This article uses data from the Health and Retirement Survey to examine the characteristics of individuals who are covered under integrated pension plans by comparing them with people covered by non-integrated plans and those with no pension plan.

Acknowledgments: Many thanks to John Woods for the many thoughtful discussions about pension integration in general and this article in particular. Thanks also go to Ben Bridges, Sharmila Choudhury, Susan Grad, Tom Hungerford, and Mike Leonesio for their helpful comments and suggestions. Finally, I want to thank Chuck Slusher for developing the basic pension data for this study.

Characteristics of Individuals with Integrated Pensions

by Keith A. Bender, Division of Economic Research, Office of Research, Evaluation, and Statistics, Office of Policy, Social Security Administration

Summary

Employer pensions that integrate benefits with Social Security have been the focus of relatively little research. Since changes in Social Security benefit levels and other program characteristics can affect the benefit levels and other features of integrated pension plans, it is important to know who is covered by these plans. This article examines the characteristics of workers covered by integrated pension plans, compared to those with nonintegrated plans and those with no pension coverage.

Integrated pension plans are those that explicitly adjust their benefit structure to help compensate for the employer's contributions to the Social Security program. There are two basic integration methods used by defined benefit (DB) plans. The offset method causes a reduction in employer pension benefits by up to half of the Social Security retirement benefit; the excess rate method is characterized by an accrual rate that is lower for earnings below the Social Security taxable maximum than above it. Defined contribution (DC) pension plans can be integrated along the lines of the excess rate method.

To date, research on integrated pensions has focused on plan characteristics, as reported to the Bureau of Labor Statistics (BLS) through its Employee Benefits Survey (EBS). This research has examined the prevalence of integration among full-time, private sector workers by industry, firm size, and broad occupational categories. However, because the EBS provides virtually no data on worker characteristics, analyses of the effects of pension integration on retirement benefits have used hypothetical workers, varying according to assumed levels of earnings and job tenure.

This kind of analysis is not particularly helpful in examining the potential effects of changes in the Social Security program on workers' pension benefits. However, data on pension integration at the individual level are available, most recently from the Health and Retirement Study (HRS), a nationally representative survey of individuals aged 51-61 in 1992. This dataset provides the basis for the analysis presented here.

The following are some of the major findings from this analysis.

- The incidence of pension integration in the HRS sample is 32 percent of all workers with a pension (14 percent of all workers). The HRS can also identify integrated DC plans, a statistic that is not available from BLS data. The rate of integration for workers with only DC plans is 8 percent.

- After controlling for other variables, several socio-demographic characteristics are significantly related to the incidence of integration. The probability of having an integrated pension is 4.6 percentage points less for men compared to women. Non-Hispanic blacks are 6.4 percentage points less likely than non-Hispanic whites to have integrated pensions. Union members are 14 percentage points less likely to have integrated pensions, while workers with less than a graduate level education are at least 15 percentage points more likely to have a pension that is integrated.
- Some earnings and pension characteristics are also significantly correlated with pension integration. Earnings are positively related, with the probability of having an integrated pension increasing by 2 percentage points for an increase of $1,000 in annual pay. An even larger effect comes from earning at or above the Social Security taxable maximum. Workers at or above this income level are 10 percentage points more likely to have an integrated plan, but for those with more than one plan the probability of pension integration goes up by 13 percentage points.

I. Introduction

In considering proposals to ensure Social Security solvency, it is important to understand the interrelationships of the three primary sources of post-retirement income—income from the Old-Age, Survivors, and Disability Insurance (OASDI) program, from savings and assets, and from employer-sponsored pensions. While there is a large literature covering the relationship between Social Security and private savings and a smaller literature regarding the interrelationships of employer pensions and private savings, research on the integration of Social Security and employer pension benefits is quite limited. The integration of benefits can take place in a variety of ways. Examples include reducing employer pension benefits by a percentage of an individual's Social Security benefit (the so-called offset integration method) or having pension benefit accrual rates that are lower for low wages than for high wages (the excess rate method). According to Bureau of Labor Statistics (BLS) data, pension integration is observed primarily in private sector defined benefit (DB) plans; about one-half of the full-time workers in large and medium sized firms who participate in DB plans are covered by a benefit formula that is integrated with Social Security.[1] Many proposed changes in Social Security could, therefore, affect the pensions of a significant number of retirees and workers and substantially increase employer pension costs. In sum, examining measures to ensure Social Security solvency without taking into account pension integration would miss an important aspect of U.S. retirement policy.

Most of the research on pension integration coverage centers on data collected from firms. Since 1980, the Employee Benefits Survey, collected and published by the BLS, has been the main data source on the incidence of integration among employer pension plans. The data are collected at the firm level, and therefore, there is very little information about the characteristics of the persons who are covered by integrated pension plans. To evaluate proposals to change Social Security, it is important to know not only *which* pension plans would be affected but also *who* would be affected.

Fortunately, the detailed pension data from the Health and Retirement Study (HRS) allow the analysis of pension integration at the individual level. Beginning in 1992, the HRS has biannually collected demographic, work history, health, and retirement information from the same group of individuals who were aged 51-61 in 1992. The 1992 wave of the HRS also contains information from employers on the pension plans to which workers or retirees belong. As a result, the 1992 HRS is a very rich data source with which to analyze the incidence of pension integration. Since the focus of this article is identifying the characteristics of persons covered by integrated pensions, it is this dataset which forms the basis for the results presented here. The next section of this article briefly details the types of pension integration and outlines previous research on integration. Section III describes the data employed in this study, while section IV contains the methodology and results.

II. Pension Integration

Brief Overview of Pension Integration[2]

Firms give several rationales for integrating their pension benefits with Social Security. First, firms are nominally required to pay half of the OASDI payroll tax. Pension integration, therefore, allows firms to offset at least part of this tax by reducing employee pension benefits. This reduction is an incentive for employers to offer a pension plan. Second, without integration, some low-paid workers could be much more likely to receive combined pension and Social Security benefits that are greater than their pre-retirement earnings, which is seen by some as unfair. Finally, integration allows higher pension benefits, within limits, for more highly paid workers; thus, it can be used as a tool by firms to retain and motivate highly skilled workers.

The first legislation for the integration of private employer pensions and Social Security benefits occurred in the Revenue Act of 1942 (Dyer 1977). However, integration provisions have changed much since then and went through the last major change in the Tax Reform Act of 1986 (TRA86). Much of the TRA86 regulation of these plans involves the scope of how much more firms can provide in benefits to high-paid workers relative to low-paid workers (called permitted disparity). These limits are part of nondiscrimination legislation, which mandates that within certain bounds, pension plans cannot discriminate against low-income workers compared with high-income workers. One consequence of these limits is that no individual can lose an entire pension benefit under an integrated pension plan, which was possible before the TRA86.

