

# Insurance Scores and Assertions of Disparate Impact

Richard A. Smith  
EPIC Consulting, LLC  
rsmith@ask-epic.com



# Disparate Effect

## Definition / Burden of Proof

- Burden is upon aggrieved party to show disproportionate impact (negative impact on minority/protected group)
- Burden then shifts to company to show that the process has a valid business purpose
- Burden then shifts back to aggrieved party to show a feasible, effective alternative

# The Challenges / Response

- External focus seems to be on:
  - Limited evidence of prevalence of certain credit report attributes
  - Average score differences
- “Slippery Slope”
  - What about other risk characteristics (e.g. territory, multi-car, multi-line, age of dwelling, driver class)?
  - How much (difference) is “too much”?
- Industry focus
  - Provide actuarial support of use of score
  - Input to suggest that studies must be done properly
  - Trade associations busy on legal issues
    - Does it apply to insurance? (NAMIC)
    - Do states have legal standing to do some studies?
    - Point out FCRA explicitly allows use
- Others (e.g., III) have also offered critiques

# Federal Government Now Involved

- Congress directs FTC (Insurance) and Fed (Banking) to analyze insurance scores
  - Effect on affordability/availability?
  - Disparate Impact?
- Will develop joint report
- Plan to use similar (if not identical) approach to analysis

# FTC is Gathering Information

- Have held series of meetings/discussions with stakeholders
  - Consumer groups
  - Individual companies
  - Trade associations
  - State Regulators
- “Aware of” the multi-state initiative and data call
- Will publish in 2005

# Details of Recent Meetings with FTC

- Plan multivariate models
  - Will probably look at average scores, but don't plan to stop there
  - Acknowledge need to control for risk differences, including geographic differences
  - Understand need to include (or model) expected losses as part of analysis
- Discussions with trades (and sponsors) to enhance multi-company database developed for 2003 study
- Use of census-block information at the policy level to assign race/income variables
- Considering that “similar lift” for sub-groups may substantially answer the issue

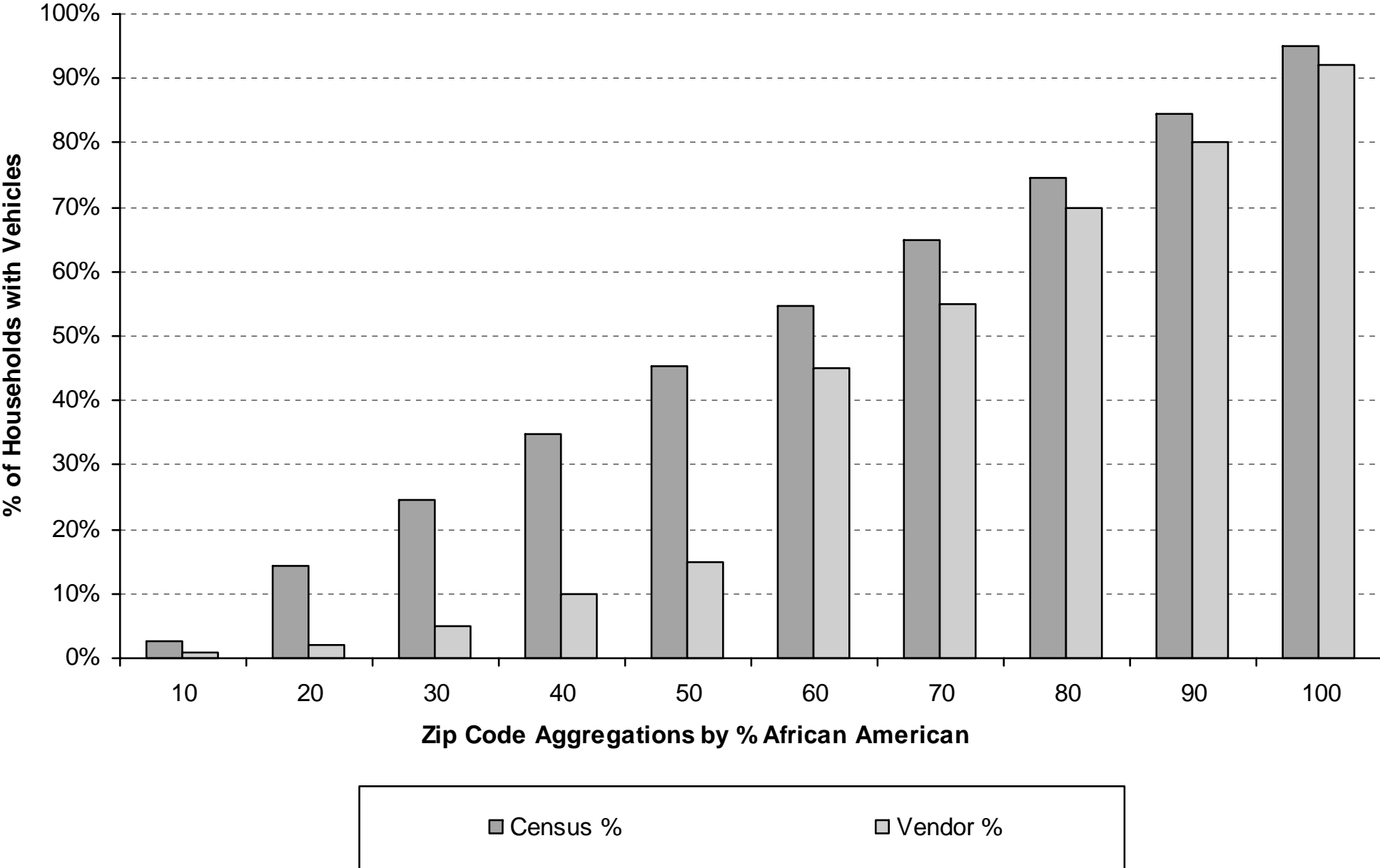
# Have Suggested an Appropriate Approach

- Suggested PD frequency as best basis for study
- Analyze the impact of insurance scores on consumer premiums
  - Described that other factors are affected by introduction of scores (Age, Tenure, Multi-car and –line, etc.)
  - Need before/after multivariate analysis to gauge impact
- Requires race and/or income information to allow aggregation and comparison after modeling

# Identifying an Individual's Race is Difficult

- One problem is proportion of minorities in predominantly minority areas
- Census block is superior to zip code
- Vendors offer ways to assign race to an individual level
  - Results appear reasonable in some cases (e.g., surname of “Hernandez” likely to be Hispanic)
  - Identification of African Americans, especially in low concentration areas, is difficult. Surname of “Smith” not useful.

