# Occupational Safety and English Language Proficiency

Marvasti, Akbar

University of Southern Mississippi

July 2008

# Occupational Safety and English Language Proficiency

Akbar Marvasti
Department of Economics and Finance
University of Southern Mississippi
118 College Drive #5072
Hattiesburg, MS 39406
Tel. (601) 266-4484
Fax (601) 266-4920
E-mail: .marvasti@usm.edu

JEL Classification: J28, J88

I am grateful for the helpful comments made by Brooks Pierce of the Bureau of Labor Statistics and by participants of my session at the 2008 Western Economic Association International meetings in Honolulu. All remaining errors are mine.

**Abstract**

Recent occupational injury data shows a rising trend, which happens to coincide with both increases in the population of foreign born in the U.S. and with changes in its composition. This study aims at exploring the presence of a statistical relationship between occupational injuries and the level of English proficiency of foreign born using cross-sectional data on the rate of injury and count of injury incidents. A cultural gap hypothesis is also examined as an alternative explanation for the rise in work injuries. While there is some support for the adverse effect of inadequate English language proficiency of foreign born, the results for the cultural gap hypothesis are more robust.

# Occupational Safety and English Language Proficiency

## I. Introduction

The objective in this study is to investigate the link between language proficiency of foreign-born workers and the prevalence of work injuries among them. The Bureau of Labor Statistics (BLS) data shows that, since 1992, while the overall number of work injuries and fatal accidents has fallen for the population, the number of work injuries and fatal accidents among the foreign-born segment of the labor force has been rising. In fact, this number has more than doubled since 1992. It is notable that in 2006, Mexican immigrants suffered 44 percent of the work injuries. Of course, one explanation for this observation is that Hispanics are disproportionately employed in high-risk industries such as construction. At the same time, Census Bureau statistics shows declining language proficiency among more recent cohorts of immigrants. Effective communication between workers and the Occupational Safety and Health Administration (OSHA), or management, is crucial in reducing workplace accidents. Work injuries impose a significant cost to workers, business and government. Establishing a clear connection between English language proficiency and work injuries, or fatalities, while taking all other related factors into account, enhances our understanding of the consequences of changes in composition of immigrant population and helps design public policy aimed at improvement of workers' English skills or even communication of safety rules to them in their native language.

Numerous studies have dealt with various aspects of the rising immigrant population in the U.S., such as housing, education, and the labor market. However, studies of the effect of foreign-born English language proficiency on work injury and fatality is scarce. Henshaw of the U.S. Department of Labor believes that risk of work injury or fatal accidents is greater for immigrants, especially Hispanics, than the rest of the population (Henshaw, 2002). Is there a statistical connection between the English language proficiency of foreign-born workers, or their culture, and work injuries? If so, what is the magnitude of such an effect? While a number of studies have dealt with occupational safety, Lanoie (1992-a) is a rare example of using ethnicity as an explanatory variable in a work injury model where he finds a positive correlation between the percentage of ethnic minorities and incidence of accidents in Canada. While OSHA suspects that lack of English proficiency is the source of the problem because of the inability of some foreign-born workers to read safety manuals and follow safety rules, there are no studies exploring the role of English proficiency in the rising occupational injuries and fatal work accidents trend among immigrants.[1]

Several studies however have examined other aspects of work injury or fatal accidents. For example, Adnett and Dawson (1998) use work accident data from the U.K. to demonstrate that accidents positively correlate with business cycles, but wage premium for high accident occupations is absent, particularly during high unemployment periods. Kahn (1987) in his analysis of job tenure in non-union firms finds that workers with more tenure actually prefer more risky jobs. Kahn also explores the non-linearities between work injuries and workers' age. Controlling for experience, he finds that as workers' age injury rate falls, which he attributes to an increasing risk aversion. Viscusi (1979) discovers that more educated workers are less likely to get injured, arguing that more educated workers have access to safer jobs, avoiding industries with high accident rate, while Chelius (1974) conjectures that the negative relation is due to lower propensities

(Chelius, 1974). Viscusi also finds lower injury rates for blacks, while Smith (1974) and McLean et al. (1978) reach an opposite conclusion. Kahn (1987) shows that males are more accident prone, which he believes to be due to their aggressive behavior. Ruser (1995) uses plant level manufacturing injury rate data to deal with the effectiveness of OSHA regulations. He finds a significant positive residual in his estimates even in the presence of OSHA, arguing in favor of OSHA inspection targeting establishments with high injury rates. Ruser (1993) uses counts of injuries data in manufacturing, to show that increases in workers' compensation benefits increases non-fatal injuries and days away from work, while reducing the frequency of fatalities. He acknowledges that workers are likely to receive compensating wages for taking high risk jobs, but due to lack data he could not deal with the simultaneity problem. Lanoie (1992-a), on the other hand, benefiting from a richer data set is able to take the endogeneity of wages into account in his analyses. Decker and Flynn (2008) link worker safety incentives to market structure in the U.S. steel industry by demonstrating that increases in competition motivate firms to increase their efforts to improve worker safety. They find that both prevalence of unions and high workers' compensation reduce work injuries, while high unemployment is associated with high work injuries.

In this paper, I attempt to explore the effect of foreign-born workers and their English language proficiency on occupational injuries using cross-sectional data from the U.S. I do so first by briefly looking at the changes in the composition of immigrants into the U.S. over time and the state of English proficiency of the foreign born. I next present the state level data set and the econometric models for the rate of injury and the count of the number of injury incidents. Then, I present the results from the least squares and

Poisson regressions measuring the impact of English proficiency and cultural differences on the incident of work injury. At the end, a summary of the findings and conclusions are presented.