Pension plans can be integrated with Social Security in several ways. Excess rate and offset plans are the two methods

of integrating defined benefit (DB) pensions. Excess rate plans have a lower benefit accrual rate for earnings under some earnings threshold, or integration level (usually the Social Security taxable maximum earnings level) than for earnings above that level. For example, earnings below the taxable maximum could accrue benefits at 1 percent of final salary per year of service, while earnings above the maximum could accrue benefits at 1.5 percent. Therefore, if a worker with 30 years of service earned $85,000 before retiring in 1998, he or she would receive (0.01*30*$68,400=) $20,520 in pension benefits for earnings below the taxable maximum ($68,400) and an additional (0.015*30*($85,000-$68,400)=) $7,470 from the earnings above the maximum. In offset plans, firms are able to subtract some portion of the recipient's initial (usually estimated) Social Security benefit from the employer pension benefit. For example, if a retiree's Social Security benefit was $8,000 per year and the pension benefit was $10,000, integration could cause the pension benefit to be cut by half of the Social Security benefit—that is by $4,000. The resulting net pension benefit would be $6,000. Although more rare, defined contribution plans can also be integrated by having an employer contribution rate that is lower for earnings below the integration level than for those above the integration level, similar to DB excess plans. For more numerical examples of benefits based on representative workers and typical integrated plans, see Wiatrowski (1991), Graham (1994), and Kollmann and Schmitt (1994).

Previous Research[3]

Firm-level research.—As mentioned above, previous data collection in the area of pension integration has occurred primarily at the firm level. The BLS conducts a biennial survey of employee benefits in medium and large firms, which includes information on whether a pension plan is integrated with Social Security. The most recent survey, from 1995, indicates that among full-time private sector workers who are employed in large and medium size firms and are covered by a defined benefit pension plan,[4] 51 percent were covered by an integrated benefit formula.[5] This rate is down from a high of over 60 percent in the late 1980s, although it is up from the 1993 figure of 48 percent. Historically, offset plans have been the dominant type of integrated plan, although by 1991 the majority of participants in integrated plans were in excess rate plans. This shift occurred, according to Kollmann and Schmitt (1994), because the TRA86 legislation made it relatively harder for offset plans to prove nondiscrimination, compared with excess rate plans. Finally, while pension integration is primarily a characteristic of private sector pension plans, state and local government pension plans can also be integrated with Social Security benefits. As reported in Graham (1994) and Foster (1997), the incidence of integration in defined benefit plans in the state and local government sector fell from 8 percent of workers in 1991 to 4 percent by 1993-94. Pension plans for federal government employees are not explicitly integrated with Social Security.[6]

Individual-level research.—The 1992 Health and Retirement Survey (HRS) is the most recent micro level data set that can be used to explore the incidence and characteristics of pension integration among individuals.[7] The only study that has investigated the issue of pension integration at the individual level is Slusher (1998). Using HRS data, he examines the incidence of pension integration by certain characteristics of individuals. His results show that, among pension-covered individuals, the incidence of pension integration is slightly higher for males and whites, those with a bachelor or associate degree, and those with high average hourly wages (over $25/hour). Furthermore, workers in large firms (over 500 employees) and those in durable goods manufacturing, finance, insurance, and real estate industries have higher rates of integration. He also finds that excess rate plans are more common than offset plans.

Pension integration and proposed changes in Social Security.—There are also recent analyses regarding how changes in Social Security may affect pension integration. Slusher (1998) discusses the potential effects of two types of proposals: the institution of individual retirement accounts and the reduction of Social Security benefits. If private accounts are initiated, it is possible that workers will demand defined benefit pensions to counterbalance the increased variability in returns in the new Social Security system. On the other hand, if Social Security benefits are cut, pension plans using the offset method would be hardest hit, since without modifications they would automatically make up some of the lost Social Security benefits. This would probably cause a further move away from offset to excess plans or possibly away from integration altogether. If the latter is the case, then it could even cause firms to stop offering pensions, since one of the justifications of allowing integration is to permit firms to recoup some of the Social Security tax as an incentive to offer a pension.

Another study has examined the potential effects of change from the employer's point of view. The ERISA Industry Group (1998) has come to several conclusions about the effects of changing Social Security on employer pensions.[8] First, most pension plans take into account Social Security benefits, even if they are not explicitly used to determine pension benefits. Implicitly, plans have lower benefits than if there were no Social Security retirement program. The implication of this is that pension benefits (whether formally integrated or not) and firms that offer pensions will likely be affected by changes in Social Security. Second, firms need time to adjust to any changes in the Social Security system, so substantial lead times are needed to minimize transition costs for firms. Third, change in the Social Security system may increase administrative costs for firms. Besides more direct costs of increasing pension benefits, a transitional system with individual retirement accounts would involve higher administrative costs for firms because they may need to design pension plans for workers under the existing Social Security system, under the transitional system, and under the permanent system. In addition, some proposals suggest that firms direct the collection and distribution of funds in privatized, individual accounts,

which would also increase administrative costs. Finally, the imposition of means testing Social Security benefits would make employer pensions less enticing to high income workers because pension benefits might cause a reduction in Social Security benefits. If pensions are used as an incentive to recruit and retain highly skilled and paid workers, means testing could lead to some firms finding alternative means of retention and possibly not offering pension plans. The key implication of some of these assertions is that, depending on the nature of changes in Social Security, there could be less incentive for employers to offer an integrated or, possibly, any pension plan.

III. The Health and Retirement Survey Data

The Health and Retirement Study (HRS) is a nationally representative panel survey of individuals who were aged 51 to 61 in 1992. In addition to collecting information on demographic, financial, and employment characteristics of individuals, the initial wave of the HRS in 1992 also collected detailed pension plan characteristics for covered respondents from their firms. Data in this Pension Provider Survey (PPS) were collected from the Summary Plan Descriptions of the employers' pension plans or from official records on the plans held by the U.S. Department of Labor. The data identifies integration provisions and benefit formulae.[9]

There is potentially two sets of pension information for each respondent in the HRS. If the respondents are currently working, they are asked if they are covered by an employer pension;[10] if the respondent is not currently working, the survey asks if they had a pension in their most recent job. In addition, both sets of respondents are asked about job and pension characteristics of a previous job that they held for at least 5 years. This study uses information about the current or most recent job. The information about the current or most recent job. The information on previous jobs is not used, however, for two reasons. First, for previous jobs, the rate of missing values for the pension (and indeed most job-specific) data was higher than for current or most recent jobs. Second and more importantly, given the emphasis in this article on job-specific covariates (industry, occupation, annual pay, and so forth), it is unclear how to characterize individuals who have had more than one job. For example, are they characterized by the pension and job attributes of the first job or second job? Since there is no clear answer for this question, information on previous jobs was not employed.

