The Journal of Risk and Insurance

© *The Journal of Risk and Insurance*, 2010, Vol. 77, No. 1, 231-260
DOI: 10.1111/j.1539-6975.2009.01343.x

Evaluating Permanent Disability Ratings Using Empirical Data on Earnings Losses

Jayanta Bhattacharya
Frank Neuhauser
Robert T. Reville
Seth A. Seabury

Abstract

Workers' compensation systems are typically designed to assign higher permanent disability benefits to workers with more severe disabilities. However, little or no scientific work exists to guide the design of ratings systems to properly account for the amount of earnings power lost due to disability. In this article, we examine the effectiveness of disability ratings using matched administrative data on ratings and earnings for a large, representative sample of permanent disability claimants in California. We find that while workers with higher ratings do experience larger earnings losses on average, there are large and persistent differences in average earnings losses for similarly rated impairments in different parts of the body. We then explore how adjusting permanent disability ratings to reflect cross-impairment differences in earnings losses can affect the equity of permanent disability benefits. Adjusting disability ratings to account for typical earnings losses reduces cross-impairment differences substantially. The adjusted ratings result in a more equitable distribution of disability benefits across workers with different impairments.

Introduction

Individuals disabled by workplace injuries or illnesses are usually eligible for permanent disability (PD) benefits from workers' compensation. Estimates suggest that there are over 500,000 PD claims nationwide per year (Blum and Burton, 2003). These workers suffer from a wide range of injuries of varying severity, and determining

---

Jayanta Bhattacharya is at Stanford University and NBER. Frank Neuhauser is at University of California, Berkeley. Robert T. Reville is at the RAND Corporation. Seth A. Seabury is at the RAND Corporation. Seth A. Seabury can be contacted via e-mail: seabury@rand.org. The data used in this paper were collected with financial support from the California Commission on Health, Safety and Workers' Compensation (CHSWC). We are grateful to Jeff Biddle, Les Boden, John Burton, and two anonymous referees for helpful comments and to Craig Martin and Sue Polich for excellent programming support. The authors take full responsibility for any errors. The views and opinions presented here are the authors', and do not necessarily reflect those of CHSWC or any institution that the authors are affiliated with.

an equitable scheme to compensate these workers is one of the most daunting policy challenges facing any workers' compensation system. Ideally, an equitable PD ratings system should be structured so that, when it is used to determine the level of PD benefits, more compensation is provided to workers with impairments that cause greater earnings losses (Berkowitz and Burton, 1987). However, basing PD benefit payments directly upon the actual earnings losses suffered by disabled workers creates work disincentives (cf. Krueger, 1990; Butler, 1994), so most states base benefits at least partially upon physician evaluations of the severity of impairment.[1]

The main purpose of PD ratings scales is to convert medical information into a form that permits the necessary comparisons of impairment severity across and within disabilities occurring in different parts of the body. However, the extent to which these scales actually produce equitable results in practice is an open question. In this article, we use matched administrative data on the disability ratings and earnings of a large sample of PD claimants in California to assess the effectiveness of PD ratings. We study the extent to which workers with similarly rated injuries suffer similar postinjury outcomes. Furthermore, we examine the extent to which incorporating information about predicted earnings losses into the disability rating system can improve the equity of compensation for PD.

An important limitation that nearly all current disability ratings systems face is a lack of scientifically validated data measuring the relationship between impairment, disability, and earnings loss (cf. Spieler et al., 2000; Cocchiarella, Turk, and Andersson, 2000). In the absence of such data, the current assignment of disability benefits is based on an implicit, unproven assumption that workers with greater medically detectible signs of impairment will also suffer more severe earnings losses. In fact, past work has cast some doubt on the ability of disability ratings systems to accurately assess the impact of different injuries on earnings capacity. Three 1968 studies that focused on Wisconsin, Florida, and California exhibited mixed results about the relationship between earnings losses and disability ratings (Berkowitz and Burton, 1987). Park and Butler (2000) use data from Minnesota to show that impairment ratings are able to explain only a small fraction of earnings losses subsequent to a disability. Using data from California, Reville et al. (2002) show that different types of musculoskeletal impairments with identical ratings suffered very different long-term outcomes. However, these studies all are limited either in their ability to measure the economic consequences of disability or in the scope of different injuries that they are able to evaluate.

Using data on over 80,000 PD claimants in California, we are able to more completely assess the relationship between disability ratings and earnings losses. Our results show that, on average, the California ratings system predicts losses reasonably well *within* injury types but fairly poorly *between* injury types. In other words, for any given type of injury, a worker with a higher disability rating tends to experience significantly higher earnings losses. However, a worker with a given rating and an injury to one part of the body may experience much different losses than a worker with the same rating but an injury to another part of the body. In an equitable

[1] There are a number of works that compare the different state systems for evaluating and compensating permanent disabilities. A few important examples include Berkowitz and Burton (1987), Barth and Niss (1999), and Burton (2005), but this list is far from comprehensive.

system, individuals with similar earnings losses should receive similar benefits, regardless of the nature of the injury. Therefore, disparities in ratings across body parts can lead to considerable inequities in the amount of compensation provided by PD benefits.

Given these results, we then examine whether or not earnings loss data can improve the rating of permanent disabilities across injuries to different parts of the body. We do this by computing a set of adjustments to the ratings of different body parts based on the average relationship between disability ratings and earnings losses. We compute two sets of adjustments: a linear adjustment that is constant within each category of injury that we consider, and a piecewise linear adjustment that changes within injury category based on the severity of the rating. This is useful to correct for any nonlinearities in the relationship between earnings losses and ratings within injury type. We show that the adjustments, particularly the piecewise linear adjustment, can improve the equity of PD compensation.

The article proceeds as follows. In the next section, we provide some background to evaluating and compensating PD in workers' compensation, with a focus on California. We then discuss the data we use in our analysis. Earnings losses from disability cannot be directly observed so it must be estimated, and we discuss our estimation technique in this section. In the "Methods for Evaluating and Adjusting Ratings" section, we discuss how to compute adjustment factors to disability ratings based on the average relationship between ratings and losses. In the "Results" section, we present our empirical results, demonstrating the disparities in ratings for different impaired body parts and how these are reduced by the adjustment factors. In the "Implications for the Equity of Permanent Disability Benefits" section, we explore the implications of our results for the equity of PD benefits. We conclude with a discussion of the policy implications of our results and some considerations for future work.

Systems for Rating Permanent Partial Disability

To tie disability benefits to the severity of injury, it is necessary to develop a systematic method to evaluate how severe an injury is. Unfortunately, there is no one system that obviously trumps all others. As a result, there is considerable heterogeneity across the United States in how different states choose to evaluate the severity of injury and the ways in which benefits are tied to these evaluations. A full accounting of the different state and federal programs for evaluating disability is beyond the scope of this article. Detailed summaries of the nature and variety of different programs can be found in Berkowitz and Burton (1987), Barth and Niss (1999), Reville et al. (2005), and others. Since we use data on California, however, we focus our discussion on the basics of the California system and how it differs from others nationwide.

In most states, PD benefit levels are determined based at least in part on physician evaluations of injured worker impairment using the American Medical Association's (AMA's) *Guides to the Evaluation of Permanent Impairment*. The AMA *Guides* represent the most widely accepted tool for converting medical evidence in a quantitative measure of impairment (on a scale from 0 to 100). Prior to January 1, 2005, California used its own ratings system that was specifically designed to measure disability.[2] The

[2] Note that our discussion is limited to evaluating permanent *partial* disabilities. Permanent total disabilities are generally treated differently, in terms of compensation and evaluation,

distinction between disability and impairment is an important one; whereas disability represents the impact of an injury on the ability to work, impairment reflects the limitations on activities of daily living (ADLs). In practice, the focus on disability in the California system led to a greater reliance on the subjective evaluations of physicians about the diminished work capacity of injured workers (e.g., determinations about the extent to which they should refrain from heavy lifting).

In California, permanent partial disability benefits are tied directly to the level of measured disability severity. This is different from states that schedule disability benefits according to the type of body part injured, or wage loss states that base benefits on postinjury labor market outcomes. The key feature of the California system is that each disability is evaluated individually by a physician, and this evaluation leads to the assignment of a number between 1 and 100 indicating the percent of disability (called a disability rating). During the time frame our data cover, the method through which physician evaluation were converted to ratings was based on a unique schedule used only in California.

The California system had a number of aspects that make it controversial. It was the only state to allow disability ratings based solely on prophylactic work restrictions, recommended limits on the amount or type of work that physicians make that are not necessarily based on "objective" organic signs of injury. The use of these restrictions led many to complain that the system was too subjective. The reliance on physician evaluations generates incentives to "shop" for physicians to receive favorable ratings (Reville et al., 2005; Seabury, Reville, and Neuhauser, 2006). A vast majority of claims in California are settled as lump sum benefits (Peterson, Reville, and Stern, 1997), so there are concerns that workers fail to adequately save benefits for future medical needs. California has historically compensated many psychiatric claims not compensable in other states, including "mental–mental" disability claims (claims with both nonorganic causal factors, such as stress, and outcomes, such as depression), although the laws governing the compensation of psychiatric claims were tightened to require at least 50 percent job causation in 1993.

The controversy surrounding the California system has led to many periods of reforms, including during the time frame of our analysis. However, most reforms tended to either target overall benefit levels, or other factors that affect employer costs (such as the cost of medical treatment). None of these reforms should necessarily affect the cross-impairment inequities that are the focus of this article.

DATA ON DISABILITY RATINGS AND EARNINGS LOSS

The purpose of a disability rating is to incorporate all available medical information into a single index of the extent to which an individual's ability to work is reduced. To evaluate these ratings, we need an alternative measure of the work limitations that result from disability. Our approach is to estimate the lost earnings capacity of individuals and then directly compare these estimates to the associated disability ratings.

and are excluded from the analysis in this article. In general, there are far more permanent partial than permanent total cases.

