Multiple Approaches to Analyzing Count Data in Studies of Individual Differences: The Propensity for Type 1 Errors, Illustrated with the Case of Absenteeism Prediction

Michael C, Sturman, Louisiana State University

The present study compares eight models for analyzing count data: ordinary least squares (OLS), OLS with a transformed dependent variable, Tobit, Poisson, overdispersed Poisson, negative binomial, ordinal logistic, and ordinal probit regressions. Simulation reveals the extent that each model produces false positives. Results suggest that, despite methodological expectations, OLS regression does not produce more false positives than expected by chance. The Tobit and Poisson models yield too many false positives. The negative binomial models produce fewer than expected false positives.

A fundamental problem with scientific research is that the way we try to solve a problem affects what kind of results we see (Kuhn, 1970). This is partly reflected in the social sciences because the statistical method used to analyze data affects what kind of relationships we observe (Schmidt, 1996). When the assumptions of the employed statistical model are met, the observed coefficients usually describe the actual relationship well (Greene, 1993). However, when these assumptions are violated, such as when using ordinary least squares (OLS) regression to analyze nonnormal data, resulting estimates may not be meaningful. This can result in true relationships not being discovered (i.e., Type II errors) or the misidentification of nonexistent relationships (i.e., Type I errors). Analyses of count data in studies of individual differences have often been criticized for the use of inappropriate methodologies. An example for human resources is research on absenteeism. Although there are many domains yielding data that may be best quantified as counts, I use the example of predicting absenteeism because this domain lends itself well to the illustration of problems of count data. As for count data in general, the largest problem stems from rates of absenteeism not following a normal distribution (Arbous & Sichel, 1954; Baba, 1990; Hammer & Landau, 1981; Harrison & Hulin, 1989; Landy, Vasey, & Smith, 1984; Martocchio& Harrison, 1993; Mikalachki & Gandz, 1982; Rushmore & Youngblood, 1979; Watson, Driver, &Watson, 1985). This results in a skewed, truncated distribution, which contradicts the assumptions of commonly employed statistical methods such as OLS (Nunnally, 1978). Yet, despite fundamental methodological flaws, correlation and multiple regression analyses dominate the absence research (Baba, 1990; Martocchio & Harrison, 1993).

The methodological issues raised by the characteristics of count data are quite noteworthy, and thus, a researcher reviewing prior empirical work would have difficulty interpreting the validity of published results. One could take an extreme view of the problem and suggest that all prior empirical work is questionable and thus empirical absenteeism research should begin anew. A more conservative approach, and perhaps one more appropriate for research, would entail evaluating the noteworthiness of the methodological problems. This article investigates the extent that various methodologies are prone to Type I errors and intends to provide information on the extent that statistically significant findings from previous analyses can be considered valid.

Methodological Alternatives

Researchers have identified alternatives to help overcome the methodological problems associated with studying individual differences by analyzing count data. These alternatives can be seen as either (a) changing the characteristics of the data to meet the assumptions of traditional statistical methods better or (b) using a statistical method that is more appropriate for the type of data collected. One way to change the data involves changing the level of aggregation. Aggregation allows for a wider distribution of values. This approach ostensibly produces distributions that better meet the assumptions of traditional analyses. Aggregation may also improve the explanatory power of models. Effect sizes are required to be reported by the editorial policies of this journal (Thompson, 1994), and the American Psychological Association (APA, 1994) publication manual “encourages” (p. 18) that effect sizes always be reported; therefore, understanding factors that influence effect size estimates is important. Effect sizes largely depend on whether counts represent single instances, such as being absent on a given day, or an individual’s average level, such as yearly attendance record (Abelson, 1985). In addition, if we assume that there are no differing effects on the dependent variable over time, then correlations at the individual level of analysis should equal correlations computed from the aggregated data (Ostroff, 1993).

However, aggregation still may create other problems that make its use undesirable (Hulin & Rousseau, 1980). For example, aggregation of voluntary absence data from 1 month to 1 year exaggerates the level of skew and kurtosis, and the truncation problem is still obvious (Harrison & Hulin, 1989). Aggregation only reduced the effect of discreteness. Thus, aggregation leaves the researcher facing the same methodological concerns. In addition, aggregation may obscure relationships because of longer cause-effect time gaps and may occlude the effects of some environmental variables (Harrison& Hulin, 1989). For example, if winter weather has a noteworthy effect on an individual's ability to attend work, aggregating over the entire year would obscure the effect that winter storms have on level of absenteeism for employees who live in snowy climates.

An alternative to traditional analysis was demonstrated by Harrison and Hulin (1989), who overcame methodological problems by using an event history model. Event history models describe the states that individuals are in, the time spent in those states, and the rates of movements from state to state. Markov models and hazard-rate models are two examples of popular event history models. Such models overcome many of the estimation problems for count data because "model estimation is built on an emerging theory called quasi-likelihood, which does not require that the data conform to a specified multivariate distribution" (Harrison & Hulin, 1989, p. 315); however, such models entail analyzing large samples over multiple observation periods and, thus, may not be applicable for more typical samples of count data.

