

**Assessment of the District of Columbia
Department of Insurance, Securities and Banking
Market Conduct Examination Report**

***“Evaluating Unintentional Bias in Private Passenger
Automobile Insurance”***

June 2024

Robert P. Hartwig, PhD, CPCU
Clinical Associate Professor, Finance Department and
Director, Center for Risk and Uncertainty Management
Darla Moore School of Business
University of South Carolina
robert.hartwig@moore.sc.edu



Executive Summary

Upon request of the APCIA, I have reviewed the District of Columbia Department of Insurance, Securities and Banking's May 2024 market conduct examination report titled "*Evaluating Unintentional Bias in Private Passenger Automobile Insurance.*"

My assessment is that the DISB conclusion that the models used by personal auto insurers to determine premiums in the District of Columbia are biased against inferred Black drivers—resulting in a "Black/white premium gap" of \$271—is not actuarially supported by the arguments presented in the report.

The DISB's reasoning in support of its conclusions generally derives from (i) deficiencies in its approved rating model—which omits key rating variables, and (ii) the modeling methodologies employed in the DISB market conduct exam itself, which are potentially ill-suited for capturing the true impact of rating variables on premium. Consequently, *the DISB should have little confidence as to the accuracy of its \$271 "Black/white premium gap."* Statistically, deriving the estimate of the "Black/white premium gap" from DISB's approved (yet actuarially flawed) rating model effectively invalidates the credibility of DISB's estimate of the "Black/white" premium gap.

I conclude that the DISB's assertion of bias (and inferences of unfair discrimination) derives from several interrelated methodological problems and statistical deficiencies in the report, primarily related to the six factors below. Each of these six factors is discussed in the sections that follow.

1. Misguided assessment of bias
2. The omission of key rating factors
3. Misinterpretation of research related to the importance of key underwriting variables, including credit-based insurance scores
4. A failure or unwillingness to interpret results in a manner consistent with established actuarial standards of practice
5. A failure to recognize deficiencies in the DISB's own permitted rating model and the inherent consequences of those deficiencies, including implicit subsidies.
6. Statistical deficiencies in the modeling methodology employed in the DISB Market Conduct Exam

1. DISB's Misguided Assessment of Bias

- The DISB report immediately concedes (Executive Summary, p. 5) that "Black drivers' average losses are 2.38 times that of white drivers" and that the loss ratio for Blacks is 0.60 while for whites it is 0.36 (see p. 8; chart reproduced below with my annotations). Taken together, this suggests/proves that whites (not Blacks) are paying a materially higher rate per expected dollar of loss. The report characterizes this concession as follows: "*From the actuarial perspective, this means Black drivers are less profitable as a group than the other groups, because although their premiums are high (relative to white drivers), their losses are even higher (relative to white drivers).*" (Exec. Summary, p. 5, para. 3)

		Avg Premium (\$)	Avg Loss (\$)		Loss/Premium
inferred_race					
API	3.5% > Whites	734.54	279.76	9.1% > Whites	0.38
Black	44.3% > Whites	1024.42	611.54	138.4% > Whites	0.60
Hisp	21.0% > Whites	858.82	370.32	44.4% > Whites	0.43
White		709.96	256.49		0.36

LR ratio for Blacks is 24 pts. higher than for whites

- Most readers would probably conclude that the DISB report should have ended with the statement above. The remainder of the report amounts to what might best be characterized as a misguided effort to demonstrate a claim of bias against inferred Black drivers even though they are charged relatively less than whites (as demonstrated by a loss ratio that is 24 points higher for Blacks than for whites).
- While the DISB report seeks to ascertain if the risk of Black and white drivers as measured by variables *DISB deems* acceptable accounts for the differences in premium, there are material actuarial concerns related to both their list of acceptable variables and their methodology for measuring their impact on premium.

