INTRODUCTION

One of the key indicators of recovery from a workplace injury is return to work (RTW). While the injury type and work context are major determinants of RTW, the broader socio-economic and demographic attributes of injured workers also play important roles (Krause et al 2001). Most quantitative research into the determinants of RTW has used either administrative databases designed to support surveillance of the compensation claims process or surveys of individuals sampled from those databases. Their disadvantages include their being restricted to work injuries and work absences for which workers compensation is paid, their limited geographic universes, and their limited information on socio-demographic factors associated with RTW. Multi-purpose, nationally representative, longitudinal ("panel") surveys that include questions on occupational injuries have the potential to overcome these disadvantages (Reville, Bhattacharya, and Sager Weinstein 2001). In the present study, we explore this potential by comparing estimates of the socio-demographic variables associated with work injuries and RTW in the National Longitudinal Survey of Youth (NLSY) 1979 Cohort and the Survey of Income and Program Participation (SIPP).

Administrative databases on workers compensation claims have the major advantages of collecting information specifically on workplace injuries and their resolution through the compensation claim process, including job and firm characteristics, and of having large numbers of cases to analyze (Cheadle et al 1994; Krause et al 1999; Oleinick and Zaidman 2004; Seland, Cherry, and Beach 2006). Administrative systems, however, track only those injuries submitted through the registration or claim process, and track those injuries only until the claim is fully settled or paid out. Administrative databases may therefore result in underestimation by 50% or more of losses and the time out of work due to injury (Reville 1999).

The problem of undercoverage of occupational injuries in administrative databases has broader implications for estimating both their total costs and the distribution of these costs in society (Weil 2001; Schulte 2005). Greater self-reported than administratively-recorded time out of work, for example, has been found to be associated with socio-economic and socio-psychological disadvantage (Pole et al 2006). Related to this, another problem with using administrative databases to analyze RTW is their very limited recording of the socio-demographic characteristics of the injured worker. For example, higher levels of education, family income, and social support have all been found to be associated with faster RTW even after controlling for occupation and the physical job demands (McKenzie et al 1998). Such variables are typically not included in workers compensation databases.

Recognizing these limitations of administrative databases for the analysis of the determinants of work injuries and RTW, researchers have designed and analyzed data collected in special-purpose surveys of injured workers sampled from administrative claims databases (Butler, Johnson, and Baldwin 1995; Turner et al 2007). These surveys collect a wider set of social, organizational, and economic variables, have questions dedicated to the topic of workplace injuries and their contexts and consequences, and are designed to include sufficiently large samples with which to explore those topics. These surveys have, however, four main disadvantages. The first is their very high cost compared to secondary data analyses. The second, related to the first, is that they are typically one-off, cross-sectional surveys. Third, their selective geography (e.g., confined to single states) makes them less suitable for addressing the national-level public health consequences of workplace injuries, including disparities in occupational health. Fourth, because their samples are drawn from administrative databases, they similarly miss workplace injuries not covered by the workers compensation system.

Multi-purpose, nationally representative surveys with questions on work injuries have the potential to overcome the deficiencies of workers compensation databases and surveys of individuals sampled from those databases. Longitudinal ("panel") surveys are especially useful for understanding the roles of factors influencing a process that occurs over time, such as recovery from a workplace injury or illness. They include a larger range of questions on the social and family context of the injured individual and antecedent and consequent work outcomes. They also include workplace injuries that resulted in time out of work but that did not result in a successful workers' compensation claim, or periods of work absence after workers compensation receipt has ended. In sum, they may be considered as representative of all workplace injuries and time out work for the cohorts in their national sampling universes.

Three national panel surveys in the U.S. have questions that identify workplace injuries (Reville et al 2001), the Health and Retirement Study (HRS, University of Michigan, of cohorts aged 50 and over, the National Longitudinal Survey of Youth (NLSY79, Bureau of Labor Statistics, aged in their 30s in the 1990s, and the Survey of Income and Program Participation (SIPP, Westat 2001) of all ages and cohorts. Studies by Zwerling et al (1996; 1998) using the HRS, and by Dembe, Erickson, and Delbos (2004) and Dong (2005) using the NLSY, explore the determinants of work injuries. None of these studies, however, analyze subsequent returns to work. DeLeire (2000, 2001) pools the 1986 through 1993 panels of the SIPP to study the employment and earnings consequences of having experienced a workplace injury, but again without explicitly studying returns to work.

