Meta-Analysis of Human Trafficking in the United States

David J Corliss and Heather M. Hill

Peace-Work, Plymouth, MI

ABSTRACT

Meta-analysis is a method for combining multiple independent studies on the same subject or question, producing a single large study with increased accuracy and enhanced ability to detect overall trends and smaller effects. This is done by treating the results of each study as a single observation and performing analysis on the set, while controlling for differences between individual studies. These differences can be treated as either fixed or random effects depending on context. This paper demonstrates the process and techniques used in meta-analysis using human trafficking studies. This problem has seen increasing interest in the past few years and there are now a number of localized studies for one state or a metropolitan area. This meta-analysis combines these to begin development of a comprehensive analytic understanding of human trafficking across the United States. Both fixed and random effects are described. All elements of this analysis were performed using SAS® University Edition.

INTRODUCTION – HUMAN TRAFFICKING IN THE UNITED STATES

In recent years, there has been increasing public awareness of the problem of human trafficking in the United States. Civic groups and social justice organizations have begun to partner with law enforcement to address this persistent and growing problem. Human trafficking is often brown down into four types: sex trafficking, forced labor, debt bondage, and domestic servitude.

Sex trafficking is the most familiar form of human trafficking, in part because it is occasionally encountered in public places, while other forms tend to operate strictly in secret. Recently, however, law enforcement and public advocacy groups have become aware of a vast underground industry in sex trafficking. Victims are often found among runaways or others without critical support structures from family and friends. The new secret sex industry in the United States often has an international dimension, as victims are tricked by traffickers with promises of a job or student programs in the United States, only to find themselves sold into trafficking and held captive by stealing their passport and physical violence.

Forced labor has many of the same characteristics of sex trafficking with different a different kind of work: labor in secret sweatshops right here in the United States. Conditions normally found in the worst situations in Third World countries are replicated in the U.S., with the same conditions: captive labor that cannot leave the job or even the premises. As in other parts of the world, forced labor in America in most common in the garment industry; Robert Ross' book *Slaves to Fashion* is an exposé of this hidden criminal industry.

Debt bondage is the modern form of indentured servitude. While the 13th Amendment to the Constitution criminalized indentured servitude as well as slavery, both were driven underground - not abolished - and persist today. It is especially common among immigrants tricked into debt bondage in the United States, who have held until the cost of being smuggled into the country is paid. Whatever the source of the putative debt, those dealing in modern-day slavery see to it that the debt is never paid, generally through charging the victim excessive amounts for rooms, board, travel and other costs. The victim remains in slavery throughout his or her life.

Domestic servitude is perhaps the hardest form of human trafficking to detect, challenging quantification and therefore many analytic-based lines of redress. Victims are often held more by threats of violence – usually against family members – than physical imprisonment. In many cases, domestic servitude victims are used as sex slaves as well.