**II. A Glance at the Foreign Born and English Proficiency**

Some attribute the declining language proficiency of the labor force to the changing pattern of migration to the U.S., which has historically been influenced by the expansion of the agricultural sector, world wars, the Great Depression, and legislative mandates. In fact, when the flow of immigrants to the U.S. has fluctuated, the composition of immigrants has also changed. Perhaps the strongest determinant of the composition has been legislative changes. Enactment of restrictive immigration laws in the 1920s created a national-origin quota system, which limited the number of immigrants. These laws also favored immigrants from the Western Hemisphere and Western Europe. However, the Immigration Act of 1965 abolished the national-origin quota system and subsequent legislation in 1986 and 1990 encouraged immigration. For example, according to the U.S. Census Bureau, in 1850, 92.2 percent of the foreign born were from Europe, 6.7 percent from North America, and only 0.1 percent from all other countries. By 2006, only 12 percent of the foreign born were from Europe and 2 percent from North America while 54 percent were from Latin America (of which, approximately 58 percent from Mexico) and 26 percent from Asia. To put the rapid growth of the population of foreign born from Latin America in perspective, in 1960, 0.9 million Latin American foreign born lived in the U.S. By 2006, this number had grown to 19 million. In fact, Latin Americans are now the largest minority in the country, surpassing African Americans.

Changes in the composition of the foreign born in the country have generated debates about their level of education and skills, particularly concerning language proficiency. Language proficiency is a key ingredient in human capital, which affects several aspects of an individual's life including job qualifications and cultural identity. English proficiency, in particular, is important to the foreign born as workers in the U.S. labor market and is likely to affect the incident of work injuries. Data on the spoken language of the foreign born shows that the percentage of foreign born (5 years and over) who speak only English has dropped steadily between 1980 and 2000. The figures for 1980, 1990, and 2000 are 30 percent, 21 percent, and 17 percent, respectively. In 2006, from 37.2 million foreign born who were 5 years and over, 13.3 million claimed to speak English very well. This means that 64 percent of the foreign born believed that they did not speak English very well. In addition, the Census Bureau data shows that 46 percent of the foreign born spoke Spanish in 2006, from which only 27 percent claimed to also speak English very well.

Based on the 2006 Census Bureau's American Community Survey, English proficiency of foreign born affects their labor force participation and the likelihood of being unemployed. In the U.S. as a whole, 31 percent of foreign born in the working age group of 18-64 do not speak English well or at all. However, 39 percent of those who are not in the labor force do not speak English well or at all. Among those who are in the labor force, 37 percent of the unemployed do not speak English well or at all. This apparent screening in the labor market based on English language proficiency might have reduced the potential adverse effects of poor English communication skills of foreign-born workers on work injury.

## II. Data

U.S. state level data, including the District of Columbia, from the BLS Survey of Occupational Injuries and Illnesses for 2006 are used in the statistical analysis. Since a few states have chosen not to participate in the survey, only 43 observations are used in this study.[2] Foreign-born English proficiency data is extracted from the American Community Survey by the Census Bureau (CB).[3] Two measures of English proficiency are computed from the CB data. The first measures the percentage of the foreign-born working population, age 18-64, who do not speak English well or at all. The alternative is the percentage of employed population age 18-64 who do not speak English well or at all. The high-risk industries variable is based on identification of top ten six-digit NAICS industries with the highest rate of injuries, then, calculating percentage of employment in these industries as a fraction of the total state employment.[4] The BLS is also the source for percentage of foreign-born workers, unemployment rate, union membership, and employment statistics data. Workers' compensation data are from the U.S. Chamber of Commerce. Obesity data is from the Center for Disease Control. The descriptive statistics are presented in Table 1.[5] Mean and standard deviation for the state for which injury data is not available shows that overall there is no sample selection bias. However, a smaller fraction of the labor force in the non-participating states is made up of foreign-born workers and a smaller fraction of the employed population has poor English communication skills, which may cause a potential sample selection bias. Because of the small sample size for the excluded states, analysis of variance is not feasible.

## III. Econometric Model

Two empirical approaches are applied to the alternative measures of work injury incidence. Since the rate of injury series is a conditional probability limited between zero and one, it follows a logistic distribution. Therefore, a logistic regression is employed where $P_s$ is the proportion of the labor force in state $s$ that has suffered from work injury. Then, the logistic transformation of $P_s$ is determined by:

$$ln(\frac{P_s}{(1-P_s)}) = F(E_s, \mathrm{X}_s) + e_s, \qquad s = 1, ...., n \qquad (1)$$

where, $E_s$ is English proficiency of the foreign-born workers, $\mathrm{X}_s$ is a vector of control variables which includes other characteristics of the workers and market conditions in each state, and $e_s$ is a random error term with $E\left[e_s\right] = 0$ and $Var\left[e_s\right] = \frac{P_s(1-P_s)}{n}$.

The primary hypothesis to be tested is whether the level of English proficiency among foreign-born workers positively affects the chances of work injury. Therefore, the coefficients of measures of English proficiency are expected to be positive and statistically significant. An indirect measure of the role of English proficiency would be testing the significance of the percentage of foreign born in the workforce. Because the new cohort of foreign born is from non-English speaking countries, particularly the Latin American countries, the coefficient of the percentage of foreign born who do not speak English well is expected to be positive and statistically significant. Considering the origin of immigrants, variations in the percentage of foreign-born workers among states may also reflect cultural differences, which influence attitude towards risk.

The vector of control variables, $\mathrm{X}$, includes the average workers' characteristics and market conditions in the state in percentage form such as gender, education, race, marital status, obesity, union membership, unemployment rate, wage rate, as well as variables such as workers' compensation and industry risk. Age is introduced in both level and squared form to capture nonlinearities in the relationship. One expects increasing risk aversion as workers get older. On the other hand, younger workers are more physically able to avoid accidents than older workers. Lanoie (1992-a) and Dillingham (1982) findings on the effect of age support the latter argument.[6] The incidence of work accident is expected to drop with more schooling either because of the educated workers' access to safer jobs, or their lower accident propensities. Increases in the percentage of minorities such as blacks or Hispanics as well as percentage of males, controlling for high-risk occupations, increases the incidence of accidents if these workers happen to be risk-takers. Thus, the expected sign for the coefficient of these variables is positive. The expected sign for the coefficient of the unemployment rate is also positive because in states with a high unemployment rate, workers are more likely to accept risky work environments. Since in the states with stronger union presence, workers' safety regulations are more likely to be reinforced and union workers are more aware of accident prevention measures, injury incidence is likely to drop. Therefore, the sign of percentage of workers in the labor unions is negative. Overweight workers tend to be less agile and prone to have work accidents. Therefore, the coefficient of obesity is expected to be positive. Finally, more generous workers' compensation reduces workers' incentive to take accident preventive measures, causing incidents of injuries to rise, which is consistent with Ruser's findings (Ruser, 1993). On the other hand, high workers' compensation motivates employers to undertake accident preventive measures, reducing the chances of work injuries.

