# Reserving in the Age of Obesity

Chris Laws and Frank Schmid

**Abstract**
**Motivation.** There is increasing evidence that obesity contributes to the cost of medical care in workers compensation, and that this contribution is significant in magnitude. For instance, a recent study of workers compensation claims of Duke University employees shows that, for the morbidly obese, the medical costs per 100 full-time equivalent employees are nearly seven times as high as for employees of recommended weight. In the following study, the evidence of the contribution of obesity to the medical costs of workers compensation is generalized to a set of claims that comprises 36 U.S. States and nine injury years. Further, it is shown how the cost difference between "obese claims" and comparable "non-obese claims" develops as claims mature—this evidence of the difference in development offers important guidance for both reserving and ratemaking. The study is confined to the effect of obesity on severity—the effect of obesity on claim frequency is beyond the scope of this analysis.
**Method.** The study makes use of a matched-pairs research framework. Every obese claim in the data set is matched with a non-obese claim. Exact matching applies to all claim characteristics, except age at injury, where proximity matching is employed. The set of matched pairs is then analyzed using a semiparametric Bayesian multilevel model, the nonparametric component of which accounts for the possible nonlinear influence of age. Aside from age, the covariates comprise the injury year, the nature of injury, the U.S. state, and the industry—these four covariates enter the model as random effects. Further, the gender of the claimant and cross-state differences in the legislative environment, as they manifest themselves in mandatory utilization review and mandatory bill review, are accounted for using indicator variables. The model is estimated by means of MCMC (Markov Chain Monte Carlo simulation). The reversible jump concept of Bayesian modeling averaging is used in determining the functional form of the nonparametric component that captures the influence of age.
**Results.** The study shows that, in the aggregate, obese claims are 2.8 times more expensive than non-obese claims at the 12-month maturity, but this cost difference climbs to a factor of 4.5 at the three-year maturity and to 5.3 at the five-year maturity. Further, the cost difference (at the five-year maturity) is less for females than for males. Mandatory utilization review and, in particular, mandatory bill review significantly reduce the cost difference between obese and non-obese claims.
**Availability**. The semiparametric multilevel model was estimated using JAGS with R. JAGS (Just Another Gibbs Sampler, http://www-ice.iarc.fr/~martyn/software/jags/) is an open-source platform for Gibbs sampling, developed by Martyn Plummer at the International Agency for Research on Cancer of the World Health Organization in Lyon, France. The reversible jump routine was written as a C++ JAGS module. R is an open-source statistical modeling platform (http://www.r-project.org/), which is administered by the Technical University of Vienna.

**Keywords**. Obesity, Multilevel Model, Partial Linear Model, Reversible Jump MCMC, Semiparametric Model, Workers Compensation

## 1. INTRODUCTION

In July 2009, the Centers for Disease Control and Prevention (CDC) held its inaugural conference "Weight of the Nation," thus calling to public attention the mounting problem that obesity poses for the health of the American people. A research paper presented at this conference by Finkelstein et al. [4] called the "link between rising rates of obesity and rising medical spending" nothing short of "undeniable." These authors estimate that obesity accounts for "almost 10 percent of all medical spending and could amount to $147 billion per year in 2008." This is about twice the cost estimate of $78.5 billion that Finkelstein, Fiebelkorn, and Wang [3] had established for the year 1998 in an earlier study.

Obesity may also have significant implications for workers compensation. In a recent study of Duke University employees, Truls, Dement, and Krause [9] found that employees in the highest obesity class, when compared with employees of recommended weight on an FTE (full-time equivalent) basis, filed twice as many claims, had 13 times as many lost workdays, and experienced medical and indemnity costs that were 7 and 11 times as high, respectively.

The Duke University study aimed at estimating the differences in medical and indemnity costs between obese and non-obese employees. To this end, the cumulative payments for every claim were tallied at the end of the study, which is December 31, 2004; for open claims, which amount to 2.8 percent of the total number of claims, estimated reserves were used that had been provided by the competent workers compensation actuary—these details have been confirmed in writing by the corresponding author of the Duke University study, Dr. Truls Østbye. Thus, for the purpose of arriving at total costs, the Duke University study does not rest entirely on observed payments but (for open claims) also includes reserve estimates.

In what follows, we take a different approach than the one pursued by Truls, Dement, and Krause [9]. For one, the nature of our data set is quite different, as will be discussed. But most importantly, it is not our objective to provide a measurement for the difference in ultimate costs between obese and non-obese claimants—such a measurement would have to make use of reserve estimates for open (and potentially re-opening) claims. Instead, we try to shed light on the difference in development between claims from obese and non-obese employees; this way, we are able to provide guidance (on a per claim basis) on the divergence in cumulative payments between these two types of claims for reserving (and ratemaking) purposes.

Because of potential dissimilarities in development between claims of obese and non-obese employees, the difference in the costs per claim may not be apparent early on (e.g., at the 12-month maturity) but may take time to reveal itself. As is shown in this study, the ratio in the medical costs per claim of obese to non-obese claimants indeed develops; whereas this ratio stands at 2.8 at the 12-month maturity, it climbs to 4.5 at the 36-month maturity, and to 5.3 at the 60-month maturity. A possible reason for such dissimilarity in development may be the longer duration of obese claims; another reason may be that the distribution of medical costs, as claims develop, differs between obese and non-obese claims as obesity may raise the likelihood of medical complications. Due to data limitations, no definite statement can be made as to why the ratio of medical costs between obese and non-obese claims increases over time. Future research on differences in claim duration between obese and non-obese claims may provide an answer to this question.

## 1.1 Research Context

The Duke University study by Truls, Dement, and Krause [9] is to date the only comprehensive statistical analysis of the effect of obesity on the cost of workers compensation. This study makes use of a longitudinal data set, which was obtained by monitoring a cohort of 11,728 employees of Duke University and the Duke University Health System from January 1, 1997, through December 31, 2004. The cohort was defined by all employees that had at least one HRA (health risk assessment) during this time period; taking an HRA is voluntary and available to employees eligible for health care benefits. (Note that the number of members in the study may have shrunk over time due to employee termination or disability.) The members of this cohort were assigned to body mass index (BMI) categories based on the first HRA they participated in during the time of the study.

At the end of the eight-year time window, the number of claims, the number of work days, and the indemnity and medical costs were tallied for each employee; then, this information was matched up with the BMI category (and other characteristics) of the claimant. There are six BMI categories, ranging from underweight to recommended weight, overweight, and three classes of obesity. The highest level of obesity is class III, which comprises the morbidly obese, identified by a BMI of 40 or higher. The Duke University study finds that for the morbidly obese employees, the medical costs are 6.8 times as high as for employees of recommended weight; at the same time, an employee in this group is twice as likely to have a claim. For obese classes II (BMI of at least 35 but less than 40) and I (BMI of at least 30 but less than 35), the medical costs per employee are (respectively) 3.1 and 2.6 times as high as for employees of recommended weight; the respective multiples for the number of claims read 1.9 and 1.5. (The numbers cited above rest on the bivariate analysis presented in Truls, Dement, and Krause [9], Table 3.)

Another way of presenting the findings of the Duke University study is on a per claim basis. Transforming the medical costs per 100 FTE employees into costs per claim shows that this amount is 3.4 (obesity class III), 1.7 (obesity class II), and 1.7 (obesity class I) times the magnitude recorded for employees of normal weight. The corresponding numbers for indemnity read 5.5 (obesity class III), 3.4 (obesity class II), and 2.9 (obesity class I). As is apparent in the data, on a per claim basis, the percentage difference between obese employees and employees of normal weight is even higher for indemnity payments than it is for medical costs. (Here, too, the numbers rest on the bivariate analysis presented in Table 3 of Truls, Dement, and Krause [9].)

## 1.2 Objective

The effect of obesity on the medical cost per workers compensation claim is quantified at different maturities, thus showing how the (percentage) cost difference between obese and non-obese claims develops as claims mature. Further, it is shown how the dissimilarity in development between obese and non-obese claims varies with the legislative environment, as such manifests itself in mandatory utilization review (MUR) and mandatory bill review (MBR). Quantifying differences in development patterns across obese and non-obese claims is important for reserving and ratemaking in workers compensation.

## 1.3 Outline

What follows is an account of how the data set was prepared, followed by descriptive statistics. Section 3 then offers a discussion of the Bayesian semiparametric multilevel model, which is followed in Section 4 by a presentation of the findings for the random effects (injury year, U.S. state, industry, and nature of injury), gender, and age. Section 5 presents estimates of the effects of the legislative environment (MUR and MBR). Section 6 concludes.

# 2. THE DATA

We use a large sample of workers compensation claims provided by select insurance companies. The data base comprises records from 36 states (AK, AL, AR, AZ, CO, CT, DC, FL, GA, HI, IA, ID, IL, IN, KS, KY, LA, MD, ME, MO, MS, MT, NC, NE, NH, NM, NV, OK, OR, RI, SC, SD, TN, UT, VA, and VT) and nine injury years (1998-2006). Claims are studied at three different maturities: 12 months, 36 months, and 60 months. Clearly, not all maturities are available for all injury years. For instance, observations at the 36-month and 60-month maturities are available only up to injury years 2004 and 2002, respectively.

For each claim, cumulative medical payments are tallied at each of the chosen maturity dates. The observed cumulative payment at a given maturity is flagged as pertaining to an obese claimant if, at that maturity, there has been a record of a co-morbidity indicator pointing to obesity. Specifically, an observation is flagged as "obese" if the three leading digits of an ICD-9 code serving as a co-morbidity indicator equal 278. We refer to such an observation as an "obese claim."