Since the article's sample is limited to individuals who are currently working or have worked, the sample members, hereafter, will be referred to as 'workers.' The sample also includes only individuals who are aged 51-61 in 1992, are civilian wage and salary workers, and have no missing observations for the key variables used in this study.[11] The reason for this latter condition is that multivariate regression estimation is used in the following section and therefore there can be no missing values in the data for the estimation. Finally, the results reported are based on HRS supplied sample weights, which adjust for nonresponse and the oversampling of minorities and individuals residing in Florida.[12]

A number of variables of interest from the HRS that are examined in the following analysis. The most important is whether a pension plan is integrated with Social Security. It is relatively easy to identify in the data whether a particular plan is integrated. A person covered by more than one pension in his or her job may have both an integrated plan and a nonintegrated plan;[13] in this study, such an individual is identified as being covered by an integrated plan. The set of covariates used here expands on Slusher's (1998) list of socio-demographic variables (such as gender, race, and education level) to include region of residence and, as suggested by Freeman (1985), union membership. Again, similar to Slusher (1998), this study examines the relationship between pension integration and firm size, annual earnings, occupation, and industry.

Other variables were derived as potentially interesting correlates of pension integration. The Social Security taxable maximum is important in the determination of the integration level of many integrated plans. Therefore, a variable was constructed to indicate whether the worker earned in the previous year (or the year prior to the last year that the respondent worked, if not currently working) at or above the taxable maximum. Given the advantage of relatively higher pension benefits in integrated DC or excess rate DB plans for these high earners, one might expect a positive correlation between being at the taxable maximum and integration. Full time status (annual hours greater than 1,500) might also be positively correlated with pension integration because full time workers are more likely to have a pension and earn more. The PPS also has a selection of pension plan characteristics. Other variables indicate whether the respondent has a DB plan only, a DC plan only, a hybrid plan,[14] more than one type of pension (at least one DB and one DC plan), and the number of pension plans on a given job. Finally, three variables indicate the projected level of benefits that individuals will get from their pensions. These include the private pension replacement ratio, the annual benefit, and the present value of the benefit.[15]

IV. Characteristics of Workers with Integrated Pension Plans

There are several ways of analyzing the correlates of integrated pension coverage. Two methodologies are used below. First, two tables examine the distribution of characteristics and the incidence of integration based on different subsamples of the data. This method is instructive in showing bivariate relationships between worker characteristics and pension integration. However, these simple cross-tabulations do not control for cross-correlation among the variables. Therefore, later in the section, the correlates of pension integration are examined using a multivariate regression methodology known as sample selection regression, which controls for the cross-correlation of the other variables in the study.[16]

Description of the Data Subsamples

A key issue in the analysis is to determine which subsamples to compare. Initially, one might consider examining the differences between those persons with integrated pensions and those with a pension that is not integrated (termed a 'standard' pension below). On the other hand, since the goal of this study is to determine how workers with integrated pensions differ from those who do not have integrated pensions, one should also consider workers who have no pensions. Therefore, to find if these three subsamples (individuals with no pension, with a standard pension, and with an integrated pension) are different, table 1 records the distributions of the dummy variables and the means of the continuous variables by the three different subgroups. Statistical measures (t-tests) are used to test for differences between the subsample of those with integrated pensions and the two other subsamples.

A comparison of the subsample with integrated pensions and of those without any pension reveals a big difference in the distribution of variables across the two subsamples. The distributions of gender, race/ethnicity, education, union status, and regions are generally significantly different between the two sectors with the integrated pension subsample being mostly males, white, more highly educated, and more unionized compared with workers without a pension. This pattern of significant differences continues when one examines the distribution of firm sizes (where the integrated group is more highly concentrated in larger firms), industries, and occupations. Annual pay, being at the taxable maximum, and full time status are significantly higher among those with an integrated pension.

Clearly these two subsamples are substantially different. However, having pension rights, rather than integration *per se*, may cause much of this difference. To isolate the integration effect, one can compare the distribution of the variables for the integration subsample and the standard pension subsample. Overall, there are fewer significant differences than for the comparison with the subsample of those not covered, although there are still quite a few statistically significant differences. Gender is not significantly different between these samples, indicating that the differences found above are due to pension coverage, rather than integration. The integrated pension subsample is more likely to be white, less likely to have a postgraduate degree, and less likely to be in a union compared with the subsample of standard pension holders. Likewise, they are more likely to be in larger firms and in manufacturing, sales, and service industries. While there is no significant difference in mean annual pay between the subsamples, the subsample of workers covered by an integrated pension are more likely to be above the taxable maximum and to be employed full time. Occupational classifications also exhibit significant differences. Because both subsamples contain information on pensions, comparisons can be made between several pension characteristics. The distribution of the type of plan (whether it is a DB plan, DC plan, and so forth) is not the same across the two subsamples. The integrated plan subsample is concentrated among those workers holding a DB only or both a DB and DC plan (which summed, constitute 94 percent of the subsample). Workers covered by integrated pensions also tend to have higher numbers of pensions and have higher replacement ratios. On the other hand, there are no significant differences between the subsamples for the level of annual benefits or in the present value of the pension benefits. Overall, the analysis shows that the three subsamples are very different, and therefore, in the subsequent analysis, all three subsamples are used to examine the incidence of pension integration.

Rates of Integration by Workers' Characteristics

The results in table 2 show the rates of integration for all workers and workers with pensions, by their characteristics.[17] The table also identifies differences in integration rates that are statistically significant. The null hypothesis is that there is no significant difference between the variable indicated by the symbol, '@,' and the other variables in that particular group. If the t-statistic is larger than the critical value, this null hypothesis will not be accepted, and the conclusion would be that there are significant differences in the rate of integration between these variables.

Integrated pensions are found in only 13.7 percent of the sample, although 32.2 percent of those covered by a pension have an integrated pensions. The latter percentage is substantially smaller than the 48 percent reported in the 1993 BLS data (the year of BLS data closest to the HRS). There are a couple of reasons for this. First, the BLS data on integration cover only DB pension plans, while the HRS data include not only DB, but also DC and hybrid plans, which are integrated much less frequently than DB plans. In addition, the BLS data are based on full time workers in medium to large firms (groups with relatively high pension coverage rates), whereas the HRS data includes all workers. When the HRS data are restricted to match the BLS characteristics, the incidence of integration for those who have pensions in this subsample increases to around 45 percent.

Socio-demographic characteristics.—Females have a slightly higher incidence of integration than males. This is true of both samples (all workers and the pension covered), although the gender differential is significant only in the subsample of all individuals. There is a pronounced racial and ethnic differential as well, which is indicated by the fact that whites are significantly more likely to have an integrated pension compared with blacks or Hispanics, regardless of the subsample. Different education levels also play a role in the incidence of integration. Those with a postgraduate education have a significantly lower incidence of integration in both groups, compared with other education levels. The rate of integration for union members is much higher than for nonmembers in the full sample, but this is primarily a pension effect, since among pension holders the rate of integration for union members is 10 percentage points less than the rate for nonmem-

bers. Finally, there is some regional variation. The rates of integration in either column for most regions are significantly different than the rate of integration in the Northeast. While it is not clear why regional variation exists, it may be due to cross-correlations with the other variables.

