Predictive Power of External Data in Pricing: Credit and Beyond

Concurrent Session PL-1
2004 CAS Ratemaking Seminar
Richard A. Smith
EPIC Consulting, LLC

Informal poll

- How many are directly involved in analyzing credit-based insurance scores?
- How many are indirectly involved? (e.g., work with legal and/or public affairs)
- How many believe that credit-based insurance scores are <u>not</u> related to expected loss?

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
- Focused on private passenger auto
- Purpose was to answer three questions
  - Correlation
  - Overlap / Interaction
  - Business purpose / Importance

Data for Study

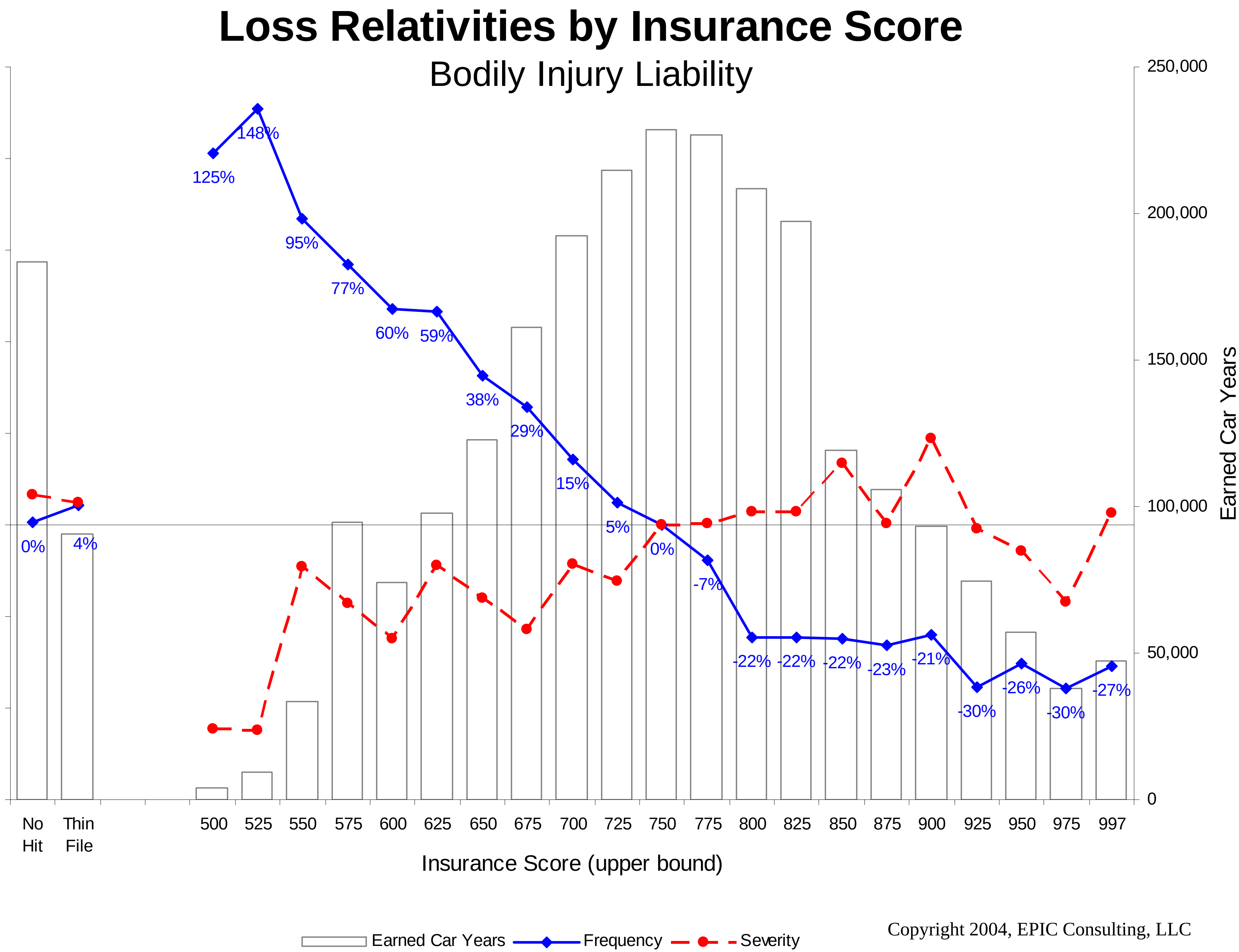
- Provided by multiple companies
- Distribution of business designed to be roughly similar to actual U.S. population distribution
- Random sample
- Approximately 2.7 million earned car years
- Six major personal auto coverages

External Data

- Needed surrogate for territories
  - Census data candidates
    - Population density
    - Vehicle density
    - Considered several others
- Wanted surrogate for symbols
  - Vendors: MSRP

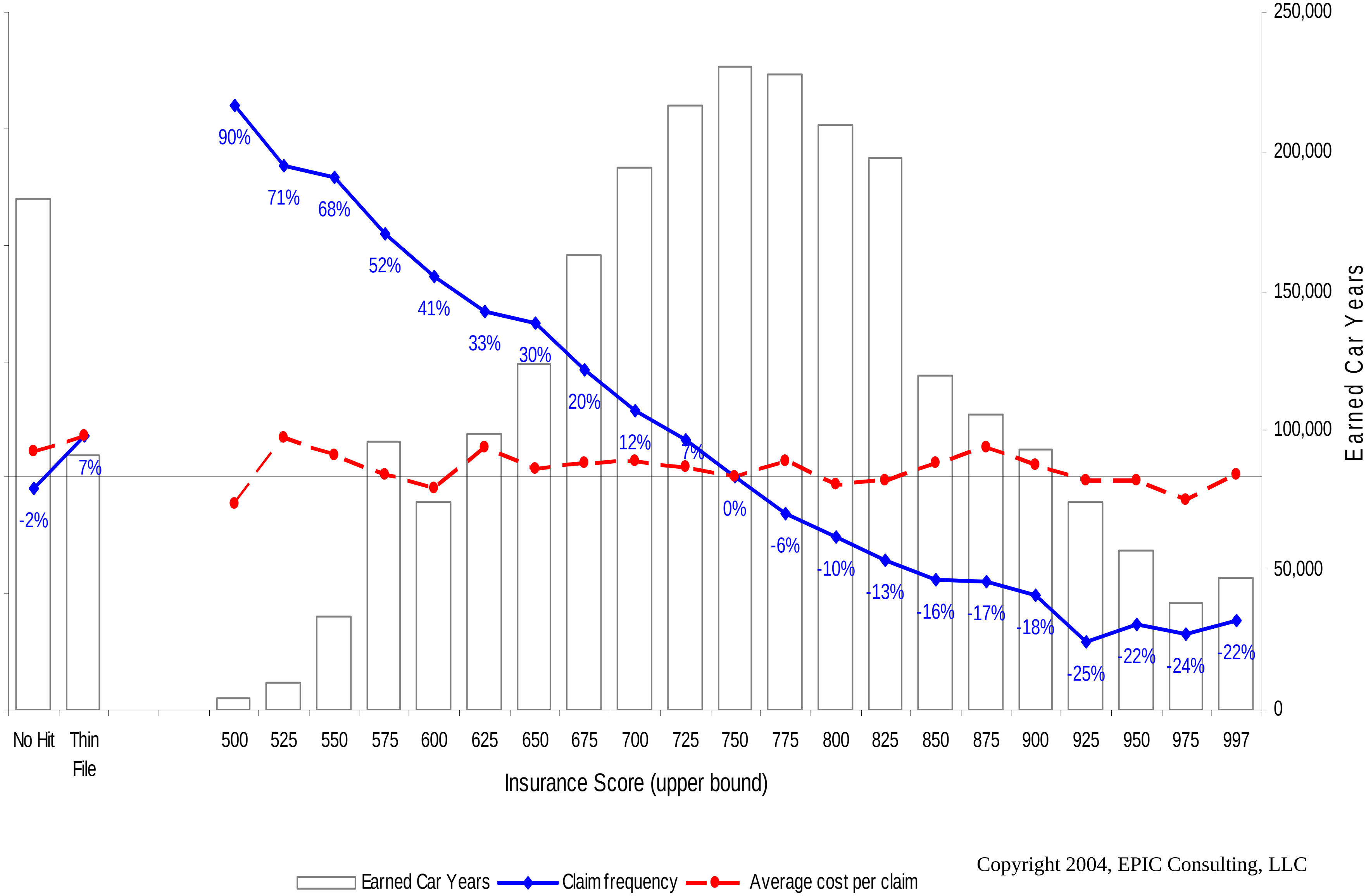
Overview: Non-Modeled Results

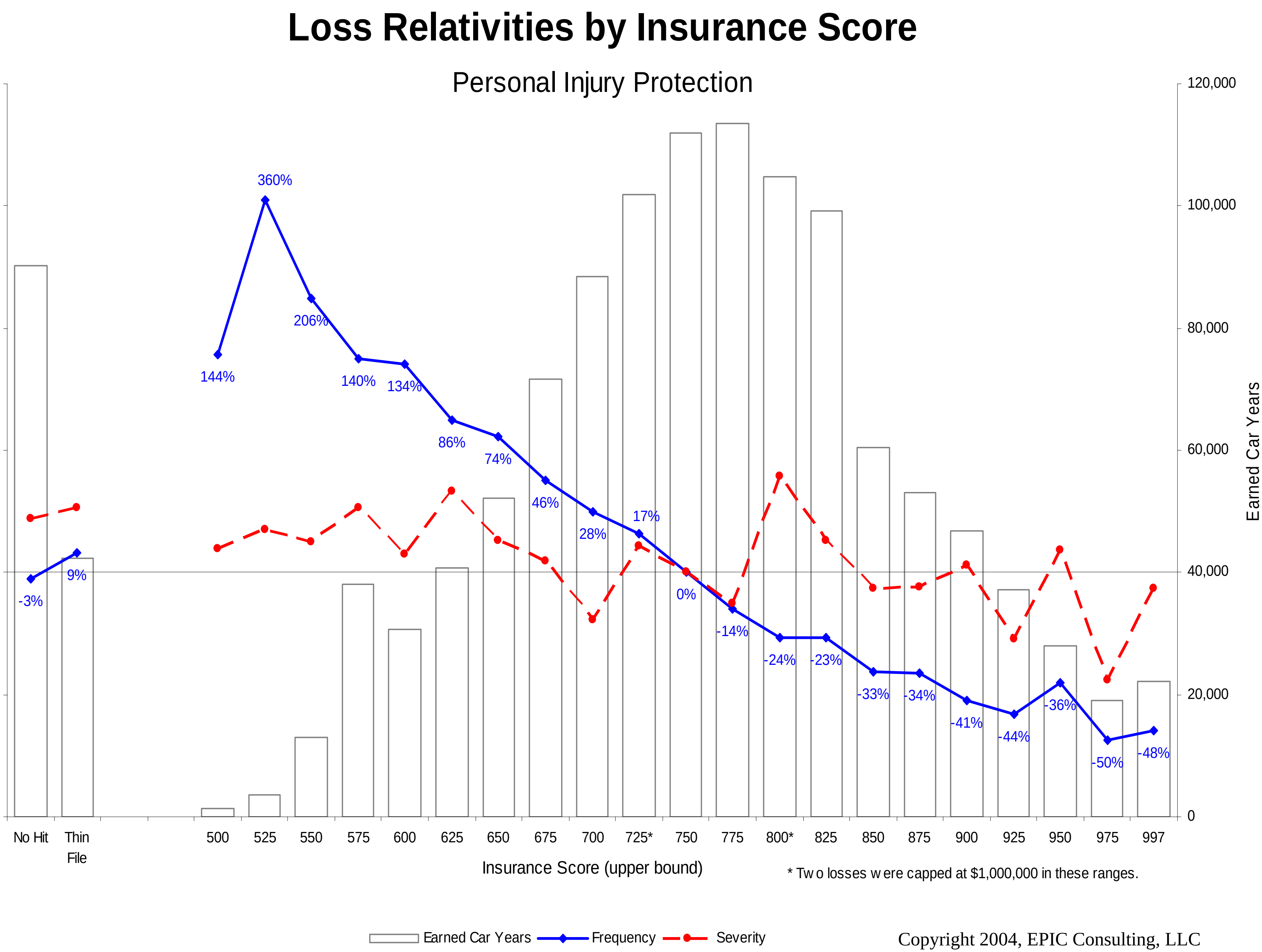
- Claim frequency primary driver of cost differences
- Frequency relationships consistent for all coverages
- Claim severity differences for some coverages
  - Collision and comprehensive: poor scores suggest higher severities
  - Bodily injury: modest reversal: poor scores suggest lower severities. This becomes insignificant during modeling process


Loss Relativities by Insurance Score
Bodily Injury Liability
148%
125%
95%
77%
60%
59%
38%
29%
15%
5%
0%
4%
0%
-7%
-22%
-22%
-22%
-23%
-21%
-26%
-30%
-30%
-27%
250,000
200,000
150,000
100,000
50,000
0
Earned Car Years
No Hit
Thin File
500
525
550
575
600
625
650
675
700
725
750
775
800
825
850
875
900
925
950
975
997
Insurance Score (upper bound)
Earned Car Years
Frequency
Severity