It is worth noting that a claim that qualifies as obese at the 36-month maturity may not qualify as obese at the 12-month maturity if the co-morbidity indicator 278 does not show up before the 12-month maturity (but shows up no later than the 36-month maturity). On the other hand, such a claim does not qualify as non-obese at the 60-month maturity, even if the co-morbidity indicator 278 fails to show up again after the 36-month maturity. In summary, a claimant is treated as obese only after actually having been diagnosed as such. There may be instances where a claimant turns obese after the injury; and there may be instances where a claimant is diagnosed as obese only after the claimant's being obese has given rise to medical complications.

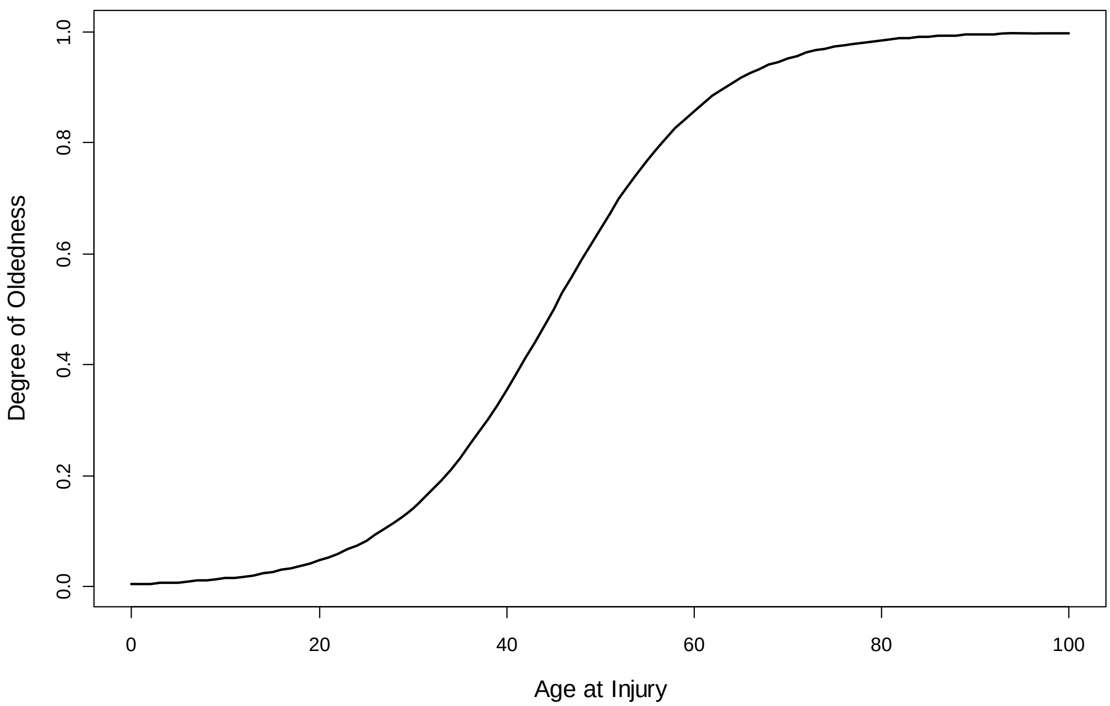
For the purpose of proper cost comparison between obese and non-obese observations, claims are dropped once a lump sum payment has been observed among the transactions. For instance, a claim may be included at the 12- and 36-month maturities, but excluded at the 60-month maturity if a lump sum payment was observed later than 36 months (but no later than 60 months) into the duration of the claim. The reason for excluding claims following a lump sum payment is that such transactions represent the present value of a stream of payments that may extend past the maturity date of interest.

The data base offers no information on the BMI, which is a standard measure of obesity (see Truls, Dement, and Krause [9]). As a result, this study does not differentiate between degrees of obesity. At the same time, it can be assumed that the co-morbidity indication identifies the claimant as severely obese, which puts him into one of the higher obesity classes. The small proportion of obese observations in the total number of claims supports this conjecture; for instance, at the 12-month maturity, the proportion of obese observations ranges between 0.1 percent (injury years 1998–2003) and 0.2 percent (injury years 2004–2006). By comparison, the proportion of morbidly obese claimants in the mentioned study of Duke University employees amounts to 4.9 percent.

The small percentage of obese claims in the total number of claims, along with the very high total number of claims (3,834,891 claims for the 12-month maturity; 2,956,285 claims for 36 months; and 2,079,225 claims for 60 months) is most suited for a matched-pairs research framework. In such an analytical setting, each obese observation in the data set is matched with a non-obese claim of the same maturity based on injury year, U.S. state, industry, ICD-9 code, gender, and age at injury. Except for age at injury, all matching criteria are categorical in nature, which allows for exact matching (thus obviating the need for propensity matching). At the same time, extending the concept of exact matching to age at injury (which is measured on a scale of whole numbers) leaves many obese claims without matches. Thus, for age at injury, we use proximity matching.

Proximity matching rests on the concept of the nearest neighbor. When exact matching for age is not feasible (for lack of exact matches when using whole numbers for years of age), researchers often resort to matching by age brackets. In matching by age bracket, an obese claim is paired up with a non-obese claim that belongs to the same (for instance) five-year age bracket (subject to being identical in all exact-matching characteristics). A disadvantage of matching by age bracket is that many claims are not matched with the closest neighbor. For instance, a 20-year-old obese claimant may be matched (within the 20-24 age bracket) with a 24-year-old non-obese claimant, but is prevented from being matched with an otherwise identical 19-year-old claimant. Proximity matching avoids this problem by looking for the nearest neighbor. At the same time, it may not be appropriate to have the concept of the nearest neighbor rest on the simple age difference. Because aging is a nonlinear process, it may be preferable to match an obese 25-year-old claimant with a 35-year-old, instead of a 16-year-old. Similarly, matching an obese 55-year-old with a non-obese 35-year-old may be more appropriate than matching this person with a non-obese 74-year-old. For this reason, we use a sigmoid function to create a fuzzy set for old age; the sigmoid function has the form $1/(1+\exp(-\sigma \cdot (h-45)))$, where $h$ is the age at injury and $\sigma$ was chosen to be equal to 0.12. (For the concept of fuzzy sets see, for instance, Kasabov [7].) Chart 1 shows the fuzzy set for old age (and its complement, young age); the degree of oldness is set to 50 percent at age 45 and then, in an "S-shaped" manner, this degree of oldness approaches zero and 100 percent as the years of age approach zero and 100, respectively. The nearest neighbor of an obese claim to an otherwise identical non-obese claim is defined by the smallest difference (in absolute value terms) in the degree of oldness.

**Chart 1:** Degree of Oldness, Defined by Sigmoid Function




To summarize, we define as neighbors to a given obese claim the set of non-obese claims that match exactly based on maturity, injury year, ICD-9 code, U.S. state, industry, and gender. In this matching process, a given non-obese claim may be used as a neighbor to more than one obese claim. Among the thus identified set of neighbors, the nearest neighbor is chosen based on the degree of oldness. This nearest neighbor may not be unique because of ties in the oldness distance, in which case we are left with a set of "tying neighbors." The highest number of tying neighbors (across all maturities) for a given obese claim equals 47; the percentage of obese claims with a single nearest neighbor equals 59.5.

Table 1 displays by maturity and injury year the number of obese claims and the total number of their (potentially tying) nearest neighbors within these categories. In the statistical analysis, for each obese observation, only one of the tying nearest neighbors is chosen, at random. We create multiple (50) data sets, each with random tie-breaking where obese observations have non-unique nearest neighbors. By analyzing these multiple data sets, we are able to provide credible intervals around the parameter estimates that reflect the uncertainty originating in the randomness of the tie-breaking process. (For details on credible intervals, see Carlin and Lewis **Error! Reference source not found.**.)

In Table 1, there are two reasons why the number of obese claims (and hence the number of non-obese neighbors) varies by maturity. First, claims are dropped following lump sum payments, as mentioned. Second, claims are added upon obesity showing up as a co-morbidity indicator in a transaction.

**Table 1:** Numbers of Obese Claims and (Potentially Tying) Non-Obese Nearest Neighbors

| Injury Year | Obese | Non-Obese | Obese | Non-Obese | Obese | Non-Obese |
|---|---|---|---|---|---|---|
| | Maturity | | | | | |
| | 12 months | | 36 months | | 60 months | |
| 1998 | 250 | 537 | 266 | 518 | 271 | 529 |
| 1999 | 282 | 565 | 313 | 573 | 304 | 545 |
| 2000 | 343 | 802 | 365 | 831 | 364 | 831 |
| 2001 | 417 | 1,058 | 430 | 1,018 | 413 | 953 |
| 2002 | 467 | 1,108 | 481 | 1,106 | 459 | 1,049 |
| 2003 | 553 | 1,290 | 517 | 1,157 | — | — |
| 2004 | 630 | 1,464 | 656 | 1,414 | — | — |
| 2005 | 722 | 1,514 | — | — | — | — |
| 2006 | 836 | 1,744 | — | — | — | — |

The concepts of matching by age bracket and matching by proximity of age share the characteristic of matching toward the center of the age distribution. These matching techniques, unlike exact matching (which, for instance, pairs up a 55-year-old obese claimant with a 55-year-old non-obese claimant when using whole numbers), tend to pair up old claimants with claimants of lesser age, and young claimants with claimants of higher age; this is because the claimants on the edges of the age distribution turns are sparse. For instance, within the age bracket 60-64, the number of observations tends to decline with age. Thus, a claimant at the center of this age bracket has a comparatively high chance of being "matched down the age distribution," as opposed to being "matched up;" this is simply because the number of potential matches within this bracket is higher at lower ages than at higher ages. Conversely, young claimants tend to be "matched up the age distribution." Note that this matching toward the center (of the age distribution) is not unique to proximity matching but is also characteristic of the traditional approach of matching by age bracket.