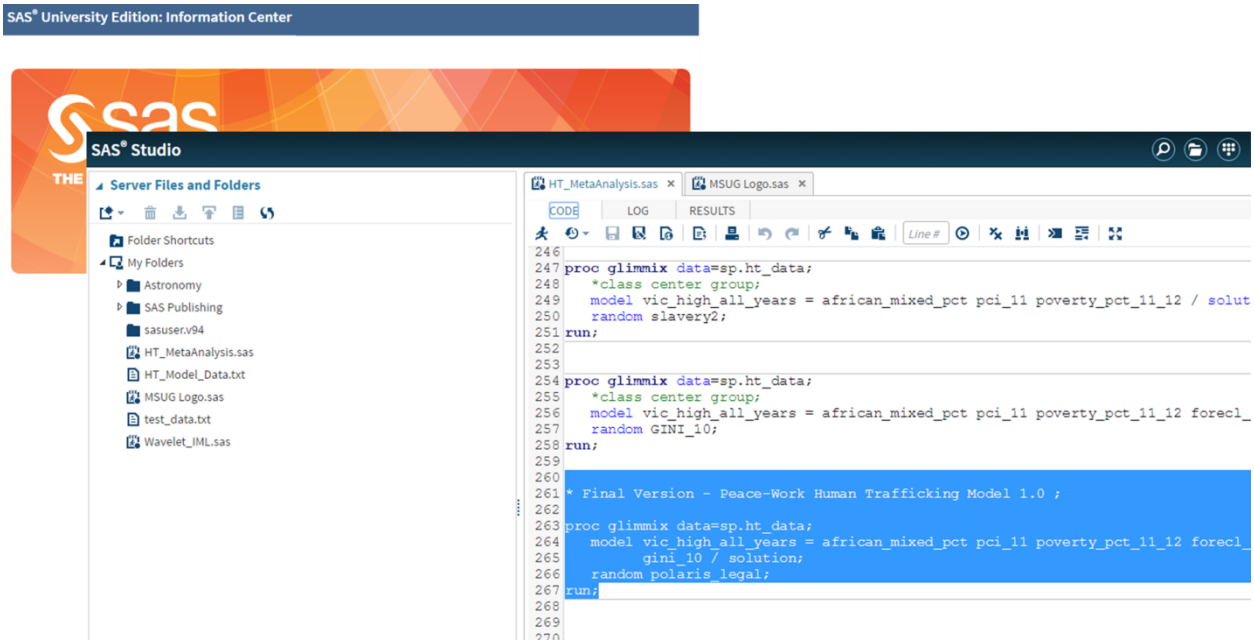
ANALYSIS OF HUMAN TRAFFICKING

In the last 10 years, there increasing public awareness of this important problem, as local victim advocacy and intervention groups sprang up across the country. Law enforcement task forces have been established in many areas. One of the most important drivers of change has been Polaris, a national not-

for-profit organization dedicated to the eradication of modern-day slavery. In 2007, Polaris established the National Human Trafficking Resource Center (NHTRC), including a call center for victims. Others who suspect human trafficking can contact the NHTRC hotline. Data from this the Polaris call center, summarized at a state level and published on their website, forms the basis of this study.

Advanced analytic methods using SAS are being applied to these data by Peace-Work, an all-volunteer cooperative of statisticians, data scientists and other researchers applying analytics to issues in poverty, education and social justice. Peace-Work is part of the growing "Data For Good" movement, where teams of data scientists, statisticians and other researchers use their analytic skills to pressing issues and problems in our society and in our world. A fairly new organization, Peace-Work operates by connecting volunteers with organizations and datasets, often from governments and other official sources, to address issues volunteers care about deeply. With no full-time staff, larger statistical project are broken into small pieces to accommodate volunteers busy schedules.

SAS University Edition as a powerful analytic tool with extensive statistical, machine learning and big data capabilities available to their researchers that is free for non–commercial use. Using virtual server technology, it provides a familiar console-based SAS coding environment with an extensive tool set of SAS procedures and data management and visualization tools. Peace-Work makes extensive use of SAS University Edition its analyses, including data ingestion, ETL, and modeling for its human trafficking project.