2. Omission of Key Rating Variables

- DISB’s prohibition on auto insurers’ use of territory in its accepted rating model and the omission of other key rating variables in its market conduct model introduce bias into any estimate of premium for Black and white drivers, thus directly biasing any estimate of a Black/white premium gap. Bias is a well-known statistical consequence of omitting actuarially valid (i.e., predictive) variables.¹ A partial list of variables omitted from DISB’s approved model and the market conduct model, include:
 - Territory (*prohibited for use in DC’s approved rating model*)
 - Vehicle Make (*omitted from market conduct model*)
 - Vehicle Model (*omitted from market conduct model*)
 - Annual Mileage (*omitted from market conduct model*)
 - Primary Use (i.e., commuting or pleasure) (*omitted from market conduct model*)
- DISB’s prohibition against the use of territory is illustrative of the problems that can arise when key rating variables are omitted from models. Territory is a rating factor proven to be highly correlated with expected loss in the personal auto line. Many cost factors are influenced by territory, including automobile theft rates, and are thus important drivers of premium. For this

¹ Econometrics with R, *Omitted Variable Bias* (Section 6.1). Available at: <https://www.econometrics-with-r.org/6.1-omitted-variable-bias.html>.

reason, territory is a permitted rating variable in auto insurance pricing models in virtually every jurisdiction in the United States—with the exception of the District of Columbia. The statistical consequence of this is that ratemaking models approved for use in DC are unambiguously *biased* and are therefore *less* accurate than they would otherwise be if territory were permitted. The models are biased because the omission of a relevant variable will attribute the effect of that missing variable (e.g., territory) to those variables that are included. Despite this, the DISB report relegates this fact to a single footnote on p. 9, which states that “DC does not allow territorial ratings; so, territory is not an explanatory factor in this analysis.” The fact that territory is not an approved rating factor in DC is a problem of the DISB’s own creation—one that guarantees that rates are not only an inaccurate reflection of cost but are intrinsically unfair. Decades of research by regulators, academics and the federal government demonstrate that regulatory prohibitions, restrictions, and limitations on actuarially sound rating factors distort pricing, reduce competition, and undermine policyholder incentives to reduce risk, thereby imposing costs on society and potentially harming the very policyholders regulators are attempting to help (Hartwig & APCA, 2023).² .

➤ **Omitted Variables: The Case of Territory and Automobile Theft**

The omission of key rating variables, including territory, results in bias in the actuarial estimation of premiums as detailed in section (iii) below. The impact of DISB’s prohibition on the use of territory as a rating factor can be vividly illustrated using data on automobile thefts, which are a significant problem in DC. According to data from the Metropolitan Police of the District of Columbia (MPDC), motor vehicle thefts skyrocketed by 82% in 2023, accounting for the vast majority of the overall 24% in property crimes in DC last year.³ The National Insurance Crime Bureau (NICB), using data from the Department of Justice’s National Crime Information Center, corroborates the MPDC, stating that motor vehicle thefts increased 64% in 2023—the largest increase of any state (including DC)—with DC’s theft rate per 100,000 people rising to 1,149.71—nearly double that of second-place Colorado (583.25).⁴ The MPDC and NICB data make clear that auto theft is an important auto insurance cost driver in DC. It should not, however, be surprising that the frequency of motor vehicle thefts varies across the District. A variety of visualization tools are available to make clear these differences across territories. The heat map in Figure 1 below from CrimeData|DC indicates that the incidence of motor vehicle thefts varied tremendously throughout the District in 2023. DISB requires auto insurers to exclude this information for the purposes of ratemaking.