Matching toward the center of the age distribution poses no difficulty for the statistical analysis (because age is included as a covariate in a nonparametric setting, which accommodates potential nonlinearities), but it does affect the interpretation of the regression results for age (and age alone). For instance, if the medical costs of workers compensation claims increase with age, then the influence of age on the cost of obesity is underestimated for young claimants and overestimated for old claimants. As a consequence, the estimated effect of age may be distorted on the edges of the age distribution, taking on an "S-shaped" form. For this reason, the nonparametric regression finding for age has to be interpreted with care.

Exact matching by maturity, injury year, U.S. state, ICD-9 code, and gender is straightforward. Matching by industry relies on an NCCI industry classification; the five industries comprise Manufacturing, Contracting, Office and Clerical, Goods and Services, and Miscellaneous.

There are a total of 1,560 obese observations (or 14.3 percent of the total number of obese observations) going unmatched, because no match is available by maturity, injury year, ICD-9 code, U.S. state, industry, and gender.

Chart 2 offers a distribution of age at injury of the set of studied (obese and non-obese) claims for all three maturities; the observations are pooled over the nine injury years.

**Chart 2:** Relative Frequency of Claims by Age at Injury and Maturity

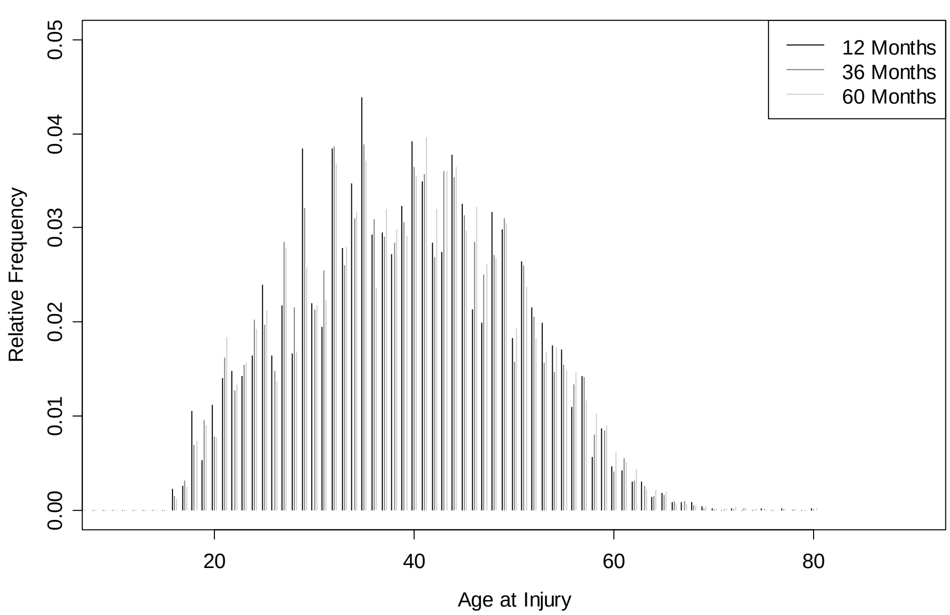


Chart 3 presents for all studied claims at the 60-month maturity a breakdown of gender by age (thus comprising only injury years 1998-2002); the majority of claimants are male, as expected. Further, Chart 4 details for all three maturities the relative claim frequency by U.S. state. Florida is the most highly represented state, and Rhode Island is the least highly represented. The representation of a state in the data set depends primarily on the size of its labor force, but also on the combined market share of the insurance companies that contribute to the mentioned data base; another contributing factor is the share of the self-insured.

**Chart 3:** Relative Frequency of Claims at 60-Month Maturity by Gender and Age at Injury

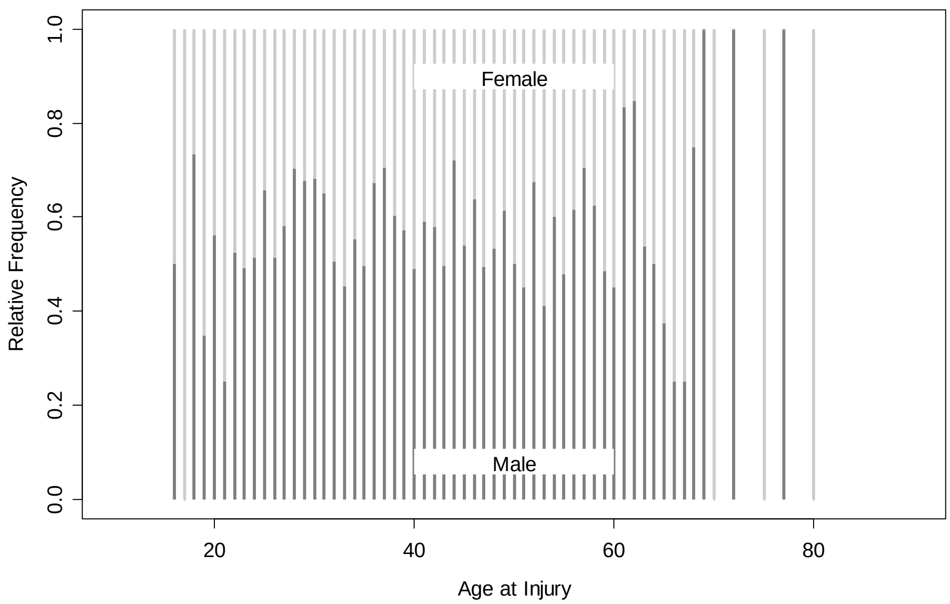
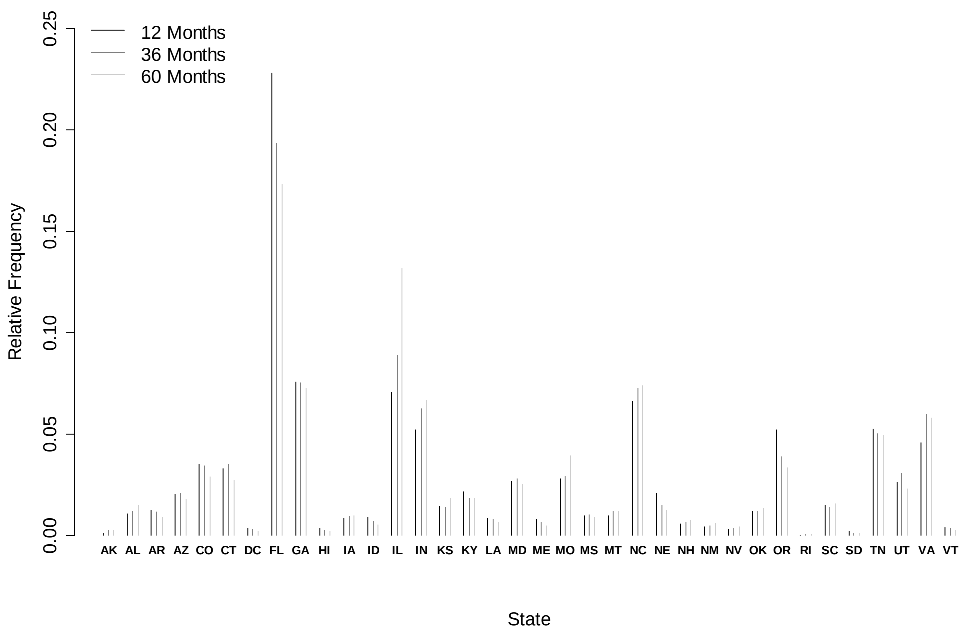
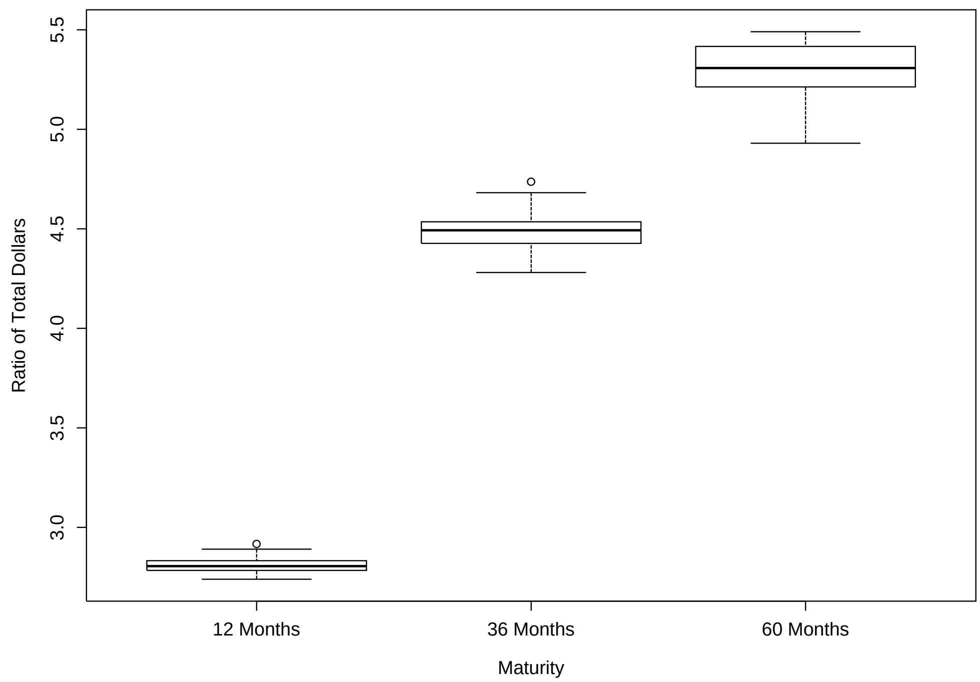


**Chart 4:** Relative Frequency of Claims by State and Maturity




Finally, Chart 5 offers an estimate of the ratio of the total cost of all obese claims, divided by the total cost of the non-obese claims these obese observations have been paired up with. The chart hosts three box plots, one for each maturity. Each boxplot represents 50 ratios, as generated by 50 randomized data sets. (As discussed, using multiple randomized data sets accounts for the uncertainty that originates in the tie-breaking of equidistant non-obese nearest neighbors when paired up with a given obese claim.)