Loss Relativities by Insurance Score

Property Damage Liability

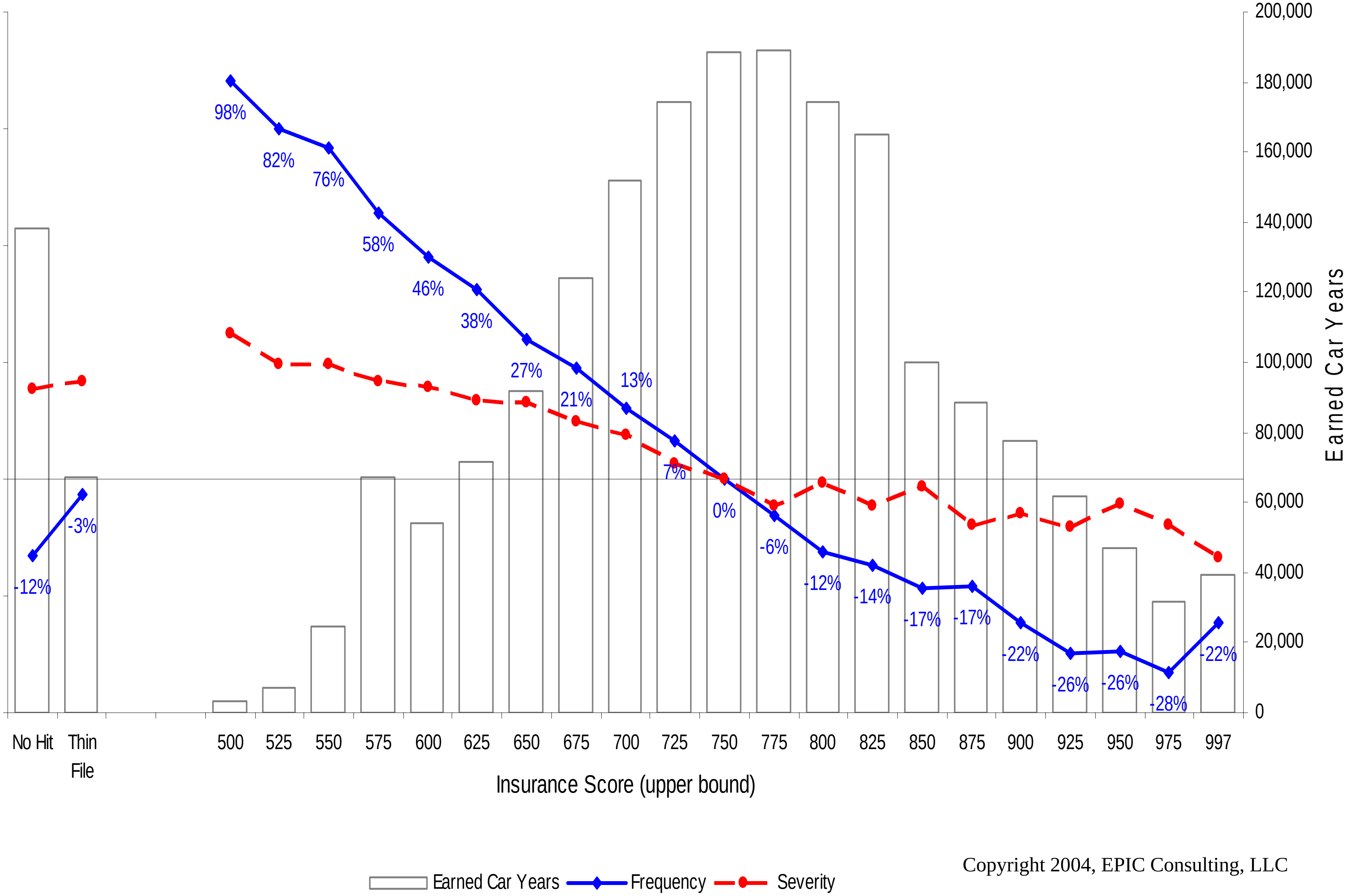


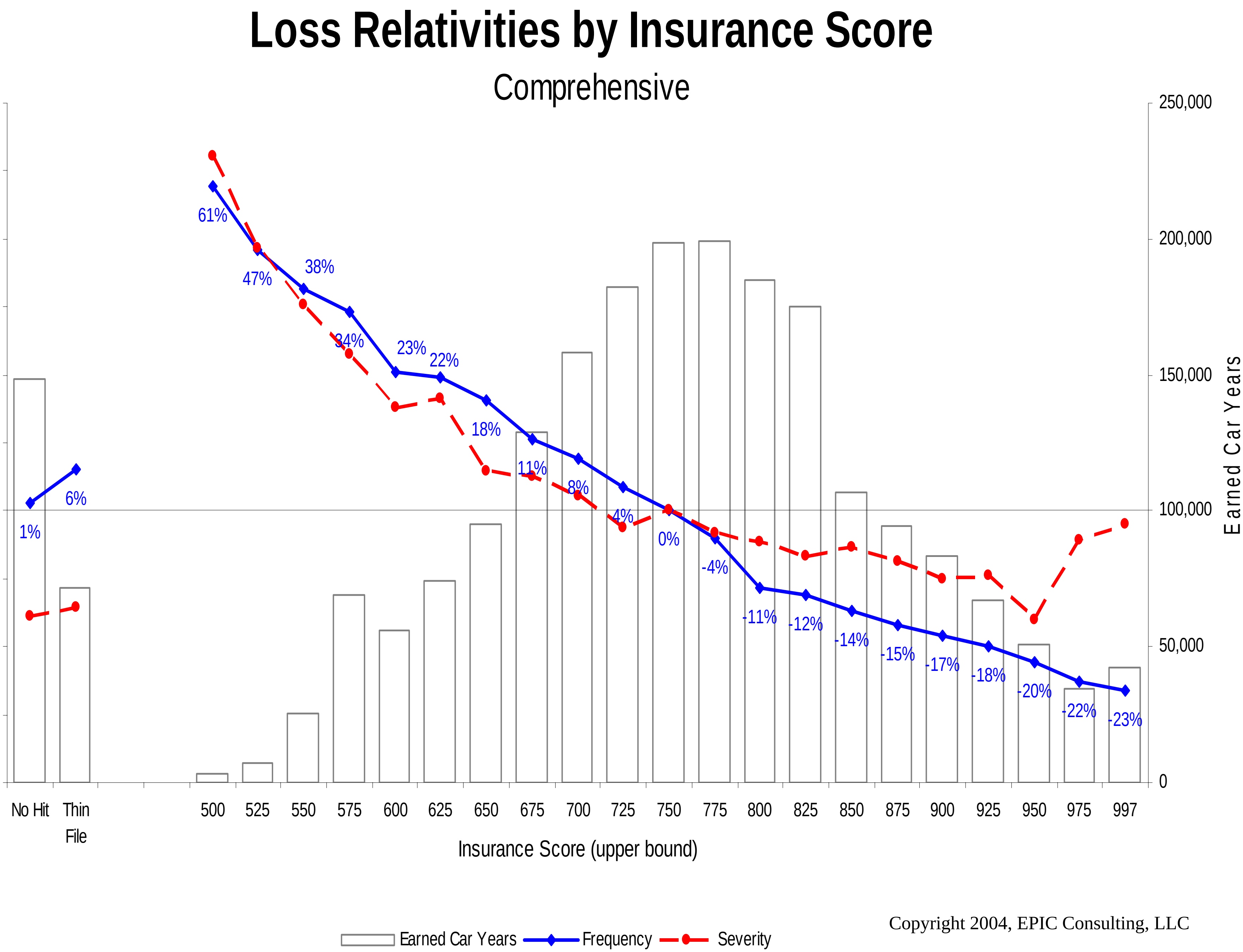


Loss Relativities by Insurance Score
Personal Injury Protection
120,000
100,000
80,000
60,000
40,000
20,000
0
Earned Car Years
360%
206%
144%
140%
134%
86%
74%
46%
28%
17%
9%
-3%
0%
-14%
-24%
-23%
-33%
-34%
-41%
-44%
-36%
-50%
-48%
No Hit
Thin File
500
525
550
575
600
625
650
675
700
725*
750
775
800*
825
850
875
900
925
950
975
997
Insurance Score (upper bound)
* Two losses were capped at $1,000,000 in these ranges.
Earned Car Years
Frequency
Severity

Loss Relativities by Insurance Score

Collision





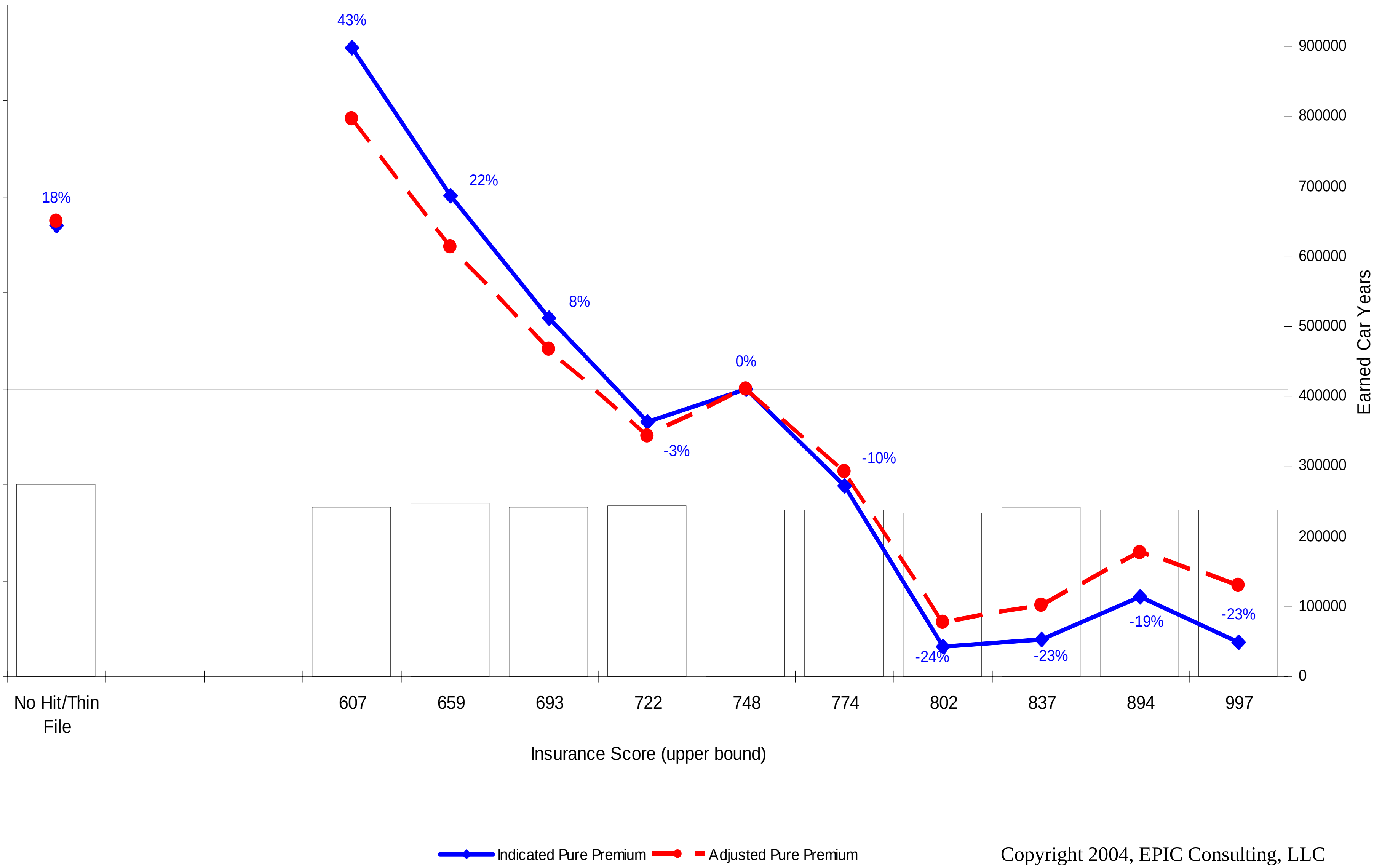
Loss Relativities by Insurance Score
Comprehensive
250,000
200,000
150,000
100,000
50,000
0
Earned Car Years
61%
47%
38%
34%
23%
22%
18%
11%
8%
4%
0%
-4%
-11%
-12%
-14%
-15%
-17%
-18%
-20%
-22%
-23%
1%
6%
No Hit
Thin File
500
525
550
575
600
625
650
675
700
725
750
775
800
825
850
875
900
925
950
975
997
Insurance Score (upper bound)
Earned Car Years
Frequency
Severity

Correlation and “Overlap”

- Modeled frequency and severity, then combined to yield pure premiums
- Included the usual “suspects”:
  - Age/gender, limits, deductibles, accidents/violations, multi-car, multi-line, tenure, model year, vehicle use, state
  - Population density as surrogate for territory
- Overlap / Interaction
  - Modeled all characteristics excluding scores
  - “Froze” those factors, and modeled the remaining variation with insurance scores

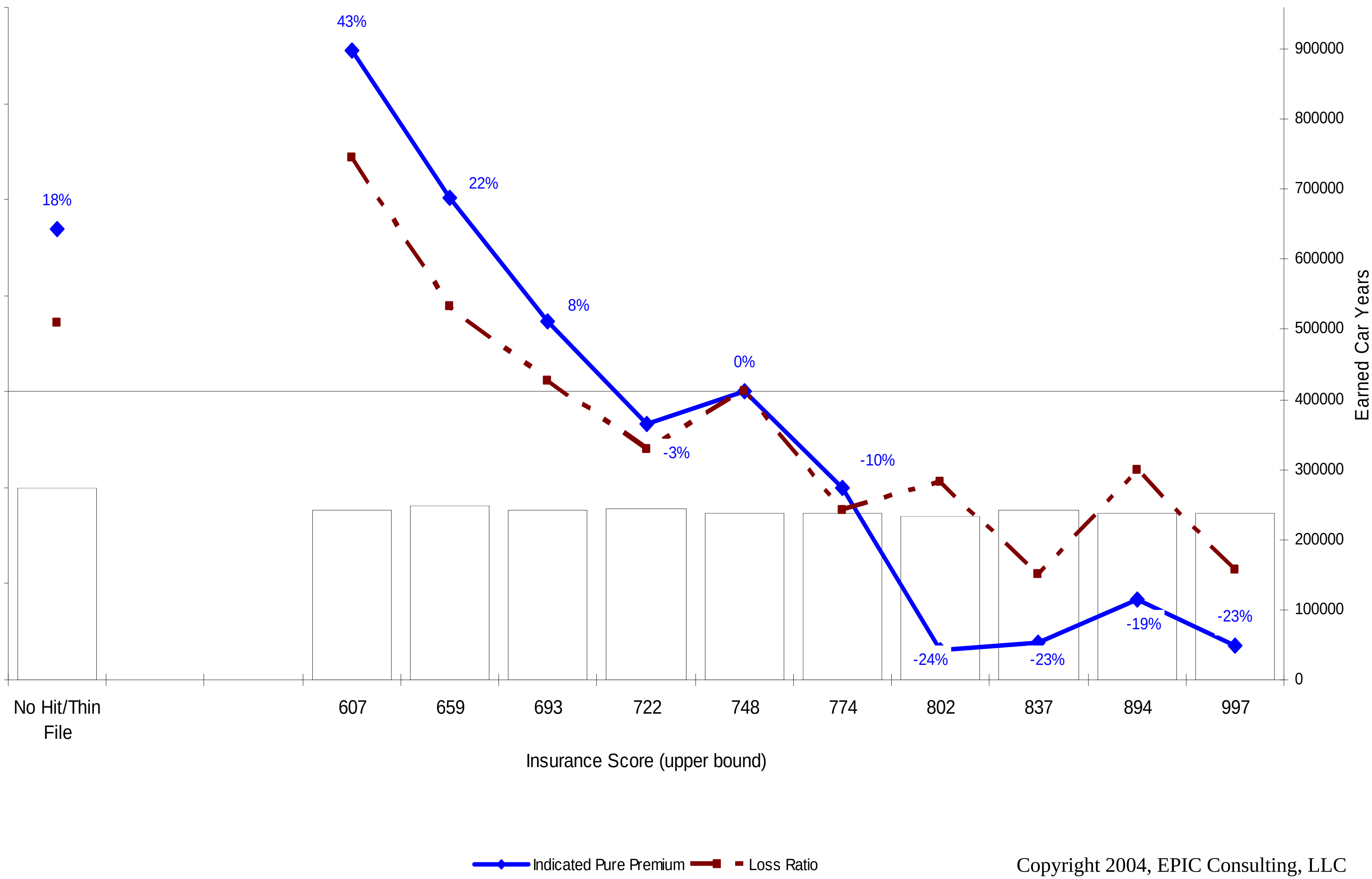
Indicated Relative Pure Premium by Insurance Score

Bodily Injury Liability



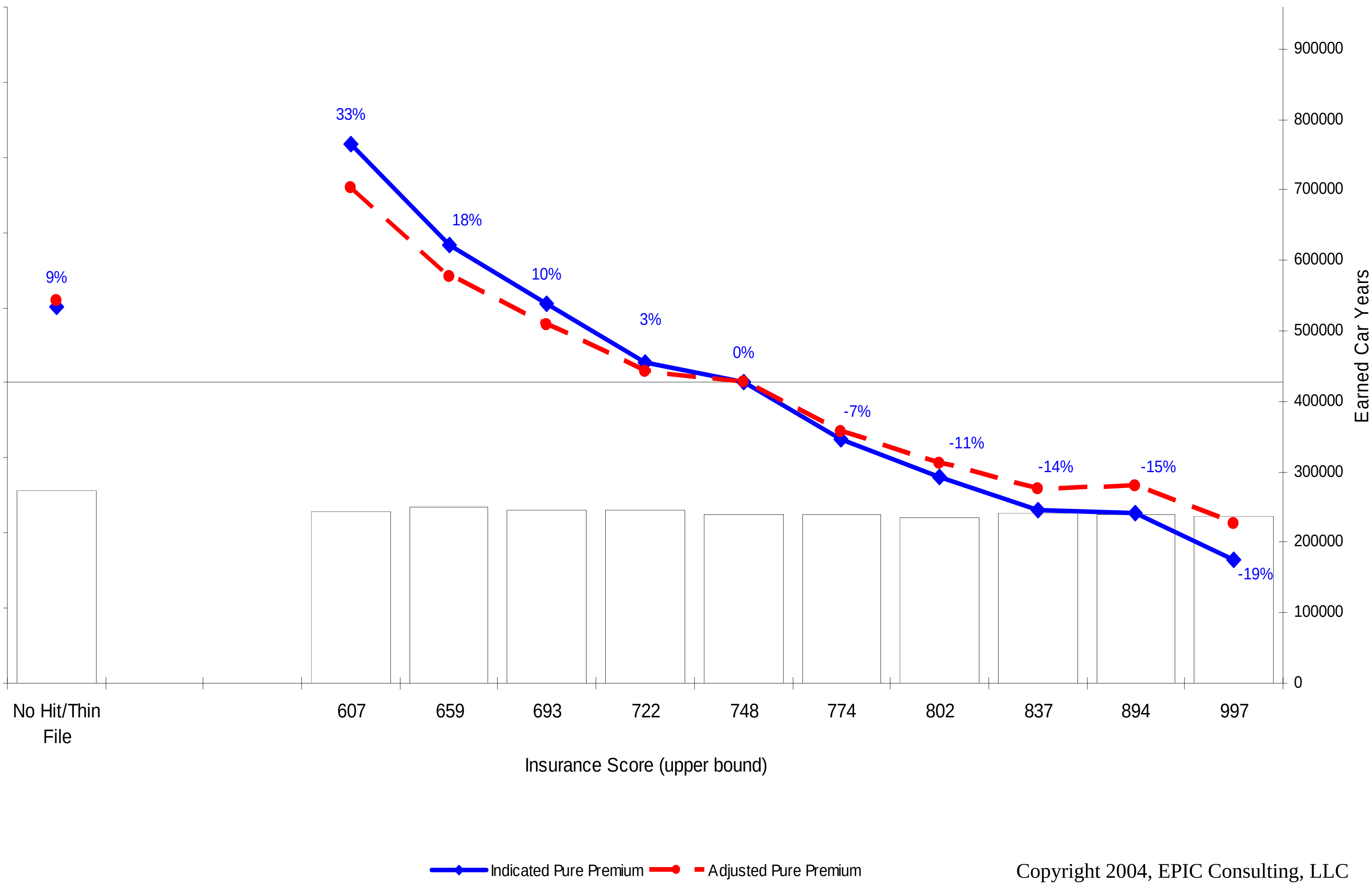
Indicated Relative Pure Premium by Insurance Score