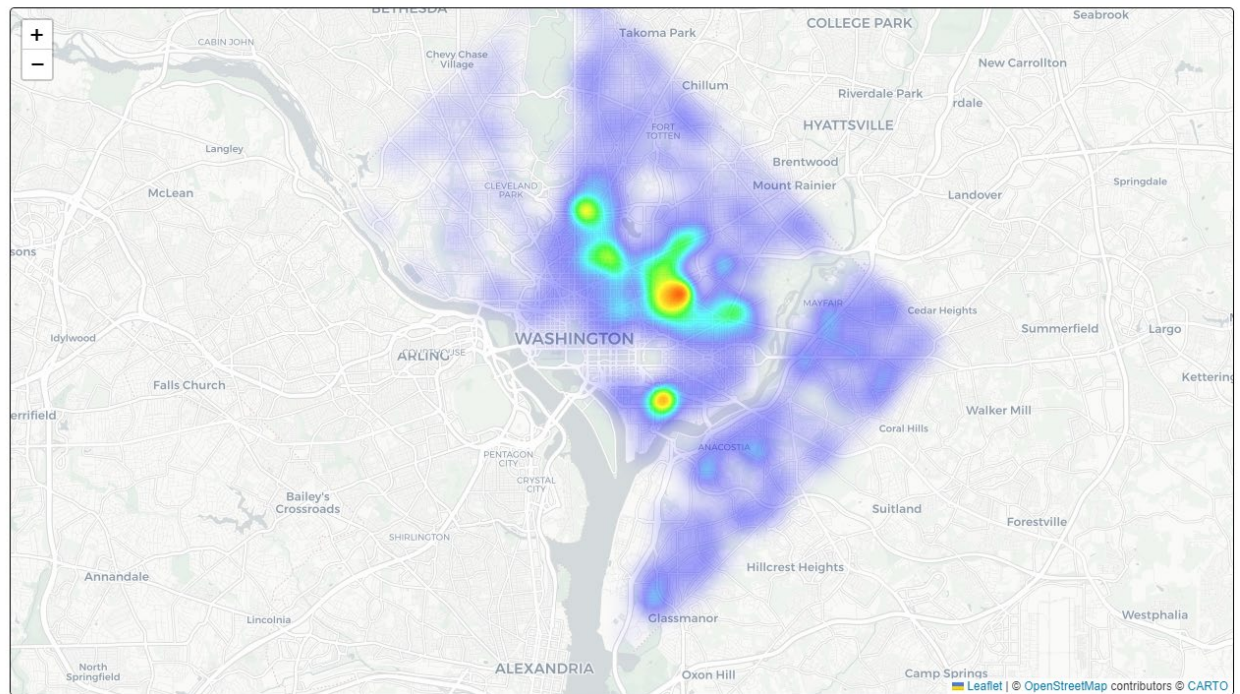
² Hartwig, R. “The Insurance Industry’s Commitment to Fairness in Pricing: A Survey of Historical and Current Research Concerning Risk-Based Pricing and Unfair Discrimination,” *APCIA Working Paper* (August 2023). Available at: <https://www.uscriskcenter.com/the-insurance-industrys-commitment-to-fairness-in-insurance-pricing-august-2023>.

³ Metropolitan Police of the District of Columbia. *District Crime Data at a Glance* (as of May 30, 2024). Accessed at: <https://mpdc.dc.gov/page/district-crime-data-glance>.

⁴ National Insurance Crime Bureau, *2023 Vehicle Theft Trends* (April 2024). Available at: <https://www.nicb.org/news/news-releases/vehicle-thefts-surge-nationwide-2023>.

Figure 1
Heat Map of Motor Vehicle Thefts in the District of Columbia in 2023

6,799 Incidents of Motor Vehicle Theft in 2023



Source: CrimeData | DC: <https://www.crimedatadc.com/heat?f=MOTOR%20VEHICLE%20THEFT>.

3. Misinterpretation of Research on the Issue of Unfair Discrimination in Insurance Pricing

- Insurers do not use race as a rating factor and in DC territory is not permitted. As a result, the DISB report theorizes that the Black/white premium gap is likely the result of insurers “using other characteristics correlated with race and/geography in determining premiums,” citing only a single study on the matter—the FTC’s 2007 report on credit-based insurance scores.⁵ Unfortunately, the DISB totally misinterprets the findings of the FTC study while simultaneously ignoring the voluminous literature on the topic, decades of which are summarized in the 2023 study referenced in footnote 1.
- With respect to the FTC report specifically, that report is widely viewed to be the most comprehensive and authoritative on the use of CBIS by insurers. The FTC conclusively determined that CBIS are “effective predictors of claims that consumers will file.” But the Commission’s findings go far beyond the well-established correlation between CBIS and insured loss, pointing out that the use of CBIS benefits all consumers, *including racial and ethnic*