**Chart 5:** Box Plots for the Ratios of Total Obese to Total Non-Obese Cumulative Payments for 50 Sets of Matched Claims at Alternative Maturities




In Chart 5, the horizontal line within a given box indicates the median value. The "hinges" of a box represent (approximately) quartiles, meaning that the box comprises about 50 percent of the observations. The distance between the hinges (that is, the height of the box) is called the inner quartile range. The whiskers located above and below the box point to extreme values. Specifically, the whisker above the box represents the highest observed value outside the box that is less than the sum of the third quartile plus 1.5 times the inner quartile range. Correspondingly, the whisker below the box signifies the lowest observed value outside the box that is greater than the first quartile minus 1.5 times the inner quartile range. Data points more extreme than those indicated by the whiskers are plotted as circles and may be considered outliers. For details on boxplots see Chambers et al. [1].

Chart 5 shows that obese claims are 2.8 times more expensive than non-obese claims at the 12-month maturity, but this cost difference climbs to a factor of 4.5 at the three-year maturity, and to 5.3 at the five-year maturity. This divergence in development between obese and non-obese claims has profound implications for reserving, as the added cost of obesity reveals itself only over time. A possible reason for such dissimilarity in development may be the longer duration of obese claims, although, at this point, due to data limitations, this cannot be confirmed with confidence. Clearly, the Duke University study points to longer durations for obese claimants.

## 3. THE STATISTICAL MODEL

The purpose of the statistical modeling is to quantify the effect of claim characteristics on the percentage cost difference between obese claims and comparable non-obese claims. Thus, the dependent variable in the statistical analysis is the natural logarithm of the ratio (or log ratio, for short) of the cost of an obese claim to the cost of a comparable non-obese claim. Comparable non-obese claims are identified by means of pair-wise matching, as detailed above.

The statistical model has a semiparametric (or, synonymously, partially linear) structure. Generically, a semiparametric model may be written as

$$y_i = \boldsymbol{x}_i \cdot \boldsymbol{\beta} + f(z_i) \ , \tag{1}$$

where $y_i$ is the dependent variable, $\boldsymbol{x}_i \cdot \boldsymbol{\beta}$ is the parametric, standard linear regression component, and $f(z_i)$ is a smoother that constitutes the nonparametric component. The purpose of the semiparametric structure is to accommodate a potentially nonlinear influence of the covariate $z$.

The semiparametric model makes use of a Bayesian multilevel (hierarchical) approach, which is estimated using MCMC (Markov Chain Monte Carlo simulation). At the first level, there is the log ratio of obese to non-obese medical costs at a given maturity. At the second level, there are four (non-nested) attributes that are modeled as random effects; these attributes are the injury year, the nature of injury (by aggregated ICD-9 code), the U.S. state, and the industry. The purpose of random effects in multilevel modeling is to shrink the multiple measurements within a second-level category (e.g., within a category in the group nature of injury) toward the weighted mean of group means (e.g., toward the weighted mean of the means of the categories within nature of injury). The fewer observations there are within a category (e.g., the fewer observations there are of a certain nature of injury) and the less precisely these observations are measured, the more the estimated mean of this category is shrunk toward the (weighted) mean of the category means within the group (e.g., the weighted mean of the means of the 22 categories within the group nature of injury). Conversely, the more observations there are within a category and the more precisely these observations are measured, the closer the estimated mean for this category is to its sample mean (see

Gelman and Hill [5]). This concept of shrinkage is closely related to the actuarial concept of credibility, as discussed by Guszcza [6].

For the purpose of statistical modeling (but not for matching), the multitude of observed ICD-9 codes are aggregated into 22 injury categories; such aggregation prevents an undue proliferation of regression coefficients. Table 2 details the aggregation rule. ICD-9 codes not covered in Table 2 are conditions that are typically not related to workplace injuries and illnesses (and, for this reason, are either very rare or do not show up at all in the raw claims data set). Among these conditions are mental disorders, complications of pregnancy, congenital anomalies, and others.

Random effects modeling is available only for categorical variables for which more than two values are observed. Gender, MUR, and MBR can take on only two alternative values and, hence, are modeled as indicator variables. These three indicator variables equal unity if (respectively) the claimant is female, is subject to MUR, and subject to MBR.

**Table 2:** Injury Category (Aggregated ICD-9 Codes)

| Category | ICD-9 Codes | Description |
|---|---|---|
| 1 | 001–289.9 and 390–629.9 | Diseases other than diseases of the musculoskeletal system and connective tissue and diseases of the nervous system and sense organs |
| 2 | 320–389.9 | Diseases of the nervous system and sense organs |
| 3 | 710–739.9 | Diseases of the musculoskeletal system and connective tissue |
| 4 | 800–829.1 | Fractures |
| 5 | 830–839.9 | Dislocation |
| 6 | 840–848.9 | Sprains and strains of joints and adjacent muscles |
| 7 | 850–854.1 | Intracranial injury, excluding those with skull fracture |
| 8 | 860–869.1 | Internal injury of thorax, abdomen, and pelvis |
| 9 | 870–897.7 | Open wounds |
| 10 | 900–904.9 | Injury to blood vessels |
| 11 | 905–909.9 | Late effects of injuries, poisonings, toxic effects, and other external causes |
| 12 | 910–919.9 | Superficial injury |
| 13 | 920–924.9 | Contusion with intact skin surface |
| 14 | 925–929.9 | Crushing injury |
| 15 | 930–939.9 | Effects of foreign body entering through orifice |
| 16 | 940–949.5 | Burns |
| 17 | 950–957.9 | Injury to nerves and spinal cord |
| 18 | 958–959.9 | Certain traumatic complications and unspecified injuries |
| 19 | 960–979.9 | Poisoning by drugs, medicinal, and biological substances |
| 20 | 980–989.9 | Toxic effects of substances chiefly nonmedical as to source |
| 21 | 990–995.94 | Other and unspecified effects of external causes |
| 22 | 996–999.9 | Complications of surgical and medical care, not elsewhere classified |

Finally, the influence of age is potentially nonlinear. (Age is the only continuous variable in the model and, hence, the only variable the influence of which may be nonlinear.) Thus, age is modeled in the nonparametric component of the model, which is implemented as a linear spline using reversible jump MCMC. Reversible jump MCMC is a concept of Bayesian model averaging, which is applied to the number of knots of the spline; knots in linear splines are locations where the linear function changes slopes. The model averages over a set of specifications with alternative numbers of knots; the location of these knots are determined by the model. For details on reversible jump MCMC in the context of linear splines, see Lunn, Best, and Whittaker [7].

We use a normal likelihood for the dependent variable (which is defined as the log ratio of obese to non-obese claim costs, by pair of matched claims, as discussed). This likelihood for claim $i$ reads

$$y_i \sim \mathrm{N}(\mu_i, \pi) \ , \tag{2}$$

where $\mu_i$ is the expected value and $\pi$ is the precision. (As is common practice in Bayesian modeling, the notation is in term of precision, which is defined as the reciprocal value of the variance.) We use a gamma prior, $\mathrm{Ga}(1, 0.001)$, for the precision.

The fixed effect of indicator variable $j$ (which, for instance, equals unity for female claimants and zero otherwise) reads

$$\beta_j \sim \mathrm{N}(0, 0.001) \ . \tag{3}$$

The random effect specification for a given set $k$ of indicator variables (such as those that represent the 22 categories of the group nature of injury) reads

$$\delta_{k,m} \sim \mathrm{N}(0, \pi_k) \ , \tag{3}$$

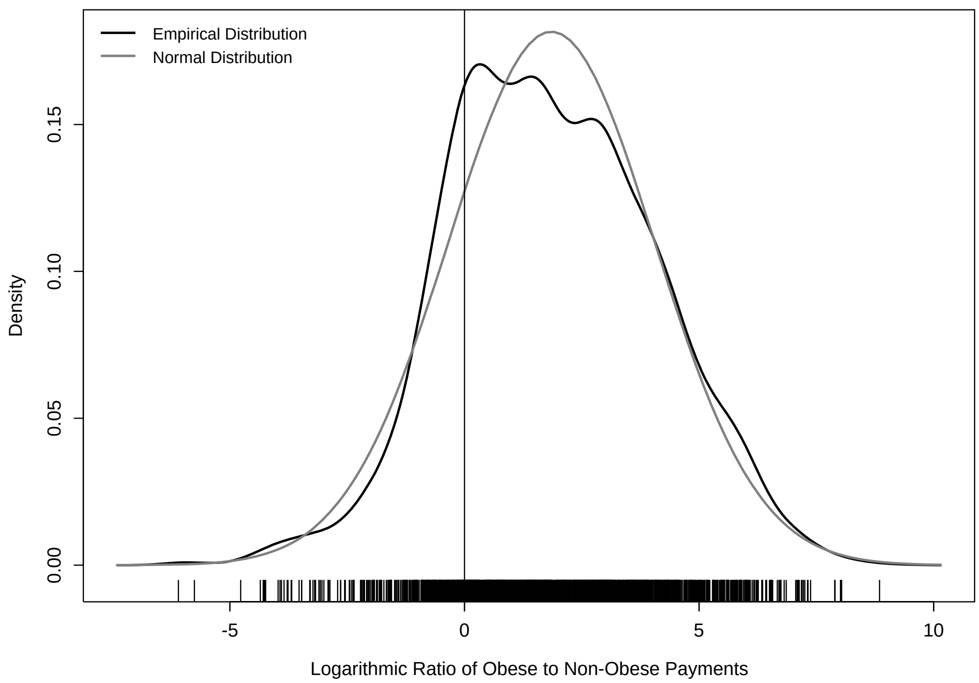
where $\delta_{k,m}$ is the effect of attribute $m$ within group $k$ (e.g., category 20 within the group nature of injury). Again, the prior distribution for the precision reads $\mathrm{Ga}(1, 0.001)$. Note that the random effects are draws from a common distribution, and that these effects are centered on zero.

