelihood techniques -- introduced in Hoching, Hackney and Speed (1978) -- provide estimates of variance components and their standard errors at the expense of imposing a rigid structure on the distribution of level effects and errors. Because there is no way to judge the appropriateness of the structure for any industry or employer, and because the purpose of this study is to investigate the characteristics of establishment differentials, a nonparametric method was preferred for this analysis. Groshen (1986) presents examples of the application of alternative ANOVA techniques to similar data.

Thus, the technique applied is a decomposition of the sum of squares of wages, rather than an explicit estimation of variance components. This approach is nonparametric and avoids ANOVA's difficulty with unbalanced data. A variance is a sum of squared

deviations divided by the appropriate number of observations or degrees of freedom. In data with an unbalanced design, the correct number of degrees of freedom is unknown, so variance estimates must rely on estimates of the correct degrees of freedom. Such estimates require the imposition of structure on the data. This method provides a measure of the collinearity arising from design imbalance instead of imposing a structure on estimated differentials to resolve the ambiguity.

The summary of the technique provided in table 2 shows how a series of ordinary least squares regressions is used to decompose the variation of wages in an unbalanced salary survey data set into occupation, employer, job-cell and individual components. Changes in the R^2 (the sum of squares explained as a proportion of total) partition wage dispersion into components and standardize the variance of wages to a value of one.

First, in the pooled sample, log wages are regressed separately on vectors of occupation and establishment indicators and then on both sets of indicators together (the full main-effects model). The marginal contribution of each set of indicators to the full main-effects model (over the equation with the other one alone) measures the portion of wage variation associated unambiguously with that factor. These correspond to minimum estimates of the relative size of the variance contributed by occupation and differentials. The difference between the R^2 of each in the equation alone and their marginal contribution to the full main-effects equation is a measure of their joint (collinear, or ambiguous) explanatory power.

Next, the contribution of the interaction (job-cell) differentials, which indicate internal labor market differences, is the difference between the explanatory power of a regression on job-cell indicators and that of the full main-effects model. The individual contribution is the share of variation unexplained by job-cell indicators.

A final issue addressed in this paper was whether salary survey job classifications could really be expected to pick up human capital as productively used. To test this, standard measures of human capital as regressors can be added to CPS wage regression with occupation indicators. After control for occupation, these measures of human capital add very little explanatory power, which suggests that occupational indicators even as coarse as CPS codes control well for traditional measures of human capital.

B. Do Wage Differences Among Employers Last? (Groshen 1989)

Findings: This paper confirms previous findings on employer wage differentials and provides strong evidence against the possibility that they may simply reflect temporary, random errors by wage-setters. The variance of wages is analyzed in a six-year AWS panel and a 36-year panel from the CSS. In addition, the results are compared

to a CPS sample in order to estimate the importance of interemployer wage variation in the economy as a whole.

The results independently confirm previous estimates of the size of employer wage differentials, even with careful control for local conditions. The most conservative estimate of employer wage differentials here yields a standard deviation of approximately 12 percent within industry, or 18 percent, including interindustry differentials.

Then, the paper identifies three key characteristics of establishment wage differences. First, the major contribution of this paper is the finding that employer wage differences, and rankings of employers by wage, are virtually stationary over six years, and strongly persistent for as long as 35 years. Thus, their determinants must be long-lived employer characteristics. Second, although wages increase with size of establishment, changes in plant size have no simple, consistent relationship with wage level. Third, employer wage differences are found in all industries, among white- and blue-collar workers, which argues for explanations that apply across-the-board to all occupations in an establishment, and to the establishments in most industries.

These results cast more light onto the nature and the plausibility of proposed explanations for employer wage differences. The evidence on size and strong stability over time presented above rejects random or temporary variations (generated or perpetuated by costly information) as the explanation of employer differentials.

Methodological issues: Two methodological challenges faced in this paper were summarizing and testing the persistence of establishment wage differences in the AWS for 6 years and in the CSS for 36 years, and comparing the AWS to the IWS and the CPS to establish whether the results were consistent with more familiar data sets.