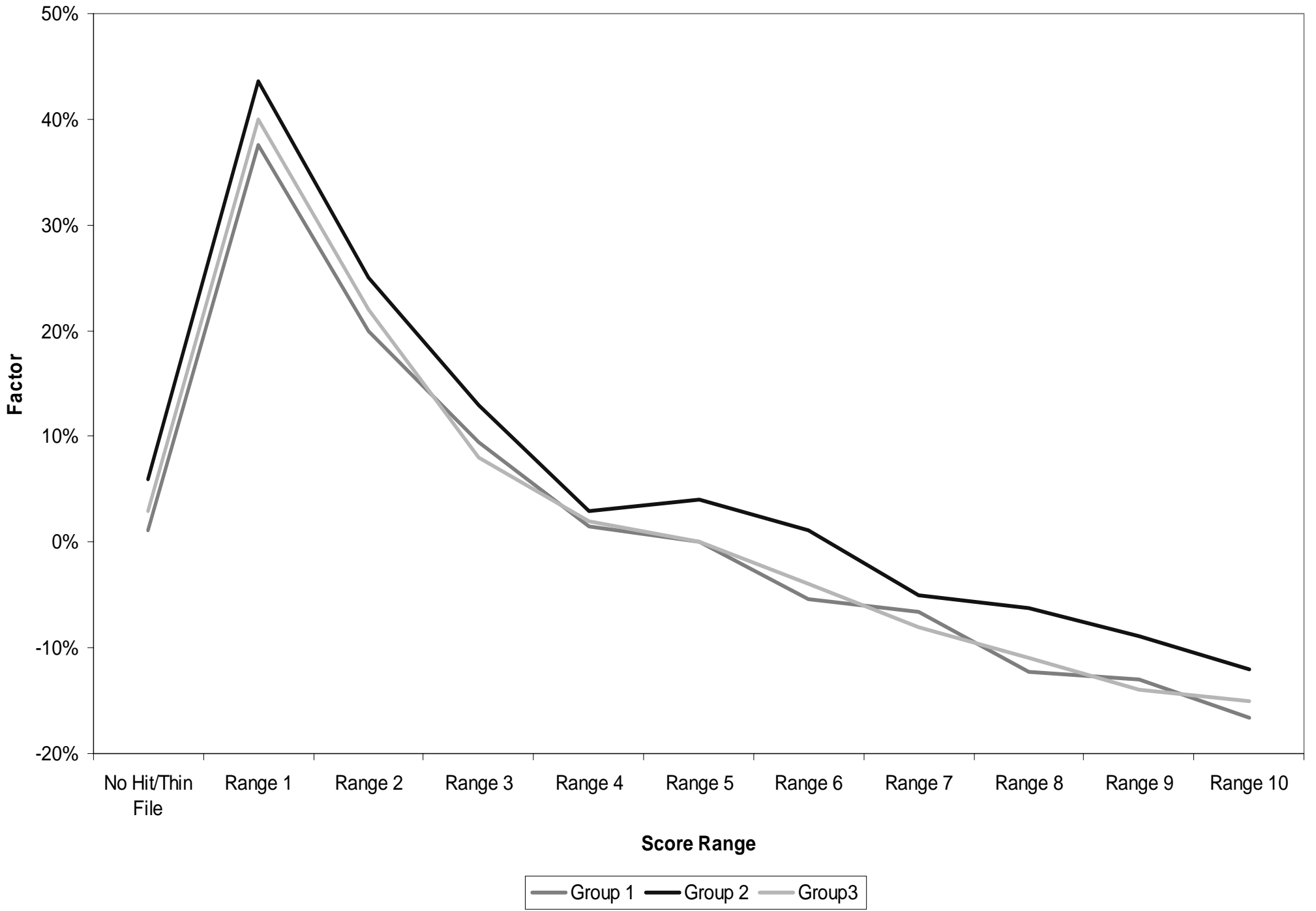
# African American Households – Not Easily Identified



# FTC Also Thinking of Another Approach

- Considering that “similar lift” for sub-groups may substantially answer the issue
  - Essentially, look at the indicated factors for the interaction of race and score
- Concerns?
  - Directly includes race, which may also affect indicated risk factors for other characteristics
  - Political dynamite if indicated factors for minorities are higher than for non-minorities?
  - (Recall that critics refer to public policy decisions to prohibit use of race-related mortality assumptions for life insurance)

# Example of "Similar Lift"



# EPIC Study on Credit-Based Insurance Scores

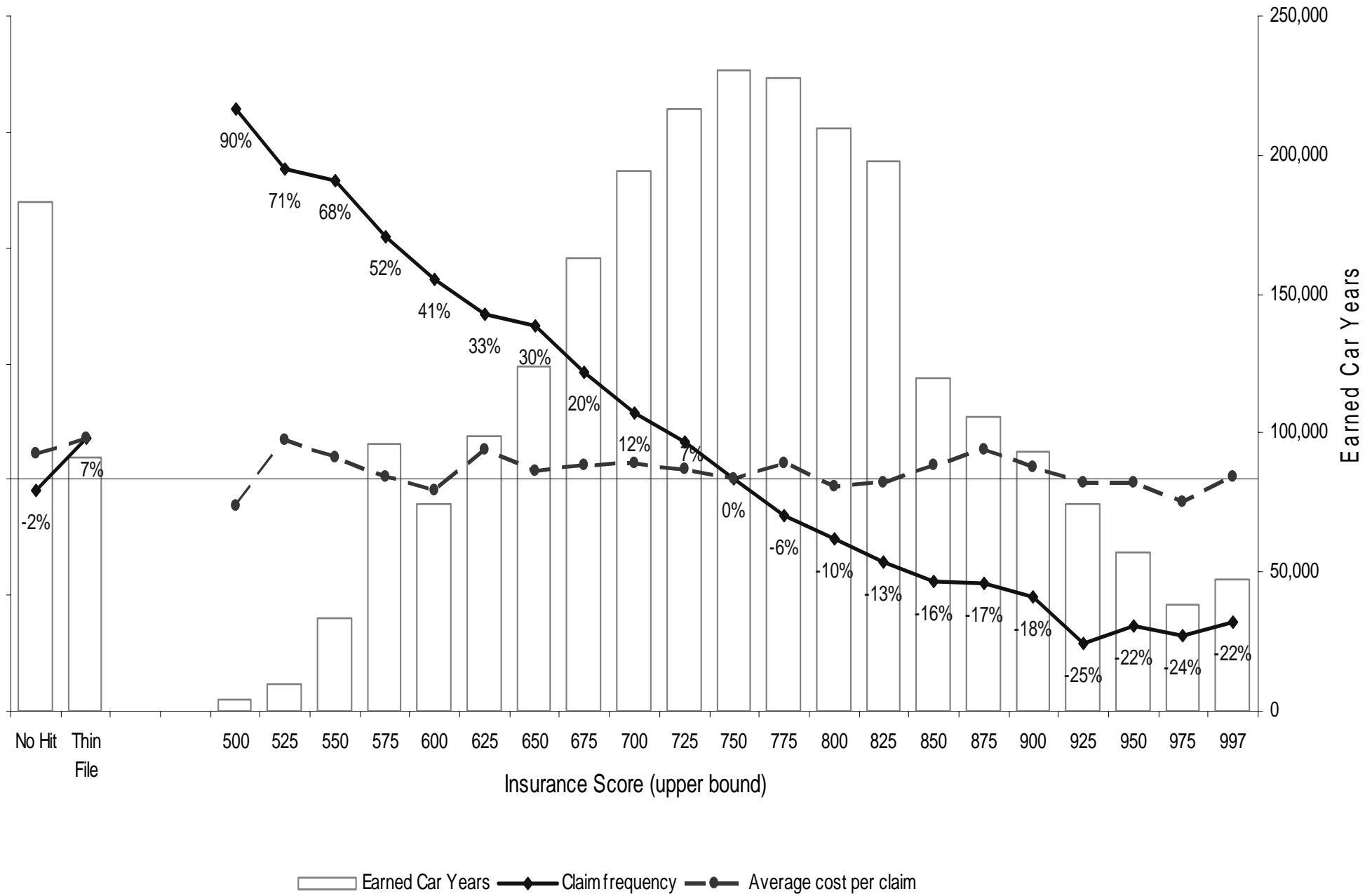
- Sponsored by four trade associations
  - Published in June, 2003
  - Authored by Michael J. Miller and Richard A. Smith
  - Available at [www.ask-epic.com](http://www.ask-epic.com)
- Private passenger auto
- Purpose was to answer three questions
  - Correlation
  - Overlap / Interaction
  - Business purpose / Importance

# EPIC Study on Credit-Based Insurance Scores

- Most of indication is driven by claim frequencies
- Collision and Comprehensive have a severity-related component
- Multivariate analysis reduces, but does not remove the “slope” of univariate results.