For the number of the injury incidence, the count data model is used for estimation of the parameters of the injury function where the probability of *y* number of occurrences of a work injury event is a non-negative integer as $e^{-\lambda}\lambda^{y} / y$ with the mean of $\lambda = \exp(E_i, X_i)$. In this *observed heterogeneity* model, the conditional expectation of the number of injuries in each state, $Y_s$, given the values of the explanatory variables or the predictors, i.e. $E(Y_s | E_s, X_s)$, must be positive. Thus a linear model, which allows negative expected values, is not suitable for count data. Instead, the Poisson maximum likelihood model seems appropriate. Therefore, a common formulation of the Poisson model in log-linear form is applied as:

$$\ln E(Y_s | E_s, X_s) = F(E_s, X_s) + \varepsilon_s \qquad (2)$$

where $\varepsilon_s$ has a gamma distribution with mean of unity and variance of $\alpha^2$.

Next, the Poisson model is modified by removing the assumption of no unobserved heterogeneity across individual states. This modification essentially corrects for misspecification in the mean function by allowing unobserved heterogeneity in the conditional mean function in the form of normally distributed error term, i.e., $\lambda = \exp(E_i, X_i, \varepsilon_i)$. This creates a *normal heterogeneity* model, where $\varepsilon_i$ reflects a random effect that is different for each worker.

Then, the possibility of underreporting work accidents is considered. Ruser (1985) and Lanoie (1992-a) have argued that work accidents might be underreported due to institutional factors. The work injuries data used here are also likely to be underreported, especially when illegal foreign-born workers are involved. This could be because both the employer and the foreign-born workers lack incentive to report work injuries due to concern about the immigration laws. A model for underreporting of count data is considered here based on replication of underreporting models developed by Winkelmann and Zimmermann (1994) and Winkelmann (1997) where the observed count of work injuries $y_i$ is dependent upon whether the count indicator $c_i$ is equal to one or zero, meaning that the underlying probability that c is 1 motivates a probability distribution of $y_i$. If the probability of a work injury that has actually occurred being reported is dependent upon the foreign-born workers covariate, $Z_i$, then:

$$c_i^* = \gamma Z_i + \mu_i \quad \text{, and}$$

$$c_i = 1 \text{ if } c_i^* > 0, \qquad (3)$$

where $\mu_i$ is an error term. Accordingly, the mean of the underlying observed work injuries is determined by the probability of work injury being observed (reported), $P_i$ , or:

$$E[y_i \mid c_i] = P_i^* E[y_i] \text{, where}$$

$$P_i = prob[c_i = 1] \text{ , or}$$

$$P_i = \phi(\gamma Z_i). \qquad (4)$$

Therefore, applying this argument to the state level data here, the mean of work injury depends upon the probability that an injured worker is a foreign born and the Poisson model is modified to one with *exogenous underreporting* as:

$$\ln E(Y_s | E_s, \mathrm{X}_s, Z_s) = \exp[F(E_s, \mathrm{X}_s)] F(\gamma Z_i) \qquad (5)$$

Before reporting the results, a battery of other tests are run to determine specification of models. The Poisson restriction of equality of the conditional mean and variance is tested using Cameron and Trivedi's overdispersion test where

$Var(Y_s | E_s, X_s) = \mu_s + \alpha\mu_s^2$ (Cameron and Trivedi, 1990). If the variance of $Y_s$ is equal $\mu_s$,

i.e. $\alpha$=0, $Y_s$ is Poisson distributed. The value of $\alpha$ is trivial in various specifications of the count model as presented in Table 3, confirming the Poisson restriction.

Examination of the simple correlation coefficients and variance inflation factor does not show any significant potential multicollinearity between explanatory variables, except between age and age squared. However, it is noted that foreign-born workers and measures of English proficiency are mildly positively correlated—ranging between 0.20 to 0.23.

The residuals of the both the logit and the Poisson models are subjected to a few tests to rule out the possibility of heteroskedasticity. In the logit model, the Breusch-Pagan LM test statistic for models with English proficiency of foreign-born working population and with English proficiency of the employed population are 10.73 and 11.04, respectively. Given the critical value of 23.68, the null hypothesis of no unknown heteroskedasticity at the five percent level is rejected. But, the White test could not be run because of an insufficient number of observations. The Park test statistic for pure hetroskedasticity with respect to the percentage of foreign-born workers as a proportionality factor is applied because, as the size of the state increases, the variability of the rate of injury might rise. Using the log of the squared residuals as the dependent variable, the t-value for percentage of foreign-born workers as a regressor for models with English proficiency of foreign-born working population, and with English proficiency of the employed population are 2.17 and 1.52, respectively. Given the outcome of the heteroskedasticity tests, to ensure robust standard error, I conclude that, for consistency, the non-linear weighted least squares method, using the percent of foreign-born workers as the weight, with the White heteroskedasticity-corrected standard errors is the proper choice for the logit models. Of course, foreign born is not used as a weight when it is introduced as an explanatory variable.