Firm size and industry.—Incidence of integration increases as firm size increases. However, while there is a statistically significant difference between "small" and "medium" size firms in integration rates for the full sample, there is no significant difference if one considers only workers covered by a pension. Rates of pension integration by industrial categories show that in the sample of all workers those in all other industries have significantly lower rates of integration, compared with individuals in manufacturing industries. However, when the sample is conditioned on pension coverage, sales and service industries have integration rates that do not differ significantly from that

Table 1.—The distribution and means of HRS variables, by pension and integration status

Variable	Integrated pension	Not covered by pension	Standard pension	Variable	Integrated pension	Not covered by pension	Standard pension
Demographic variables				*Firm size and industry (1 digit)*			
Female	45.6	55.1***	44.4	Below 100	2.0	36.6***	4.4***
Male	54.4	44.9***	55.6	Between 100-499	6.0	6.4	13.5***
White (non-Hispanic)	86.4	79.9***	83.0**	More than 499	81.2	19.5***	64.0***
Black (non-Hispanic)	8.3	10.2*	11.3**	Size missing	10.8	37.5***	18.1***
Other race (non-Hispanic)	2.2	2.7	2.1	Manufacturing	32.0	13.6***	17.3***
Hispanic	3.1	7.3***	3.8	Natural resources	3.8	13.0***	4.3
No high school diploma	33.9	43.4***	28.7***	Transportation	10.6	4.6***	12.0
High school diploma	39.3	39.2	37.1	Sales industries	10.2	22.8***	5.5***
Associate degree	4.4	3.7	3.5	Service industries	12.3	22.5***	6.1***
Bachelor degree	14.8	8.9***	13.0	Professional industries	24.5	21.3**	40.2***
Postgraduate education	7.6	4.8***	17.7***	Public administration	5.6	1.5***	13.9***
Not a union member	65.6	93.3***	54.4***	Missing industries	1.0	.7	.7
Union member	34.4	6.7***	45.6***				
Northeast	4.2	5.4	5.7	*Pay and hours data*			
Mid Atlantic	14.7	14.8	17.9**	Annual Pay(in thousands)	$35.0	$22.8***	$34.0
East North Central	14.6	14.6	17.1*	Below taxable maximum	83.3	91.9***	89.1***
West North Central	6.0	9.1***	8.9***	At or above taxable maximum	16.7	8.1***	10.9***
South Atlantic	22.3	20.6	18.3**	Annual hours<1500	7.5	30.0***	10.3**
East South Central	9.7	5.6***	5.1***	Annual hours>1500	92.5	70.0***	89.7**
West South Central	4.3	9.8***	9.7***				
Mountain	3.0	5.3***	5.7***	*Occupation (1 digit)*			
Pacific	21.2	14.7***	11.7***	Managerial	18.1	12.8***	16.5
Pension Data				Professional	17.3	9.2***	26.3***
DB plan only	48.9	0	52.4*	Sales	5.8	14.4***	3.2***
DC plan only	4.6	0	26.7***	Clerical	23.3	13.6***	18.7***
Hybrid plan only	1.9	0	2.4	Service	7.1	25.5***	11.8***
Both DC and DB plan	44.6		18.6***	Mechanics	5.7	2.7***	5.3
Number of pension plans#	1.5	0	1.2***	Construction	1.5	3.8***	2.6*
Replacement ratio#	47.7	0	44.3**	Precision tools	5.6	3.0***	3.0***
Annual benefit(in thousands)	18.8	0	17.3	Operations	15.5	14.6	12.3***
Present value of benefit				Missing	.1	.3	.2
(in thousands)#	$14.40	0	14.4	Number of observations (unweighted)	872	4,008	1,968

Note: Data are from the 1992 wave of the HRS and based on sample weights. "Integration" indicates that at least one of the individual's pension is integrated. Except for the continuous variables, denoted by a '#', all variables are indicator (dummy) variables and are recorded as percentages.
*,**, and *** indicate 10 percent, 5 percent, and 1 percent statistical significance of the difference between the integrated subsample and each of the other two subsamples for the particular variable.

for manufacturing. All the others have significantly lower levels of integration.

Pay and hours.— Compared with workers without an integrated pension, those with integrated pensions earn substantially more on average (approximately $35,100 compared to $26,500 in 1992). This is primarily due to the fact that part timers, who tend not to be covered by pensions, are a significant proportion of the subsample of all individuals. Once the sample is conditioned on pension coverage, the difference drops dramatically and becomes insignificant; those who are covered by non-integrated pensions tend to earn around $34,000 per year. Individuals who earn or once earned at or above the taxable maximum have higher rates of integration, compared with those with earnings below the taxable maximum. In fact, over 40 percent of the pension covered, high earners have an integrated pension. Full time status is also an indicator of higher integration in both subsamples.

Occupation.—In the overall sample, professional, sales service, construction, and operations occupations have significantly lower rates of integration compared with managers,

Table 2.—The distribution and means of HRS variables, by pension and integration status

Variable	All individuals	Pension holders	Variable	All individuals	Pension holders
Integrated plan	13.7	32.2	*Firm size and industry (1 digit)—Continued*		
Sociodemographic			Manufacturing	25.5	46.8
Female	15.1	32.7	Natural resources	5.6***	29.4***
Male	12.3***	31.7	Transportation	19.2***	29.5***
White (nonHispanic)	14.5	33.1	Sales industries	8.7***	46.7
Black (nonHispanic)	11.1***	25.8***	Service industries	10.3***	48.8
Other race (nonHispanic)	12.0***	32.8	Professional industries	12.4***	22.4***
Hispanic	7.6***	28.8***	Public administration	13.5***	16.0***
No high school diploma	12.3	35.9***	Missing industries	19.0***	39.1***
High school diploma	13.9***	33.4***			
Associate degree	16.4***	37.8***	*Pay and hours data*		
Bachelor degree	18.5***	34.9***	Below taxable maximum	12.7	30.7
Postgraduate education	11.8	17.0	At or above taxable maximum	22.7***	42.1***
Not a union member	11.5	36.4	Annual hours<1500	4.8	25.7
Union member	21.7***	26.3***	Annual hours>1500	16.1***	32.8***
Northeast	10.8	25.9			
Mid Atlantic	12.9***	28.1*	*Occupation (1 digit)*		
East North Central	13.0***	28.7**	Managerial	17.0	34.2
West North Central	9.6**	24.5	Professional	15.5**	23.7***
South Atlantic	15.1***	36.5***	Sales	8.0***	45.9***
East South Central	21.9***	47.3***	Clerical	19.4***	37.1**
West South Central	6.5***	17.3***	Service occupations	5.1***	22.1***
Mountain	8.1***	20.3***	Mechanics	20.2***	34.0
Pacific	19.7***	46.3***	Construction	6.4***	21.0***
			Precision tools	22.9***	47.2***
Firm size and industry (1 digit)			Operations	15.2***	37.5***
Below 100	1.2	17.7	Missing occupation	6.6***	22.3***
Between 100-499	9.8***	17.5	Number of observations		
More than 499	27.3***	37.5***	(unweighted)	6,848	2,840
Size missing	5.2***	22.0***			

Note: Data are from the 1992 wave of the HRS and based on sample weights. *, **, *** indicate a 10 percent, 5 percent, and 1 percent statistical significance for the bivariate t-test of whether the particular dummy variable is indicated by "@" are significantly different. for annual pay, the average for those with an integrated pension is 35.1, while the mean pay level for anyone in the sample without an integrated pension is 26.5 and the mean level of pay for those with a standard pension in 34.0. The difference between the first two averages is statistically significant at the 1 percent level, while the difference in income between standard and integrated pension holders is not statistically significant. All other continuous variables are not reported here, as they pertain only to workers with a pension.

while the other occupations have higher rates. The same pattern is shown in the pension only group (except for the mechanics occupation, which is insignificantly different from managers in the incidence of integration) and sales and operations, which now have significantly higher rates of integration.