OBJECTIVES

The main objective of this study is to compare estimates of the relationship between social and economic characteristics and occupational injury outcomes in two multi-

purpose population representative surveys. Comparisons are made for of the social and demographic predictors of incidence of occupational injury and the hazards of return. These comparisons are drawn using twelve years of prospective data on occupational outcomes among adults in the National Longitudinal Survey of Youth 1979 Cohort (1988-2000) and three panels of prospective data following adults for as long as four years from the Survey of Income and Program Participation (1996, 2001, and 2004).

METHODS

Survey Description and Sample

Data come from two nationally representative panel surveys. The first survey is the National Longitudinal Survey of Labor Market Experience, Youth Survey 1979 (NLSY). This is an ongoing study of educational and labor force attainments, commissioned by the Bureau of Labor Statistics (BLS) and conducted by the Center for Human Resource Research (CHRR) at Ohio State University. The NLSY has followed a cohort of the civilian population annually from 1979 through 1994 and biennially from 1994-2006. At baseline, respondents were 14-21 years-old. The NLSY is particularly appropriate to studying social disparities in the labor force outcomes because it over-sampled African Americans, Hispanics and, until 1990, economically disadvantaged non-minority men and women. After the 1990 sample cuts, 9,964 individuals remained in the survey. Retention rates (excluding the deceased respondents) are above 90% for the years when the survey was annual (1979-1994), and are only slightly lower once the survey became biennial (1996-2004) (http://www.bls.gov/nls).

The second survey is the Survey of Income and Program Participation (SIPP), a study conducted by the Census Bureau as a series of panels of between 2 and 4 years duration, with survey interview “waves” 4 months apart (Westat 2001; Weinberg 2002). It was conducted from 1984 through 1993 as a series of annual panels. Each panel ran

for between 2 and 3 years and individual panel samples were of around 20,000 households. The sample sizes at any given year are typically twice this size, due to overlap between panels. There was a break in 1994 and 1995 when no new panel was initiated. Instead, a larger sample of 40,000 households was included in the 1996 panel, which ran for 4 years. A larger sample was similarly included in the 2001 panel, which ran for three years, and again in the 2004 panel. In the present study, we use only the three largest and most recent panels (1996, 2001, and 2004). Attrition in the SIPP is generally higher than in the NLSY. However, analysis does not show it to be a major problem for estimating general labor market models (Zabel 1998).

Measurement of Occupational Injury

In both the NLSY and the SIPP, an occupational injury is defined among respondents who report a work-limiting health condition on the basis of questions about injuries at work that influence the kind and/or amount of work that a respondent is able to conduct.

In the NLSY, questions about work-limitation and work injuries are asked in each wave of the panel between 1988 and 2000. The data on work-limiting health conditions and work injuries are combined to identify occupational injuries that are most equivalent to those reported in the SIPP. Injuries that lead to a work-limiting health condition are identified through the coincidence of reports of work-limiting health condition and of an injury at work. Specifically, this entails a positive response to either of the following two questions about work-limiting health conditions*"[are you/would you be] limited in the kind of work you (could) do on a job for pay because of your health?*" or "[*are you/would you be] limited in the amount of work you (could) do because of your health?*", and it entails a positive response to the following question about work injuries, "*during the past 12 months/since the last survey], have you had an incident at any job we previously discussed that resulted in an injury or illness to you?"* Refinement of coincident reporting

of occupational injury is conducted for analyses of return to work that employs the reported month and year that the work limitation and injury occurred, and this is described below. On the basis of these questions, there were 805 occupational injuries reported in the NLSY, 190 of which were observed among respondents age 35-40 in the years 1996, 1998, and 2000 that overlap most closely with the SIPP[1].

The SIPP included an occupational injury history module from its 1986 panel onwards, always in Wave 2, four months after the beginning of the panel. This history is asked of all respondents reporting a work-limiting condition. The questions on work limitation are asked as part of the core survey in every wave, but these reports are linked explicitly to work injuries only in Wave 2. Occupational injuries are identified from questions in SIPP as follows. To define being work limited, respondents must positively answer, *"do you have a physical, mental or other health condition that limits the kind or amount of work you can do?"* Then among these respondents, injuries at a job are identified using the following series of questions "*we have recorded that your health or condition limits the kind or amount of work you can do. Is that correct?"; "were you employed at the time your work limitation began?"; "was this condition caused by an accident or injury?";* and lastly *"where did the accident or injury take place?" 1. on the job; 2. Armed Forces; 3. in the home; 4. somewhere else?".* On the basis of these questions, there were 1,200 respondents who reported occupational injuries in the 1996, 2001 or 2004 SIPP panels, 359 of whom were among respondents ages 35-40 matching those of the NLSY.