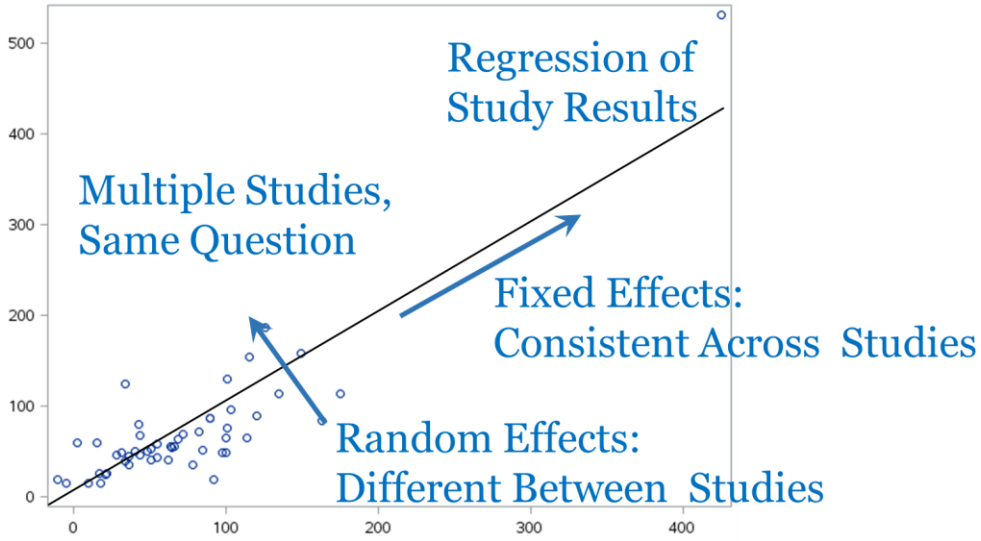


META-ANALYSIS

Meta-analysis is a commonly used analytic method to combine multiple, related statistical studies into a single, unified analysis. By incorporating data from all of the earlier studies, it is possible to reduce statistical error, gain broader insights across larger populations, geographies and other factors, and understand the factors associated with differences between individual studies on a given subject. Meta-analysis perhaps is most familiar from medical, especially pharmaceutical analyses, combining multiple studies on a given drug, medical condition or treatment. These studies increase the sample size of often very limited studies and bring into focus the difference between the populations in different studies associated with differences in outcomes.

Mixed models – regression-type models combining both fixed and random effects - are especially well-suited to meta-analysis. In these models, factors that are *consistent* across the different studies, processes and populations become the fixed effects. *Variations between studies* are explained by the random factors. In this way, meta-analysis can be a study in how to avoid Simpson's Paradox. For

example, two studies on angioplasty by two different university hospital systems. The fixed effects in the mixed model analyze angioplasty itself, while the random factors provide insight into the differences in the environment and context between the two universities and how the data are captured and processed.



In SAS, one choice for mixed models is PROC GLMMIX. In this syntax, the fixed effects are placed in the MODEL statement, much like a simple regression model. Random effects are explicitly called out in a RANDOM statement:

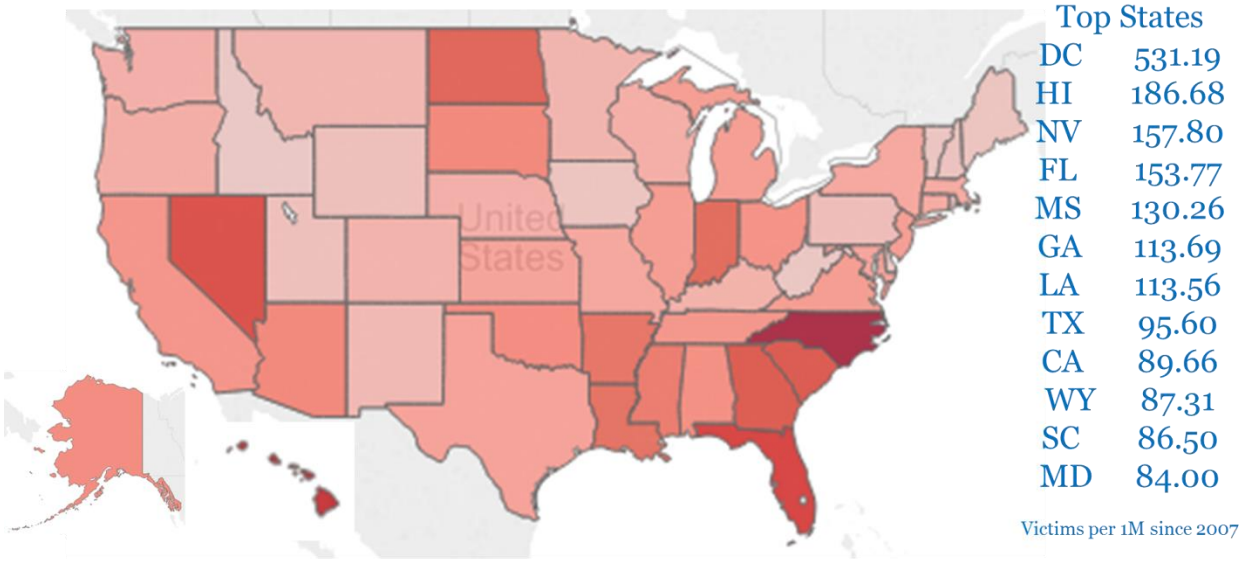
```
proc glimmix data=pharma_study;
   class study treatment;
   model serum_level =  x1  x2  x3  x1*x2  x1*x3
        / solution;
   random treatment(study);
run;
```

META-ANALYSIS OF HUMAN TRAFFICKING IN THE UNITED STATES

The first phase of study applied meta-analysis to state-level summary data. NHTRC call center data provided the outcome, using the number of contacts with rated as “highly probable” to be human trafficking victim from the creation of the database in 2007 until the present, divided by the state population in millions, used as the key metric. 78 socio-economic and demographic factors were considered for fixed effects. PROC GLMSELECT with LASSO and LAR were used to screen candidate factors. The final model included:

- Combined African-American and mixed-race percentage from the Census Bureau. Higher percentages of both racial group designations are associated higher levels of modern day slavery, leading to combining them into a single factor
- Per capita income – the 2011 value was selected to be near the middle of time period of the outcome data
- Percent of households below the poverty line
- GINI index by state, indicating income disparity as an associated factor.
- The year-over-year rate of change in the home foreclosure rate – this poverty factor is closely associated with worsening poverty and new homelessness, suggesting a possible link between becoming homeless and being a victim of human trafficking

- Historical prevalence of slavery. It was noted during data exploration that states with a history of legal slavery prior to 1865 often had high per capita rates of slavery today, including 9 of the top 12 states. Two geographic factors were tested. The first is a simple binary factor with a value of 1 for states with legal slavery at the time of the Civil War (without regard to secession) and 0 for free states. A second factor attempts to classify states ordinal scale, with a value of 2 for states in the Deep South – TX, LA, MS, AL, GA, FL, SC and NC – 1 for the remaining slave states and 0 for the free states. The second factor was found to be more predictive and strengthened the model, if marginally, by its inclusion.

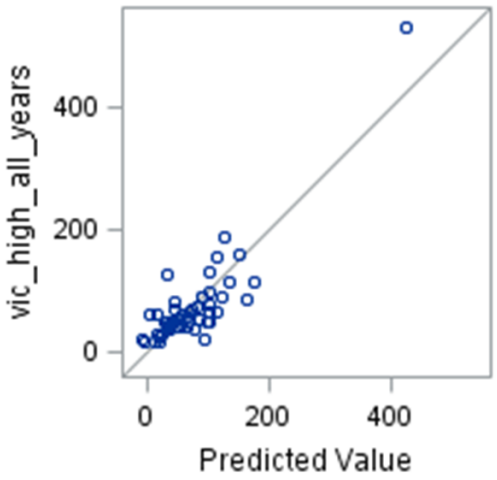


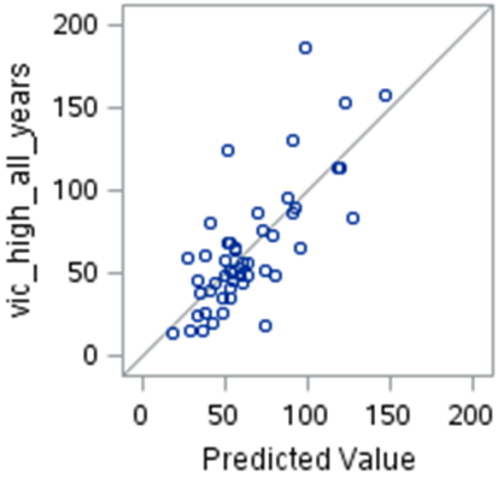
Here is the SAS code for the national mixed model:

```
proc glimmix data=sp.ht_data;
   model vic_high_all_years =
       african_mixed_pct pci_11 poverty_pct
       GINI_10 forecl_change_rate slavery2 / solution;
   random state;
run;
```

Model Results

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
african_mixed_pct	1	43	5.36	0.0255
pci_11	1	43	48.06	<.0001
Poverty_pct_11_12	1	43	25.11	<.0001
forecl_change_rate	1	43	13.53	0.0006
Slavery2	1	43	2.36	0.1319
GINI_10	1	43	14.02	0.0005





50 States and DC
R-Squared: 0.7705

50 States / not DC
R-Squared: 0.5615

GEOGRAPHIC VARIATIONS

The mixed effects model selected single random effect: State. While the outcome data measuring the level of modern-day slavery in each state has been compiled in a single national database, the state-to-state differences in environment are considerable, with difference in local organizations reaching out to victims, varying practices among law enforcement that tend to be standardized within individual states, and especially the legal environment toward human trafficking in each state. Some states will do more than others in terms of reporting standards and legal treatment of victims. While the fixed effects point to certain socio-economic factors as key predictors of human trafficking levels, state-by-state variations impact the degree to which different states address this issue and report data.