Regression results[18]

The bivariate analysis in the previous tables does not control for the other variables when identifying a statistical relationship between the incidence of integration and a particular characteristic. While there are several potential methodological candidates, an alternative methodology that does control for the cross-correlations of other variables is a multivariate regression technique called sample selection regression. This type of analysis can incorporate a dual decision-making process that can model the determination of pension integration status. First, workers decide whether or not to take a job with a pension. Then, assuming the worker decides to take a job with a pension, the worker decides whether to take a job, or the firm decides to offer a job, with an integrated or a nonintegrated pension. Therefore, a bivariate probit regression is run. One equation of this regression incorporates the choice of obtaining a job with a pension. The dependent variable in this equation is whether or not the individual has a pension. The second equation of the regression models the choice of obtaining a job with an integrated pension among the sample of individuals with pensions, taking into account the sample selection of the pension sample. The dependent variable in this second equation is whether or not the individual has an integrated pension.[19] The regression uses the full sample and allows for correlations across the two equations.

Table 3 records the results from the regressions. Two points should be noted regarding the interpretation of this table. First, because bivariate probit regression is used in the estimation procedure, the regression coefficients are not readily interpretable. As a result, the regression coefficients have been transformed into marginal probabilities. Therefore, for a change in a particular independent variable (assuming other variables at their mean values), the marginal probability is the associated increase or decrease in the probability of having a pension or an integrated pension, after controlling for all other variables. Second, the table presents three sets of results. The first set contains the results from the pension coverage equation, while the second set records the integration equation results. The final set sums the results from the two equations to find the full effect of a variable on pension integration.[20] For instance, the coefficient on the male indicator variable (column 1) shows that men have a 0.59 percentage point greater probability of having a pension than women.[21] However, once the subsample is conditioned on having a pension (column 2), men are 5.2 percentage points less likely than women to have an integrated pension.

Among all workers (column 3), males are 4.6 percentage points less likely than females to have integrated pensions, taking into account differing rates of pension coverage.[22] Interestingly, while both the coefficients on the male variable are statistically significant, the combination of the two is significant only at the 90 percent level of confidence (since the two effects work in opposite directions). This indicates that we are only marginally sure of the point estimate of the difference in the rates of integration between the genders.[23]

Socio-demographic characteristics.—Compared with whites, black workers are significantly less likely to have an integrated pension, both among workers covered by pensions (6 percentage points) and among all workers (6.4 percentage points). While other races and Hispanics are less likely to have a pension (although the difference is small), they do not differ statistically from whites regarding pension integration. Postgraduates degree holders are a little more likely to have a pension than individuals with other educational degrees. However, if those without a postgraduate degree do have a pension, they are significantly more likely to have an integrated pension. This effect dominates, so that for the combined effect, the less educated people are more likely to have an integrated pension than those with a postgraduates education; these differences in probabilities are relatively large and range from 15.7 to 23.9 percentage points.[24] As expected from the findings in Freeman (1985), union membership significantly increases the likelihood that individuals have pensions by 2.3 percentage points, but for those with a pension, it decreases the likelihood of having an integrated pension by 15.9 percentage points. Combining these probabilities leads to a decreased likelihood of having an integrated pension by 13.6 percentage points for union covered individuals. Lastly, after controlling for all the other variables, there is some residual regional variation. Compared to the northeast region, workers in the west south central and mountain regions are 19.7 and 16.6 percentage points less likely to have an integrated pension, while workers in the Pacific region are 14.1 percentage points more likely to have an integrated pension.

Firm size and industry.—There is a positive correlation between firm size and pension coverage, there is no relationship between pension integration and firm size once the sample is conditioned on a person having a pension nor are the combined coefficients significant.[25] There is significant variation within the set of industry indicators. Once the sample is conditioned on having a pension, compared with people in manufacturing, workers in natural resources, transportation, professional, and public administration categories have smaller probabilities of having integrated pensions (ranging from 13 to 39 percentage points).

Pay and hours.—By far the most important variable in terms of magnitude and significance in this subset is the taxable maximum indicator. While it has no effect on the rate of pension coverage, it does increase the likelihood of integration for pension holders by 11.1 percentage points. Based on the combined probabilities, an individual with earnings at or above

the taxable maximum is 10.9 percentage points more likely than an individual below the taxable maximum to have an integrated pension. There is a weaker relationship between annual pay and integration. An increase in annual pay of $1,000 results in an extremely small, but statistically significant, increase in the probability of having a pension. For workers covered by a pension, an additional $1,000 in annual pay increases the probability of having an integrated pension by 0.2 percentage points. The combined effect is also 0.2 percentage points. Full time status (annual hours $\geq$ 1,500) increases the probability of having a pension, although it has no significant effect on pension.

Pension characteristics.—Among those covered by a pension, the type of plan is an important correlate of integration status. Compared to participating in a single DB plan, individuals with only a DC plan are nearly 50 percentage points less likely to have an integrated plan. On the other hand, having a hybrid plan or both a DB and a DC plan does not significantly change the probability of having an integrated pension compared with having only a DB plan. Furthermore, the more pension plans that people have, the more likely they are to have one that is integrated. Replacement rates and annual benefits have little or no statistical relationship with pension integration. The higher pension wealth (as measured by the present value of annual benefits) is, the less likely the worker is to have an integrated plan. Since the benefits are summed across all plans that a worker has, this result implies that those with integrated pensions tend to receive less in total benefits compared with workers who do not have an integrated pension.