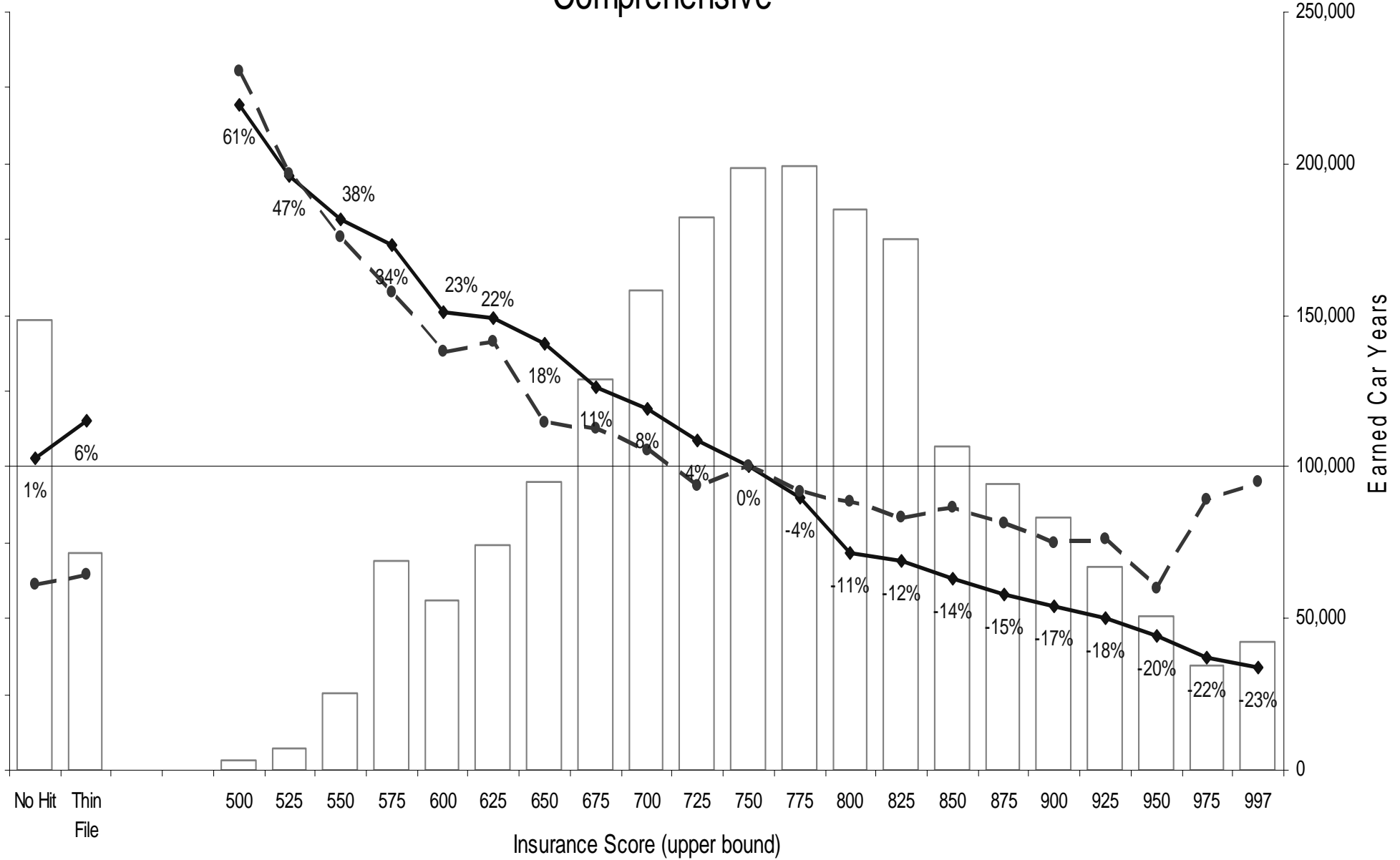
# Loss Relativities by Insurance Score

## Property Damage Liability



# Loss Relativities by Insurance Score

## Comprehensive



Legend:  Earned Car Years   Frequency   Severity

# Obviously, Other Factors Affect Score Indications

- Reviewed results when score is paired with other factors
- Study shows results when all are considered
- Also showed the residual effect
  - All but score is modeled
  - Factors are “frozen”
  - Residual results reviewed
- No surprises, Insurance Scores are important!

# Missouri Study

- Univariate analyses on aggregated data
  - Credit score percentiles (race/income)
  - % exposures in lower three quartiles (race/income)
  - Multivariate analysis to analyze other socioeconomic variables
  - Extrapolation of aggregate analysis to individuals
    - Neighborhood model
    - Goodman's regression
    - Ecological Inference
- Claims disproportionate impact in each case

# Issues with Missouri Study

- Describes as “. . .Impact on Minority . . .”
  - Gives a disclaimer (“the impact ... on pricing and availability ... is not directly examined in this study” )
  - Regulators are using/interpreting otherwise
  - A given difference in average scores is NOT equivalent to the same difference in premium!
- Realistically, univariate analyses only
- Treats all scores as the same
  - Relies on “normalization” to compare companies
  - Some company results hard to criticize even with their approach

# Issues with Missouri Study (cont.)

- Poor data call design (two files)
  - Aggregated data at the ZIP level
  - Year, policy form, ZIP, average & median score, exposures by interval
  - Year, age range, gender, average & median score, exposures by interval
  - Unable to look at age range, etc. by ZIP -Why didn't they ask for one file that combined the information????
- Design introduces biases
  - Multiple year policies are counted multiple times
  - Multiple car policies are counted multiple times
- Their (poor) approach more properly done on a policy-level basis

# Issues with Missouri Study – Univariate Analyses

- Use of percentiles “to allow comparisons”
  - Not unusual, but almost certainly overstates actual differences
  - Highlighted results based on extrapolation to endpoints
- Distribution of exposures into range quintiles
  - Based on range of scores, not quintiles of exposure (third quintile of one score probably near midpoint of exposure distribution)
  - Is disconnected from real-world use – need context of how scores are used in rating (cut-offs, etc.)
- Low  $R^2$  on analysis of aggregate data; presumably much smaller on individual-level data

# Issues with Missouri Study – Multivariate Analyses

- Bad analysis at multivariate level
  - Poorly defined data call prevents meaningful approach; no risk-related variables included
  - Census data used (badly): e.g., average age Also left out other potentially useful variables
  - Ignored loss data available through other MO calls
- Attempts to overcome “Aggregation Bias”
  - No consensus on usefulness/accuracy of methods
  - Seemingly concentrates on EI method, considered very controversial by some, another method gave only minor differences
  - Methods heavily influenced by preliminary findings/assumptions (thwarted by GIGO)

# Example of Aggregated Data Problem

	Number in Worst Score Group	Number in Best Score Group	Totals
Minority Population	Unknown	Unknown	Known
Non-minority Population	Unknown	Unknown	Known
Totals	Known	Known	Known