Finally, reversible jump MCMC is implemented using a linear spline on age at injury; the prior for the number of knots is a uniform categorical distribution on the integers in the interval [0,trunc(range(*age*)/2)+1], where *age* is the set of observed values for age at injury (measured in whole years), "range" determines the difference between the maximum and minimum values, and "trunc" rounds down to the nearest integer. The prior distribution for the location of knots is uniform on the interval defined by the minimum and maximum observed age.

It is worthy of note that in the model outlined above, the (percentage) effect of obesity is allowed to vary with all covariates; this is because the dependent variable is the log ratio of the cost of an obese claim to the cost of a comparable non-obese claim. Thus, instead of postulating a uniform percentage effect of obesity across all claims, the percentage cost impact of obesity is allowed to vary by injury year, nature of injury, U.S. state, industry, gender, and age (and, where applicable, by the legislative environment). This approach is in keeping with the approach taken by Truls, Dement, and Krause [9] in their multivariate modeling of the claim costs of Duke University employees.

Before applying the statistical model, we plot the empirical distribution of the dependent variable (for one of the 50 randomized data sets) against the normal. As Chart 6 shows, the normal likelihood is a fair assumption for modeling the log ratio of obese to non-obese claim costs at the 60-month maturity, in spite of some degree of skewness; the distributions for the shorter maturities are similarly close to normal. (The whiskers at the bottom of Chart 6 indicate the locations of the observations.)

**Chart 6:** Kernel Density Estimate for a Randomized Set of Log Ratios (for Pairs of Matched Claims) of Obese to Non-Obese Cumulative Payments at the 60-Month Maturity




## 4. FINDINGS FOR RANDOM EFFECTS, GENDER, AND AGE

We estimate two versions of the model. In a first version, we include as covariates the injury year, the nature of injury, the U.S. state, and the industry; these four covariates are modeled as random effects; further, we include as covariates gender and age at injury, but we do not account for differences (across states and over time) in the legislative environment. In a second version of the model, the findings of which are presented in Section 5, we add indicator variables for MUR and MBR; the explanatory power of these covariates will subtract from the measured random effects of the injury year and the state. That is because without the covariates MUR and MBR, variations in the legislative environment over time and across states are absorbed by the random effects of injury year and state, respectively.

All findings shown in this section pertain to the 60-month maturity. Chart 7 offers a graphical exposition of the estimated random effects of the injury years 1998 through 2002 (which are the only injury years available for the 60-month maturity). The gray dots and whiskers indicate the mean and 80 percent credible intervals for 50 randomized data sets. The black dots and whisker signify the mean and 80 percent credible intervals after aggregating across all 50 data sets. Clearly, there is variation across injury years, and there appears to be a mild time trend in the ratio of obese to non-obese medical claim costs. Remember that, by definition, random effects are centered on zero.

**Chart 7:** Mean and 80 Percent Credible Intervals for the Random Effect of the Injury Year in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims at a Maturity of 60 Months

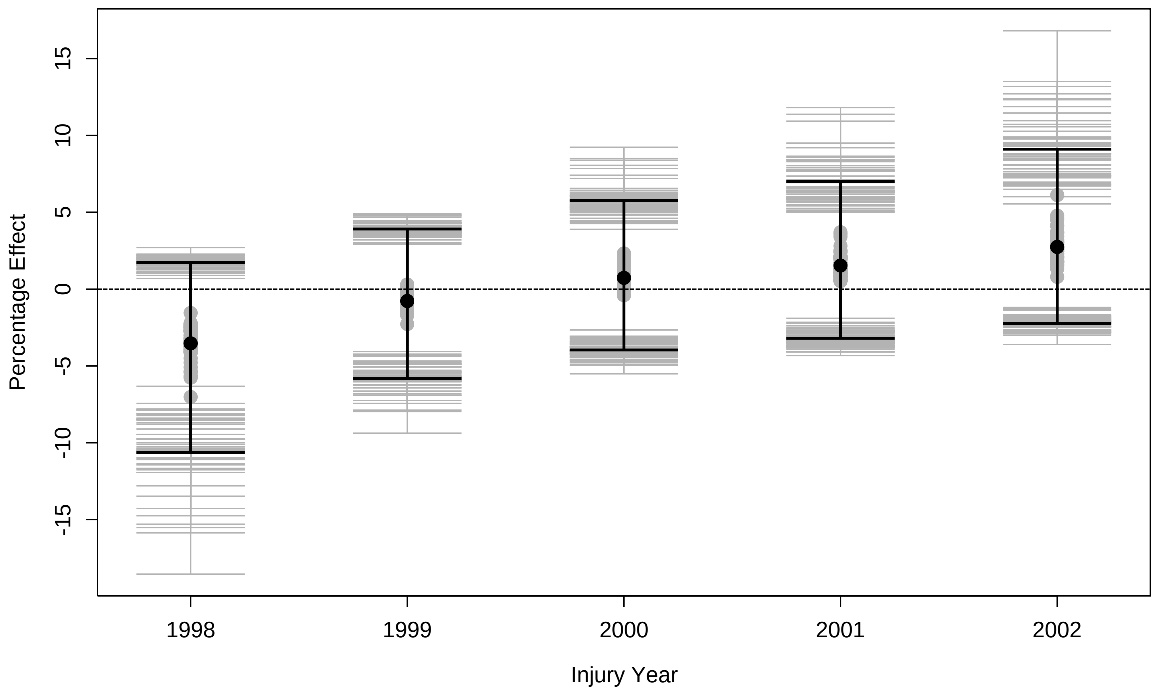



Chart 8 presents the random effect estimates for the 22 injury categories (of which only 19 are observed at the 60-month maturity). The mean of the estimated effect of a given injury category is indicated by the location of the number of the category. The fewer observations there are in a given category and the "noisier" these observations are, the wider the credible intervals. The injury categories with the highest percentage cost contributions to obesity are 7 (intracranial injury, excluding those with skull fracture), 9 (open wounds), 11 (late effects of injuries, poisonings, toxic effects, and other external causes), 14 (crushing injury), 17 (injury to nerves and spinal cord), and 18 (certain traumatic complications and unspecified injuries). At the same time, some of these injury categories are among those with the widest credible intervals; exceptions are categories 9 and 18. The scale of differences among injury categories is quite large. For instance, an estimated displayed effect of 50 percent for a given injury category implies that for this category, the effect of obesity on the ratio of claim costs is 50 percent higher than the average across categories.

**Chart 8:** Mean and 80 Percent Credible Intervals for the Random Effect of the Injury Category in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims at a Maturity of 60 Months

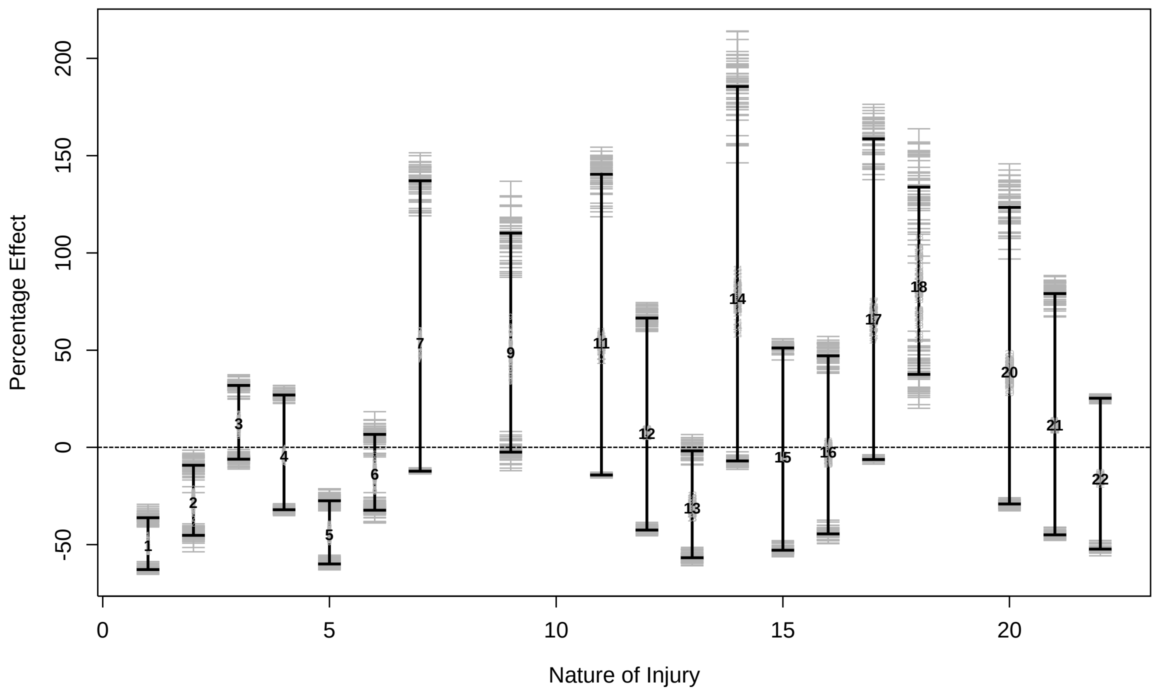



Chart 9 displays the random effect estimates for the 36 U.S. states included in the analysis. Here again, the states with above or below-average estimated effects of obesity also tend to be the ones with large credible intervals. Finally, Chart 10 shows the random effect estimates for the five NCCI industries.