When the Poisson model is subjected to heteroskedasticity tests, the LM test statistic is 10.32 and 9.46 for the two alternative measures of English proficiency. Given the critical value of 23.68 at the five percent level, the Breusch-Pagan test rejects the null hypothesis of no unknown hetroskedasity. Again, the White test could not be run because of insufficient number of observations. Therefore, the hypothesis of no unknown heteroskedasticity could not be rejected for the Poisson model either. The results from the Park test for the employment proportionality factor confirm the presence of heteroskedasticity at the five percent level—2.15 and 2.62 for models with English proficiency of foreign-born working age population, and with English proficiency of the employed population, respectively. Thus, the Poisson models are also estimated with employment as weight.

Wage rate is a suspect to be endogenous because workers are expected to receive compensating wages for taking high risk jobs, which can potentially create a bias in the estimated parameters. Lanoie (1992-b) and Black and Kniesner (2003), faced with the same problem, use simultaneous models in their analysis. Ruser (1993), on the other hand, uses a single equation in his study of occupational injuries because he does not find any instrumental variable which is correlated with the wage rate, but not with the incident of work injury. For the same reason, I proceed with estimation of single equation models for both the logit and the Poisson models.

## IV. Results

Because of the continuity of the logit transformation of the rate of injury, the least squares method is used for estimation of the parameter of the logit model. Table 2 presents the results for various specifications of the model. It is notable that the estimated coefficients of the weighted least squares regression here do not measure the marginal effects, or the change in the probability of work injury. They rather reflect the effect of a change in an independent variable on $ln(\frac{P_s}{(1-P_s)})$. Also note that since relation between $P_s$ and its logit transformation is inverse, for convenience of interpretation of the results, the reported sign of the estimated parameters are reversed in Table 2. The adjusted $R^2$ indicates that at least seventy five percent of the variations in injury rates across states are explained by the explanatory variables in the reported models. The first two models include only one of the measures of English proficiency. In model 3, language proficiency measures are replaced with foreign-born workers. Then, the last two models include both an English proficiency variable and foreign-born workers. The coefficients of most of the control variables are statistically significant with correct signs. But, the focus of this study is on the English proficiency of the labor. Estimated parameters of English proficiency in models 1 and 2 are in support of the hypothesis that inadequate language proficiency of foreign-born workers increases the likelihood of work injury. Model 3 shows that increases in the percentage of foreign-born population raise the rate of work injury in the states, be it because of English language incompetence or a cultural gap. Including foreign-born workers and one of the measures of English proficiency in models 4, and 5, reduces both the size of the coefficient of English proficiency and its statistical significance. In model 5, even the sign of English proficiency coefficient is reverse, though it is not statistically significant.

Table 3 reports the initial weighted maximum likelihood estimates of the parameters of the Poisson models. Specifications of the alternative models correspond with the estimates of the logit models in Table 2. In the first two models where the percentage of foreign-born workers is absent, the expected adverse effect of the lack of English proficiency is observed. Parameter estimates in model 3 show that the number of injuries rises in states where the percentage of foreign-born workers is high, signifying the language proficiency issue or the cultural gap. In the next two models, foreign-born workers effect is sustained, while the impact of language proficiency is reversed. Clearly the results for foreign-born workers are more robust than the language proficiency variables' results. In other words, it appears that foreign-born workers variable explains the adverse effect of changes in the composition of immigrants on work injury better than the language proficiency of foreign born. In fact, when a similar issue is addressed regarding productivity of foreign-born teaching assistants, Marvasti (2007) does not find any evidence of an adverse effect of poor communication skills of foreign-born teaching assistants on the students' grades, while the presence of foreign born, apparently due to cultural gap, lowers students' grades.

Parameter estimates of the modified weighted Poisson models are reported in Table 4.[7] The first group of models are based on *normal heterogeneity*. The coefficients of both measures of English proficiency and foreign born are smaller than the estimates from more simple weighted Poisson models in Table 3. Parameter estimates of other explanatory variables are mostly stable. The last two models are based on exogenous underreporting with the foreign-born workers as the underreporting factor. The estimated coefficients for the measures of English proficiency show that taking underreporting into account has reduced the size of the coefficient for English proficiency of the foreign-born working age population, but has increased the coefficient for English proficiency of the employed population.

Of course, the coefficients here are not the marginal effects. Marginal effects of the three variables of interest in this study from selected logit and Poisson estimates are presented in Table 5. The reported marginal effects are based on the sample mean of the population. Note that although the value of the coefficient influences the calculation of the marginal effect, the sign and the magnitude of the marginal effect need not bear any relationship to the coefficient (Greene, 2003). In the logit model, a one percent increase in the percentage of working age foreign-born population who do not speak English well or at all raises the rate of injuries in a state by 0.00045. Similarly, a small increase in the rate of injuries is expected as a result of an increase in the percentage of employed population of working age who do not speak English well or at all, or as a result of a one percent increase in the foreign-born workers population-- 0.00034 and 0.00033, respectively. In the basic Poisson models, a one percent increase in the percentage of foreign born of working age who do not speak English well or at all raises the number of annual work injuries in a state by approximately 35, while a one percent increase in the percentage of employed population of working age who do not speak English well or at all raises the number of annual work injuries in a state by approximately 24. The marginal effect of a one percent increase in foreign-born workers in the state is much higher, leading to an increase of approximately 91 cases of annual work injuries. In the weighted Poisson models with normal heterogeneity, the marginal effects are larger for the measures of English proficiency as well as for foreign-born workers than the marginal effects in the basic Poisson model. After the possibility of underreporting, because of the presence of foreign-born workers, is taken into account, the marginal effects for both measures of English proficiency drop substantially.

**V. Summary and Conclusions**

To explore the impact of deterioration in English proficiency of immigrants into the U.S. on the incident of work injury at the state level, two measures of English proficiency are selected-- one based on English proficiency of working age foreign born, and the other based on English proficiency of employed population. Also, the cultural gap hypothesis is tested by examining the effect of rising foreign-born population on the incident of work injury. Two available measures of work injury incidents are employed as alternative dependent variable. The empirical model for the rate of injury is based on the application of weighted least squares regression to the logistic transformation of the rate of injuries. For the count of injury incidents, after ruling out the possibility of overdispersion, the weighted Poisson regression model is employed. Then, the weighted Poisson model is modified by removing the assumption of no unobserved heterogeneity. Finally, the possibility of underreporting of work injury incidents is considered, using foreign-born workers as the underreporting factor.

