-- Working Paper Do Not Cite Without Permission of The Authors--

Measuring the Effect of Disability on Worklife and Earnings

George M. McCollister and Christopher C. Pflaum
Spectrum Economics, Inc.

Introduction

Worklife is the number of years an individual is active in the labor force from a specified age until final departure from the workforce due to full retirement or death and includes periods during which a worker is unemployed but looking for work. Worklife is a key measure of individual productivity and a critical consideration in evaluating income support and retraining programs for the disabled as well as estimating economic loss resulting from injury or death.

The measurement of worklife is old in the art, the most widely accepted method being the increment – decrement model introduced by Smith in 1982 and refined and expanded in a series of articles by numerous authors.[1] The impact of disability on worklife, on the other hand, has been relatively unstudied.

In this paper, we first discuss the data problems that have impeded the development of worklife tables for the disabled. These problems are well known and due to shortcomings and biases of the widely used surveys, the *Current Population Survey* (CPS) and the *Survey of Income and Program Participation* (SIPP).[2] Next, we present examples of worklife and earnings models for the disabled that, we believe, overcome most objections raised to current approaches. We do this by applying enumeration and statistical modeling techniques to data from the *National Health Interview Survey* (NHIS), which includes detailed information on medical conditions and physical impairments and limitations. We explore the possibility that self selection into the labor force may bias the coefficients for disabilities in the earnings equation using Heckman's sample selection bias test to test for and remove this bias.

We find that disability affects both worklife and earnings and that the effect is very dependent on the level of education of the injured person and the work that they do. For the uneducated, a physical impairment is likely to have a profound impact on likelihood of working and earnings. For the educated, the effects are much smaller. The effect of disabilities on earnings is reduced and less statistically significant when the bias of self-selection is removed.

Previous Research

A threshold challenge to all studies of the effects of physical limitations on work is differentiating impairment from disability. An inability to lift heavy weights above one's head would be disabling for a construction worker but vocationally irrelevant for a computer programmer.

Studies of the effects of injury characteristics on return to work found in the vocational rehabilitation, spinal cord injury and traumatic injury literatures tend to use data with precise indicia of physical impairment while those found in the economics literature rely on self selection into a disabled category. These medically-focused data sets, however, have frequently been too small or specialized for statistical inference to the general population and the studies using them have not generally focused on the questions that most interest economists. Finally, studies of labor force reentry reported in the rehabilitation literature, while informative, do not provide guidance regarding the duration of post injury employment. This leaves unanswered the key question of the extent to which programs that train and place the physically impaired in jobs fully restore their human capital.

Economists have studied workforce participation and earnings of the disabled using data from the CPS and SIPP. Because the CPS and SIPP rely on self-reported rather than objectively measured limitations, and indeed the CPS is not designed to measure the effect of disability on employment[3], the results of these studies are suspect. As Kirchner has noted, there has been a tendency for researchers to let the available data drive the analysis, however inappropriate to the task the data may be.[4] The widely recognized problems with these data sets are reviewed in Currie and Madrian.[5]

There is an inherent bias towards over-reporting disability by those who do not work due to the moral hazard created by income support programs and the "justification hypothesis," the tendency of the unemployed to use the socially acceptable excuse of disability to rationalize not working. Two specific factors that create a bias toward over reporting disability are the loss of federally provided health insurance and the difficulty of reestablishing disabled status if the return to work is unsuccessful in the long run. Another problem with the classification of disability is that early retirement and participation in a disability program can be due to an occupational limitation rather than the inability to work (e.g. railroad employees, military). The payments under these programs establish a high "reservation wage" and provide a strong disincentive for workforce reentry in an alternative occupation for which an individual is qualified. This phenomena also causes the sample for estimation of earnings for the disabled to be censored and the resulting coefficients biased. Comparing earnings estimates with and without correction for sample selection bias, we find that transfer payments are a strong disincentive to work.

There is also a tendency by those who are employed to underreport disability. An employed individual adequately coping with a biomedical limitation may not consider himself "disabled" and report his condition as not disabled. Hence we have a problem of the denominator of the employment rate calculation inflated and the numerator deflated.

Forensic economists encounter the additional problem that the broad categories presented in worklife studies to date are of limited usefulness is assessing an individual. Specifically, factors found important in the rehabilitation literature such as cohort employment rates, immigration status, marital status and medical and vocational indicia of disability have not been considered in previous studies found in the economic literature.

The Data

The National Health Interview Survey (NHIS) is an annual, cross-sectional survey of approximately 100,000 individuals that has been conducted since 1957.[6] The primary purpose of the survey is to collect heath related data, such as chronic medical conditions, substance abuse, physical limitations and types of injuries suffered. The survey also collects demographic information including family structure, education, income, and employment status.

The medical data collected include both need for assistance in activities of daily living (ADL) and difficulty in performing standardized physical tasks such as those reported for occupations in the *Dictionary of Occupational Titles* and routinely measured and used by rehabilitation and vocational counselors such as sitting for two hours, walking four hundred yards, stooping, squatting, etc.. ADLs include bathing, dressing, toileting, moving, and walking. Examples of chronic health problems are body mass index (BMI), heart disease, chronic back pain, injuries and substance abuse.

The number of individuals surveyed each year is large and multiple years of data can be combined to provide a larger sample. These attributes make the NHIS, in our opinion, the best broad based data source for the study of the effects of physical impairment on work and earnings.

The NHIS data relevant to the study of disability should be less affected by the biases found in the CPS and it provides an expansive set of information regarding health and physical limitations[7] and does not use self-reported disabled status as a screening tool. However, we do not claim that the NHIS is free from the biases inherent in self reporting. Specifically, economists have long known that health problems are used to rationalize not working and this behavior biases surveys of self-reported medial ailments by those not in the labor force. Recently, in a study comparing self-reported disability with data from respondents medical records, Baker, Stabile and Deri found significant over reporting of disability in the Canadian National Population Health Survey, the Canadian equivalent of the NHIS.[8]

Survey questions about employment status during the previous week include work related activities performed, the number of hours worked or usually worked, reasons for not working, occupation and industry, and for persons who worked less than 35 hours, whether they usually worked fulltime. Possible reasons for not working include on layoff, retired, on a planned vacation or leave from work, going to school, off-season from a job or disabled. Though we have chosen to model employment, the question set permits specification of labor force participation and employment status consistent with BLS categories. Questions about employment status during the previous calendar year include the number of months the person had a job, and the amount of earnings. Data are also collected on spousal earnings and non-wage income and its sources.

We limited the data used in our analysis to that collected from 1997 to 2002 for adults between the ages of 25 and 65, who were not full time students, in the military, mentally retarded or institutionalized. In this group, there were 291,862 respondents. Males comprised 47.5% of the unweighted sample.

The weighted distribution of marital status for each gender and the percent working full time is shown in Table 1. Note that marriage increases the odds that a man will work, but has the opposite effect for woman.