The main challenge in estimating earnings loss is measuring potential earnings—what a worker would have earned had the disabling injury not occurred. Obviously potential earnings are unobservable, so to estimate the impact of disability on earnings we must estimate potential earnings. One estimation approach used in Park and Butler (2000) as well as the 1968 studies described by Berkowitz and Burton (1987) is to use preinjury earnings, combined with additional controls such as age. Such estimates are limited, however, because preinjury earnings ignore earnings growth that workers experience over time. We use an alternative approach that involves identifying matched controls consisting of uninjured workers who are observably similar to injured workers. This approach follows directly from Reville (1999), Reville and Schoeni (2001), Boden and Galizzi (2002), and a number of related RAND studies, so we discuss it only briefly.[3]

The key to this estimation approach is to find workers with close enough characteristics that suffering a permanently disabling injury (and filing a claim) can be seen a random event. Under an appropriate match, the earnings of the uninjured control workers are the estimate of an injured worker's potential earnings, and the difference between potential earnings and actual earnings provides an unbiased estimate of earnings losses. Formally, we can define the cumulative earnings loss, denoted L, between the time of injury, $t = 0$, and some future date, T, with Equation (1)

$$L = \sum_{t=0}^{T} \beta^t \left(y_t^{\mathrm{U}} - y_t^{\mathrm{I}}\right), \tag{1}$$

where β is the discount rate, y^{U} is the average of the control worker earnings, and y^{I} are the earnings of the injured worker.

For our primary outcome measure we use proportional losses: cumulative earnings losses divided by cumulative potential earnings. Formally, proportional losses, denoted PL, can be defined as

$$PL = \frac{\sum_{t=0}^{T} \beta^t \left(y_t^{\mathrm{U}} - y_t^{\mathrm{I}}\right)}{\sum_{t=0}^{T} \beta^{\mathrm{t}} y_t^{\mathrm{U}}}. \tag{2}$$

From a conceptual standpoint, the use of proportional losses can be problematic if there is specification error in our estimate of potential earnings. Control worker earnings y^{U} represent an estimate of true potential earnings, so suppose that $y^{\mathrm{U}} = y^* + e$, where y^* is the true value of potential earnings and e is a random, mean zero error term. If we consider only the level of earnings loss, given by Equation (1), this error term is additively separable. As long as the error term is "well behaved," then e acts as a classical measurement error term and will not affect the results of OLS

[3] This approach was inspired by the training program evaluation literature (cf. Dehejia and Wahba, 1999; Heckman and Hotz, 1989; Lalonde, 1986).

when the quantity in Equation (1) is used as the dependent variable (Greene, 2002). However, if we substitute $y^{\mathrm{U}} = y^{*} + e$ into Equation (2) it is clear that the error term is not additively separable. This will potentially bias our estimate of proportional losses and could corrupt estimates using OLS if the quantity in Equation (2) is the dependent variable.

Note that the magnitude of any bias introduced by this problem will diminish as the specification error goes to zero. We have experimented with using the level of earnings losses as the dependent variable in the regressions we employ later on and find nearly identical qualitative results. This suggests that the size of the specification error is small, and the issue is probably not too important empirically. Thus, despite the potential for bias, we report results using proportional losses for two reasons: they are more relevant to the public policy discussions in California and they are the most directly comparable to disability ratings.

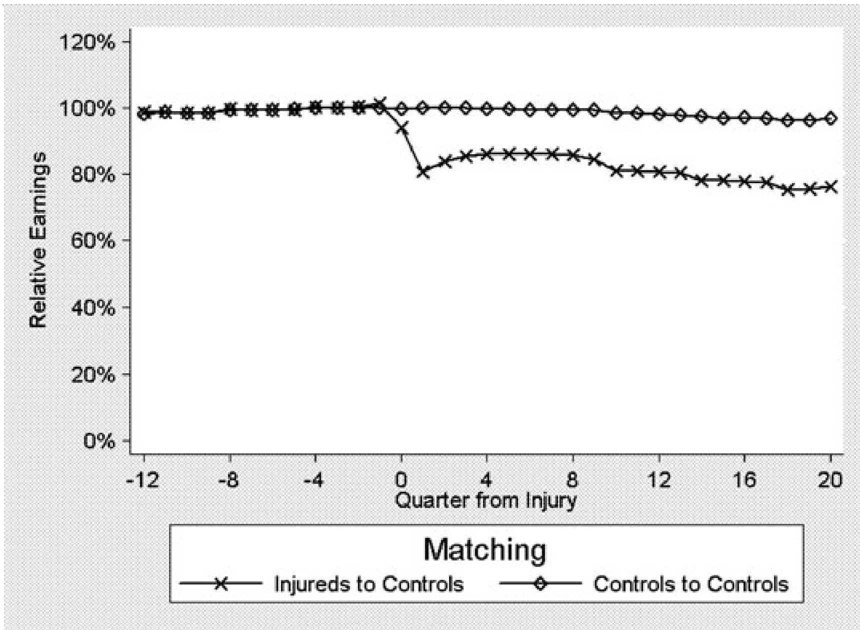
We estimate proportional losses using earnings data from quarterly unemployment insurance (UI) rolls collected by the California Employment Development Department (EDD). The data on earnings come from the Base Wage file maintained by the California EDD. Every quarter, employers covered by UI in California are required to report the quarterly earnings of every employee to the EDD, and the RAND data cover every worker from 1991 to first quarter 1999.

For each injured worker, we pick two to five uninjured workers at the same firm and with similar preinjury earnings and tenure at the date of the injured worker's injury but who had not themselves suffered a disabling workplace injury up to that point. We required injured workers to have at least two controls because potential earnings with only one control were highly variable (thus potentially biasing proportional losses, as discussed earlier). Workers with a single control also appeared particularly sensitive to a small number of observations with extreme negative proportional losses (a problem that can arise because proportional losses are mechanically bound by 1 but unbounded below). The EDD data allow us to select uninjured control workers with very similar earnings to the injured worker for a fixed period of time prior to the injury (we use four quarters). Importantly, the EDD data allow us to match injured and control workers *within firm*. Using workers within the same firm allows us to control for a number of important characteristics (such as industry and local economic conditions) and allows us to condition on employer heterogeneity.

On the other hand, the EDD data are limited because they do not allow us to match according to other key demographic factors such as age or gender, because these data are not available in unemployment insurance records (so we do not have them for the uninjured workers). In addition, our matching strategy is only effective for injured workers with four quarters of preinjury earnings. Workers who are injured early on the job likely have lower earnings loss on average, because they have lower tenure and thus lower earnings, which may cause us to overstate losses for workers on average (although it is less clear if this would affect proportional losses).

A key issue with matching estimators is to ensure there are no systematic differences in earnings profiles prior to the match, and the only systematic differences subsequent to the match should be driven by the injury. We only use a limited period of time (four quarters) prior to the injury for the purposes of matching. If the control group is valid,

FIGURE 1
Evaluating the Quality of the Match Between Injured and Control Workers



we would expect to find no difference between earnings of injured workers and their controls prior to the matching period. To test for systematic differences subsequent to the match, we randomly select one control worker from each group and compare its earnings to the earnings of the other controls.[4] For the controls matched to other controls, we would expect to find no systematic differences either before or after the match period if the assumptions underlying the matching process are valid.

The results of these tests are illustrated in Figure 1. The figure illustrates relative earnings, the percent of average earnings of the treatment workers relative to the average earnings of the control workers (so that 100 percent is equivalent to identical earnings), for the injured workers matched to controls and for the randomly selected control workers matched to controls. We illustrate 3 years (12 quarters) prior to injury and the 3 years after injury. Quarter 0 denotes the quarter of injury.

Figure 1 shows that the injured workers' wages are very close to 100 percent of the control workers' wages (i.e., they are very close to equal) for the entire 3 years prior

[4] Because our matching procedure called for selecting workers with similar earnings and similar tenure at the same firm, matching a randomly selected control worker to other controls who were matched to the same injured worker is analogous to replicating our matching efforts on a random sample of noninjured workers who are observably similar to injured workers prior to an injury. Note that we could not match a fully random sample of noninjured workers, such as those at firms with no injured workers, because we were not granted access to data for these.

to injury. Even though the matching is only based on quarters 1 to 4, we see no systematic difference in earnings prior to the match period. However, there is clearly a systematic difference in earnings for the injured workers and their controls, with relative earnings of about 80 percent. Moreover, relative earnings are reasonably flat between quarters 12 and 20, suggesting that focusing on average losses 3 years after injury will be representative of longer-term losses.

The line illustrating the earnings of controls matched to controls shows no systematic differences in earnings either prior to or subsequent to the match. Relative earnings are close to 100 percent for these workers over the entire time period. This is suggestive that our control group is valid, and the systematic decline in earnings of injured workers is a direct result of their injuries.

The central identifying assumption for our earnings loss estimates to be unbiased is that earnings four quarters prior to injury summarize all unobserved heterogeneity that is correlated with the postinjury earnings of injured workers. Strictly speaking, having an unbiased estimate of earnings losses is not critical for our purposes. If earnings losses have some bias, nearly all of the qualitative results we discuss in this article will hold as long as the bias is uncorrelated with the disability rating that a worker receives. Given that the disability ratings we use are entirely functions of injury characteristics, this seems like a plausible assumption. For a further discussion of factors that could confound earnings loss estimates, see Reville and Schoeni (2001).

We focus on proportional losses 3 years after injury in our empirical implementation. Three years is an appropriate amount of time because the proportional losses of most permanently disabled workers stabilize by then, as can be seen in the flatness of the line in Figure 1 from quarter 12 on. Longer time periods result in fewer observations due to limited postinjury earnings data.[5] Finally, we note that all dollar values are reported in 2003 dollars and we use an annual discount rate of 2.3 percent (Reville, 1999).