Parameter estimates from the logistic and basic Poisson models, in the absence of foreign-born workers in the equation, confirm the adverse effect of inadequate English proficiency of foreign born on the incidence of work injury. However, when foreign-born wokers is added to the model as an explanatory variable, the effect of English proficiency either disappears or is reversed. The results for the foreign-born workers factor itself are more robust suggesting that perhaps cultural factors play a more important role in the incident of work injury than language proficiency. However, the marginal effects from the Poisson models show a stronger effect from changes in proportion of foreign born than from changes in language proficiency of the work force. The rising proportion of Hispanics in the population is generally suspected to have contributed to the rising work injuries. Correcting for the high risk industries, the results from the logistic models do not lend much support for this hypothesis. Estimates from the Poisson models, on the other hand, indicate that the increase in the percentage of Hispanics in the population is a factor in the rising incident of work injury. Though, when foreign-born workers variable is added to the model, the effect of Hispanic is reversed. It appears that foreign born masks the effect of both language proficiency and Hispanics.

While examination of rate of occupational injury has found some evidence supporting the argument that the characteristics of the new cohort of immigrants, be it English proficiency or cultural gap, play a role in rising incident of work injury, the high degree of aggregation in the data is likely to have prevented exposing a more clear relationship between the incident of work injury and English proficiency. Industry level analysis of work injury incidence may generate more robust results. Nevertheless, this study provides some evidence in support of English language programs targeting the foreign-born segment of the work force in order to reduce work injuries, which are costly in terms of loss of workers' productivity and increases in health care costs.

**Endnotes:**

1. English proficiency is used in a study by Chowdhury and Peduce (2007), which deals with ethnic enclaves and labor market. The authors find a negative effect of ethnic concentration on earning in CA, but the effect of English proficiency on immigrants' income varies by type of ethnicity. Presence of a large foreign-born population may have reduced the negative effect of English proficiency because of the larger number of foreign-born businesses in California. When English proficiency of the foreign born improves, they are likely to move out of ethnic enclaves. Edin (2003), on the other hand, finds a positive, 'warm support,' effect of ethnic enclaves and wages in Sweden.

2. States that have decided not to participate in the survey include Colorado, Idaho, Mississippi, New Hampshire, North Dakota, Ohio, Pennsylvania, and South Dakota. For more information on the history of BLS safety and health statistics programs, including the Survey of Occupational Injuries and Illnesses, visit www.bls.gov/iif/oshhist.htm.

3. While the CB data on language proficiency and the BLS injury data overlap for a few years, because of lack of fluctuations in the series from year to year the empirical analysis here is limited to the 2006 data.

4. The selected high risk industries include non-metallic mineral mining and quarrying, taxi and limousine service, metal ore mining, truck transportation, non-specified type of mining, sawmills and wood preservation, non-specified food industries, leather tanning and products except footwear, sewage treatment facilities, and automobile dealers.

5. Percentage of wages was considered as an alternative to the maximum weekly workers' compensation by the state. Since the rate is rather constant across states-- typically 66.67% of the gross earning or 80% of the after tax earning-- the maximum compensation was selected instead.

6. Because of lack of data, the effect of age could not be controlled by experience here.

7. Convergence of the modified Poisson models in LIMDEP occurs only when the Newton algorithm is used.

**References:**

Adnett, Nick, and Dawson, Alistair. 1998. The Economic Analysis of Industrial Accidents: A Re-Assessment, *International Review of Applied Economics*, Vol. 12, No. 2, 241-256.

Black Dan A., and Kniesner Thomas J. 2003. On the Measurement of Job Risk in Hedonic Wage Models, *The Journal of Risk and Uncertainty*, Vol. 27, No. 3, 205-220.

Cameron, A. Colin, and Pravin K. Trivedi. 1990. Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests, *Journal of Applied Econometrics*, Vol. 1, No. 1, 29-53.

Chowdhury, Mussaddeq, and Pedace, Roberto. 2007. Ethnic Enclaves and Labor Markets: An Analysis of Immigrant Outcomes in California, *Contemporary Economic Policy*, Vol. 25, No. 2, 238-249.

Decker, Christopher S., and David T. Flynn. 2008. Work-Related Accidents and the Level of Market Competition: An Analysis of Worker Injury Rates at U.S. Steel Corporation, 1907-1939, *Economic Inquiry*, Vol. 46, No. 3, 438-53.

Dillingham, Allan E. 1983 The Effect of Labor Force Age Distribution on Workers' Compensation Cost, *Journal of Risk and Insurance*, Vol. 50, No.2, 235-49.

Edin, P., P. Fredriksson, and O. Aslund. 2003. Ethnic Enclaves and the Economic Success of Immigrants-- Evidence from a Natural Experiment, *Quarterly Journal of Economics*, Vol. 118, No.1, 329-57.

Gibson, Campbell J., and Jung, Kay. 2006. Historical Census Statistics on the Foreign-Born Population of the United States: 1850-2000. U.S. Census Bureau, Working Paper No. 81.

Greene, William H. 2003. *Econometric Analysis*, Fifth Ed., Upper Saddle River, N.J., Prentice Hall.

Lanoie, Paul. 1992-a. Safety Regulation and the Risk of Workplace Accidents in Quebec. *Southern Economic Journal* 58 (4): 950-65.

Lanoie, Paul. 1992-b. The Impact of Occupational Safety and Health Regulation on the Risk of Workplace Accidents, *Journal of Human Resources*, Vol. 27, No. 4, 643-660.

Kahn, Shulamit. 1987. Occupational Safety and Worker Preferences: Is There a Marginal Worker? *Review of Economics and Statistics*, Vol. 69, No. 2, 262-68.