Measurement of Return to Work and Duration of Work Disability

[1] Occupational injuries are restricted to those occurring after the first date of the NLSY survey in 1988.

All respondents identified above as having an occupational injury have been, by definition, unable to work for some period of time due to the injury. Following the occupational injury, respondents may or may not return to work. Among those respondents that do return to work, the period of time between the occupational injury start date and the date the respondent returned to work is defined as the duration of the work disability, or time of exposure to RTW. For respondents who are never observed as returning to work, the duration of the work disability is censored and the exposure time is the interval between the occupational injury start date and the date of censoring (when the respondent is last observed in the survey).

The measurement of occupational injury start date comes from a reconciliation of the dates of the work-limiting health condition, exit from the workplace, and the date of the job injury. Respondents are excluded whose injury is reported as occurring after or more than 12 months prior the onset of a work limiting condition. The date (month and year) of the work limiting health condition in the NLSY comes from the question "*since what month and year have you had this limitation"* and in the SIPP from the question "*when [month and year] did you become limited in the kind or amount of work you can do"*. The date of the injury in the NLSY comes from the question *"in what month and year did the most recent incident occur that resulted in an injury or illness to you"*, and in the SIPP from the question *"when (month and year) did you become unable to work at a job*".

The exposure time for the duration of work disability is measured so as to produce equivalent measures for the two datasets, despite the differences in the structure of the questionnaires. In the NLSY, the exposure time is reported directly via the questions, *"did the (injury/illness) cause you to miss one or more scheduled days of work, not counting the day of the incident?"* and *"not counting the day of the incident, how many days was this?"*. For respondents with missing data for exposure time or for

whom they are still out of work at the time of the interview, the number of days from the occupational injury start date to the return to work is calculated using a weekly work history. This weekly work history file is produced by the CHHR using data from the retrospective work history questions asked at each wave of the NLSY, and it contains a cleaned accounting of weekly labor force status (i.e. in the categories: employed with reported job number; employed but missing data on job number; unemployed; out of labor force; active military; or not working but undetermined if unemployed or out of labor force) for all respondents in the NLSY. The occupational injury start date is reported in month and year and is located within the work history file as the first week of the month of respective injury date. Respondents are followed forward in the work history until they return to work (i.e. labor force status changes to 'employed') or they are censored from the work history panel. Respondents are excluded who have never worked. In the NLSY censoring occurs at the date of the last interview wave. From the 805 respondents observed with occupational injuries in the NLSY, data on exposure time can be coded for 407 respondents, three of whom are censored before returning to work[2].

In the SIPP, there is no question asked about the number of days out of work, so the duration of work disability is obtained using a retrospective work history and work history file similar to that used for the NLSY. The work history file in the SIPP is more complete in terms of work status—covering both labor force status and leaves of absence—but it does not cover as long of a period of time as that in the NLSY. Due to the structure of the SIPP questionnaire, in which the occupational injury questions are asked only in wave 2 of the survey, the measurement of exposure time must be divided

[2] Respondents who are not included in the return to work analysis included 146 respondents with work limitation dates prior to the job injury, and 216 respondents with work limitation dates more than one year after the job injury.

into two parts. The first piece of the exposure time entails time out of work prior to the wave 2 interview when the occupational injuries were first reported, and the second piece of the exposure time entails the exposure time following the wave 2 that the respondent remained out of work. Respondents are excluded from analysis if they have never worked.

The first piece of the exposure time is the number of weeks from the date of the occupational injury to the wave 2 interview. Respondents are excluded who have not been continuously out of work throughout this period. The first step in identifying whether a respondent has been continuously out of work, involves the use of the SIPP work history file. The SIPP work history file accounts for the weekly work status of respondents (i.e. in the categories: with job or business and working; with job or business and not on lay off but absent without pay; with job or business and absent with pay; no job or business and looking for work or on layoff; no job or business and not looking for work and not on layoff) and it covers the four months prior to wave 1 of the survey and all subsequent weeks until the last observed wave of the survey. Respondents are excluded if they have not been continuously out of work from the date of the occupational injury or (if it is prior to the beginning of the work history) the first week in the work history file though to the date of the wave 2 survey. Respondents with an out of work series that is broken by a return to work before wave 2 or missing data are also excluded. All respondents whose injury occurs within the work history panel and who are identified as continuously out of work are given a wave 2 exposure time of the number of weeks between the occupational injury start date and the date of the wave 2 interview. For respondents whose occupational injury date occurs prior to the start of the work history file there is a second exclusionary step. They are not only excluded if they are observed as not continuously out of work though the work history file, but also if their last reported work date before the work history panel is not coincident with the

injury date. The last reported work date comes from the SIPP Employment History Supplement conducted in wave 1. Respondents whose last work date occurs before the injury date are excluded, as are those whose last work date is more than twelve months after the injury date. All respondents identified as continuously out of work from the injury date to the wave 2 interview are assigned a wave 2 exposure time that equal to the number of weeks between the occupational injury start date and the date of the wave 2 interview.