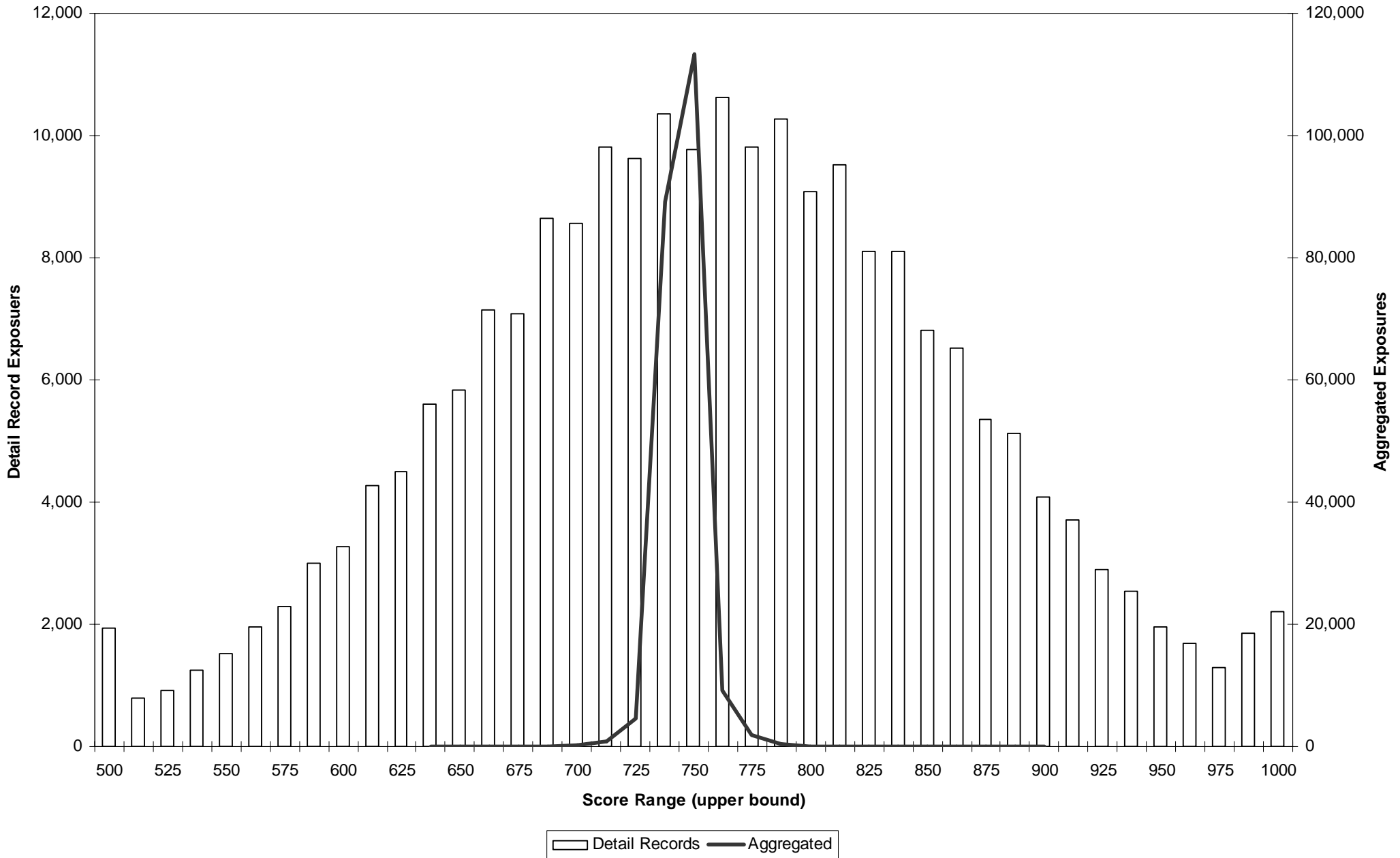
# Dangers of Looking at Average Score

- Remember: multi-car/-year policies over-counted
  - Both have above-average scores
- On a univariate basis, average scores show substantial variation for a number of sub-groups
- Controlling for distributional differences of other characteristics often reduces/removes these differences

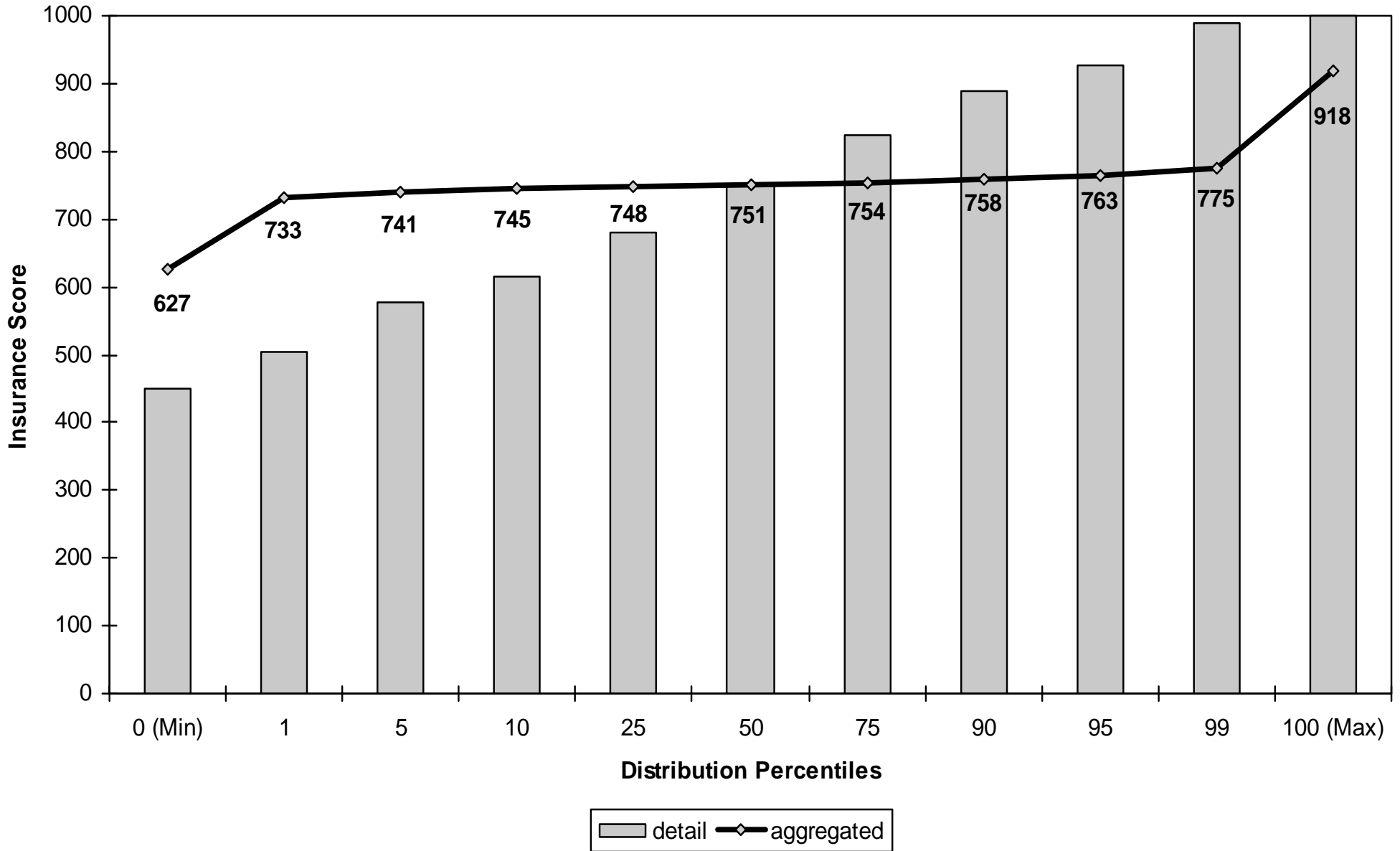
## Issues with Missouri Study – Percentiles

- Study example:
  - 100% minority zip: 18<sup>th</sup> percentile
  - 0% minority zip: 57.3 percentile
- On simulated data, this yields about a 6-point difference in score probably not meaningful in terms of impact on premiums