**Chart 9:** Mean and 80 Percent Credible Intervals for the Random Effect of the U.S. States in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims at a Maturity of 60 Months

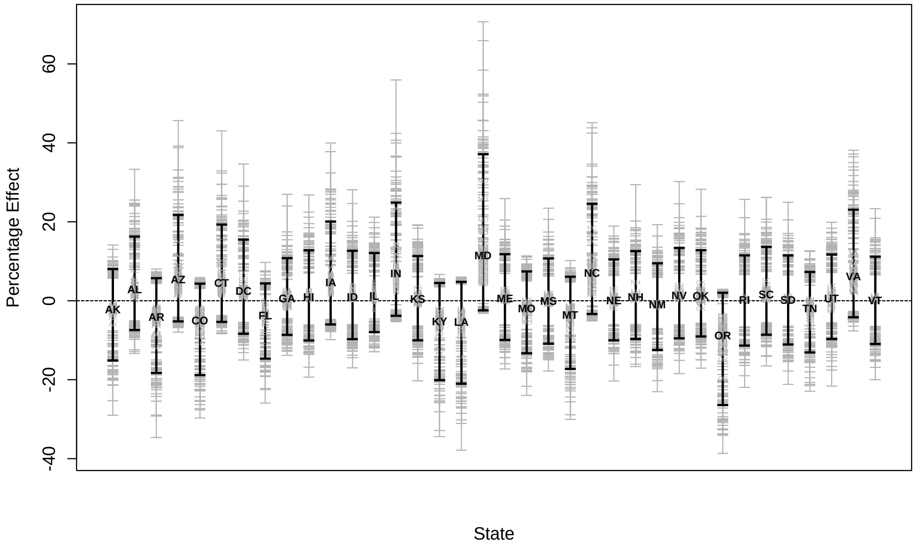
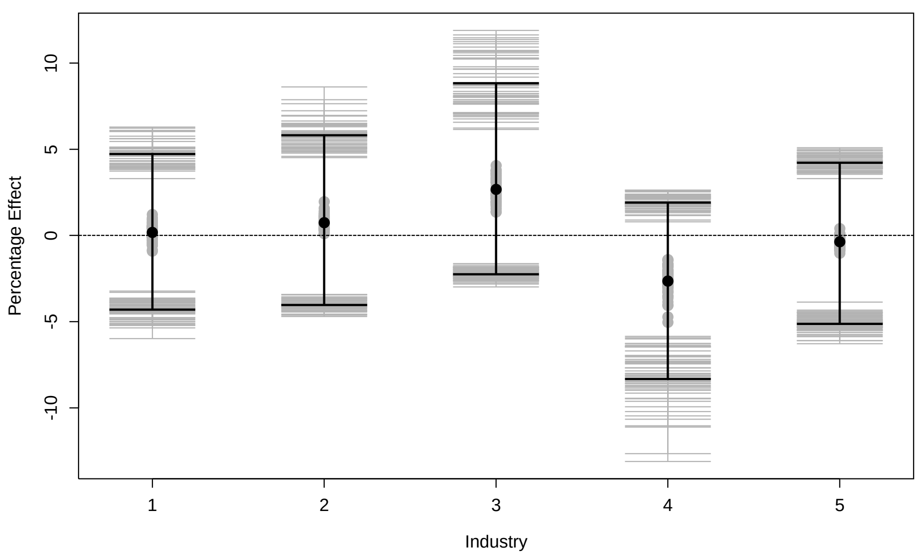



**Chart 10:** Mean and 80 Percent Credible Intervals for the Random Effect of the Industry in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims at a Maturity of 60 Months. Industry codes: (1) Manufacturing; (2) Contracting; (3) Office and Clerical; (4) Goods and Services; (5) Miscellaneous



The random effects for the injury year and the industry are quite small compared to those of the injury category. This finding is summarized in Chart 11, which offers a comparison of the variances of the four random effects. As shown in this chart, the injury year and the industry have comparatively little explanatory power. There is more variation across states than there is over time and across industries. But the most explanatory power originates in the injury categories.

**Chart 11:** Boxplots for the Means of the Posteriors of the Variances of the Random Effects

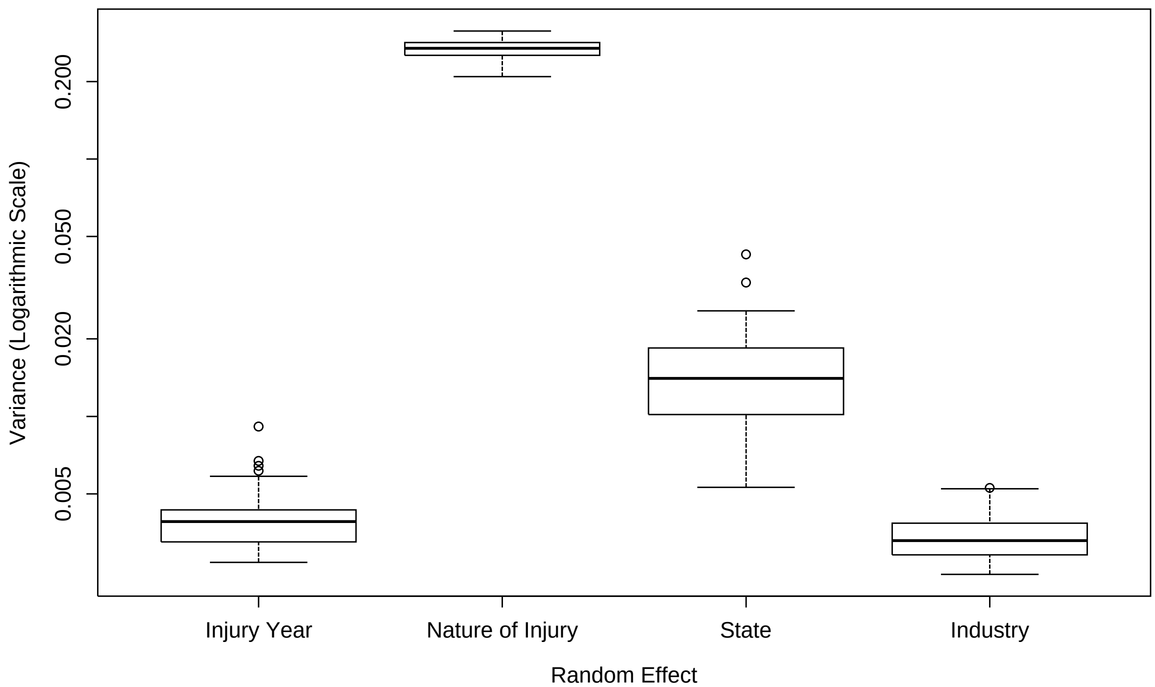
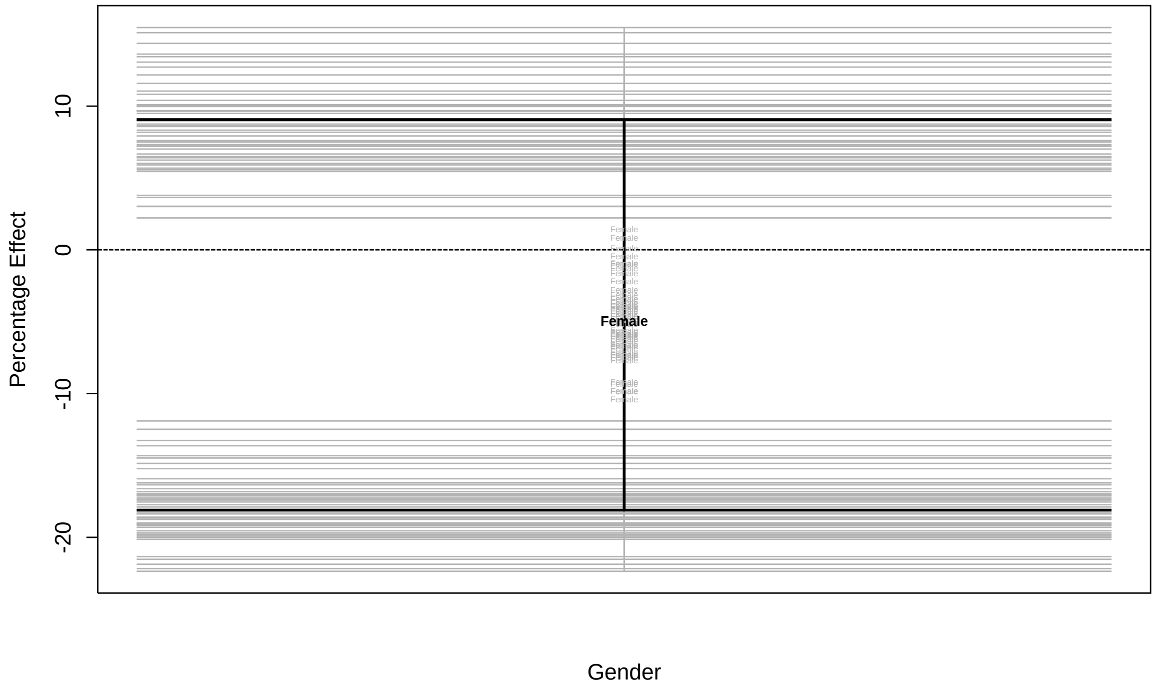