The second piece of the exposure time is the number of weeks after the wave 2 interview that a respondent remains out of work. Respondents are followed in the work history panel from the wave 2 interview to the week that the employment status changes to "working" or until the work history series is broken by missing data or an end of the survey panel. The date of the 'return to work' or censoring from missing data or the end of the survey panel is identified and the number of weeks between this date and the wave 2 interview are reported in the final exposure time. From the 1,200 respondents observed with occupational injuries in the SIPP, data on exposure time can be coded for 380 respondents, 270 of whom are censored before returning to work.

Measurement of Social and Demographic Factors

In the NSLY and the SIPP, fixed demographic factors include the respondent's date of birth, age at injury (for the RTW analyses), sex, and racial and ethnic origin. In addition, educational attainment, marital status, family size and metro residence are as reported in wave 2 of the SIPP and at each survey wave of the NLSY. Ideally we would code these at each survey wave of exposure to RTW after wave 2 also in the SIPP. In practice, however, they are expected to change little over the period of exposure to RTW after wave 2. All variables are constructed to be categorized consistently in the NLSY and the SIPP.

Statistical Analyses

We conduct comparisons to establish the similarity or dissimilarity between the NLSY and the SIPP in the overall proportions experiencing occupational injuries, their hazards of return to work (RTW), and the social and demographic predictors of injury and RTW. Analysis of the rate of workplace injury is conducted first (Table 1). The rate of workplace injury is calculated in both data sets and tabulated by social and demographic factors. For these analyses, we restrict the sample to respondents from the same cohort years, matching ages and years in the SIPP panels with the NLSY cohort's ages over the same years. In specific this entails the restriction of the sample to adults age 35-40 years old observed in the NLSY in 1996, 1998, or 2000, and in the SIPP in the 1996, 2001, or 2004 panels. All injuries are restricted to those occurring after the first observation period of the NSLY in 1988.

The second set of analyses entails a comparison of social and demographic predictors of the duration of work disability before return to work. The length of time until return to work is estimated as a discrete hazard function, where the "hazard" of return to work is estimated monthly among for persons who reported a work-limiting injury that resulted in the person being unable to work:

$$P(Y) = f(\text{exposure time, demographic factors, social factors}) \quad (1)$$

The dependent variable Y equals 1 if the individual reports return to work, and equals 0 if he/she continues to report being out of the workplace/workforce and work disabled. Demographic factors are fixed in time and social factors are fixed in time in the SIPP (observed at the wave 2 interview) and are time-varying in the NLSY (observed in the year of the exposure time in which return to work is observed). The hazard of return to

work is observed for each month of the total number of weeks observed in the NLSY between the injury date and the return to work or censoring, and in the SIPP between the wave 2 and the return to work or censoring. Both exposure variables and the returns to work are defined as described above. It is noteworthy that in the SIPP the hazards of return to work are not observed from the onset of injury but rather from the date of the wave 2 interview when the occupational injury questions were asked. Thus, respondents enter the hazards regression with the exposure time observed at wave 2 and are followed with a monthly time-varying exposures until return to work or exit subsequent to wave 2. By restricting the analyses of return to work in this way for the SIPP respondents, we are able to address the left-truncation of the exposure time data that ensues from the structure of the questionnaire.

The logistic function for f( ) is used, implying the regressors are linear in the log-odds (Maddala 1983). The probability P(Y) then corresponds to the "hazard" of return to work. This discrete-hazard function allows for flexible specification of the relationship between RTW and the exposure time (duration of time out of work), and for a flexible specification of the interaction between that exposure time and the other regressors. Two different hazards models are fit to both the NLSY and the SIPP data and reported in Table 2. The first Model includes exposure time, age at injury and gender, and the second model includes a broader set of social and demographic variables. The differences between the coefficients estimated in the NLSY and the SIPP are tested using a pooled model that combines the data from both surveys and tests for differences in the effects of all variables in the model with interactions.