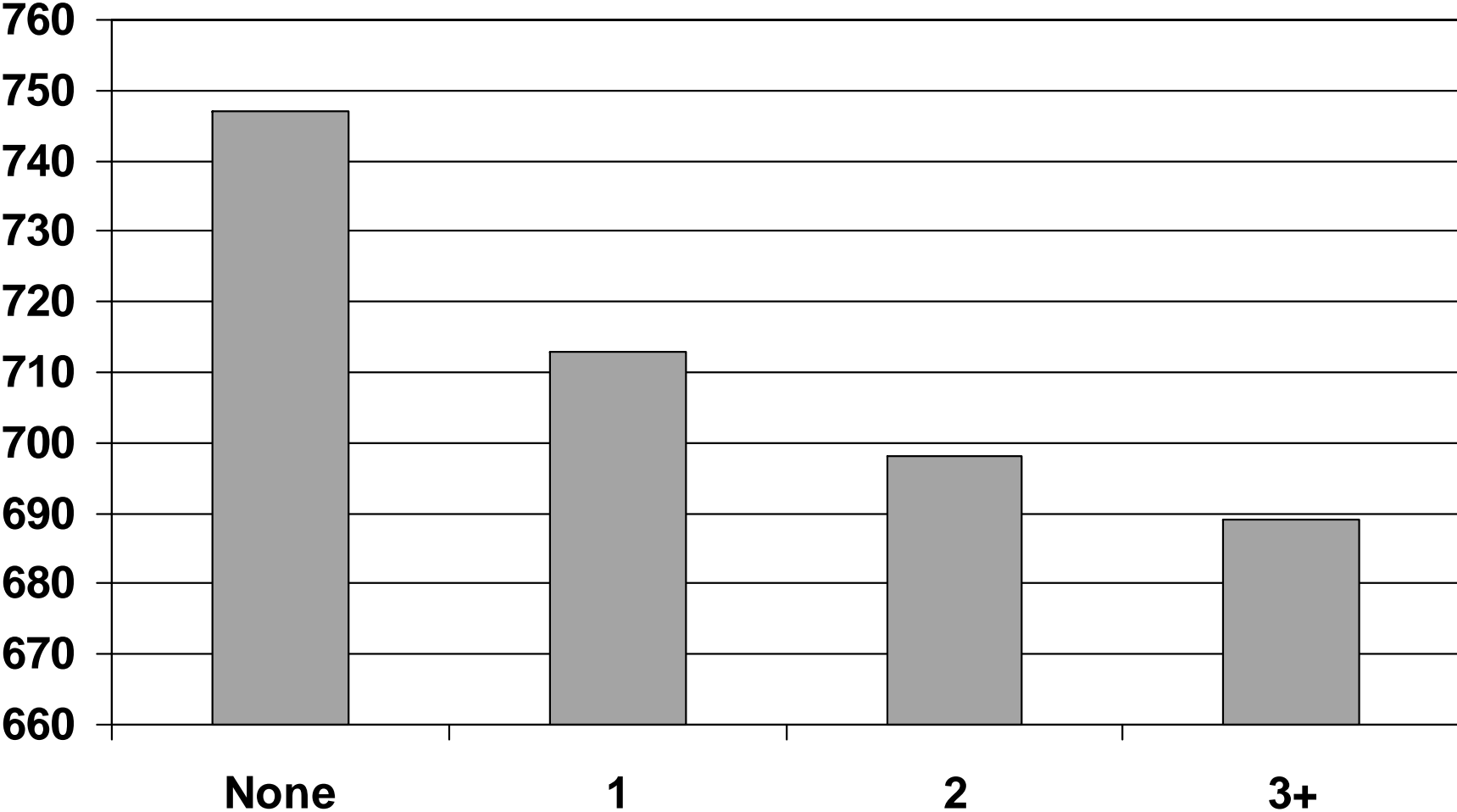
# Simulated Score Distributions



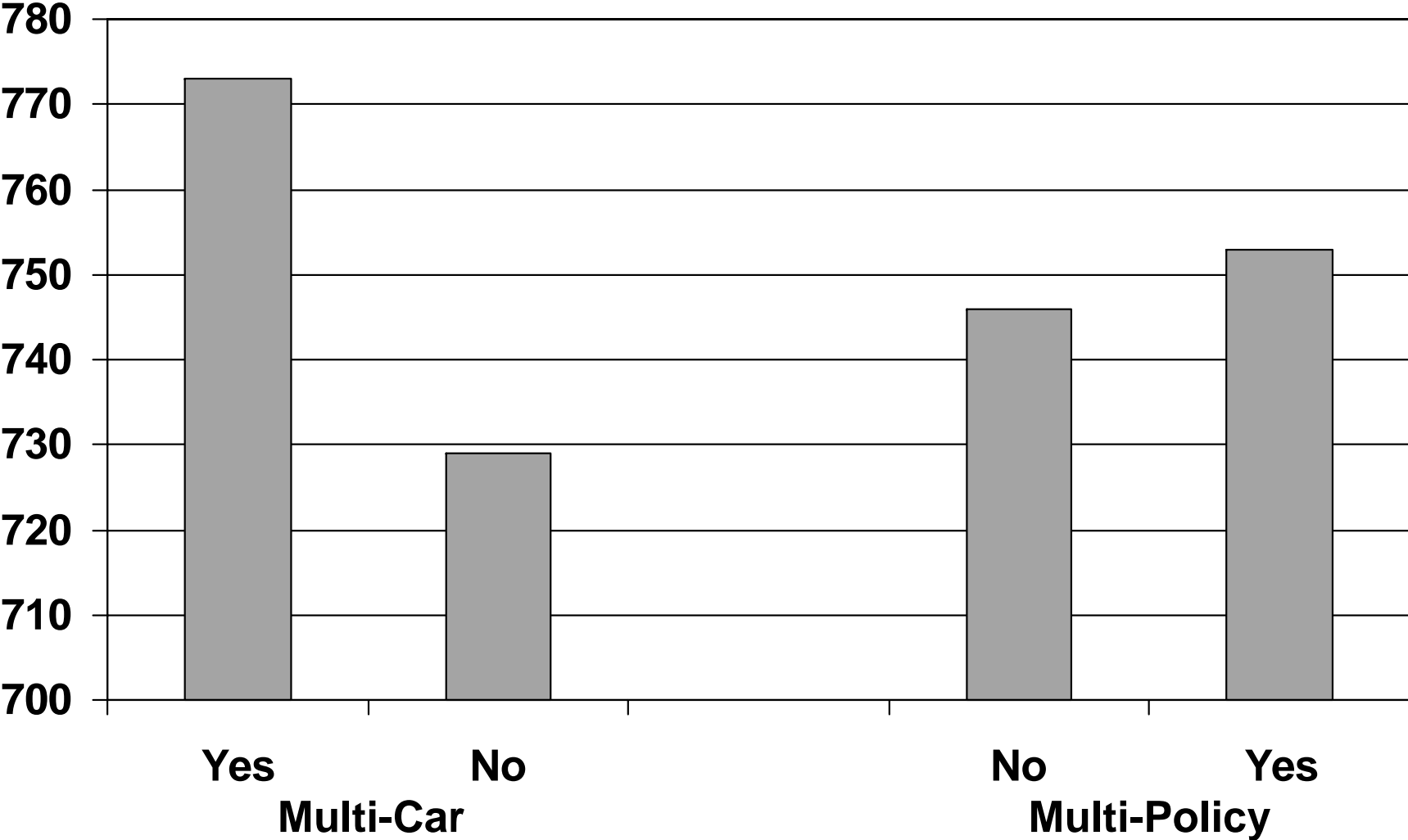
# Simulated Percentiles for Detail vs. Aggregate Data