Table 1
Fulltime Employment by Marital Status

	Males		Females	
	% of Total	Working	% of Total	Working
Married	69.9%	86%	66.7%	53%
Widowed	0.8%	58%	3.4%	42%
Divorced	10.0%	76%	13.1%	71%
Separated	2.2%	74%	3.3%	58%
Never Married	17.1%	76%	13.4%	69%
Total	100.0%	82%	100.0%	57%

The distribution and percent working by level of education in the weighted sample is presented in Table 2. Note that the odds of being employed increases with education. Also, a GED is not as helpful as a high school diploma in obtaining fulltime employment.

Table 2
Fulltime Employment by Education

	Males		Females	
Education	% of Total	Working	% of Total	Working
No Diploma	14.4%	69%	13.5%	37%
HS Diploma	27.2%	82%	27.8%	55%
GED	2.8%	73%	2.6%	37%
Some College	17.4%	83%	18.8%	60%
AA Technical	6.5%	87%	6.9%	63%
AA Academic	2.9%	87%	4.0%	64%
BA	18.3%	89%	17.9%	65%
MA, MBA	6.7%	87%	6.7%	70%
MD, DDS, JD	2.1%	90%	1.1%	68%
Ph.D.	1.6%	88%	0.7%	77%
Total	100.0%	82%	100.0%	57%

In Table 3 the percentage of the adult population afflicted with multiple physical impairments as measured by the number of standard tasks that are very difficult or impossible to perform is presented. The numbers are reported separately for adult males and females between the ages of 25 and 65. Note that females are much more likely to suffer from multiple physical limitations even though their jobs are typically less physically demanding. Of particular interest is that as the number of limitations increases, the percentage of female workers dropping out of the labor force is much smaller than for men. This is a reflection of the more physical nature of the work typically performed by men and the relatively less severe impact that disability has on those with sedentary occupations.

Table 3
Fulltime Employment by Number of Difficult Tasks

	% of Sample		Working Fulltime	
Tasks	Males	Females	Males	Females
0	92.3%	87.6%	86%	60%
1	2.8%	4.7%	62%	50%
2	1.2%	2.2%	43%	38%
3	0.9%	1.5%	30%	32%
4	0.7%	1.1%	23%	26%
5	0.7%	1.0%	20%	18%
6	0.7%	0.9%	12%	13%
7	0.5%	0.6%	13%	8%
8	0.3%	0.4%	6%	8%

In Table 4 we present the distribution of limitations and corresponding employment rates in the sample by age and education. Three facts are apparent with important implications for observing the effects of physical disabilities in the general population. First, physical limitations become more common as people age, are more common among those with less education and, as we saw previously, are more common among females than males. Second, the effects of impairment on employment rates are also greater in these same groups. Third, employment rates in the general population are also lower in these groups. Thus, physical limitations are more common and have a greater impact in groups that naturally have lower employment rates, even in the absence of any disability. Age, education and gender are all confounding factors to the effects of physical limitations on employment for which we must control when assessing their impact. Overlooking any of these factors will cause the impact of physical limitation on employment to be overstated.

Table 4
Fulltime Employment for ADLs and Standard Tasks

	% of Sample Afflicted							Working Fulltime						
		Age			Education				Age			Education		
	All	<46	46-55	55+	<HS	HS	BA+	All	<46	46-55	55+	<HS	HS	BA+
Need Help From Others to														
Bath	0.3%	0.2%	0.4%	0.8%	0.6%	0.3%	0.1%	7.0%	10.9%	6.3%	3.6%	2.5%	7.3%	21.3%
Dress	0.3%	0.2%	0.4%	0.7%	0.6%	0.3%	0.1%	9.0%	12.5%	8.1%	5.7%	2.4%	10.4%	20.4%
Eat	0.1%	0.1%	0.1%	0.2%	0.1%	0.1%	0.0%	8.7%	12.5%	5.1%	7.6%	0.7%	11.3%	16.4%
Get in or out of beds or chairs	0.3%	0.2%	0.3%	0.5%	0.5%	0.3%	0.1%	7.9%	10.8%	7.3%	4.7%	1.8%	9.3%	19.6%
Use the toilet	0.2%	0.1%	0.2%	0.4%	0.3%	0.2%	0.1%	6.3%	10.9%	4.7%	2.6%	0.3%	6.3%	22.6%
Move around the home	0.2%	0.1%	0.3%	0.5%	0.4%	0.2%	0.1%	9.2%	12.7%	8.2%	6.2%	1.3%	10.3%	26.3%
Very Difficult or Impossible to														
Walk 1/4 mile	4.2%	2.1%	5.7%	10.6%	9.5%	4.0%	1.4%	23.5%	30.9%	26.6%	14.6%	13.1%	27.3%	45.3%
Climb 10 steps	2.9%	1.4%	4.1%	7.5%	7.0%	2.8%	0.8%	18.9%	24.3%	22.4%	11.4%	9.2%	24.0%	37.9%
Stand for 2 hours	5.3%	2.8%	7.4%	12.4%	10.6%	5.3%	2.2%	25.7%	33.0%	29.8%	14.6%	11.6%	29.8%	47.6%
Sit for 2 hours	2.6%	1.7%	3.7%	4.8%	5.4%	2.6%	1.0%	28.9%	38.3%	28.4%	15.9%	14.5%	34.3%	49.7%
Stoop, bend or kneel	5.4%	2.9%	7.6%	12.4%	10.2%	5.5%	2.3%	32.3%	41.3%	36.8%	18.9%	16.6%	36.7%	54.4%
Reach over the head	1.7%	0.9%	2.4%	4.0%	3.9%	1.6%	0.5%	22.0%	33.7%	22.3%	11.3%	10.8%	27.0%	41.4%
Lift and carry 10 pounds	2.7%	1.5%	3.8%	6.1%	6.1%	2.6%	0.9%	18.1%	25.1%	19.4%	9.8%	7.9%	21.5%	41.3%
Pushing large objects	4.9%	2.6%	6.8%	11.1%	9.9%	4.9%	1.9%	24.3%	32.9%	28.0%	12.0%	10.8%	28.1%	48.2%
With none of the above	89.9%	94.0%	86.1%	78.5%	81.6%	89.6%	95.2%	73.4%	75.8%	77.7%	53.3%	63.1%	73.6%	78.4%

Both genders, ages 25 to 65.

Methods of Estimating Worklife

Worklife probabilities are usually estimated in either of two ways: conditional on the previous year's employment status or unconditionally. Conditional probabilities, also known as transition probabilities, are estimated using the increment-decrement method. Increment – decrement worklife can be difficult to estimate because, as data sources are subdivided based on previous employment status, the smaller samples are more likely to produce insignificant or aberrant results. Another problem is that conditional probabilities are very sensitive to economic conditions. For example, if the economy is entering a recession, the probability of transitioning from employed to unemployed states will be overstated. Consequently, conditional probabilities are often estimated by blending conditional and non-conditional results or using a multi-year period thought to either cover a business cycle or represent a "normal" economic period[9]. Since the NHIS does not provide employment status one-year prior, we can only estimate unconditional probabilities.