Table 3.—Marginal effects from sample selection regression results on the incidence of pension integration

Variable	Pension effect (1)	Integration effect (2)	Combined effect (3)	Variable	Pension effect (1)	Integration effect (2)	Combined effect (3)
Sociodemographic variables				*Firm size and industry (1 digit)*			
Male	0.006***	-0.052**	-0.046*	Natural resources	-0.007***	-0.134***	-0.141***
White (non-Hispanic)	.004***	-.060**	-.064**	Transportation	.003	-.185***	-.183***
Black (non-Hispanic)	-.013***	.039	.026	Sales industries	-.011***	.066	.055
Other race (non-Hispanic)	-.009***	-.035	-.043	Service industries	-.009***	.019	.011
Hispanic	-.012***	.208***	.197***	Professional industries	.076***	-.211***	-.203***
No high school diploma	.008***	.179***	.172***	Public administration	.019***	-.389***	-.370***
High school diploma	-.010***	.249***	.239***	Missing industries	.012	-.072	-.071
Associate degree	-.007***	.164***	.157***				
Bachelor degree	.023***	-.159***	-.136***	*Pay and hours data*			
Union member	.002	-.055	-.053	At or above taxable maximum	-0.019	0.111***	0.109***
Mid Atlantic	.003	-.100**	-.096*	Annual pay (in thousands)	7.6E-5***	.019**	.020**
East North Central	.006**	-.121**	-.115*	Annual hours>1500	.017***	-.005	.012
West North Central	.002	.021	.024				
South Atlantic	.007**	.066	.073	*Pension data*			
East South Central	.004	-.200***	-.197***	Has DC plan only	(1)	-0.499***	-0.499***
West South Central	.001	-.166**	-.166**	Hybrid plan only	(1)	.007	.007
Mountain	-.001	.143***	.141***	Both DC and DB plan	(1)	-.014	-.014
Pacific				Number of pension plans	(1)	.134***	.134***
Firm size and industry (1 digit)				Replacement ratio	(1)	.008*	.008*
Between 100-499	0.035***	0-.057	-0.022	Annual benefit (in thousands)	(1)	.012	.012
Above 499	.045***	.020	.065	Present value of benefit	(1)		
Size missing	.019***	-.008	.011	(in thousands)	(1)	-.054***	-.054***
				Number of observations (unweighted)	6,848	2,840	6,848

Note: Data are from the 1992 wave of the HRS and are based on sample weights. Excluded categories are: female, white race, postgraduate education, non-union, northeast region, firm size less than 100, manufacturing industry, has a DB plan only (only in the integration equation), and managerial occupation (only in the pension equation). Constant terms were also included in each of the regressions but are not reported. Occupational controls were included in the pension determination regression but are not reported. Log likelihood = -3998.869. *, **, and *** indicate 10 percent, 5 perceent, and 1 percent statistical significance, respectively. The estimated error correlation, ρ, is a -0.108, which has a t-statistic of -0.511 and is not statistically significant at conventional levels. See text and technical appendix for further details.

[1]Not included in the equation.

V. Conclusions

Until recently, there has been little research on employer pensions that are integrated with Social Security benefits, and what little research existed used data collected from samples of firms. This study is one of the first to examine pension integration at the individual level, using survey and pension plan data from the 1992 Health and Retirement Study. Socio-demographic, pension, earnings, and industry characteristics were examined to determine the correlates of pension integration. It is likely that changes in the Social Security system will affect integrated pension plans covered by them. The regression results presented here show that females, whites, workers with less than a postgraduate education, and those not in unions are more likely to have integrated pensions. Having higher earnings or a defined benefit pension plan are also positively correlated with pension integration. There is also some variation in pension integration across regions and industries, with workers in the Pacific region and in manufacturing being more likely to be covered by an integrated pension.

Notes

[1]The percentage of the total working population covered by integrated plans is much lower. See Section III for a more thorough discussion.

[2]For the most detailed analysis of pension integration, see Schulz and Leavitt (1983). For other early analyses of integration see Schiller and Snyder (1982) and Bell and Hill (1984); for a more current review, see Kollmann and Schmitt (1994), McGill, *et al.* (1996), and Employee Benefit Research Institute (1997).

[3]Besides the empirical studies mentioned in this subsection, one study examines the theoretical reasons for pension integration. Merton, *et al.* (1987) argue that since the benefits from integrated pensions are negatively correlated with Social Security benefits, employers are insuring their employees against adverse changes in Social Security benefits. This point is also mentioned, but not rigorously examined, in Gustman, *et al.* (1994).

[4]There are no published data on the prevalence of integration in defined contribution plans. However, the Health and Retirement Survey (HRS) data used below indicate that in 1992, the rate of integration of those who have just a defined contribution pension is 7.6 percent.

[5]It is important to note that this number is somewhat misleading in terms of overall coverage. Since only 52 percent of full-time workers in these firms had a DB pension plan in 1995, only 26.5 percent of full-time workers in these firms have an integrated DB plan. The percentage would be even lower if one considered the entire working population, which would include government workers, part time workers, and those employed by small firms.

[6]On the other hand, Social Security benefits for federal retirees under the Civil Service Retirement System pension plan may be reduced by a proportion of their pension plan benefits.

[7]While the 1992 HRS data does have information on pension integration, the other waves of the panel do not have current information on integration because the HRS collected detailed pension information only in 1992. However, the HRS plans to do a follow-up survey of pension plans after the 1998 wave.

[8]Gregory (1998) offers a summary of the findings from the ERISA report. Crenshaw (1998) also references the ERISA report as well as examines other aspects of the effects of change on Social Security on firms.

[9]Since the new regulations were not enforced until January of 1994, the integration provisions recorded in the 1992 PPS comply with the pre-TRA86 law but not necessarily with the TRA86 legislation (Kollmann and Schmitt 1994).

[10]This is found in question F37 in the HRS data documentation where the respondent is asked, 'Are you included in any … pension, retirement, or tax-deferred plan with [your] employer?' In addition, for those not currently working, some are actually receiving pensions from their most recent jobs.

[11]The data from the HRS employed here differ in two ways from the data used by Slusher (1998). First, Slusher uses the job and pension information from previous jobs, as well as from current or most recent jobs. Second, he does not exclude observations when they have missing values for some of the variables of interest. Slusher can do this because his analysis uses cross tabulations rather than regression analysis. Naturally, by excluding these observations with missing variables, this study is based on a sample that is smaller than the one employed by Slusher (1998).

[12]See the Technical Appendix at the end of this study for more discussion of the data.

[13]By law, an employee with multiple pension plans from the same employer can only have one of those plans integrated.

[14]Hybrid plans are pensions which have characteristics of both defined benefit and defined contributions plans.

[15]These three variables are calculated using the HRS/University of Michigan Pension Calculator, Version 4, which uses pension plan characteristics, estimated earnings histories, and demographic information to calculate replacement rates, annual pension benefits, and the total present discounted value of those benefits. For those who have not yet retired, it was assumed that these workers would retire at age 65. If an individual had more than one plan from a single employer, the benefits were summed across plans.

[16]There are potentially several ways of statistically modeling pension integration using multivariate regression techniques. See the Technical Appendix for further discussion of these competing methodologies.

[17]Note that there are no results recorded from the pension variables since they are defined only for workers with pensions and are therefore not applicable to the entire sample.

[18]Note that in the results below, there will be no discussion of the occupational variables. See the Technical Appendix for the reasons for this.

[19]See the Technical Appendix for a discussion of methodology and the estimation procedure.