Bodily Injury Liability



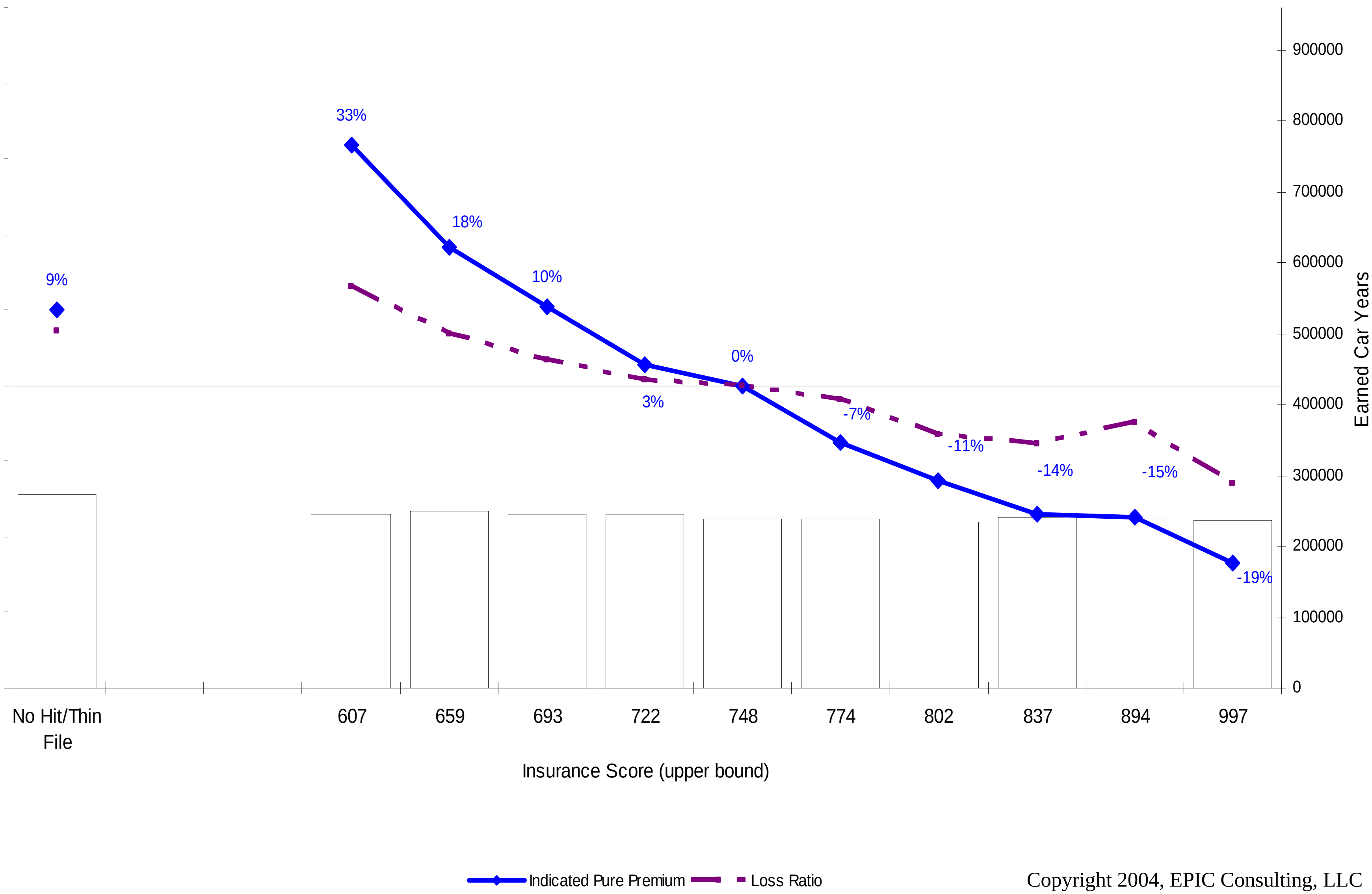
Indicated Relative Pure Premium by Insurance Score

Property Damage Liability

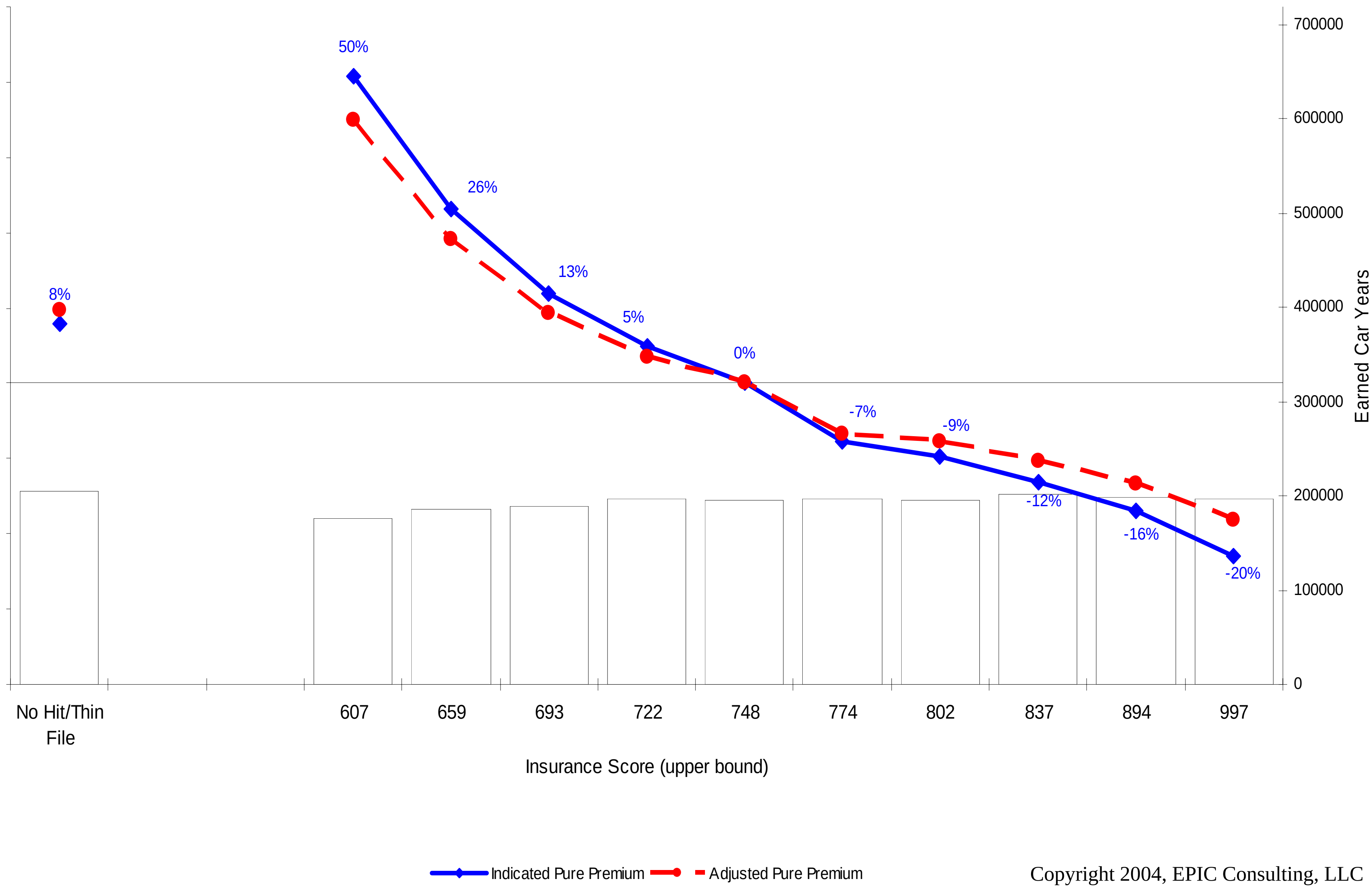


Indicated Relative Pure Premium by Insurance Score

Property Damage Liability

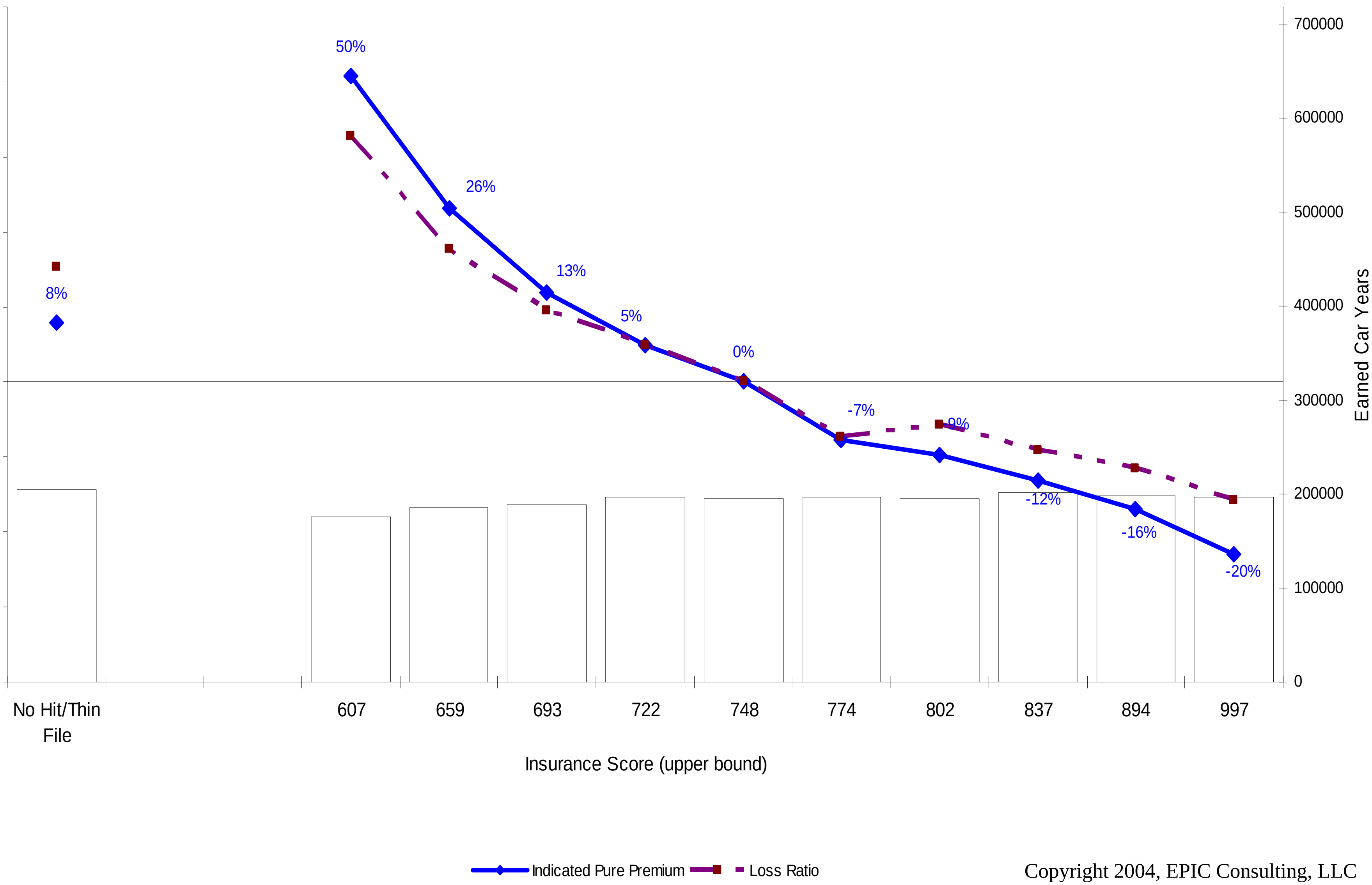


Indicated Relative Pure Premium by Insurance Score
Collision



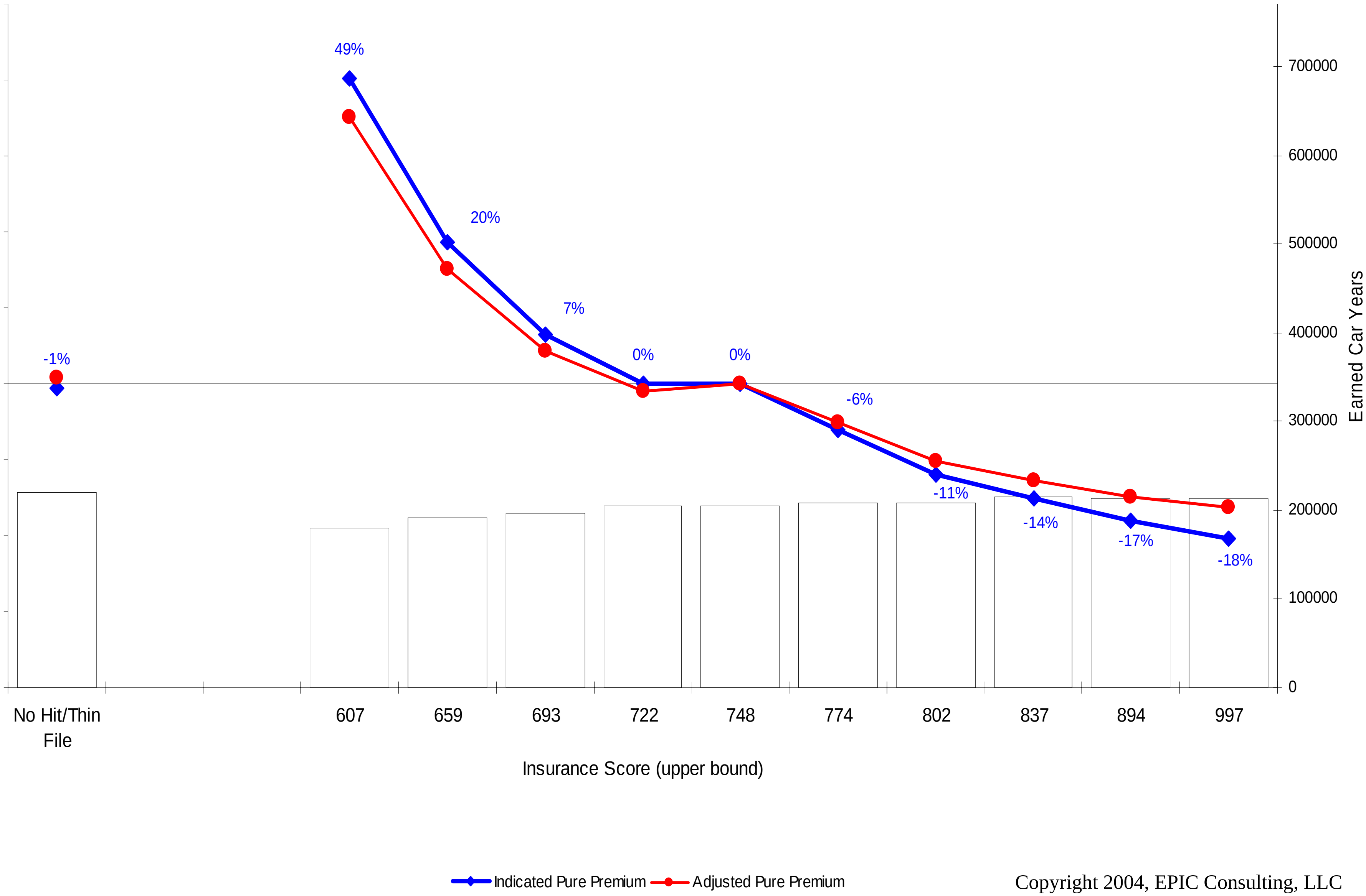
Indicated Relative Pure Premium by Insurance Score