For information on disability severity, we match the earnings loss data to information on disability ratings from the California Disability Evaluation Unit (DEU). The DEU performs between 60,000 and 80,000 ratings of permanent disabilities each year, and RAND sampled all evaluations done between January 1, 1991 and April 1, 1997. These data provide a substantially complete quarterly earnings history for California PD claimants.

We sampled only disability evaluations that were performed by neutral physicians (i.e., physicians who were not selected by either the applicant or the defense) because nonneutral physicians have been shown in past work to be biased, and their ratings do a worse job of predicting losses (Boden, 1992a, b; Seabury, Reville, and Neuhauser,

[5] Others have found that serious cases can take as long as 5 years to stabilize, suggesting that 3 years is possibly too short a time horizon (cf. Cheit, 1961). Our data suggest that while this is undoubtedly true in some cases, earnings loss appear to stabilize *on average* after 3 years. Relative earnings of injured workers in our data are approximately 83 percent in quarter 12, 81 percent in quarter 16, and 84 percent in quarter 20 (although, as suggested earlier, we have fewer observations for these last two quarters).

TABLE 1
Match Rate of Injured Workers by Size of At-Injury Employer

Number of Employees	% of Injured Workers Matched
0–24	3
25–49	16
50–99	33
100–249	59
250–499	79
500–999	88
1,000–4,999	95
5,000+	99
All	68

2006). Also, we kept only cases that involved individuals reporting only one disability over our study time frame and involved only a single impairment.[6] In our sample, we are able to match approximately 68 percent of injured workers to at least two uninjured control workers. The driving factor behind a failure to match is the size of the at-injury employer. This is demonstrated in Table 1, which shows the fraction of injured workers who are matched to control workers by the size of the at-injury employer. Workers at firms with less than 50 workers are matched rarely because it is difficult to find other workers at the same firm with similar earnings. As firm size increases, however, the match rate grows considerably, with approximately 99 percent success for workers at firms with 5,000 employees or more.

In principle, the failure to match in small firms could bias our results, since it is known that workers at small firms tend to differ according to wages and other characteristics (cf. Oi and Idson, 1999), an issue that has largely gone unaddressed in past studies. We can correct for potential bias, however, using techniques suggested in Little and Rubin (1987). Specifically, we can model the probability of a match for an individual as a function of observable characteristics as

$$\Pr[match_i = 1] = X_{it}\beta + \lambda_t + \varepsilon_{it}, \tag{3}$$

where X is a vector of individual characteristics, such as preinjury earnings and firm size, for worker i who was injured in time t; λ is a fixed time effect; and ε is a random error term. We estimate this relationship using a probit model, to get predicted values of p, the probability of a match. Using these predicted values, $\hat{p}$, we can generate weights $w = 1/\hat{p}$; in other words, we place greater weight on observations that are less likely to be sampled. In practice, Little and Rubin (1987) argue that using the actual predicted value as the weight may introduce variability into the estimates.

[6] For multiple impairment cases we included cases that listed impairments in separate specific body parts or that simply claimed general upper or lower extremity impairments. These latter cases were excluded because they can involve multiple impaired body parts (e.g., a shoulder and an elbow).

They suggest partitioning the data into quantiles of the predicted values, so we group the data into quintiles and use the quintile average of the inverse of the estimated propensity score as the weight. As long as a failure to match is driven by selection on observable characteristics, this will produce unbiased estimates of earnings loss. All the subsequent results (and the previous results in Figure 1) are calculated using these weights.[7]

METHODS FOR EVALUATING AND ADJUSTING RATINGS

Disability ratings systematically incorporate beliefs about the impact of impairment in different parts of the body on an individual's ability to work by providing different scales based on the type of injury. For instance, in the rating schedule used in our data, "major immobility of a single elbow joint in 70° to 100° flexion" is given a standard rating of 35 percent. The closest analogy in the lower extremities, "immobility of one knee joint in 90° flexion," is given a standard rating of 65 percent (California Department of Industrial Relations, 1990).[8] This example suggests that the designers of the California system believed immobility in the knee to be more disabling than immobility in the elbow. Similar implicit judgments are made throughout the rating schedule.

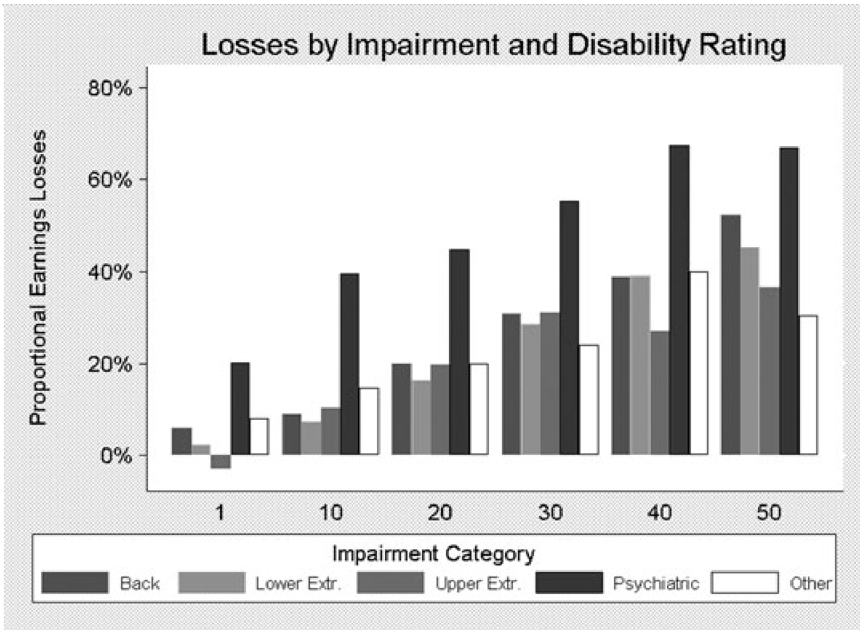
If disability ratings are able to successfully rank the impact of all medical information on an individual's ability to work, then we would expect to see identical, or at least very similar, earnings losses for identically rated claims in different parts of the body. Figure 2 casts some doubt on the success of disability ratings in this regard. In the figure, we compare the average proportional losses of claimants with injuries in different body parts by disability rating. We average disability ratings in intervals of 10 ratings point, so the category with disability rating 10 would actually present the average proportional losses of ratings from 6 to 15 percent. The lowest category, denoted 1, represents ratings between 1 and 5. Also, the category 50 includes all ratings higher than 45, because the sample size of ratings for specific different body parts becomes quite small for ratings over 50.

The figure demonstrates substantial differences in average losses for different types of impairments. The most striking difference is for psychiatric claims, which have average earnings losses close to 20 percent even for the lowest rated claims. Although the differences are smaller among the other impairment types, they are still persistent. Impairments to the lower extremity appear to have comparatively low losses, except for the highest rated claims. What is interesting, however, is that for all impairment types there does appear to be a positive correlation between ratings and losses. This suggests that ratings do have predictive power, but it appears that there could be problems in the predictions across different types of impairments.

[7] In practice, however, the qualitative results are essentially the same with our without weighting, suggesting that the oversampling of workers at larger firms does not have a significant impact on the results.

[8] These ratings are for cases with "objective" medical evidence of disability in the affected joints. If the physician has a "subjective" belief about an injured workers' diminished capacity for work, this could result in ratings different from those described here. In our data, we have the actual rating assigned based on the physician's evaluation.

FIGURE 2
Average Proportional Losses for Different Impairments by Disability Rating



Evaluating the Predictive Power of Ratings

Although the results of Figure 2 are suggestive, it fails to provide a complete evaluation of the ratings system. For one, it ignores the possibility of confounding factors such as correlation between injuries to different parts of the body and the type of industry in which individuals work. It also only considers aggregates of the impairment categories. Finally, it gives us no information as to the extent to which the differences observed are significant in a statistical sense.

To complete the analysis, we examine the link between earnings losses and disability ratings in California using a simple OLS model. The dependent variable is the 3-year discounted proportional earnings loss, and the key independent variable in the regression is the California disability rating, reported as a percentage. We use the California *standard* rating, the rating without adjustments for age or occupation, which is most similar to the impairment rating obtained by the AMA *Guides* (although it incorporates additional subjective factors that are not used by the *Guides*). If the disability rating system were perfectly predictive of disability, as measured by 3-year proportional losses, then the regression coefficient for the disability rating variable would be exactly equal to 1. Although it is unlikely that the rating is a perfect predictor of losses, we can say that the closer the estimated coefficient is to one the more predictive is the ratings schedule.

In the regression, we include binary indicators for 10 separate impairment categories: back, hip, knee, ankle, shoulder, elbow, hand or wrist, psychiatric, hearing, and all other single impairments. These injury fixed effects allow us to examine the extent to which the average value of losses varies across different types of injuries, conditional on the disability rating. In addition to the fixed effects, we include slope interaction terms between the disability rating and each of the injury dummy variables. This allows us to see how the predictive power of ratings differs across injury categories.

To account for regional economic conditions that may affect the earnings capacity of the disabled, we break the state into three regions: Southern California, the Central Coast region, and Northern California. We control for firm heterogeneity by including selected firm characteristics in the regression. These characteristics come from the EDD data and include the industry of the firm, the number of employees at the firm, and whether or not the employer was self-insured.[9] Firm characteristics are potentially important because, even though we difference out the outcomes of control workers at the same firm at the time of injury, the firm characteristics may reflect differences in employer policies that reflect postinjury outcomes (e.g., larger employers may have policies that promote return to work and are more likely to bring workers back to their preinjury job, lessening long-term earnings). We also include preinjury wages in the quarter prior to the injury to control for the incentives of workers to return to work, and include dummy variables for the year of injury. Our final sample includes 82,556 observations.[10]

Table 2 provides some summary statistics for the data in our regression sample. We see that the average preinjury quarterly earnings are \$10,067 in year 2003 dollars, which corresponds to annual income of \$40,268. The average claimant in our sample worked at a firm with about 5,441 employees at the time of injury. Approximately, 17 percent of workers were at self-insured employers. We expect that workers with higher wages will have greater incentives to return to work, although larger employers and self-insured employers will have greater ability or desire to accommodate injured workers, so we expect all three of these variables to be negatively related to earnings losses.