Importance of Knowing Type I Error Rates

For research involving the prediction of count data, such as for predicting number of absences, there are two major gaps in the literature. First, when faced with a data set that does not lend itself to sophisticated analysis such as event history models due to limits in sample size or other factors, it is unclear what quantitative modeling strategy is most appropriate. Second, it is unclear how to interpret the results of prior research that has used a variety of modeling strategies but has not been evaluated in terms of its appropriateness for count data. The present study begins bridging these gaps by exploring the extent to which various methodologies are prone to Type I errors when analyzing count data. Although any error in data analysis is undesirable, we focus on Type I errors for three reasons. First, although researchers have argued that Type I errors have been overemphasized and statistical significance tests have inappropriately and all-too-often become the end-all of scientific inquiry (Cohen, 1994; Kirk, 1996; Thompson, 1996), the practice of such statistical testing is still prevalent. An understanding of the quantitative methods that dominate the field is necessary to accurately interpret the results of research performed in this style or, more important, for interpreting prior research. Second, even if research follows the practice of examining point estimates and confidence intervals of estimates in place of statistical significance tests (e.g., Cohen, 1994; Loftus, 1991; Schmidt, 1996; Schmidt & Hunter, 1996), research is still necessary to investigate the extent to which different methodological techniques yield estimates with inappropriate beta values and biased standard errors for these coefficients, and third, although many have argued that Type II errors are often ignored by researchers and have been, in general, underemphasized (e.g., Greenwald, 1975; Rosenthal & Rubin, 1985; Schmidt, 1996), calculation of power for any statistical technique must be based on accurate estimates of the probability of Type I errors. In sum, an understanding of the likelihood of Type I errors is an essential first step to understand the implications of various statistical techniques for analyzing count data. The present study addresses this point by comparing eight different analysis methods: OLS regression, OLS regression with the dependent variable transformed, the Tobit model, Poisson regression, overdispersed Poisson regression, the negative binomial model, ordinal logistic regression, and ordinal probit regression. A simulation is employed to evaluate the models' propensities for Type I errors. This simulation demonstrates the extent to which these models are likely to detect statistically significant relationships when these relationships do not really exist and explores the sensitivity of the models to characteristics of the dependent variable's distribution, characteristics of the independent variables' distributions, and sample size. This article is intended to contribute to the understanding of the implications of various methodological approaches for analyzing count data. Although the example of absences is used throughout to help demonstrate the implications of this piece for a human resource problem, this work has implications beyond this single domain.

Modeling Strategies

OLS Regression

When faced with nonnormal data, we can always ignore the problem and proceed as if the data were normally distributed. Although not recommended for clear statistical reasons (Johnson & Wichern, 1998), this is by far the most common practice. For absenteeism research, this is illustrated by the preponderance of OLS regression (Baba, 1990; Harrison & Hulin, 1989; Martocchio & Harrison, 1993). One implication of this is that because OLS regression does not account for the data being truncated at zero, it can predict negative values that are clearly meaningless. In addition, the validity of hypothesis tests in OLS regression is purported to depend on assumptions that are unlikely to be met in typical count data (Gardner, Mulvey, & Shaw, 1995). As a result, sampling statistics (i.e., mean and variance) may appreciably differ from the true population parameters, which could lead to a loss of power and Type II errors (Hammer & Landau, 1981). Similarly, a skewed distribution can lead to heteroskedasticity, which can severely affect standard errors and lead to Type I or Type II errors (Hammer & Landau, 1981). Because for count data the absolute value of the residuals almost always correlate positively with the predictors, the estimated standard errors of the regression coefficients are smaller than their true value, and thus the *t* values associated with the regression coefficients are likely to be inflated (Gardner et al., 1995). Thus, OLS regression seems prone to Type I errors for analyzing count data.

Although these assertions are methodologically sound and have face validity, the noteworthiness of this problem has not received much empirical attention. There have been some exceptions to this. OLS regression has been compared to the Tobit model (Hammer & Landau, 1981) and the Poisson and negative binomial models (Cameron & Trivedi, 1986; Hausman, Hall, & Griliches, 1984). These comparisons are discussed after the specific models have been introduced.

Data Transformations and OLS Regression

An alternative to ignoring the nonnormality problem is to transform the data to make it "more normal looking." Transformations are nothing more than expressing the data in different units (Johnson & Wichern, 1998); thus, they are commonly recommended to help satisfy statistical assumptions. For count data, the square root is a recommended transformation function (Johnson & Wichern, 1998).Unfortunately, there are a number of potential problems with transforming count data. Most obviously, data transformations do nothing to compensate for the fact that the data are truncated at zero and, thus, negative values can still be predicted (Hammer & Landau, 1981; Harrison & Hulin, 1989). In addition, Harrison and Hulin (1989) assert that "even transformed measures of absenteeism data are often inadmissible as dependent variables in linear regression, because linear regression assumes normality of the marginal distributions of the dependent variable" (p. 300). The modal value of the distribution still falls near or at the bottom of the scale, and although transformations may make the distribution look more normal, heteroskedasticity is still likely to be a problem. This can lead to reduced correlation coefficients and loss of power or inflated *t* values (Gardner et al., 1995; Hammer & Landau, 1981; Harrison & Hulin, 1989). Transforming data also may cause false detection of interactions (Busemeyer & Jones, 1983). Nonetheless, transformation of count data has often been suggested (e.g., Johnson & Wichern, 1998); thus, the appropriateness of the technique merits investigation.