Collision



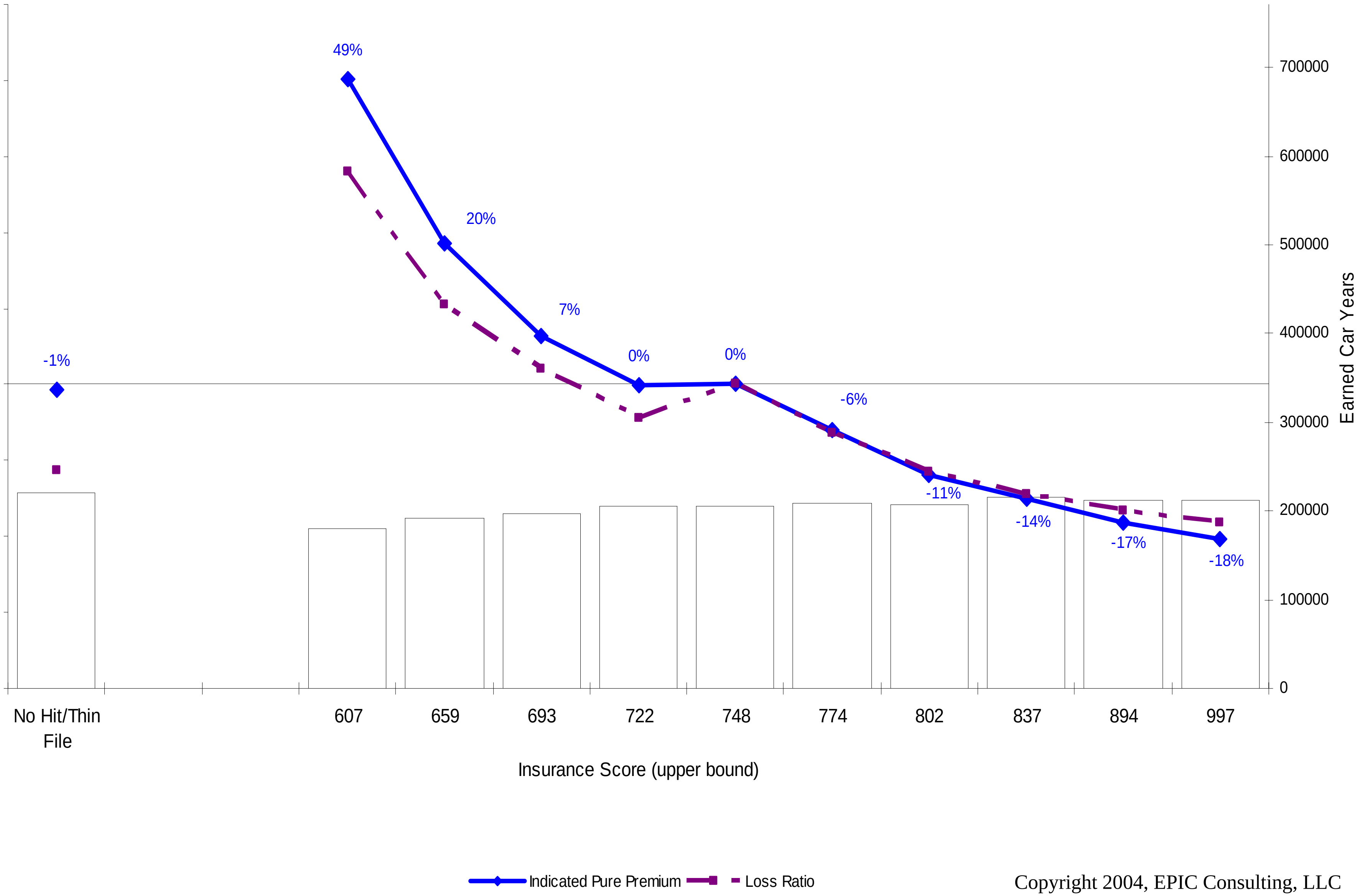
Indicated Relative Pure Premium by Insurance Score

Comprehensive



Indicated Relative Pure Premium by Insurance Score

Comprehensive



Importance of Insurance Scores*

Coverage	Factor 1	Factor 2	Factor 3
BI Liability	Age/Gender	Ins. Score	Pop. Density
PD Liability	Age/Gender	Model Year	Ins. Score
Pers. Inj. Prot.	Ins. Score	Pop. Density	Yrs. Insured
Med Pay	Ins. Score	Limit	Age/Gender
Comprehensive	Model Year	Age/Gender	Ins. Score
Collision	Model Year	Age/Gender	Ins. Score

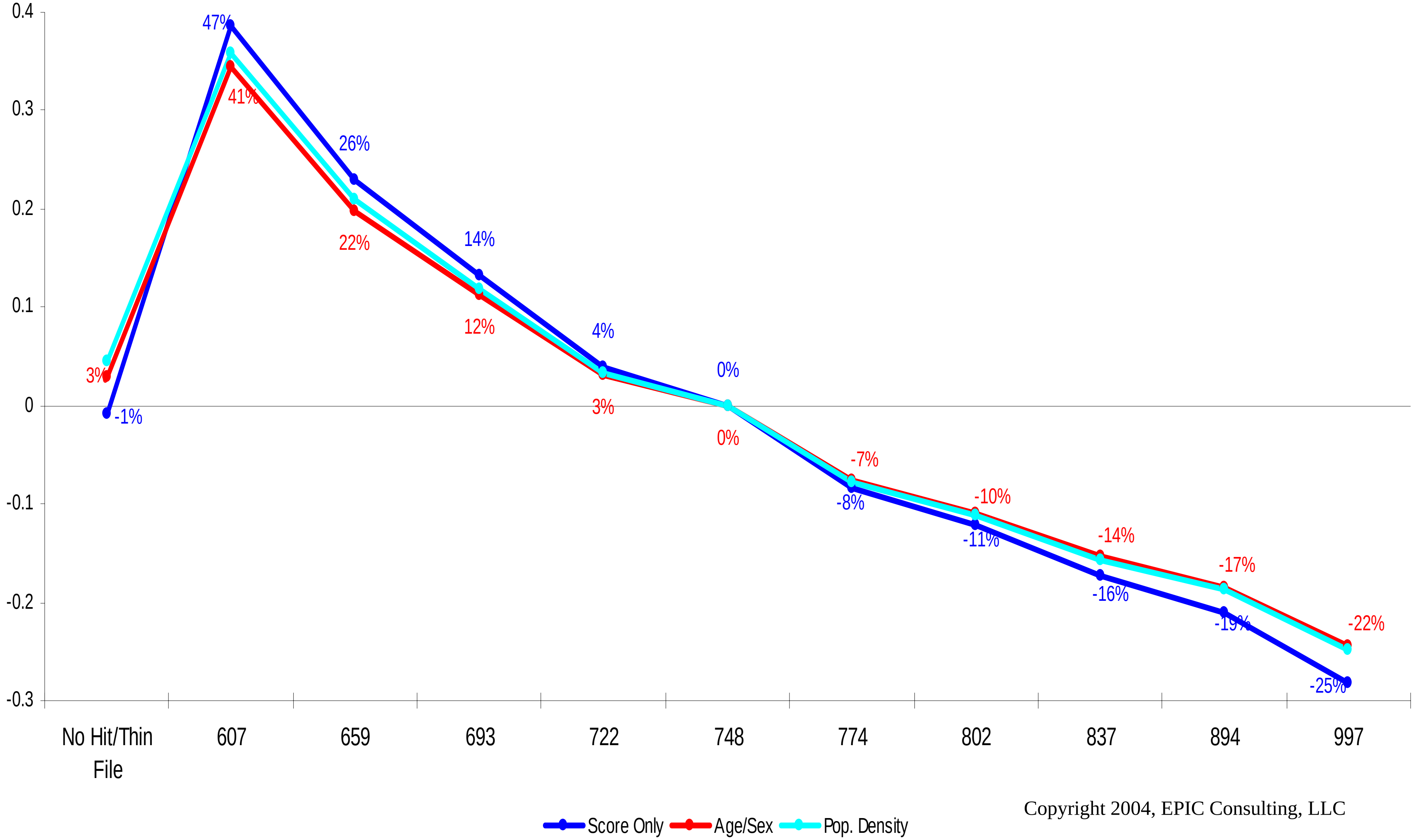
* Based on average absolute differences

Back to External Data

- Use to supplement, not replace
- Census variables can act as surrogates for some traditional characteristics
  - Percentage of multi-car penetration / zip
  - Percentage of home ownership / zip
  - Age distributions by zip
- Ultimately chose to keep things simple
  - Addition of variables had declining impact on insurance score factors

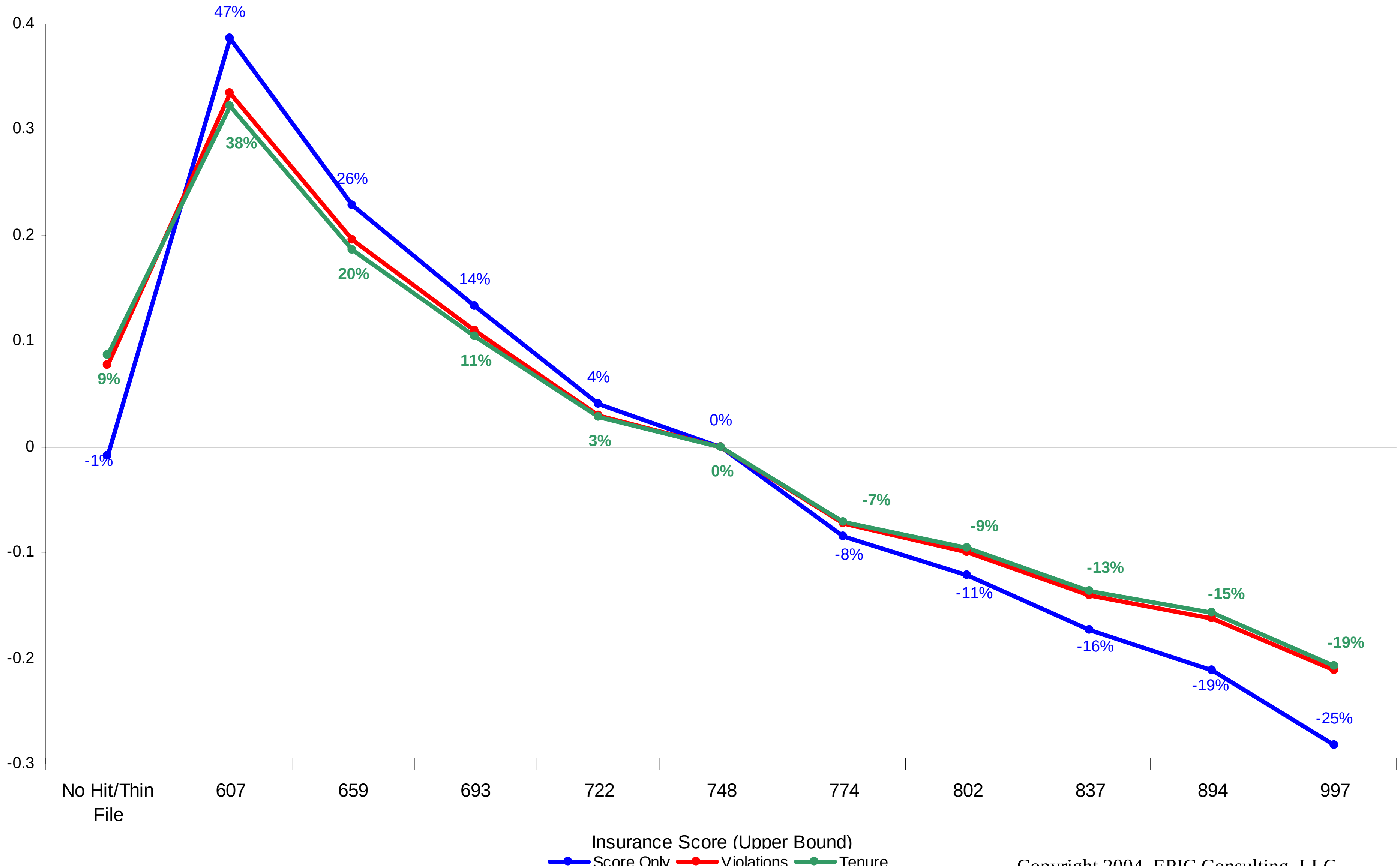
Impact on Indicated PD Frequency Score Relativities

From Introducing Additional Variables



Impact on Indicated PD Frequency Score Relativities

From Introducing Additional Variables

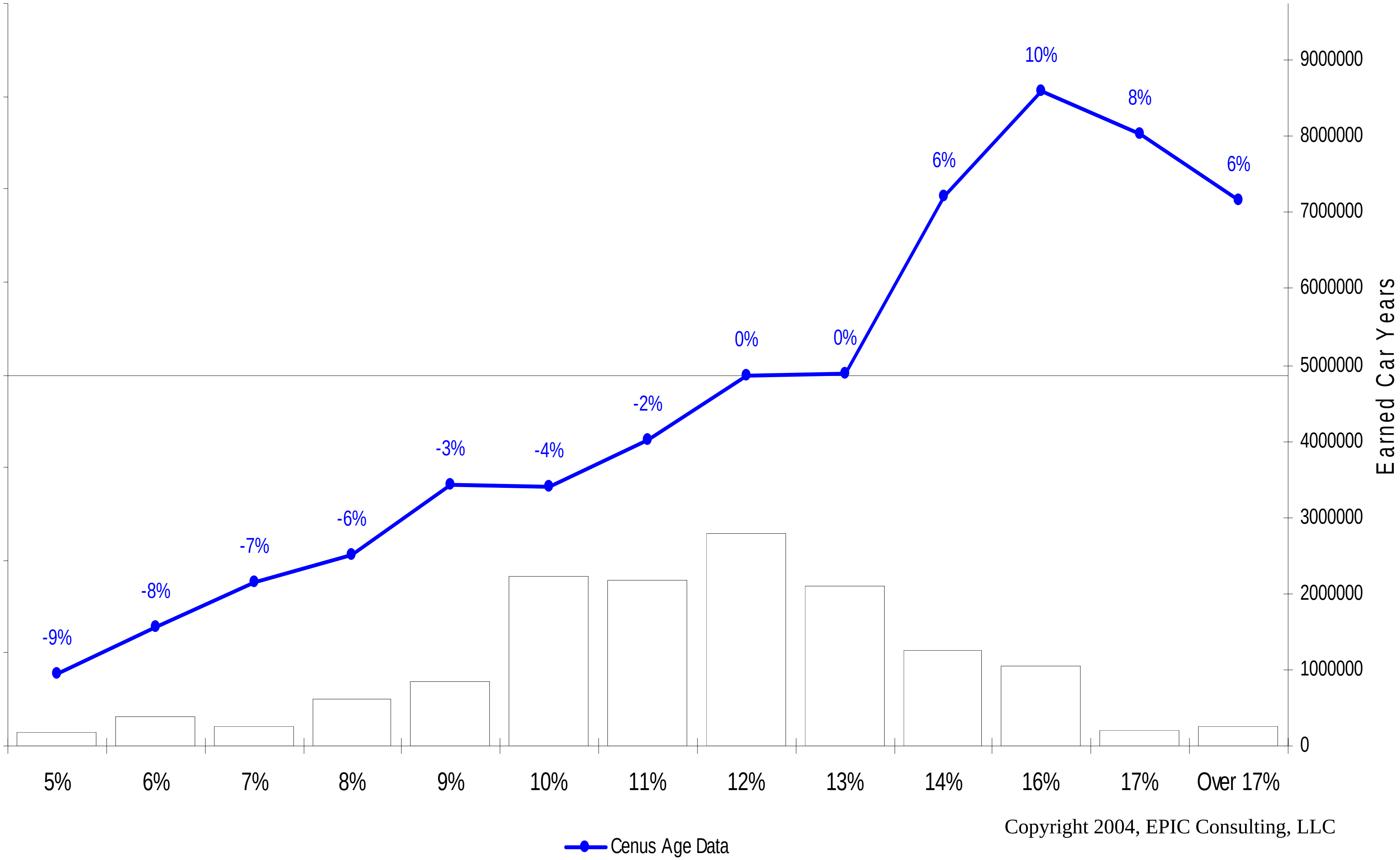


Example - Texas Collision Frequencies

- Modeled territory versus
  - Driver classification
  - Census age range statistics
- External data is predictive
- Census age factors not as effective as actual information
- Not surprisingly, indicated territorial factors differ, in some cases substantially