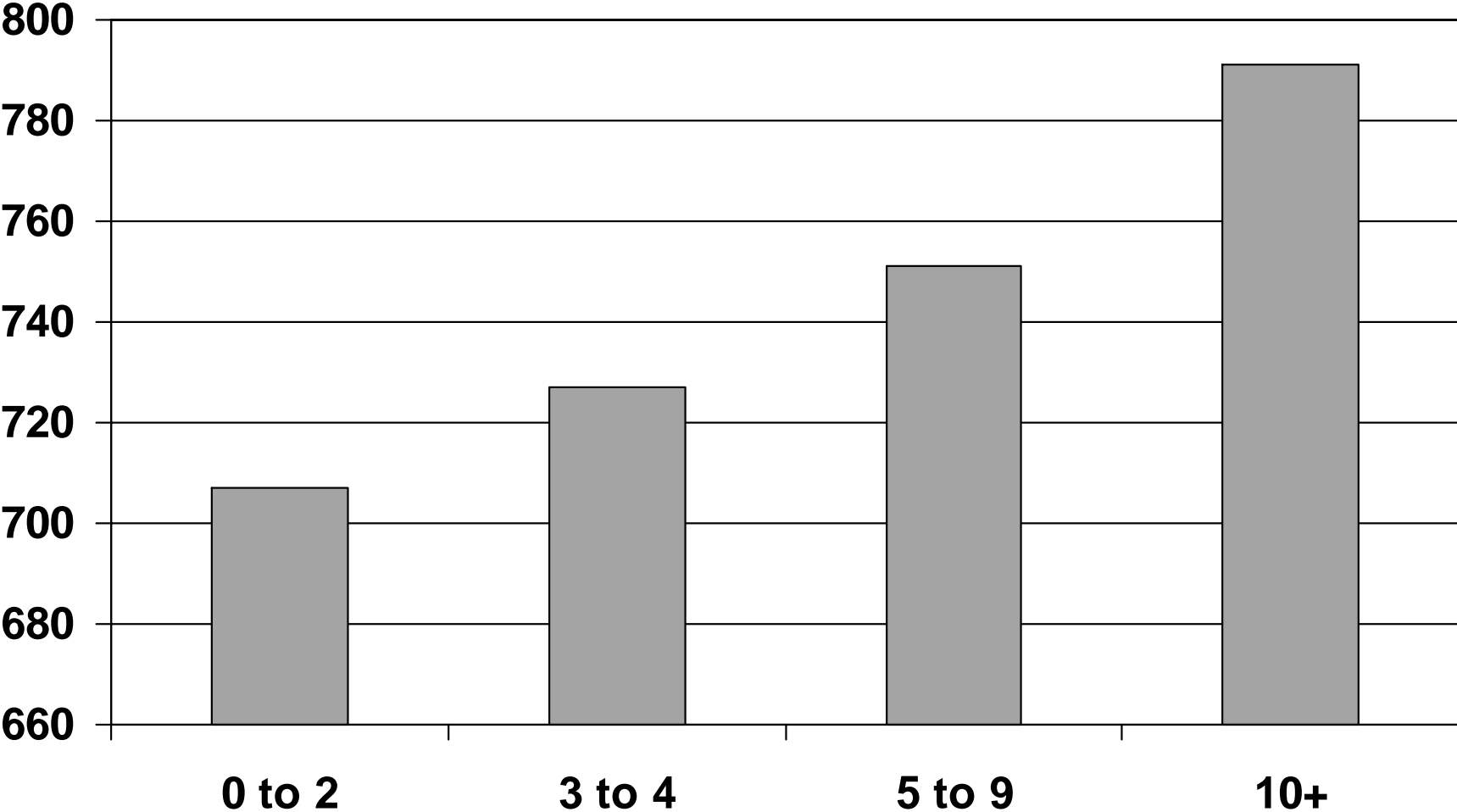
# Average Score by Prior Accidents



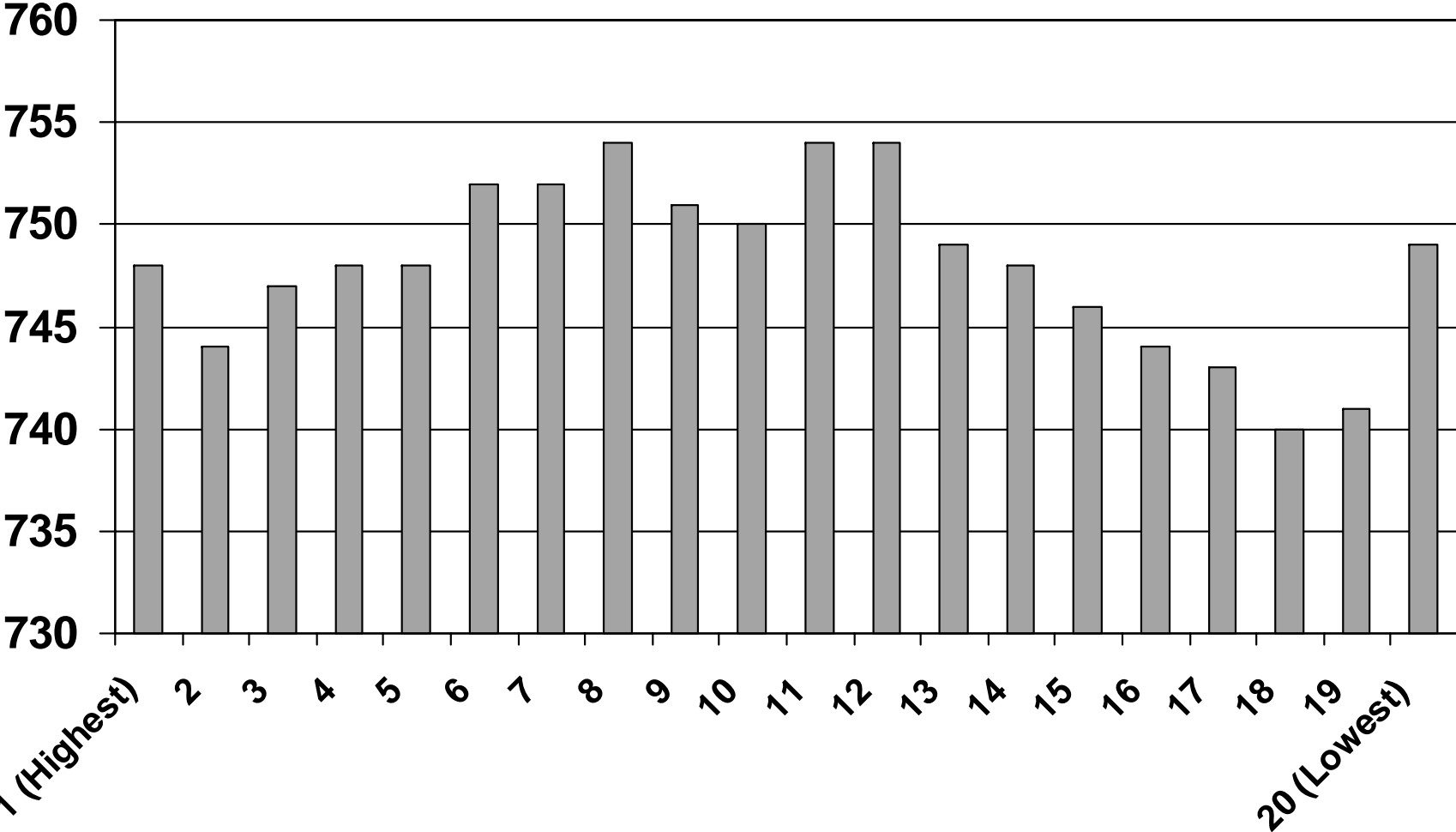
# Average Score by Multiple Car/Policy



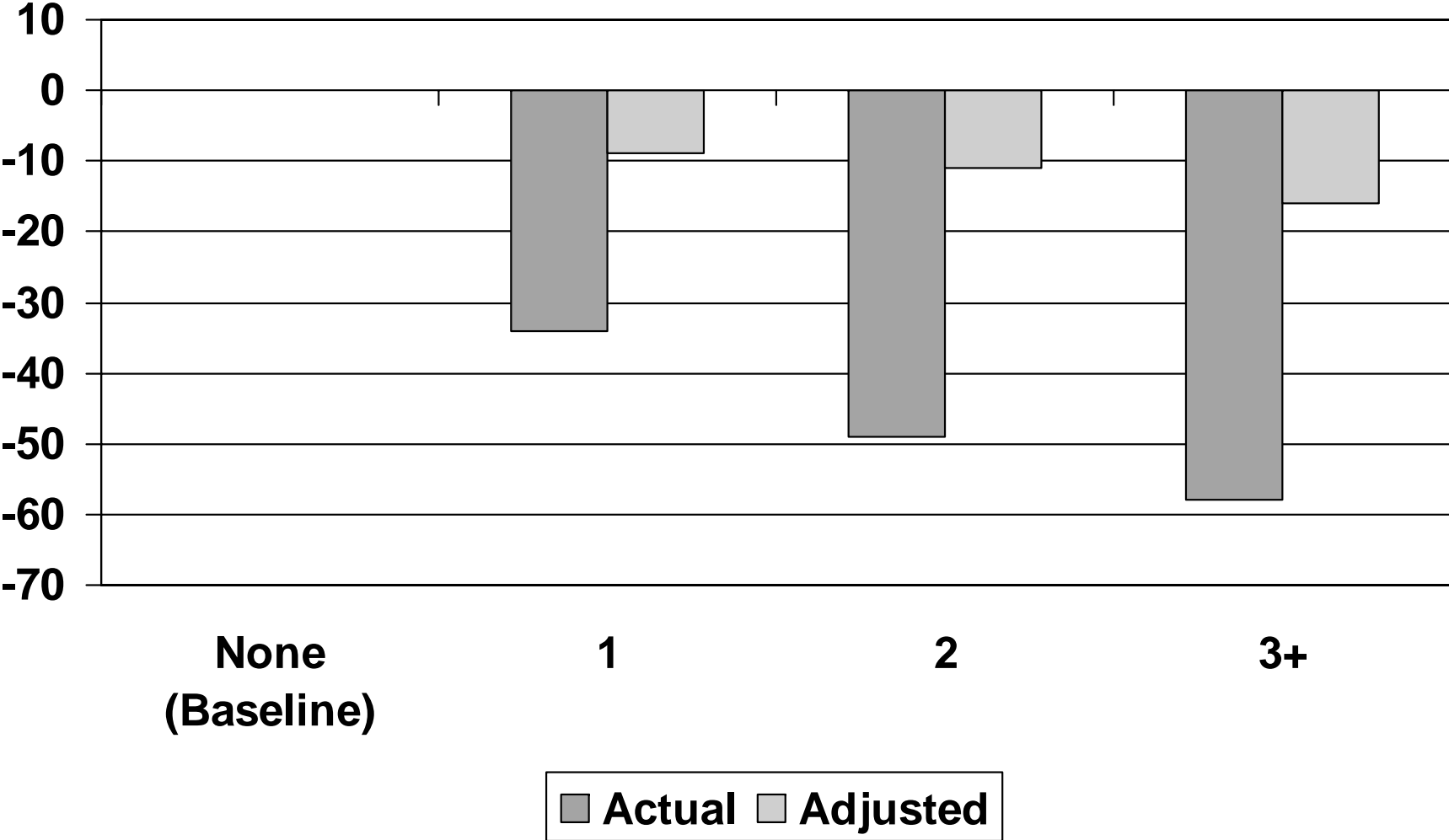
# Average Score by Policy Tenure (Years)