Marvasti, A. 2007 Foreign-Born Teaching Assistants and Student Achievement: An Ordered Probit Analysis, *The American Economist*, Vol. 51, No. 2, 61-71.

OI, Walter Y. 1974. On the Economics of Industrial Safety, *Law and Contemporary Problems*, Vol. 38, No. 4, 669-698

Ruser, John W. 1985. Workers' Compensation Insurance, Experience-Rating, and Occupational Injuries, *Rand Journal of Economics*, Vol. 16, No. 4, 487-503.

Ruser, John W. 1993. Workers' Compensation and the Distribution of Occupational Injuries, *The Journal of Human Resources*, Vol. 28, No. 3, 594-617.

Ruser, John W. 1995. Self-Correction Versus Persistence of Establishment Injury Rates, *The Journal of Risk and Insurance*, Vol. 62, No. 1, 76-92.

Wiatrowski, William J. 2005. Occupational Safety and Health Statistics: New Data for a New Century, *Monthly Labor Review*, Vol. 128, No. 10, 3-10.

Winkelmann, R. and K. F. Zimmermann. 1994. Count Data Models for Demographic Data, Working Paper

Winkelmann, R. 2003. *Econometric Analysis of Count Data*, Fourth Edition, New York: Springer.

**Table 1. Descriptive Statistics**

| Variables | Included States and D.C.<br>Mean<br>(Standard Dev.) | Non-participating States<br>Mean<br>(Standard Dev.) |
|---|---|---|
| Rate of injuries and work-related illnesses per 100 full-time workers in private industry-- total recordable cases ($P_s$) | 0.047<br>(0.02) | - |
| $Y_s$= Log ($P_s$/(1- $P_s$ )) | -3.03<br>(0.25) | - |
| Number of injuries and work-related illnesses in private industry-- total recordable cases (in hundred) | 828.51<br>(844.14) | - |
| Percentage of foreign-born working population age 18-64 who do not speak English well or at all | 25.87<br>(10.08) | 20.63<br>(11.54) |
| Percentage of employed population age 18-64 who do not speak English well or at all | 3.19<br>(2.49) | 1.44<br>(1.24) |
| Percentage of foreign-born workers | 11.36<br>(7.94) | 5.44<br>(3.13) |
| Mean age of working population | 38.89<br>(2.05) | 37.98<br>(2.13) |
| Percentage with a bachelor degree | 11.90<br>(2.06) | 12.19<br>(2.42) |
| Percentage black | 9.92<br>(9.46) | 7.05<br>(10.63) |
| Percentage male | 49.23<br>(0.77) | 49.47<br>(0.79) |
| Percentage Hispanic | 28.69<br>(14.39) | 29.34<br>(16.29) |
| Percentage married | 40.61<br>(3.54) | 42.01<br>(2.63) |
| Percentage in the labor union | 11.58<br>(5.70) | 8.70<br>(3.50) |
| Unemployment rate | 4.46<br>(0.97) | 4.30<br>(1.28) |
| Maximum weekly workers' compensation by the state law (in hundreds) | $7.19<br>(2.12) | $6.75<br>(2.28) |
| Percentage of high-risk industry in the state (top 10 high risk industries) | 2.61<br>(0.74) | 6.75<br>(2.28) |
| Median wage rate | 18.36<br>(2.79) | 17.25<br>(1.83) |
| Percentage of population who is obese (BMI>30) | 25.09<br>(2.79) | 24.91<br>(3.92) |
| Number of workers employed (in thousands) | 2266.43<br>(2453.35) | - |
| N | 43 | 8 |

**Table 2. Parameter Estimates of the Least Squares Estimates of the Logistic Model***

| Variables | Model 1 (Weighted) Coefficient (St. Error) | Model 2 (Weighted) Coefficient (St. Error) | Model 3 Coefficient (St. Error) | Model 4 Coefficient (St. Error) | Model 5 Coefficient (St. Error) |
|---|---|---|---|---|---|
| English Proficiency of Working- | 0.0101[b] | - | - | 0.0036[c] | - |
| Age Foreign-Born Population | (0.0054) | | | (0.0025) | |
| English Proficiency of Employed | - | 0.0076 | - | - | -0.0006 |
| Population | | (0.0068) | | | (0.0096) |
| Foreign-Born Workers | - | - | 0.0073[b] | 0.0053[c] | 0.0074[b] |
| | | | (0.0037) | (0.0035) | (0.0037) |
| Bachelor Degree | -0.0032 | 0.0158 | -0.0097 | -0.0162 | 0.0104 |
| | (0.0332) | (0.0327) | (0.0189) | (0.0211) | (0.0251) |
| Age | -0.6752 | -0.5884 | -0.4031 | -0.4419 | -0.4133 |
| | (0.6754) | (0.7379) | (0.3971) | (0.4029) | (0.3653) |
| Age squared | 0.0093 | 0.0081 | 0.0051 | 0.0057 | 0.0052 |
| | (0.0087) | (0.0095) | (0.0049) | (0.0050) | (0.0045) |
| Male | -0.1074[a] | -0.1065[a] | -0.0879[a] | -0.0901[a] | -0.0885[b] |
| | (0.0354) | (0.0405) | (0.0330) | (0.0318) | (0.0373) |
| Hispanic | 0.0016 | 0.0037[c] | 0.0017 | 0.0013 | 0.0016 |
| | (0.0025) | (0.0027) | (0.0020) | (0.0020) | (0.0021) |
| Black | 0.0015 | -0.0071[c] | 0.0071[b] | 0.0082[b] | 0.0072[b] |
| | (0.0058) | (0.0051) | (0.0045) | (0.0045) | (0.0045) |
| Married | 0.0153[c] | -0.0347[a] | -0.0181[b] | 0.0143[b] | -0.0178[c] |
| | (0.0110) | (0.0120) | (0.0093) | (0.0088) | (0.0127) |
| Labor Union | -0.0066 | -0.0106[b] | -0.0053 | -0.0045 | -0.0053[b] |
| | (0.0058) | (0.0059) | (0.0062) | (0.0064) | (0.0063) |
| Unemployment Rate | -0.0838[b] | -0.0924[b] | -0.0495[b] | -0.0543[b] | -0.0488[b] |
| | (0.0401) | (0.0452) | (0.0275) | (0.0295) | (0.0290) |
| Maximum Workers' Compensation | -0.0315[a] | -0.0351[a] | -0.0205[b] | -0.0208[b] | -0.0205[b] |
| | (0.0114) | (0.0133) | (0.0107) | (0.0102) | (0.0109) |
| Wage Rate | 0.0928[a] | 0.0829[a] | 0.0459[b] | 0.0568[a] | 0.0461[a] |
| | (0.0312) | (0.0290) | (0.0197) | (0.0213) | (0.0206) |
| Obesity (BMI) | 0.0291[c] | 0.0489[b] | 0.0118 | 0.0090[c] | 0.0111[b] |
| | (0.0206) | (0.0265) | (0.0111) | (0.0112) | (0.0165) |
| Percentage of High-Risk Industry | 0.1317[a] | 0.1558[a] | 0.0759[a] | 0.0860[a] | 0.0754[b] |
| | (0.0529) | (0.0602) | (0.0309) | (0.0333) | (0.0344) |
| Intercept | 18.7470[c] | 17.6715 | 15.0242[b] | 15.4176[b] | 15.2620[b] |
| | (13.0998) | (14.8568) | (7.8145) | (7.8430) | (7.4632) |
| Adj. $R^2$ | 0.84 | 0.82 | 0.76 | 0.76 | 0.75 |
| SSE | 0.3796 | 0.4235 | 0.4054 | 0.3937 | 0.4053 |