This draws a distinction between the actual level of human trafficking in a state and the degree to which it is reported. As an example, the fixed effects when employed as an ordinary regression model indicate Michigan and Ohio should have similar victim rates. However, in actual practice, Ohio has a higher reported rate. This can be explained by a legal and legislative in Ohio that is more favorable to the identification and reporting of human trafficking victims:

Michigan

17.9 Victims per 1M

5.5 out of 12 legislative actions recommended by the NHTRC

Ohio

22.0 Victims per 1M

9 out of 12 legislative actions recommended by the NHTRC

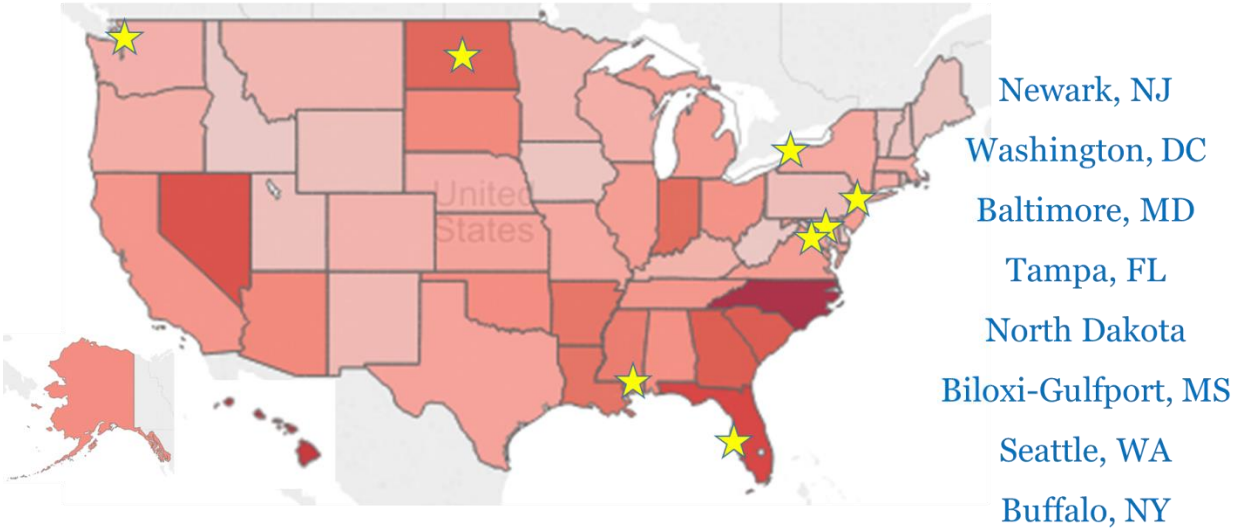
Polaris data, victims identified with high confidence, 2015

The prevalence and impact of local victim advocacy and outreach groups and variations in the legal and legislative environment impact the reporting of victims, apart from the organic economic drivers of the true level of human trafficking in each state. While the NHTRC call center has produced a single database, the information in it should not be viewed as a single object but rather as a mosaic individual state-level reporting systems. For this reason, meta-analysis is an appropriate analytic methodology to apply to this problem, the state as a random effect.

TARGETING LOCATIONS

Individual metropolitan areas may be targeting using a Decision Tree model, leveraging the fixed effects identified by the National mixed model along with local factors such as proximity to major freeway routes and international borders, both of which facilitate the transport close to demand and far from potential rescue. This study initially identified 18 metropolitan areas where human trafficking is expected to be high. Many of these locations had already been identified and are the focus of critical partnerships between law enforcement and victim advocacy and outreach organizations. However, eight locations are noteworthy for not being identified of potential centers of human trafficking activity or are severely lacking in resources to combat it:

Candidate Locations of Undetected or Under-Resourced Human Trafficking Centers



CINCINNATI – A GEOGRAPHICALLY TARGETED ANALYSIS

One metropolitan area where a local task force is making a significant impact is Cincinnati. While statistical modeling identified Cincinnati as an area of serious concern, End Slavery Cincinnati – a community coalition of more than 40 organizations and individuals – has been working to stop human trafficking and support victims since 2007.