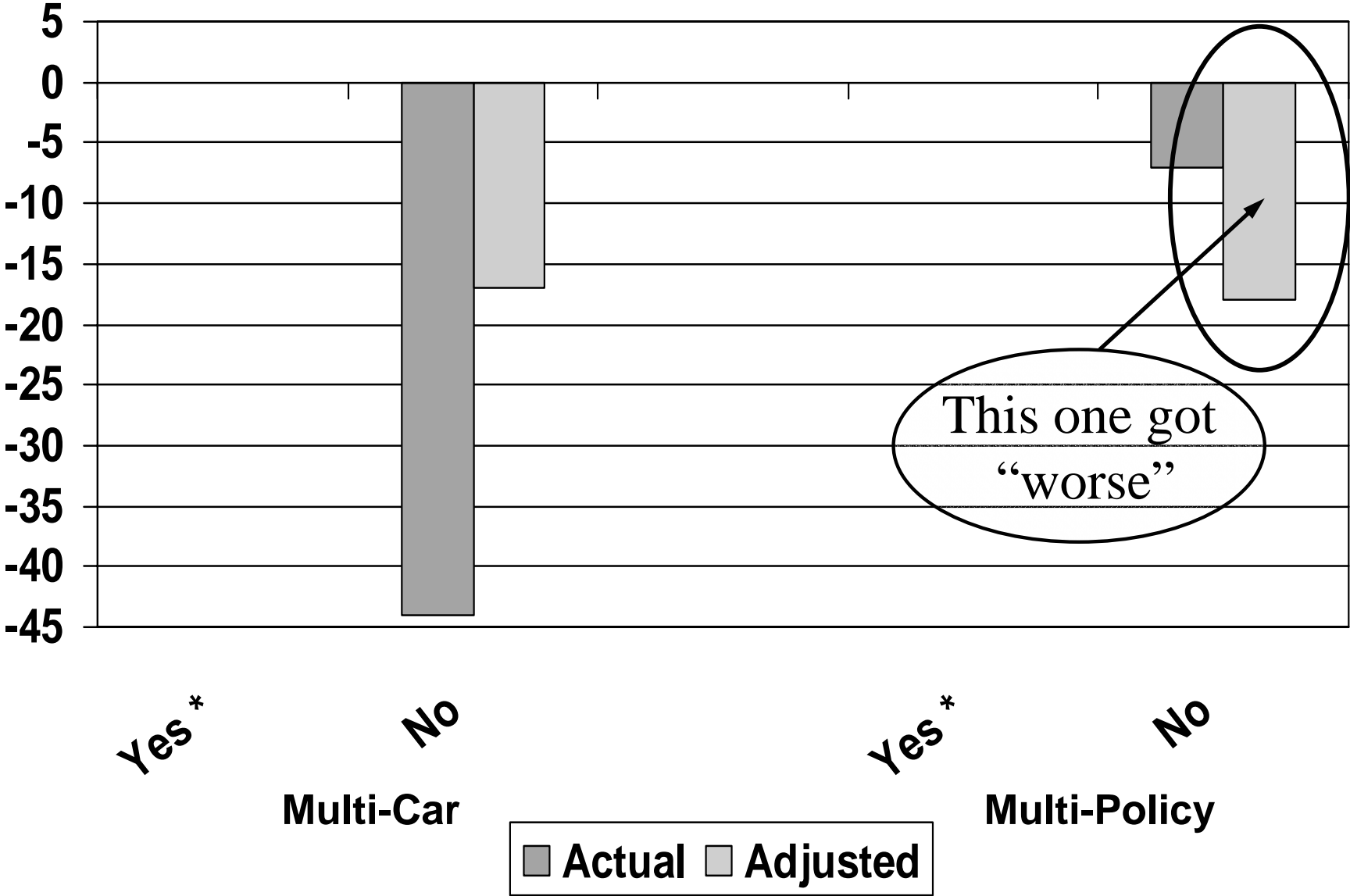
# Average Score by Population Density Group