⁵ Federal Trade Commission, “Credit-Based Insurance Scores: Impacts on Consumers of Automobile Insurance,” (July 2007). Available at: https://www.ftc.gov/sites/default/files/documents/reports/credit-based-insurance-scores-impacts-consumers-automobile-insurance-report-congress-federal-trade/p044804facta_report_credit-based_insurance_scores.pdf.

minorities. Indeed, the FTC found that “credit-based insurance scores appear to have little effect as a “proxy” for membership in racial and ethnic groups in decisions related to insurance. In this sense, CBIS achieves two critical objectives of any ratemaking plan: efficiency and equity. Elements of each are central to the FTC’s key findings, summarized below:

- Scores effectively predict the number of claims consumers file and the total cost of those claims. Their use is likely to make the price of insurance better match the risk of loss that consumers pose. Thus, on average, as a result of the use of scores, higher-risk consumers pay higher premiums and lower-risk consumers pay lower premiums.
- Use of scores may result in benefits for consumers. For example, scores permit insurers to evaluate risk with greater accuracy, which may make them more willing to offer insurance to higher-risk consumers for whom they otherwise would not be able to determine an appropriate premium. Scores also may allow insurers to grant and price coverage more efficiently, producing cost savings that could result in lower premiums.
- Although scores vary among racial and ethnic groups, scores effectively predict risk of claims within those racial and ethnic groups.
- The Commission could not develop an alternative scoring model that would continue to predict risk effectively yet decrease the differences in scores among racial and ethnic groups. The results of these efforts indicate that there is no readily available alternative scoring model that would achieve those results.²⁵

Policymakers and consumer advocates generally predicate their calls to prohibit or restrict insurer use of CBIS and other rating variables on equity concerns centered on race, ethnicity, and income. The FTC’s findings concerning CBIS suggest that these concerns are largely misplaced. Further, the Commission cautions that prohibitions on the use of insurance scores can lead to pricing distortions and market disruptions that can harm the very populations such actions are intended to benefit.

- The DISB report is in direct contradiction to the FTC’s principal findings (and virtually all other research conducted over the prior 30 years), by alleging that insurer use of CBIS is unfairly discriminatory to certain racial and ethnic groups.

4. Failure to Interpret Results in a Manner Consistent with Actuarial Standards of Practice: *The Question of Correlation vs. Causation*

- Again, the DISB (p. 14, Section 5) concludes that the “premium gap” between inferred Black drivers and inferred white drivers is likely the result “other factors not in our data.” DISB cites credit-based insurance scores (discussed previously), discounts and payment modes as among those factors not included in their data and suggests that these factors are “likely responsible for sizable differences in premium rates.” Having acknowledged that CBIS, discounts and payment modes are all “legitimate factors commonly used by insurers,” the DISB nevertheless takes issue with the use of these factors, asserting that they are “arguably not causally related to auto insurance claims.”

- *The DISB comment about causality is irrelevant.* Actuarial standards of practice require a demonstration that a factor is correlated with (i.e., predictive of) loss. There is no actuarial standard requiring a demonstration of causality.
- The DISB is dismissive of actuarial evidence that factors such as credit-based insurance scores have, in fact, been demonstrated to be correlated with certain risky driving behaviors. A 2021 report (Hartwig and APCIA) uses telematics data to demonstrate that hard braking behaviors are strongly associated with the drivers with a lower credit-based insurance scores.⁶ The cost to insure the highest risk drivers (i.e., those with a high relative frequency of hard-braking behavior) is approximately 28 percent higher than for the overall driver population. Conversely, the cost to insure drivers with the strongest CBIS (and lowest relative frequency of hard-braking behavior) is about 30 percent less than the overall driver population. The telematics data therefore demonstrate that credit-based insurance scores are strongly predictive of auto insurance claims costs. This latter finding that has been corroborated in many independent studies over the past two decades.