Enumeration Model of Worklife

The enumeration method is a standard method for computing life and worklife expectancy. The basic methodology is to collect percentages dying, in the workforce, etc. by age group at a point in time and then apply those percentages to a starting cohort of some assumed number.

In developing worklife estimates, enumeration works well for common injuries and occupations where there are large numbers of observations. Results from enumeration for different age groups are often smoothed with statistical regression so that the distinction between enumeration and regression-based methods is less than might be surmised.

An Example Using Enumeration

Back injury is one of the most common physical ailments effecting workers. It has been estimated that back injury costs employers in the United States $15 billion per year in lost worker time and worker compensation payments.[10] As a practical matter, chronic back pain is also associated with other physical ailments such as obesity and lack of fitness.

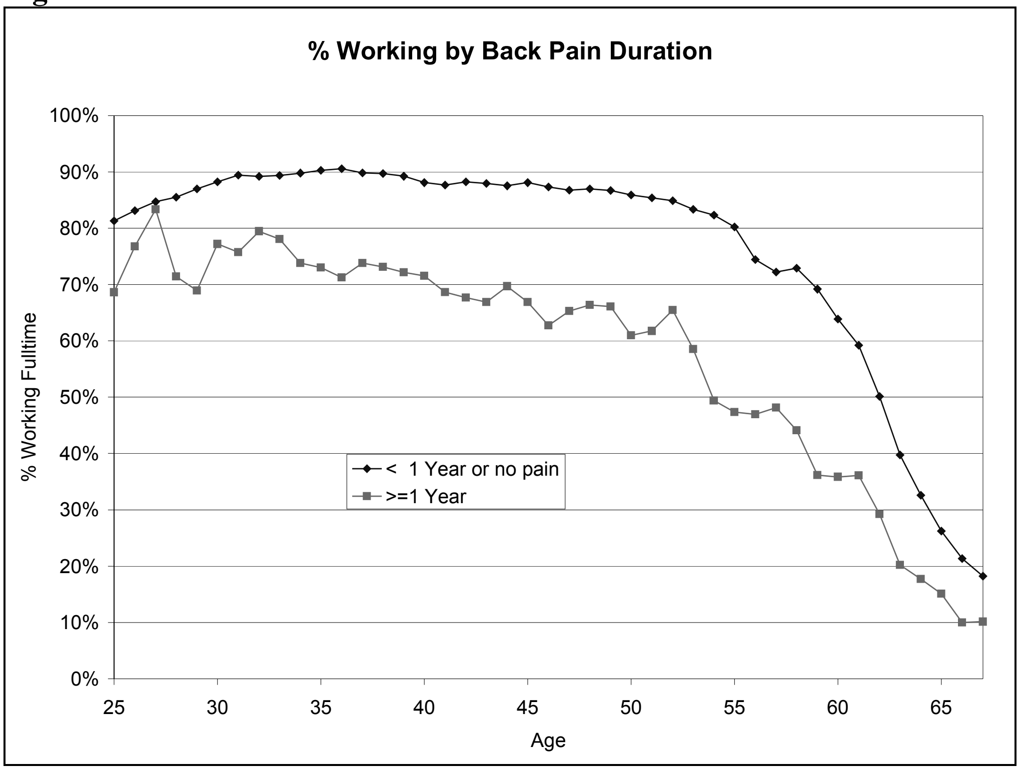
Using the NHIS, we computed the percentage of males with different levels of education who worked at each age, both with and without chronic back pain that limited some activities, where chronic indicates a condition lasting one year or longer. The results are summarized in Table 5. A high school diploma or BA degree causes a large increase in worklife expectancy compared to the next lower level shown. An AA degree causes a smaller increase. These increases are much larger for those with chronic back pain. The trend for all levels of education combined is shown in Figure 1.

Table 5

Worklife Expectancy by Education and Back Pain								
	< HS Diploma		HS Diploma		AA Degree		BA Degree	
Age	No Injury	Back Pain	No Injury	Back Pain	No Injury	Back Pain	No Injury	Back Pain
25	27.3	15.9	31.1	24.2	32.0	25.9	33.1	31.4
35	19.4	10.5	22.7	16.5	23.3	17.6	24.6	23.0
45	12.0	5.5	14.2	9.5	14.5	10.6	15.6	14.5
55	5.2	2.2	6.2	3.7	6.2	4.1	7.0	6.4

Source of mortality data used in computing worklife was United States Life Tables, 2001, National Vital Statistics Reports, 52(14), Centers for Disease Control and Prevention.

Figure 1



The effect of back pain on worklife and the time trend of the effect are particularly interesting. The difference in the probability of being employed at younger ages is ten to fifteen percent but the difference grows dramatically as the worker ages and is seventeen percent at age forty and thirty percent after age fifty. One could conclude from this finding that, even if an injured individual were able to continue employment in their original occupation post-accident, worklife is still shortened. This finding is important in designing reeducation programs and career paths for workers who sustain back injury early in life. Though many younger workers function with chronic back pain, the ability to do so appears to decline with age. As expected, back pain has a much smaller effect on the worklife of the educated as their jobs are more likely to be sedentary.

Statistical Model of Worklife

The alternative to enumeration is statistical regression, which can be used to incorporate a greater breadth of information on individuals. This is particularly important when many confounding factors are present. Since a subject is either working or not, a binary variable, the assumptions of ordinary least squares (OLS) are violated and logistical regression is the favored method.

The logistic model takes the form:

$$p(work) = \frac{1}{1+e^{-Z}}$$

where Z is the linear combination of explanatory variables $X_1,\ldots,X_n$, and regression

coefficients $B_0,\ldots,B_n$,

$$Z = B_0 + \sum_{i=1}^{n} B_i X_i \,.$$

Information that is potentially relevant in predicting employment status includes highest educational degree, race, marital status, age, the number and ages of children living at home and any physical limitations. Our model to predict the probability of fulltime employment is similar to one we developed previously for spinal cord injury victims.[11] The dependent variable in the logit model is fulltime employment, which is defined as 1 for those working 35 hours or more per week and otherwise as 0.

We divided the sample into six groups based on three levels of education and gender and modeled the probability of fulltime employment for each group separately. The three levels of education were: less than a high school diploma; a high school diploma or AA degree; and an undergraduate or advanced college degree. The sample included ages 25 to 70.

To measure the effects of disabilities we included several medical conditions and standard tasks that the subject stated were either very difficult or impossible to perform. We also included one ADL, difficulty walking without special equipment. The medical conditions were insulin dependent diabetes and a history of stroke. Binary variables are used to indicate two-state conditions with a value of 1 if the condition is present and 0 if not. The standard tasks were listed previously in Table 4.

Since the sample was stratified based on factors that are included in the model, the regression was unweighted.

The coefficients are shown in Table 6 along with the level of significance and the odds ratio, which shows the ratio of probabilities resulting from a one unit change in the explanatory variables. An odds ratio of 0.5 for a disability would indicate that a person is only half as likely to work fulltime with that disability all else being equal. A significance value of 0.0500 indicates that a variable is significant at the 5% level. Stepwise regression was used to eliminate demographic variables that were not significant at the 5% level.