Table 2 also shows us that impairment to the neck, spine, or pelvis (a category that we simply label as "back") is the dominant type of impairment in our data, accounting for approximately 41 percent of all cases. The next largest category is impairment in the hand or wrist, which accounts for 25 percent of cases. Psychiatric impairments are a very small category, less than 2 percent of cases, but bear special consideration given the high losses they experience and the disproportionate degree of interest that these cases inevitably generate.

Probably the most important fact about the industry distribution in our data is that we see a fairly wide variety of industry types. In fact, the service industry is the single most common industry type for injured workers in our sample, representing 28 percent of cases. The second most common is manufacturing, with approximately

[9] Because we have repeated observations on many firms, an alternative approach is to include firm fixed effects in the analysis. We have run alternate specifications not reported in the article that include firm fixed effects and have found that they have no impact on the qualitative results.

[10] This is less than the number of matched claimants because we do not have 3 years of postinjury earnings for all workers.

TABLE 2
Selected Summary Statistics ($N = 82{,}556$)

	Mean	SD
Preinjury quarterly wage ($)	10,067	6,777
At-injury firm size (workers)	5,441	12,306
At-injury employer self-insured (%)	16.8	37.4
Impairment Category	**Number**	**%**
Back	33,466	40.5
Hip	412	0.5
Knee	10,770	13.1
Ankle	3,587	4.4
Shoulder	6,044	7.3
Elbow	2,412	2.9
Hand or wrist	20,941	25.4
Psychiatric or other mental	1,281	1.6
Hearing	1,477	1.8
Other impairment type	2,166	2.6
Industry (Firm Level)	**Number**	**%**
Agriculture, forestry, and fishing	3,121	3.8
Mining	523	0.6
Construction	3,518	4.3
Manufacturing	17,676	21.4
Transportation and utilities	7,556	9.2
Wholesale trade	4,542	5.5
Retail trade	10,333	12.5
Finance, insurance, and real estate	3,501	4.2
Services	23,100	28.0
Public administration	8,688	10.5
Region of the State	**Number**	**%**
Southern California	39,199	47.5
Central Coast Region	14,042	17.0
Northern California	29,314	35.5

21 percent of cases. Finally, we note that nearly half of the injured workers in our sample worked in Southern California at the time of injury.

Adjusting Ratings Using Earnings Loss Data

In addition to analyzing the unadjusted disability ratings schedule, we apply the same methodology to ask how equitable the ratings schedule would have been had disability ratings been adjusted to reflect earnings losses. The 2004 California reforms require that disability ratings reflect "the average percentage of long-term loss of income resulting from each type of injury for similarly situated employees."[11] We consider

[11] California Labor Code, Section 4660, Part b, Paragraph 2.

two adjustment schemes: a simple linear adjustment and a piecewise linear adjustment.[12] The simple linear adjustment is computed as a constant, impairment-specific multiplier based on the average ratings and earnings losses of a given impairment type relative to that of some baseline impairment. The piecewise linear adjustment is calculated the same, except we use a different adjustment factor based on the reported severity of the impairment (reflecting the fact that the relationship between earnings losses and ratings may not be constant over injury severity). For a baseline, we choose back impairments because they are the most common.

Formally, we can define the adjustment factors as

$$A_i = \frac{R_B/L_B}{R_i/L_i}, \tag{4}$$

where A_i is the adjustment factor for injury type i, R_B is the average rating for the baseline injury type, L_B represents the average proportional losses for the baseline injury type, and R_i and L_i are the average ratings and losses, respectively, for injury type i. For the linear adjustment factor, we calculate a single adjustment factor for each injury type. The piecewise linear adjustments are calculated exactly the same, except we compute adjustments using the average ratings and losses for cases reporting severity in each of five separate ranges: 0–10, 10–20, 20–35, 35–50, and over 50.

Table 3 provides the necessary information to compute the simple linear adjustment factors for all 11 impairment categories. In the table, we present the average ratings and proportional losses for each impairment category. Hip, back, and psychiatric impairment are the highest rated on average, each with an average standard rating of approximately 20 percent. Psychiatric impairments have the highest average losses by far, which is no surprise given Figure 2. In Table 3, we also list the maximum PD award that individuals are eligible for, based on the individual disability ratings. The highest PD awards are given to hip impairments, averaging over \$19,514 in year 2003 dollars.[13]

We illustrate the calculation of the adjustment factors using two impairments: shoulders and backs. In our data, shoulder impairments have average proportional losses of 0.179 and average ratings of 0.099. Back impairments have average proportional losses of 0.213 and average ratings of 0.199. Suppose for a moment that the average

[12] We do not incorporate the actual rating scheme used in California because the reforms also required the simultaneous adoption of the AMA *Guides*. As we have no data on earnings losses matched to AMA *Guide* ratings, we instead focus on a hypothetical exercise that examines how adjustments would have affected the old California ratings. See the Appendix for a comparison of the ratings we use to a sample of ratings under the new California system.

[13] PD benefits are distributed weekly, with the number of weeks reflecting the disability rating, so the totals in Table 2 incorporate quarterly discounting at an annual rate of 2.3 percent (consistent with the earnings losses). The weekly value of PD benefits are two-thirds of wages subject to a cap, but the cap is small enough that about 94 percent of workers qualify for the benefit maximum. Therefore, we simply assume workers qualify for the maximum when calculating the implied benefits. We also impute the value of a life pension that is associated with disability ratings of 70 or higher. For those eligible for a life pension, we assume that the injured worker has a life expectancy of 80 and compute the total discounted present value based on the number of years between current age and 80 years.

TABLE 3
Comparison of Ratings, Losses, and Benefits by Impairment Category

Impairment Category	Assigned Disability Rating	Average Proportional Losses[a]	PD Award[b]
Back	0.199	0.213	15,178
Hip	0.222	0.230	19,514
Knee	0.148	0.133	10,959
Ankle	0.142	0.127	10,990
Shoulder	0.099	0.179	6,221
Elbow	0.094	0.092	5,922
Hand or wrist	0.109	0.075	7,421
Psychiatric and other mental	0.212	0.456	19,384
Hearing	0.107	0.198	8,098
Other	0.159	0.126	13,210
All impairments	0.154	0.159	11,459

[a]Proportional earnings loss is defined as the ratio of 3-year cumulative postinjury earnings over the 3-year cumulative "potential" earnings. Potential earnings are estimated with the cumulative earnings of uninjured "control" workers, workers employed at the same firm as injured workers with similar tenure and preinjury wages.
[b]The PD award is calculated as the total amount an injured worker is presumptively entitled to based on the final disability rating (which incorporates statutory adjustments for age and occupation). Actual disability benefits may vary because of settlement, apportionment, or challenges to the claim. The benefits are reported in 1997 dollars and are discounted quarterly using an annual rate of 2.3 percent.

proportional losses were the same. Taking back impairments as the baseline, we must multiply the ratings of shoulder impairments by 2.01 to equalize them with back ratings on average. However, it makes little sense to equalize the ratings if the losses are not equal. The average proportional losses of back impairments are approximately 1.19 times higher than the average losses of shoulder impairments. Therefore, in this example the appropriate adjustment factor for shoulder impairments is 1.690 (approximately equal to 2.01/1.19), raising the average rating to 0.169. Because backs are the baseline impairment, their ratings are unchanged.

Repeating this exercise for each of the impairment categories generates a full set of adjustment factors. Table 4 lists the full set of linear and piecewise linear adjustments. Note the numbers in the table and the text differ due to rounding.[14] Back injuries are the baseline category, so the adjustment factor is 1 in all cases. For the linear adjustment, the factors range from a low of 0.641 for the hand or wrist to a high of 2.005 for the psychiatric and other head cases. The piecewise linear adjustment factors vary significantly, but in no clear pattern. The psychiatric adjustments are quite high for the less severe injuries but become lower for the more severe cases. The same pattern holds even more strongly for shoulder injuries or hearing impairments, for

[14] Note that the adjustments in Table 4 were calculated using up to 15 significant digits. This means that calculating the adjustments manually using the data reported in Table 3 will lead to adjustments that differ from those in Table 4 by more than what we might typically expect from rounding error.

TABLE 4
Disability Rating Adjustment Factors by Impairment Category

	Linear Adjustment[a]	Piecewise Linear Adjustment[b]				
		Injury Severity Range				
		0 to 10	10 to 20	20 to 35	35 to 50	50 and Up
Back	1.000	1.000	1.000	1.000	1.000	1.000
Hip	0.971	2.052	1.188	1.137	0.296	0.862
Knee	0.842	0.863	0.684	0.931	0.955	0.909
Ankle	0.840	0.813	0.865	0.862	0.930	0.810
Shoulder	1.690	2.193	1.408	1.261	0.893	0.716
Elbow	0.917	0.856	0.922	0.909	0.319	0.934
Hand or wrist	0.641	0.020	0.854	0.931	0.692	0.686
Psychiatric	2.005	4.415	2.470	1.650	1.528	1.092
Hearing	1.723	3.944	1.612	1.140	0.706	0.354
Other	0.742	1.516	0.205	0.728	0.531	0.734

[a]The linear adjustment factor is calculated in two steps. First, we divide the average proportional losses for a given impairment category by the average disability rating for that category. Then we divide this result by the ratio of the average proportional losses to the average disability rating for back impairments, approximately equal to 1.070.
[b]The piecewise adjustment is calculated identically to the linear adjustment, except separate factors are calculated for injuries with reported ratings in each of the five different severity ranges.

