AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

WAGE STRUCTURE
IN THE
SUPERMARKET INDUSTRY
1984 - 1993

A PLAN B PAPER

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

ROBERT C. JOHANSSON

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

JAY S. COGGINS

Name of Faculty Adviser

Jay S. Coggins

Signature of Faculty Adviser

July 21, 1997

Date

378.776
D346
J646

ABSTRACT

A number of studies have used estimation procedures to model wage equations to account for union self-selection endogeneity. However, no recent study has examined the supermarket industry using these various procedures using an extended year data set. This Plan B Paper uses standard ordinary least squares, full-information maximum likelihood, and an instrumental variables approach to estimate the coefficients in the standard wage determination equation. It uses data from the Current Population Survey from 1984 - 1993, a period of declining real wages and of declining union influence.

ACKNOWLEDGMENTS

I would like to take a moment to thank the many people that have helped me in this endeavor. Firstly, I am thankful for my colleagues' timely suggestions and ideas that have helped shape this paper; especially Mona Sur, Amyaz Moledina, Asitha Sandanayake, and Jeff Stensland. Secondly, I am thankful for the guidance of the members of my committee, Professor Ben Senauer, Professor John Budd and especially Professor Jay Coggins, my advisor.

In addition I am extremely grateful for the love, patience and understanding of my family over this past year. They have helped me keep perspective. Lastly, I would like to thank my wonderful wife, Siobhan, without whom none of this would have been possible.

TABLE OF CONTENTS

TABLES AND GRAPHS

I. INTRODUCTION

The last rigorous national examination of the structure of retail food industry wages was conducted in 1977 by the U.S. Department of Labor and by The Executive Office of the President: U.S. Council on Wage and Price Stability under the Carter administration. Since 1977 there have been many targeted studies of small sections of the supermarket industry, but for time periods of less than three years. Obviously there have been many changes in the retail food industry over the last 20 years and many of these changes should be reflected in the industry's wage structure. For example, at the time of the Department of Labor survey in 1977, union influence was near its peak (see Table 1). Since that time union membership and influence have declined. In 1993 a reported 25.7% of supermarket employees listed themselves as union members, a decline from 29.4% in 1984 (Current Population Survey: National Bureau of Economic Research CD-Rom, 1995).

It is the intention of this study to review the changing supermarket industry and the current prevailing literature on wage determination and to examine the structure of wages in the supermarket industry from 1984 through 1993. More specifically, the effect of unionization on wages, controlling for demographic variables such as sex, age, education, job category, and region, will be closely examined. This analysis will draw on supermarket wage data from the Current Population Survey (CPS) of the U.S. Census Bureau (1984 - 1993). This data set was compiled by the National Bureau of Economic

Research in October, 1994[1]. Using econometric techniques, a log wage regression will be estimated using various explanatory variables from the CPS. This study will describe the results and implications of the regressor coefficients in terms of the changing supermarket work force demographics.

Another important issue that will be addressed in this paper is that of local union density effects; or "spill-over" effects of union percentage rates on wage rates in the union and nonunion sectors. It is widely accepted that the higher the union percentage in local labor markets the higher the wages for those in the union; the corresponding effect has not been closely examined for the nonunion sector, but has also been theorized to be positive.

[1] See Data section for a further description of the data sources used.

II. CHANGES IN THE SUPERMARKET INDUSTRY AND U.S. ECONOMY

It is important to review the structural changes occurring in the supermarket industry as well as in the U.S. economy to provide a theoretical basis for an estimation of the effects those changes on supermarket wages. What one can deduce from these changes will be essential in determining the correct model specification and in the verification of results; i.e. does the model yield results that are consistent with theoretical expectations?

Supermarket Industry

The grocery store, outside of the workplace or school, is one of the most frequently visited public establishments in the United States, and with all the changes in supermarket structure intended to attract customers through associated services[2] (Kinsey and Senauer, 1996) will remain so into the twenty-first century. However, soon the share of the American food dollar spent on grocery food will have fallen to less than fifty percent. Conventional supermarkets will thus continue to increase square footage per employee, implement labor saving technology, and incorporate non-traditional services and departments. The implications of these changes on the structure of supermarket wages should be reflected in an analysis of supermarket employee wages over time.

It can be seen by examining the Statistical Summaries of the U.S. economy over the last fifteen years that retail trade has become increasingly important in the total U.S.

[2] Such services include: video rental, florists, banks, etc.

economy, and that retail food has become more important within retail trade. In 1987 retail trade accounted for 22% of total employment (18+ million employees). In 1988 the Bureau of Labor Statistics (BLS) reported that 3.1 million Americans were employed in retail food stores (Table 1), an increase of 60% since 1970. In 1990 total grocery sales equaled $369 billion with the larger stores averaging sales of $10 million/year.

Table 1. Selections from Statistical Abstract of the United States, 1994

	1970	1980	1990	1993
Total Employment	78,678	99,803	117,914	119,306
Manufacturing Emp.	20,746	21,942	21,184	19,557
Retail Trade Emp.	12,336	16,270	19,618	20,163
Agriculture Employment	3,463	3,464	3,186	3,074

Table No. 641: in (1000's of workers)

	1979	1989	1993
Membership UFCW[3]	1,123,000	999,000	997,000

$1994 in billions	1980	1990
Total GDP	2,708	5,546
Manufacturing	588	1025
Retail*	245	516
Agriculture	67	112

682Table no.68

***Breakdown of Retail: in 1990**	
1529.7	(1000) retail trade establishments
186.1	(1000) retail food stores
19,815	(1000) employees in retail
3,124	(1000) employees in food stores
241.7	billion total payroll
35.8	billion food store payroll

Source: U.S. Census Bureau: Statistical Abstract of the United States, 1994. Washington, D.C. Governmental Printing Office.

[3] United Food and Commercial Workers International Union: 1775 K Street, NW, Washington, D.C. 20006

Such trends also reflect the changing culture of work in the United States. There has been a movement away from manufacturing towards the service industries. The supermarket industry itself has become more service-oriented in response to changes in the U.S. labor force and cultural shifts in grocery shopping and dining, illustrating the fact that the supermarket industry reflects major changes in the economy.

These trends are outlined in-depth in John Walsh's, *Supermarkets Transformed: Understanding Organizational and Technological Innovations.* In addition I have drawn insights from several Retail Food Industry Center[4] publications, several Food Marketing Institute (FMI) publications, and from *The Progressive Grocer.*

The transformation from small family-owned grocery stores to enormous retail centers has involved many structural changes in the supermarket industry. The changes occurring in the supermarket industry can be seen as an adaptation to the changing demographic, economic, and social environment in the United States following WWII. For example, the increasing number of women in the work force has been a major cause of many of these changes. The increase in demand for service (such as keeping stores open for more hours and increasing the availability of pre-prepared foods) has been driven by the average U.S. family's need to have two working parents. In addition, as the population growth slows from post WWII highs the competitive pressure on firms dictate a more consumer aware approach to supermarket retailing (Kinsey and Senauer, 1996).

---

[4] An Alfred P. Sloan Foundation Industry Studies Center in the Department of Applied Economics.

Another reason for change in supermarket structure is that the percentage of disposable income spent specifically on food for consumption at home has fallen from 9.1% in 1980 to only 6.9% in 1994 (this compares to small decline from 4.4% to 4.2% for food-away-from-home)[5].

These changes include increases in food choices such as seafood and organically grown produce to cater to changing consumer tastes, as well as increases in specialty services like banks, florists, and video rental outlets that appeal to two-working parent families. The supermarket industry has increased productivity by introducing new technology into their stores (such as ECR[6] and scanner technology). These computer and mechanical innovations have automated many tasks previously done by hand, such as scanners and meat packaging techniques increasing productivity greatly.

The changes in labor organization over the last 50 years have also been mirrored in the retail food industry. The change to self-service shifted work from clerks to customers in the 1930's and 1940's (Walsh, 1993). In meat departments, increasing centralization of packaging has occurred enhancing productivity, where as the increase in store size has led to a decentralization in store workers first to store managers and then to shop-floor employees.

---

[5] Food Marketing Institute publication, "The Food Marketing Industry SPEAKS 1995"; Food Marketing Institute, 1995.

[6]Efficient Consumer Response (ECR) technology refers to the use of bar coding to more closely track and cater to consumer demands as well as store inventory, shipping, and stocking.

One way supermarkets have become more competitive in the increasingly lean market environment is through enhancing productivity by increasing store square footage at a greater rate than the increase in store employees. As reported by the FMI's *Annual Financial Review 1994-95*, net profit for supermarkets averaged about 1.14% on each dollar of sales. In addition, average gross margin[7] was 25.8% of sales for those companies with warehouses and 23.4% for those without. Of that approximately 50% went to labor cost in 1994[8]. This decrease in profit margins and the high percentage of those profits going to labor coupled with the increase in the number of non-traditional grocery store components being adopted by modern supermarkets, have been shaping the structure of the supermarket labor environment over this changing landscape.

It is interesting to examine the statistics from an actual supermarket to see how some of these changes have become accepted and profitable in the supermarket industry (see Table 2). It can be noted that the dairy and pharmacy are the most profitable areas in the store, whereas video rentals, the salad bar and floral departments are the least. The low profitability of these latter departments reflect their relative infancy in supermarkets. One might expect the trend towards customer self-service to reduce the ratio of employee costs to sales as these areas become more competitive as have other traditional departments.

---

[7] Defined as the difference between what a supermarket pays for its goods and the amount for which it sells them.

[8] In one supermarket surveyed in 1995, the reported ideal ratio of labor hours to dollars of sales per week was: 1 hour labor to $100 in sales per week.

Departments	Hours/Week	Sales/Week	Sales / Hour
Front End (cashiers and office staff)	1847		
Baggers	905		
		$500,000	
Grocery	680		
Frozen Foods	80		
Dairy	88	$230,000	$271
Produce	320	$32,000	$100
Floral	103	$3,000	$29
Salad Bar	143	$4,000	$28
Meat	450	$45,000	$100
Seafood	65	$3,000	$46
Deli	540	$22,000	$41
Drug G/M	260	$82,000	$315
Video	120	$3,000	$25
Cosmetics	80	$4,500	$56
Bookstore	40	$5,000	$125
Pharmacy	210	$75,000	$357
Total Store	4800	$500,000	$104

Table 2. Supermarket Spreadsheet: Store Size: 56,160 square feet; Average Sales: $500,000 per week. Source: Anonymous Chicago Supermarket, 1994.

The U.S. Economy

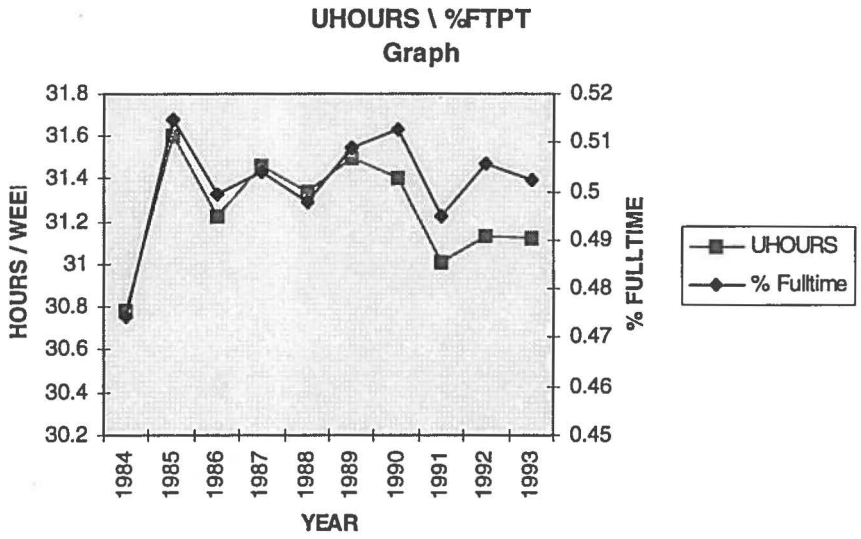
It is important to view the changing supermarket industry and its corresponding impact on its wage structure in the context of the changes that have been occurring in the economy as a whole. Income inequality is rising in the United States. From 1973 to 1993 men's median earnings fell 11%, from $34,048 to $30,407, even though the earnings of the top 20% of men grew steadily and the real, per-capita gross domestic product (GDP) rose 29%. The increasing numbers of working women during this period however contributed to a slowly increasing household median income until 1989, at which point median real wages for women working full-time year-round began to fall as well. Between 1989 and 1994 median household incomes have fallen more than 7% , from $35,585 to $31,241 after correcting for inflation and family size (Thurow, 1995). At the same time the share of total net worth of the top one-half of 1% of the U.S. population rose from 26% to 31% between 1983 and 1989. By the early 1990's the share of wealth in the hands of the top 1% of the population was more than 40%, double what it had been in the mid-1970's. These figures are similar to those of the late 1920's, before the introduction of progressive taxation. The Gini index for U.S. households, a measure of inequality that takes values from 0 (perfect equality) to 1 (perfect inequality), rose from 0.379 in 1950, to 0.396 in 1970, to 0.428 (Nelson, 1994). Mean real household income for the lowest quintile of American households rose from $6,304 to $7,195 in 1990, an

increase of 14%. Over the same period, mean real household income for the highest quintile rose from $68,622 to $87,137, an increase of 27% (U.S. Census Bureau, 1997).

Comparing mean wage values in the supermarket industry to those from the U.S. economy, it can be noted that both have steadily declined over the last twenty years in real terms. Supermarket wages have, however, traditionally been larger than the mean hourly wage for the economy as a whole, and although this was evident in 1993, the premium for supermarket jobs had fallen. The average hourly real wage for non-supervisory, non-production workers in all U.S. industries has fallen from $6.34 in 1972 to $4.97 in 1993 (in $1982), a decline of 21.6%. This compares to a decline of 29.7% in real, non-supervisory supermarket wages (BLS, 1997).

There are many reasons for the downturn in wages over the last twenty years. First, changes in the proportion of full-time workers might help to explain falling wages and earnings (see Graph 1). Over the period from 1972 to 1993 average weekly hours worked fell from 33.0 to 29.8 for non-supervisory grocery workers. Part-time workers receive lower wages and non-income benefits in general than the full-time co-workers. Minimum wage levels, which have not kept pace with inflation, may help also explain falling real wages. Technological advances also contribute to the decline in real wages by increasing the wage differential between skilled and unskilled workers. Furthermore, the fact that developing nation wage competition has undercut unskilled labor pay in free-trading, mobile economies such as the United States could partially explain the decline in

real wages. It can be seen that new technology has increased the demand for skilled workers, driving the wages further apart increasing the distributional inequity. Indeed, it was reported in 1976 that in future collective bargaining negotiations employers in the retail food industry may try to trade off wage increases for changes aimed at increasing productivity through automated check-out stands, new ECR technology, and the centralization of meat processing and packaging[9].