The persistence of employer wage differentials in the AWS was shown two ways. First, interacting the ANOVA components from table 2 with time showed that establishment interactions with time were small. Second, when employer wage differences are estimated independently for each year, rank and Pearson correlations matrices can reveal persistence over time. In the AWS, these reported correlations were casewise: one estimated coefficient for each possible pairing of years. The length of time under study in the CSS favored use of the unusual, but parsimonious, method of pairwise deletions (rather than a 36x36 matrix of casewise deletions) to summarize the decay pattern of employer and occupation wage differentials. That is, starting with employer differentials for each year (calculated as described above), a correlation coefficient is calculated for all pairs of observations one year apart, two years apart, and so on, up to 35 years apart.

The decay of CSS employer and occupation effects over time is plotted in Groshen (1989). The employer effect correlation begins at 0.9 for all pairs of observations one

year apart. The correlation then declines perceptibly to the neighborhood of 0.6 as the time between observations lengthens to 14 years. In the subsequent 20 years, the correlations show only a slight tendency to decline further. By contrast, city-specific occupation effects decay more slowly than do employer effects. Correlation coefficients on occupation one year apart have a correlation coefficient of 0.98, and decline slowly for 19 years. For observations 19 to 34 years apart, the correlations are essentially flat and bounded between 0.8 and 0.9.

The second methodological issue addressed in the paper is comparison to the CPS. Table 3 illustrates the differences in results of wage decompositions among the CPS for May 1977 and two salary surveys: the IWS from Groshen (1991b) and the AWS used in Groshen (1989). May 1977 falls within the ranges of both the AWS and IWS studied here. The IWS estimates are the simple means from ANOVA of the wages of production workers in six manufacturing industries. The AWS estimates are repeated from table 5 of Groshen (1989)--except that all interactions with time have been removed.

Since these three data sources differ substantially, adjustments for the differences are necessarily speculative. For instance, the standard deviation of wages in the AWS (0.40 log points), is double the mean for the six IWS surveys (0.20 log points). As noted above, area surveys cover a broader mix of occupations, both blue- and white-collar, and capture the effects of interindustry wage variation. The CPS includes all sources of variation already mentioned, in addition to covering the full range of occupations in the economy.

The first two rows of table 3 present the numbers least comparable across the three surveys: standard deviation estimates for total dispersion and those due to occupation, sex, region, and industry differentials. Reported AWS and IWS figures allocate the entire joint (collinear) occupation-establishment wage effects to occupation. In the IWS, the variance in the first row includes regional variation, but not interindustry variation, while the opposite is true of the AWS.

In the CPS, the first row captures industry, regional, occupational, and gender-related wage variation. The level of detail of region and industry are roughly the same in the AWS and CPS, but CPS three-digit occupations lack the detail of the job classifications in the IWS and AWS. In the last column of the first row, we see that CPS wage variation in the occupations covered by the AWS is about the same as that of the AWS. This suggests that wage variation within the CPS occupational categories is greater than the variation between regions of the country. Lack of occupational specificity leaves more wage variation unexplained than the addition of regional controls can capture.

Another way to judge the impact of the coarse occupations in the CPS is to note that in the plastics industry, contraction of the 42 BLS job classifications into 12 CPS

occupational categories reduces the R^2 of the regression equation by one half, from 0.49 to 0.25. In an ANOVA as shown, at least half of this difference -- judging from the size of the contribution "joint" to occupation and establishment -- might then be claimed by establishment differentials, raising the estimated employer effect in the CPS.