TABLE 5
The Impact of Adjustments on Disability Ratings and Benefits

	Linear Adjustment			Piecewise Linear Adjustment		
	Adjusted Rating	Adjusted PD Award	Increase on Unadjusted Award	Adjusted Rating	Adjusted PD Award	% Increase on Unadjusted Award
Back	0.199	15,178	0%	0.199	15,178	0%
Hip	0.215	18,345	−6%	0.223	17,877	−8%
Knee	0.125	8,596	−22%	0.125	8,889	−19%
Ankle	0.119	8,478	−23%	0.121	8,603	−22%
Shoulder	0.167	12,469	100%	0.160	10,690	72%
Elbow	0.086	5,299	−11%	0.082	4,975	−16%
Hand or wrist	0.070	4,193	−44%	0.068	4,596	−38%
Psychiatric	0.388	46,766	141%	0.428	43,432	124%
Hearing	0.181	16,650	106%	0.179	12,621	56%
Other	0.118	8,713	−34%	0.122	8,770	−34%
All impairments	0.148	11,107	−3%	0.148	10,995	−4%

which the adjustments are greater than 1 for the less severe injuries and less than 1 for the more severe injuries.

In Table 5, we illustrate the impact of the adjustments on the maximum potential PD awards that claimants are eligible for based on their assigned ratings. From the table, we see that the overall impact of the adjustments on PD benefits is small (a 3 or 4 percent reduction). However, this small average effect masks considerable heterogeneity in the affect on certain types of injuries. Psychiatric injuries receive the by far largest increases, equal to 141 and 124 percent with the linear and piecewise linear adjustments, respectively. Hand or wrist injuries receive the most substantial decreases in benefits, with 44 and 38 percent declines under the linear and piecewise linear adjustments, respectively. The effect on average ratings would also be larger if we used another category as the baseline, one in which ratings and proportional losses differed more on average.

RESULTS

Table 6 presents our results on the estimated predictive power of disability ratings. The table reports the results of six regressions. Columns I–III present results for regressions including the unadjusted disability rating, the linearly adjusted rating, and the piecewise linearly adjusted rating, respectively, as independent variables. The last three columns present results from analogous regressions, except they also include individual age and occupation groups to simulate the final rating used in California over this time period (which adjusted for these factors). Preinjury wages, year fixed effects, and fixed effects for region of the state are also included, although coefficients for these variables are not reported here. All models report "robust" standard error estimates, and the models with firm characteristics allow for clustering of standard errors at the firm level.

The key variables in the regressions are the variables indicating the type of impairment, the disability rating, and the impairment type interacted with the disability rating. Note that back injuries are the omitted category for the type of impairment. This means that the coefficients for the different impairment categories indicate the predicted differences in earnings losses between a given type of impairment and back impairments. Likewise, the coefficient for the disability rating represents the predicted relationship between disability ratings and earnings losses for back impairments, and the sum of this coefficient with the coefficients for the various interaction terms represents the relationship between disability ratings and earnings losses for the other impairments.

The coefficient on the disability rating variable is practically identical for all three models (there are differences if we examine results to more than three decimal places). This is largely because the coefficient for the rating reflects the relationship for back impairments, which are unchanged by the adjustments; the only reason there is any difference between the coefficients at all is because of colinearity between the independent variables. However, the coefficients for both the impairment category dummies and the interactions between these dummies and the disability rating change substantially, which is not surprising given the differences in the predicted values described in the text. Each rating type predicts earnings losses fairly well on average, with a 10 percentage point increase in the disability rating associated with approximately 8.4–8.6 percentage points of higher earning losses.

These correlations are much stronger than those reported by Park and Butler (2000), who found impairment ratings in Minnesota to have very little predictive power for earnings losses. We can think of two possible explanations for this disparity. First,

TABLE 6
Predicted Impact of Injury Categories on Earnings Losses With and Without Ratings Adjustments

Dependent Variable: Proportional Earnings Losses	I	II	III	IV	V	VI
Hip	5.601e-02	5.607e-02	1.745e-03	5.690e-02	5.694e-02	9.575e-03
	(3.630e-02)	(3.630e-02)	(4.269e-02)	(3.660e-02)	(3.660e-02)	(4.328e-02)
Knee	−0.037	−0.037	−0.026	−0.027	−0.027	−0.016
	(0.014)**	(0.014)**	(0.014)+	(0.015)+	(0.015)+	(0.014)
Ankle	−0.034	−0.034	−0.035	−0.023	−0.023	−0.024
	(0.017)*	(0.017)*	(0.017)*	(0.018)	(0.018)	(0.018)
Shoulder	0.037	0.033	−0.007	0.036	0.033	−0.005
	(0.014)**	(0.014)*	(0.016)	(0.015)*	(0.015)*	(0.017)
Elbow	−0.033	−0.033	−0.039	−0.025	−0.025	−0.029
	(0.023)	(0.023)	(0.023)+	(0.022)	(0.022)	(0.022)
Hand or wrist	−0.077	−0.077	−0.039	−0.072	−0.072	−0.035
	(0.031)*	(0.031)*	(0.025)	(0.032)*	(0.032)*	(0.026)
Psychiatric	0.300	0.264	0.184	0.319	0.282	0.200
	(0.031)**	(0.034)**	(0.054)**	(0.032)**	(0.035)**	(0.055)**
Hearing	0.073	0.071	0.030	0.084	0.082	0.040
	(0.032)*	(0.033)*	(0.056)	(0.033)*	(0.034)*	(0.057)
Other impairment	−0.012	−0.012	−0.033	0.001	0.001	−0.020
type	(0.034)	(0.034)	(0.045)	(0.035)	(0.035)	(0.046)
Disability rating	0.839	0.839	0.839	0.855	0.855	0.855
	(0.031)**	(0.031)**	(0.031)**	(0.033)**	(0.033)**	(0.033)**
Hip * rating	−0.477	−0.466	−0.234	−0.513	−0.503	−0.302
	(0.145)**	(0.149)**	(0.180)	(0.152)**	(0.156)**	(0.188)
Knee * rating	−0.036	0.115	0.025	−0.054	0.097	0.008
	(0.059)	(0.067)+	(0.062)	(0.061)	(0.070)	(0.064)
Ankle * rating	−0.071	0.075	0.075	−0.093	0.053	0.052
	(0.075)	(0.087)	(0.087)	(0.077)	(0.090)	(0.090)
Shoulder * rating	0.011	−0.308	−0.030	0.016	−0.320	−0.060
	(0.098)	(0.063)**	(0.087)	(0.096)	(0.064)**	(0.090)
Elbow * rating	0.050	0.129	0.247	0.050	0.131	0.218
	(0.149)	(0.162)	(0.172)	(0.151)	(0.164)	(0.173)
Hand or wrist	0.153	0.711	0.193	0.131	0.685	0.168
* rating	(0.143)	(0.224)**	(0.127)	(0.144)	(0.226)**	(0.126)
Psychiatric	−0.304	−0.454	−0.303	−0.311	−0.461	−0.308
* rating	(0.101)**	(0.064)**	(0.112)**	(0.107)**	(0.067)**	(0.116)**
Hearing * rating	−0.713	−0.756	−0.524	−0.716	−0.764	−0.525
	(0.123)**	(0.082)**	(0.214)*	(0.123)**	(0.083)**	(0.218)*
Other * rating	−0.307	−0.122	0.026	−0.333	−0.152	−0.000
	(0.150)*	(0.200)	(0.224)	(0.151)*	(0.201)	(0.226)

Continued

TABLE 6
Continued

Dependent Variable: Proportional Earnings Losses	I	II	III	IV	V	VI
Ratings adjustment	None	Linear	Piecewise linear	None	Linear	Piecewise linear
Age and occupation controls	No	No	No	Yes	Yes	Yes
Observations		82,556			78,769	

Notes: All regressions include fixed effects for year and region of the state as well as the earnings in the quarter prior to the date of injury. Columns IV–VI, which do not include firm fixed effects, also include the size, industry, and self-insurance status of the at-injury firm. Robust standard errors are presented in parentheses. The standard errors in Columns IV–VI also allow for clustering at the firm level.
**Statistical significance at the 1% level.
*Statistical significance at the 5% level.
+Statistical significance at the 10% level.

Park and Butler did not include the same level of detail with regards to the injured part of the body, which could bias the coefficient toward zero. The second explanation is that the disability ratings used in California, which rely on subjective evaluation of work restrictions by physicians, are better predictors of lost earnings than the impairment ratings used in Minnesota. We do note, however, that our results are consistent in the sense that the overall R^2 on our estimates are quite low (only about 0.01 in all cases). These results suggest that earnings losses appear dominated by unobserved individual (and perhaps employer) heterogeneity and disability ratings are noisy signals of losses.

Although the ratings seem to predict losses well on average, there are clear differences in both the *level* and *slope* of earnings losses for different ratings. The impairment fixed effects demonstrate that various disabilities are associated with higher or lower losses for a given rating, even controlling for preinjury earnings and employer heterogeneity. In addition, the interaction terms suggest that disability ratings are clearly more predictive for some types of impairments than others. The most striking results appear to be for hearing and psychiatric impairments. In both cases, impairments are associated with higher losses for a given rating and the predictive power of ratings appears to be much lower.

In Table 7, we summarize the effect of the adjustments on both the slope and level effects of different impairment categories. We do this by comparing the F-statistics for joint significance of the constant and slope interaction terms of the different impairment categories. These comparisons are valid because the degrees of freedom are the same in each model. We compare separate F-statistics for the models with and

Table 7
Comparing the Influence of the Injury Effects for the Adjusted and Unadjusted Ratings

	Unadjusted		Linear Adjustment		Piecewise Linear Adjustment	
	F-Statistic for Constant Injury Effects	*F*-Statistic for Injury Slope Effects	*F*-Statistic for Constant Injury Effects	*F*-Statistic for Injury Slope Effects	*F*-Statistic for Constant Injury Effects	*F*-Statistic for Injury Slope Effects
No age and occupation controls	16.59	6.16	12.05	18.82	2.68	2.40
With age and occupation controls	16.51	6.21	11.93	17.57	2.45	2.24

Note: All *F*-statistics are statistically significant at the 1 percent confidence level. The *F*-statistics associated with the regressions with no controls for age and occupation have 10 numerator and 82,505 denominator degrees of freedom. Those associated with the regressions with age and occupation controls have 10 numerator and 78,653 denominator degrees of freedom. The large decline in denominator degrees of freedom is attributable to the adjustment for firm-level clustering in the latter case.

without age and occupation controls. All *F*-tests based on the statistics reported in the table are significant at the 1 percent confidence level.