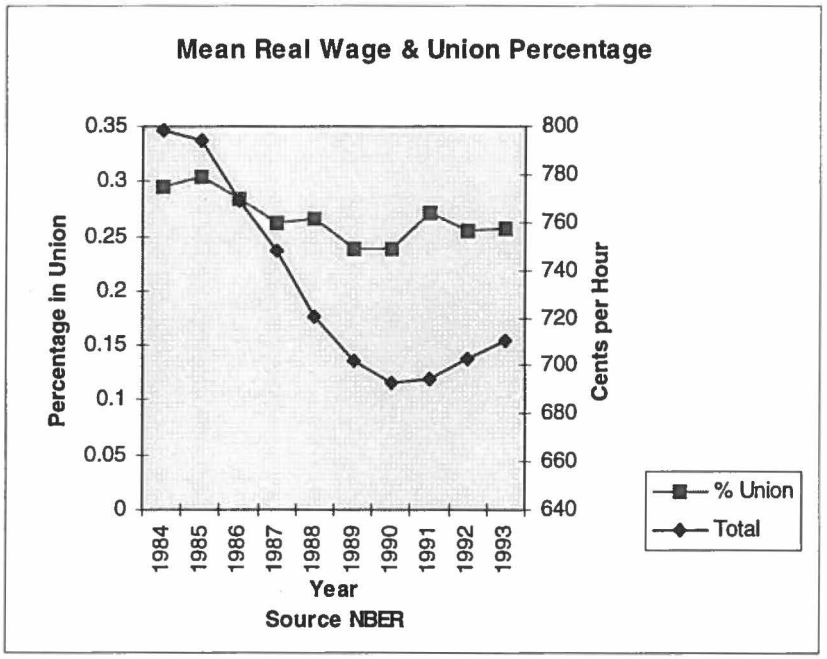


Graph 1: Usual Hours -vs- Percentage of Full-time Supermarket Employees. Source NBER, 1995.

Finally, the proportion of an industry's workers belonging to trade unions may be an important determinant of real income and relative income inequality: only 12 percent of workers are organized in the United States (down from 30 percent in 1970), compared to 40 percent in Germany (see Graph 2). However, in the supermarket industry the

[9] Executive Office of the President. Council on Wage and Price Stability. January 1976.

percentage of workers in a union is much larger than the mean, which perhaps could explain some of the differences between the economy as a whole and the supermarket industry. Past studies, including recent work by Harvard economist Richard Freeman, have indicated a correlation between lower income inequality and relatively high unionization (Freeman and Medoff, 1984).



Graph 2. Mean Real Wages -vs- Union Percentage for Supermarket Employees. Source: NBER, 1995.

III. WAGE DETERMINATION

"The Lord said to Moses, 'Say to the people of Israel, . . .your valuation of male from twenty years old up to sixty years old shall be fifty shekels of silver, according to the shekel of the sanctuary. If the person is a female, your valuation shall be thirty shekels . . . . And if the person is sixty years old and upward, then your valuation for a male shall be fifteen shekels, and for a female ten shekels.'"

The Bible, Revised Standard Version, Leviticus 27: 3-7

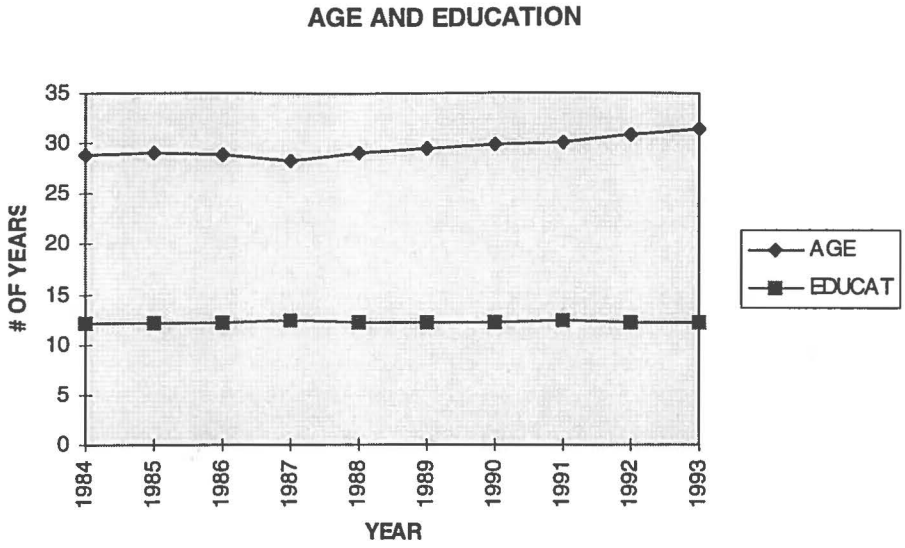
The economic basis for wage determination studies is derived in most part from human capital theory. Roots of this can be traced back to *The Wealth of Nations* (Adam Smith, 1776, ch. viii). The reason for differing earnings across different occupations stems from the different costs and benefits associated with those occupations. The equilibrium wage will depend on the labor supply for a certain occupation and the labor demand from industry. The logical conclusion of this reasoning is that an individual's wage will depend on his/her marginal product of labor in a competitive market economy.

Education can be therefore viewed as an individual's investment in his/her future earnings potential. Optimizing individuals will not pay for increased education unless the cost of education incurred is balanced by future higher earnings. Education should increase the individual's marginal product of labor and consequently, an employer will be

willing to pay a higher wage. Therefore, one would expect the correlation between an individual's wage and their educational level to be positive. In this study *experience* is used as a proxy for the inverse of *education*[10].

Similarly, one would expect the coefficient of age to be positive: as one ages it is generally accepted that one acquires additional human capital through the acquisition of experience. This is true to a point at which human capital in the form of experience no longer shows a positive wage effect and could possibly show a negative effect. This reflects the fact that the initial investments for education will most likely be paid by younger workers in the form of lower wages. However, as a worker ages he/she will be less likely to incur the cost of training and education to combat the depreciation of his/her human capital, due to the fact that there are fewer working years to recover those costs in the form of higher wages. This effect could be captured by a age-squared explanatory variable. One might add that as the average age of U.S. workers increases as the baby-boomers head into their fifties and sixties the supply of young labor to work entry-level supermarket jobs will be decline. Correspondingly, the average age of grocery store workers should be increasing with age (see Graph 3).

[10] See section on Data for a more complete description of explanatory variables. In this case experience = age - years of formal education - 6. This is similar to other log wage regressions (Perloff and Sickles, 1987).



Graph 3: Age - vs- Education for Supermarket Employees by year. Source NBER, 1995.

In fact, examining a crude graph of mean age and education over the ten year sample, it can be seen that although the mean age is tending upwards, the mean level of education in the supermarket industry has remained at a high school diploma. This can be interpretted in many ways. If one considers that the mean age and unionization rates are significantly lower for part-time employees as compared to full-time collegues, the age and education results could indicate that there is a low turnover rate in the industry, amoungst full-time union employees, but that there is evidence supporting the claim that supermarkets cannot attract more experienced employees because of high non-wage benefits and low hourly wages for part-time positions.

Model Specification

This study of factors explaining wages in the retail food industry is based upon a reduced-form expression of the standard wage determination equation:

$$(1)\quad \log wage_i = \beta_0 + \beta_1 X_i + \beta_2 U_i + \varepsilon_i$$

where X_i is a vector of explanatory demographic, industrial, and human capital variables for that worker, U_i is a dummy variable representing union membership, and ε_i is an unobservable disturbance term. This is also known as a statistical earnings function.

Ordinary Least Squares (OLS) has been used extensively in the regression of wages against independent variables (as in the statistical earnings function). The majority of literature using this econometric model has been used to determine the union membership wage premium across industries and demographic groups. An important early study cited in wage determination literature is Lewis (1963) who examined competitive and monopoly unionism. Lewis found that the union wage premium was approximately 10-15% in the period 1957-1958. Subsequent studies using more extensive data sets estimated the union premium to be much higher, as high as 30% in some studies (Fuchs, 1968). These later estimates seemed to be too high to account for the relatively low observed union organization.

Ashenfelter and Johnson (1972) began to examine the determination of wage levels considering level of labor quality and extent of unionization as endogenous to the OLS regression model, i.e. that labor quality (or education) is an employer choice

variable and therefore endogenous to a system determining relative wages and the degree of unionization. The implication is that previous wage determination estimations were flawed due to endogeneity selection bias in the wage equation. This problem will cause independent regressors to be correlated with the error term, which will lead to inconsistent estimates. To study this question they estimated the bias that might result from endogeneity problems and then estimated model coefficients under a two-stage least squares specification. They found in examining a data source of manufacturing industries that under OLS estimation the highly significant union coefficient was 0.46 , but was 0.19 under the Two Stage Least Squares (2SLS) model. In addition the marginal effect of education also dropped from 18% to 12%[11].

However, Freeman and Medoff (1981, 1984) examined the question of union wage premiums quite extensively. In their 1981 article they studied the percentage organized as an independent variable affecting wages. They argued for an econometric specification similar to the statistical earnings function estimated using OLS. They maintained that a certain degree of co-determination of economic variables always exists, and that there was no basis for the theory that wage levels should be positively associated with union organization. Therefore, they state the omitted industry variable bias due to partial correlation with the percentage organized and wages need not make OLS an unsatisfactory estimation procedure. It is their conclusion that a wage equation can be

[11] Under the LogWage - statistical earnings function; coefficient estimates that are less than .25 can be viewed as percentage point changes in logWage. This is due to the fact that for r sufficiently small: ln (1+r) = r.

estimated by traditional OLS given a sufficiently large set of industry controls in the union and nonunion sectors. This is because, a priori, one cannot readily determine the possible bias direction. Longitudinal studies of the union wage effect show smaller differentials than cross-sectional studies. Cross sectional effects due to unobserved personal characteristics can increase bias. Correction by adding inverse Mills' ratios or fitting a system of equations including a unionization equation does not appear to yield useful results for analysis of union effects (see Freeman and Medoff, 1981). They further argue that estimates resulting from simultaneous equation and sample selection models are not stable with respect to small sample changes. The OLS parameter estimates, ignoring the endogeneity issue, are much less sensitive to changes in specification and data sources; and therefore are preferable to the other models which are unsatisfactory.

Addressing the Issue of Endogeneity

Schmidt and Strauss (1975) corrected for union endogeneity in wage determination using individual observations in a similar fashion to Ashenfelter and Johnson (1972). Their innovation in econometric technique to account for continuous wage variable and the binary membership observation, involved a mixed logit approach. The first equation in their model estimated the membership decision by standard logit specification followed by maximum likelihood estimation to determine the effect of membership on wage. They found that in the wage equation, the union membership coefficient was insignificantly positive and the wage coefficient on the probability of

union membership (logit equation) was significantly positive; indicating that higher earnings make one more likely to join a union rather than the reverse. This result suggests that the causality between wages and union status may be not be correctly supported by an OLS specification with an exogenous union variable.

Kahn (1977) furthered the examination of endogenous explanatory variables. He maintained that to account for union endogeneity a system of simultaneous equations can be used. He argued that the long run impact of unions on wages, quit rates, education and skill level resulted from short term direct results (wage increases collectively bargained). This implies that that labor costs have risen, all else equal, while capital costs have remained constant. The firm has then substituted capital for labor. Following directly from human capital theory, increased mechanization has led to the demand for workers with higher skill levels, which in turn has led to a demand for workers with a high level of education. These movements have in turn theoretically caused further wage increases to compensate increased skill levels (increases in training costs) and to retain those skilled employees higher wages have been offered. This cumulative rise in wage equilibrium level was the result of the long run effect of unionization on wages, quits, education, skill. Kahn found that firms have more incentive (due to unions) than otherwise to offer incentives of an internal labor market: higher wages, job security, good working conditions, and chances for advancement. In addition unions were found to lead

to higher wages; although higher wages may attract unionization which may have caused an existing wage differential to become wider.

Lee (1978) and Heckman (1976) each published articles using the "inverse Mills' ratio" to correct for self-selection endogeneity bias in estimation procedures. These articles developed the method of splitting the specification equation into two categories (union sector and nonunion sector) and separately estimating these equations using the inverse Mills' ratio (IMR) to correct for the missing individuals from the omitted sector in each equation.

This method is a common strategy for dealing with union endogeneity due to self-selection bias. It involves estimating the omitted variable and then applying ordinary least squares techniques to estimate the wage equation including the estimated omitted variable as a regressor. In this study the omitted variable is the self-selection decision of whether or not to become a union member.

The IMR method involves estimating the following equations in the second stage following a first stage probit estimation of the probability of union membership.

$$(2) \qquad \ln W_{ui} = \beta_0 + \beta_1 X_{ui} + \beta_2 \{-\frac{\phi(X_i \hat{P})}{\Phi(X_i \hat{P})}\} + \eta_{ui}$$

$$(3) \qquad \ln W_{ni} = \beta_0 + \beta_1 X_{ni} + \beta_2 \{\frac{\phi(X_i \hat{P})}{1-\Phi(X_i \hat{P})}\} + \eta_{ni}$$

In (2) and (3), $\Phi(*)$ and $\phi(*)$ are the cumulative distribution function and density function, respectively, of a standard normal variable. $\hat{P}$ is the probability estimate from the first stage probit. The inverse Mills' ratios are given by $\phi(*) / \Phi(*)$. Consistent estimates require normally distributed errors. The null hypothesis for no endogeneity is to estimate the model and test whether the β_2 terms are significantly different from zero. If they are different from zero in their respective sectors (union and nonunion)[12] one would reject the null (i.e. indication of endogeneity).

Duncan and Leigh (1985) developed a relatively new approach to wage determination specification. This entailed using an "instrumental variables" (IV) procedure for estimating the parameters of the standard wage equations. They presented this alternative to the IMR method as one that did not rely on distributional assumptions. The drawback being that the process generating the union and nonunion error terms must be assumed to be the same. Their study using the 1971 National Longitudinal Survey sample of white men (between the ages of 50 to 64 years old) and the IV procedure resulted in estimates that better indentify union endogeneity in a Hausman endogeneity test than do IMR estimates.

The first step in this version of instrumental variables is to estimate the probability that individual i is a union member , by regressing the observational variable *union* (0 or

[12] See Heckman (1976) and Lee (1978) for a further discussion of the Inverse Mills Ratio technique to correct for endogeneity bias.

1) on the explanatory variables. This probability will then be interacted with all of the independent regressor variables in the second equation. The second equation is an OLS regression using all of the original OLS variables, but interacted with the probability of being a union member if union member or the probability of not being a union member if nonunion member.

The model can be represented by three equations[13]:

(4) $q_i = X_i\alpha + \varepsilon_i$

(5) $\ln W_{ui} = X_{ui}\beta_u + \eta_{ui}$

(6) $\ln W_{ni} = X_{ni}\beta_n + \eta_{ni}$

The X_i's are exogenous regressors; $\ln W_{ui}$ and $\ln W_{ni}$ are the log of wages (self-reported in the CPS) in the union and nonunion sectors, q_i is unobserved utility gain from union membership. Union status is determined by the equation:

(7) $U_i =$ 1 if $q_i > 0$ (meaning the *i*th worker is in a union),

0 if $q_i \leq 0$ (meaning the worker is in the nonunion sector).