5. The Current DISB Permitted Rating Model is *Already Unfair, by Design* and the DISB’s Estimate of a “Black/white Premium Gap” Has No Statistical Credibility

- The prohibition of territory by DISB introduces statistical bias into its existing, approved rating model(s). Specifically, the prohibition creates a network of implicit subsidies benefiting drivers who live in territories associated with high expected loss. Inclusion of territory would increase the accuracy of rating models and control for bias, allowing other traditional rating variables to be more accurate predictors of loss. *DISB’s prohibition against the use of territory makes any assertion of unfair pricing by insurers highly suspect. Specifically, the DISB should have little-to-no confidence as to the accuracy of its \$271 “Black/white premium gap.”* Statistically, deriving the estimate of the “Black/white premium gap” from DISB’s approved (yet biased) rating model effectively invalidates the credibility of DISB’s estimate of the “Black/white” premium gap.

6. Statistical Deficiencies in the DISB’s Market Conduct Exam Model⁷

The model used by DISB to arrive at its conclusion of bias suffers from several statistical deficiencies. Consistent with the discussion in previous sections, the validity of that model is in doubt for a variety of reasons, several of which are described below:

- Use of an ordinary least squares (OLS) linear regression model (as described in Appendix D of the report) does not reflect standard industry practice for the modeling of premiums. A more flexible technique known as Generalized Linear Models (GLM) is more commonly applied by actuaries. One major advantage of the GLM approach is that it does not require a linear relationship between the model’s explanatory variables (e.g., vehicle age) and the dependent variable (i.e., premiums). In practice, many explanatory variables have a *nonlinear* relationship

⁶ Hartwig, R. and APCIA. “Behavioral Validation of Auto Insurance Rating Variables,” (November 2021). Available at: <https://www.uscriskcenter.com/presentations/white-paper-behavioral-validation-of-auto-insurance-rating-variables/>.

⁷ The DISB retained the services of an outside party, O’Neil Risk Consulting and Algorithmic Auditing (ORCAA), to “assist and provide subject matter expertise,” including the construction of the model used in the report.

with the variable being modeled.⁸ The GLM approach also does not require that the dependent variable be normally distributed.

- Insurers do not collect information related to the race or ethnicity of policyholders. The DISB report therefore relies upon an indirect technique known as Bayesian-Improved First Surname Geocoding (BIFSG) to *infer* the race of policyholders. It is unclear if the use of BIFSG is a reasonable approach for inferring race for the purposes of ascertaining the race of single vehicle/single driver auto insurance policies, which collectively account for barely half (54%) of aggregate premium. Specifically, the distribution of policyholders by race and age included in the report (see p. 9) likely differs from race and age distributions derived from U.S. Census data on which the BIFSG inference algorithm is based.

Conclusion

The District of Columbia’s Department of Insurance, Securities and Banking June 2024 market conduct exam alleges that Black drivers in the District pay \$271 more per year, on average, for auto insurance than white drivers of similar risk. The DISB characterizes this “Black/white premium gap” as prima facie evidence of unintentional discrimination by automobile insurers.

Upon review of the DISB’s permitted auto insurance rating model and the premium model developed for the market conduct exam, it is clear that the DISB’s assertion of bias (and inferences of unfair discrimination) is unfounded and instead derives from several interrelated methodological problems and statistical deficiencies in the report. Those methodological problems and statistical deficiencies are material and are discussed in detail in this assessment of the DISB report.

DISB’s approved rating model and the model constructed for the purposes of the market conduct exam both omit key rating variables that are known to be predictive of loss. Their omission, combined with potentially actuarially deficient model specifications, suggest that any estimates of premium are likely to be inaccurate. Consequently, the DISB should have little confidence as to the accuracy of its \$271 “Black/white premium gap.”

⁸ Penn State University, Eberly College of Science, *Introduction to GLMs*, Accessed at: <https://online.stat.psu.edu/stat504/lesson/6/6.1>.