The second row shows the remaining variation for each sample. These are quite similar for the AWS and IWS: a standard deviation of about 0.16. The CPS, however, retains a standard deviation of 0.31, almost twice as high. The next two rows speculate on the size of the within-industry establishment effect in the CPS, in order to provide bounds for the probable contribution of establishment to CPS wage variation. The first method takes the point estimate of standard deviation from the IWS and AWS: 0.11. Although this is a large portion of the unexplained standard deviation of 0.31, the estimate is conservative because CPS occupations are very broad. The large joint component of variation in the IWS and AWS would shrink with these broad occupations, increasing the estimated establishment impact on variation. In addition, the IWS and AWS oversample large establishments. Since estimated establishment variance is highest among the smallest establishments, employer wage variation may be higher in the CPS than in the IWS and AWS because the CPS samples evenly from all sizes of employers. (Assigning the AWS establishment percentage of total wage variation to establishment in the CPS converts this to a standard deviation of 0.13.) The second (less conservative) method assigns to establishment the same percentage of remaining variation (after occupation, industry, etc.) as found in the AWS. That converts to a standard deviation of 0.20.

In order to see if the limited number of occupations surveyed in the AWS accounted for these results, the last column of table 3 presents the same exercises on the subsample of CPS observations for workers in AWS occupations -- 24 percent of the CPS sample. The variance is lower in the subsample, but the entire decrease occurs in the between-occupation portion of variance -- leaving estimates of the employer effect unchanged.

But how much of the remaining variation is noise? CPS wage reports probably have a larger noise-to-signal ratio than BLS wage surveys for four reasons: 1) CPS average hourly earnings are imprecisely defined (they include overtime or shift premia and second jobs earnings); 2) CPS respondents' memories are more subject to error than the employer records used by the BLS; 3) CPS data-cleaning is less thorough than BLS efforts; and; 4) CPS occupations are subject to large reporting error (see Mellow and Sider 1986), so, the nonoccupation variation in the CPS is biased upwards.

Thus, compared to total wage variation in the CPS, estimated variation due to establishment differentials is large, even by conservative measures.

C. Rising Inequality in a Salary Survey: Another Piece of the Puzzle (Groshen 1991d).

Findings: Recent studies of wage inequality conclude that rising inequality in the recent past has made U.S. family income less equally distributed. This paper uses data from the CSS to examine the role of occupational distinctions and employer compensation practices in the recent rise in U.S. wage dispersion.

Most previous studies of rising wage inequality in the U.S. have been based on household surveys--particularly the CPS--with two consequences. First, this highly publicized phenomenon has not been confirmed in many alternative data sources. Second, the reasons for rising inequality are still not fully understood because only part of the increase can be associated with CPS variables. This paper reaches beyond the CPS to study the time path of wage dispersion in data with fine detail on both occupation and employer.[6] The results show that the wages of two hypothetical nonproduction workers who differed in both occupation and employer would have pulled apart over the past three decades, even if neither changed occupation or employer.

During the 1960s, inequality rose mainly as a result of increasing occupational wage differentials and internal labor market variations, and this pattern continued throughout the 1970s. In addition, wage differences among employers underwent a large, apparently permanent increase in dispersion as union and industry wage differentials expanded in the late 1970s. During the 1980s, the only evident source of rising inequality was the widening of occupational wage differentials, which can be linked to increased returns to general education. Finally, despite reports suggesting otherwise, growing use of merit raises had no noticeable impact on wage variation during the 1980s or before.

These results confirm the existence of rising inequality and reject one important hypothesis. Because wage disparity among nonproduction workers increases even when companies and occupations are held constant over time, the rise is not primarily attributable to the direct effects of the net creation of unusually unequal jobs.

Methodological Issues: The CSS data allow a close focus on the role of changing returns to occupation and employer attributes, but are not well suited to a study of the changing composition of jobs over time, since the sample is not randomly drawn (that is, entry and exit from the sample do not necessarily reflect the birth or death of jobs in the economy). Thus, the main methodological goal was to purge the data of any impact of changes in composition, in order to focus attention on the role of changing wages for

---

[6] Two other studies of widening wage inequality that use employer-based data contribute importantly to our understanding of this phenomenon, but they are limited by their inability to control for occupation. See Leonard and Jacobson (1990) and Davis and Haltiwanger (1991).

constant jobs. A second challenge was to identify the characteristics of occupations with rising wage differentials.