The *F*-statistics for the joint significance of the constant impairment effects are over 40 without adjustment, whereas the statistics for the slope effects are over 7. When we introduce the linear adjustment, we see an interesting pattern. The *F*-statistics for the constant effects drop to 30–35, whereas the *F*-statistics for the slope effects increase markedly. This suggests that the relationship between earnings losses and ratings across different impairment categories is sufficiently nonlinear that the adjustment can actually worsen the predictive power of the ratings. This is reinforced by the results for the piecewise linear adjustment, which drops the *F*-statistics for the constant effects down to about 7 or 8 while dropping those for the slope effects down to 4 or 5. The implications are clear; the piecewise linear adjustment does the best job of predicting the 3-year proportional earnings losses resulting from different types of disabilities.

Labor Market Responses

A key assumption underlying our analysis is that the positive relationship between observed earnings losses and disability ratings is driven entirely by more severe impairment for the higher rated claims. If higher disability ratings influenced behavior in such a way that it inflated earnings losses, it could confound our evaluation. One reason we might worry about this kind of blurred causality is because higher disability ratings lead to higher PD benefits. If the higher benefits induced people to reduce their labor supply then this could generate a positive relationship between losses and ratings like we see in Table 6, regardless of the true relationship between functional limitations, ratings, and loss of earnings capacity.

PD benefits act as an increase in unearned income, or wealth, to recipients because one can work and still receive PPD. In our data, the average, discounted PD benefits for an individual with a 15 percent disability rating (the average in our sample) were approximately $9,263 in 2003 dollars. For individuals with ratings of 25 percent, the average benefits were $18,636 (approximately a 102 percent increase). Now compare these to the implied increases in proportional losses from Table 6. From Table 3, we know that the average proportional losses in our sample are about 16 percent, so an increase from a rating of 15 to 25 percent would be associated with an increase of proportional losses to about 24.4 percent (a 52 percent increase).

We can use these numbers to impute the wealth elasticity of labor supply that would be required to explain the observed decline in labor force participation if it were driven *entirely* by higher workers' compensation indemnity benefits. The wealth elasticity, denoted η, can be defined as the percent drop in labor force participation due to a 1 percent increase in wealth; that is, we can define $\eta = (\%\Delta L)/(\%\Delta B)$. Using the numbers discussed earlier, we have implied wealth elasticities of labor supply of approximately -0.51. There is a wide literature on the impact of wealth on labor supply, but unfortunately the range of results of these studies tends to be too wide to say whether these implied elasticities are "too strong." They are fairly elastic relative to studies that focus on a randomized treatment design, such as the negative income tax (NIT) experiments or lotteries, in which the elasticities have tended to range from 0 to -0.25 (Imbens, Rubin, and Sacerdote, 2001). Obviously, this is a crude method for approximating the impact of disability benefits on labor supply and does not allow us to rule out the possibility that at least some of the relationship we observe is driven by benefits. Still, it suggests at least that not all of the relationship can be attributed to labor supply effects.[15]

Another possibility we have not considered is that wages might adjust to reflect the changes to the benefit structure. In theory, compensating wage differentials should reflect expectations of uncompensated earnings losses, so workers who appeared undercompensated prior to the adjustment may have been receiving a wage differential to offset this. Thus, an adjustment to increase compensation for these workers could leave them overcompensated (from the perspective of optimal insurance). One potential long-term solution to this problem would be updating of the adjustments to reflect changes in earnings losses (which would presumably lower compensation for any of the overcompensated groups).

IMPLICATIONS FOR THE EQUITY OF PERMANENT DISABILITY BENEFITS

In Table 5, we showed that adjusting disability ratings to reflect earnings losses has a significant effect on the PD benefits received by injured workers. If the earnings loss

[15] Reville and Seabury (2007) examine the role of PPD benefits in the labor supply decisions of disabled workers in California. They find that the labor supply responses are small, on average, but significantly larger for older workers and those with low preinjury wages. Thus, the relationship between ratings and losses is probably not driven by benefits on average, although it may be for certain subpopulations.

adjustments improve the ability of ratings to predict the loss of earnings capacity, then we would expect them to also improve the equity of disability benefits. This improvement would occur because workers with injuries to different body parts but with similar earnings losses would receive more similar benefit payments. In this section, we examine the extent to which this is the case.

We construct an *ad hoc* measure of the equity of disability benefits by taking the ratio of the maximum PD award for which a worker would be eligible to the predicted amount of earnings lost by injured workers with different impairments. In a perfectly equitable system, different impairments with the same losses should receive identical benefits. We predict the dollar value of earnings losses for different impairment categories for the average disability rating (15 percent) by multiplying the predicted 3-year proportional losses of each by the overall average discounted 3-year cumulative earnings of uninjured control workers ($119,592 in 2003 dollars). The predicted proportional losses come from the regressions of earnings losses on ratings reported in Table 6.[16]

Disability benefits are computed by applying the schedule to the observed rating (assuming workers have high enough weekly wages qualify for the benefit maximum, which we estimate is true for the majority of the sample, but we cannot say for sure because we do not have data on actual weekly wages), and then converting into a present value at the time of injury using an annual rate of 2.03 percent (the same as for losses). Those with ratings of 70 percent or higher are eligible for a small life pension, and we assume that the injured worker has a life expectancy of 80 and compute the total discounted present value based on the number of years between age at injury and 80 years.

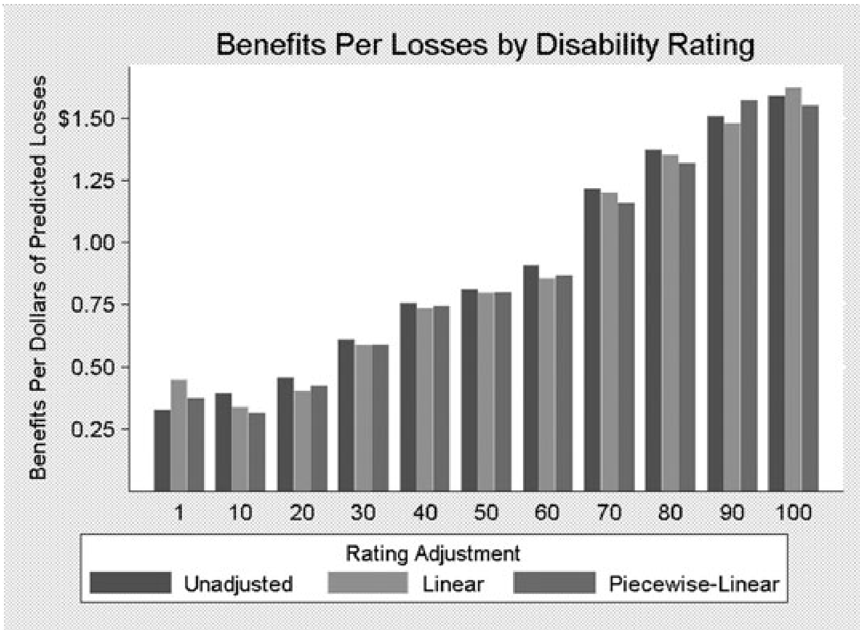
We complete our measure of equity by dividing the maximum PD award individuals are eligible for (given their ratings) by their predicted earnings losses. We do this separately for the unadjusted, linearly adjusted, and piecewise linearly adjusted ratings. It would be easy to try and interpret this measure as an estimated replacement rate of lost income, but this is inappropriate for two reasons. First, we do not have information on what benefits workers actually receive, since they may settle for an amount different than what they are eligible for based on this rating and because there are other benefits (such as temporary disability benefits) that are not accounted for. Second, we are comparing the present value of lifetime PD benefits against the present value of 3 years of earnings losses. To determine a replacement rate, we would need to compare the present value of lifetime (total) benefits against the present value of lifetime losses. As such, we can infer nothing about the adequacy of benefits by looking at the value of this ratio, and our discussion focuses on equity.[17]

---

[16] We only use the results from Columns I–III of Table 6, without the age and occupation adjustments, but the results are nearly identical if we use the coefficients reported in Columns IV–VI.

[17] Other potentially important sources of compensation for which we have no data are California or Social Security Disability Insurance (DI) benefits. If workers who were relatively undercompensated by PPD benefits were more likely to apply for these other benefits, we might find that the inequities in the system would be at least partially offset. However, we have no data to support or refute this possibility.