In the case of Cincinnati, important socio-economic and demographic factors driving human trafficking include very high levels of poverty – two to three times the level of Cleveland and Columbus – combined with significant income disparity with highly affluent areas, and a high percentage of persons of color, who are more often victimized. Local factors include the intersection of three interstate highways, especially I-75; an international airport, a state border complicating collaboration between law enforcement and other organizations, manufacturing centers most people never see, and a large number of mobile and transient persons so encountering people from out of town is commonplace.

The statistical model based on the NHTRC published data has been applied to the population, demographics and economic characteristics found in the Cincinnati area to develop a picture of human trafficking in the area.

Based on the above analysis, estimated summary statistics on human trafficking in the greater Cincinnati area are as follows:

- Number of contacts to the National Human Trafficking Resource Center per year: 224
- Number of cases reported: 61
- Number of victims contacted: 70-80

Cincinnati is a large Midwestern city and metropolitan area with significant economic drivers of high levels of human trafficking, including high levels of poverty and economic disparity. There are also significant logistic factors facilitating this criminal industry, including important freeway routes and other travel routes, a large but declining manufacturing center, and a state border adding a level of complication to law enforcements and coordination of agencies. These factors are mitigated by a strong local task force, End Slavery Cincinnati, coordinating the efforts of dozens of advocacy and intervention organizations, community service groups, law enforcement and other government agencies. The inter-state nature of this task force is a particular strength of the local response to human trafficking in the Cincinnati metropolitan area.

While sex trafficking in the urban core receives much public concern, sex slavery likely flourishes in moderately affluent suburbs with little notice. Domestic servitude in moderately affluent areas and forced labor in the urban core may be an issue but little hard data exists for these forms of trafficking in Cincinnati. Given the relative paucity of reports of domestic and forced labor to the NHTRC hotline - just 7.3% of reported cases – indicates further efforts in these areas, along with sex trafficking in more affluent areas, may be very productive.

CONCLUSION

Peace-Work volunteers, using data from the National Human Trafficking Resource Center (NHTRC), have identified and modeled factors associated with elevated levels of human trafficking in the United States. Key factors are poverty, economic disparity and race, with geographic areas with above average percentage of persons of color having higher levels of modern-day slavery, even after accounting for poverty factors often associated with race.

Variation was seen between states, which have different legal and legislative environments leading to some variability in the probability of reporting cases. This variation led to the use of a mixed model to approach the problem, with demographic and economic factors as fixed effects and state as a random effect. R-squared for a regression model containing the fixed effects is 0.56, indicating both some progress made and more work needed to refine the model.

A decision tree model was developed using the fixed effects from the mixed model, supporting prediction of the level of the relative level of human trafficking in smaller areas such as cities, counties and metropolitan areas. This model was applied to the Cincinnati metropolitan area to predict human trafficking levels, success and challenges in the Cincinnati area.

Recommended next steps include: independent review of the data, methods and results by other statisticians. Similar work can be performed for other areas of the country to better understand the successes and challenges faced of human trafficking victim advocacy and intervention organizations to guide the allocation of scarce resources. Meta-analysis can be applied to other social justice concerns and issues, such as homelessness and to human trafficking internationally.

While we can be encouraged that more action is being taken on sex trafficking – even as more efforts are needed - there is another side of this public perspective that gives more concern. While the plight of people caught and held in human trafficking for sex tugs at heart-strings and motivates many to action, there is less concern among the general public over those held captive and forced to work in fields, factories, kitchens and even homes across the United States. Forced labor, debt bondage and domestic

servitude are areas where volunteer analytics in the Data For Good movement can make a great impact on an often overlooked problem. Greater involvement by professionals in statistics and data science offers many excellent opportunities to make a significant impact on many important issues in our society and our world.

ACKNOWLEDGEMENTS

Thanks are due to the SAS Institute for making available SAS University Edition at no cost to for non-commercial use, including this project.

REFERENCES

Schabenberger, O., 2005, “Introducing the GLIMMIX Procedure for Generalized Linear Mixed Models”, Proc. SUGI 30

Schall, R. (1991), “Estimation in Generalized Linear Models with Random Effects,” Biometrika, 78:719–727

Whitehead, A. “Meta-Analysis of Controlled Clinical Trials, Wiley 2002

Banks, D. and Kyckelhahn, T. (2011), “Characteristics of Suspected Human Trafficking Incidents, 2008-2010”, USDOJ

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

David J Corliss
Peace-Work
davidjcorliss@peace-work.org