[20]See Green (1997, p. 901ff) for an explanation of this methodology and the calculation of marginal probabilities in bivariate probit regression models.

[21]Whenever there is a group of mutually exclusive and exhaustive set of indicators in this (or, indeed, any multivariate) regression, one of the group of indicator variables has to be left out of the regression to alleviate multi-collinearity. All of the coefficients in a group of these variables are interpreted in relation to the excluded indicator variable in that group. Note that for comparison with previous

results, the excluded variables are the ones that formed the basis for the bivariate comparisons in table 2.

[22]Note that this result is different from that found in Slusher (1998) who finds that males are more likely to have an integrated pension. The differences are likely due to the different samples, as explained above, and the use of a regression methodology that controls for cross-correlation between variables.

[23]A one-tailed test of this coefficient, however, shows that we are at least 95 percent certain of the negative sign of the difference, even though we cannot be sure of the exact size of the estimated difference.

[24]This result may seem to be counterintuitive because of the previously hypothesized positive relationship between earnings and integration. However, given that earnings are controlled for in the regression, there is no economic theory about the relationship between pension integration and educational attainment, *ceteris paribus*. This is an area of future research.

[25]This is in direct contrast to Parsons (1994), who finds some evidence that large firms offer integrated pensions because of economies of scale. However, his analysis is restricted to bivariate comparisons (like in table 2 in this article, which also suggests a positive relationship between firm size and the incidence of pension integration).

Bibliography

Bell, Donald and Diane Hill. 1984. "How Social Security Payments Affect Private Pensions," *Monthly Labor Review*, Vol. 107, No. 4 (May), pp. 15-20.

Crenshaw, Albert B. 1998. "Social Security Overhaul Could Jar Private Plans," *The Washington Post*, (August 16) Section H, pp. 1-2.

Dyer, John K. 1977. "Concept of Pension-Social Security Integration." In *Social Security and Private Pension Plans: Competitive or Complementary?* Dan M. McGill (ed.). Pension Research Council, Philadelphia: University of Pennsylvania Press, pp. 123-33.

Employee Benefit Research Institute. 1997. *Fundamentals of Employee Benefit Programs*. Fifth Edition. Washington, DC: Employee Benefit Research Institute.

ERISA Industry Committee. 1998. *The Vital Connection: An Analysis of the Impact of Social Security Reform on Employer-Sponsored Retirement Plans*. Washington, DC: ERISA Industry Committee.

Foster, Ann C. 1997. "Public and Private Sector Defined Benefit Pensions: A Comparison," *Compensation and Working Conditions*, (Summer), pp. 37-43.

Freeman, Richard B. 1985. "Unions, Pensions, and Union Pension Funds." In *Pensions, Labor, and Individual Choice*. David A. Wise (ed.). Chicago: University of Chicago Press, pp. 89-121.

Graham, Avy D. 1994. "Coordinating Private Pension Benefits with Social Security," *Monthly Labor Review*, Vol. 117, No. 3 (March), pp. 35-38.

Greene, William H. 1997. *Econometric Analysis*. Third Edition. New York: Macmillan Publishing Company.

Gregory, Janice M. 1998. "Social Security Reform: The Facts from the Employer's Viewpoint," *Compensation and Benefits Review*, (Sept/Oct), pp. 56-59.

Gustman, Alan L.; Olivia S. Mitchell; Andrew A. Samwick; and Thomas L. Steinmeier. 1998. "Evaluating Pension Entitlements." Paper presented at April 1998 Pension Research Council Symposium.

Gustman, Alan L.; Olivia S. Mitchell; and Thomas L. Steinmeier. 1994. "The Role of Pensions in the Labor Market: A Survey of the Literature," *Industrial and Labor Relations Review*, Vol. 47, No. 3 (April), pp. 417-38.

Kollmann, Geoffrey and Ray Schmitt. 1994. "Effect of Pension Integration on Retirement Benefits." *Congressional Research Service Report for Congress, 94-974*. Washington, DC: Congressional Research Service.

McGill, Dan M.; Kyle N. Brown; John J. Haley; and Sylvester J. Schieber. 1997. *Fundamentals of Private Pensions*. Seventh Edition. Philadelphia: Pension Research Council and University of Pennsylvania Press.

Merton, Robert C.; Zvi Bodie; and Alan J. Marcus. 1987. "Pension Plan Integration as Insurance Against Social Security Risk." In *Issues in Pension Economics*. Zvi Bodie, John B. Shoven, and David A. Wise (eds.). Chicago: University of Chicago Press, pp. 147-69.

Parsons, Donald O. 1994. "Recent Trends in Pension Coverage Rates." In *Pension Coverage Issues for the '90s*. Richard P. Hinz, John A. Turner, and Phyllis A. Fernandez (eds.). Washington, DC: U.S. Department of Labor, pp. 39-52.

Schiller, Bradley R. and Donald C. Snyder. 1982. *Linkages between Private Pensions and Social Security Reform: An Information Paper*, Special Committee on Aging, United States Senate, Washington, DC: GPO.

Schulz, James H. and Thomas D. Leavitt. 1983. *Pension Integration: Concepts, Issues and Proposals*. Washington, DC: Employee Benefit Research Institute.

Slusher, Chuck. 1998. "Pension Integration and Social Security Reform," *Social Security Bulletin*, Vol. 61, No. 3, pp. 20-7.

Wiatrowski, William J. 1991. "New Survey Data on Pension Benefits," *Monthly Labor Review*, Vol. 114, No. 8 (August), pp. 8-22.

Technical Appendix

This section describes some of the more technical aspects of the analysis. The first subsection details the sample from the HRS dataset that is used in this study and examines the potential problems of missing employer pension data. The second subsection describes, in more detail than the text, the regression technique used in Section III of the study.

Data Issues

Some discussion of the sample needs to be made. Appendix table A1 summarizes how the sample size of the dataset was determined. Of the 9,825 age-eligible respondents in the HRS, the sample was first limited to those who were current workers

or had worked and were not self-employed (for whom no pension information was collected). Further exclusions were those in the armed forces (since none have an integrated pension), those with negative or unusually large (over 200 percent) replacement ratios, and those with negative or zero annual pay. While this sample of 8,439 respondents would be the optimal dataset, there are a number of workers with pension coverage who have no detailed, employer supplied information on pension characteristics. Therefore, the final cut excluded these observations. The resulting dataset includes 6,848 observations, of which 2,840 have pension coverage with the detailed pension information.

A potentially serious problem concerns this missing pension information. Even though 4,431 HRS respondents in the 8,439 person sample reported being covered by a pension, only 2,840 (64 percent) have the detailed pension plan data. Because there is no information from the employer on these individuals, and therefore no information on whether they are covered by integrated pensions, they cannot be included in the analysis. However, if these individuals differ in nonrandom ways from the sample of pension holders for whom there is information, these differences could lead to biased results in either the analysis of means or in the regression results.