Tobit Model

Although the application of alternatives to OLS is rare, one technique receiving some attention in absenteeism research is the Tobit model (Baba, 1990; Hammer & Landau, 1981). The Tobit model (Tobin, 1958) is a regression model designed to handle truncated data, in which the truncated value occurs with a high probability and the variable is continuously distributed beyond that point (Baba, 1990; Greene, 1993). Tobit models are espoused to provide more consistent, reliable, and less biased estimates than the OLS model (Baba, 1990; Leigh, 1985; Maddala, 1984).

Some research has suggested that the Tobit model is more sensitive than OLS regression (Baba, 1990; Hammer & Landau, 1981). Although this assertion has received some support in the form of more statistically significant coefficients resulting from Tobit models than OLS models (Baba, 1990), this does not necessarily prove the value of the Tobit model. First, the additional statistically significant results obtained using the Tobit model may be false positives (i.e., Type I errors). Second, if we assume that the new statistically significant values from the Tobit model are not Type I errors (or, in other words, the lack of significance from the OLS regression constitutes Type II errors), this implies that findings from previous research relying on OLS regression are still valuable, although some effects will still have been missed in this period research. Indeed, this seems to be what researchers espousing the merits of the Tobit model are implying: They suggest using the Tobit model as a check on an OLS solution (Baba, 1990; Hammer & Landau, 1981). Nonetheless, the need still exists for further comparisons of the Tobit model to OLS regression (Baba, 1990). The Tobit model also should be compared to OLS with transformed data. In addition, because the Tobit model assumes a continuous dependent variable and count data are by definition discrete, the Tobit model should be compared to count models.

Poisson Regression

For data in which the dependent variable is a discrete count, Poisson regression is a natural model choice (Cameron & Trivedi, 1986, 1990; Gurmu, 1991; Hausman et al., 1984; Lee, 1986). Poisson models are particularly attractive for modeling count data because the model has been extended into a regression framework (Lee, 1986), it has a simple structure, and it can be easily estimated (Greene, 1993; Lee, 1986).However, this simplicity is the result of some limiting assumptions, violations of which may have substantial effects on the reliability and efficiency of the model coefficients. The most noteworthy criticism of the Poisson model is its assumption that the variance of the dependent variable equals its mean (Cameron & Trivedi, 1986; Greene, 1993; Lee, 1986). Poisson regression also assumes that each occurrence is independent of the number of previous occurrences and the expected number of occurrences is identical for every member of the sample. Research has addressed some of these limitations by developing tests for overdispersion (e.g., Cameron & Trivedi, 1986, 1990; Gurmu, 1991; Gurmu& Trivedi, 1992; Lee, 1986). Overdispersion occurs if the distribution's variance is greater than the distribution's mean (Greene, 1993). Overdispersion causes the estimates of standard errors to be less than their true value, which leads to inflated *t* coefficients and Type I errors.

The level of overdispersion in part depends on the data aggregation (e.g., by the month, year, etc.). One approach for dealing with overdispersed data is to use a less constrained model (Cameron &

Trivedi, 1986; Gurmu, 1991; Lee, 1986), such as the negative binomial model. Another approach is to correct the *t* values based on an estimate of the dispersion. Despite its limiting assumptions and the availability of alternatives, the implications of Poisson regression for research using count data should be considered.

Overdispersed Poisson Regression Model

If overdispersion exists, one correction method, described by Gardner et al. (1995), is overdispersed Poisson regression. This technique entails estimating a dispersion parameter and using it to modify the *t* tests resulting from a Poisson regression. The overdispersion term is a function of the squared deviation from its expected value (see Gardner et al., 1995, p. 397):

$$\theta = (N - J)^{-1} \{(y_i - \mu[X_i, d_i]^2 / \mu[X_i, d_i]\}$$

where N represents number of cases, J represents number of independent variables, y_i is the observed value, and $\mu[X_i, d_i]$ is the predicted value.

Each squared deviation is divided by that score's variance, assuming that the standard Poisson model were true (Gardner et al., 1995). If the variance equals the mean, then as the sample size approaches infinity, the deviation score approaches one. Because the $\mu[X_i, d_i]$ estimates are chosen to fit the specific sample of y_i, the deviation score sum needs to be adjusted so the dispersion term still approximates 1 if the assumptions of the Poisson model are met with less than infinite sample size (Gardner et al., 1995). The value of ø can be tested to see if the overdispersion is statistically significant (Gardner et al., 1995), but there are also many other methods for testing overdispersion (e.g., Cameron & Trivedi, 1990). Of more immediate use, the value of ø can be used to modify the results from a Poisson regression. If statistically significant overdispersion exists but ø is assumed to equal 1, then the estimated variance of the regression coefficients would be smaller than their true values (Gardner et al., 1995), resulting in inflated *t* tests for the regression coefficients. The Poisson regression results can be corrected by multiplying the Cov (B) by ø; alternatively, the *t* tests computed by the Poisson regression can be divided by the square root of ø (Gardner et al., 1995).

