

1.1.1 Proxy methodology white paper

Fair Lending collaborates closely with the Office of Research, in the Bureau's Division of Research, Markets and Regulations, to conduct statistical analyses of the data we receive in our fair lending work. In order to evaluate a lender's compliance with fair lending laws outside of mortgage lending, the Bureau uses a proxy methodology, as do other federal supervisory agencies and many researchers and private companies. On September 17, 2014, the Bureau published a white paper, entitled *Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity*, that details the methodology the Bureau uses to calculate the probability that an individual is of a specific race and ethnicity based on his or her last name and place of residence. In connection with the release of the report, the Bureau made available the statistical software code and publicly available census data used to build the proxy to enable lenders to replicate the analysis performed by the Bureau.¹ Links to these materials are available on our website.

The Bureau has published white papers on a variety of topics, and welcomes responses to our publications. In this instance our white paper prompted interest in the topic from industry commenters, all of whom appear to agree that disparities in dealer markup exist, but differ over the degree of those disparities. Some industry responses expressed concern about the proxy's tendency to overestimate the number of minority borrowers, particularly African Americans, when compared to a data set of mortgage borrowers. The Bureau's study also observed overestimation, but noted the fact that mortgage applicants are not representative of the general population may account for the overestimation. When the proxy is applied to data where the applicants are more representative of the general population, such as auto lending data, there may not be overestimation or it may decrease significantly.

The Bureau's own analysis demonstrates that *its the DISC proxy probability*, which assigns an individual probability of inclusion in a prohibited-basis group based on *both geography and surname*, is more accurate than a geography-only or surname-only proxy in its ability to predict individual applicants' reported race and ethnicity and generally more accurate than a geography-only or surname-only proxy at approximating the overall reported distribution of race and ethnicity.

Industry responses to the Bureau's white paper also urged the Bureau to incorporate controls based on broad assumptions about the auto lending market and lending practices into our analyses of discretionary markup, in order to explain disparities identified when comparing the interest rates paid by similarly-situated minority and non-minority consumers. Because the Bureau takes a data-driven approach to its work, and given the requirements of the law, we must evaluate the relevance of any proposed controls on a lender-by-lender basis to determine whether

Comment [CL1]: Are there examples beyond the AFSA paper? If you are just referring to the AFSA paper, suggest deleting "all of," which seems to refer to multiple commenters.

Comment [CL2]: Again, is this the AFSA paper/association letter, or are there other responses?

Formatted: Font: Italic

Comment [PAF3]: Cleared answer from [REDACTED]

Comment [CL4]: Is this the AFSA paper or other responses?

¹ <https://github.com/cfpb/proxy-methodology>.

they are legitimate and actually incorporated into decisions about discretionary markup of interest rates on auto loans.²²

Moreover, ~~if when~~ during an examination or investigation, we identify potential disparities, we ~~will continue to share our initial findings with the institution before they are finalized, and to solicit the institution's response and explanation.~~ As part of this process, we have considered, on a case-by-case basis, many of the controls and recommendations listed in the AFSA paper offered by industry and ~~Many of the controls and recommendations are already, where supported by the facts of a particular case, have incorporated them into our analysis what we do, both to test the robustness of our results and to anticipate (and respond to) lender concerns.~~ This process is an ongoing dialogue between specific institutions and the Bureau.

As we stated in white paper, the Bureau is committed to continuing our dialogue with other federal agencies, lenders, advocates, and researchers regarding the Bureau's methodology, the importance of fair lending compliance, and the use of proxies when self-reported race and ethnicity is unavailable. We expect that the methodology will continue to evolve as enhancements are identified that improve accuracy and performance.

Comment [CL5]: Is there additional support for the assertion in the footnote regarding the appropriateness of portfolio analysis that could be included here? Alternatively, it might be worth expanding the discussion of our reasoning regarding this point.

Comment [CL6]: Suggestion.

Comment [CL7]: Findings are currently shared.

Comment [CL8]: Are there mark up analyses in which we have included controls suggested by the institution? Because the previous paragraph references mark up analysis, that context seems implicit here.

Comment [PAF9]: Cleared answer from [REDACTED]

²² Another concern voiced by industry relates to the Bureau's analyses of a lender's entire portfolio, rather than more disaggregated slices. The lender's portfolio is generally the appropriate scope for analyzing whether discrimination has occurred at that lender. CFPB Bulletin 2013-02, Indirect Auto Lending and Compliance with the Equal Credit Opportunity Act (March 2012) ("The disparities triggering liability could arise either within a particular dealer's transactions or across different dealers within the lender's portfolio."). Of course, as described above, the Bureau considers on a case-by-case basis any controls reflecting legitimate, nondiscriminatory, demonstrated factors.