Of particular note among the demographic variables, black females who graduated from college are much more likely to seek work than their white counter parts. This is also true for high school graduates, but to a lesser extent. Black males without a high school diploma are 44% less likely to find work than males of other races. An AA degree increases the odds of employment by 29% and 33% for males and females relative to those with only a high school diploma. Single males in all education categories are much less likely to work than married males. For females, marriage has the opposite effect. Having children living at home greatly decreases the odds that a female will work. Except for males without a high school diploma, immigrating to the US within the last five years decreases the odds of finding work. This effect increases with the level of education perhaps because language skills are more important in occupations that require more education.

The effects of age and education on the probability of working fulltime are illustrated in Figure 2. These probabilities were computed using the means of the other explanatory variables.

Figure 2

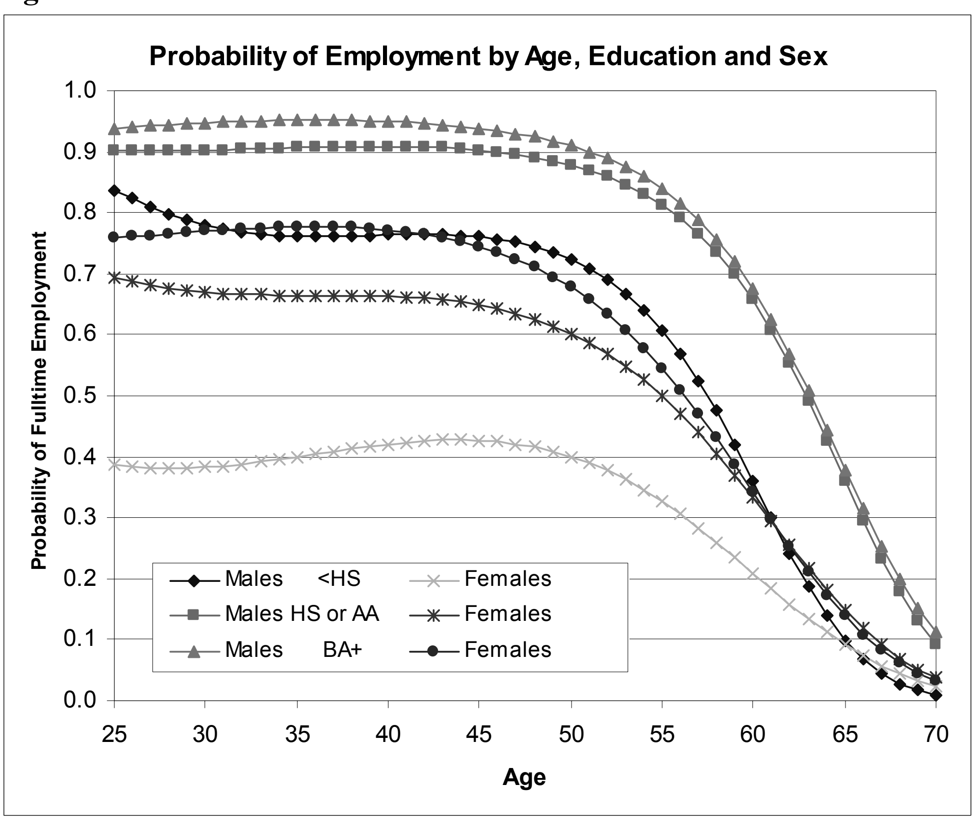


Table 6
Coefficients in Logit Model of Fulltime Employment

		COEFFICENTS						SIGNIFICANCE						ODDS RATIO					
		MALES			FEMALES			MALES			FEMALES			MALES			FEMALES		
		EDUCATION CATEGORY																	
DESCRIPTION	VARIABLE	<HS	HS	BA+	<HS	HS	BA+	<HS	HS	BA+	<HS	HS	BA+	<HS	HS	BA+	<HS	HS	BA+
Constant	Constant	1.8605	1.9634	2.9335	-0.9785	0.9466	1.7476	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000						
White	WHITE		0.3861				-0.1748		0.0000				0.0166		1.47				0.84
Black	BLACK	-0.5858				0.2339	0.5141	0.0000				0.0000	0.0000	0.56				1.26	1.67
Age	AGE		0.0087		0.0184				0.0044		0.0000								
$(Age-36)^2$	AGE2	0.0018	-0.0008	-0.0028			-0.0016	0.0001	0.0175	0.0000			0.0001						
$(Age-36)^3$	AGE3	-0.0002	-0.0001	0.0000	-0.0001	-0.0001	-0.0001	0.0000	0.0000	0.0038	0.0000	0.0000	0.0000						
No. of grades completed up to 12	EDUCSEC				0.0483						0.0000						1.05		
GED	EDUCGED				0.3073						0.0000						1.36		
AA degree or higher	EDUCAA		0.2508			0.2889			0.0000			0.0000			1.29			1.33	
MD, DDS, DVM, JD	EDUCDR			0.2202						0.0316						1.25			
Ph.D.	EDUCPHD			0.5318			0.7121			0.0000			0.0000			1.70			2.04
Never married	SINGLE	-0.9054	-0.8546	-0.4679	0.1507	0.4979	0.7431	0.0000	0.0000	0.0000	0.0060	0.0000	0.0000	0.40	0.43	0.63	1.16	1.65	2.10
Widowed	WIDOWED	-0.4538	-0.4999	-0.4476		0.2711	0.3098	0.0021	0.0000	0.0190		0.0000	0.0042	0.64	0.61	0.64		1.31	1.36
Divorced or Separated	DIVSEP	-0.3679	-0.4390		0.5742	0.9927	1.0112	0.0000	0.0000		0.0000	0.0000	0.0000	0.69	0.64		1.78	2.70	2.75
Number of children living at home	KIDS			0.1573		-0.0804				0.0000		0.0000				1.17		0.92	
$(children < age 13)^{.5}$	CHILD12				-0.4875	-0.6156	-1.1694				0.0000	0.0000	0.0000				0.61	0.54	0.31
Living in the US < 5 years	RECENTUS		-0.5896	-0.8484	-0.3816	-0.9603	-1.4373		0.0000	0.0000	0.0002	0.0000	0.0000		0.55	0.43	0.68	0.38	0.24
Difficulty walking w/o equipment	ADLWALK	-1.6024	-0.9931	-0.7527	-0.8742	-0.8753	-0.8513	0.0000	0.0000	0.0010	0.0000	0.0000	0.0000	0.20	0.37	0.47	0.42	0.42	0.43
Walk 1/4 mile	DIFQMILE	-0.4499	-0.5442	-0.4598	-0.5140	-0.2955	-0.2048	0.0008	0.0000	0.0536	0.0000	0.0001	0.2424	0.64	0.58	0.63	0.60	0.74	0.81
Stand for 2 hours	DIFSTAND	-1.3208	-1.1890	-0.6378	-0.7634	-0.5333	-0.8600	0.0000	0.0000	0.0005	0.0000	0.0000	0.0000	0.27	0.30	0.53	0.47	0.59	0.42
Sit for 2 hours	DIF2HSIT	-0.6190	-0.3317	-0.5372	-0.3273	-0.3612	-0.2493	0.0002	0.0076	0.0362	0.0179	0.0000	0.1631	0.54	0.72	0.58	0.72	0.70	0.78
Stoop, bend or kneel	DIFSTOOP	-0.4954	-0.6067	-0.2587	-0.2435	-0.1450	-0.1346	0.0001	0.0000	0.1708	0.0139	0.0171	0.2983	0.61	0.55	0.77	0.78	0.87	0.87
Reach over the head	DIFREACH	-0.8749	-0.9264	-0.5863	-0.2804	-0.4744	-0.6333	0.0000	0.0000	0.1063	0.0939	0.0000	0.0056	0.42	0.40	0.56	0.76	0.62	0.53
Lift and carry 10 pounds	DIFCARRY	-1.3935	-1.0549	-0.6015	-1.0265	-0.9287	-0.4303	0.0000	0.0000	0.0694	0.0000	0.0000	0.0197	0.25	0.35	0.55	0.36	0.40	0.65
Insulin dependent diabetes	INSULIN	-0.8305	-0.5328	-0.3299	-0.6661	-0.4515	-0.1201	0.0000	0.0000	0.1077	0.0000	0.0000	0.6011	0.44	0.59	0.72	0.51	0.64	0.89
History of stroke	STROKE	-0.8649	-1.0244	-0.9275	-0.7405	-0.5109	-0.5368	0.0000	0.0000	0.0003	0.0003	0.0000	0.0562	0.42	0.36	0.40	0.48	0.60	0.58
Observations / Nagelkerke R^2		11,678	31,242	16,293	14,475	39,678	17,212	.437	.143	.386	.046	.337	.106						