Use of the overdispersed Poisson regression model, as demonstrated by Gardner et al. (1995), resulted in *t*-test values that were dramatically lower than those generated by a Poisson model. The results of the overdispersed Poisson regression model were also similar to those from a negative binomial regression model. Yet, it is still unclear how the overdispersed Poisson regression model would compare to the other models described in this article.

Negative Binomial Model

Although the overdispersed Poisson regression model addresses the assumption of Poisson regression that the mean equals the variance, characteristics of count data may yield further violations of assumptions, which may produce flaws in both the Poisson and overdispersed Poisson models. For example, absenteeism research has shown that past absences are one of the best predictors of future absences (Ivancevich, 1985; Morgan & Herman, 1976; Waters & Roach, 1979), thus calling into question the independence assumption of Poisson models. In addition, theoretical models of absenteeism (e.g., Blau & Boal, 1987; Gibson, 1966; Nicholson, 1977; Rhodes & Steers, 1990; Steers & Rhodes, 1978) commonly suggest that absenteeism is a function of the construct of ability to attend work, which includes illness and accidents, family responsibilities, and transportation problems. These models suggest that individual characteristics cause (or at least correlate with) absenteeism; thus, the expected number of absences for individuals differs. When the mean level of absences is expected to differ across cases, the negative binomial model may be more appropriate (Gardner et al., 1995).

The negative binomial model is one of the more general count models (Cameron & Trivedi, 1986; Gurmu, 1991; Gurmu & Trivedi, 1992; Lee, 1986). In fact, the Poisson model is a special case of the negative binomial model (Cameron & Trivedi, 1986). Negative binomial models can take a number of forms. Commonly, they are categorized through one of two specifications of the dependent variable's variance: $Var(y) = (1 + a)E(y)$ or $Var(y) = E(y)*(1 + a)E(y)$, where a is positive. The former case implies a constant variance to mean ratio; the latter case implies a variance to mean ratio that is linear (Cameron & Trivedi, 1986). Clearly, there are even more possibilities; however, the present study focuses on the first case, in part because it is simpler but primarily because this study has a practical focus and only the former model was included within a commonly available statistics package (e.g., Greene, 1992). As noted earlier, there is a dearth of literature comparing the effects of different types of modeling methods. However, there are also some exceptions to this for the negative binomial model. Cameron and Trivedi (1986) compared the methodological implications and results of analyzing count data for a number of models, including OLS, Poisson, and negative binomial models. Out of 12 independent variables at an alpha level of .05, the OLS model revealed four statistically significant coefficients. The Poisson model revealed the same four statistically significant coefficients in addition to five others; the negative binomial model revealed the same four as the OLS model, two statistically significant coefficients also found by the Poisson model, and one statistically significant coefficient that was not detected by either the OLS or Poisson regression models. The authors concluded by recommending a sequential modeling strategy in which one begins with the basic Poisson model and proceeds to increasingly flexible and data-coherent models.

A similar illustration can be found in an article that focused on developing and adapting models of counts for panel data (Hausman et al., 1984). Although comparing model types was not the focus of the study, that piece did show results for OLS, Poisson, and negative binomial models. In one instance, the OLS model showed two of seven coefficients to be statistically significant. The Poisson model showed five of the seven to be statistically significant, and the negative binomial model showed three of the seven to be statistically significant. In a second instance, the OLS model showed one of two coefficients to be statistically significant, whereas both the Poisson and negative binomial models showed both coefficients to be statistically significant. Although Hausman et al. (1984) do not specifically discuss the importance of these differences, they do point out that the negative binomial model better represents the data as illustrated by the log-likelihood values.

Both of these studies are useful demonstrations of alternative modeling strategies. Unfortunately, they do not shed much light on the meaning of statistically significant coefficients or the implications of the various approaches. A more explicit comparison of modeling techniques to demonstrate the implications of various models for research using count data is needed.

Ordinal Logit and Ordinal Probit

Another way to model count data is to take advantage of its ordered nature. The ordered logit and probit models have come into fairly wide use as a means for analyzing discrete, ordered data (Zavoina & McElvey, 1975). The models are built around a regression framework and, thus, are relatively easy to estimate. In addition, the binomial probit and binomial logit models can be seen as special cases of their corresponding ordinal models. The differences between the ordinal logit and probit models are similar to the differences between their binomial counterparts. The underlying distributions of both models are similar in the middle of the distribution, but the logit model is considerably heavier at the tails (Greene, 1993). Although some researchers have discussed justification of one model over another (Amemiya, 1981), the issue remains unresolved.

A disadvantage of these models is that interpretation of their coefficients is highly complex (Greene, 1993). Furthermore, although count data are ordered, ordinal analysis may not be meaningful. For example, the difference between four absences and three absences (i.e., 1 day) is equal to the difference between six absences and five absences. In addition, ordinal models require that there are instances of each level. So, if modeling yearly absences, which may range from 0 to 40, if no one was absent, say, 25 times, the ordinal model cannot be estimated. This can be corrected by either truncating the distribution or renaming the instances of absences greater than 25 to be one less. The first remedy opens an entirely different issue about the effect of truncation on models of counts, an issue that is not addressed by the present article. The second remedy also works but would make interpretation of the coefficients even more confusing and would make predictions of future levels of the dependent variable confusing and perhaps impossible. In sum, ordinal models may not always be applicable to count data but nevertheless can be useful in certain circumstances. The author found no comparisons of ordinal models to other models for count data for more than the bivariate case.