One way to gauge the growth of inequality in the CSS is to compare annual between-cell standard deviations of wages, with a weight of one per cell, which controls for the effect of changes in the number of workers in cells over time. However, since occupations and employers are added and deleted from the sample over time, simple between-cell measures of variance do not control for the possibility that the survey now includes more-diverse occupations and firms than previously. To control for sample changes, this paper uses a "rolling sample" technique. Between any two years, the change in variation is measured only for the subsamples of job cells that are present in both years.

These changes are then added to the cumulative sum of previous changes plus the initial variance. The square roots of those estimated variances are the estimates of the standard deviation of wages for an unchanged job-cell. The alternative method of controlling for compositional changes is to study only the job-cells that remain in the sample for the whole 36 years. However, this latter approach retains very few observations in long-lived data such as the CSS, while the preferable rolling sample technique minimizes the number of observations eliminated.

Which types of occupations gained relative to others during this period? The general approach is to look for evidence of an increase in the returns to both formal education and skill in the widening occupation differentials. To do this, I merged information on job attributes with the CSS data. Although many attributes could be examined, two generally explain 60 to 70 percent of the variation in occupational wage differentials. These are "specific vocational preparation" and the average of "general education development" of three types: reasoning, mathematical, and language. To discern changes in the rewards to these factors over time, I regressed occupational differentials on these characteristics in each year. The CPS finding of increased returns to education is confirmed in the CSS and seems to explain much of the increase in wage variation among occupations.

D. Synopses of Other Research Using Salary Surveys

Do Hostile Takeovers Reduce Extramarginal Wage Payments? (Gokhale et al. 1995). Part of the shareholder gains from hostile takeovers may stem from breaches of implicit contracts, leading to transfers of extramarginal wage payments from workers to shareholders. Our tests of this expropriation hypothesis, using the CSS, improve on existing research by using employer (not industry) level data, and by performing both ex ante and ex post tests. We test two forms of this hypothesis, that acquirers extract shared

rents from workers, and that they extract quasi-rents from senior workers. The ex ante results provide no evidence that extramarginal wages paid to all or more-tenured workers in a firm are associated with subsequent hostile takeovers. Furthermore, we find ex post evidence that employer wage differentials rise after a hostile takeover, in contradiction either of the hypothesis that hostile takeovers lead to the expropriation of shared rents, or that these differentials represent shared rents. However, most of our ex post results are consistent with the hypothesis that hostile takeovers (but not other types of mergers) reduce extramarginal wage payments to senior workers.

The Structure of the Female/Male Wage Differential: Is It Who You Are, What You Do, or Where You Work? (Groshen 1991c): Any policy to reduce wage differences between men and women addresses specific components of the gap, and the policy's potential efficacy depends on the magnitude of the component(s) targeted. The role of workplace segregation in wage inequality has been relatively neglected, even though prior work suggest that it may be a large part of the wage gaps among races or the sexes (see Blau 1977, Buckley 1971, and McNulty 1967). This study uses five IWS samples to further the evaluation of policy by jointly estimating all four components of the female/male wage gap within industries -- the individual effect and the effects of segregation by occupation, establishment, and job-cell, the latter two for the first time. Although men and women who work together in a job-cell earn about the same amount, such integration is rare. Occupations are either mostly male or female, and within establishments, occupations are almost totally segregated. So, even people who choose integrated occupations work primarily with members of their own sex. Each type of segregation tends to raise men's wages and lower women's.

The Effects of Inflation on Wage Adjustments in Firm-Level Data: Grease or Sand? (Groshen and Schweitzer 1996): This paper explores the impact of inflation on the labor market with an eye toward distinguishing positive effects (greasing the wheels by facilitating real wage adjustments to intermarket shocks) from the negative ones (throwing sand in the gears by distorting intramarket relative wages). We study wage changes in the CSS from 1956 through 1992. This paper's strength -- the unusually tight link we forge between our analytic approach and common compensation adjustment practices -- is made possible by studying a long salary survey. The analysis interprets intramarket (employer) wage adjustments as likely to include errors and corrections or deviations in speed of adjustment, while inflation-induced inter-market (occupational) wage changes should display a higher concentration of responses to real shocks. Relying on this distinction to interpret our results, we estimate the relationship between the standard deviation of employer and occupational wage adjustments and measures of inflation. In support of the