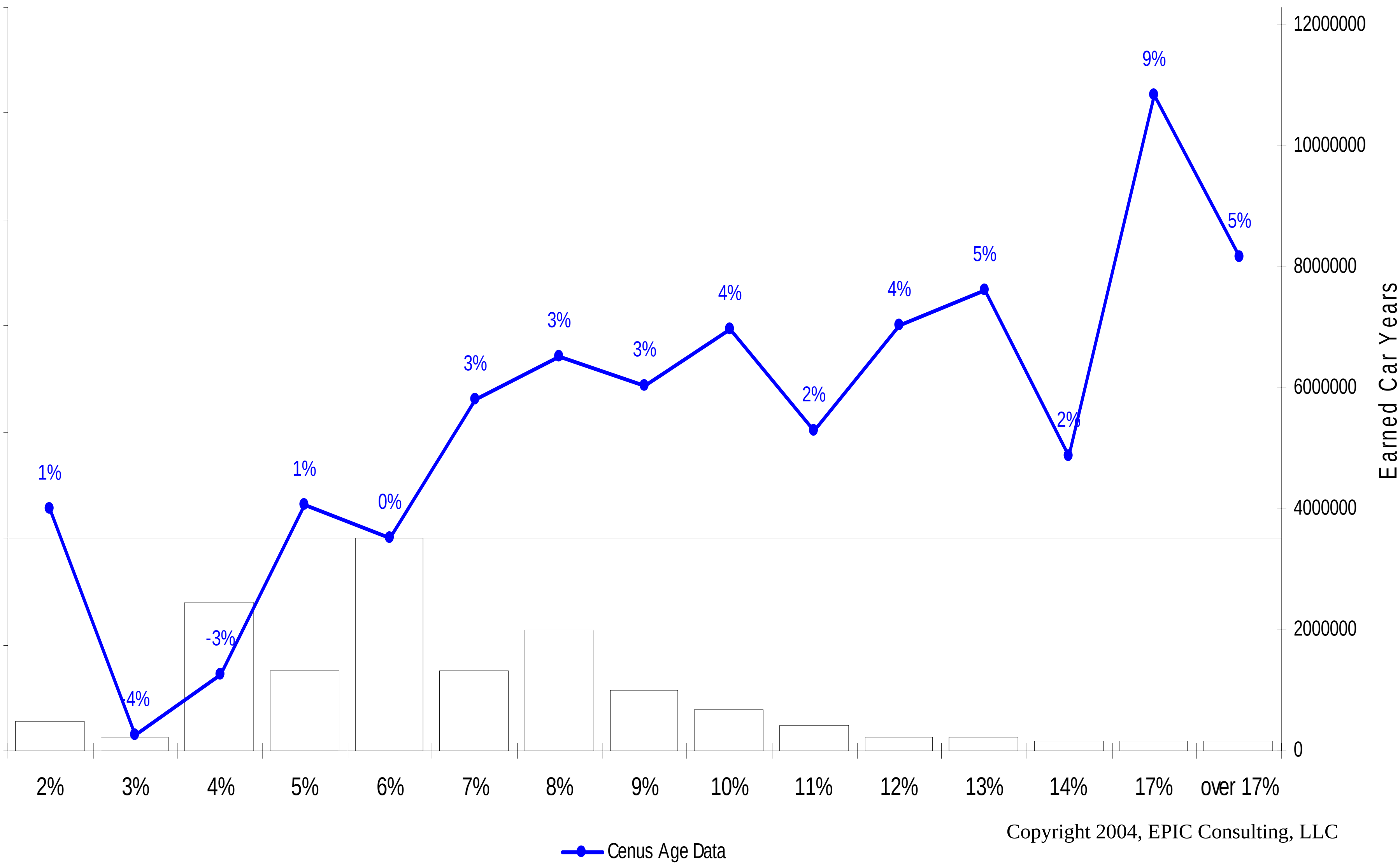
External Data Illustration

Percent of ZIP aged 15-20



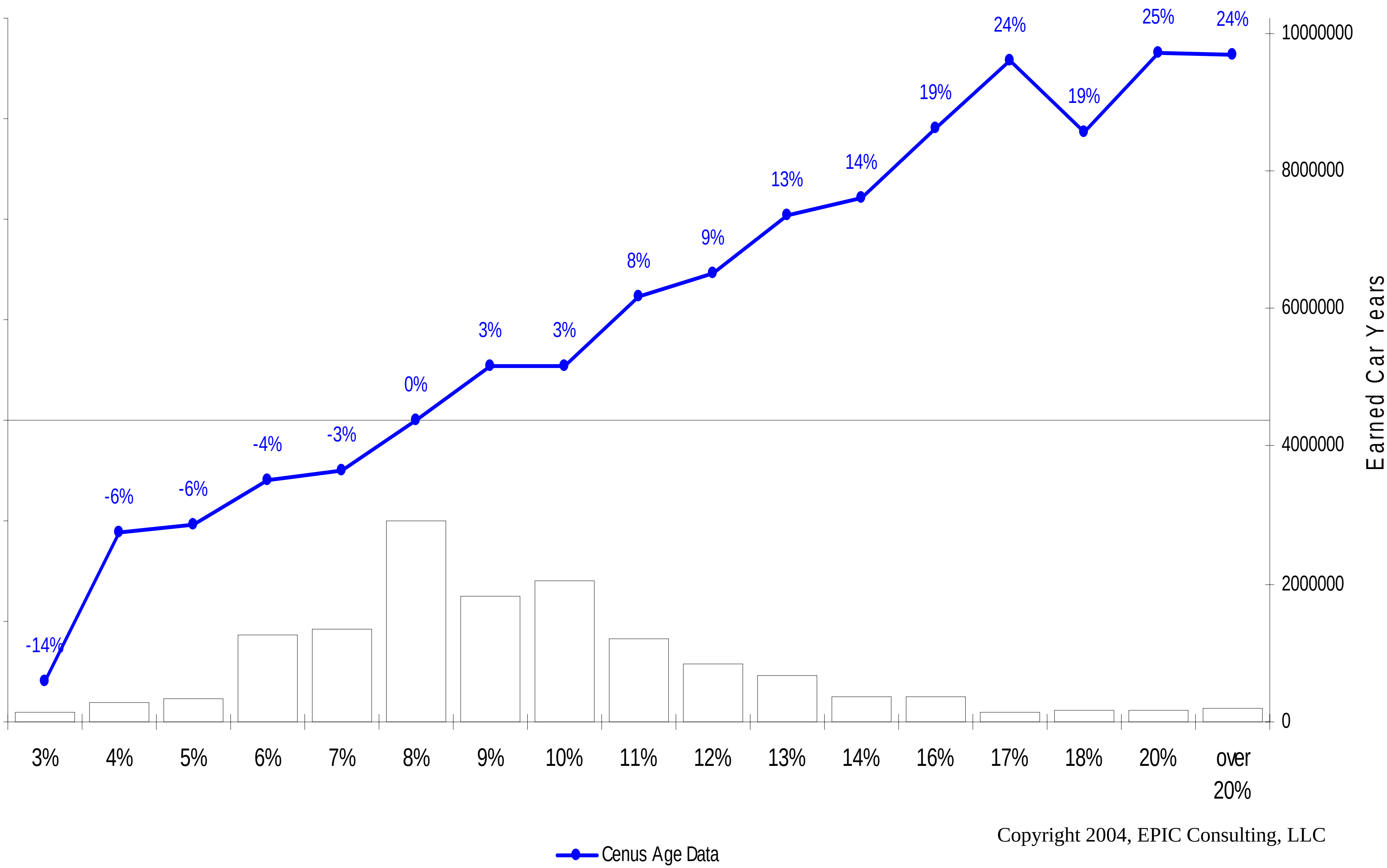
External Data Illustration

Percent of ZIP aged 21-24



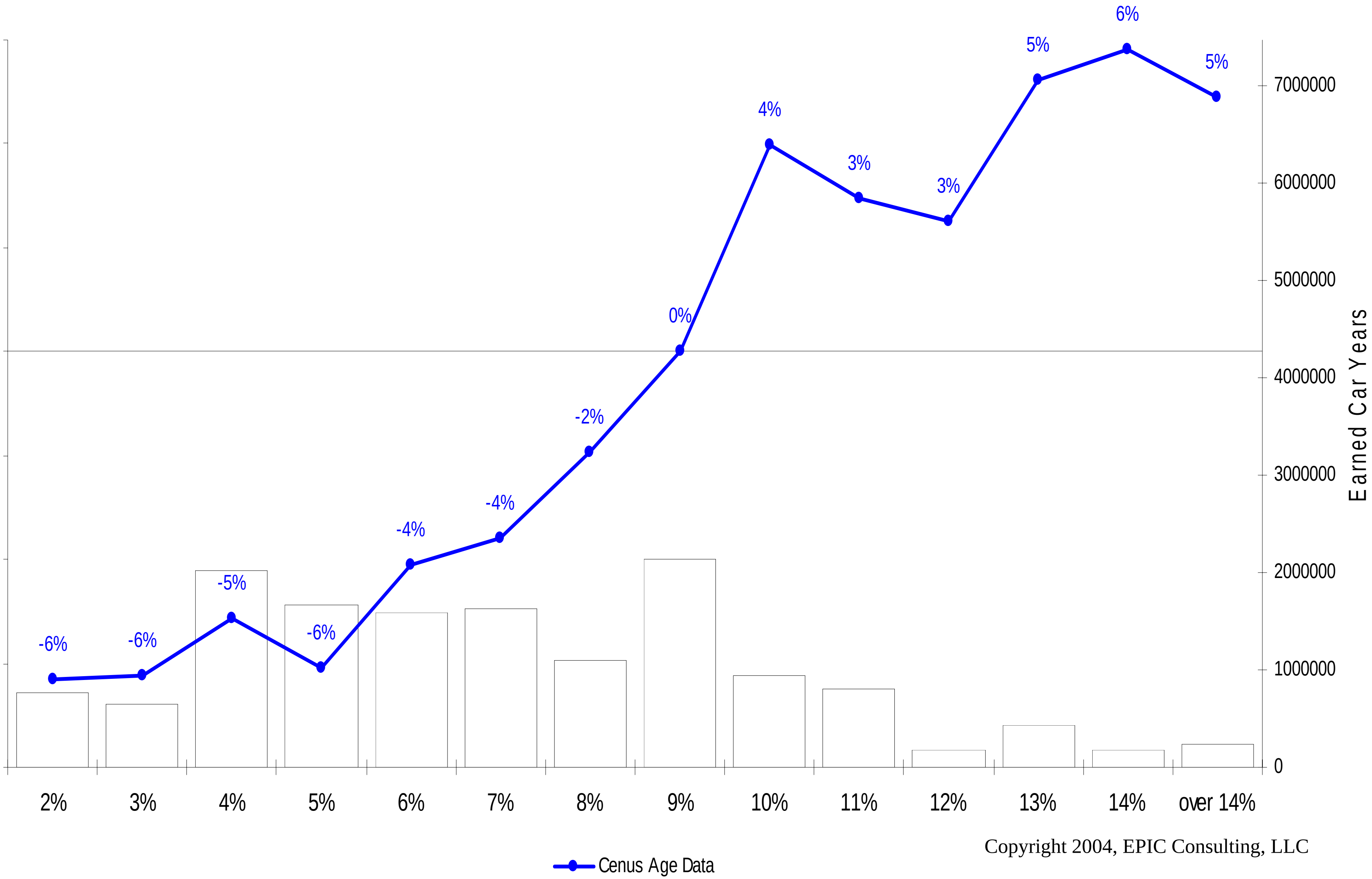
External Data Illustration

Percent of ZIP aged 25-29



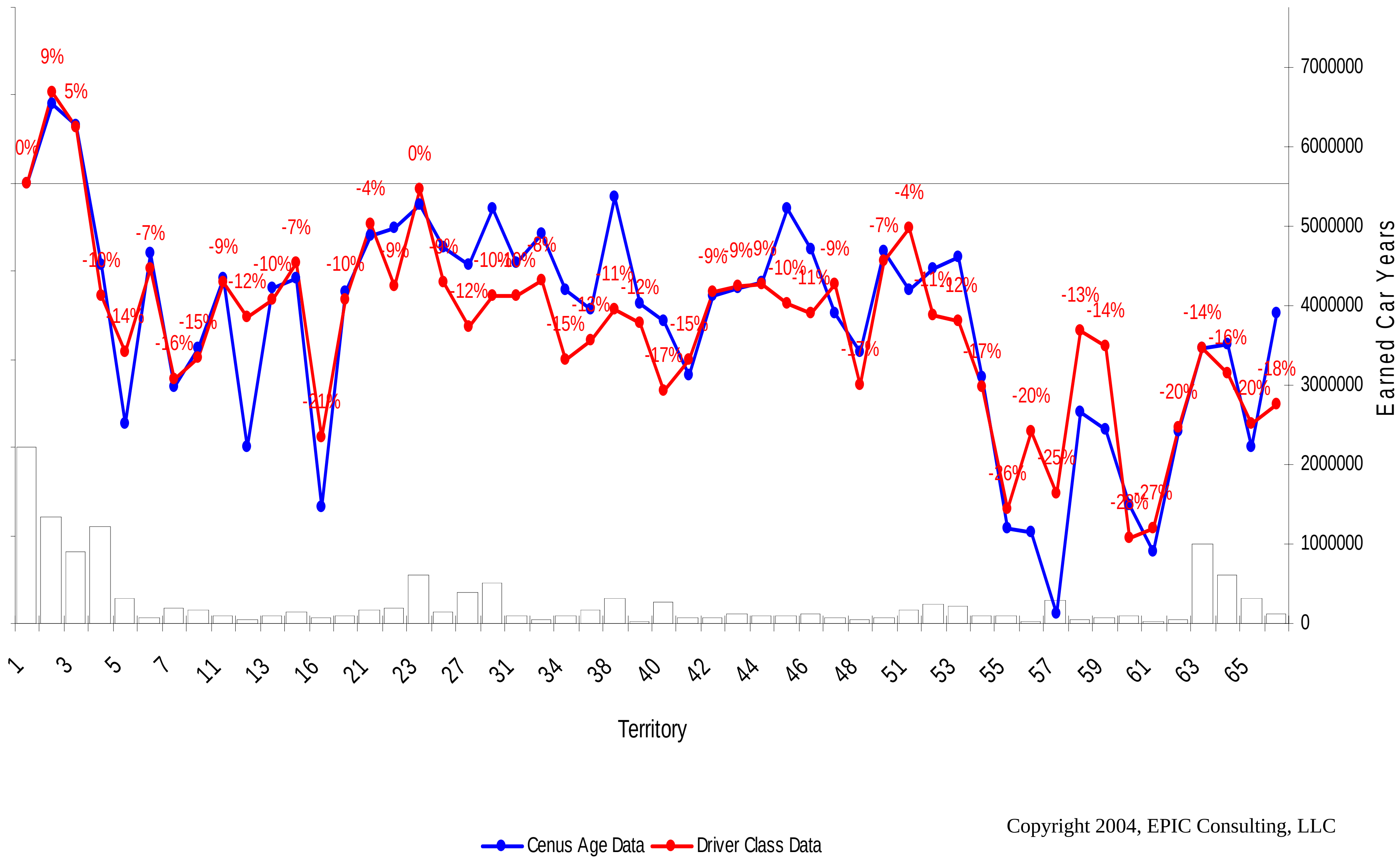
External Data Illustration

Percent of ZIP aged 65-74

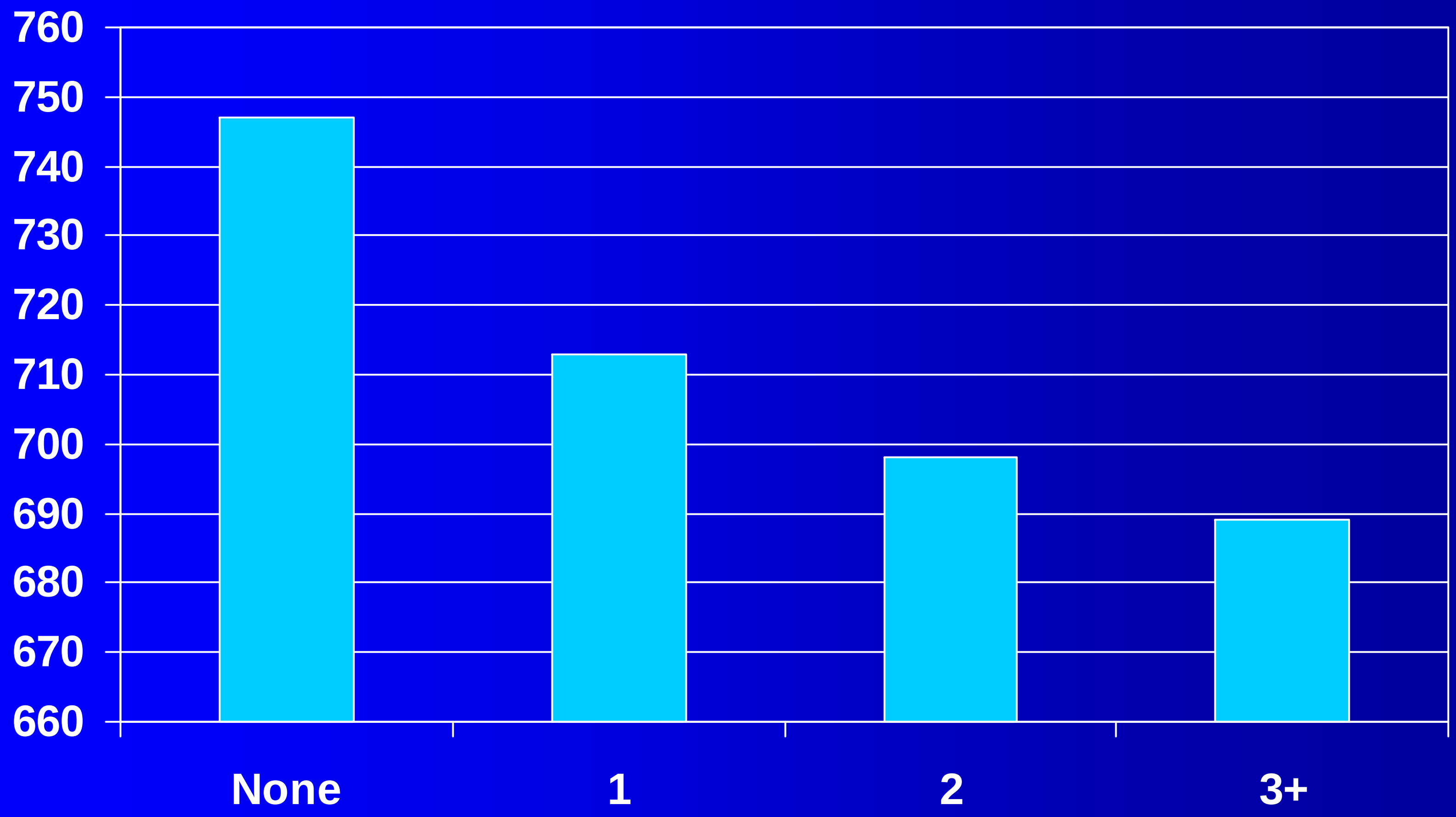


External Data Illustration

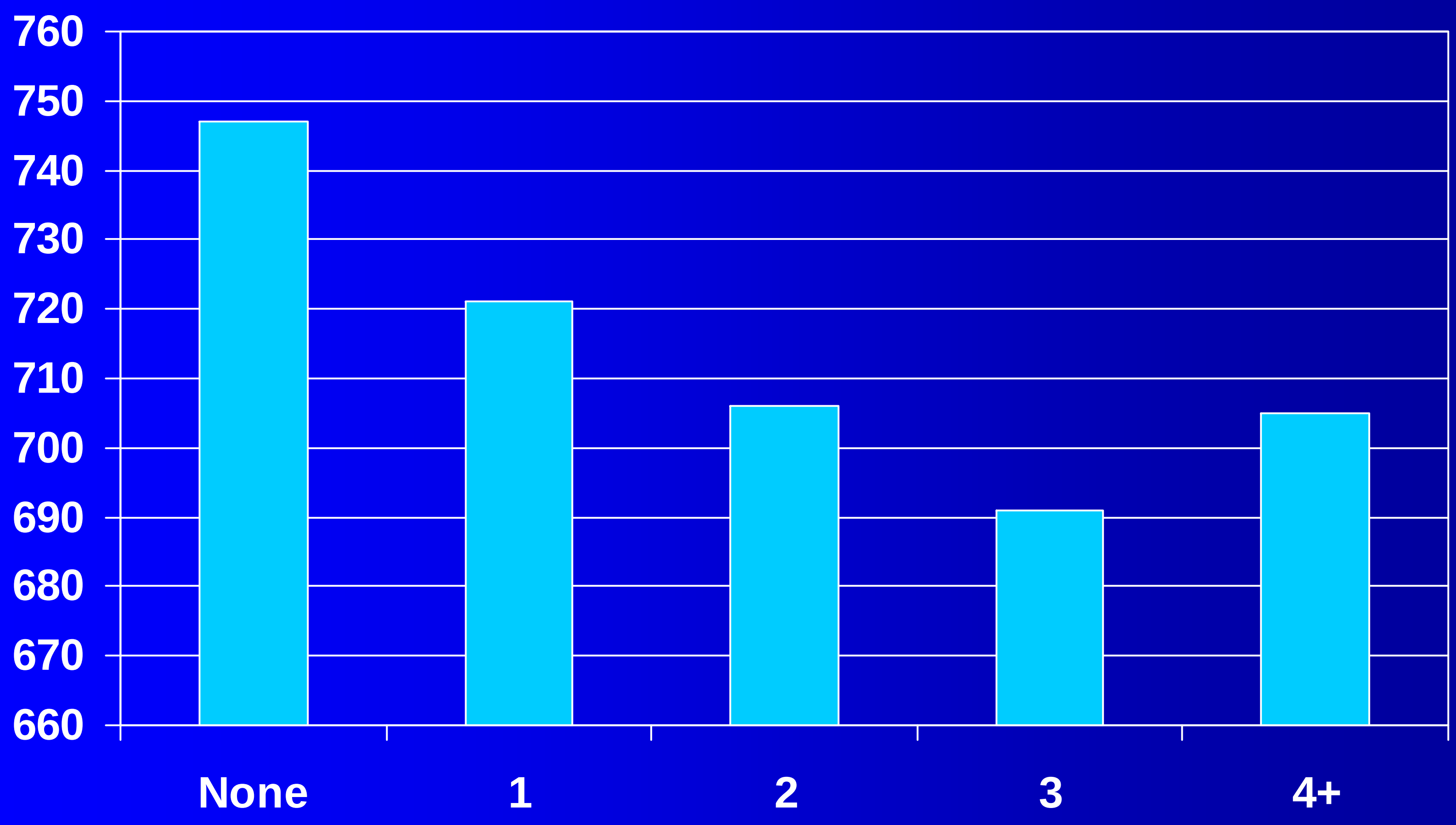
Indicated Factors using Census versus Driver Classification