Simulation Tests

The studies that have compared various methodologies have all revealed that different models vary somewhat in terms of what coefficients are identified as statistically significant (Baba, 1990; Cameron & Trivedi, 1986; Hausman et al., 1984). However, simply the existence of more statistically significant coefficients does not necessary mean a model is better. Indeed, this discrepancy harks back to the issue of Type I versus Type II errors. For example, in a comparison of two models, one of which shows one more statistically significant coefficient, either one model is exhibiting a Type I error or the other model is exhibiting a Type II error. The stress in social science research has generally been on minimizing the chance of Type I errors. In addition, accurate estimates of the likelihood of Type I errors are necessary to compute the level of power in a study. Therefore, the present article concentrates on evaluating the extent to which these various models are prone to Type I errors; in other words, it compares how likely these models are to yield false positives.

To compare the sensitivity of the above models (OLS, transformed dependent variable OLS, Tobit, Poisson, overdispersed Poisson, negative binomial, ordered logit, and ordered probit), a simulation was employed. Specifically, simulation was used to count the number of times that each model incorrectly identified a statistically significant relationship.

The dependent variable in the simulations was a randomly generated count. Because the way the dependent variable is generated may affect the results, we do not rely on any common mathematical distribution (e.g., a chi-square, squared normal) or published theoretical model to generate the variable. Rather, to illustrate a realistic case, and thus continuing the example of predicting absenteeism, this piece used distributions reported in previous articles on absenteeism (Hammer & Landau, 1981; Harrison & Hulin, 1989; Rushmore & Youngblood, 1979). The five distributions that were used include a variety of aggregations (monthly to 3 years) and had varying degrees of skewness and kurtosis. The simulation employed distributions reported in existing absenteeism research to ensure generalizability. Because these distributions are based on estimates from published figures rather than from the original data, the distributions in this piece may not exactly match those on which they are based. Furthermore, the distributions may have been changed slightly to facilitate simulation, such as truncating the distribution after the probability of individual levels of absences dropped below 1%. The distributions used in this article and their summary statistics are shown in Figure 1.

The simulation estimates the likelihood of observing Type I errors attributable to the effects of (a) model type, (b) distribution of the dependent variable, (c) sample size, and (d) distribution of the independent variables. To determine if sample size influences the sensitivity of these models to errors, simulations were run with samples of 100 and 1,000. Because research on counts may employ lagged variables, such as using past absences to predict future absences (e.g., Clegg, 1983; Ivancevich, 1985), some simulations are run with the independent variable following the same distribution as the dependent variable, while others were run with normally distributed independent variables. This allows the simulation to determine if the distribution of the independent variable affects the likelihood of Type I errors.

The simulation variables were completely crossed, which permitted the assessment of the independent effects of each factor on the likelihood of Type I errors. Crossing the factors yields 20 simulations scenarios (5 distributions of the dependent variable × 2 sample sizes × 2 shapes of the distribution for the independent variables). For each simulation, the dependent variable was generated from one of the distributions shown in Figure 1. Each model had 10 independent variables drawn either from a normal distribution ($M = 0$, $SD = 1$) or the same distribution as the dependent variable. Either 100 or 1,000 cases are randomly generated, with each variable being generated independently. Thus, there is no a priori relationship between the random independent variables and the dependent variable. For each scenario, 50 simulations were performed (for a total of 1,000 simulations). For each set of generated data, each of the eight were used to determine the relationship between the independent variables and the dependent variable. Any statistically significant findings are spurious relationships and, thus, are Type I errors.

The number of Type I errors was recorded at the $p < .10$, $p < .05$, and $p < .01$ levels. If more than the expected number of Type I errors occur, such as more than 10% at alpha equal to or less than .10, then we can conclude that the model is prone to Type I errors. Although not a direct test of Type II errors, if fewer coefficients than expected are statistically significant, the *t* tests may be too conservative and the models may be prone to Type II errors.