To estimate wage equations (5) and (6) the probability of U_i being either 1 or 0 must be calculated. This probability is then interacted with the regressors in equations (5) and (6); which are combined to estimate the log wage equation for the entire sample:

(8) $\ln W_i = U_i X_i\beta_u + (1-U_i)X_i\beta_n + [U_i\eta_{ui} + (1-U_i)\eta_{ni}]$

[13] This model is fully described in Duncan and Leigh (1980).

A logit estimation of equation (4) yields the probability of being a union member. The instruments created for the Duncan and Leigh instrumental variables method are generated from the interaction of this probability with the union member characteristics or one minus the probability multiplied by the nonmember characteristics. The error term in equation (8), $v_i = U_i\eta_{ui} + (1-U_i)\eta_{ni}$ is associated with several assumptions that allow consistent and unbiased OLS estimation:

$$Cov[U_i, \eta_{ui}] = 0, Cov[U_i, \eta_{ni}] = 0, Var(\eta_{ui}) = Var(\eta_{ni}).$$

Robinson (1989) shows that in general the assumptions of zero covariances in general will not hold. Duncan and Leigh, however, assume that:

$g(\varepsilon_i, \eta_{ui}) \equiv g(\varepsilon_i, \eta_{ni})$, where g(*) represents the joint density function.

This assumption of equal joint densities implies that $E[V_i] = 0$, which makes their instrumental variable approach consistent. If the joint densities, g(*), are not identical the IV estimates will not be consistent.

This method was used by Ashraf (1992) to estimate a standard wage determination model accounting for union endogeneity. In this study using data from the 1977-1981 *Panel Study of Income Dynamics*, Ashraf estimated the wage equation using OLS, the IMR method, and the IV method. All the estimates were consistent across the three methods except for the coefficient marginal effect sign on workers with a college degree. In addition it appeared as though the IV method yielded estimates that were more

closely related to OLS estimates than the IMR method. Ashraf concluded that his study confirmed the consistency of estimates using the IV method and that due to having less restrictions than the IMR method that it was a superior estimation procedure.

Perloff and Sickles (1987) employed a full-information maximum likelihood (FIML) procedure to estimate three endogenous variables: union wages, hours and earnings differentials. This method is similar to that of the FIML estimation technique used in this paper. They allowed all the exogenous variables to interact with the endogenous union dummy variable to examine the union wage premium across demographic groups. Included in their sample was both employed and unemployed construction workers (also from the May CPS files from 1973 - 1975).

They maintained that the consistent two-step estimators used in past studies to model wage determination and endogenous regressors were rather unstable because of the severe collinearity between the regressors, such as the percentage unionized in a local labor market and the instruments for the right hand side endogenous union dummy. To circumvent this problem a FIML estimation was used. The FIML estimated union wage premiums using the same data had the expected sign, but were larger than previous studies using less efficient estimation techniques. They expected to find that in the nonunion sector the relation between wages or hours and demographics was determined competitively; but were largely determined via the collective bargaining process in the union sector.

Their model consisted of three equations: a probit estimation of the union status decision; a tobit estimation of the hours worked; and a full-information maximum likelihood equation for the log wage equation. They found that is was necessary to look at earnings not just wage, due to the fact that there is a negative union effect on hours worked by union members. Another difference was that previous OLS studies of CPS data that treated union status as exogenous, resulted in lower union markups than if using FIML. The higher FIML markup was due to the negative covariance between union status and log wages. This negative covariance indicated all else equal nonunion craftsmen had higher levels of ability than did union members (contrary to what might be expected in the supermarket industry). Perloff and Sickles theorized that this may have resulted from using highly skilled key workers in the nonunion sector who had supervisory skills not captured by the other explanatory exogenous variables.

A recent examination of CPS wage data was conducted by Budd and Na (1994), who used the 1983 - 1991 CPS outgoing files to study the union wage premium of members and covered nonmembers. They began by initially using the general OLS method of estimation. However, as they pointed out this method is inappropriate if union membership is correlated with the error term stemming from employee behavior or measurement error. To address probable union endogeneity stemming from possible self-selection or measurement error they utilized three methods of empirical estimation: (1) inverse Mills ratio selection correction (this employs a probit model to estimate union

membership status); (2) Duncan and Leigh's (1985) instrumental variables estimator (using logit estimation of union membership status); (3) the possible correlation is modeled assuming that it is resultant from an individual fixed effect where *t* indexes time using longitudinal data[14]. The results of these different techniques yielded different coefficients: 2SLS ≈ 12% union premium; cross-sectional tests for exogeneity reject the null and resulted in estimates ≈ 2 - 7 % higher than in the OLS case. The longitudinal estimates were significantly smaller by ≈ 2 - 5 %[15].

Examining Percentage Unionized

Freeman and Medoff (1981, 1984) examined the question of union wage premiums quite extensively. In their 1981 article they studied the percentage organized as an independent variable affecting wages. They argued for an econometric specification similar to the statistical earnings function that is estimated using OLS. They maintained that a certain degree of co-determination of economic variables always exists, and that there was no basis for the theory that wage levels should be positively associated with union organization. Unionization is likely to occur in high rent industries, therefore creating the potential for collective wage increases without reduction in employment, although they regard management as capable of offsetting this effect in these industries.

---

[14] With the exception of the longitudinal correction method, this paper follows closely the format of the Budd and Na study.

[15] The longitudinal estimator is reported to be susceptible to measurement error problems, which makes this approach less desirable when using CPS data over time due to the changing nature of this survey. See Documents for Use with the NBER CPS Labor Extracts for a description of the survey changes.

They used the CPS May outgoing files from 1973 - 1975 for individuals and the Expenditures for Employee Compensation Surveys (EEC) from the BLS (1968, 1970, 1972), and concluded that there is either no spill-over effect or that there existed a small positive effect on percentage unionized in local labor markets and compensation for nonunion workers, although the compensation differentiation would be expected to grow as union organization increases.

Belman and Voos (1993) examined the criticism of the empirical literature on union coverage effect on wage; that there is omitted variable bias. Their argument was that higher union organization within an industry yielding higher union wages, could be reflecting in part the incentives of organizing in that industry rather than the effects of the actual union percentage. This argument is supported by the fact that inter-occupational and inter-area studies show less effect between unionization and wage gap than have inter-industry studies. To examine this effect Belman and Voos used CPS outgoing rotation files for 1987 and 1989 for the supermarket and aerospace industries (local or regional labor markets versus national or international). They viewed OLS estimates to be consistent but biased downwards due to the aggregate error component. To estimate their wage determination equation they used three methods: OLS; selection corrected models, which follow the methodology of Lee (1978); and error components models following Mundlak's (1978) random effects model in correcting for city-level error

components. The important outcome in utilizing these three methods was that it made no difference to the results of the effect of union density on wage, although the selection correction method or error components altered the magnitude and statistical significance of some other coefficients. It is important to note however that they found that the inverse Mills ratio coefficients are statistically significant for both sectors in the supermarket industry but not for either sector in the aerospace industry. The reason for the self-selection being more significant in one industry than the other was not revealed. Importantly, in the supermarket industry Belman and Voos found that higher union percentages are associated with higher union wages and that higher union percentages are associated with lower nonunion wages (although insignificantly)[16].

[16] In their study, Belman and Voos omitted observations from CMSA's with less than 40 observations, similar to the Spill40 definition. They comment that their results only vary slightly if all individuals are included (page 371).

IV. METHODS AND DATA

In reviewing the previous wage determination and union effect studies, one finds several methods used to address the union endogeneity issue and the resulting bias of regression coefficients and error terms. It is agreed that the traditional OLS wage regression, treating union membership as an exogenous variable, is consistent with theory but biased due to the correlation between union membership and the error term stemming from employee behavior or measurement error. In the more recent literature, union endogeneity in traditional wage determination is an important question. There is agreement as to the endogeneity of union membership, but the form of endogeneity has yet to be adequately determined. Observed union status has been regarded as resulting from an individual's choice based on probable earnings as a member of a union and those received in the nonunion sector. Therefore, I propose to estimate the wage determination equation using several different estimation procedures following a preliminary OLS regression, in order to yield more robust estimates of variable coefficients. These other estimation procedures to correct for union endogeneity bias are: (1) Full Information Maximum Likelihood (FIML) employing inverse Mills ratios in the iterative maximization routine; and (2) the Duncan and Leigh Instrumental Variable (IV) method.

It should be noted that bias other than that of self-selection for union membership is not accounted for in these estimation procedures. Such bias could be caused by

omitting other parameters measuring individual's abilities for example. Many would argue that therefore, since abilities are generally agreed to be positively correlated with years of education, omitting the abilities measures from the econometric model will cause upward bias to the returns of schooling[17]. Such ability measures such as tenure would have been valuable to include in this study should they have been available.

Probit / Logit Estimations of Union Membership

An individual seeking employment in the supermarket industry may choose either to work in a unionized store or a nonunionized store. Once this decision is made, the individual faces the decision of whether or not to become a member. An individual in this sample either is a member or is not [Y=1 or Y=0][18]. A set of factors such as age, sex, education, percentage unionized, etc. is used to describe the vector, X , which is used to explain the individual's decision to join a union or not. Hence:

(9) Prob[Y=1] = F(X,b)

(10) Prob[Y=0] = 1 - F(X, b).

The set of parameters b reflect the impact of changes in X on the probability. In principle, any proper, continuous probability distribution defined over the real line will be

---

[17] Berndt (1991) argues that the differences in abilities does not appear to account for a sizable proportion of earning premiums amongst individuals with differing levels of education. See also Belman and Voos (1993) and Shaw (1984) for a more in-depth discussion of omitted variable bias.

[18] The possibility of Covered Nonmembers is discussed in Budd and Na (1994).

sufficient to produce consistent predictions. If one uses the Normal Distribution then probit modeling is indicated.

(11) Prob[Y=1] = $\int_{-\infty}^{\beta' X_i} \phi(t)dt$, which = $\Phi(\beta`x)$. The function $\Phi(.)$ indicates the standard normal distribution[19].

Data: The Current Population Survey

The Current Population Survey (CPS) is a monthly household survey of employment and labor markets. It is conducted by the U.S. Census Bureau on behalf of the U.S. Bureau of Labor Statistics (BLS). The survey data has been compiled on a NBER (National Bureau of Economic Research) CPS Labor Extracts: Annual Earnings File Extracts: 1979 - 1993 CD-Rom.

The survey provides information on the labor status of nonmilitary and non institutionalized participants. About 60,000 households are surveyed monthly, with a reference person from each household asked to report on his or her labor status and for those of other household members. The CD includes data for approximately 30,000 individuals aged 16 or greater, each month for 180 months. The time period covered is a calendar week.

---

[19] The question of which distribution to use, a Probit or a different model (Logit for example), often arises. The variance of the logistic function is similar to the normal except in the tails, which are heavier. There are practical reasons for choosing one or the other, but it is difficult to theoretically justify, but seems not to make much difference. See Greene, 1990.

The sample is designed to yield accurate measures of labor force participation by state. For each monthly sample, eight representative sub samples or rotation groups are included. Each household in a rotation group is interviewed a total of 8 times: four consecutive months of interviews, ignored for eight months, and then interviewed again for four months. If the occupant moves they are not followed, rather the new occupants are interviewed. Since 1979 outgoing participants have been questioned with respect to usual weekly earnings and usual weekly hours. These participants in outgoing rotation groups (i.e. in interview month 4 and month 16) are used to compile the BLS Annual Earnings File (Merged Outgoing Rotation Groups File). The Annual Earnings File is used to produce the CPS Labor Extracts CD-Rom. Each observation in the CPS extracts includes:

(1) Misc. Variables: household identification, outgoing month, year;

(2) Geography: state, metropolitan status, central city;

(3) Demographics: age, sex, race and ethnicity, marital status, education, veteran status, household relationship;

(4) Wages: earnings per hour, earning per week, usual hours;

(5) Employment: worker category, employment status, industry classification, part-time / full-time status, student;

(6) Union Variables: member, non member, covered by collective agreement.

For the sample, this study selected those workers who reported themselves as working in Grocery Stores (variable ind80 = 601). After creating separate data samples for each year between 1984 through 1993 inclusive the study merged the ten sample years. To avoid repetition of individuals both in individual year sample sets and also in the ten-year sample set, those observation that had identical household identification numbers and *minsamp*[20] values of four were eliminated (i.e. for repeated individuals from month 4 and month 16, only the later observation [*minsamp* = 8] were used).

After compiling a unique list of grocery store workers from 1983 - 1993, those observations were discarded with missing data in the major variable categories: *LogWage, sex, union, region, occupation*, etc. (see NBER CPS labor extracts for definition of terms). In addition, to avoid measurement error those observations listing wage as less than or equal to $1.00 per hour were discarded. This excludes those individuals who have mistakenly reported their earnings in dollars and not in cents as demanded by the survey. The resulting sample contains roughly 20,000 individual observations for the ten years 1984 -1993, for those workers in the grocery store industry. Several variables were then created for use in the regression analysis[21].

Union: This dummy variable was created to represent union membership or not (0=nonmember; 1=member).

---

[20] Minsamp refers to whether the individual was in the fourth or sixteenth month of the survey.
[21] See Also Table #2: Variable Summary Statistics.

Sex : This is just the standard dummy to represent being male or female (0=female; 1=male).

LogWage: This variable (*LogWage*) was created using the *earnhr* variable multiplied by the aforementioned CPI multiplier, with the base year being 1993.

Educ: The variable *educ* was created to reconcile the differences between survey year variables *grade92* and *gradeat*. The variable *educ* represents the amount of formal education an individual has received in years of schooling attended on a spectrum of 0 through 18 (0 = no formal education; 18 = advanced graduate degree).

Occupation: This variable was created to describe the class of job within the grocery store industry: 1 = manager, supervisor; 2 = sales, cashier; 3 = support staff (security, secretarial, etc.); 4 = skilled employees (butchers, truck drivers, etc.); and 5 = clerking. Following this classification, each category was made into a dummy (0-1) and denoted by *super, sales, support, skill,* and *stock.*

Region: A region variable was created to delineate four country regions: 1 = east; 2 = midwest; 3 = south; 4 = west. Following this classification, each

category was made into a dummy (0-1) and denoted by *East, Midw, South,* and *West.*

Full-time / Part-time: The variable *ftpt* is a full-time / part-time dummy (0 = less than 35 hours per week; 1 = more than or equal to 35 hours per week).

Race: The dummy *rac* was created: *rac* = 0 if not white; *rac* = 1 if white.

Experience: This variable was created to proxy labor market experience and tenure. This was generated using the common expression : experience = age - years of education - 6 (see Schmidt and Strauss 1976, Perloff and Sickles 1987). It should be noted that other measures of experience may more accurately reflect this human capital investment[22].

Percu: This is the variable representing the percentage of grocery store workers in a local labor market that are unionized. As organization in that market increases there should be a resultant decrease in the elasticity of demand for organized workers due to the decrease in the opportunity to substitute nonunion products for union products. Through collective bargaining union wages should be higher with increasing coverage. This could either induce wage increases in the nonunion sector due to increases in relative cost of

---

[22] Katheryn L. Shaw (1984) argues that the concept of occupational investment which takes into account that different occupations will show differences in human capital investments. However in the context of this study which examines across the eight categories, the experience proxy can be considered sufficient.

union labor; or it could cause a decrease in nonunion wages due to reduced employment in the union sector. By relating nonunion wages to percu, we provide direct estimates of the extent to which nonunion workers in organized markets receive higher or lower wages than nonunion in other areas. This variable reflects the union percentage in the supermarket industry for each observation's local labor market.