In order to get a handle on the problem, the variable means are compared between the two samples: the pension covered with plan information and the pension covered with no plan information. These means are found in Appendix table A2. This table gives an indication of whether the observed characteristics are the same across the two samples. As can be seen,

Table A1.—Determination of HRS sample size

Dataset or restriction	Sample size
Full age eligible sample	9,825
Nonself-employed (current or recent workers)	9,029
Cutting armed forces occupation	9,013
Cutting replacement ratios<0 or >200	8,911
Cutting negative or zero earnings	8,439
Detailed pension data	6,848
Pension covered	2,840

Note: Sample numbers are unweighted.

Table A2.—Distribution and means of variables of persons who report having a pension plan

Variable	Plan information	No plan information	Variable	Plan information	No plan information
Socio-demographic variables			*Firm size and industry (1 digit)*		
Male	0.552	0.576	Below 100	0.037	0.139
White (non-Hispanic)	.840	.844	100-499	.112	.123
Black (non-Hispanic)	.104	.086	Above 499	.693	.469
Other race (non-Hispanic)	.021	.022	Firm size missing	.159	.269
Hispanic	.036	.049	Natural resources	.041	.088
No high school diploma	.303	.361	Manufacturing	.218	.297
High school diploma	.378	.404	Transportation	.116	.081
Associate degree	.038	.050	Sales industries	.069	.142
Bachelor degree	.136	.106	Service industries	.080	.125
Postgraduate education	.146	.078	Professional industries	.354	.242
Union member	.422	.309	Public administration	.114	.018
Northeast	.052	.077	Missing industries	.008	.008
Mid Atlantic	.169	.183	*Occupation (1 digit)*		
East North Central	.163	.168	Managerial	.170	.179
West North Central	.080	.102	Professional	.235	.154
South Atlantic	.195	.184	Sales	.040	.088
East South Central	.065	.046	Clerical	.201	.173
West South Central	.080	.061	Service occupations	.104	.071
Mountain	.049	.039	Mechanics	.054	.034
Pacific	.146	.140	Construction	.023	.049
Pay Data			Precision	.038	.046
Annual pay (in thousands)	34.331	58.069	Operations	.133	.204
At or above taxable maximum	.127	.131	Missing occupation	.002	.002
Annual hours>1500	.906	.926	Number of observations (unweighted):	2,840	1,591

Note: Data are from the 1992 wave of the Health and Retirement Survey and based on HRS sample weights. After excluding annual earnings values of a million dollars or greater (resulting in the loss of 4 observations), the mean annual earnings of the "no plan information" subsample is $37,100.

the two samples have relatively the same means for many of their common variables, with differences of only a couple of percentage points. A strikingly big difference is between earnings levels. However, the large average pay for individuals with no plan information is caused by the inclusion of four observations with extremely large earnings (a million dollars or more). After dropping these observations, the means are much more comparable, as reported in the note to the table. Individuals with no plan information also tend to be from smaller sized firms.[1] The biggest differences come in the manufacturing, sales, professional, and public administration industries, and in professional and operations occupations. However, the overall majority of observed variables have similar means, which may indicate that there is relatively little sample selection bias, although this analysis cannot rule out this possibility.

Modeling Pension Integration

As mentioned in the text, there are several different ways of investigating the correlates of pension integration. The simplest method is using probit (or logit) analysis, examining whether the respondent has an integrated pension or not. Unfortunately, this forces the selection of one of two comparison groups: comparing pension holders with and without integrated pensions and comparing those with integrated pensions and everyone else. The first has the disadvantage of ignoring those not covered by a pension, while the second cannot distinguish between the effects of integration and coverage.

Another contending methodology is the multinomial logit. Here, the dependent variable takes on three discrete values, one each corresponding to having no pension, having a standard pension, and having an integrated pension. This type of modeling allows for all the data to be analyzed, not just the sample with a pension. The potential problem with the multinomial logit is that the results are not easily interpretable since the generated coefficients are the relative effects of one outcome compared with another. For example, if those with no pensions were the comparison group, the coefficients for individuals with standard pensions and with integrated pensions would be in relation to those with no pension. The interpretation of these coefficients is made even more complex due to the large number of dummy variables that are included in the regressions, since the coefficients are also interpreted in relation to the excluded variable.

A methodology that avoids some of these interpretational issues is the sample selection modeling utilized in this article. This model avoids the use of having to omit one of the outcomes in the regression analysis. The methodology involves estimating two probit regressions, one which describes the pension choice while the other describes whether the pension is integrated. If the two regressions were estimated separately, there would be a good argument to be made that the pension integration equation, which uses the sample of pension holders only, might have biased coefficients due to the fact that covered workers are somehow nonrandomly different from workers with no pensions. Therefore, in the estimation procedure, a bivariate probit model (two probit regressions with correlated errors) is estimated, adjusting for sample selection in the integration equation (for more details, see Greene 1997). This procedure includes estimating the following two equations:

$$P_i = X_i^p \beta^p + \varepsilon_i^p \text{ and } I_i = X_i^I \beta^I + \varepsilon_i^I$$

where for individual i, P_i is an indicator variable showing whether the individual has a pension or not, I_i is an indicator variable showing whether the individual has an integrated pension (conditional on having a pension), X_i^p and X_i^I are vectors of correlates for the two equations, which may or may not include the same variables, β^p and β^I are coefficients to be estimated, and ε_i^p and ε_i^I are correlated errors where the estimated correlation of these errors is ρ.[2] The equations are estimated using full information, maximum likelihood on the full sample. Finally, marginal effects were calculated from the estimated coefficients based on the methodology found in Greene (1997, pp. 906-911).

One final issue with the sample selection estimation procedure concerns the variable specification of the model. To statistically identify the model, there must be variables in at least one equation that do not appear in the other equation. These exclusion restrictions are in fact found in both equations. In the case of the pension determination equation, it is easy to see that the pension characteristics variables are the natural choice to exclude, since any respondent without a pension will not have any pension characteristics. A more difficult choice is which variables to exclude from the integration equation. After some investigation it was found that when the occupational controls were included in the integration equation *without* the union status indicator, several were statistically significant. However, once the union variable was added, none of the occupational characteristics individually or as a group were significant. A theoretical reason for this finding is that the occupational indicators were proxying for union status. Given that pension integration is a firm (or firm and union) choice, there is no theoretical reason why, once we control for unionization, occupations should be correlated with integration. Therefore, these occupational indicators were excluded from the pension integration equation.

Note

[1]See Gustman, *et al*. (1998, footnote 1) for more details about linkages rates by firm size for the HRS detailed pension data.

[2]In table 3 of this article, the results regarding ρ are presented. As can be seen, it is statistically insignificant, indicating that there is no information to be gained by allowing the errors to be correlated. However, Green (1997) indicated other statistical tests that can be used to test for significance. Following the procedure for a likelihood ratio test, the calculated chi-squared statistic is 73.9, easily significant at the 1-percent level. Given this ambiguity, the author decided to follow the more general model of the bivariate probit.