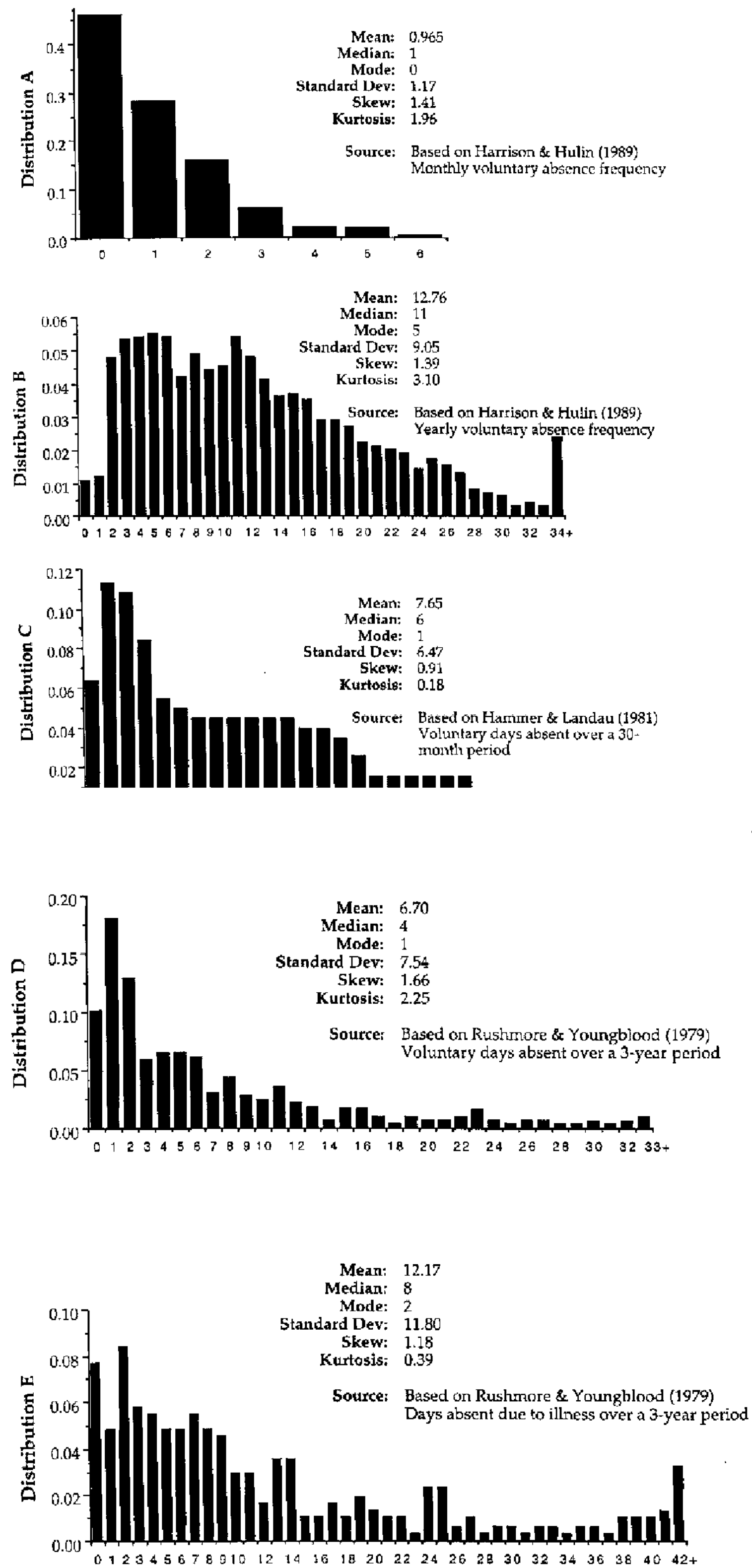

Figure 1. Distribution of absenteeism variables and summary statistics.

Results

The frequencies of Type I errors by model are shown in Table 1. For a number of models, the number of Type I errors appreciably differed from the expected level. Most obviously, the Poisson model incorrectly identified far more statistically significant relationships than expected given the alpha levels (i.e., 0.465, 0.391, and 0.270). The Tobit model also had more false positives than would be expected. On the other hand, some models identified fewer false positives than the alpha levels would suggest. The ordered logit models had fewer Type I errors at alpha = 0.10 (0.088) and the ordered probit model had fewer Type I errors at the alpha = 0.10 (0.087) and alpha = 0.05 (0.043) levels. The negative binomial model had fewer false positives at all three alpha levels. Of note in these results is that OLS regression yielded the number of Type I errors expected by chance. This implies that previous results using OLS regression are valid or at least are not Type I errors.

MANOVA was used to determine if the characteristics of the analysis affect the number of Type I errors. The dependent variables were the number of false positives at all three alpha levels. Simulation characteristics and first order interactions were the independent variables. All of the independent variables were treated as categorical. Results of the MANOVA are shown in Table 2.

Results suggest that the model type, distribution of the dependent variable, sample size, and distribution of the independent variables all affect the likelihood of Type I errors. In addition, this likelihood depends on the interactions between the distribution of the dependent variable and the model type, the sample size, and the distribution of the independent variables. The effect of sample size also partially depends on the distribution of the independent variables.

Conclusion

The results of the present study provide some insights that may benefit future research analyzing count data. The simulation results suggest that OLS regression is not overly sensitive to false positives. Thus, this study suggests, for example, that researchers of absenteeism do not need to ignore the statistically significant findings from previous analyses. Another implication of this study is that researchers should not use Poisson regression to analyze count data. Indeed, these results indicate that the Poisson regression can lead to a large number of false positives. Similarly, despite previous results suggesting that the Tobit model is more sensitive than OLS (Baba, 1990; Hammer & Landau, 1981), these results suggest this “sensitivity” may be Type I errors. Although the Tobit model is not nearly as prone to false positives as the Poisson model, these results suggest that better methods exist to analyze count data.