Several definitions of this variable were used for the three econometric estimation methods. For the *Full Sample*, *percu* was defined as the percentage of union members in each cmsarank[23] city. If however that city did not have 40 observations or if that particular individual did not live in one of the 252 cmsarank cities the state union percentage was assigned (using only those observations not falling into the cmsarank percentage category). For example, if there were not at least 40 workers in the supermarket industry from 1984 - 1993 in Kankakee, Illinois (cmsarank 252) then those observations that were from Kankakee would be pooled into the Illinois union percentage[24]. If, however, there happened to be at least 40 workers in the CPS sample (see NBER 1995) from Kankakee the *percu* variable would only consist of those observations and the state percentage would not include those observations. For *Spill40, percu* is defined at cmsarank percentage if there were 40 or more observations from that city, if there were not then the observation was

[23] This is a ranking of the most populated 252 U.S. cities.
[24] This is similar to the definition used by Perloff & Sickles (1987) where the percent unionized is the average unionization percent in a worker's standard metropolitan statistical area (SMSA) if in one of the 20 largest SMSAs or the average state percentage otherwise.

deleted from consideration. This is similar to the definition used in the Belman and Voos study. For *Spill20, percu* was defined as in *Full Sample* with the exception of assigning cmsarank percentage to individuals in cities with at least 20 observations over the 10 year sample period.

Right-to-Work: In many states as a condition of employment in many supermarkets an employee must be a member of the local UFCW. In other states it is illegal to make union membership a necessary condition for employment in a certain store. A dummy variable distinguishing those states that have Right-to-Work legislation was created (0 if right-to-work legislation in that state, 1 if no right-to-work legislation in that state).

States with Right to Work Legislation:

Washington D.C., North Dakota, South Dakota, Wyoming, Idaho, Nevada, Arizona, Utah, Texas, Nebraska, Kansas, Iowa, Arkansas, Louisiana, Mississippi, Alabama, Florida, Tennessee, Georgia, South Carolina, North Carolina, Virginia.

No Right-to-Work Legislation:

Minnesota, Wisconsin, Montana, Washington, Oregon, California, New Mexico, Colorado, Oklahoma, Missouri, Illinois, Ohio, Alaska, Hawaii, Michigan, Kentucky, Indiana, Maine, Vermont, New Hampshire, West Virginia, Maryland, New York, Massachusetts, Connecticut, Rhode Island, Pennsylvania, Delaware, New Jersey.

Indexing term: From the Bureau of Labor Statistics this study used the "All Items" values to index the CPS earnings observations, choosing base date = 1993 (see Appendix).

Variable	Definition	Obs	Mean	Std. Dev.
LogWage	Log of indexed wage	23400	6.504768	0.419914
lnwu	" for Union Indiv.	6304	6.831249	0.448335
lnwnu	" for Nonunion Indiv.	17096	6.384382	0.336809
age	Age in years	23400	29.62389	12.83043
age2	Age squared in yrs.	23400	1042.188	968.5579
exp	Experience in yrs.	23400	11.45201	13.04145
	= age - educ. - 6			
educ	Years of formal educ.	23400	12.17188	1.710902
sx	= 1 if male	23400	0.455556	0.498031
rac	= 1 if white	23400	0.80359	0.397291
ftpt	= 1 if full-time	23400	0.500983	0.50001
Super	= 1 if supervisor	23400	0.115556	0.319698
Sales	= 1 if sales	23400	0.404359	0.490778
Suppor	= 1 if support	23400	0.123333	0.328827
Skill	= 1 if skilled	23400	0.080043	0.271365
Stock	= 1 if clerk	23400	0.276709	0.447381
East	= 1 if East	23400	0.223462	0.416574
Midw	= 1 if Midwest	23400	0.252051	0.4342
South	= 1 if South	23400	0.312564	0.463548
West	= 1 if West	23400	0.211923	0.408679
yr1	1984	23400	0.123333	0.328827
yr2	1985	23400	0.130214	0.336546
yr3	1986	23400	0.095769	0.294281
yr4	1987	23400	0.079615	0.270703
yr5	1988	23400	0.079487	0.270503
yr6	1989	23400	0.083504	0.276649
yr7	1990	23400	0.088974	0.284713
yr8	1991	23400	0.093718	0.291442
yr9	1992	23400	0.097222	0.296267
yr10	1993	23400	0.128162	0.334278
percu	Percentage of local	23400	0.261255	0.196292
	labor market unionized			
union	= 1 if union	23400	0.269402	0.443659
RTW	= 1 if right-to-work	23400	0.601197	0.489663
	state			
Full Sample:	sample with percentage unionized determined by 40+ observations in a CMSA; state percentage if not.			
Spill40:	sample with percentage unionized determined by 40+ observations in a CMSA; dropped if not.			
Spill20:	sample with percentage unionized determined by 20+ observations in a CMSA; state percentage if not.			

Table 3.Description of the Variables, Observation Number, Means, and Standard Deviations

V. RESULTS

When comparing results between the three estimation procedures it is important to distinguish the sample that was used. Both *Full Sample* and *Spill20* have a total of 23400 individual observations. *Spill40* only has 8719 observations and should be regarded as those supermarket workers living in a large urban center. Therefore, it is interesting to examine the various results in the context of how does urban wage structure differ from the total.

Consistent across methods and samples are the positive effect of age and negative effect of age squared, which is also consistent with theory. The negative effect of increasing experience has been determined to represent the inverse of education. As educational levels increase, all else equal, the lower the experience. A positive effect of education on wages is therefore proxied by the negative experience effect. In future studies drawing on this paper, this variable should be changed to directly represent education. For all significant results, being male positively affects wage, which is consistent will previous studies. It should be noted that a self-selection bias for work force participation may exist for women who choose not to work. Therefore, these results should be interpreted as being in reference to workers in the supermarket industry and not the population as a whole.

In all of the samples across the three methods, being a supervisor, a support staff, or in a skilled job category positively affects wage when compared to being a clerk. In addition, full-time employees significantly earn more than their part-time cohorts.

Preliminary OLS

In this framework I assume that the disturbance term is normally distributed with mean zero and constant variance. I therefore use the log wage as the dependent variable rather than just wage. This remedies the problem of having only positive wage observations as well as correcting for the large distribution of low wages and high wages in the sample.

Referring to appendix Table A.i-vi, we can examine the results of the preliminary OLS regression of LogWage against explanatory variables. For the regression, the variable for *stock*; the region variable for the *South*; and the dummy variable for 1993, *yr10*; have been omitted and should be regarded as the base individual, i.e. when examining the effect of being in a supervisory position in the supermarket industry the coefficient of 0.218753 reflects a comparison (all else equal) with an individual in a clerking position.

As expected the *union* variable is strongly positive in significance and effect: 0.317166 with t-value = 66.134. The results from the OLS regressions by year (appendix Table A.iv) the decrease in union influence is mirrored in the declining effect of the dummy *union* over the years 1984-1990, but its positive affect on wages begins to

increase after 1990. Interesting to note is the positive *percu* variable coefficient of .01177 with t=15.344. This is of course what has been noted in previous studies, and what would be expected from a preliminary OLS regression. It can be noted when one examines appendix Table A.v-vi that *percu* has a much stronger influence in the union sector than in the nonunion sector.

The *experience* coefficient is negative in all three samples, and is highly significant[25]. Other notable results include a negative coefficient on the regional *Midwest* variable of -.02801 indicating that, all else equal, workers in the *South* report higher wages than those in the *Midwest*. This result is also significant with a t-value of -5.301.

We can see from the dummy year variables that in comparison to 1993, real wages were highest in 1984 and declined in each successive year until 1991, at which point they began to rise again. It is interesting to note that both in 1989 and 1991 there were increases in the Federal Minimum wage: to $3.75 per hour in 1989 and to $4.25 in 1990 (in constant 1995 dollars). In addition there is a positive effect on LogWage all else constant if the individual is white, male, middle aged, a full-time employee, and living in the West.

There are no significant differences in preliminary OLS regressions using the alternate samples with *percu* defined differently. In *Spill40* the dummy for 1985 shows a

[25] Experience serves as a proxy for the inverse of education, which is a large reason for this negative result. As education increases experience will decrease.

negative coefficient indicating that all else equal wages were lower in 1985 than in 1993, but the t-value is not significant at the alpha =.05 level.

Full Information Maximum Likelihood

Using the statistical program Stata, the Heckman procedure will be used to first estimate via a Probit equation union membership dependent upon age, age squared, exp, educ, race, sex, full-time / part-time status, job classification, right-to-work legislation by state, year, mean wage premium (union-nonunion) and percentage unionized by locale[26]. Stata uses a two-step procedure for estimating a selection model including a maximum-likelihood probit model and then a regression using inverse Mills' ratios similar to the method employed by Perloff and Sickles (1987). They maintain that other consistent two-step estimators are unstable due to the collinearity between regressors (such as percentage unionized in a local labor market) and the instruments for the right-hand-side endogenous union dummy. To avoid this instability Perloff and Sickles employ a full-information maximum likelihood approach leading to larger sample mean wage union markups.

This model assumes an observation is observed if:

$X_1\beta_1 + u_1 > 0$; where u_1 has a standard normal distribution. In addition there is a regression equation: $y = X_2\beta_2 + \sigma u_2$ where u_2 is potentially rho[27] correlated with u_1.

---

[26] See Stata Reference Manual (Release 4). Volume 2.page 438-445.
[27] Rho denotes the correlation between u1 and u2.

When rho $\neq$ 0 standard regression results are biased. The Heckman method provides consistent, asymptotically efficient estimates for such models. Stata uses inverse Mills' ratio estimates for starting values. In the iterative maximization routine, derivatives of the likelihood are computed numerically to obtain the maximum-likelihood results which are presented in the appendix.

When checking the rho correlation results, it can be seen for the union sector in *Full Sample* and *Spill40* that rho is significantly different from zero; however in the nonunion sectors and in the union sector in *Spill20* the rho correlation is not significantly different from zero.

As with the OLS results, *experience* negatively influences real wages. This is significant and consistent over the three samples. Once again experience can be shown to be a proxy for the inverse of education, and positive educational influence are to be expected. Race shows no significant difference from zero in either sector over the three samples. The regional variables once again show inconsistent results as compared to preliminary OLS analysis. In all three samples *Midwest* shows negative influence in both sectors; *West* shows positive influence in both sectors; and *East* shows positive influence in the nonunion sector and negative influence in the union sector.

The *percu* variable shows insignificant negative influence in *Full Sample*; significant negative influence in S*pill40*; significant positive influence in *Spill20.* Its resulting coefficients in the nonunion sector are completely opposite to the union sector:

positive-significant in *Full Sample*; positive-insignificant in *Spill40*; negative-insignificant in *Spill20.*

Other interesting results are that all job categories show positive influence as compared to stock positions in the union sector, but sales positions (e.g. cashiers) in the nonunion sector show negative influence as compared to stock positions.

Instrumental Variable Method

For the instrumental variable logit the variable RTW (right-to-work) is used to identify the endogenous union variable when calculating the probability that *i*th individual will choose to join a union, given his/her characteristics. Theoretically one would expect that RTW will be correlated with the union dichotomous decision to join a union, but will not be correlated with the error term in the log wage equation. If this is true then we can use RTW as a identifying variable for union.

In comparing the coefficients resulting from the IV-method it is useful to keep in mind the definition (hence data sample) of *percu* that is being used. Notable differences between the three samples (see appendix C.i-iii) are that experience is now positive in both the union and nonunion sector taking into account the endogeneity of union membership as opposed to being negative in the preliminary OLS and in the FIML estimation. This is consistent in both sectors across all three samples (although not significant in the union sector for *Spill40*).

Across all three samples accounting for union endogeneity using the instrumental variable approach, being white has a negative effect on *LogWage* (not significantly different from zero in the union sector). Being male has a consistently and significantly positive effect on *LogWage* across all samples. The regional variables in the union sector reflect that *South* reports the highest wages (not significant in *Full Sample*). The regional variables in the nonunion sector are consistent with OLS results being positive for the *East* and *West* and negative in the *Midwest*.

Corresponding to the regional variable results in the union sector are the results for the coefficients on the *percu* variable. This variable shows significant negative percentage unionized in the local market influence on *LogWage* for the union and nonunion sectors in all three samples. This indicates surprisingly that all else equal both union and nonunion individuals report higher wages in areas with a lower percentage of union workers in the supermarket industry. This obviously is contrary to the theory that the local labor market union percentage positively affects wages for both union and nonunion workers. It should be noted that for all of the samples, the constant logWage intercept is the same for union and nonunion sectors and is significantly higher than in either sector or sample set when estimating via OLS or FIML. This method of estimating the entire sample but having only one constant may be one reason explaining the inconsistencies with the other methods.

The dummies for 1984 through 1993 reflect the general trend as revealed by preliminary OLS that wages have fallen through the 1980's and have rebounded somewhat by 1993, although not quite back to original levels in real terms.

VI. SUMMARY AND CONCLUSIONS

This examination of supermarket wages relies on the self-reported usual earnings as collected in the CPS. This may introduce measurement bias[28]. In addition, the estimations of wage used for this study omit the non-wage benefits as part of the estimation procedure, which over the ten year sample period could mask many employer/employee relations, especially when examining the union effects on wage. As reported by the FMI in 1995, an average store pays its employees about 25 cents in fringe benefits out of every dollar spent on salary. In addition it remains despite technological and structural changes in supermarkets that food marketing remains highly labor intensive as compared to other national industries[29]. In addition, the union effect on fringe benefits has been found to be much larger than the union wage effect (see Freeman & Medoff, 1984). An example of this would be the importance that union-negotiated pension funds have become in capital markets.

Also not examined as a determinant of wage are the hidden employer effects, which could cause omitted variable biases. In fact Freeman and Medoff (1984) cite informally that when questioned about treatment of union and nonunion employees in a

---

[28] Other studies that have used the Current Population Survey as a data source for calculating similar estimates of wage determinants have reported the estimated percentage gains in wages due to collective bargaining to be lower than estimates obtained using other sources of data (see Freeman and Medoff, 1984, pg.46).

[29] In 1994 the average gross margin of supermarkets was between 23.4% and 25.8% (gross margin being defined as the difference between the supermarket's buy and sell price). Of that percentage roughly fifty percent as reported by the FMI's 1994 *Operations Review*, is spent on store labor costs.

firm, some firms admit that they change the pay of their nonunion workers at different times to disguise the union influence (see Freeman and Medoff, 1984. page 152).

However, despite the drawback of incomplete data sets, these estimation methods study the marginal effects of a variety of explanatory variables on wages in the supermarket industry from 1984 - 1993. They can reveal a wealth of information regarding employee wages in this industry and also about the economy as a whole (in so far as the supermarket industry is major component of the economy as a whole).