Most of the disabilities were significant within the six groups, but less so for those with a BA degree. Only two disabilities show a decreasing impact with higher education for both sexes. These include the ability to lift and carry 10 pounds and being dependent on insulin. For males only, this trend is evident for difficulty walking without special equipment and standing for two hours. For females only, this trend is evident for walking a quarter mile. Accounting for other confounding factors has reduced the effect of education on disability that was apparent previously in Table 4.

Models of Earnings

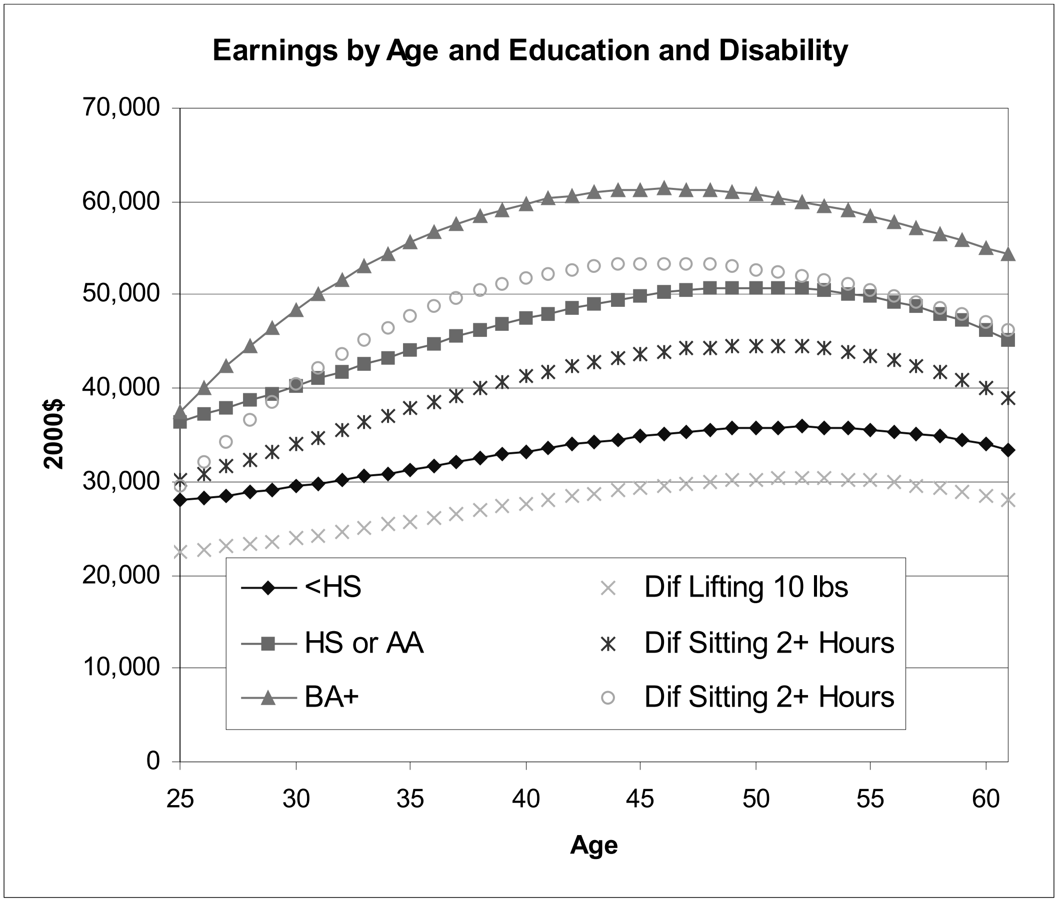
Annual earnings in 2000$ was modeled separately for each group with ordinary least squares regression. The sample included respondents who worked 12 months during the previous year. Most of the variables that were significant in predicting fulltime employment were also significant in predicting earnings. The results are shown in Table 7. Stepwise regression was used to eliminate variables that were not significant at the 5% level.

Of note among the demographic variables, whites tend to earn more, and blacks less, than other races. One exception is for college educated females – white females earn less than their counterparts. An AA degree increases annual earnings by $5,511 for males and by $4,791 for females who work fulltime. A masters degree increases earnings by about $7,000. An MD adds another $20,000 or more to earnings. A Ph.D. adds about $14,000 for males and $19,000 for females. Married males earn more than single males. The opposite is true for females. Recent immigrants earn less, other things being equal.

Fewer disabilities are significant in the model of earnings than in the model of fulltime employment, which suggests that disability affects the likelihood of obtaining employment or holding a job more than it affects the amount of earnings.

The trend in predicted men's earnings as each group ages in shown in Figure 3 by education and gender. Earning are compared for those with and without a difficult task. Clearly, this model shows that disability has a significant effect on earnings. For example, a college graduate who has difficulty sitting for two hours will earn, on average, almost $10,000 less per year than a similar individual without this physical limitation. Similar reductions in earnings are found for other significant disabilities and educational levels.