RESULTS

Comparison of Socio-demographic Predictors of Occupational Injury

Table 1 displays the sample characteristics of the two datasets employed in the comparison of occupational injury, as well as the rates of occupational injury observed in total and by social and demographic characteristics of the population. The sample universes used to compare the NLSY and the SIPP are identified to be as equivalent as possible in terms of age and survey years. In both surveys, adults age 34 to 40 are selected. The sample in the SIPP is observed between the years 1996 and 2004. The sample in the NSLY is observed between 1996 and 2000[3]. All findings are weighted for the complex survey sampling in the NLSY and the SIPP.

The NLSY and the SIPP both have a nearly equal distribution of men and women, with 50.8% of the population male in the NLSY and 49.4% of the population male in the SIPP. They are also similar in terms of racial and ethnic distribution. The proportion of the population that is white (non-Hispanic and non-black) in the NLSY is 79.4% and the respective proportion in the SIPP is 74.9%. Hispanics are a larger proportion of the SIPP due to the later periods (1996, 2001, and 2004) in which the sample was initiated compared to the 1979 year of the NLSY. Educational attainment is slightly higher in the SIPP, with 59.6% reporting at least high school graduation compared with the 47.5% in the NLSY that had completed more than 12 years of schooling. In both samples, the majority of the adult population is currently married; however the relative proportion of never married and formerly married in the two samples differ slightly. The marital status distribution in the NLSY and the SIPP, respectively, entails 64.2% and 66.3% currently married, 19.1% and 15.8% formerly married, and 16.7% and 17.8% never married.

[3] In order to compare injuries occurring in the years of the SIPP that are observed for the NLSY (i.e. 1996, 1998, and 2000), we combined the data for the NLSY sample over these three years using the cross-sectional sample weights in each year.

The rate of occupational injury for adults is nearly equivalent, with no statistically significant difference in magnitude between the two surveys (p=0.726). In the NLSY 1.4% adults 35-40 years old who were surveyed in 1996, 1998 or 2000 reported an occupational injury. In the SIPP 1.3% of the adults 35-40 years old surveyed in the 1996, 2001 or 2004 reported an occupational injury. Recall that in both surveys, reported injuries were included only if they occurred on or after the first date of observation in the NLSY panel (i.e.,1988).

Socio-demographic differences in injuries are similar for the NLSY and the SIPP with respect to race, education and marital status, but differ with respect to gender and ethnicity. Gender differences in the rate of injury are observed in the SIPP but not the NLSY. In the NLSY, the rate of injury is 1.4 % for men and 1.5% for women. In the SIPP, the odds of injury are 1.8 times greater among men than women, with a rate of 1.6% for men and 0.9% for women. The difference in the rate of injury among females between the SIPP and the NLSY is statistically significant (p=0.040). Racial differences in injury are equivalent across the two surveys, with both finding a1.3 times greater odds of injury among non-Hispanic blacks than non-Hispanic whites. Hispanics have greater odds of injury than non-Hispanic whites in the SIPP, but not the NLSY. The difference in the rate of injury among Hispanics between the SIPP and the NLSY is statistically significant (p=0.038). In both surveys the rate of injury is lowest among currently married adults and highest among formerly married adults. In the NLSY formerly married adults have 1.6 times greater odds of injury than currently married, while in the SIPP they have 1.9 times greater odds.

Comparison of Socio-demographic Predictors of Return to Work

Table 2 displays findings from the discrete time hazards models that predict the likelihood of returning to work among persons who have exited the workplace due to

occupational injury. These models use all available data from the NLSY and the SIPP on occupational injuries that result in an exit from work. Thus the person-month sample sizes in these models are large because they include adults of all ages who report an occupational injury that lead to an exit from work in any of the waves of the NLSY between 1988 and 2000, or in any of the waves following wave 2 in the three panels of the SIPP.

The likelihood of returning to work is assessed in monthly intervals and related to a monthly time-varying assessment of the amount of time a respondent has been out of work since the occupational injury (exposure time). Returns to work are observed from the date of the injury in the NLSY through 2000, and they are observed in the SIPP from the date of the wave 2 interview (when work injury questions were asked) through to the end of the respective SIPP panel (i.e. four years later for the 1996 panel, three years later for the 2001 panel, and two years later for the 2004 panel). Models that estimate the likelihood of return to work as a function of exposure time and social and demographic characteristics of the respondents are fit separately using data from the SIPP and the NLSY. Then, these data are pooled and the differences between the coefficients estimated in each survey are statistically tested using a pooled hazards model that combines the data on returns to work from each dataset.