model, we find that large, independent employer and occupation components in wage changes. In the CSS, moderate inflation (below about 4 percent) appears to increase the speed of transmission of interoccupational wage adjustments. But inflation also exacerbates potentially confusing errors and corrections, or lagged adjustments by employers. The costs of inflation have the steeper slope and a later peak over the range observed in this study, suggesting that inflation's costs continue to rise long after its potential benefits have been exhausted.

VII. Conclusion

This paper introduces readers to the contributions of, and challenges posed by, the use of American employer salary surveys for labor market research. These underused surveys are now ripe for exploitation, due to advances in theory, computing power, and econometric technique.

Admittedly, these data have several drawbacks compared to the well-known household surveys: obtaining access requires contacts and ingenuity, their preparation for analysis can be a labor-intensive investment, and their analysis often requires application of unfamiliar statistical techniques. Nevertheless, they frequently provide valuable information about the workings of the labor market that cannot be duplicated by any other source -- particularly in the U.S., where the federal government does not sponsor a broad employer-based survey of wages. The lines of research mentioned here were made possible by the unique data contained in salary surveys. Further advances in understanding the role of employers will depend on the ability of labor economists to continue forming contacts with the business and policy communities. These contacts will allow researchers to better understand the intentions and activities of employers, and to gain access to the data necessary to gauge the impact of these activities on labor market outcomes.

References

Abowd, John, Francis Kramarz, and David Margolis. "High-Wage Workers and High-Wage Firms," INSEE Working Paper 1994.

Akerlof, George A. "Jobs As Dam Sites." *Review of Economic Studies*, 48 (January 1981), 37-49.

Akerlof, George A. and Janet L. Yellen. Efficiency Wage Models of the Labor Market. Cambridge, UK: Cambridge University Press, 1986.

Atrostic, B.K. "Alternative Pay Measures and Labor Market Differentials." Working Paper No. 127, U.S. Department of Labor Office of Research and Evaluation, Bureau of Labor Statistics, March 1983.

Blau, Francine. Equal Pay in the Office. Lexington, Mass: D.C. Heath and Co., 1977.

Brown, Charles and James L. Medoff. "The Employer Size Wage Effect." *Journal of Political Economy*, 97 (October 1989), 1027-59.

Buckley, John E. "Do Area Wages Reflect Area Living Costs?" *Monthly Labor Review*, 102 (November 1979), 24-29.

Bulow, J.I., Summers, L.H. (1986). -- "A Theory of Dual Labor Markets with Applications to Industrial Policy, Discrimination, and Keynesian Unemployment", *Journal of Labor Economics*, 4, pp. 376-414.

Davis, Steve J. and John Haltiwanger. "Wage Dispersion Between and Within US Manufacturing Plants: 1963-86; Comments and Discussion." *Brookings Papers on Economic Activity*, 1991, 115-200.

Dickens, W.T., Katz, L.F. (1987). -- "Industry Wage Differences and Theories of Wage Determination," NBER Working Paper No. 2271. Cambridge MA: National Bureau of Economic Research.

Dunlop, John T. "Fundamentals of Wages and Labor Markets." Unpublished paper, Harvard University. August 30, 1982.

Dunlop, John T. "The Task of Contemporary Wage Theory." In G. Taylor and F. Pierson, eds., New Concepts in Wage Determination. New York: McGraw-Hill, 1957.

Freedman, A. (1976). -- *The New Look in Wage Policy and Employee Relations*, New York: The Conference Board Report #865.

Gokhale, Jagadeesh, Erica L. Groshen, and David Neumark. "Do Hostile Takeovers Reduce Extramarginal Wage Payments?" *Review of Economics and Statistics*, (August 1995).