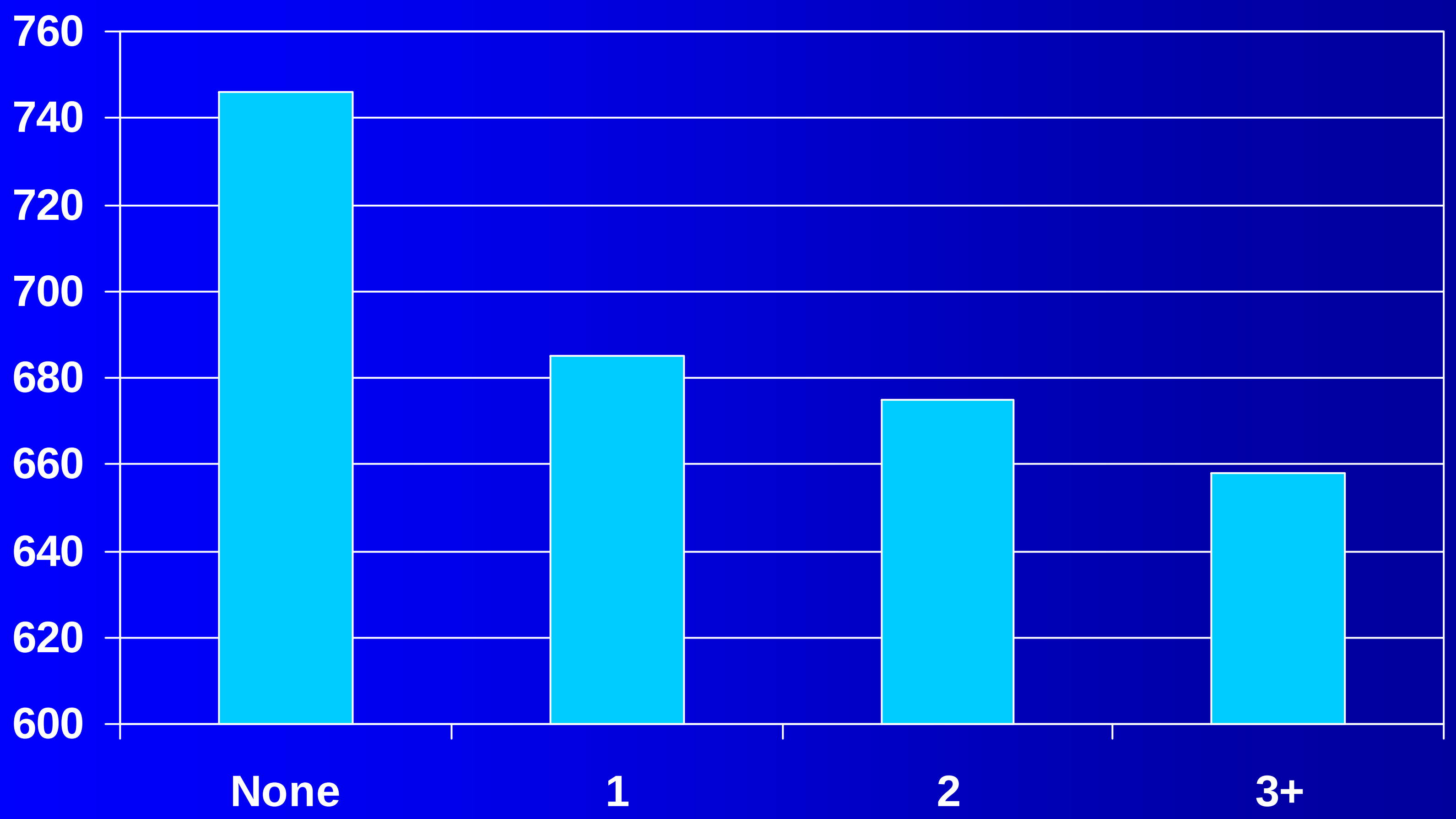
Average Score by Prior Accidents



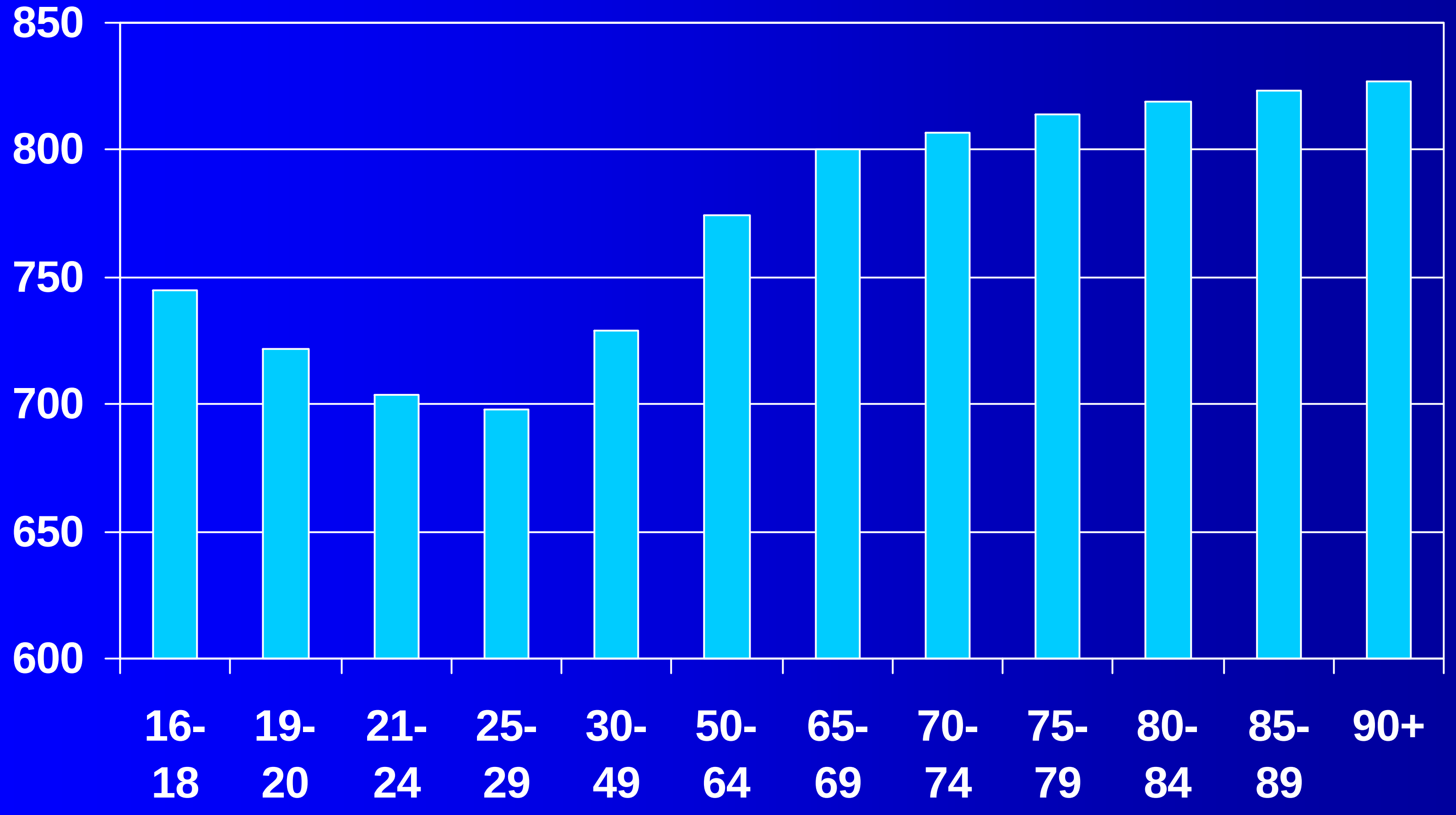
Average Score by Prior Minor Violations



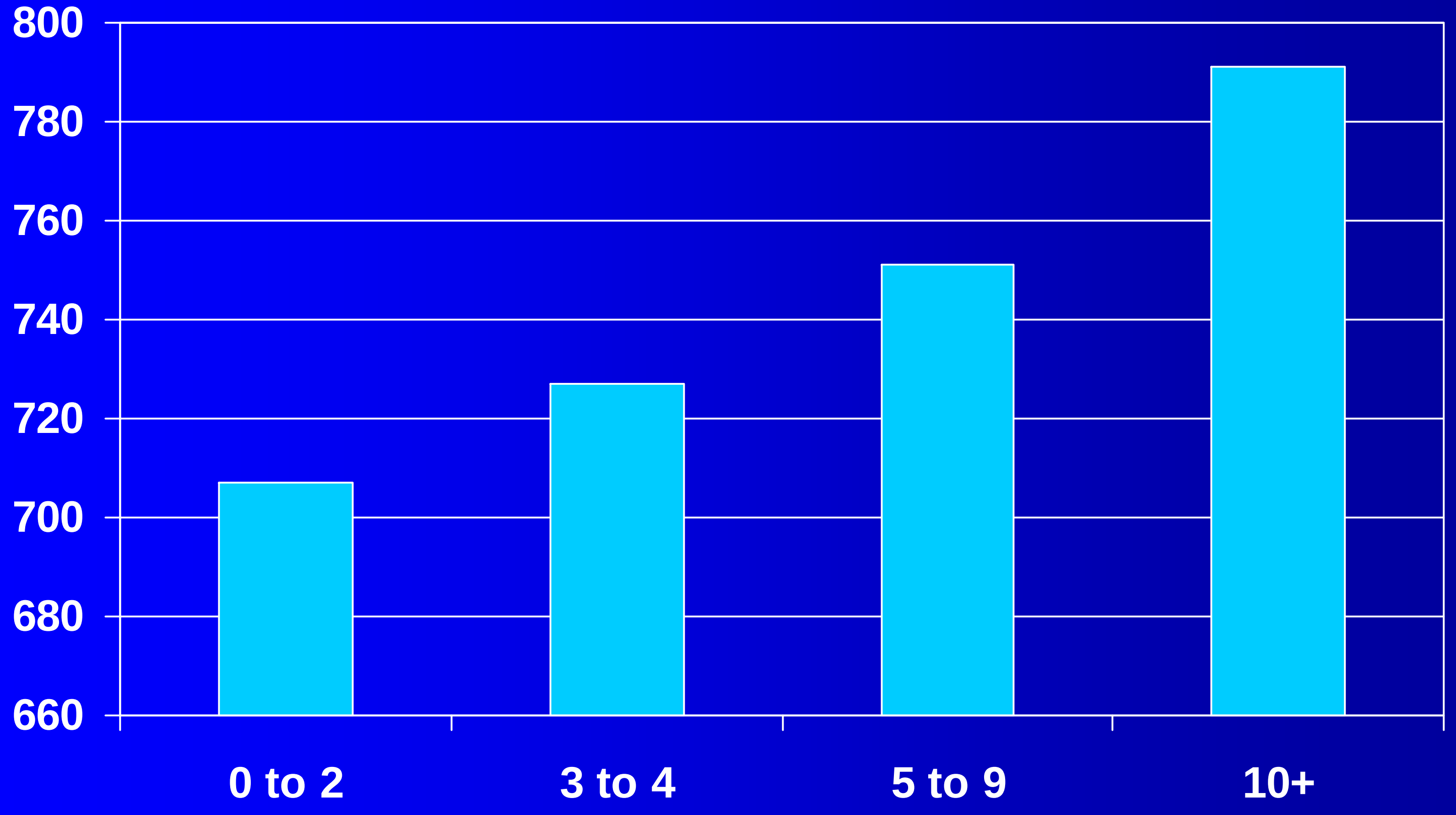
Average Score by Prior Major Violations



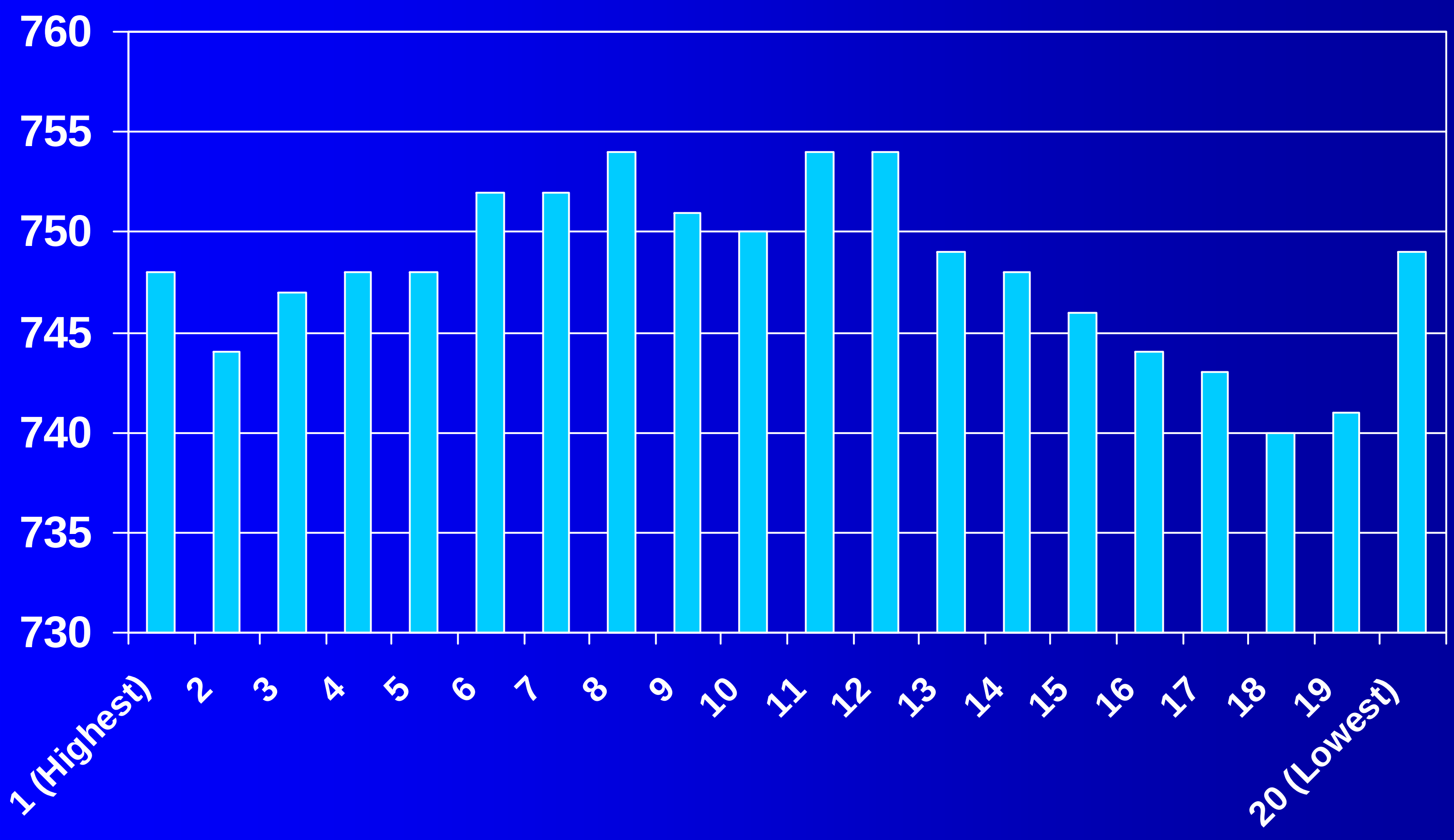
Average Score by Driver Age



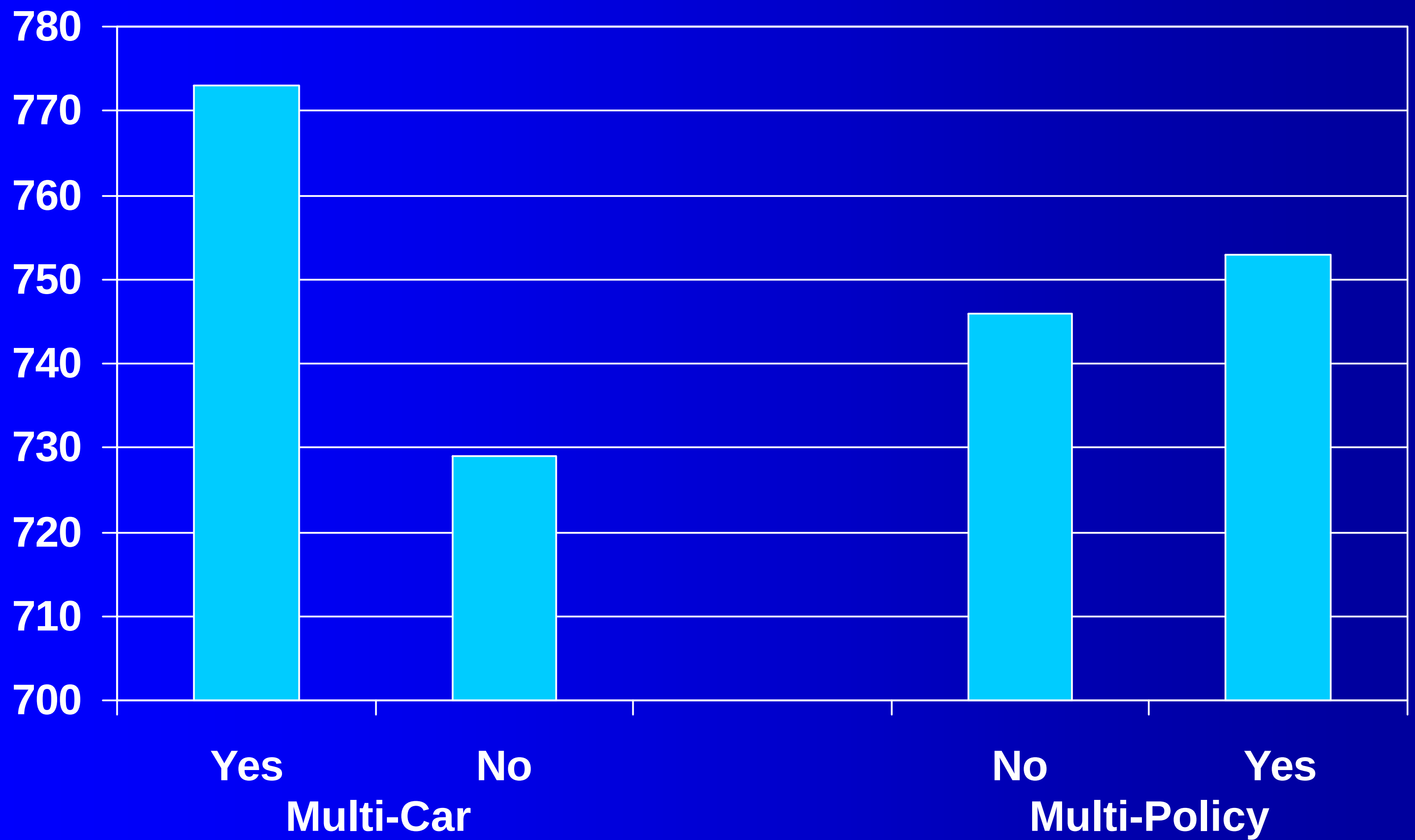
Average Score by Policy Tenure (Years)



Average Score by Population Density Group