Model 1 assesses the relationship between exposure time and the likelihood of returning to work, controlling for the gender of the respondents and their age at the time of the injury. In both the SIPP and the NLSY, the likelihood of returning to work is negatively associated with longer exposure time out of work and highly statistically significant ($p<0.001$). The negative relationship between exposure and return to work is larger in the NLSY than it is in the SIPP and the difference in magnitude of the coefficient between the two surveys is statistically significant ($p<0.001$). Age at the injury is also negatively associated with return to work, and is highly statistically significant in

both surveys. Additionally, the coefficients for age are nearly equivalent in the two surveys, with no statistically significant difference in magnitude (p=0.954).

In contrast with these consistent findings between the two surveys, are the findings on the overall likelihood of return to work and the association with gender. The overall hazard of returning to work estimated, as indicated by the coefficient for the constant, is not equal in the two surveys. The likelihood of returning to work, controlling for exposure, age and gender, is lower in the SIPP than the NLSY, and this difference is statistically significant (p<0.001). The direction of the association between return to work and gender also differs between the two surveys. There is a greater likelihood of returning to work observed for men in the SIPP but a lower likelihood for men in the NLSY. Despite this difference in direction, the association with gender is only statistically significant in the NLSY (p=0.042), and the difference in the association between the two surveys reaches statistical significance only at the 0.10 level (p=0.072).

Model 2 evaluates whether there are social and demographic differences in the likelihood of return to work, and whether these differences vary across the two surveys. As before we find that the likelihood of returning to work decreases as exposure time increases and that this effect is highly statistically significant in both surveys, though larger in the NLSY than in the SIPP (p=0.003). We also find again that the likelihood of returning to work decreases with age at the time of the injury and that the small differences in the magnitude of this effect in the NLSY and the SIPP are not statistically significant (p=0.427). The overall lower likelihood of returning to work is also again found to be lower in the SIPP than the NLSY and the difference is statistically significant (p=0.012).

Comparing across the NLSY and the SIPP, there are (with few exceptions) no social disparities in return to work that are statistically significant. The two exceptions are the statistically significant negative effect of male gender in the NLSY and the

statistically significant positive effect of having a racial category of “other” in the SIPP. Neither of these effects is observed in both datasets, and the difference for the respective coefficients is not statistically significant when the data is pooled.

DISCUSSION

In this study, we compared prevalence of having experienced an occupational injury over an approximately 12-year period, and subsequent time until return to work (RTW) among those experiencing occupational injuries, in two multi-purpose, population-representative panel surveys. Our findings demonstrate that the overall rate of occupational injury is consistently observed between the two surveys for adults age 35-40 in the late 1990s and early 2000s, at a rate of 1.3 to 1.4%. In both surveys, the rates of occupational injury are higher for adults with low educational attainment (below high school) and for adults who are formerly, rather than currently married. In further work, we will estimate a hazard regression model to understand differences in rates of occupational injury by socio-demographic characteristics but also controlling for job characteristics. This will be conducted with equations estimated from the surveys in parallel, using the 1988 through 2004 SIPP panels.

The likelihood of return to work was predicted in the reported analyses here using hazard regression analyses also with SIPP panels from 1996 through 2004 only. Samples size constraints on this analysis will be addressed in part with estimation from added data on the full set of available SIPP panels, including observations of return to work from occupational injuries in the 1988 through 2004 SIPP panels. Across both surveys in the reported analyses, longer exposure to past spells of work disability are associated with lower hazards of returns to work, and older age is associated with lower hazards of return to work. The negative effect of age on returns to work is a strong predictor in both surveys, consistent with previous findings (Krause et al. 2001). Other

demographic and social factors are not found to be strong predictors of returns to work. This differs from previous work in which, for example, lower education is associated with slower RTW (McKenzie et al 2006; Turner et al 2007). The small sample sizes of our study may be a problem here, as we did find socio-demographic differentials in rates of occupational injury, for which the "exposed" sample sizes are much larger.

The biggest differences in RTW estimates between the surveys are that the overall likelihood of return to work is lower in the SIPP than in the NLSY and that the magnitude of the negative effect of additional duration of time out of work on RTW is stronger in the NLSY than in the SIPP. That is, while the socio-demographic determinants in RTW are not different between the two surveys, the baseline hazards are different. The findings from the comparison of occupational injuries by socio-demographic factors, in which rates of experiencing occupational injuries between the surveys were similar between the NLSY and SIPP, suggest that differences in the sample of occupational injuries are less likely to be at stake than differences in the reporting of returns to work. This is not surprising given that consistent definition of RTW has been identified as an ongoing challenge for the field (Krause et al. 2001). The potential for RTW differences between the two surveys here follows from their different structures of panel versus retrospective reporting of the work injury (between-wave reports in the NLSY versus retrospective wave 2 reports only in the SIPP), and the more detailed definition of work status in the SIPP than in the NLSY.