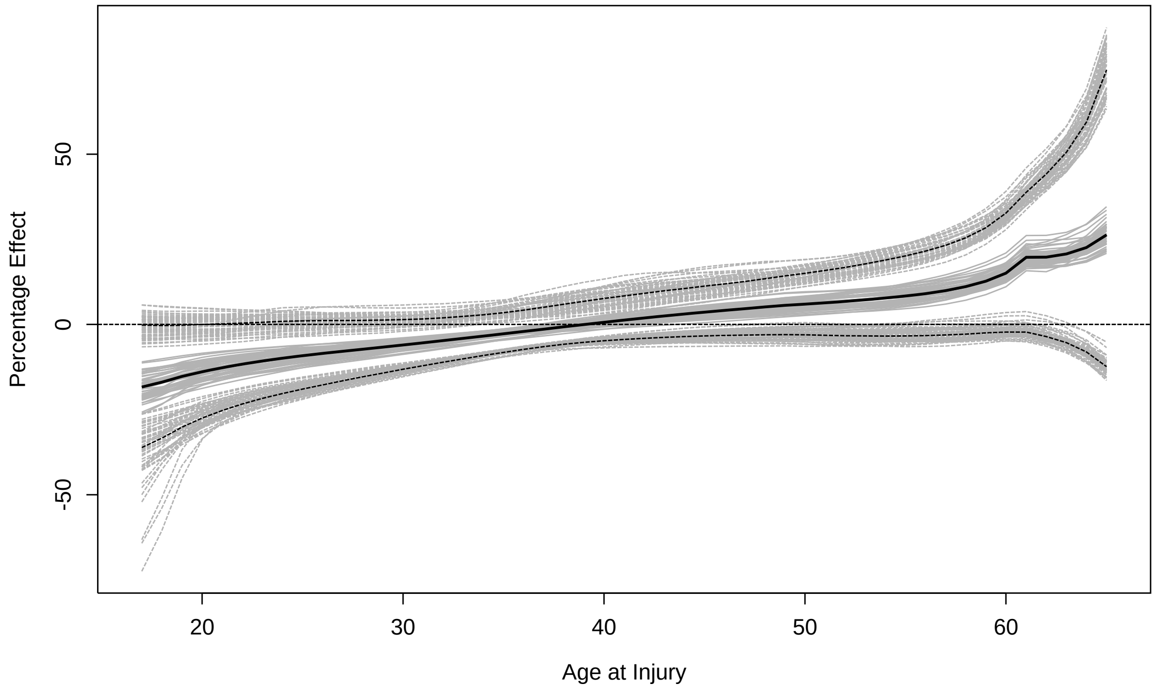

Chart 12 displays the effect of the female gender on the difference between obese and comparable non-obese claims. Here again, the gray whiskers (and "Female" label indicating the mean) pertain to the 50 individual, randomized data sets, whereas the black whiskers (along with the black "Female" label) indicate the results obtained when aggregating over these 50 runs. Although the mean estimate of the female gender is negative, the credible interval includes the zero value. Then again, the negative effect of the female gender on the ratio of obese to non-obese claim costs agrees with the findings of Truls, Dement, and Krause [9] in the Duke University study.

**Chart 12:** Mean and 80 Percent Credible Intervals for the Effect of Gender in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims at a Maturity of 60 Months




Finally, Chart 13 presents the effect of age. (Note that this effect is centered on zero.) As discussed, if the medical costs of workers compensation increase with age, then, due to matching toward the center, the effect of age shows up as "S-shaped." This is in fact the case. As a consequence, the displayed effect should be read as an effect of matching toward the center, instead of an effect of age, on the percentage cost difference between obese and comparable non-obese claims.

**Chart 13:** Mean and 80 Percent Credible Interval for the Influence of Age in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims at a Maturity of 60 Months




## 5. FINDINGS FOR THE LEGISLATIVE ENVIRONMENT

As discussed, in a second version of the model, we add MUR and MBR as covariates. Information on MUR and MBR is provided by the Workers Compensation Research Institute (WCRI) for calendar years 1997 and 2001. For injury years 2001 and later, the 2001 value applies; for earlier injury years, the 1997 value is used. The covariates are coded as indicator variables that are equal to unity for U.S. states with MUR and MBR (respectively) in place, and zero otherwise.

WCRI [10] puts a state in the "mandated utilization review/management category" if such jurisdiction mandates that "payers review claims for proper medical care utilization…or if the workers' compensation agency or exclusive state fund conducts utilization review on its own initiative (either for all claims or those that meet certain criteria)." Further, WCRI [10] credits a state with a bill review program if in such jurisdiction "the workers' compensation agency routinely examines for proper charges all bills or cases that meet explicit criteria, such as a certain number of days lost from work or a dollar limit on medical care costs; the statute mandates the examination of bills by payers; or an exclusive state fund does regular bill review."

Table 3 details the MUR and MBR categorization of states for the two available years. In 1997, twelve states had MUR, eight states had MBR, and six jurisdictions had both legislative provisions in place. By 2001, the number of states with MUR had dropped to nine, those with MBR had risen to ten, and those with both provisions had decreased to five.

Chart 14 offers boxplots for the estimated 50 values of the influence of MUR, as obtained when analyzing the 50 randomized data sets. Clearly, MUR significantly reduces the ratio in the medical costs per claim of obese to non-obese claimants. For the 12-month maturity, MUR reduces this difference by 15.8 percent. For the 36-month maturity, the effect of MUR grows to 18.8 percent, before reaching 28.5 percent at the 60-month maturity. This evidence attests to the importance of the legislative setting for the medical cost of workers compensation claims.

Chart 15 offers boxplots for the 50 estimated effects of MBR. Although the cost containment effect of MBR is substantial, its influence is weaker than MUR. For the 12-month maturity, MBR reduces the ratio in the medical costs per claim of obese to non-obese claimants by 9.4 percent. For the 36-month maturity, the effect of MUR grows to 16.6 percent, before dropping back to 12.2 percent at the 60-month maturity. Clearly, the maximum cost containment effect is achieved where both of the two legislative provisions are in place: The combined effect MBR and MUR (not shown) at the 12-month maturity amounts to a negative 24.0 percent, before growing to a negative 33.0 percent at the 36-month maturity and reaching a negative 37.4 percent at the 60-month maturity.

**Table 3:** MUR and MBR by State

| State | Year 1997 MUR | Year 1997 MBR | Year 2001 MUR | Year 2001 MBR |
|---|---|---|---|---|
| AL | No | No | No | No |
| AK | No | No | No | No |
| AZ | No | No | No | No |
| AR | Yes | Yes | No | Yes |
| CO | Yes | Yes | Yes | Yes |
| CT | No | No | No | No |
| DC | No | No | No | No |
| FL | Yes | Yes | Yes | Yes |
| GA | No | No | No | No |
| HI | No | No | No | No |
| ID | No | No | No | No |
| IL | No | No | No | No |
| IN | No | No | No | No |
| IA | No | No | No | No |
| KS | No | No | No | No |
| KY | Yes | No | Yes | Yes |
| LA | Yes | Yes | Yes | Yes |
| ME | Yes | No | Yes | No |
| MD | No | No | No | No |
| MS | Yes | Yes | Yes | Yes |
| MO | No | No | No | No |
| MT | Yes | No | No | No |
| NE | No | No | No | No |
| NV | Yes | Yes | No | Yes |
| NH | No | No | No | No |
| NM | Yes | No | Yes | No |
| NC | No | No | No | Yes |
| OK | No | No | No | No |
| OR | No | Yes | No | Yes |
| RI | No | No | No | No |
| SC | No | Yes | No | Yes |
| SD | No | No | No | No |
| TN | Yes | No | Yes | No |
| UT | Yes | No | Yes | No |
| VT | No | No | No | No |
| VA | No | No | No | No |

Source: WCRI [10][11].

**Chart 14:** Box Plots for the Means of the Influence of MUR in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims

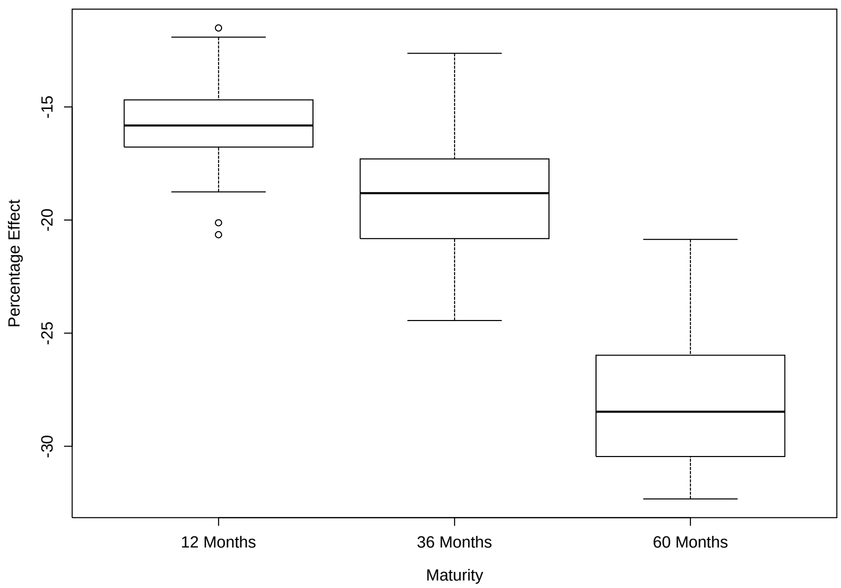
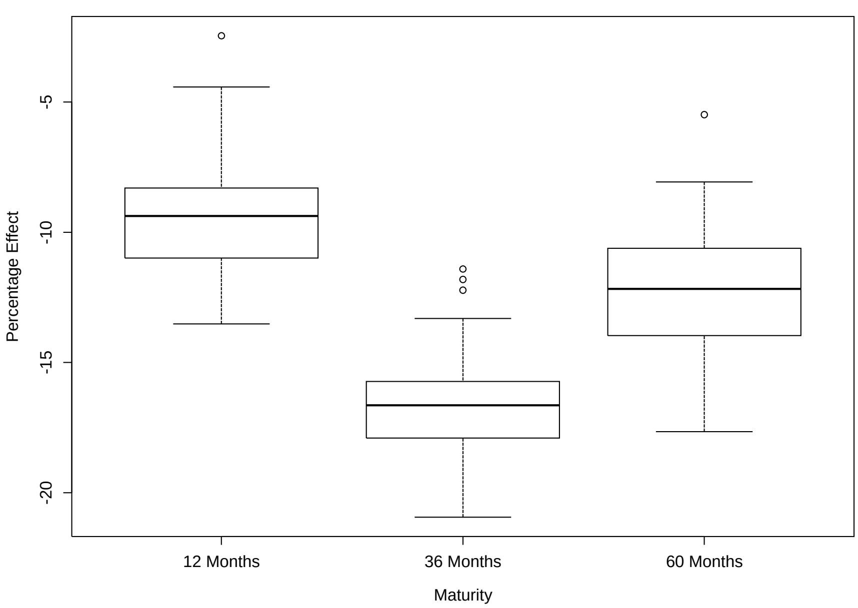