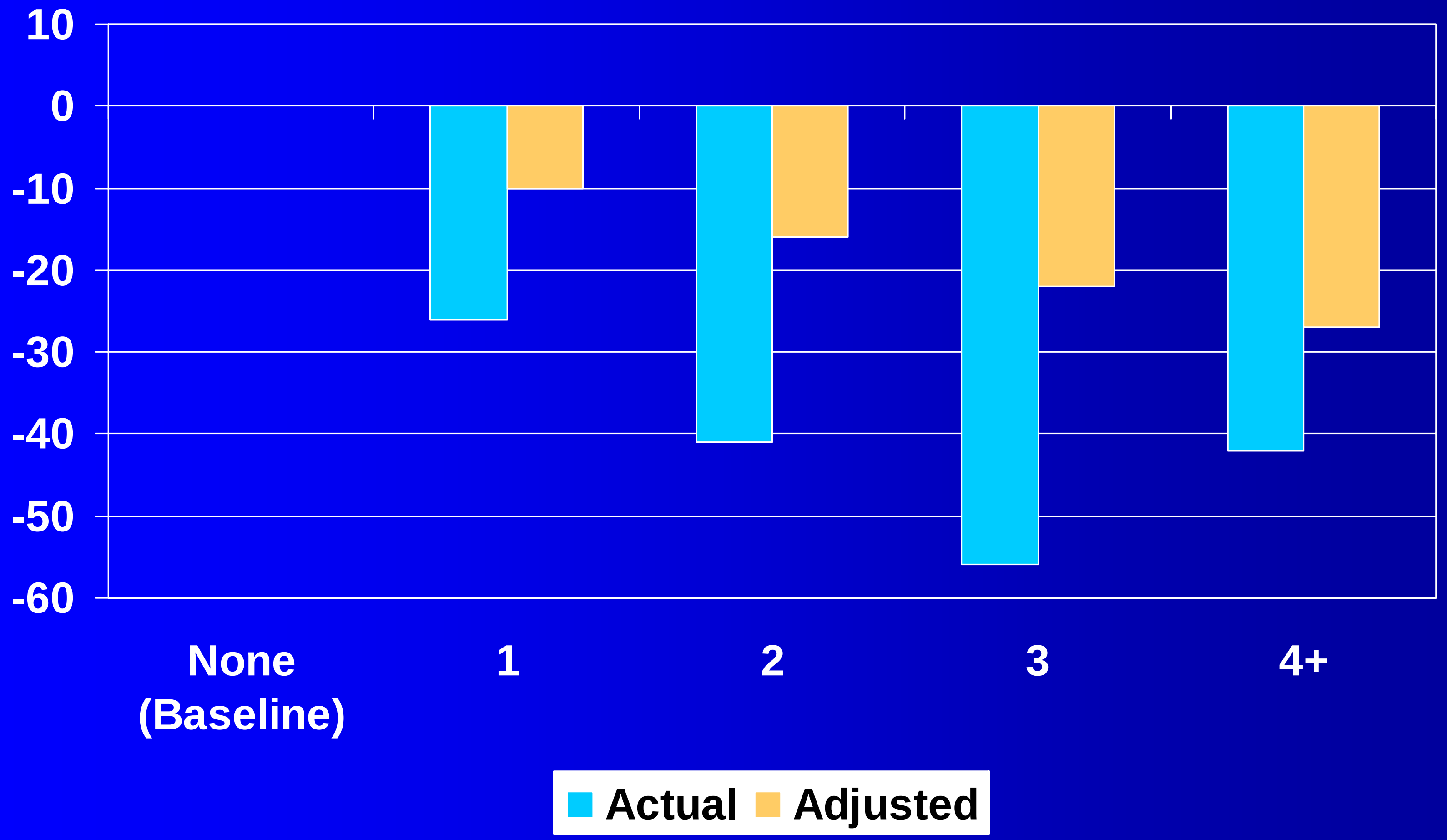


Average Score by Multiple Car/Policy

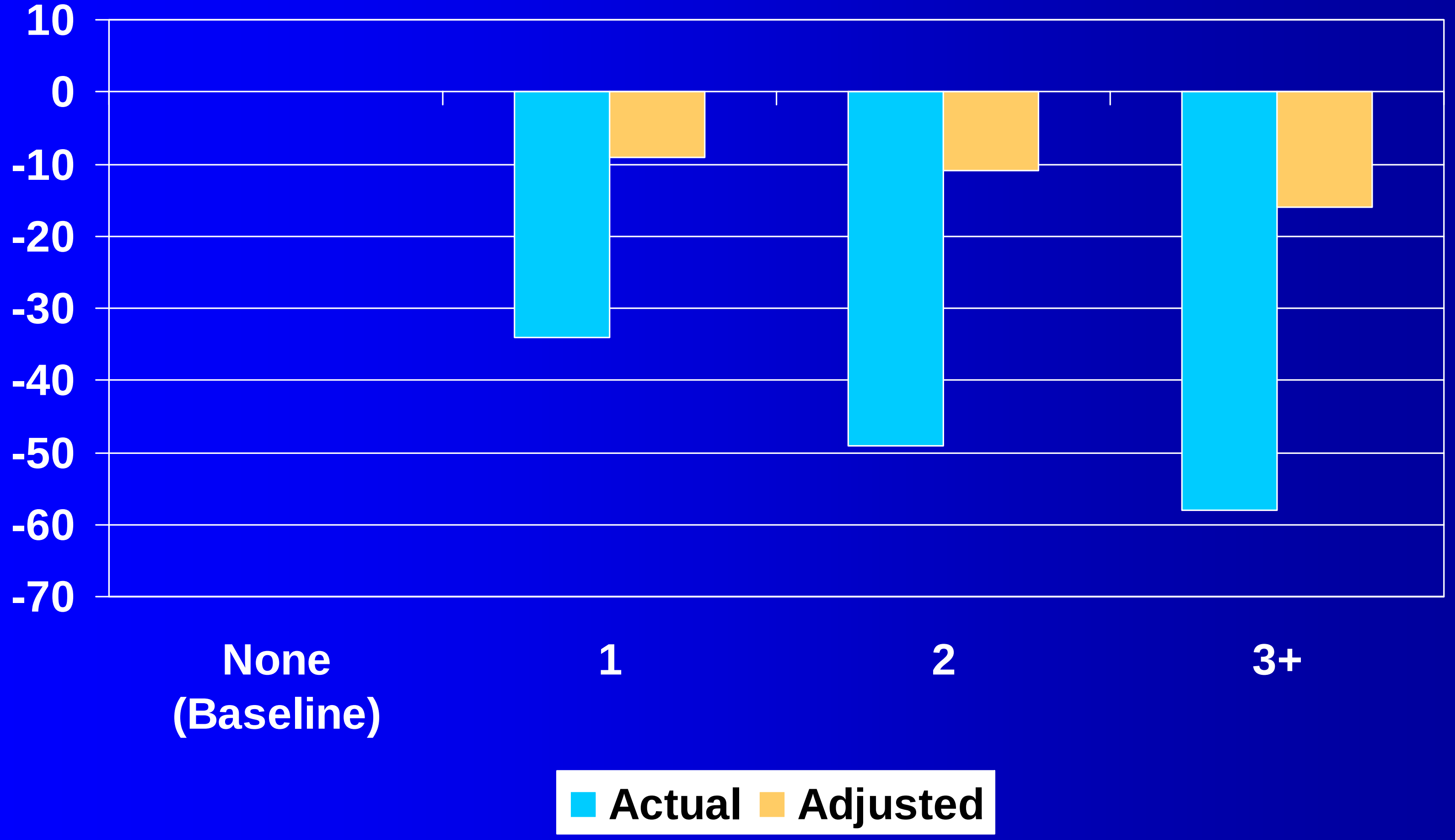


- The following slides show actual and “adjusted” average scores
- The adjusted averages “control for” correlated differences in other characteristics
- Controlling for these differences substantially lessens apparent differences in scores.
- Adjusted differences are generally small and not significant in the real world.