References

Butler, R.J., W.G. Johnson, and M.L.Baldwin (1995) Managing work disability: Why first return to work is not a measure of success Industrial and Labor Relations Review 48:452-469.

Cheadle, A., G. Franklin, C. Wolfhagen, et al (1994) Factors affecting the duration of work-related disability: A population-based study of Washington state workers' compensation American Journal of Industrial Medicine 34(2):190-196.

DeLeire, T. (2000) The wage and employment effects of the Americans with Disabilities Act Journal of Human Resources 35(4):693-715.

DeLeire, T. (2001) Changes in wage discrimination against people with disabilities: 1984-93 Journal of Human Resources 36(1):144-158.

Dembe, A.E., J.B. Erickson, and R. Delbos (2004) Predictors of work-related injuries and illnesses: National survey findings Journal of Occupational and Environmental Hygiene 1:542-550.

Dong, X.W. (2005) Long work hours, work scheduling, and work-related injuries among construction workers in the United States Scandanavian Journal of Work Environment and Health 31(5):329-335.

Krause, N., L.K. Dansinger, L.J. Deegan, R.J. Brand, and L. Rudolph (1999) Alternative approaches for measuring duration of work disability after low back injury based on administrative workers' compensation data American Journal of Industrial Medicine 35:604-618.

Krause, N., J.W. Frank, L.K. Dansinger, T.J. Sullivan, and S.J. Sinclair (2001) Determinants of duration of disability and return-to-work after work-related injury and illness: Challenges for future research American Journal of Industrial Medicine 40:464-484.

Maddala, G.S. (1983) Limited-Dependent and Qualitative Variables in Econometrics. New York, NY: Cambridge University Press.

Oleinick, A., and B. Zaidman (2004) Methodologic issues in the use of workers' compensation databases for the study of work injuries with days away from work.

I. Sensitivity of case ascertainment American Journal of Industrial Medicine 45:260-274.

Pole, J.D., R. France, S. Hogg-Johnson, M. Vidmar, and N. Krause (2006) Duration of work disability: A comparison of self-report and administrative data American Journal of Industrial Medicine 49:394-401.

Reville, R.T. (1999) The impact of a disabling workplace injury on earnings and labor force participation," pp. 147-173 in Haltiwanger, J., and J. Lane (eds) The Creation and Analysis of Linked Employer-Employee Data: Contributions to Economic Analysis New York: Elsevier Science, North-Holland.

Reville, R.T., J. Bhattacharya, and L.R. Sager Weinstein (2001) New methods and data sources for measuring economic consequences of workplace injuries American Journal of Industrial Medicine 40:452-463.

Schulte, P.A. (2005) Characterizing the burden of occupational injury and disease Journal of Occupational and Environmental Medicine 47(6):607-621.

Seland, K., N. Cherry, and J. Beach (2006) A study of factors influencing return to work after wrist or ankle fractures American Journal of Industrial Medicine 49:197-203.

Turner, J.A., G. Franklin, D. Fulton-Kehoe, et al (2007) Early predictors of chronic work disability associated with carpal tunnel syndrome: A longitudinal workers' compensation cohort study American Journal of Industrial Medicine 50:489-500.

University of Michigan (no date) The Health and Retirement Study http://hrsonline.isr.umich.edu/

Weinberg, D.H. (2002) "The Survey of Income and Program Participation: Recent History and Future Developments." Washington, DC: US Census Bureau

Westat (2001) Survey of Income and Program Participation User Guide, 3rd Edition Washington, DC: US Census Bureau.

Zabel, J.E. (1998) An analysis of attrition in the Panel Study of Income Dynamics and the Survey of Income and Program participation with an application to a model of labor market behavior Journal of Human Resources 33(2):479-506.

Zwerling, C., N.L. Sprince, R.B. Wallace, C.S. Davis, P.S. Whitten, and S.G. Heering (1996) Risk factors for occupational injuries among older workers: An analysis of the Health and Retirement Survey American Journal of Public Health 86(9):1306-1309.

Zwerling, C., N.L. Sprince, R.B. Wallace, C.S. Davis, P.S. Whitten, and S.G. Heering (1998) Occupational injuries among older workers with disabilities: A prospective cohort study of the Health and Retirement Survey, 1993 to 1994 American Journal of Public Health 88(11):1691-1695.