As one would expect from the US economic trends during the eighties, the real wage for supermarket workers fell significantly during this period. There has been a rebound in the value of real wages, however in 1992 and 1993. Correspondingly, UFCW membership as a percentage of workers in the supermarket industry follows the trend of real wages (or vice versa) across time, and we can see that membership increased in the early nineties.

One would expect to see that the movement towards technology to increase productivity would probably increase the premium for skilled labor. This hypothesis appears to be true in that every *occupation* variable positively affects earnings as compared to general clerking labor. However, there are negative experience coefficients. This could be explained by collinearity in the specification (i.e. between *age* and *experience*), or perhaps because there is little tenure effect (or high turnover) in the

supermarket industry. In addition it may be that graduating high school seniors now have sufficient technological knowledge for employment in the supermarket industry.

The federal increases in minimum wage in 1989 and 1990 do not seem to show up in any dummy year coefficient. In fact mean real wage in the supermarket industry for employed workers seems to have continued to decrease. It may be that without the federal increases that the decrease would have been more dramatic.

The values for the *percu* coefficients are certainly interesting. The significant negative effects on LogWage in the nonunion sector could be due to the market response involving the fact that high union wages and possibly higher product prices could tend to reduce the number of workers in the union sector and thereby increase the supply of labor in the nonunion sector (Berndt, 1991, p. 177). This market effect would tend to support the results that imply that highly unionized local labor markets for the supermarket industry would all else equal adversely affect the wages for nonunion workers in that market. The negative effects on LogWage in the union sector could perhaps indicate that there are certain thresholds that characterize a unions collective bargaining strength in the supermarket industry. It could be that high union percentages in the supermarket industry will tend to occur, all else equal, in areas where wages are, for other reasons lower, than elsewhere.

It is interesting to compare these results with the theoretical implications discussed in the Chapter II. Firstly, there does not appear to be a positive spill-over

effect. In other words, for nonunion workers the marginal effect of working in a more highly unionized area, all else equal, does not appear to be positive. Secondly; it does not appear in examining both supermarket data and that of other industries that unions raise non-wage benefits to such an extent that supermarket stores cannot afford to compete for qualified employees against higher hourly-wage / lower nonwage-benefit employers such as Wal Mart (as had been hypothesized by industry leaders). This was determined by running separate estimations on data from other industries. An industry with similar entry-level skill requirements, such as department store retailing, that has lower union membership, was found to show lower mean wage values than the supermarket industry, all else equal. It may be that the data from such retailers as Wal Mart may be hard to distinguish from the data sets of the country as a whole and that these effects may be masked by other industry noise. The answers to both these questions are important enough to warrant further inquiry of the supermarket wage structure. Due to limitations in data sources (no nonwage benefit data) and to estimation techniques (possible omitted variable biases) the conclusions drawn in this study are by no means complete, but are solely based upon the data analysis subject to the aforementioned constraints.

One reason for the seemingly ambiguous results could be that this study examined wages over a extended period: ten years. During this period it is evident both that the number of employees in the supermarket industry increased and also that the percentage of those workers in the UFCW decreased. These two factors could have affected the other

regression coefficients longitudinally. Therefore, it is important to further study the longitudinal effects on the coefficient values in the various models used to analyze the wage data to have a more complete picture of what is actually happening (see appendix Table A.iii.). This effect could be described as a longitudinal disturbance which left undetected could cause bias in the estimation methods and corresponding coefficients.

It should be noted that the econometric models used to estimate the wage equations are limited by the assumption of normally distributed errors. There is reason to believe that the assumption of homoskedastic errors is not valid. When a Cook-Weisberg test for homoskedasticity is performed on the FULL Sample OLS model, it is found that the null hypothesis of homoskedasticity is rejected. Of course if the errors are normally distributed with constant variance the FIML and IV methods will yield consistent estimations. If this strong assumption is not made however, there remains another more general estimation technique that yields consistent results without the assumption of normality. This is known as the General Method of Moments (GMM), and consists of solving simultaneously a system of equations that utilize the Central Limit Theorem and The Law of Large Numbers to assure asymptotically unbiased and consistent estimates. Therefore, a future avenue for further analysis is the verification of these estimates by GMM estimation.

VII. REFERENCES

Ashenfelter, Orley, and George E. Johnson. "Unionism, Relative Wages, and Labor Quality in U.S. Manufacturing Industries," *International Economic Review*, 13.5 (1972), 488-507.

Ashraf, Javed. "Union Wage Premiums in an Instrumental Variables Framework", *Journal of Labor Research*, XIII.2 , (Spring 1992), 230-235.

Belman, Dale L. and Paula B. Voos. "Wage Effects of Increased Union Coverage: Methodological Considerations and New Evidence". *Industrial and Labor Relations Review*, 46.2. (1993), 368 - 380.

Budd, John W. and In-Gang Na. "The Union Membership Wage Premium For Employees Covered By Collective Bargaining Agreements," Industrial Relations Center: Working Paper Series No. 94-09. University of Minnesota. 1994.

Duncan, Gregory M. and Duane E. Leigh. "Endogeneity of Union Status: An Empirical Test". *Journal of Labor Economics*, vol. 3 no.3, (1985). 385 - 402.

Food Marketing Institute. "The Food Marketing Industry Speaks: 1995." Food Marketing Institute publication, Washington, D.C., 1995.

Food Marketing Institute. "1995 Supermarket Facts." Food Marketing Institute publication, Washington, D.C., 1995.

Freeman, Richard B. and Medoff, J. L. "The Impact of the Percentage Organized on Union and Nonunion Wages." *The Review of Economics and Statistics.* 63, (1981). 561 - 572.

Freeman, Richard B. and Medoff, J. L. What Do Unions Do?. Basic Books, Inc., Publishers. New York. 1984.

Fuchs, Victor R., The Service Economy. New York: National Bureau of Economic Research, 1968.

Greene, William H. Econometric Analysis. New York University. Macmillan Publishing. 1990.

Heckman, J., "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement,* 5, (1976), 475-492.

Kahn, Lawrence M., "Union Impact: A Reduced Form Approach." *The Review of Economics and Statistics.* 59, 1977.

Kinsey, Jean and Ben Senauer. "Consumer Trends and Changing Food Retailing Formats," AAEA meeting paper: May 1996.

Kinsey, Jean and Ben Senauer. "Food Marketing in an Electronic Age: Implications For Agricultural Producers," Working Paper 96-2, University of Minnesota, June 1996.

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

Lee, Lung-Fei. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables," *International Economic Review.* 19.2, (1978), 415-433.

Mundlak, Yair. "On Pooling of Time Series and Cross Section Data," *Econometrica,* 46.1 (1978), 69-85.

National Bureau of Economic Research. Documents for use with the NBER CPS Labor Extracts: Annual Earnings File Extracts: 1979 - 1993 CD-ROM. Harvard University. 1995.

Nelson, Charles T. "Levels of and Changes in the Distribution of U.S. Income." The Changing Distribution of Income in an Open U.S. Economy. J.H. Bergstrand et al, Notre Dame, Indiana. 1994. ch. 2.

Perloff, Jeffrey M. and Robin C. Sickles. "Union Wage, Hours, and Earnings Differentials in the Construction Industry". *Journal of Labor Economics*, 5.2, (1987), 174-210.

Robinson, Chris. "Union Endogeneity and Self-Selection". *Journal of Labor Economics*, 7.1, (1989),106-112.

Schmidt, Peter and Robert P. Strauss. "The Effect of Unions on Earnings on Unions: A Mixed Logit Approach". *International Economic Review*, 17.1, (1976), 204-212.

Shaw, Katheryn L. "A Formulation of the Earnings Function Using the Concept of Occupational Investment," *The Journal of Human Resources*, XIX.3 (1984), 319-340.

Stata. Reference Manual 1-2-3 (Release 4). Stata Press, College Station, Texas, 1985-1995.

Thurow, Lester. "Why Their World Might Crumble." *The New York Times Magazine.* November 19, 1995. 78 - 79.

U.S. Census Bureau. @ *http://www. census. gov/hhes/income/histinc/ index. htm/.*, 1997.

U.S. Department of Labor. BLS Handbook of Methods. Washington, D.C.: U.S. Government Printing Office, BLS Bulletin 2414. 1992.

U.S. Department of Labor. BLS @ *http://www.stats.bls.gov.* 1997.

Voos, Paula B. and Lawrence R. Mishel. "The Union Impact on Profits: Evidence from the Supermarket Industry," University of Wisconsin - Madison. Working Paper-88, October 1985.

Walsh, John P. Supermarkets Transformed: Understanding Organizational and Technical Innovations. Rutgers University Press, New Jersey. 1993.

APPENDIX

No. 761. Consumer Price Indexes, by Major Groups[1982-84=100. Represents annual averages of monthly figures. Reflects buying patterns of all urban consumers. Minus sign (-) indicates decrease. See text, section 15. See Historical Statistics, Colonial Times to 1970, series E 135-173 for similar data]

YEAR	All Items	Medical Care	Energy	Food	Shelter	Apparel & Upkeep	Trans-portion
1960	29.6	22.4	30.0	25.2	45.7	29.8	22.3
1961	29.9	22.5	30.4	25.4	46.1	30.1	22.9
1962	30.2	22.6	30.6	25.8	46.3	30.8	23.5
1963	30.6	22.6	31.1	26.1	46.9	30.9	24.1
1964	31.0	22.5	31.5	26.5	47.3	31.4	24.6
1965	31.5	22.9	32.2	27.0	47.8	31.9	25.2
1966	32.4	23.3	33.8	27.8	49.0	32.3	26.3
1967	33.4	23.8	34.1	28.8	51.0	33.3	28.2
1968	34.8	24.2	35.3	30.1	53.7	34.3	29.9
1969	36.7	24.8	37.1	32.6	56.8	35.7	31.9
1970	38.8	25.5	39.2	35.5	59.2	37.5	34.0
1971	40.5	26.5	40.4	37.0	61.1	39.5	36.1
1972	41.8	27.2	42.1	38.7	62.3	39.9	37.3
1973	44.4	29.4	48.2	40.5	64.6	41.2	38.8
1974	49.3	38.1	55.1	44.4	69.4	45.8	42.4
1975	53.8	42.1	59.8	48.8	72.5	50.1	47.5
1976	56.9	45.1	61.6	51.5	75.2	55.1	52.0
1977	60.6	49.4	65.5	54.9	78.6	59.0	57.0
1978	65.2	52.5	72.0	60.5	81.4	61.7	61.8
1979	72.6	65.7	79.9	68.9	84.9	70.5	67.5
1980	82.4	86.0	86.8	81.0	90.9	83.1	74.9
1981	90.9	97.7	93.6	90.5	95.3	93.2	82.9
1982	96.5	99.2	97.4	96.9	97.8	97.0	92.5
1983	99.6	99.9	99.4	99.1	100.2	99.3	100.6
1984	103.9	100.9	103.2	104.0	102.1	103.7	106.8
1985	107.6	101.6	105.6	109.8	105.0	106.4	113.5
1986	109.6	88.2	109.0	115.8	105.9	102.3	122.0
1987	113.6	88.6	113.5	121.3	110.6	105.4	130.1
1988	118.3	89.3	118.2	127.1	115.4	108.7	138.6
1989	124.0	94.3	125.1	132.8	118.6	114.1	149.3
1990	130.7	102.1	132.4	140.0	124.1	120.5	162.8
1991	136.2	102.5	136.3	146.3	128.7	123.8	177.0
1992	140.3	103.0	137.9	151.2	131.9	126.5	190.1
1993	144.5	104.2	140.9	155.7	133.7	130.4	201.4
1994	148.2	104.6	144.3	160.5	133.4	134.3	211

A.i.

OLS REGRESSION
FULL SAMPLE

```
Number of obs =   23400
F( 24, 23375) = 1115.36
Prob > F      =  0.0000
R-squared     =  0.5338
Adj R-squared =  0.5334
Root MSE      =  .28685
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0535661	.0013153	40.725	0.000	.050988	.0561442
age2	-.0003604	.000011	-32.771	0.000	-.0003819	-.0003388
exp	-.0197401	.0011398	-17.319	0.000	-.0219742	-.017506
ftpt	.1599166	.004377	36.536	0.000	.1513374	.1684958
rac	.0065918	.0047682	1.382	0.167	-.0027542	.0159377
sx	.0818698	.0045526	17.983	0.000	.0729465	.0907931
job1	.2187528	.0071471	30.607	0.000	.204744	.2327615
job2	-.0089087	.0053869	-1.654	0.098	-.0194674	.0016501
job3	.0582952	.0068471	8.514	0.000	.0448744	.071716
job4	.2238482	.0078891	28.374	0.000	.2083851	.2393114
reg1	.035476	.00568	6.246	0.000	.0243428	.0466093
reg2	-.0280131	.0052843	-5.301	0.000	-.0383707	-.0176556
reg4	.108241	.0059102	18.314	0.000	.0966567	.1198253
yr1	.1156502	.0075131	15.393	0.000	.1009239	.1303764
yr2	.0926718	.0074065	12.512	0.000	.0781545	.1071891
yr3	.0802498	.0080339	9.989	0.000	.0645028	.0959968
yr4	.0573838	.0084867	6.762	0.000	.0407493	.0740183
yr5	.0220183	.0084856	2.595	0.009	.005386	.0386507
yr6	.0140231	.008356	1.678	0.093	-.0023553	.0304015
yr7	-.0023816	.0081978	-0.291	0.771	-.0184498	.0136866
yr8	-.0072297	.0080682	-0.896	0.370	-.0230439	.0085845
yr9	-.0067001	.007977	-0.840	0.401	-.0223355	.0089353
percu	.1806084	.0117704	15.344	0.000	.1575376	.2036791
union	.3171656	.0047958	66.134	0.000	.3077656	.3265657
_cons	5.153676	.0239623	215.074	0.000	5.106709	5.200644

A.ii.

OLS REGRESSION
SPILL40

```
Number of obs =    8719
F( 24,  8694) =  418.05
Prob > F      =  0.0000
R-squared     =  0.5358
Adj R-squared =  0.5345
Root MSE      =  .29575
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.05937	.0021228	27.968	0.000	.0552089	.0635312
age2	-.0003796	.0000184	-20.617	0.000	-.0004157	-.0003435
exp	-.0235689	.0017786	-13.251	0.000	-.0270553	-.0200824
ftpt	.1811352	.0075119	24.113	0.000	.1664101	.1958603
rac	.0209239	.0076281	2.743	0.006	.005971	.0358768
sx	.0619057	.0074762	8.280	0.000	.0472506	.0765609
job1	.2172804	.0120123	18.088	0.000	.1937335	.2408274
job2	-.0088868	.0090416	-0.983	0.326	-.0266105	.0088369
job3	.0369272	.0110614	3.338	0.001	.0152442	.0586102
job4	.2323958	.0130932	17.749	0.000	.20673	.2580616
reg1	.0106727	.0095588	1.117	0.264	-.0080648	.0294102
reg2	-.0686525	.0104541	-6.567	0.000	-.089145	-.04816
reg4	.0870921	.0108338	8.039	0.000	.0658553	.1083289
yr1	.0893595	.0939266	0.951	0.341	-.0947588	.2734778
yr2	-.1914998	.2095024	-0.914	0.361	-.6021742	.2191745
yr3	.0717213	.0116978	6.131	0.000	.0487908	.0946518
yr4	.0652087	.0123481	5.281	0.000	.0410035	.0894138
yr5	.0257821	.0125541	2.054	0.040	.0011732	.0503911
yr6	.0282107	.0125336	2.251	0.024	.0036418	.0527796
yr7	-.010468	.0119654	-0.875	0.382	-.0339231	.012987
yr8	-.0126915	.0118102	-1.075	0.283	-.0358423	.0104592
yr9	-.0036969	.0117556	-0.314	0.753	-.0267406	.0193468
percu	.1206361	.0191878	6.287	0.000	.0830235	.1582487
union	.2956755	.0075594	39.114	0.000	.2808572	.3104937
_cons	5.117058	.0384898	132.946	0.000	5.041609	5.192507

A.iii.