Groshen, Erica L. "Five Reasons Why Wages Vary Among Employers." *Industrial Relations*, 30, 3 (Fall 1991a), 350-381.

Groshen, Erica L. "Sources of Intra-Industry Wage Dispersion: How Much Do Employers Matter?" *Quarterly Journal of Economics*, (August 1991b), 869-884.

Groshen, Erica L. "The Structure of the Female/Male Wage Differential: Is It Who You Are, What You Do, or Where You Work?" *Journal of Human Resources*, 26, 3, (Summer 1991c), 457-472.

Groshen, Erica L. "Rising Inequality in a Salary Survey: Another Piece of the Puzzle," Federal Reserve Bank of Cleveland Working Paper no. 9121 (1991d).

Groshen, Erica L. "Ability to Pay, Rent Capture, and Salaries in the Private Sector." In Proceedings of the Fortieth Annual Meeting, Madison, WI: Industrial Relations Research Association, (1990), 186-194.

Groshen, Erica L. "Do Wage Differences Among Employers Last?" Working Paper No. 8906, Federal Reserve Bank of Cleveland, (1989) revised 1991.

Groshen, Erica L. and Alan B. Krueger. "The Structure of Supervision and Pay in Hospitals," *Industrial and Labor Relations Review*, 43 (February 1990), S143-46.

Groshen, Erica L. and Mark E. Schweitzer. "The Effects of Inflation on Wage Adjustments in Firm-Level Data: Grease or Sand?," Federal Reserve Bank of New York Staff Reports No. 9, (January 1996).

Groshen, Erica L. "Sources of Wage Dispersion: How Much Do Employers Matter?" Ph.D. Dissertation. Harvard University, 1986.

Henderson, C. R. "Estimation of Variance and Covariance Components." Biometrics, (June 1953).

Hocking, Ronald R., O. P. Hackney and F. M. Speed. "The Analysis of Linear Models with Unbalanced Data." in David, H. A., ed., Contributions to Survey Sampling and Applied Statistics: Papers in Honor of H. O. Hartley. New York: Academic Press, 1978.

Kerr, Clark. "The Intellectual Role of Neorealists in Labor Economics." *Industrial Relations*, 22 (Spring 1983), 298-318.

Kremer, Michael and Eric Maskin. "Segregation by Skill and the Rise in Inequality." unpublished paper, Harvard University, February 1994.

Krueger, Alan B. "Ownership, Agency, and Wages: An Empirical Analysis," Princeton University, Unpublished paper, April 1987.

Leonard, Jonathan S. "Wage Structure and Dynamics in the Electronics Industry," *Industrial Relations*, 28 (Spring 1989), 251-75.

Leonard, J.S., Jacobsen, L. (1990). -- "Earnings Levels and Job Turnover," *The American Economic Review*, 80, pp. 298-302.

Levine, David I. "Fairness, Markets, and Ability to Pay: Evidence From Compensation Executives." *The American Economic Review*, vol. 83, no. 5 (November 1993) pp. 1241-1259.

Levy, F., Murnane, R. (1992). -- "Earnings Levels and Earnings Inequality: Review of Recent Trends and Proposed Explanations," Journal of Economic Literature, 30,3, pp. 1333-1381.

Rees, A., Shultz, G.P. (1970). -- *Workers and Wages in an Urban Labor Market*, Chicago: University of Chicago Press.

Reynolds, Lloyd G. The Structure of Labor Markets. New York: Harper & Brothers, 1951.

Salop, Steven C. "A Model of the Natural Rate of Unemployment," *The American Economic Review*, 69 (March 1979), 117-25.

Searle, S.R. Linear Models. New York: Wiley, 1971

Segal, Martin. "Post-Institutionalism in Labor Economics: The Forties and Fifties Revisited," *Industrial and Labor Relations Review*, 39 (April 1986), 388-403.

Van Giezen, Robert W. "A New Look At Occupational Wages Within Individual Establishments." *Monthly Labor Review*, 105 (November 1982), 22-28.

Ward, Virginia L. "Measuring Wage Relationships Among Selected Occupations." *Monthly Labor Review*, 103 (May 1980), 21-25.