Figure 3



Earnings by Age and Education and Disability
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
2000$
25
30
35
40
45
50
55
60
Age
<HS
HS or AA
BA+
Dif Lifting 10 lbs
Dif Sitting 2+ Hours
Dif Sitting 2+ Hours

Table 7

Coefficients in Regression Model of Earnings

		COEFFICIENTS						SIGNIFICANCE LEVEL					
		MALES			FEMALES			MALES			FEMALES		
		EDUCATION CATEGORY											
DESCRIPTION	VARIABLE	<HS	HS	BA+	<HS	HS	BA+	<HS	HS	BA+	<HS	HS	BA+
Constant	CONST_	-3,467	11,429	20,525	3,953	14,680	19,490	0.0243	0.0000	0.0000	0.0045	0.0000	0.0000
White	WHITE	1,803	4,386	3,653		1,289	-2,182	0.0082	0.0000	0.0001		0.0000	0.0001
Black	BLACK	-3,048	-1,648	-3,215	-1,167			0.0006	0.0216	0.0127	0.0058		
Age	AGE	391	720	1,003	148	349	734	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$(Age-36)^2$	AGE2		-11.2	-60.1		-19.3	-49.6		0.0001	0.0000		0.0000	0.0000
$(Age-36)^3$	AGE3	-0.5	-0.7	0.6	-0.3		0.4	0.0000	0.0000	0.0003	0.0000		0.0089
No. of grades completed up to 12	EDUCSEC	1,739			1,018			0.0000			0.0000		
AA degree or higher	EDUCAA		5,511			4,791			0.0000			0.0000	
MA or MBA or higher	EDUCMA			6,900			7,175			0.0000			0.0000
MD, DDS, DVM, JD	EDUCDR			20,052			22,114			0.0000			0.0000
Ph.D.	EDUCPHD			14,124			19,276			0.0000			0.0000
Never married	SINGLE	-4,538	-6,137	-7,143	-1,392	2,019	1,497	0.0000	0.0000	0.0000	0.0023	0.0000	0.0067
Divorced or Separated	DIVSEP	-2,009	-3,860	-5,565		2,404	1,323	0.0003	0.0000	0.0000		0.0000	0.0188
Number of children living at home	KIDS			920	-734	-1,824	-2,836			0.0005	0.0000	0.0000	0.0000
Living in the US < 5 years	RECENTUS	-7,378	-13,427	-9,629	-2,290	-7,598	-10,406	0.0000	0.0000	0.0000	0.0454	0.0000	0.0000
Stoop, bend or kneel	DIFSTOOP		-3,240	-5,392		-1,284			0.0010	0.0156		0.0242	
Sit for 2 hours	DIF2HSIT		-6,478	-9,448	-2,439	-2,606	-6,006		0.0000	0.0073	0.0185	0.0020	0.0086
Lift and carry 10 pounds	DIFCARRY	-5,525			-3,940	-1,727	-5,506	0.0163			0.0002	0.0499	0.0247
Walk 1/4 mile	DIFQMILE		-4,715			-3,321			0.0005			0.0000	
Insulin dependent diabetes	INSULIN		-4,778	-5,144		-2,295	-8,745		0.0006	0.0448		0.0285	0.0012
Number of observations / R^2		6,395	22,841	13,036	4,952	23,338	12,134	0.118	0.091	0.135	0.075	0.064	0.116

Self Selection Bias

A model of earnings is a censored regression since earnings is observed only for those who participate in the labor force. Potential earnings can certainly be expected to influence participation causing biased coefficients in an OLS model of earnings. This is a classic problem in labor economics and econometrics. Heckman (1974) demonstrated that self-selection into the labor force can bias coefficients in a model of earnings.[12] He showed that the reservation wages of woman determined labor force participation and were correlated with earnings.[13] Heckman (1976) proposed a two-step method for correcting self-selection bias when selection is determined by a probit model and earnings by a linear regression equation.[14] The inverse Mills ratio is computed with the results of the probit model and included in the OLS regression on earnings. Maximum likelihood (ML) estimation can also be used to estimate the coefficients for both equations simultaneously and is preferred over the two-step procedure.[15]

Self-selection causes correlation between the residuals in the two equations and is measured by either the two-step or ML methods. Two factors allow the correlation to be measured. One is the non-linearity of the inverse Mills ratio and the other is the existence of any variables that appear in the model of employment but not the model of earnings.

To test for self-selection bias in our model of earnings, we used the ML method to estimate our models for employment and earnings. Since both models use a nearly identical set of explanatory variables, we introduced two additional variables in the employment model to serve as instruments. These variables were the earnings of other family members and family income from sources other than earnings. Neither variable should affect earnings for the respondent. We would expect that other sources of income would decrease the need to be employed and thus would have a negative coefficient in the model of labor force participation. These were variables were multiplied by age and age^2. The coefficients were then plotted as a function of age. The patterns showed that these sources of income created increasing disincentives to work after ages 40 and 50, for non-earnings income and earnings from other family members, respectively.

The coefficients for the earnings model when ML estimation is used to correct self-selection bias are presented in Table 8. The model for males with a BA degree or higher did not converge in 1000 iterations. In all but one of the other cases, the correlation coefficient, RHO, is negative, and the coefficients for disabilities are smaller compared to the OLS estimates shown above. In these cases, the significance level of the disabilities is also less. The case that is an exception is females with a high school diploma. RHO is positive and three of the five disability coefficients are larger. In the equation for females without a high school diploma, the significance level of rho is only 0.1058.

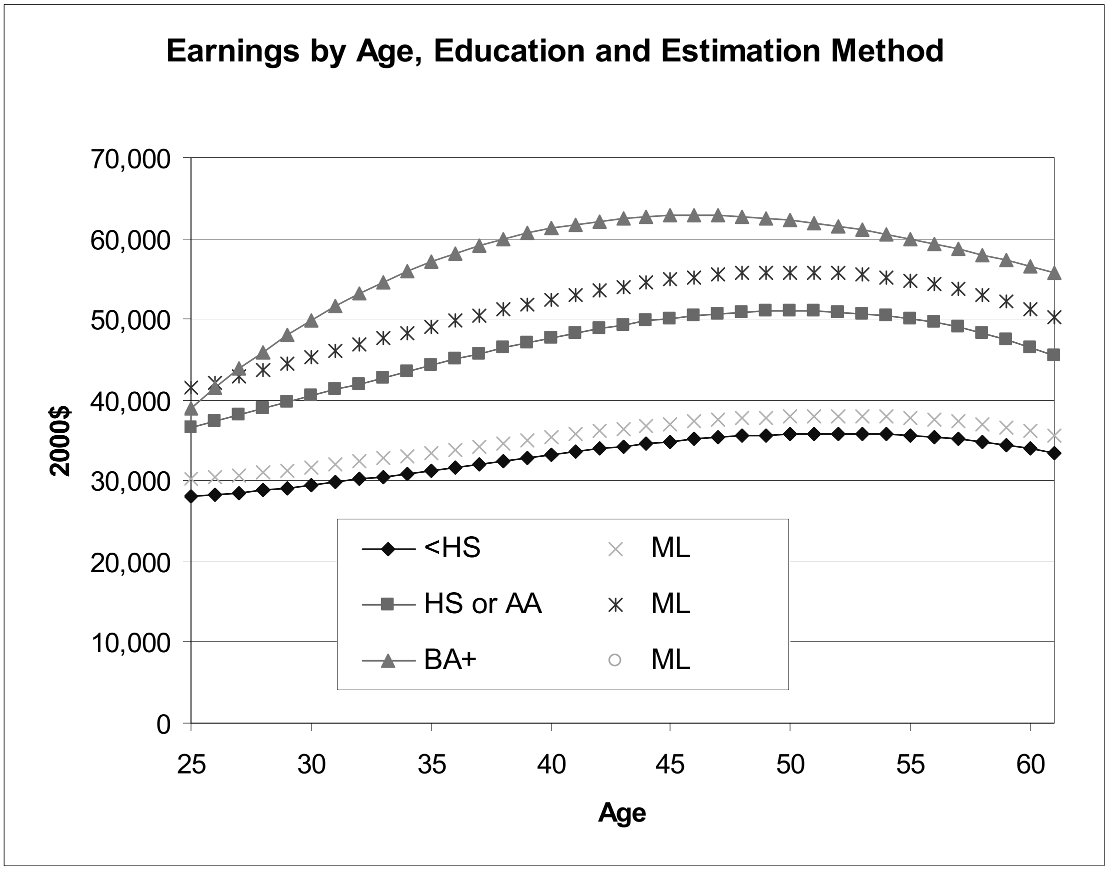
Further analysis will be required to identify the computing problem. We believe that the model for females without a high school diploma is solving at a local rather a global maximum.