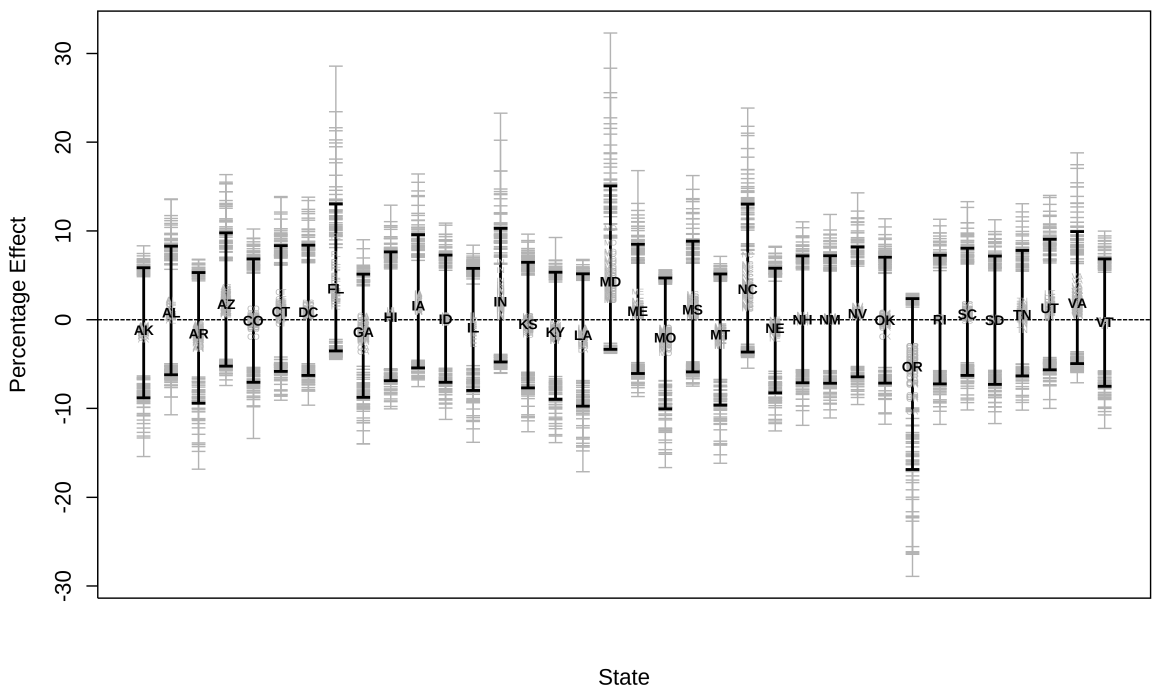

**Chart 15:** Box Plots for the Means of the Influence of MBR in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims



As discussed, including MUR and MBR as covariates subtracts from the random effects of injury year and state. This is because some of the variation over time and across states is due to these two legislative provisions. It is therefore of interest to revisit the random effects of states and to investigate how much of the variation across states is due to these two covariates.

Chart 16 displays the random effects for the states when MUR and MBR are included in the regression equation. Clearly, the cross-state variation of the ratio in the medical costs per claim of obese to non-obese claimants is now much smaller than previously. In fact, the variance of the state random effects drops to 0.00579 from the previous level of 0.01406, which implies that MUR and MBR explain about 59 percent the cross-state variation in the percentage cost effect of obesity.

**Chart 16:** Mean and 80 Percent Credible Intervals for the Random Effect of U.S. States in a Partial Linear Multilevel Regression on the Ratio of Cumulative Payments of Obese to Non-Obese Claims for 50 Sets of Pairs of Matched Claims at a Maturity of 60 Months when Controlling for Differences in the Legislative Environment




## 6. CONCLUSIONS

Studying a large data set of obese claims, which comprises 36 U.S. states and nine injury years, we are able to provide evidence on the effect of obesity on the medical cost per claim. We showed that the effect of obesity is substantial, and that the entirety of the effect of obesity reveals itself only over time, as claims mature. Most importantly, we were able to quantify the effect of the legislative environment on the effect of obesity on claim costs.

Our methodology differs in important ways from the Duke University study by Truls, Dement, and Krause [9]. In part, this dissimilarity in approach is necessitated by a difference in the nature of the data set, but also by a difference in objective. Although our data set is considerably larger (as it comprises millions of claims), it is also more limited in the scope of the data items. Most importantly, we have no information on the BMI. Instead of having an objective criterion for obesity (and having obesity differentiated by degree), our data set offers an obesity categorization that relies on the physician's decision to list obesity as a co-morbidity indication.

This difference between subjective and objective categorization of obesity has potential implications for the legislative findings of our model. For instance, it may be argued that the decision of the physician to diagnose a claimant as obese is influenced by the legislative environment, thus causing an endogeneity bias in our measurement of the effects of MUR and MBR. Although we cannot refute such a proposition with confidence, it is important to note that the potential endogeneity is likely to cause the effects of MUR and MBR to be underestimated (instead of overestimated). This is because if the increased scrutiny of MUR and MBR raises a physician's propensity of coding claimants as obese (in order to provide better documentation supportive of the treatment), the implied broader definition of obese claims diminishes the recorded cost difference between obese and comparable non-obese claims in the data set.

Finally, the proportion of claims with obesity as a co-morbidity indicator is comparatively small (0.1 to 0.2 percent) when compared with the proportion of the obese in the workforce. It is likely that obesity serves as a co-morbidity indicator primarily where complications from obesity are highly probable (e.g., when claimants are morbidly obese) or have already materialized. From this perspective, the measured effect of obesity on claim severity may be viewed as an upper bound.

## Acknowledgment

Thanks to Uriel Carrasquilla, John Potter, Tanya Restrepo, John Robertson, Harry Shuford, and Martin Wolf for comments and to Chun Shyong for research assistance.

## 7. REFERENCES

[1] Carlin, Bradley P., and Thomas A. Lewis, *Bayes and Empirical Bayes Methods for Data Analysis.* 2nd ed., Boca Raton: Chapman & Hall/CRC, 2000.
[2] Chambers, John M., William S. Cleveland, Beat Kleiner, and Paul A. Tukey, *Graphical Methods for Data Analysis,* New York: Chapman and Hall, 1983.
[3] Finkelstein, Eric A., Ian C. Fiebelkorn, and Guijing Wang, "National Medical Spending Attributable to Overweight and Obesity: How Much, and Who's Paying," *Health Affairs* 22, May 14, 2003, pp. w3-219-w226.
[4] Finkelstein, Eric A., Justin G. Trogdon, Joel W. Cohen, and William Dietz (2009) "Annual Medical Spending Attributable to Obesity: Payer- and Service-Specific Estimates," *Health Affairs* 28(5), pp. w822-w831, DOI 10.1377/hlthaff.28.5.w822.
[5] Gelman, Andrew, and Jennifer Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models,* Cambridge, MA: Cambridge University Press, 2007.
[6] Guszcza, James, "Hierarchical Growth Curve Models for Loss Reserving," Casualty Actuarial Society *E-Forum,* Fall 2008, http://www.casact.org/pubs/forum/08fforum/7Guszcza.pdf.
[7] Kasabov, Nicola K., *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*, Cambridge, MA: MIT Press, 1996.
[8] Lunn, David J., Nicky Best, and John C. Whittaker, "Generic Reversible Jump MCMC Using Graphical Models," *Statistics and Computing* 19(4), pp. 395-408, DOI 10.1007/s11222-008-9100-0.
[9] Truls, Østbye, John M. Dement, and Katrina M. Krause, "Obesity and Workers' Compensation: Results from the Duke Health and Safety Surveillance System," *Archives of Internal Medicine* 167*,* 2007, pp. 766-773.
[10] Worker Compensation Research Institute, *Managed Care and Medical Cost Containment in Workers' Compensation: A National Inventory, 2001-2002*, Cambridge, MA: WCRI, 2001.
[11] Worker Compensation Research Institute, *Managed Care and Medical Cost Containment in Workers' Compensation: A National Inventory, 1997-1998*, Cambridge, MA: WCRI, 1997.

**Abbreviations and notations**

BMI, Body Mass Index
FTE, Full-Time Equivalent
HRA, Health Risk Assessment
ICD-9, International Classification of Diseases, Ninth Revision
MBR, Mandatory Bill Review
MCMC, Markov Chain Monte Carlo simulation
MUR, Mandatory Utilization Review
NCCI, National Council on Compensation Insurance
WCRI, Workers Compensation Research Institute

## Biographies of the Authors

**Chris Laws** is a Senior Actuarial Analyst at the National Council on Compensation Insurance, Inc.
**Frank Schmid**, Dr. habil., is a Director and Senior Economist at the National Council on Compensation Insurance, Inc.