Table 1

Type I Errors by Model

Model	Number of Simulated Independent Variables	Mean Frequency of Type I Errors at $p < .10$		at $p < .05$		at $p < .01$	
Ordinary least squares (OLS)	1,000	0.099	(0.096)	0.047	(0.067)	0.010	(0.033)
Transformed OLS	1,000	0.096	(0.095)	0.050	(0.070)	0.010	(0.032)
Tobit	1,000	0.112[a]	(0.106)	0.057[a]	(0.078)	0.014[a]	(0.39)
Poisson	1,000	0.465[a]	(0.215)	0.391[a]	(0.209)	0.270[a]	(0.186)
Overdispersed Poisson	1,000	0.104	(0.101)	0.054	(0.073)	0.012	(0.037)
Negative binomial	1,000	0.071[b]	(0.087)	0.033[b]	(0.062)	0.005[b]	(0.025)
Ordered logit	571	0.088[b]	(0.091)	0.045	(0.070)	0.010	(0.032)
Ordered probit	571	0.087[b]	(0.092)	0.043[b]	(0.068)	0.008	(0.029)
Expected frequency		0.100		0.050		0.010	

Note. Each simulation contained 10 independent variables. Standard deviations of means are presented in parentheses. Because the ordered logit and ordered probit required that there be no missing values in the middle of a distribution, the models were incalculable in a number of instances, particularly when the simulation was conducted with the smaller sample size.

a. The frequency of errors was statistically significantly greater at $p < .05$ than the number expected due to chance at this alpha level.

b. The frequency of errors was statistically significantly less at $p < .05$ than the number expected due to chance at this alpha level.

Table 2

Multivariate Analysis of Variance Results

Source	*df* Numerator/ *df* Denominator	*F*	Wilks's Lambda
Main effects			
Analysis model (8 levels)	21/20,316	296.28	0.46**
Distribution of dependent variable (5 levels)	12/18,719	13.00	0.98**
Sample size (2 levels)	3/7,075	3.58	1.00*
Distribution of independent variable (2 levels)	3/7,075	4.69	1.00**
Interactions			
Analysis model × Distribution of Dependent Variable	84/21,166	47.73	0.60**
Analysis model × Sample Size	21/20,316	1.23	1.00
Analysis model × Distribution of Independent Variable	21/20,316	0.81	1.00
Distribution of dependent variable × Sample Size	12/18,719	3.04	1.00**
Distribution of dependent variable × Distribution of Independent Variable	12/18,719	2.39	1.00**
Sample size × Distribution of Dependent Variable	3/7,075	4.83	1.00**
Total	192/21,214	77.28	0.20**

$^{*}p < .05$. $^{**}p < .01$.

Although OLS regression did not yield more than the expected number of false positives, its violated assumptions still suggest that other models may be more appropriate. Overdispersed Poisson regression (see Gardner et al., 1995) seems promising as both an easy and accurate way of modeling count data. The overdispersed Poisson model is designed for count data, overcomes the major limitation of Poisson regression, and yields the expected number of false positives.

The negative binomial, ordered logit, and ordered probit models all yielded fewer false positives than expected given the specified alpha levels. All three can be seen as conservative tests. Although this may mean that the models are prone to Type II errors, researchers may have faith that findings that these models say are statistically significant are indeed statistically significant. Yet, because of the nature of count data, it seems that the negative binomial model is more theoretically appropriate than the ordered models. Hopefully, future research analyzing count data will move beyond simple OLS regression. As suggested by other researchers (Baba, 1990; Hammer & Landau, 1981), the use of multiple methods of analysis appears to hold some hope for coping with the difficulties of research using count data. For example, OLS may be desirable to use because of its familiarity and because it does not yield more false positives than expected. The overdispersed Poisson regression model may also be used because it is relatively easy to compute and matches the assumptions of the data better than OLS while still yielding the appropriate level of Type I errors. Finally, the negative binomial can serve as a conservative check of the results because it is likely to better match with the characteristics of the data but may be prone to Type II errors.

Future research should address the methodological issues of count data research by exploring other potential models or methods of model comparisons, which perhaps may lead to a definitive best way to analyze count data. More extensive simulation may provide researchers with a better understanding of which models are most appropriate for certain circumstances. In addition, more thought needs to be given to the issue of the sensitivity of these models to Type II errors. But perhaps most important, research involving count data should begin using additional, or perhaps alternative, methodologies to perform hypothesis tests. The results from this article confirm that previous empirical findings from OLS regression are likely still valid; however, a number of alternatives to OLS regression exist that may make the way we look for results more accurate.

References

Abelson, R. P. (1985).A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, *97*, 129-133.

Amemiya, T. (1981). Qualitative response models: A survey. *Journal of Economic Literature*, *19*, 481-536.

American Psychological Association. (APA). (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.

Arbous, A. G., & Sichel, H. S. (1954). New techniques for the analysis of absenteeism data. *Biometrica*, *41*, 77-90.

Baba, V. V. (1990). Methodological issues in modeling absence: A comparison of least squares and Tobit analysis. *Journal of Applied Psychology*, *75*, 428-432.

Blau, G., & Boal, K. (1989). Using job involvement and organizational commitment interactively to predict turnover. *Journal of Management*, *15*, 115-127.

Busemeyer, J. R., & Jones, L. E. (1983). Analyses of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, *93*, 549-562.

Cameron, A. C., & Trivedi, P. K. (1986). Econometric models based on count data: Comparisons and applications of some estimators and tests. *Journal of Applied Economics*, *1*, 29-53.