Table 8
Coefficients For Earnings After Correcting For Self Selection

	COEFFICIENTS					SIGNIFICANCE LEVEL				
	MALES		FEMALES			MALES		FEMALES		
VARIABLE	<HS	HS	<HS	HS	BA+	<HS	HS	<HS	HS	BA+
C	-1126	17099	6048	-1885	27921	0.5609	0.0000	0.0019	0.0342	0.0000
WHITE	1645	3419		2099	-3350	0.0548	0.0000		0.0000	0.0001
BLACK	-2906	-2100	-1085		-2109	0.0064	0.0130	0.0150		0.0528
AGE-1	376	648	126	516	602	0.0000	0.0000	0.0001	0.0000	0.0000
(AGE-37)^2	-6	-14	0	-15	-46	0.1758	0.0000	0.9696	0.0000	0.0000
(AGE-37)^3	0	0	0	-1	1	0.0457	0.0014	0.0272	0.0000	0.0016
EDUCSEC	1725		985			0.0000		0.0000		
EDUCAA		5235		6712			0.0000		0.0000	
EDUCMA					7095					0.0000
EDUCPHD					18527					0.0000
EDUCDR					22504					0.0000
SINGLE	-3951	-5170	-1414	5202	941	0.0000	0.0000	0.0045	0.0000	0.1311
DIVSEP	-1725	-3252		6576	517	0.0016	0.0000		0.0000	0.4140
KIDS			-712	-2595	-2275			0.0000	0.0000	0.0000
RECENTUS	-6532	-11137	-1993	-16712	-8317	0.0027	0.0000	0.3856	0.0000	0.0000
DIFSTOOP		-1510		-5351			0.1538		0.0000	
DIF2HSIT		-5865	-1969	-2599	-5434		0.0005	0.1727	0.0000	0.0439
DIFCARRY	-3376		-3407	-10780	-3230	0.2413		0.0149	0.0000	0.2462
DIFQMILE		62		-9082			0.9671		0.0000	
INSULIN		-4496		-937	-8383		0.0029		0.3414	0.0193
Rho	-0.15	-0.31	-0.09	0.96	-0.23	0.0042	0.0000	0.1058	0.0000	0.0000

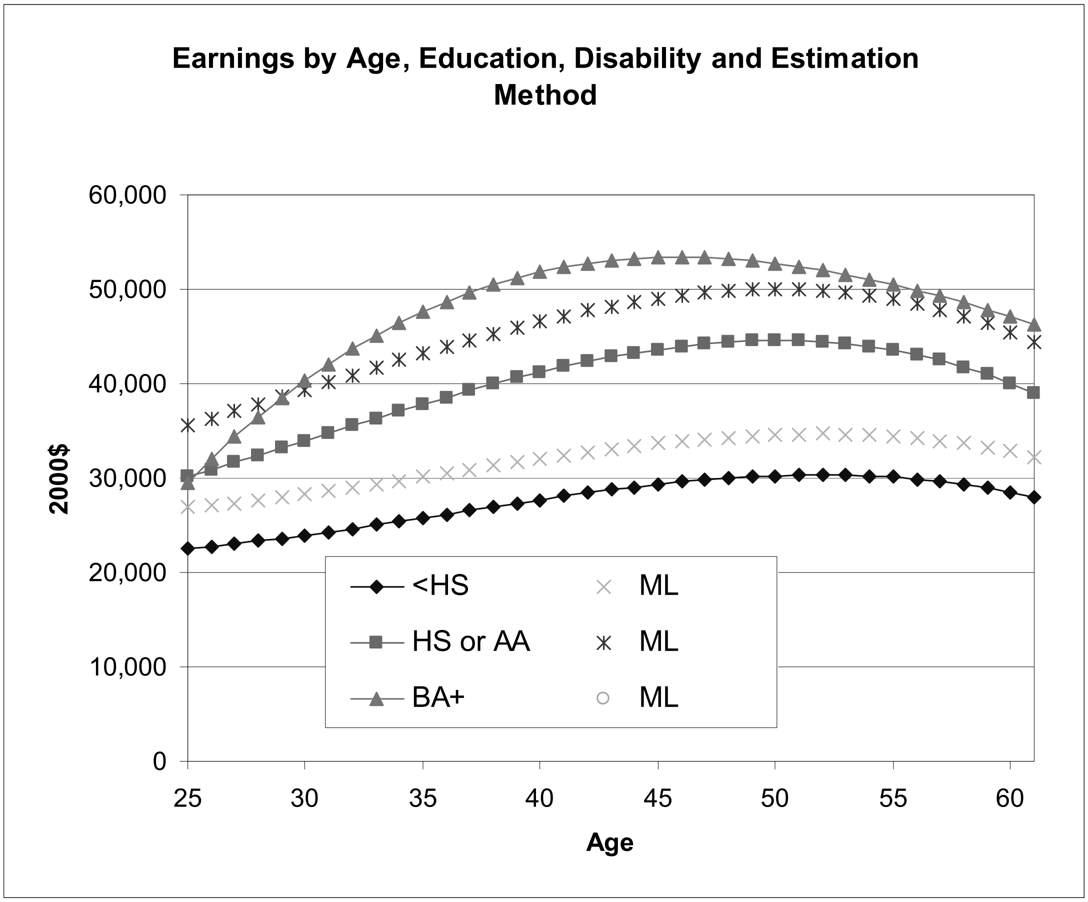
Figure 4 compares predicted earnings assuming no disabilities for the two estimation methods. The differences for a given age category are, then, attributed to the estimation method used, OLS or ML. The ML method consistently produces higher estimates of earnings indicating the sample bias causes the earnings to be underestimated when using OLS.