Weitzman, M. (1986). *The Share Economy: Conquering Stagflation.* Cambridge, MA: Harvard University Press.

Table 1

Comparison of Major Sources of U.S. Employer Microdata

	Longitudinal Research Database	**Unemployment Insurance Records**	**Employer Salary Surveys**
Coverage	Manufacturing	All industries	Each limited by area, industry or occupation
Plant Information	Highly detailed, can merge in more	Industry, size, location, ownership	Name of employer plus some other characteristics, can merge in or ask more
Worker Information	Mean production and non-production worker earnings, no occupation	Quarterly mean establishment or individual earnings, no occupation	Individual wages by detailed job description

Table 2

A Technique for Partitioning the Sum of Squares of Wages
in Unbalanced Salary Survey Data

Source of Variation	Share of Total Sum of Squares*
1. Occupation (controlling for establishment)	$R^2_C - R^2_B$
2. Joint (collinear) occupation and establishment	$R^2_A + R^2_B - R^2_C$
3. Establishment (controlling for occupation)	$R^2_C - R^2_A$
4. Job-cell (controlling for occupation and establishment)	$R^2_D - R^2_C$
5. Total between job-cells	R^2_D
6. Individual	$1 - R^2_D$
TOTAL	1

* The subscripts on the coefficients of determination correspond to the regression models listed below. For ease of exposition, occupation, sex, region, and incentive (which are all available in BLS Industry and Area Wage Surveys, but not in most others) are listed simply as occupation.

A. $w_{ijk} = \mu + X_i\alpha + \varepsilon_{ijk}$

B. $w_{ijk} = \mu + Y_j\beta + \varepsilon_{ijk}$

C. $w_{ijk} = \mu + X_i\alpha + Y_j\beta + \varepsilon_{ijk}$

D. $w_{ijk} = \mu + X_i\alpha + Y_j\beta + X_iY_j\gamma + \varepsilon_{ijk}$

where w_{ijk} = ln wage of individual k in occupation i at establishment j

X_i = vector of occupation indicator variables for occupation i

Y_j = vector of establishment indicator variables for establishment j

X_iY_j = interaction indicator variables for occupation i in establishment j, that is, for job-cell ij, and

$\mu, \alpha, \beta, \gamma, \varepsilon$= estimated parameters.

Table 3

Industry and Area Wage Survey Standard Deviation Components
Compared to Current Population Survey Log Wage Variation

Source of Variation of Log Wage	**Industry Wage Surveys** Mean Suggested Standard Deviation	**Area Wage Survey** Suggested Standard Deviation[b]	**Current Population Survey May 1977**[c] All Occupations	AWS Occupations[d]
Total standard deviation	0.20	0.40	0.48	0.42
Occupation, sex, region, and/or industry[a]	0.12	0.36	0.36	0.25
Total remaining	0.16	0.16	0.31	0.33
Establishment (known)	0.11	0.11	-	-
Establishment (estimated)				
1)AWS & IWS point estimate	-	-	0.11	0.11
2)AWS % of remaining	-	-	0.20	0.21
Internal labor market[e]	0.06	0.10	-	-
Individual	0.09	0.07	-	-

Notes:
a. For Industry Wage Survey and Current Population Survey, includes city and region indicators. For Industry and Area Wage Surveys, includes incentive indicator and joint effects. In Current Population Survey, uses 3-digit occupation. Current Population Survey and Area Wage Survey totals include 2-digit industry.
b. Effects of interactions with year have been excluded from Area Wage Survey results.
c. The Current Population Survey sample includes all private-sector full-time workers between the ages of 18 and 65 with reported average hourly earnings of more than $1.75.
d. Including only observations for occupations included in the Area Wage Survey sample.
e. Occupation-establishment interaction, or common job-cell component.

Source: Author's tabulations from U.S. Bureau of Labor Statistics Industry and Area Wage Surveys (see Groshen 1989 and 1991b), and May 1977 Current Population Survey.