Cameron, A. C., & Trivedi, P. K. (1990). Regression-based tests for overdispersion in the Poisson model. *Journal of Econometrics*, *46*, 347-364.

Clegg, C.W. (1983). Psychology of employee lateness, absence, and turnover: A methodological critique and an empirical study. *Journal of Applied Psychology*, *68*, 88-101.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997-1003.

Gardner, W., Mulvey, E. P., & Shaw, E. S. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin*, *118*, 392-404.

Gibson, J. (1966). Toward a conceptualization of absence behavior of personnel in organizations. *Administrative Science Quarterly*, *11*, 107-133.

Greene, W. H. (1992). *LIMDEP*. Castle Hill, Australia: Econometric Software.

Greene, W. H. (1993). *Econometric analysis*. New York: Macmillan.

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, *82*, 1-20.

Gurmu, S. (1991). Tests for detecting overdispersion in the positive regression model. *Journal of Business and Economic Statistics*, *9*, 215-222.

Gurmu, S., & Trivedi, P. K. (1992). Overdispersion tests for truncated Poisson regression models. *Journal of Econometrics*, *54*, 347-370.

Hammer, T. H., & Landau, J. (1981). Methodological issues in the use of absence data. *Journal of Applied Psychology*, *66*, 574-581.

Harrison, D. A., & Hulin, C. L. (1989). Investigations on absenteeism: Using event history models to study the absence-taking process. *Journal of Applied Psychology*, *74*, 300-316.

Hausman, J., Hall, B. H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, *52*, 909-938.

Hulin, C. L., & Rousseau, D. M. (1980). Analyzing infrequent events: Once you find them your troubles begin. In K. H. Roberts & L. Burstein (Eds.), *Issues in aggregation: New directions for methodology of social and behavioral science* (Vol. 6, pp. 1-16). San Francisco: Jossey- Bass.

Ivancevich, J. M. (1985). Predicting absenteeism from prior absence and work attitudes. *Academy of Management Journal*, *28*, 219-228.

Johnson, R. A., &Wichern, D.W. (1998). *Applied multivariate statistical analysis*. Englewood Cliffs, NJ: Prentice Hall.

Kirk, R. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56*, 746-759.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Landy, F. J., Vasey, J. J., & Smith, F. D. (1984). Methodological problems and strategies in predicting absence. In P. S. Goodman & R. S. Atkin (Eds.), *Absenteeism: New approaches to understanding, measuring, and managing employee absence* (pp. 110-157). San Francisco: Jossey-Bass.

Lee, L. F. (1986). Specification test for Poisson regression models. *International Economic Review*, *27*, 689-706.

Leigh, J. P. (1985). The effects of unemployment and the business cycle on absenteeism. *Journal of Economics and Business*, *37*, 159-170.

Loftus, G. R. (1991). On the tyranny of hypothesis testing in the social sciences. *Contemporary Psychology*, *36*, 102-105.

Maddala, G. S. (1984). *Limited-dependent and qualitative variables in econometrics*. Cambridge, UK: Cambridge University Press.

Martocchio, J. J., & Harrison, D. A. (1993). To be there or not to be there?: Questions, theories, and methods in absenteeism research. *Research in Personnel and Human Resources Management*, *11*, 259-328.

Mikalachki, A., & Gandz, J. (1982). *Managing absenteeism*. Ontario, Canada: University of Western Ontario.

Morgan, L., & Herman, J. (1976). Perceived consequences of absenteeism. *Journal of Applied Psychology*, *61*, 738-742.

Nicholson, N. (1977). Absence behavior and attendance motivation: A conceptual synthesis. *Journal of Management Studies*, *14*, 231-252.

Nunnally, J. M. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Ostroff, C. (1993). Comparing correlations based on individual-level and aggregated data. *Journal of Applied Psychology*, *78*, 569-582.

Rhodes, S. R., & Steers, R. M. (1990). *Managing employee absenteeism*. Reading, MA: Addison-Wesley.

Rosenthal, R., & Rubin, D. B. (1985). Statistical analysis: Summarizing evidence versus establishing facts. *Psychological Bulletin*, *97*, 527-529.

Rushmore, C. H., & Youngblood, S. A. (1979). Medically-related absenteeism: Random or motivated behavior? *Journal of Occupational Medicine*, *21*, 245-250.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods*, *1*, 115-129.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*, 199-223.

Steers, R. M., & Rhodes, S. R. (1978). Major influences on employee attendance: A process model. *Journal of Applied Psychology*, *62*, 319-327.

Thompson, B. (1994). Guidelines for authors. *Educational and Psychological Measurement*, *54*, 837-847.

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher*, *25*(2), 26-30.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, *26*, 24-36.

Waters, L. K., & Roach, D. (1979). Job satisfaction, behavioral intention, and absenteeism as predictors of turnover. *Personnel Psychology*, *32*, 393-397.

Watson, C. J., Driver, R.W., &Watson, K. D. (1985). Methodological issues in absenteeism research: Multiple absence measures and multivariate normality. *Academy of Management Review*, *10*, 577-586.

Zavoina, R., & McElvey, W. (1975, Summer).A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, pp. 103-120.