Figure 4



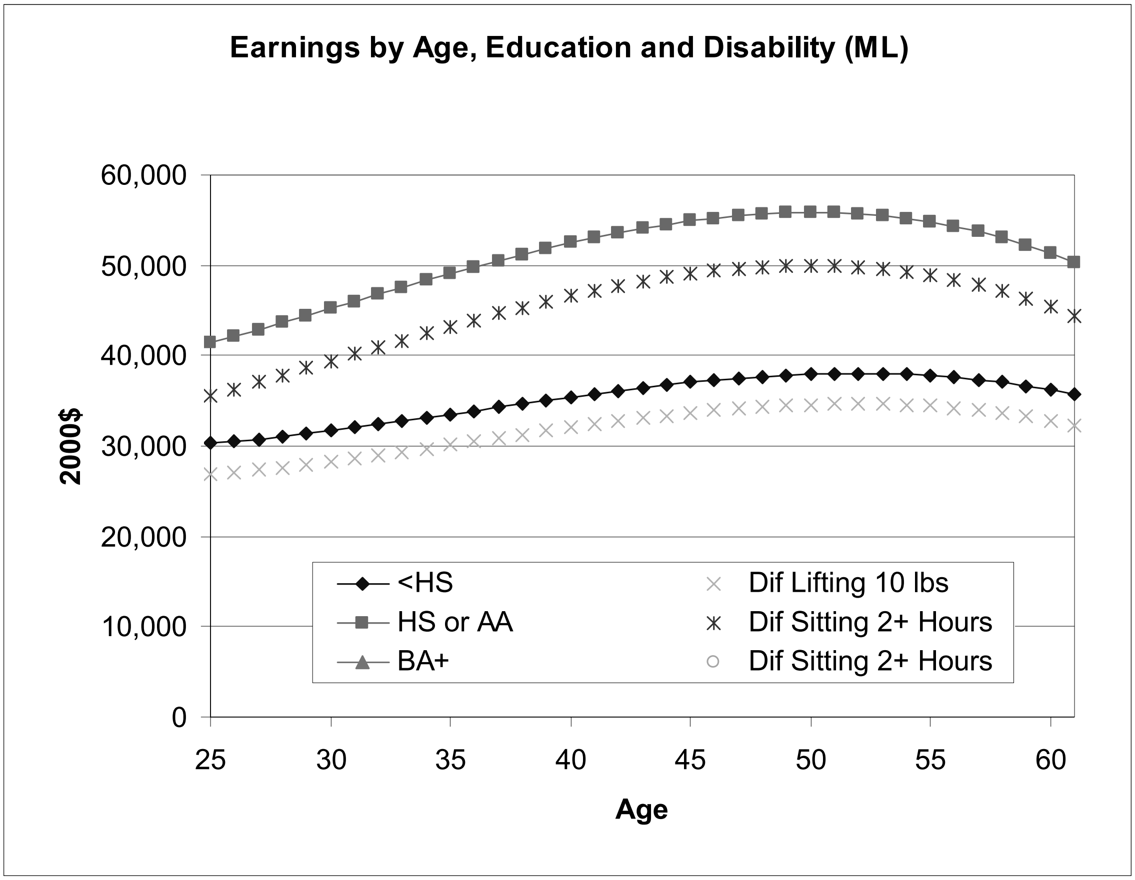
In Figure 5 we compare OLS and ML, both for the same disability that shown in Figure 3. The difference between the OLS and ML estimates is considerably larger than our estimates without disability. The sample selection bias causes the estimate of income for an individual with the disability modeled to be underestimated if using OLS.

Figure 5



In Figure 6 we compare predictions of earnings for an individual with one disability using the model estimated with ML. The difference between the earnings of the individual with an impairment and one without is less when estimated with ML than with OLS.

Figure 6



Conclusion

Modeling worklife and earnings for the physically impaired using a health-focused data set and proper econometric technique provides significantly different findings than either enumeration or simple regression using data from the CPS or the SIPP. We find that the disabled are more likely to find work and enjoy higher earnings than previous authors, using inapt data and simplistic statistics, have reported.

[1] Smith, Shirley J., *Tables of Working Life, The Increment-Decrement Model*, Bulletin 2135, US Department of Labor, Bureau of Labor Statistics, November 1982; Smith, Shirley J., *Worklife Estimates: Effects of Race and Education*, Bulletin 2254, US Department of Labor, Bureau of Labor Statistics, February, 1986; Boudreaux, Kenneth J., "A Further Adjustment Needed to Estimate Lost Earning Capacity," *Monthly Labor Review* 106(10), 1983, 30-31; Richards, Hugh, *Life and Worklife Expectancies*, Lawyers & Judges Publishing Company, Inc., Tucson, AZ, 1999.

[2] In March 2004, the Census Bureau published a paper on its website, "Uses and limitations of CPS data on work disability" which explains that questions pertaining to disability were designed for screening and not for measuring disability.

[3] Hale, Thomas W., "The lack of a disability measure in today's current population survey," *Monthly Labor Review* 2001;124(6):37-40. The BLS is currently developing methods for better incorporating disability status into the CPS.

[4] Kirchner, Corinne, "Looking Under the Street Lamp: Inappropriate Uses of Measures Just Because They Are There," *Journal of Disability Policy Studies*, 1996, 7(1), 77-90.

[5] Currie, Janet and Brigitte C. Madrian. 1999. "Health, Health Insurance and the Labor Market." In *Handbook of Labor Economics* vol. 3C, ed. Orley Ashenfelter and David Card, 3309-3416. New York: Elsevier Science.

[6] See www.cdc.gov/nchs/nhis.htm .

[7] Hale, Thomas W. and Kirchner cited previously. The newly developed screening questions to be used in the CPS are also based on ADL's and standardized tasks.

[8] Baker, Michael, Mark Stabile and Catherine Deri, "What do Self-Reported, Objective, Measures of Health Measure?" *The Journal of Human Resources,* 2004, 39(4), 1067-1093.

[9] Richards, *Life and Worklife Expectancies.*

[10] Stewart, Walter F., PhD, MPH; Judith A. Ricci, ScD, MS; Elsbeth Chee, ScD; David Morganstein, MS; Richard Lipton, MD, "Lost Productive Time and Cost Due to Common Pain Conditions in the US Workforce," *Journal of the American Medical Association,* 290(18), November 12, 2003.

[11]Pflaum, Christopher C., George M. McCollister, David J. Strauss, Robert M. Shavelle, and Michael J. DeVivo "Disability and Worklife: The Case of Spinal Cord Injury." Under review A*rchives of Physical Medicine and Rehabilitation*.

[12] Heckman, James J., (1974) "Shadow Wagers, Market Prices, and Labor Supply," *Econometrica,* 42, 679-94.

[13] This example is used extensively in Heckman's article on "Selection Bias and Self-Selection" in *The New Palgrave A Dictionary of Economics, Volume 4,* 1987, The Macmillan Press Limited.

[14] Heckman, James J., (1976) "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, 475-92.

[15] Davidson, Russell and James G. MacKinnon, *Estimation and Inference in Econometrics*, 1993, Oxford University Press.