# Average Score Differences by Prior Accidents

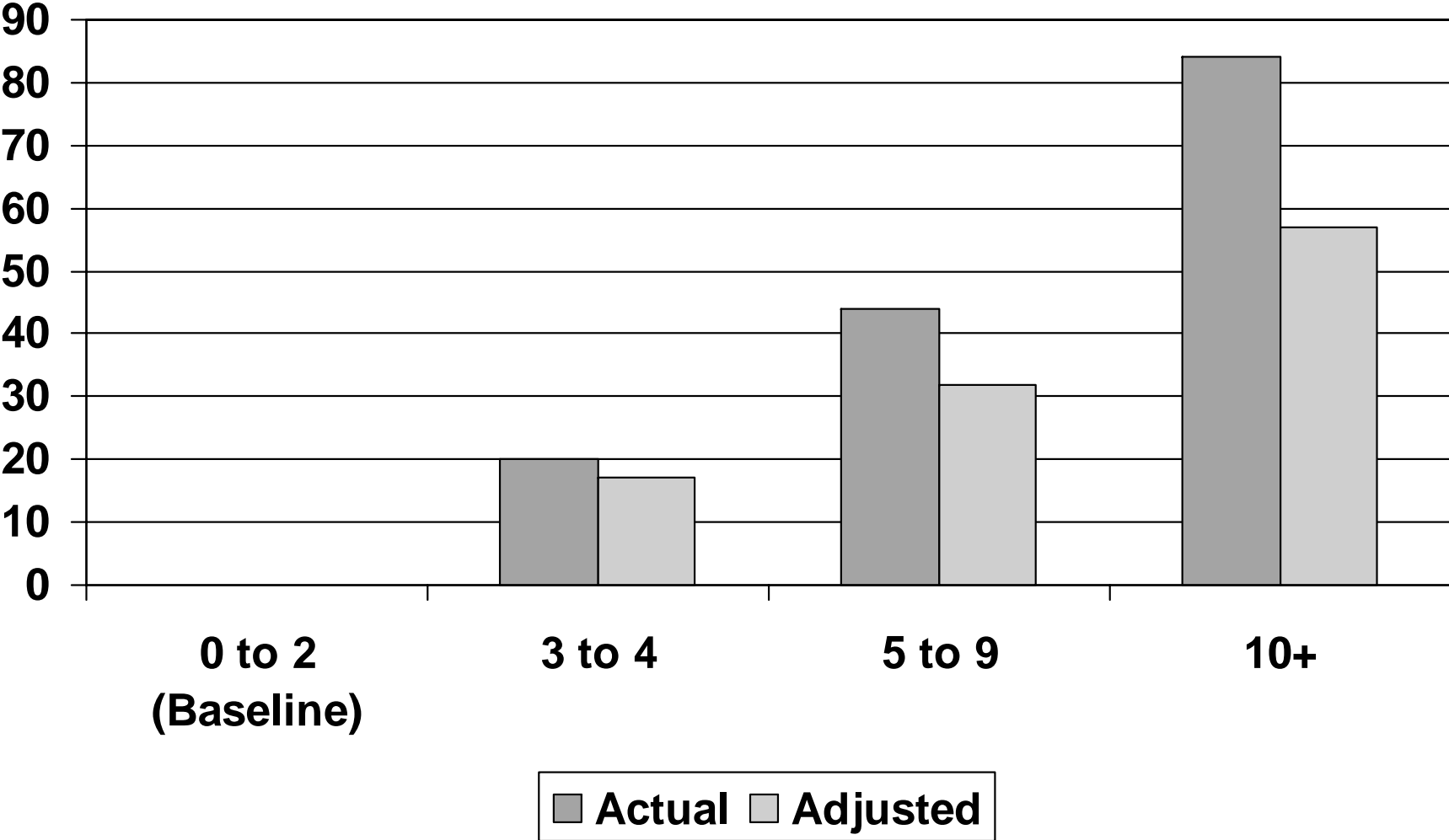


# Average Score Differences by Multiple Car/Policy

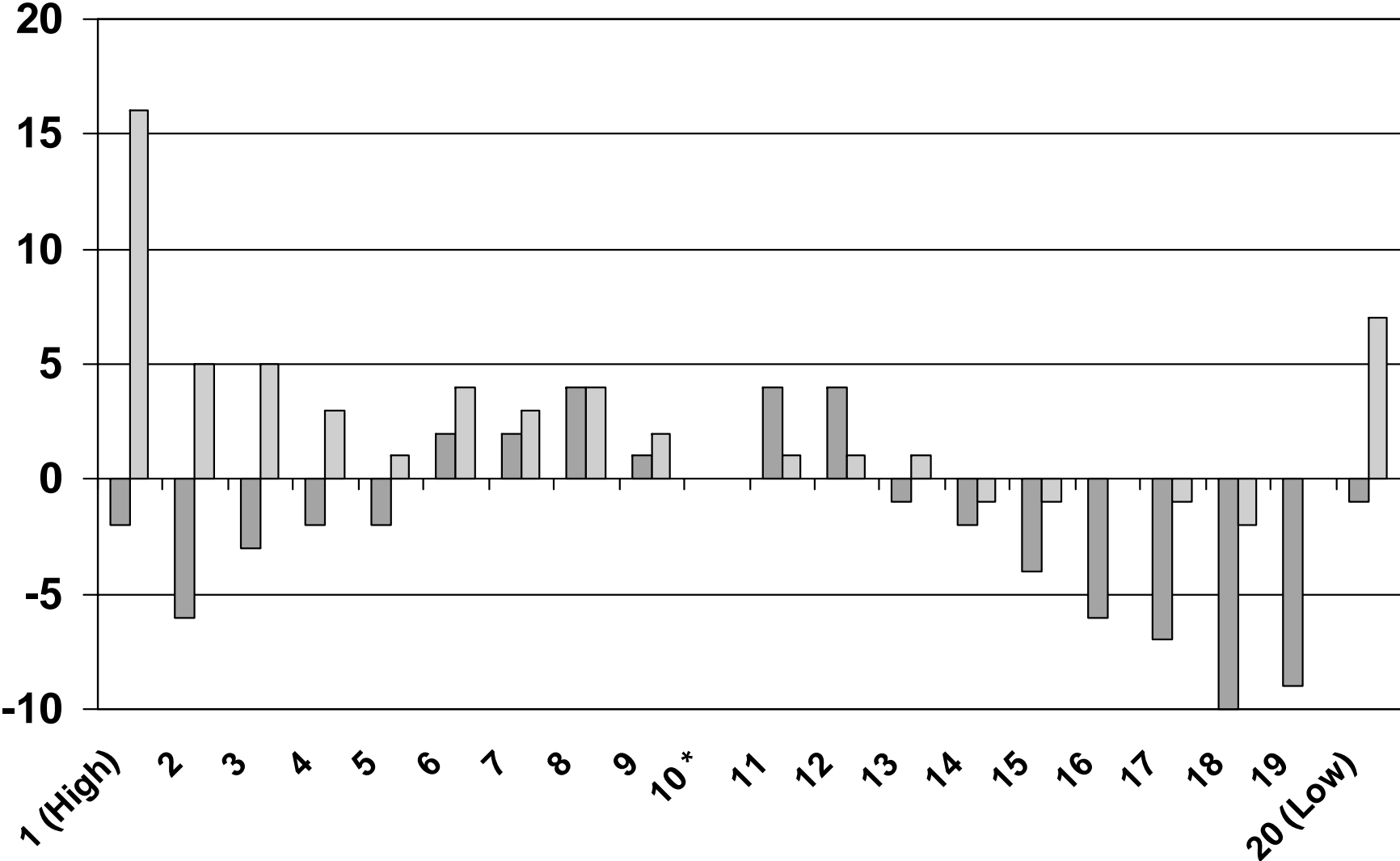


\* Baseline for Comparison

# Average Score Differences by Policy Tenure (Years)



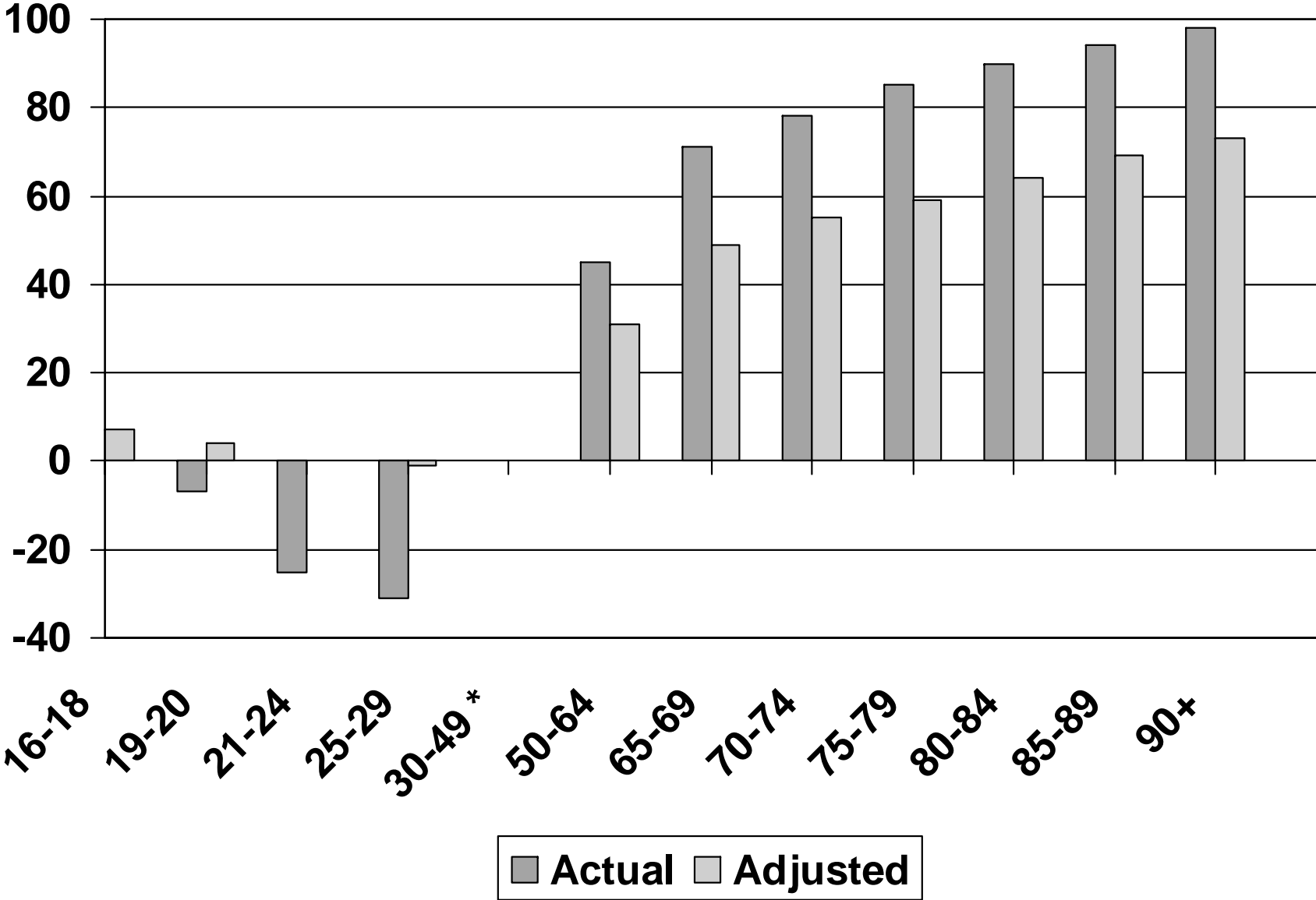
# Average Score Differences by Population Density Group



\* Baseline for Comparison

Actual Adjusted

# Average Score Differences by Driver Age



\* Baseline for Comparison

# Summary

- Missouri study likely to be (has already been?) misused
  - Impact on affordability/availability not studied
  - Use of percentiles likely to mislead readers
- Data call design introduces biases
- Statistical issues
  - Univariate study based on voting-rights court decision
  - Missing important variables bias results
  - Low  $R^2$  even on aggregated data
- EPIC critique offers examples of how these problems can affect this type of analysis

# Conclusions

- Issue like to remain contentious
- States continue to pursue their own studies
- FTD/FED study will hopefully bring relief