NAMIC
NATIONAL ASSOCIATION OF MUTUAL INSURANCE COMPANIES

317.875.5250 | [F] 317.879.8408
10 W. 106th St., Suite 100, Indianapolis, Indiana 46290

202.628.1558 | [F] 202.628.1601
20 F Street N.W., Suite 510 | Washington, D.C. 20001

June 30, 2026

The Honorable Karmina M. Woods, Commissioner
District of Columbia Department of Insurance, Securities and Banking
1050 First Street, NE, Suite 801
Washington, DC 20002

Re: Draft Initiative to Evaluate Unintentional Bias in Private Passenger Automobile Insurance

Commissioner Woods,

On behalf of our members, the National Association of Mutual Insurance Companies (NAMIC)[1], it is our recommendation that the District of Columbia Department of Insurance, Securities & Banking (DISB) should withdraw or substantially revise the proposed automobile insurance balancing test provisions because they are methodologically unsound, operationally unworkable, and risk undermining actuarially sound, risk-based pricing without a clear legal or empirical foundation.

Introduction and Overarching Concerns
Our members strongly support risk-based, non-discriminatory insurance pricing. However, the proposed balancing test appears to be a lightly modified version of stepwise variable selection, a methodology that modern statistical and actuarial practice generally discourages because of instability, order sensitivity, and weak theoretical foundations. Embedding this approach into rate regulation without demonstrated validity would erode confidence in the ratemaking process and risk undermining risk-based pricing that is central to affordability, availability, and competition in the District.

Methodological Flaws in the Balance Testing
The proposal describes a "balancing test" that relies on a ratio of a predictive-power metric ("How much information about risk does it add?") to a "fairness" or "gap reduction" metric ("How much of the gap would it explain away?"), but many components of both the numerator and denominator are left undefined or are explicitly deferred to future clarification. There is no theoretical justification or empirical simulation evidence presented to show that this procedure reliably distinguishes legitimate, risk-based variables from alleged proxies or actually reduces purported, unexplained "premium gaps" by protected class.

[1] *The National Association of Mutual Insurance Companies (NAMIC) is the foremost trade association representing the property/casualty insurance industry. Serving more than 1,300 member companies—including local and regional insurers as well as some of the nation's largest carriers—NAMIC members collectively write $467 billion in annual premiums, representing 61% of the homeowners and 53% of the automobile insurance markets. For more than 130 years, NAMIC has been the leading voice advancing public policy solutions and regulatory frameworks that promote a strong, competitive market and protect our members and their policyholders.*

Moreover, the testing process, as described, is a small modification of stepwise variable selection: variables are added and assessed in sequence, and the result can depend heavily on the order in which predictors are tested. This order dependence raises major concerns about consistency, explainability, and repeatability across companies and over time, especially in a regulatory context where like situations should be treated alike. Because different carriers use different modeling structures, variable coding, and interaction specifications, the same data could yield very different regulatory outcomes under this procedure, which increases perceived arbitrariness rather than enhancing fairness.

Lack of Transparency and Clarity in the "Gap Reduction" Denominator

The most consequential element of the balancing test is the "gap reduction" denominator, yet this is precisely where the proposal is least transparent. The draft refers to "how much of the gap would it explain away," but does not specify the exact outcome measure, functional form, or statistical model used to define and quantify the premium "gap," nor how race or other protected characteristics will be imputed and controlled.

In both the numerator and denominator, the proposal acknowledges missing calculation details and leaves key modeling choices (such as definition of the harm metric, control variables, and race imputation method) to be determined later. Without explicit, reproducible definitions, insurers cannot implement the test, cannot replicate DISB's analyses, and cannot meaningfully respond to any identified disparities. This uncertainty also exposes DISB to challenge because decisions would rest on methods that are not specified in rule or guidance and are not verifiable by affected parties.

Notwithstanding our position that DISB withdraw the proposed test, if the test moves forward, we respectfully request that DISB provide:

- A precise mathematical definition of the predictive-power metric and the gap-reduction metric, including model specification, treatment of interactions, and outcome variable.
- A clear description of how race or other protected characteristics will be estimated or inferred, including data sources, assumptions, and limitations.
- Worked examples, using realistic multi-factor insurance models rather than single-variable toy examples, showing how the test behaves under different scenarios.

Absent such detail, it is premature to adopt this framework into binding regulatory practice.

Risk to Risk-Based Pricing and Rate Review Timelines

Risk-based pricing is a foundational principle of property-casualty insurance: premium should reflect expected loss and expenses so that higher-risk policyholders do not unduly subsidize lower-risk policyholders. Put another way, an "unfair" outcome in insurance occurs when rate diverges from risk, regardless of the policyholder's membership in a protected class. Because not all risks are the same, neither are all rates.

By design, the proposed balancing test contemplates limiting or discarding variables that meaningfully improve risk prediction where the fairness metric fails a threshold, but the proposal does not address how the resulting loss of predictive power will affect overall rate adequacy, segmentation, and cross-subsidies.