Table 1 Prevalence of Occupational Injury among cohorts aged 35 to 40 in the National Longitudinal Survey of Youth (NLSY) and the Survey of Income and Program Participation (SIPP)

	NLSY (1996, 1998, 2000)					SIPP (1996, 2001, 2004)					
	% of Population age 35-40 years old	% injured of population age 35-40 years old (n=14,793)	Injured (unweighted) Yes	Injured (unweighted) No	Odds Ratio	% of population age 35-40 years old (n=28,824)	% injured of population age 35-40 years old Yes	Injured (unweighted) Yes	Injured (unweighted) No	Injured (unweighted)	Statistical Test of Difference (NLSY vs SIPP)
Total		1.4	190	14603			1.3	359	28,465		
Sex											
Female	49.2	1.5	89	7468	.-	50.6	0.9	131	14,959	.-	**
Male	50.8	1.4	101	7135	0.9	49.4	1.6	228	13,506	1.8	
Race											
White	79.4	1.2	99	7364	.-	74.9	1.1	247	21,698	.-	
Black	14.1	1.5	57	4462	1.3	12.1	1.4	49	3,458	1.3	
Hispanic	6.5	1.1	34	2777	0.9	13.0	1.9	63	3,309	1.7	**
Education											
<12	9.4	2.6	36	1739	1.5	11.2	2.2	72	3,202	1.4	
12	43.0	1.7	94	6389	.-	29.2	1.6	132	8,305	.-	
>12	47.5	0.8	60	6475	0.2	59.6	0.9	151	16,826	0.6	
Marital											
Current	64.2	1.2	96	8419	.-	66.3	1.1	204	19,143	.-	
Former	19.1	1.9	61	3066	1.6	15.8	2.1	97	4,561	1.9	
Never	16.7	1.4	34	3117	1.2	17.8	1.2	58	4,761	1.1	

Notes: Definition of "injured" in SIPP: current work-limiting condition due to a workplace accident or injury since 1988. Definition of "injured" in NLSY: current work-limiting condition and workplace injury in 2000, or workplace injury on or after 1988. All “% of population” and “% injured” estimates are weighted for complex sampling scheme in NLSY and SIPP. Tests of statistical significance reported as: ** $p < .01$; * $p < .05$.

Table 2 Logistic Regression of the Monthly Hazard of Return to Work

	Model 1: Demographics						Model 2: Social and Demographic Factors					
	SIPP		NLSY		Difference (SIPP-NLSY)		SIPP		NLSY		Difference (SIPP-NLSY)	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Exposure time (months)	-0.019**	0.000	-0.058**	0.000	0.040**	0.000	-0.019**	0.000	-0.052**	0.000	0.033**	0.003
Age at injury	-0.039**	0.000	-0.038**	0.004	-0.001	0.954	-0.043**	0.000	-0.029**	0.033	-0.014	0.427
Male	0.177	0.373	-0.237**	0.042	0.414*	0.072	0.174	0.395	-0.274**	0.026	0.448*	0.060
Race (ref=white)												
Black							0.138	0.625	-0.003	0.982	0.141	0.652
Other							0.938**	0.033	0.253	0.369	0.684	0.191
Education (ref=high school)												
<12 years							-0.197	0.408	-0.224	0.083	0.027	0.921
Marital Status (ref=married)												
never married							-0.249	0.401	0.211	0.194	-0.460	0.174
previously married							0.016	0.948	-0.047	0.757	0.064	0.828
Family size							0.060	0.344	-0.048	0.216	0.107	0.146
Metro							-0.377	0.066	-0.068	0.642	-0.309	0.222
Constant	-1.950**	0.000	0.325	0.460	-2.276**	0.000	-1.629**	0.007	0.318	0.520	-1.946**	0.012
Person-months	8355		2015		10370		8344		1881			
Persons with work injuries	380		407				378		394			

Notes: Socio-demographic predictor variables in the SIPP are observed at Wave 2 of the 1996, 2001, and 2004 panels, when work-limited respondents are asked about a work injury that resulted in that injury (all ages). Returns to work are estimated for months following Wave 2 for respondents who did not work between the injury and Wave 2. Socio-demographic predictor variables in the NLSY are observed in the wave or month of exposure to return to work after injuries in the years 1988 to 2000. . All data is weighted for complex sampling scheme in NLSY and SIPP, with tests of statistical significance reported as: * $p < .01$; ** $p < .05$.