

FACT Act Scores Study, Matter P044804

Our input regarding “the prescribed methodology and research design of the study” follows:

1. Any study that is undertaken should not just address the issue of whether any of the "protected classes" under the Equal Credit Opportunity Act (the ECOA) have, on average, lower credit scores than non-protected classes. It should address the issue of whether credit scores are significantly related to future losses, and, if so, whether that relationship is similar for protected classes as for non-protected classes.

2. In a similar vein, any study undertaken should not just address the issue of whether any of the protected classes pay, on average, significantly different *premiums* for similar coverage than do non-protected classes. It should also address the issue of whether any of the protected classes have, on averages, significantly different *losses* than do non-protected classes. After all, it is widely agreed that different premiums are justified in the presence of different loss ratios.

3. Any numerical manipulations of the data prior to analysis must not be susceptible to distorting the magnitude of any differences found. (For example, converting original measurements to percentiles is likely to exaggerate the magnitude of any differences present in data with relatively little variability, while damping the magnitude of any differences present in data with much variability.)

4. Whatever statistical techniques are to be used in the data analysis (i.e. multiple regression), there MUST be conscious examination of whether the underlying assumptions of the technique are reasonably satisfied, BEFORE proceeding to an interpretation of the results. For example, are there notable biases in sampling? Do quantities of interest nearly follow the assumed distributions (if any) used to develop the statistical technique?

5. In multiple regression situations, the extent of multicollinearity should be examined when determining which variables to retain in the final model.

6. Concerning availability of data, it is our understanding not only that insurance companies do not have data on ethnicity, color, religion, etc. of their insureds, but that it would be illegal for them to gather such information. So it seems that the gathering of data may require selecting a random sample of the population, gathering such sensitive data directly from the individuals, and then contacting their current insurers for premium and loss information on their policies (under strictest confidentiality of course). And this obviously only works if the sample members have insurance currently.