FIGURE 3
Benefits as a Function of Predicted Earnings Losses by Disability Rating

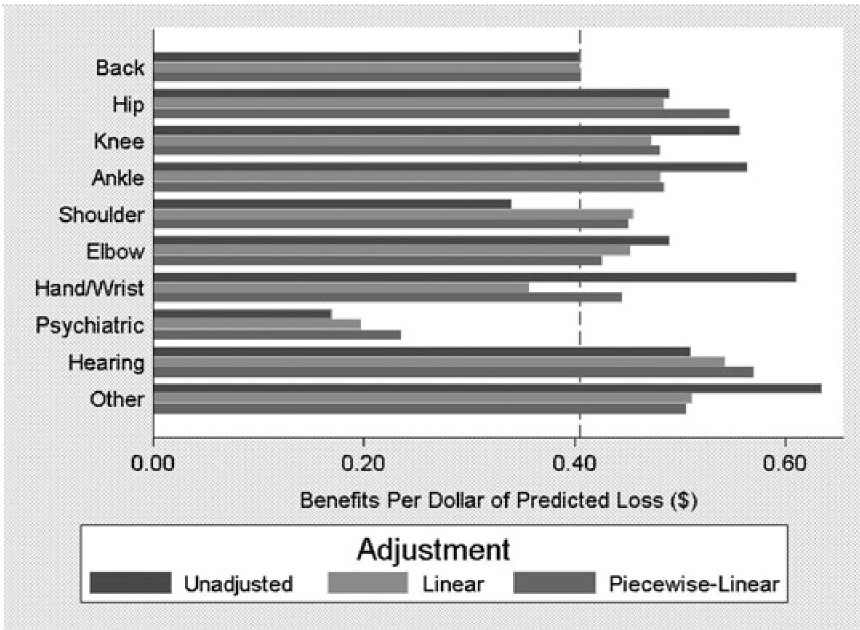


In Figure 3, we plot the maximum PD benefits per dollar of losses against the disability rating. Here we average by ratings in groups of 10 (similar to Figure 2, except here we break down the higher ratings into separate groups). From the figure we see that the average amount of benefits per dollar of losses is increasing in the disability rating. This is because benefits are a convex function of the disability rating, although we have modeled losses as linear.[18] Generally, the distribution of benefits per dollar of loss is very similar across all three ratings. This is not surprising because from Table 5 we saw that the average change in benefits across all types of impairments is close to zero. Since the type of impairment varies at each level of the rating reported in Figure 3, the impact of the adjustments on the overall level of benefits for a given rating is small.

Given that we know there are cross-impairment differences in earnings losses, a more important test of equity (for the purposes of this article) is how well the adjusted ratings equalize benefits per dollar of losses across impairment categories. Average ratings differ across impairments, however, and Figure 3 tells us that this influences benefits per losses. So, to obtain a purer test of equity we predict average losses by

[18] In other specifications we found that there does appear to be a slight concavity to the relationship between ratings and losses, which would mean that the relationship presented in Figure 3 would understate the benefits per dollar of loss at higher levels. However, the concavity of the estimated relationship was quite small, so we kept with the linear specification because it was more convenient when including the interaction terms between ratings and disabilities.

FIGURE 4
Permanent Disability Benefits per Dollars of Predicted Earnings Losses—15% Disability Rating



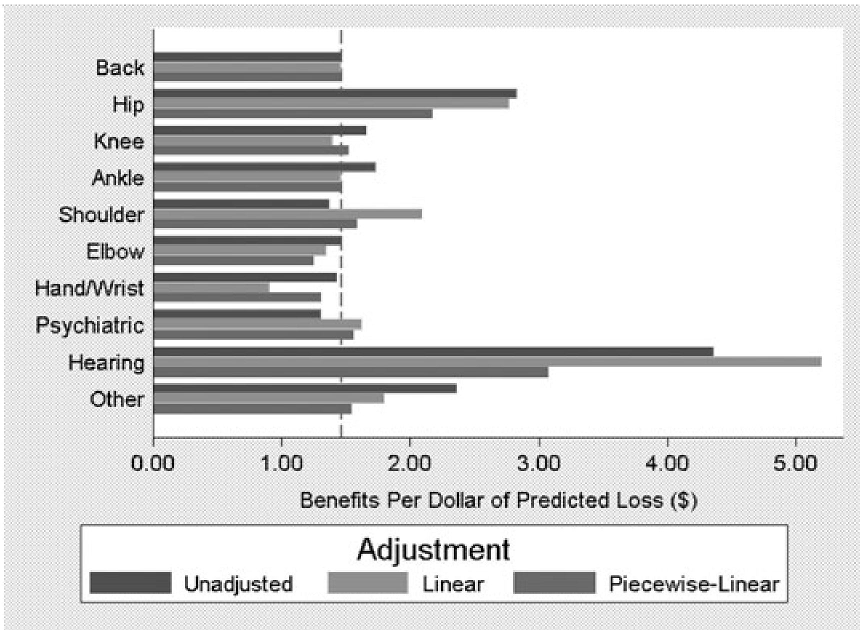
type of disability using the same regressions but holding all independent variables except the type of impairment constant. We construct this measure for ratings of 15 percent and 70 percent to see how the relationship changes between moderate and severe injuries.

In Figure 4, we present the results assuming a rating of 15 percent (suggesting benefits of $9,263, as we saw before). As noted earlier, in an equitable system, workers with different types of impairments with similar severity should be compensated in proportion to their losses. Thus, the adjustments should equalize the compensation of different impairment categories with that for the baseline category (back impairments).

The figure suggests that there are considerable deviations from equity with the unadjusted ratings. We estimate that workers with back impairments assigned a 15 percent rating are eligible for up to approximately $0.41 for every dollar of earnings loss. This is considerably less than workers with impairments to hearing, the hand or wrist, ankle, knee, or elbow, all which are eligible for at least $0.49 in benefits per losses. However, it is considerably more than workers with psychiatric impairments, who average just $0.17. Overall, the benefits per dollar of predicted loss vary from $0.17 (psychiatric) to $0.64 (other impairments).

Both adjustments substantially reduce the variation in compensation across impairments in most categories. Benefits per losses for impairments to the hand or writs fall

FIGURE 5
Permanent Disability Benefits per Dollars of Predicted Earnings Losses—70% Disability Rating



to approximately \$0.36 with the linear adjustment and \$0.45 with the piecewise linear adjustment. Likewise, the ratio of benefits to losses falls for impairments to the knees, ankles, and the miscellaneous category, becoming much closer to back impairments (although benefits for hip impairments actually increase with the piecewise linear adjustment). The relative award for psychiatric impairments is still lower than for back impairments with the piecewise linear adjustment, but the increase to \$0.24 makes up close to one-third of the difference. For most types of injuries, both adjustments appear to do a similar job at making benefits more equitable.

Figure 5 presents the same results as in Figure 4, only this time we assume a 70 percent disability rating. This leads to an estimated present value of disability benefits equal to about \$122,297. In this case, compensation for back impairments is about \$1.46 per dollar of losses on average. As before, there are several impairments that receive considerably higher levels. The most noticeable cases are impairments to the hip, which receive approximately \$2.86 per dollar of loss, and hearing impairments, which receive over \$4.00. The miscellaneous category also receives more than \$2.00. Although the extent of compensation below back injuries is somewhat smaller, psychiatric impairments still receive about 15 cents less compensation per dollar of losses, whereas shoulder impairments receive about 9 cents less.

As before, both adjustments appear to improve equity on average. With the piecewise linear adjustment, compensation for the three categories that most overcompensate relative to back impairments falls considerably. Compensation for hip impairments falls about \$0.75 per dollar of losses, hearing impairments by over \$1.00, and for impairments in the miscellaneous category it falls to just 9 cents more than for back impairments. There are also improvements for impairments to the knees and ankles, although compensation falls for elbow impairments. Interestingly, compensation for psychiatric and shoulder impairments is slightly higher than for backs with the adjustments.

One thing that stands out in Figure 5, and is different from Figure 4, is that the linear adjustment does worse than the piecewise linear adjustment. Compensation for hearing impairments actually rises, and the compensation for hip impairments barely declines. Also, compensation for shoulder impairments overshoots that for back impairments \$0.12 per dollar of loss with the linear adjustment.

Overall, the results of this section suggest that both the linear and piecewise linear adjustments can improve the equity of PD benefits. However, the linear adjustment appears to be an inferior policy tool for improving equity, which is not surprising given the results of Table 7. Thus, taken with the regression results of the previous section, our analysis appears to support the idea that the piecewise linear adjustment is the appropriate one to use for targeting compensation to match expected earnings losses across different types of impairments. However, we note that even with the adjustments, there is still considerable variation across impairment types in the estimated benefits per dollars of losses, suggesting that the adjustments here do not lead to a fully equitable compensation system.

CONCLUSION

The ability to accurately assess the severity of disability, and predict the impact of these disabilities on an individual's ability to earn in the labor market, is essential to guarantee that the victims of workplace injuries receive equitable compensation. In this article, we used matched administrative data on earnings and disability ratings to show that the California workers' compensation system suffered from a lack of empirical evidence on earnings losses to guide the disability ratings process. Our evidence shows that the equity of PD awards for workers with different impairments is significantly improved when adjustments are made for mean proportional earnings losses. Allowing the factors to differ over ranges of severity enhances the effectiveness of the adjustments. Future efforts could refine the adjustments even more, for example, with a more direct tie between medical evidence and lost earnings capacity.

We were primarily concerned with how a ratings adjustment of the kind we consider would affect the equity of PD benefits. It is still important, however, to consider who the winners and losers of such a system might be. For any adjustment that is approximately revenue neutral, which ours is, will have some losers because someone has to have benefits cut (in expectation). Those individuals with injury categories in Table 4 associated with adjustments less than 1 would all face lower expected benefits after an adjustment was made. This is not strictly necessary, in that adjustments could be tailored to raise everyone's benefit, but this would have other consequences such as raising (potentially by a great deal) expected workers' compensation claims and costs. Ultimately, although adjustments might make ratings more equitable, this does not mean they necessarily satisfy the conditions for a Pareto improvement, and policymakers should be aware of the potential trade-offs.

Our study is limited in that it only contains data from California. Of course, California is of particular interest because it is the first state to attempt to adjust its PD ratings system for earnings losses. But because the AMA *Guides* are more widely used, a study relating the earnings losses to PD claimants to impairment ratings using the AMA *Guides* would have greater direct applicability. However, the methodology we develop can be applied more generally to other rating systems, including the AMA *Guides*. Because the AMA *Guides*, like the prereform California system, do not incorporate information about proportional earnings losses, it is quite likely that use of the *Guides* results in inequitable compensation for workers with different types of injuries. On the other hand, if the *Guides* were able to predict impairment more reliably than the old California system, then adjusting the *Guide* ratings using new data on earnings loss could reduce some of the inequities that we observed in our data even after applying the adjustments. Although data do not currently exist to test for these inequities, in the Appendix we use some preliminary ratings data to show that there likely are inequities but that they may differ significant from those found here. Ultimately, adjustments to any system would need to be based on earnings losses from individuals in that system, to accurately reflect differences in institutional and labor market factors that could influence losses.