* White heteroskedasticity-Consistent Standard Errors.
[a] Significant at the 1% level.
[b] Significant at the 5% level.
[c] Significant at the 10% level.

**Table 3. Parameter Estimates of the Weighted Poisson Model**

| Variables | Model 1 Coefficient (St. Error) | Model 2 Coefficient (St. Error) | Model 3 Coefficient (St. Error) | Model 4 Coefficient (St. Error) | Model 5 Coefficient (St. Error) |
|---|---|---|---|---|---|
| English Proficiency of Working-Age Foreign-Born Population | 0.0389[a] | - | - | -0.0123[a] | - |
| | (0.0010) | | | (0.0013) | |
| English Proficiency of Employed Population | - | 0.02674[a] | - | - | -0.0457[a] |
| | | (0.0028) | | | (0.0031) |
| Foreign-Born Workers | - | - | 0.1070[a] | 0.1126[a] | 0.0199[a] |
| | | | (0.0012) | (0.0013) | (0.0012) |
| Bachelor Degree | 0.0555[a] | 0.1667[a] | 0.0548[a] | 0.0786[a] | 0.0247[a] |
| | (0.0823) | (0.0081) | (0.0082) | (0.0086) | (0.0085) |
| Age | 4.4150[a] | 5.4794[a] | 3.2863[a] | 3.4452[a] | 2.6076[a] |
| | (0.1484) | (0.1477) | (0.1578) | (0.1594) | (0.1660) |
| Age squared | -0.0536[a] | -0.0681[a] | -0.0421[a] | -0.0446[a] | -0.0338[a] |
| | (0.0019) | (0.0019) | (0.0020) | (0.0020) | (0.0021) |
| Male | -0.0727 | 0.0193[a] | 0.2653[a] | -0.2695 | -0.2985[a] |
| | (0.0109) | (0.0107) | (0.0129) | (0.0129) | (0.0130) |
| Hispanic | 0.0264[a] | 0.0350[a] | -0.0003[b] | -0.0007[a] | -0.0005[b] |
| | (0.0007) | (0.0006) | (0.0008) | (0.0008) | (0.0008) |
| Black | -0.0246[a] | -0.0407[a] | -0.0263[a] | -0.0302[a] | -0.0161[a] |
| | (0.0014) | (0.0015) | (0.0014) | (0.0014) | (0.0015) |
| Married | 0.0619[a] | -0.0050 | 0.0859[a] | 0.0701[a] | 0.1148[a] |
| | (0.0041) | (0.0040) | (0.0046) | (0.0049) | (0.0051) |
| Labor Union | -0.0151[a] | -0.0222[a] | -0.0340[a] | -0.0357[a] | -0.0262[a] |
| | (0.0017) | (0.0017) | (0.0017) | (0.0017) | (0.0018) |
| Unemployment Rate | -0.1249[a] | -0.1118[a] | -0.0924[a] | -0.1056[a] | 0.1486[a] |
| | (0.0084) | (0.0091) | (0.0087) | (0.0088) | (0.0096) |
| Maximum Workers' Compensation | 0.0024 | -0.0094[c] | 0.0219[a] | 0.0202[a] | 0.0152[a] |
| | (0.0030) | (0.0031) | (0.0031) | (0.0031) | (0.0031) |
| Wage Rate | 0.2859[a] | 0.1801[a] | 0.0397[a] | 0.0034 | 0.0417[a] |
| | (0.0085) | (0.0078) | (0.0086) | (0.0095) | (0.0086) |
| Obesity (BMI) | 0.2067[a] | 0.2470[a] | 0.2303[a] | 0.2366[a] | 0.1872[a] |
| | (0.0051) | (0.0059) | (0.0055) | (0.0056) | (0.0063) |
| Percentage of High-Risk Industry | 0.3693[a] | 0.3302[a] | 0.1348[a] | 0.1260[a] | 0.1407[a] |
| | (0.0128) | (0.0126) | (0.131) | (0.0132) | (0.0133) |
| Intercept | -95.6054[a] | -116.2667[a] | -56.4967[a] | -57.5831[a] | -41.1498[a] |
| | (2.9437) | (2.9855) | (3.0799) | (3.0894) | (3.2680) |
| Log Likelihood | -7978.03 | -8618.65 | -3593.09 | -3550.22 | -3483.83 |
| Overdispersion Test: | | | | | |
| α | 0.007[b] | 0.007[b] | 0.007[b] | 0.004[b] | 0.007[b] |
| | (0.003) | (0.003) | (0.003) | (0.001) | (0.003) |

[a] Significant at the 1% level.
[b] Significant at the 5% level.
[c] Significant at the 5% level.