Figure 1

Annual Wage-Setting Mechanism for Large Employers

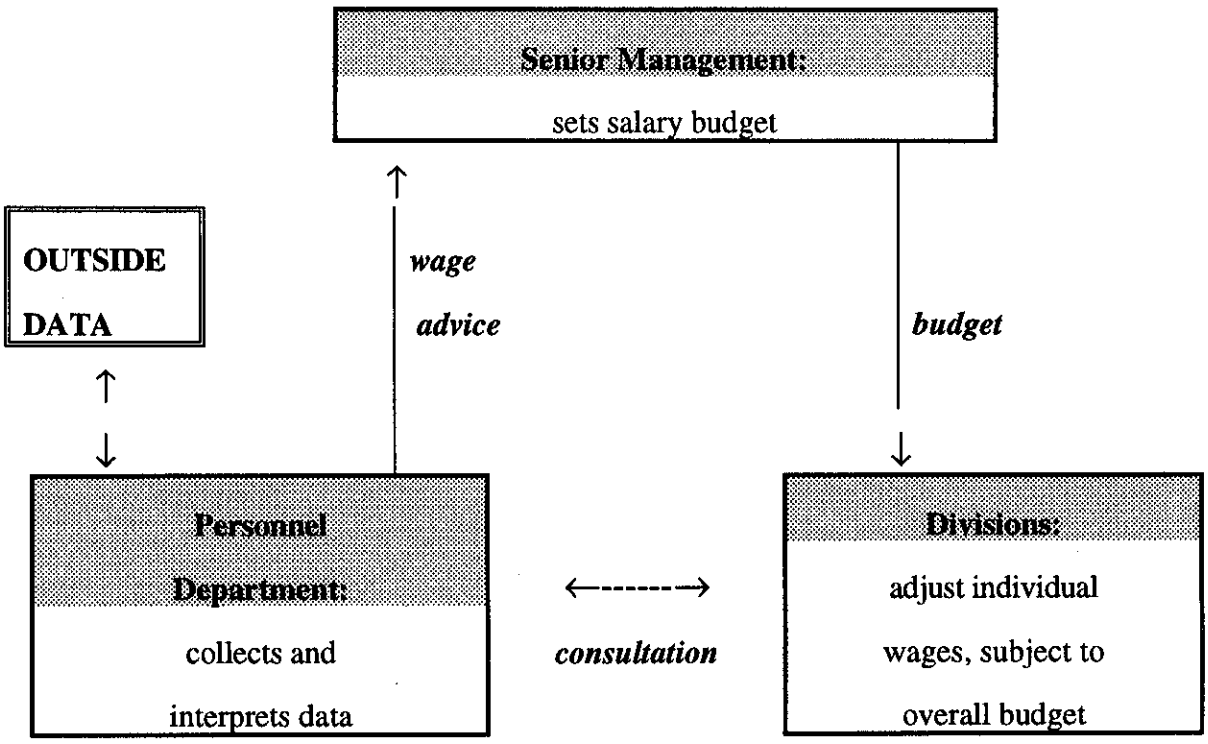

Figure 2

Coverage Comparison:
Household Survey Versus Employer Salary Survey

Occupation		Employer									Unemployed/ Out of Labor Force
		A	B	C	D	E	F	G	H	I	
O	1	o●	oo	oo	o●		o●	●●		ooo	oo
c	2		o	●	o●						
c	3	o				o	●		o●	o●●	●
u	4	●	o	o	oo	o●●				o	
p	5				o	●	o			o	o
a	6	o●		ooo			o●	o	o●		ooo●
t	7	o		o●	o●	o			oo		
i	8						ooo	●		o	o
o	9		o	o	●		●		o		●
n	10	●●●			o			o		oo●	

● = Household survey sample coverage (e.g., Current Population Survey, Census)

▒ = Employer salary survey coverage (e.g., Industry and Area Wage Surveys)

Note: In the case of unemployed or put of the labor force workers, occupation does not strictly apply, but can be defined based on previous or intended occupation.