Extending our methodology to evaluate the AMA *Guides* would have implications beyond workers' compensation. The tort system is filled with cases in which evidence of disability is used to justify the need for compensation. The ability to predict the loss of earnings capacity from different types of impairments as evaluated in the AMA *Guides* would provide a disability ratings schedule that could eliminate the need for juries to make uninformed and abstract determinations about the potential lost earnings of plaintiffs.

Although we show in this article that adjusting ratings to reflect earnings losses could improve the equity of disability benefits, we do not ask whether or not it would be *efficient* to do so. Consider the case of psychiatric impairments. These are considered to be disabilities in which the actual extent of functional limitation is very difficult to verify, and hence are constantly perceived to have a high potential for fraud. If this is true, then increasing the compensation per dollar of earnings loss for psychiatric impairments could provide workers with the incentives to fake psychiatric impairments to obtain disability benefits. Future work should explore the extent to which these moral hazard effects could reduce, or even eliminate, the benefits of adjusting disability ratings to reflect differences in earnings losses for different types of impairments.

APPENDIX: ADJUSTING RATINGS IN THE POST-2005 CALIFORNIA SYSTEM

Effective January 1, 2005, California SB 899 instituted a new PD rating system based on the AMA *Guides*. This raises an obvious question as to how the results of this article would change if we were to conduct a similar analysis using data on workers' compensation claims rated under the new system. Although we do not have data on earnings losses for newly rated claims, we were able to obtain information on average ratings by disability category from the new system. Using these average ratings, we can calculate linear adjustments under the assumption that earnings losses for each category are the same in the post-2005 system as in the pre-2005 system.

TABLE A1
Estimating Adjustments for the Post-2005 California Disability Rating System

	Observations in New System	Ratings		Linear Ratings Adjustments	
		Pre-2005 System	Post-2005 System	Pre-2005 System	Post-2005 System
Back	1,577	0.199	0.102	1.000	1.000
Hip	22	0.222	0.086	0.971	1.284
Knee	818	0.148	0.051	0.842	1.238
Ankle	144	0.142	0.045	0.840	1.357
Shoulder	678	0.099	0.051	1.690	1.689
Elbow	78	0.094	0.038	0.917	1.168
Hand or wrist	707	0.109	0.050	0.641	0.722
Psychiatric	60	0.212	0.191	2.005	1.138
Hearing	58	0.107	0.056	1.723	1.694

The average ratings and adjustments using the post-2005 system are compared to the pre-2005 system in Table A1. The sample sizes for some of the injury categories are small, although we have at least 50 ratings for all but hip impairments. One fact that immediately stands out from the table is that ratings under the AMA *Guides* are considerably lower in every category than in the pre-2005 system.[19] In general, the adjustments also vary considerably across the two systems. For three of the impairment categories, shoulders, hands/wrists, and hearing, the adjustments are approximately the same. Impairments in the hips, knees/ankles, and elbows were overcompensated relative to backs in the pre-2005 system (adjustments less than 1), but are undercompensated relative to backs in the new system (adjustments greater than 1). Interestingly, psychiatric claims received the largest adjustment in the pre-2005 system but receive the smallest adjustment in the post-2005 system.

Table A1 suggests clearly that adjustments based on ratings from the old system are of limited use with a new ratings system. But it confirms the general point that the use of earnings losses could improve the equity of ratings in other systems than just the one we consider in this article. If the adjustment factors under the post-2005 system were essentially 1, this would suggest that the between-impairment variation we found might simply have been an artifact of idiosyncrasies in the pre-2004 California rating system. However, although the exact values of the adjustments differ, the magnitude suggests that there are similar inequities in the post-2005 system as there were in the pre-2005 system. This suggests that other states and other systems could benefit from empirically based adjustments.

An important caveat to this is that we are making a critical assumption that earnings losses for the different impairment categories are the same in the post-2005 period as they were in the pre-2005 period. This assumption may fail, particularly since the AMA *Guides* are likely more restrictive in the nature of conditions that can be rated

[19] The ratings for psychiatric claims are actually not done using the AMA *Guides* but rather are rated under a separate system that is still different from the pre-2005 ratings.

as a disability (e.g., it is our understanding that the AMA *Guides* do not allow an impairment rating to be granted in the case of chronic pain with no organic signs of injury). However, even if the average earnings loss in each category did differ for ratings in the post-2005 system, the differences would have to take a very specific pattern to systematically drive the adjustment factors toward 1.

REFERENCES

Barth, P., and M. Niss, 1999, *Permanent Partial Disability Benefits: Interstate Differences*, WC-99-2 (Cambridge, MA: Workers Compensation Research Institute).

Berkowitz, M., and J. Burton, 1987, *Permanent Disability Benefits in Workers' Compensation* (Kalamazoo, MI: W.E. Upjohn Institute for Employment Research).

Blum, F., and J. Burton, 2003, Workers' Compensation Benefits: Frequencies and Amounts 1995-1999, *Workers' Compensation Policy Review*, 3(6): 2-32.

Boden, L. I., 1992a, *The AMA Guides in Maryland: An Assessment* (Cambridge, MA: Workers Compensation Research Institute).

Boden, L. I., 1992b, Dispute Resolution in Workers' Compensation, *Review of Economics and Statistics*, 74(3): 493-502.

Boden, L., and M. Galizzi, 2002, Income Losses of Men and Women Injured at Work, *Journal of Human Resources*, 38(3): 722-757.

Burton, J. F., Jr., 2005, Permanent Partial Disability Benefits, in: K. Roberts, J. F. Burton, Jr., and M. Bodah, eds., *Workplace Injuries and Diseases: Prevention and Compensation. Essays in Honor of Terry Thomason* (Kalamazoo, MI: W.E. Upjohn Institute for Employment Research).

Butler, R., 1994, Economic Determinants of Workers' Compensation Trends, *Journal of Risk and Insurance*, 61(3): 383-401.

Cheit, E. F., 1961, *Injury and Recovery in the Course of Employment* (New York: John Wiley & Sons).

Cocchiarella, L., M. Turk, and G. Andersson, 2000, Improving the Evaluation of Permanent Impairment, *JAMA*, 283: 532-533.

Dehejia, R. H., and S. Wahba, 1999, Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs, *Journal of the American Statistical Association*, 94: 1053-1062.

Greene, W. H., 2002, *Econometric Analysis*, 5th edition (Upper Saddle River, NJ: Prentice Hall).

Department of Industrial Relations, 1990, *Schedule for Rating Permanent Disabilities*, California Division of Workers' Compensation. World Wide Web: http://www.dir.ca.gov/dwc/PDR1997.pdf.

Heckman, J. J., and V. J. Hotz, 1989, Choosing Among Alternative Non-Experimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training, *Journal of the American Statistical Association*, 84(408): 862-874.

Imbens, G., D. Rubin, and B. Sacerdote, 2001, Estimating the Effect of Unearned Income on Labor Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players, *American Economic Review*, 91(4): 778-794.

Krueger, A., 1990, Incentive Effects of Workers' Compensation Insurance, *Journal of Public Economics*, 41(1): 73-99.

Lalonde, R., 1986, Evaluating the Econometric Evaluations of Training Programs, *American Economic Review*, 76: 604-620.

Little, R. J., and D. B. Rubin, 1987, *Statistical Analysis With Missing Data* (New York: John Wiley).

Park, Y., and R. Butler, 2000, Permanent Partial Disability Awards and Wage Loss, *Journal of Risk and Insurance*, 67(3): 331-349.

Peterson, M., R. Reville, and R. K. Stern, 1997, *Compensating Permanent Workplace Injuries: A Study of California's System*, MR-920-ICJ (Santa Monica, CA: RAND).

Reville, R., 1999, The Impact of a Disabling Workplace Injury on Earnings and Labor Force Participation, in: J. Haltiwanger, and J. Lane, eds., *Contributions to Economic Analysis* (Amsterdam: Elsevier Science, North-Holland), pp. 147-173.

Reville, R., F. Neuhauser, J. Bhattacharya, and C. Martin, 2002, Comparing Severity of Impairment for Different Permanent Upper Extremity Musculoskeletal Injuries, *Journal of Occupational Rehabilitation*, 12(3): 205-221.

Reville, R., and R. Schoeni, 2001, *Disability From Injuries at Work: The Effects on Earnings and Employment*, DRU-2554 (Santa Monica, CA: RAND).

Reville, R., and S. Seabury, 2007, *Estimates of the Wealth Elasticity of Labor Supply for the Disabled*, WR-466-ICJ (Santa Monica, CA: RAND).

Reville, R. T., S. A. Seabury, F. Neuhauser, J. F. Burton, Jr., and M. Greenberg, 2005, *An Evaluation of California's Permanent Disability Rating System*, MC-258-ICJ (Santa Monica, CA: RAND).

Seabury, S. A., R. T. Reville, and F. Neuhauser, 2006, Physician Shopping in Workers' Compensation: Evidence from California, *Journal of Empirical Legal Studies*, 3(1): 47-77.

Spieler, E. A., P. S. Barth, J. F. Burton, Jr., J. Himmelstein, and L. Rudolph, 2000, Recommendations to Guide Revision of the *Guides to the Evaluation of Permanent Impairment*, *JAMA*, 283: 519-523.

Walter, Y., and T. L. Idson, 1999, Firm Size and Wages, in: O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Volume 3a (Amsterdam, Netherlands: Elsevier), pp. 2165-2214.