**Table 4. Parameter Estimates of the Weighted Poisson Model with Modifications**

| | Normal Heterogeneity | | | Exogenous Underreporting | |
|---|---|---|---|---|---|
| Variables | Model 1 Coefficient (St. Error) | Model 2 Coefficient (St. Error) | Model 3 Coefficient (St. Error) | Model 1 Coefficient (St. Error) | Model 2 Coefficient (St. Error) |
| English Proficiency of Working-Age Foreign-Born Population | 0.0455[a] | - | - | 0.0328[a] | - |
| | (0.0011) | | | (0.0001) | |
| English Proficiency of Employed Population | - | 0.0336[a] | - | - | 0.0306[a] |
| | | (0.0030) | | | (0.0003) |
| Foreign-Born Workers | - | - | 0.1291[a] | - | - |
| | | | (0.0019) | | |
| Bachelor Degree | 0.0707[a] | 0.1892[a] | 0.0693[a] | 0.1214[a] | 0.2511[a] |
| | (0.0099) | (0.0075) | (0.1132) | (0.0007) | (0.0007) |
| Age | 4.8874[a] | 6.0316[a] | 1.614[a] | 3.9886[a] | 5.415[a] |
| | (0.1623) | (0.1522) | (0.1858) | (0.0200) | (0.0221) |
| Age squared | -0.0593[a] | -0.0749[a] | -0.0207[a] | -0.0484[a] | -0.0672[a] |
| | (0.0021) | (0.0020) | (0.0024) | (0.0003) | (0.0003) |
| Male | -0.0090 | 0.0259[a] | 0.3391[a] | -0.0652[a] | 0.0315[a] |
| | (0.0093) | (0.0115) | (0.0211) | (0.0009) | (0.0010) |
| Hispanic | 0.0226[a] | 0.0361[a] | -0.0126[b] | 0.0225[a] | 0.0288[a] |
| | (0.0005) | (0.0005) | (0.0011) | (0.0001) | (0.0001) |
| Black | -0.0172[a] | -0.0454[a] | -0.0313[a] | -0.0253[a] | -0.0430[a] |
| | (0.0014) | (0.0015) | (0.0018) | (0.0002) | (0.0002) |
| Married | 0.0700[a] | 0.0217[a] | 0.0527[a] | 0.0461[a] | -0.0126[a] |
| | (0.0043) | (0.0034) | (0.0056) | (0.0005) | (0.0005) |
| Labor Union | -0.0131[a] | -0.0342[a] | -0.0184[a] | -0.0112[a] | -0.0215[a] |
| | (0.0016) | (0.0014) | (0.0027) | (0.0001) | (0.0002) |
| Unemployment Rate | -0.1651[a] | -0.0546[a] | -0.0318[a] | -0.0806[a] | -0.0823[a] |
| | (0.0074) | (0.0074) | (0.0129) | (0.0008) | (0.0010) |
| Maximum Workers' Compensation | -0.0050[b] | -0.0123[a] | 0.0276 | -0.0043[a] | -0.0196[a] |
| | (0.0028) | (0.0020) | (0.0036) | (0.0002) | (0.0002) |
| Wage Rate | 0.2905[a] | 0.1953[a] | -0.0186[a] | 0.1920[a] | 0.0945[a] |
| | (0.0090) | (0.0083) | (0.0129) | (0.0009) | (0.0009) |
| Obesity (BMI) | 0.2108[a] | 0.2434[a] | 0.2547[a] | 0.2067[a] | 0.4543[a] |
| | (0.0052) | (0.0062) | (0.0072) | (0.0006) | (0.0008) |
| Percentage of High-Risk Industry | 0.3875[a] | 0.3231[a] | 0.0602[a] | 0.3262[a] | 0.2735[a] |
| | (0.0106) | (0.0166) | (0.0138) | (0.0010) | (0.0010) |
| Intercept | -109.2265[a] | -126.8440[a] | -17.0649[a] | -85.4643[a] | -114.5149[a] |
| | (3.2182) | (3.1758) | (3.7221) | (0.3975) | (0.4301) |
| Log Likelihood | -354.71 | -365.50 | -351.06 | -760.69 | -756.74 |
| σ | 0.3524[a] | 0.4009[a] | 0.2611[a] | - | - |
| | (0.0043) | (0.0028) | (0.0039) | | |
| Underreporting Factor: Percentage of Foreign-Born Workers | - | - | - | 0.0095[a] | 0.0068[a] |
| | | | | (0.0001) | (0.0001) |

[a] Significant at the 1% level.
[b] Significant at the 5% level.

**Table 5. Marginal Effects of the Poisson Models***

| Specification | English Proficiency of Working-Age Foreign-Born Population | English Proficiency of Employed Population | Foreign-Born Workers |
|---|---|---|---|
| Logit Model | 0.00045<br>(0.00055) | - | - |
| | - | 0.00034<br>(0.00000) | - |
| | - | - | 0.00033<br>(0.00000) |
| Basic Poisson with Weight | 34.53<br>(1.32) | - | - |
| | - | 23.87<br>(2.91) | - |
| | - | - | 91.10<br>(2.51) |
| Poisson with Weight and Normal Heterogeneity | 42.01<br>(0.95) | - | - |
| | - | 32.23<br>(2.87) | - |
| | - | - | 100.29<br>(1.44) |
| Poisson with Weight and Exogenous Underreporting | 16.63<br>(5.56) | - | - |
| | - | 15.80<br>(4.29) | - |

*Standard errors are in parenthesis.