OLS REGRESSION
SPILL20

```
Number of obs =   23400
F( 24, 23375) = 1102.61
Prob > F      =  0.0000
R-squared     =  0.5310
Adj R-squared =  0.5305
Root MSE      =  .28773
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.053491	.0013193	40.544	0.000	.050905	.056077
age2	-.0003603	.000011	-32.663	0.000	-.0003819	-.0003387
exp	-.0196696	.0011433	-17.204	0.000	-.0219106	-.0174287
ftpt	.1593361	.0043902	36.293	0.000	.150731	.1679413
rac	.0040601	.0047782	0.850	0.395	-.0053055	.0134258
sx	.0841115	.0045628	18.434	0.000	.0751681	.093055
job1	.2194865	.0071687	30.617	0.000	.2054353	.2335377
job2	-.0079121	.0054036	-1.464	0.143	-.0185034	.0026793
job3	.0598522	.0068676	8.715	0.000	.0463912	.0733132
job4	.2237086	.0079134	28.270	0.000	.2081979	.2392193
reg1	.0387891	.0060212	6.442	0.000	.026987	.0505911
reg2	-.0205722	.0053042	-3.878	0.000	-.0309688	-.0101755
reg4	.1120677	.0063772	17.573	0.000	.0995679	.1245674
yr1	.1134731	.007534	15.061	0.000	.0987059	.1282403
yr2	.0903476	.0074269	12.165	0.000	.0757904	.1049048
yr3	.0807795	.0080584	10.024	0.000	.0649844	.0965745
yr4	.0580048	.0085127	6.814	0.000	.0413193	.0746902
yr5	.0223092	.0085116	2.621	0.009	.005626	.0389925
yr6	.0142885	.0083828	1.705	0.088	-.0021423	.0307193
yr7	-.0012121	.0082224	-0.147	0.883	-.0173284	.0149043
yr8	-.007678	.0080929	-0.949	0.343	-.0235406	.0081846
yr9	-.0066269	.0080017	-0.828	0.408	-.0223109	.009057
percu	.1335737	.0139715	9.560	0.000	.1061887	.1609587
union	.3287767	.0047619	69.043	0.000	.3194431	.3381104
_cons	5.16353	.0240277	214.899	0.000	5.116434	5.210626

A.iv.

OLS Regressions of Total Sample By Year

-> year= 84

Number of obs = 2886
F(15, 2870) = 245.89
Prob > F = 0.0000
R-squared = 0.5624
Adj R-squared = 0.5601
Root MSE = .29229

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0648833	.0039728	16.332	0.000	.0570935	.0726731
age2	-.0004907	.0000338	-14.508	0.000	-.000557	-.0004243
exp	-.0209571	.003472	-6.036	0.000	-.0277649	-.0141492
sx	.0854859	.0137668	6.210	0.000	.0584921	.1124797
rac	.0143153	.0153296	0.934	0.350	-.0157429	.0443734
ftpt	.1426895	.0129149	11.048	0.000	.1173662	.1680129
job1	.199326	.0210133	9.486	0.000	.1581234	.2405286
job2	-.0152951	.0154777	-0.988	0.323	-.0456437	.0150534
job3	.033045	.0215224	1.535	0.125	-.0091559	.0752459
job4	.2052488	.0229098	8.959	0.000	.1603275	.2501701
reg1	-.0515727	.0174336	-2.958	0.003	-.0857563	-.0173892
reg2	-.024261	.0153093	-1.585	0.113	-.0542793	.0057573
reg4	.1220747	.0180599	6.759	0.000	.086663	.1574865
percu	.095806	.039398	2.432	0.015	.0185549	.1730572
union	.3982178	.013407	29.702	0.000	.3719295	.4245061
_cons	5.105704	.0710007	71.911	0.000	4.966486	5.244921

-> year= 85

Number of obs = 3047
F(15, 3031) = 268.02
Prob > F = 0.0000
R-squared = 0.5701
Adj R-squared = 0.5680
Root MSE = .29122

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0611287	.0038814	15.749	0.000	.0535183	.0687392
age2	-.0004486	.0000339	-13.244	0.000	-.000515	-.0003822
exp	-.0194083	.0032861	-5.906	0.000	-.0258515	-.0129651
sx	.099034	.0130068	7.614	0.000	.0735309	.124537
rac	.0096734	.0141938	0.682	0.496	-.018157	.0375039
ftpt	.1618207	.0124039	13.046	0.000	.1374998	.1861415
job1	.1882365	.019749	9.531	0.000	.1495137	.2269592
job2	.002315	.0152534	0.152	0.879	-.027593	.032223
job3	.0709405	.0197761	3.587	0.000	.0321647	.1097164
job4	.2177475	.0215804	10.090	0.000	.1754338	.2600613
reg1	-.0248864	.0168252	-1.479	0.139	-.0578764	.0081036
reg2	-.0095996	.014895	-0.644	0.519	-.0388048	.0196057
reg4	.1187793	.0180668	6.574	0.000	.0833548	.1542037
percu	.2070531	.0383974	5.392	0.000	.1317655	.2823407
union	.3653435	.0129922	28.120	0.000	.3398691	.3908178
_cons	5.083299	.0691816	73.478	0.000	4.947652	5.218947

```
-> year=      86
Number of obs =    2241
F( 15,  2225) =  165.17
Prob > F      =  0.0000
R-squared     =  0.5269
Adj R-squared =  0.5237
Root MSE      =  .29938
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0536385	.0046301	11.585	0.000	.0445588	.0627183
age2	-.0004438	.0000387	-11.458	0.000	-.0005197	-.0003678
exp	-.0124948	.0040546	-3.082	0.002	-.0204459	-.0045437
sx	.0983857	.0156502	6.287	0.000	.0676952	.1290762
rac	.0037704	.0160122	0.235	0.814	-.0276301	.0351708
ftpt	.1411422	.0145236	9.718	0.000	.112661	.1696234
job1	.2409329	.0247546	9.733	0.000	.1923883	.2894774
job2	.0140301	.0181972	0.771	0.441	-.0216552	.0497154
job3	.0436363	.0242511	1.799	0.072	-.0039209	.0911935
job4	.2228733	.0257025	8.671	0.000	.1724699	.2732767
reg1	-.0328432	.0186889	-1.757	0.079	-.0694926	.0038062
reg2	-.0229032	.0181945	-1.259	0.208	-.0585832	.0127768
reg4	.0821535	.0198132	4.146	0.000	.0432992	.1210078
percu	.0816007	.0380557	2.144	0.032	.0069724	.156229
union	.3704507	.0159959	23.159	0.000	.3390822	.4018191
_cons	5.262731	.0826285	63.691	0.000	5.100694	5.424768

```
-> year=      87
Number of obs =    1863
F( 15,  1847) =  154.87
Prob > F      =  0.0000
R-squared     =  0.5571
Adj R-squared =  0.5535
Root MSE      =  .29059
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0539074	.0047998	11.231	0.000	.0444938	.0633211
age2	-.0004094	.0000415	-9.862	0.000	-.0004908	-.000328
exp	-.0156165	.0042682	-3.659	0.000	-.0239875	-.0072455
sx	.088118	.016726	5.268	0.000	.0553142	.1209218
rac	.0327081	.0166449	1.965	0.050	.0000633	.0653529
ftpt	.1620184	.0155988	10.387	0.000	.1314253	.1926116
job1	.2731665	.0262292	10.415	0.000	.2217246	.3246085
job2	.0097502	.0198303	0.492	0.623	-.0291419	.0486423
job3	.086098	.0255816	3.366	0.001	.0359262	.1362699
job4	.2068619	.0286804	7.213	0.000	.1506125	.2631112
reg1	.0020488	.0200019	0.102	0.918	-.0371799	.0412774
reg2	-.0666274	.0188073	-3.543	0.000	-.1035132	-.0297416
reg4	.0853967	.0207485	4.116	0.000	.0447037	.1260898
percu	.1771441	.0398026	4.451	0.000	.0990813	.2552068
union	.3527165	.0175743	20.070	0.000	.3182489	.387184
_cons	5.179273	.085446	60.615	0.000	5.011692	5.346853

-> year= 88

Number of obs = 1860
F(15, 1844) = 138.01
Prob > F = 0.0000
R-squared = 0.5289
Adj R-squared = 0.5251
Root MSE = .29367

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0563719	.0046765	12.054	0.000	.0472	.0655438
age2	-.0003653	.0000402	-9.091	0.000	-.0004441	-.0002865
exp	-.022087	.0039857	-5.542	0.000	-.0299039	-.01427
sx	.1005811	.0168174	5.981	0.000	.0675979	.1335643
rac	-.0221602	.016811	-1.318	0.188	-.0551308	.0108104
ftpt	.1627135	.0159827	10.181	0.000	.1313673	.1940597
job1	.2970547	.0250291	11.868	0.000	.2479662	.3461431
job2	.0310256	.019973	1.553	0.121	-.0081465	.0701977
job3	.075258	.025694	2.929	0.003	.0248656	.1256505
job4	.3013988	.0290802	10.364	0.000	.2443653	.3584323
reg1	.1306462	.0202082	6.465	0.000	.0910128	.1702796
reg2	.0162542	.0194951	0.834	0.405	-.0219806	.0544891
reg4	.1316634	.0215111	6.121	0.000	.0894747	.1738522
percu	.1660554	.0422858	3.927	0.000	.0831224	.2489883
union	.2537444	.0175985	14.419	0.000	.2192293	.2882594
_cons	5.086686	.0825972	61.584	0.000	4.924693	5.24868

-> year= 89

Number of obs = 1954
F(15, 1938) = 153.26
Prob > F = 0.0000
R-squared = 0.5426
Adj R-squared = 0.5390
Root MSE = .28115

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0598125	.0044847	13.337	0.000	.0510171	.0686079
age2	-.0003979	.0000374	-10.626	0.000	-.0004713	-.0003244
exp	-.0225538	.0038899	-5.798	0.000	-.0301826	-.014925
sx	.0674666	.015371	4.389	0.000	.0373211	.0976121
rac	-.0040682	.0155484	-0.262	0.794	-.0345615	.0264251
ftpt	.1924093	.0150449	12.789	0.000	.1629034	.2219152
job1	.1477404	.0245972	6.006	0.000	.0995007	.1959801
job2	-.0492526	.0182659	-2.696	0.007	-.0850755	-.0134297
job3	.0301693	.023633	1.277	0.202	-.0161795	.0765182
job4	.1919243	.0273105	7.028	0.000	.1383633	.2454852
reg1	.1037589	.0188544	5.503	0.000	.0667819	.1407359
reg2	-.0296638	.0178285	-1.664	0.096	-.0646289	.0053013
reg4	.0895736	.0194638	4.602	0.000	.0514015	.1277457
percu	.20641	.0390539	5.285	0.000	.129818	.2830019
union	.2808418	.0169343	16.584	0.000	.2476304	.3140533
_cons	5.075132	.0790953	64.165	0.000	4.920011	5.230253

```
-> year=      90
Number of obs =    2082
F( 15,  2066) =  159.95
Prob > F      =  0.0000
R-squared     =  0.5373
Adj R-squared =  0.5340
Root MSE      =  .27401
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0569361	.0039382	14.457	0.000	.0492128	.0646594
age2	-.0002272	.0000326	-6.972	0.000	-.0002911	-.0001633
exp	-.0343097	.0032632	-10.514	0.000	-.0407092	-.0279102
sx	.0468142	.0143333	3.266	0.001	.0187051	.0749233
rac	-.0069902	.0149686	-0.467	0.641	-.0363454	.022365
ftpt	.2060183	.0139993	14.716	0.000	.178564	.2334726
job1	.2222487	.0236695	9.390	0.000	.1758302	.2686673
job2	-.0182084	.0170005	-1.071	0.284	-.0515484	.0151316
job3	.0316412	.0213842	1.480	0.139	-.0102957	.0735781
job4	.2449523	.0264716	9.253	0.000	.1930385	.2968661
reg1	.1195353	.0183913	6.500	0.000	.0834678	.1556028
reg2	-.015169	.0172362	-0.880	0.379	-.0489712	.0186332
reg4	.0973791	.0183335	5.312	0.000	.0614251	.1333331
percu	.2145468	.036543	5.871	0.000	.1428819	.2862117
union	.278743	.0158956	17.536	0.000	.2475699	.3099161
_cons	5.069939	.070676	71.735	0.000	4.931335	5.208542

```
-> year=      91
Number of obs =    2193
F( 15,  2177) =  150.57
Prob > F      =  0.0000
R-squared     =  0.5092
Adj R-squared =  0.5058
Root MSE      =  .27065
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0425969	.0038408	11.090	0.000	.0350648	.050129
age2	-.0003057	.0000326	-9.366	0.000	-.0003697	-.0002417
exp	-.013681	.0032916	-4.156	0.000	-.020136	-.0072261
sx	.0664595	.0138869	4.786	0.000	.0392265	.0936925
rac	.0153002	.0143421	1.067	0.286	-.0128256	.0434259
ftpt	.1480998	.0133532	11.091	0.000	.1219134	.1742862
job1	.2137164	.0225169	9.491	0.000	.1695596	.2578732
job2	-.0087586	.0163528	-0.536	0.592	-.0408272	.02331
job3	.0675595	.0203503	3.320	0.001	.0276514	.1074676
job4	.225413	.02494	9.038	0.000	.1765043	.2743216
reg1	.0599848	.0174051	3.446	0.001	.0258524	.0941172
reg2	-.046639	.0164064	-2.843	0.005	-.0788129	-.0144652
reg4	.1156921	.018054	6.408	0.000	.0802871	.151097
percu	.2306074	.034994	6.590	0.000	.1619823	.2992326
union	.25709	.0147474	17.433	0.000	.2281695	.2860106
_cons	5.351925	.0692704	77.261	0.000	5.216082	5.487768