If the balancing test leads to removal of important, actuarially indicated factors, higher-risk drivers may be subsidized by lower-risk drivers, increasing anti-selection pressures and potentially raising average premiums. This dynamic may also trigger more restrictive underwriting in the District as insurers seek to manage portfolio risk, reducing availability for some consumers. Additionally, implementing a novel, multi-step statistical test as part of every major filing will materially lengthen the rate review process, as both regulators and insurers work through unfamiliar methods, documentation, and potential disputes about modeling choices.

We urge DISB to conduct and publish a detailed regulatory impact assessment on review timelines, data requirements, model governance, and resource needs before moving forward, and to consider piloting any fairness testing approach in a voluntary, consultative phase rather than embedding it immediately as a binding requirement in rate filings.

Actuarial Jurisdiction, CAS Resources, and Definition of "Common Sense" Factors
The draft draws from the Explainable Fairness article by O'Neil, Appel, and Sargeant, which itself emphasizes the need for clear definitions of "legitimate" factors and for explainable fairness standards that regulators and industry can operationalize. However, the DISB proposal references "less common sense" factors without defining what this means, how those factors will be identified, or how they relate to existing actuarial standards that already require that rating variables be predictive of expected loss and reasonably related to risk.

In current practice, when a company files a new product, the insurer must justify each rating variable as predictive of loss and consistent with applicable law, and the regulator may object to any variable that appears unjustified or unfair. We ask DISB to clarify which rating factors it views as "less common sense" in this context, and how that list was developed. If DISB is to proceed with the proposal, we also encourage DISB to reference and build upon the growing body of work from the Casualty Actuarial Society (CAS) and the American Academy of Actuaries on measuring and mitigating bias and improving fairness in insurance pricing, many of which discuss practical methods for balancing fairness and predictive accuracy.

By anchoring its approach in actuarial literature and professional standards, DISB can promote fairness in a way that is consistent with sound ratemaking practice and can be implemented reliably across carriers. We stand ready to work with DISB to identify appropriate CAS and Academy materials that could provide a more robust, transparent framework.

Key Technical Questions About the Balancing Test
Our members, particularly pricing actuaries, request clarification on the following points, which are essential to assessing both feasibility and impact:

1. Ideal value and objective of the test
   - The proposal suggests that higher balancing-test values reflect better risk prediction with relatively small changes to the premium race gap, but it does not specify what range or threshold would be considered acceptable in practice.
   - Is the primary regulatory goal to reduce premium gaps when segmenting by race, even where rating variables are predictively strong and actuarially justified, or is the goal to identify outlier variables where there is a large fairness impact relative to modest predictive gain? Clarifying the objective is critical to designing an appropriate test.

2. Multi-variable models and interactions
   - The simplified example in the Explainable Fairness article uses a single rating factor (e.g., height) overlaid on a pie chart of high-risk drivers, but real-world auto rating plans rely on dozens of variables and interactions.
   - How will the proposed test handle multiple rating variables simultaneously, including interactions where an individual factor may not align neatly with the high-risk region on its own but the combination of factors does? Without a clear multi-variable framework, there is a risk that the test misattributes predictive value or fairness impact to individual variables that primarily work through interactions.

3. Assumptions about the distribution of high-risk drivers
   - In the illustrative pie-chart example, the authors assume that the ratio of non-white to white high-risk drivers equals the ratio of non-white to white drivers in the overall population, implemented by placing the high-risk circle at the center of the pie.
   - What empirical research does DISB rely on to support this assumption in the context of the District's auto insurance market, and on what basis is this assumption actuarially sound? If the goal is to narrow the premium gap between white and non-white drivers, it is critical to understand whether the ratio of non-white to white high-risk drivers actually matches the ratio in the total driving population. Otherwise, the framework may conflate underlying risk differences with model-based disparities.

4. Documentation and reproducibility requirements
   - Given the complexity of the proposed test, what documentation will DISB require from insurers to demonstrate compliance, and how will differences in implementation (e.g., modeling platforms, sampling approaches, or convergence criteria) be handled?
   - Will DISB provide standard code templates, test datasets, or reference implementations to ensure consistent application across carriers? Without such tools, small technical choices could lead to materially different results and regulatory determinations.

Recommendations and Path Forward
To advance fairness in a way that is both effective and feasible, we recommend that DISB:

- Abandon or pause adoption of the proposed balancing test as a binding regulatory requirement until key methodological details are fully specified, empirically validated, and tested in collaboration with industry and actuarial experts.
- Convene a technical working group—including DISB staff, practicing pricing actuaries, CAS and Academy representatives, consumer advocates, and data scientists—to vet candidate fairness metrics and auditing approaches, focusing on methods that are transparent, stable, and consistent with actuarial standards of practice.
- Pilot any new fairness testing framework through voluntary studies or parallel reporting (e.g., informational filings) before tying it to rate approval decisions, so that all stakeholders can observe real-world behavior and refine the approach.

We appreciate DISB's willingness to take comments on this topic, we share the goal of ensuring that rating and underwriting is based in risk-based pricing, transparent, and consistent with District law. We look forward to working collaboratively on an approach that preserves the benefits of accurate, risk-based pricing for District consumers.

Sincerely,

Gina Rotunno

Gina Rotunno
Regional Vice President
Mid-Atlantic