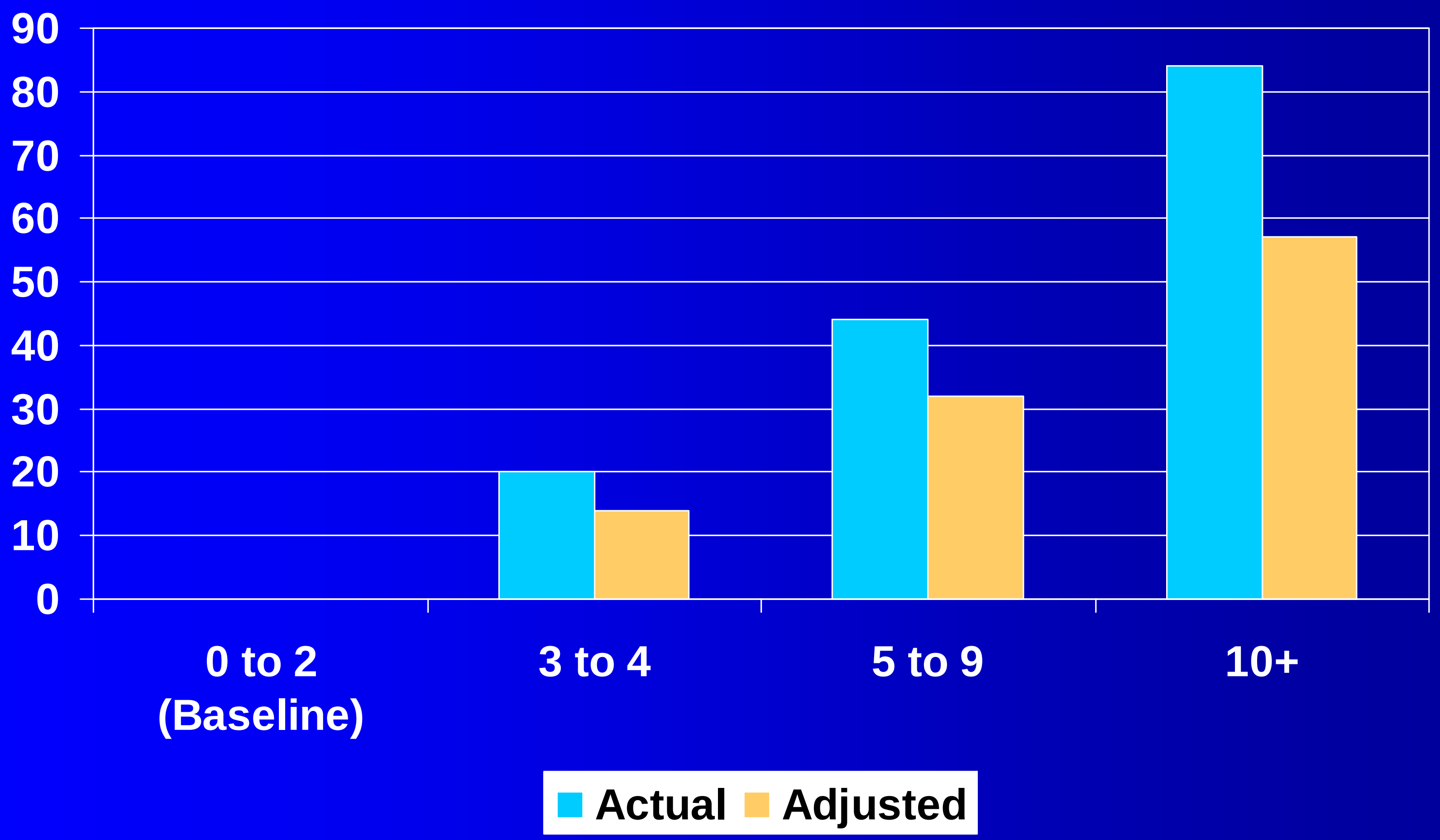
Average Score Differences by Prior Minor Violations



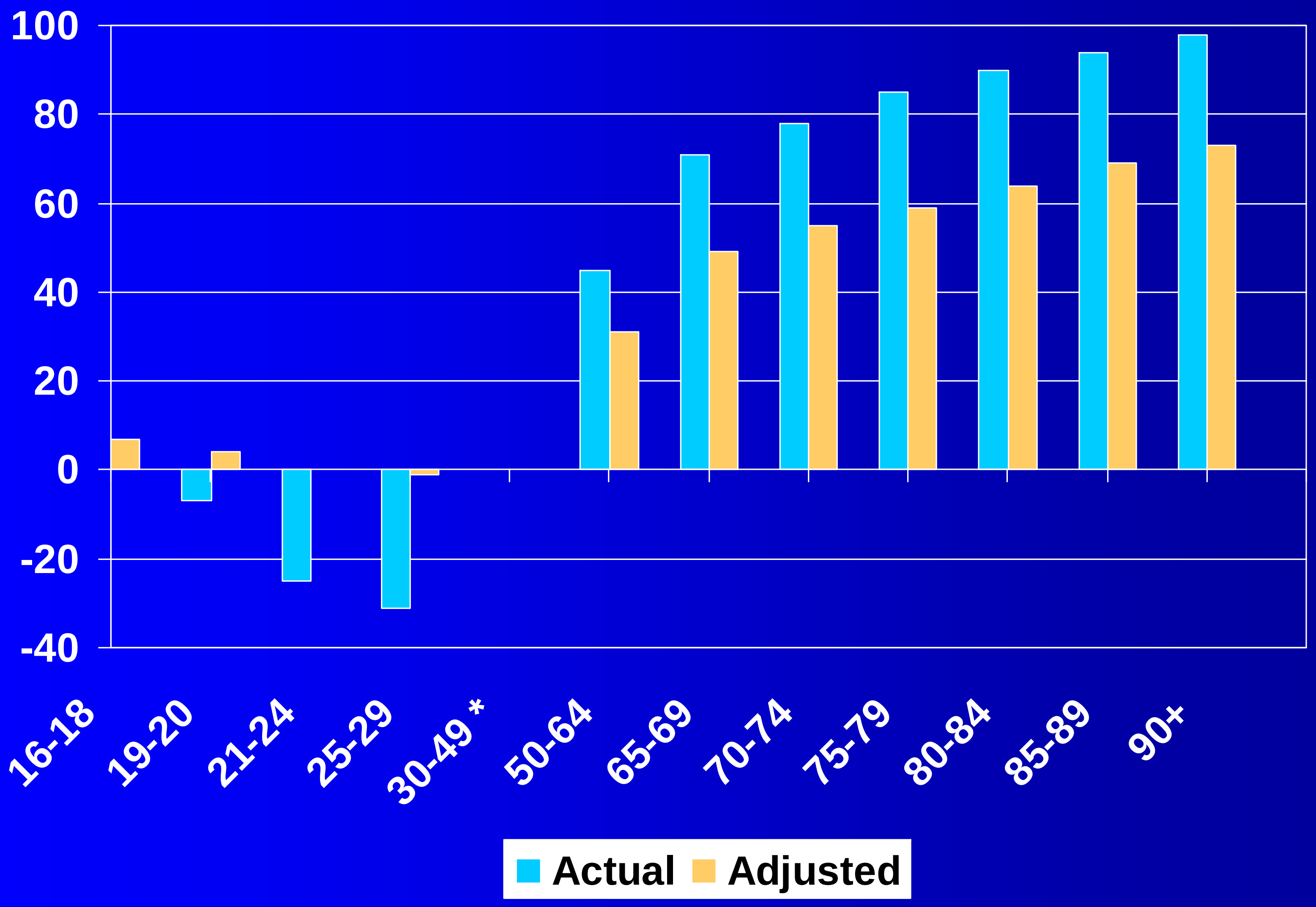
Average Score Differences by Prior Accidents



Average Score Differences by Policy Tenure (Years)

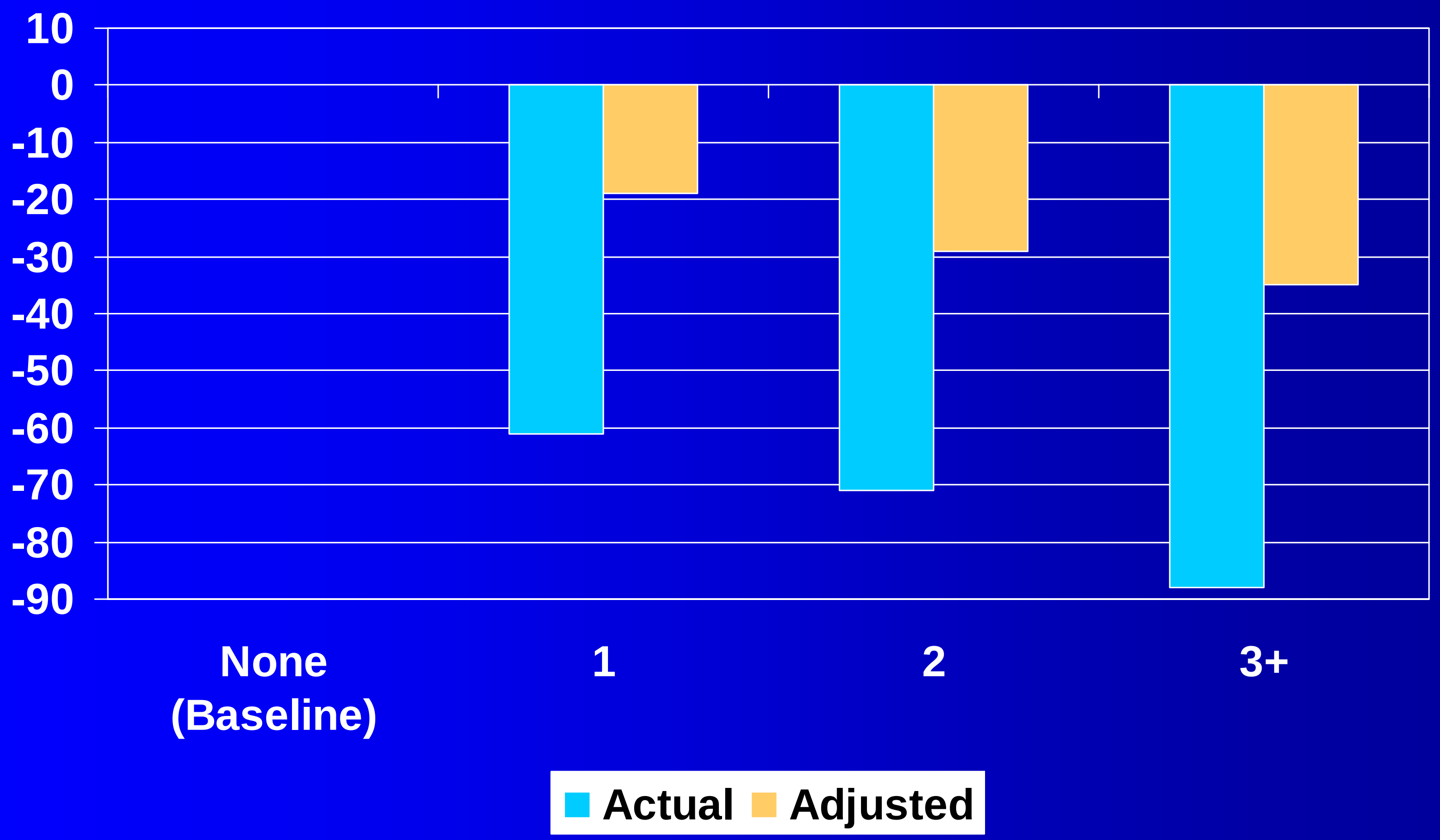


Average Score Differences by Driver Age

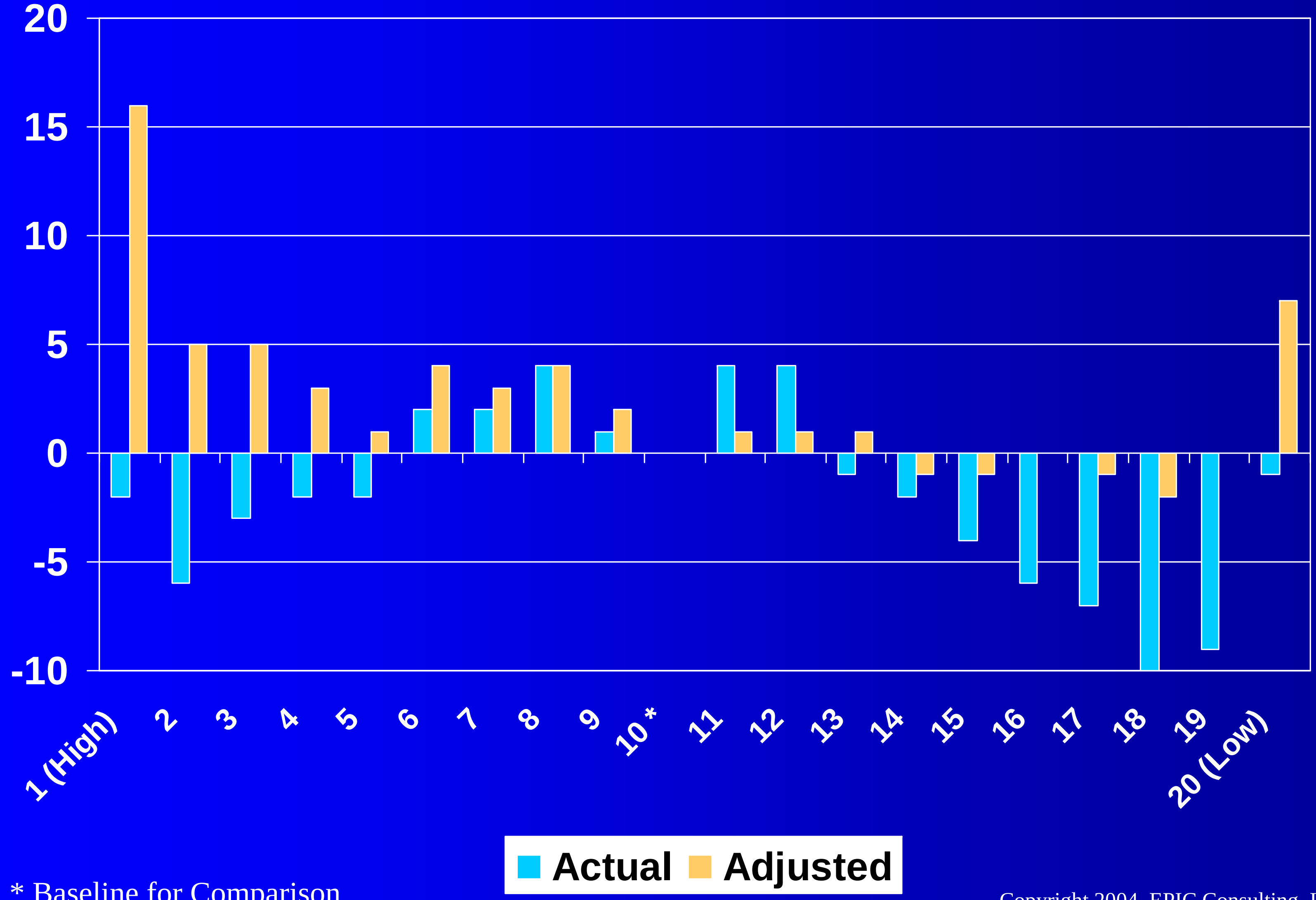


* Baseline for Comparison

Average Score Differences by Prior Major Violations

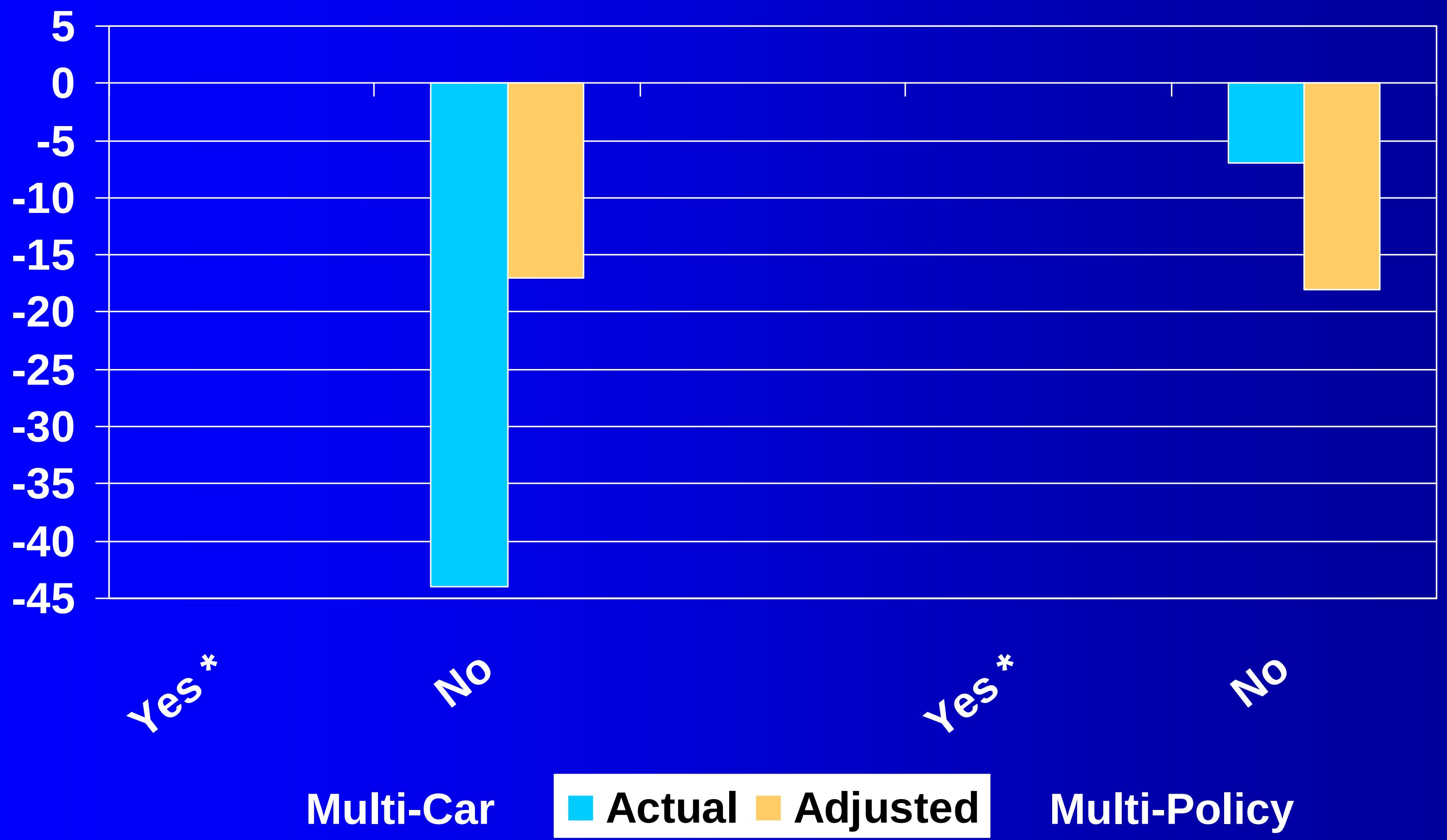


Average Score Differences by Population Density Group



* Baseline for Comparison

Average Score Differences by Multiple Car/Policy



* Baseline for Comparison