```
-> year=       92
Number of obs =    2275
F( 15,  2259) =  152.29
Prob > F      =  0.0000
R-squared     =  0.5028
Adj R-squared =  0.4995
Root MSE      =  .27685
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0388331	.0040181	9.665	0.000	.0309536	.0467126
age2	-.0002368	.0000329	-7.197	0.000	-.0003013	-.0001723
exp	-.0154995	.003611	-4.292	0.000	-.0225807	-.0084183
sx	.090341	.0139739	6.465	0.000	.0629381	.117744
rac	.0191216	.0148747	1.286	0.199	-.010048	.0482911
ftpt	.1440311	.0136321	10.566	0.000	.1172983	.1707639
job1	.2237224	.0218405	10.243	0.000	.1808928	.2665519
job2	-.0248433	.0171422	-1.449	0.147	-.0584595	.0087728
job3	.0769414	.0200505	3.837	0.000	.0376221	.1162607
job4	.2243869	.0247033	9.083	0.000	.1759434	.2728304
reg1	.0765729	.0177603	4.311	0.000	.0417448	.1114011
reg2	-.040015	.016123	-2.482	0.013	-.0716324	-.0083975
reg4	.115514	.0182426	6.332	0.000	.0797399	.1512881
percu	.2470499	.0362641	6.813	0.000	.1759355	.3181642
union	.2431152	.0155084	15.676	0.000	.2127029	.2735274
_cons	5.39884	.0711051	75.928	0.000	5.259402	5.538278

```
-> year=       93
Number of obs =    2999
F(  15,  2983) =  230.71
Prob > F      =  0.0000
R-squared     =  0.5371
Adj R-squared =  0.5347
Root MSE      =  .26817
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0487771	.0035519	13.733	0.000	.0418128	.0557414
age2	-.0002853	.0000276	-10.339	0.000	-.0003394	-.0002312
exp	-.0219169	.0031857	-6.880	0.000	-.0281634	-.0156705
sx	.0759535	.0115598	6.570	0.000	.0532875	.0986195
rac	.0110476	.0121506	0.909	0.363	-.0127768	.0348721
ftpt	.1559576	.0115459	13.508	0.000	.1333189	.1785963
job1	.2097178	.0183678	11.418	0.000	.1737029	.2457328
job2	-.0250805	.0141009	-1.779	0.075	-.052729	.002568
job3	.0737457	.0169788	4.343	0.000	.0404544	.1070371
job4	.2293786	.020476	11.202	0.000	.1892302	.2695271
reg1	.0482985	.0147364	3.277	0.001	.0194039	.0771931
reg2	-.0421286	.013817	-3.049	0.002	-.0692204	-.0150368
reg4	.109203	.0155036	7.044	0.000	.0788043	.1396018
percu	.2665365	.0297935	8.946	0.000	.2081187	.3249543
union	.2699434	.0128808	20.957	0.000	.2446872	.2951995
_cons	5.24102	.0631822	82.951	0.000	5.117135	5.364905

A.v.

OLS Regression
Union Individuals
Total Sample All Years

```
Number of obs =    6304
F( 23,  6280) =  324.20
Prob > F      =  0.0000
R-squared     =  0.5428
Adj R-squared =  0.5412
Root MSE      =  .30369
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0801075	.0028743	27.871	0.000	.0744729	.085742
age2	-.0007177	.0000252	-28.535	0.000	-.000767	-.0006684
exp	-.014881	.0024532	-6.066	0.000	-.0196902	-.0100719
sx	.1153607	.0092401	12.485	0.000	.0972469	.1334745
rac	.0090903	.00984	0.924	0.356	-.0101995	.0283801
ftpt	.2276442	.0093857	24.254	0.000	.2092451	.2460434
job1	.1504628	.0135894	11.072	0.000	.1238229	.1771026
job2	.0598038	.0107804	5.547	0.000	.0386705	.0809372
job3	-.0011198	.0139998	-0.080	0.936	-.0285641	.0263246
job4	.1511043	.0140187	10.779	0.000	.1236228	.1785857
reg1	-.0636673	.0129511	-4.916	0.000	-.0890558	-.0382788
reg2	-.075033	.0129463	-5.796	0.000	-.1004121	-.0496539
reg4	.0900247	.0130933	6.876	0.000	.0643573	.1156921
yr1	.2172164	.0152229	14.269	0.000	.1873743	.2470585
yr2	.1800252	.0149119	12.073	0.000	.1507927	.2092577
yr3	.1400914	.0163307	8.578	0.000	.1080776	.1721052
yr4	.1165712	.0176362	6.610	0.000	.0819982	.1511442
yr5	.0255749	.0175856	1.454	0.146	-.0088989	.0600488
yr6	.0377764	.01791	2.109	0.035	.0026666	.0728862
yr7	.0029161	.0175253	0.166	0.868	-.0314395	.0372718
yr8	-.0093996	.0166129	-0.566	0.572	-.0419666	.0231674
yr9	-.0252474	.0166998	-1.512	0.131	-.0579847	.0074899
percu	.4147837	.0220615	18.801	0.000	.3715357	.4580318
union	(dropped)					
_cons	4.836457	.0541701	89.283	0.000	4.730265	4.942649

A.vi.

OLS Regression
Non Union Individuals
Total Sample All Years

```
Number of obs =   17096
F( 23, 17072) =  446.59
Prob > F      =  0.0000
R-squared     =  0.3756
Adj R-squared =  0.3748
Root MSE      =  .26631
```

LogWage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.0452943	.0014001	32.351	0.000	.04255	.0480386
age2	-.0002574	.0000116	-22.278	0.000	-.0002801	-.0002348
exp	-.0210227	.0012184	-17.254	0.000	-.0234109	-.0186344
sx	.06233	.0049829	12.509	0.000	.0525629	.072097
rac	.0040989	.0051604	0.794	0.427	-.006016	.0142138
ftpt	.1391303	.0047191	29.483	0.000	.1298804	.1483801
job1	.2452358	.0080059	30.632	0.000	.2295433	.2609283
job2	-.0251292	.0059422	-4.229	0.000	-.0367766	-.0134818
job3	.0859938	.0074764	11.502	0.000	.0713392	.1006484
job4	.2469609	.0091811	26.899	0.000	.228965	.2649567
reg1	.084451	.006133	13.770	0.000	.0724297	.0964722
reg2	-.0153915	.0055248	-2.786	0.005	-.0262205	-.0045624
reg4	.1015164	.0064646	15.704	0.000	.0888452	.1141876
yr1	.08014	.0081985	9.775	0.000	.06407	.09621
yr2	.0601283	.0081041	7.420	0.000	.0442435	.076013
yr3	.0589657	.0087461	6.742	0.000	.0418225	.0761089
yr4	.0354299	.0091593	3.868	0.000	.0174767	.053383
yr5	.0207904	.0091728	2.267	0.023	.0028108	.03877
yr6	.0064126	.0089291	0.718	0.473	-.0110893	.0239144
yr7	-.0033988	.0087683	-0.388	0.698	-.0205856	.013788
yr8	-.0024858	.0087396	-0.284	0.776	-.0196164	.0146447
yr9	-.0005623	.0085885	-0.065	0.948	-.0173967	.0162721
percu	.072133	.013329	5.412	0.000	.0460068	.0982593
union	(dropped)					
_cons	5.345329	.0254861	209.735	0.000	5.295374	5.395284

B.i.

Heckman				
UNION	**Full Sample**			
	Coeff.	**Std. Error**	**z**	**P>|z|**
age	0.062064	0.003524	17.61	0
age2	-0.00058	2.98E-05	-19.439	0
exp	-0.0089	0.002658	-3.348	0.001
sx	0.100739	0.009933	10.141	0
rac	-0.00127	0.010501	-0.121	0.904
ftpt	0.230167	0.009896	23.259	0
job1	0.183639	0.014788	12.419	0
job2	0.103569	0.01226	8.448	0
job3	0.053935	0.015932	3.385	0.001
job4	0.148363	0.015047	9.86 0	
reg1	-0.09621	0.013143	-7.32	0
reg2	-0.10764	0.013174	-8.171	0
reg4	0.073424	0.012944	5.672	0
yr1	0.185	0.016429	11.261	0
yr2	0.143792	0.01619	8.881	0
yr3	0.124077	0.017376	7.141	0
yr4	0.106749	0.018702	5.708	0
yr5	0.014583	0.01864	0.782	0.434
yr6	0.034345	0.018909	1.816	0.069
yr7	0.009638	0.018538	0.52	0.603
yr8	-0.02272	0.017648	-1.288	0.198
yr9	-0.02842	0.017686	-1.607	0.108
percu	-0.02176	0.051384	-0.424	0.672
_cons	5.581338	0.09456	59.025	0

B.ii.

probit	Full Sample Union			
	Coeff.	Std. Error	z	P>|z|
age	0.125533	0.006905	18.179	0
age2	-0.00098	5.98E-05	-16.459	0
exp	-0.0403	0.005856	-6.883	0
sx	0.092253	0.023334	3.954	0
rac	0.053715	0.024849	2.162	0.031
ftpt	-0.01003	0.022598	-0.444	0.657
RTW	0.380282	0.025863	14.704	0
job1	-4.07116	0.138775	-29.336	0
job2	-4.11194	0.131891	-31.177	0
job3	-4.17388	0.135131	-30.888	0
job4	-3.79142	0.136809	-27.713	0
job5	-3.81022	0.129803	-29.354	0
yr1	0.156345	0.043964	3.556	0
yr2	0.181871	0.043464	4.184	0
yr3	0.060538	0.046666	1.297	0.195
yr5	0.032775	0.048994	0.669	0.504
yr6	-0.03116	0.048941	-0.637	0.524
yr7	-0.10785	0.048356	-2.23	0.026
yr8	0.021638	0.047062	0.46	0.646
yr9	-0.04111	0.047154	-0.872	0.383
yr10	-0.06486	0.044476	-1.458	0.145
percu	2.578678	0.058829	43.833	0

B.iii.

Heckman	Full Sample			
NONUNION				
	Coeff.	Std. Error	z	P>|z|
age	0.044429	0.001478	30.072	0
age2	-0.00025	1.21E-05	-20.736	0
exp	-0.02073	0.001228	-16.878	0
sx	0.061701	0.004994	12.356	0
rac	0.003755	0.005163	0.727	0.467
ftpt	0.139255	0.004719	29.51	0
job1	0.246981	0.008061	30.639	0
job2	-0.02307	0.006047	-3.815	0
job3	0.088657	0.007616	11.641	0
job4	0.246515	0.009182	26.848	0
reg1	0.083097	0.006175	13.456	0
reg2	-0.01613	0.005536	-2.913	0.004
reg4	0.10097	0.006467	15.614	0
yr1	0.078498	0.008247	9.519	0
yr2	0.058291	0.008165	7.139	0
yr3	0.058062	0.008759	6.629	0
yr4	0.034948	0.009162	3.814	0
yr5	0.020133	0.009179	2.193	0.028
yr6	0.006205	0.008929	0.695	0.487
yr7	-0.00319	0.008768	-0.364	0.716
yr8	-0.0031	0.008745	-0.355	0.723
yr9	-0.00069	0.008588	-0.08	0.936
percu	0.046045	0.019498	2.362	0.018
_cons	5.359725	0.026668	200.976	0

B.iv.

probit	**Full Sample Nonunion**			
	Coeff.	Std. Error	z	P>|z|
age	-0.12491	0.006912	-18.072	0
age2	0.000975	6.01E-05	16.216	0
exp	0.040381	0.005849	6.904	0
sx	-0.09917	0.023343	-4.248	0
rac	-0.05555	0.024979	-2.224	0.026
ftpt	0.006906	0.022574	0.306	0.76
RTW	-0.40358	0.026632	-15.154	0
job1	4.066062	0.139034	29.245	0
job2	4.118006	0.132151	31.161	0
job3	4.189077	0.135279	30.966	0
job4	3.796065	0.137119	27.684	0
job5	3.818289	0.129919	29.39	0
yr1	-0.15003	0.04415	-3.398	0.001
yr2	-0.17702	0.043657	-4.055	0
yr3	-0.05262	0.046882	-1.122	0.262
yr5	-0.02798	0.049238	-0.568	0.57
yr6	0.029897	0.049165	0.608	0.543
yr7	0.110164	0.048598	2.267	0.023
yr8	-0.02112	0.047288	-0.447	0.655
yr9	0.041388	0.047394	0.873	0.383
yr10	0.061457	0.044695	1.375	0.169
percu	-2.56473	0.059959	-42.775	0

B.v.

Heckman				
UNION	Spill40			
	Coeff.	Std. Error	z	P>|z|
age	0.057549	0.004463	12.895	0
age2	-0.0005	3.82E-05	-12.988	0
exp	-0.01112	0.003625	-3.067	0.002
sx	0.118206	0.01423	8.307	0
rac	0.0036	0.014504	0.248	0.804
ftpt	0.221097	0.014519	15.228	0
job1	0.192727	0.021409	9.002	0
job2	0.121276	0.017463	6.945	0
job3	0.054805	0.021817	2.512	0.012
job4	0.143034	0.022329	6.406	0
reg1	-0.07123	0.018121	-3.931	0
reg2	-0.1394	0.019115	-7.293	0
reg4	0.085125	0.019273	4.417	0
yr1	-0.01742	0.153227	-0.114	0.91
yr2	0	.	.	.
yr3	0.098781	0.021813	4.528	0
yr4	0.109732	0.023287	4.712	0
yr5	-0.00755	0.023578	-0.32	0.749
yr6	0.027023	0.023914	1.13	0.258
yr7	0.000476	0.022921	0.021	0.983
yr8	-0.03402	0.022356	-1.522	0.128
yr9	-0.03165	0.022078	-1.434	0.152
percu	-0.29139	0.060855	-4.788	0
_cons	5.859777	0.100483	58.316	0

B.vi.

probit	Spill40 Union			
	Coeff.	Std. Error	z	P>|z|
age	0.124179	0.010225	12.145	0
age2	-0.00096	9.02E-05	-10.595	0
exp	-0.04117	0.008436	-4.88	0
sx	-0.05086	0.035579	-1.43	0.153
rac	0.063176	0.036122	1.749	0.08
ftpt	-0.01345	0.035665	-0.377	0.706
RTW	0.199289	0.047684	4.179	0
job2	-0.10831	0.054338	-1.993	0.046
job3	-0.13726	0.061271	-2.24	0.025
job4	0.318754	0.068178	4.675	0
job5	0.287281	0.056222	5.11	0
yr3	-0.12635	0.392775	-0.322	0.748
yr4	-0.14682	0.393195	-0.373	0.709
yr5	-0.09032	0.393354	-0.23	0.818
yr6	-0.15843	0.393308	-0.403	0.687
yr7	-0.26537	0.393015	-0.675	0.5
yr8	-0.16811	0.392889	-0.428	0.669
yr9	-0.17219	0.392832	-0.438	0.661
yr10	-0.25194	0.392296	-0.642	0.521
percu	2.960447	0.102089	28.999	0
_cons	-3.80334	0.43858	-8.672	0

B.vii.

Heckman	**Spill40**			
NONUNION				
	Coeff.	**Std. Error**	**z**	**P>|z|**
age	0.04881	0.002523	19.347	0
age2	-0.00025	2.17E-05	-11.289	0
exp	-0.02531	0.002024	-12.508	0
sx	0.034415	0.00877	3.924	0
rac	0.015595	0.008923	1.748	0.081
ftpt	0.16162	0.008796	18.375	0
job1	0.257145	0.014938	17.214	0
job2	-0.02788	0.011054	-2.523	0.012
job3	0.079706	0.013355	5.968	0
job4	0.280785	0.016925	16.59	0
reg1	0.056254	0.010847	5.186	0
reg2	-0.03481	0.011974	-2.907	0.004
reg4	0.070173	0.012486	5.62 0	
yr1	0.071503	0.123894	0.577	0.564
yr2	-0.16509	0.195421	-0.845	0.398
yr3	0.048938	0.013861	3.531	0
yr4	0.033905	0.014516	2.336	0.02
yr5	0.036739	0.014833	2.477	0.013
yr6	0.025763	0.014612	1.763	0.078
yr7	-0.00945	0.013901	-0.68	0.496
yr8	-0.0027	0.013845	-0.195	0.845
yr9	0.012383	0.013857	0.894	0.372
percu	0.04172	0.030162	1.383	0.167
_cons	5.33611	0.045007	118.562	0

B.viii.

probit	Spill40 Nonunion			
	Coeff.	**Std. Error**	**z**	**P>|z|**
age	-0.12196	0.010301	-11.84	0
age2	0.000939	9.13E-05	10.28	0
exp	0.040389	0.008467	4.77	0
sx	0.037864	0.035805	1.058	0.29
rac	-0.06765	0.036649	-1.846	0.065
ftpt	-0.0013	0.035846	-0.036	0.971
RTW	-0.24437	0.053358	-4.58	0
job2	0.135329	0.05471	2.474	0.013
job3	0.190902	0.061692	3.094	0.002
job4	-0.29708	0.069321	-4.286	0
job5	-0.25495	0.057232	-4.455	0
yr3	0.223382	0.399388	0.559	0.576
yr4	0.227897	0.399847	0.57	0.569
yr5	0.19263	0.399988	0.482	0.63
yr6	0.240026	0.399969	0.6	0.548
yr7	0.356102	0.39966	0.891	0.373
yr8	0.255846	0.399548	0.64	0.522
yr9	0.253915	0.399539	0.636	0.525
yr10	0.325738	0.398999	0.816	0.414
percu	-2.92808	0.107488	-27.241	0
_cons	3.687358	0.445149	8.283	0

B.ix.

Heckman				
UNION	**Spill20**			
	Coeff.	**Std. Error**	**z**	**P>|z|**
age	0.082434	0.004986	16.533	0
age2	-0.00073	4.04E-05	-18.156	0
exp	-0.01577	0.00279	-5.655	0
sx	0.119895	0.009991	12	0
rac	0.004506	0.009875	0.456	0.648
ftpt	0.222842	0.009437	23.613	0
job1	0.149816	0.015609	9.598	0
job2	0.054469	0.014727	3.699	0
job3	-0.00784	0.018371	-0.427	0.67
job4	0.151497	0.014054	10.78	0
reg1	-0.1161	0.015016	-7.732	0
reg2	-0.06918	0.014496	-4.772	0
reg4	0.043173	0.014623	2.952	0.003
yr1	0.210818	0.016495	12.781	0
yr2	0.173248	0.016529	10.481	0
yr3	0.146327	0.016996	8.61	0
yr4	0.121628	0.017905	6.793	0
yr5	0.032833	0.017935	1.831	0.067
yr6	0.046147	0.018025	2.56	0.01
yr7	0.008593	0.017542	0.49	0.624
yr8	-0.01148	0.016866	-0.681	0.496
yr9	-0.01874	0.016777	-1.117	0.264
percu	0.58881	0.111068	5.301	0
_cons	4.7525	0.174867	27.178	0

B.x.

probit	**Spill20 Union**			
	Coeff.	**Std. Error**	**z**	**P>|z|**
age	0.123681	0.006868	18.009	0
age2	-0.00096	5.94E-05	-16.113	0
exp	-0.04033	0.005827	-6.921	0
sx	0.118716	0.02318	5.121	0
rac	0.022359	0.02474	0.904	0.366
ftpt	-0.01954	0.022446	-0.871	0.384
RTW	0.306773	0.029143	10.526	0
job1	-3.9836	0.138075	-28.851	0
job2	-4.04142	0.13122	-30.799	0
job3	-4.11088	0.134395	-30.588	0
job4	-3.71244	0.13619	-27.259	0
job5	-3.74271	0.129081	-28.995	0
yr1	0.111312	0.043877	2.537	0.011
yr2	0.137311	0.043367	3.166	0.002
yr3	0.047337	0.046408	1.02	0.308
yr5	0.01797	0.048724	0.369	0.712
yr6	-0.02943	0.04866	-0.605	0.545
yr7	-0.10834	0.048054	-2.254	0.024
yr8	-0.00472	0.046973	-0.1	0.92
yr9	-0.04818	0.046881	-1.028	0.304
yr10	-0.08249	0.044283	-1.863	0.062
percu	2.861598	0.074794	38.26	0

B.xi.

Heckman	Spil120			
NONUNION				
	Coeff.	Std. Error	z	P>|z|
age	0.04421	0.001478	29.911	0
age2	-0.00025	1.21E-05	-20.705	0
exp	-0.02054	0.001229	-16.708	0
sx	0.06299	0.005001	12.596	0
rac	0.002769	0.00516	0.537	0.592
ftpt	0.139246	0.004724	29.474	0
job1	0.247418	0.008066	30.674	0
job2	-0.02163	0.006052	-3.574	0
job3	0.090525	0.007619	11.881	0
job4	0.246094	0.009191	26.775	0
reg1	0.092607	0.006378	14.52	0
reg2	-0.00987	0.005545	-1.781	0.075
reg4	0.110908	0.006877	16.128	0
yr1	0.078308	0.008241	9.503	0
yr2	0.058097	0.008156	7.123	0
yr3	0.058296	0.008769	6.648	0
yr4	0.034863	0.009173	3.801	0
yr5	0.020335	0.009189	2.213	0.027
yr6	0.005887	0.008939	0.659	0.51
yr7	-0.00269	0.008775	-0.307	0.759
yr8	-0.00296	0.008752	-0.338	0.735
yr9	-0.00117	0.008596	-0.136	0.892
percu	-0.01324	0.021134	-0.626	0.531
_cons	5.369533	0.0266	201.862	0

B.xii.

probit	Spil120 Nonunion			
	Coeff.	Std. Error	z	P>|z|
age	-0.12388	0.00687	-18.033	0
age2	0.00096	5.95E-05	16.144	0
exp	0.040277	0.005827	6.913	0
sx	-0.12079	0.023178	-5.212	0
rac	-0.02136	0.024742	-0.863	0.388
ftpt	0.019002	0.02241	0.848	0.396
RTW	-0.30643	0.029156	-10.51	0
job1	3.988264	0.138092	28.881	0
job2	4.047545	0.131287	30.83	0
job3	4.115368	0.134399	30.621	0
job4	3.719434	0.136195	27.31	0
job5	3.748131	0.129132	29.026	0
yr1	-0.11247	0.043874	-2.564	0.01
yr2	-0.13885	0.043379	-3.201	0.001
yr3	-0.04695	0.046421	-1.011	0.312
yr5	-0.01976	0.048732	-0.405	0.685
yr6	0.027712	0.04867	0.569	0.569
yr7	0.107079	0.04806	2.228	0.026
yr8	0.004095	0.046978	0.087	0.931
yr9	0.047194	0.046885	1.007	0.314
yr10	0.081524	0.04429	1.841	0.066
percu	-2.86515	0.074892	-38.257	0

C.i.

FULL SAMPLE
IV REGRESSION

	UNION		NONUNION	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**
AGE	0.056495	0.003231	-0.01455	0.000866
AGE^2	-0.00075	4.84E-05	-6.90E-05	1.49E-05
EXPER	0.013861	0.003871	0.022805	0.001131
FTPT	0.448555	0.019196	0.189556	0.006486
RACE	*-0.03138*	*0.019357*	-0.03804	0.006922
SEX	0.123256	0.018587	0.018133	0.006811
SUPER	0.387194	0.025173	0.34728	0.011194
SALES	0.257433	0.022647	-0.02419	0.008359
SUPOR	0.134258	0.031123	0.117894	0.010439
SKILL	0.189154	0.024278	0.341348	0.013614
EAST	-0.40593	0.031903	0.11295	0.008789
MIDW	-0.38068	0.03212	-0.02048	0.007353
WEST	*-0.00574*	*0.032549*	0.132314	0.008894
1984	0.183109	0.030249	0.033573	0.011299
1985	0.119405	0.029403	*0.00305*	*0.011195*
1986	0.154769	0.031606	*0.013306*	*0.011899*
1987	0.164016	0.03468	*-0.01279*	*0.012339*
1988	*-0.03336*	*0.034814*	-0.04577	0.012393
1989	*0.028667*	*0.036882*	-0.0541	0.011929
1990	*-0.01*	*0.036141*	-0.05936	0.011703
1991	-0.08014	0.032985	-0.05886	0.011852
1992	-0.09591	0.032848	-0.0601	0.011504
Percu	-1.35302	0.076122	-0.17624	0.028362
	CONST	6.48508	0.009014	

NB: Italics indicate P>|z| greater than alpha = 0.05

C.ii.

SPILL40
IV REGRESSION

	UNION		NONUNION	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**
AGE	0.062353	0.004522	-0.01133	0.001557
AGE^2	-0.00071	6.01E-05	*-2.90E-05*	*2.81E-05*
EXPER	*0.004411*	*0.004931*	0.017019	0.00209
FTPT	0.399562	0.023623	0.240841	0.013022
RACE	*-0.02961*	*0.023799*	-0.04144	0.0128
SEX	0.188483	0.02301	*-0.01486*	*0.012758*
SUPER	0.343542	0.031386	0.374502	0.022093
SALES	0.248767	0.02912	-0.03539	0.016289
SUPPOR	0.105261	0.037731	0.108249	0.019497
SKILL	0.158442	0.030374	0.417513	0.026741
EAST	-0.24819	0.037965	0.088919	0.015929
MIDWES	-0.29378	0.038177	-0.04333	0.017505
WEST	0.101406	0.039498	0.103832	0.018334
1984	*-0.23258*	*0.224707*	*0.014205*	*0.21766*
1985	(dropped)		(dropped)	
1986	0.136314	0.034278	*0.02059*	*0.020357*
1987	0.158917	0.036584	*-0.01714*	*0.021026*
1988	*-0.06315*	*0.037161*	*-0.03812*	*0.021555*
1989	0.015024	0.038572	*-0.03475*	*0.021117*
1990	*-0.01862*	*0.037954*	-0.06756	0.020105
1991	-0.07415	0.035815	-0.06241	0.020042
1992	-0.09451	0.034475	-0.05195	0.020133
Percu	-1.61859	0.126024	-0.09738	0.043586
	CONST	6.463768	0.014334	

NB: italics indicate that P>|z| greater than alpha = 0.05
Note: 1985 dropped due to insufficient observations.

C.iii.

SPILL20
IV REGRESSION

	UNION		NONUNION	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**
AGE	0.051132	0.003263	-0.01448	0.000886
AGE^2	-0.00072	5.06E-05	-6.50E-05	1.51E-05
EXPER	0.017038	0.004008	0.022352	0.001139
FTPT	0.494621	0.020066	0.19088	0.006551
RACE	*-0.02052*	*0.0202*	-0.03795	0.006974
SEX	0.100549	0.019357	0.016572	0.006864
SUPER	0.409979	0.026256	0.353226	0.011316
SALES	0.283176	0.023467	-0.02192	0.008434
SUPPOR	0.163119	0.032691	0.120947	0.010528
SKILL	0.189962	0.025076	0.351311	0.013775
EAST	-0.52722	0.037676	0.120517	0.008944
MIDWES	-0.47596	0.037623	-0.01517	0.007344
WEST	-0.1845	0.038706	0.145523	0.009591
1984	0.208021	0.030827	0.03131	0.011336
1985	0.139253	0.02998	*0.00062*	*0.011226*
1986	0.133285	0.033509	*0.01137*	*0.012024*
1987	0.12967	0.036743	*-0.01511*	*0.012458*
1988	*-0.04521*	*0.037081*	-0.04692	0.012517
1989	*0.009077*	*0.039251*	-0.05603	0.012043
1990	*-0.01928*	*0.038539*	-0.06077	0.011817
1991	-0.09414	0.034518	-0.0595	0.011934
1992	-0.08651	0.035119	-0.06059	0.011629
Percu	-0.99826	0.086139	-0.21296	0.028432
	CONST	6.483438	0.009226	

NB: italics indicate P>|z| greater than alpha = 0.05

C.iv.
LOGIT

Full Sample

	Coeff.	Std. Error
AGE	0.217756	0.012014
AGE^2	-0.00168	0.000105
EXPER	-0.0716	0.010068
SEX	0.16614	0.040381
RACE	0.101066	0.043423
FTPT	-0.0124	0.039066
SUPER	-7.03307	0.24358
SALES	-7.12734	0.231954
SUPPOR	-7.24924	0.237328
SKILL	-6.56449	0.239817
STOCK	-6.59292	0.2277
1984	0.252894	0.076294
1985	0.302454	0.075373
1986	0.090963	0.081284
1988	0.041492	0.085479
1989	-0.0579	0.08561
1990	-0.19677	0.084588
1991	0.036201	0.082077
1992	-0.06925	0.082147
1993	-0.10796	0.077535
RTW	0.736731	0.048686
PERCU	4.27718	0.103387

Spill40

	Coeff.	Std. Error
AGE	0.21031	0.017572
AGE^2	-0.0016	0.000156
EXPER	-0.07157	0.014301
SEX	-0.06636	0.060566
RACE	0.117872	0.061902
FTPT	-0.0022	0.060338
dropped due to collinearity		
SALES	-0.2275	0.092418
SUPPOR	-0.31671	0.104205
SKILL	0.510071	0.117127
STOCK	0.443621	0.096618
1986	-0.61978	0.70272
1987	-0.6336	0.703365
1988	-0.56804	0.703658
1989	-0.64785	0.70355
1990	-0.84757	0.703105
1991	-0.6676	0.702917
1992	-0.66693	0.702841
1993	-0.79072	0.701937
RTW	0.459225	0.094413
PERCU	4.889436	0.187323

Spill20

	Coeff.	Std.Error
AGE	0.214689	0.011901
AGE^2	-0.00165	0.000103
EXPER	-0.0711	0.010019
SEX	0.2025	0.03986
RACE	0.036956	0.042805
FTPT	-0.02913	0.038563
SUPER	-6.86711	0.240876
SALES	-6.97677	0.229487
SUPPOR	-7.0888	0.234721
SKILL	-6.40287	0.237091
STOCK	-6.44068	0.225311
1984	0.182116	0.075488
1985	0.23139	0.074548
1986	0.084176	0.07984
1988	0.021603	0.084095
1989	-0.05451	0.084223
1990	-0.19275	0.083103
1991	0.000555	0.080756
1992	-0.08096	0.080687
1993	-0.14109	0.076222
RTW	0.599164	0.052175
PERCU	4.72359	0.127643