ANDREW
DAVIDSON
& CO

THE IMPACT OF MOVING AWAY FROM THE TRI-MERGE STANDARD

February 2026

WRITTEN BY

Joni Baker
Behavioral Modeling

Sanj Chatterjee
Behavioral Modeling

Richard Cooperstein
Alliances & Policies

Andrew Davidson
President

In July 2025, the US Federal Housing Finance Agency (FHFA) announced that the government-sponsored entities (the Enterprises or GSEs), Fannie Mae and Freddie Mac, would permit lenders to choose between Classic FICO and VantageScore 4.0 credit score models for loans sold to the GSEs. FHFA also stated in a social media post that the tri-merge standard would be maintained for mortgage underwriting. Nevertheless, some mortgage industry stakeholders recommend moving away from the tri-merge standard for GSE mortgages in favor of a single or bi-merge report standard.

To better understand the impact of changing the tri-merge standard to single or bi-merge on the mortgage ecosystem, Andrew Davidson & Co., Inc. (AD&Co) conducted a study to explore the potential impacts on decision-making, based on an examination of the differences in VantageScore 4.0 credit scores across credit bureaus. The analysis is based on an examination of a unique data set of VantageScore 4.0 credit scores of a very broad range of consumers constructed from data provided by the Nationwide Consumer Reporting Agencies (Equifax, Experian, and TransUnion). We used VantageScore 4.0. because this model is the same across all three NCRAs; any variations in scores reflect differences in the underlying credit data rather than inconsistencies that may occur between different scoring formulas.

Results of the study demonstrated that moving away from the tri-merge standard could potentially increase the risk that originators and consumers score shop during the origination process by choosing the credit score (or lender) that produces the lending outcome they desire. Even in the absence of score shopping, such a moving from the tri-merge could lead to less accurate pricing and mortgage qualification. Minority[1] and lower-scoring borrowers would be more heavily impacted. Ultimately, if investors require higher compensation for greater uncertainty, mortgage rates could be higher for everyone.

1. Market Differentiation and Credit Score Variations

Credit scores model the predicted short-term delinquency performance of consumers by placing them into credit score bands, or cohorts, based on credit characteristics such as utilization and delinquency history. Credit scores are computed based on proprietary algorithms that leverage the data from the Nationwide Consumer Reporting Agencies (NCRAs). These credit scores often differ among the NCRAs, in part because each NCRA may have different information about a given consumer, including timing differences in reporting, processing, and correcting the furnished data, due to varying relationships with data providers. Additionally, the NCRAs may acquire unique data elements such as residential property rental and buy now, pay later (BNPL) data that the other NCRAs do not have. However, even when starting with the exact same tradelines containing the exact same raw data, each NCRA has a proprietary set of rules that is applied to the data to transform the data into a finished product for use in the credit scoring models; in addition, while some credit model algorithms, such as VantageScore 4.0, are consistent across NCRAs, other models, such as Classic FICO, use different formulas at each NCRA. Thus, even for consumers with the same raw data inputs reported to each NCRA, different data rules and/or different scoring algorithms at each NCRA can drive score differences for those consumers. These data and processing differences are the result of a competitive NCRA market, in which NCRAs compete to obtain data and deliver the most effective solutions using those data components. Thus, each credit score is an estimate of a consumer's credit risk, subject to variations due to the data differences among the underlying consumer credit

---

[1] Consumer credit reports do not include demographic data such as gender, race, age, nationality, relationship status, education, or religion. For the purpose of this study, proprietary and anonymized third-party demographic data was used for evaluation.

files at each of the NCRAs used to calculate the credit scores. In mortgage lending, the tri-merge standard accounts for any potential credit score variations, thereby providing a more complete assessment of delinquency risk. The tri-merge standard minimizes potential variations in mortgage pricing and underwriting outcomes, compared with relying on just one or two scores and corresponding data, by ensuring that all the available data for a borrower from each NCRA (especially if unique to an NCRA) is included in the underwriting process. This study examined the effects that a policy change from the current tri-merge to a bi-merge or single bureau requirement could have on consumers and investors.

2. Key Findings - Overview

Our analysis shows that credit score uncertainty and mortgage pricing differences increase under a single or bi-merge standard compared to using the traditional tri-merge standard, which utilizes the median of three scores. Section 3 below provides additional definitions and context for our findings, while Figures 3–8 (in Sections 6–9) summarize the metrics arising from our research.

Key findings include:

- Individual score differences were frequent and large enough to have a meaningful impact on consumers and investors; average differences or standard deviations do not tell the whole story. Thirty-five percent of the 245 million scored consumers represented by the study dataset had at least one score that differed from the tri-merge standard by at least 10 points, 18% had a score that differed from it by at least 20 points, and 7% had a score that differed by 40 or more points (see Figures 3–5). These percentages were higher for consumers in the credit score ranges where it matters most (600–779).

- A 20-point difference guarantees that for consumers in the 640–779 range, the loan would move into a higher or lower GSE pricing bucket based on Loan-Level Pricing Adjustment categories, or LLPAs; when applicable, such movements also impact the price of mortgage insurance. However, for consumers in this range, even lower score variances of 10 or more points resulted in a move into different pricing bucket about 83% of the time. As an example, for a $350,000 GSE loan with a 90% loan-to-value (LTV) ratio, moving between consecutive pricing bins can raise or lower the combined cost of borrowing and mortgage insurance (MI) by $3,000 to $5,000 in present value (PV) over the life of the loan.

- The potential for pricing variances due to reduced information is even greater for lower-scoring (i.e., credit scores of 600–639) and minority[1] borrowers, about a quarter of whom were found in this study to have had at least one credit score that differed from the tri-merge standard by at least 20 points.

- In a non-tri-merge landscape, lending and pricing decisions that could be based on different credit scores may create an opportunity for originators to score shop during the origination process by choosing the credit score that produces the lending outcome they desire; consumers choosing between lenders would also, implicitly, be shopping for the best score outcome. Based on study data, about 20% of consumers could increase their purported credit score by 10 or more points from what the tri-merge standard would otherwise show, while 9% could achieve as much as a 20-point increase. In such cases, credit risk would be underpriced, reducing credit investor income below fairly priced risk, and potentially leading to consumers appearing qualified for mortgages they cannot afford.

- Establishing a score cutoff such as 700 to determine whether a tri-merge is required does not eliminate the existence of meaningful score discrepancies. For example, 18% of consumers in the 700–779 range had at least one score that differed from the tri-merge standard by 20 or more points, and a randomly

chosen score resulted in consumers within this range moving to a different LLPA bin (either higher or lower than the tri-merge standard) about 16% of the time—or 13%, if random outcomes below 700 were assumed to result in a tri-merge. More troubling, almost 6% of consumers with median below 700 had a maximum score of 700 or more, with the percentage varying by credit score bin (e.g., 8% of consumers in the 660–679 bin had a maximum score above 700).

3. GSE Loans, Credit Scores, and LLPAs

Currently, when the GSEs purchase a loan, they require a tri-merge credit report that combines tradeline data from all three NCRAs (Equifax, Experian, and TransUnion). For a given credit score model, such as FICO Classic (which until recently was the only approved model), this standard requires that lenders use the median of the three credit scores as the representative credit score for the borrower. The FHFA has considered changing the tri-merge requirement to allow lenders to provide just two credit reports and scores to the GSEs. Under this dual score, or bi-merge requirement, a consumer's representative score would be the average of the two scores pulled. In this paper, the term "tri-merge score" ("tri-merge" or "3B") always refers to the median of three scores, while "dual score" or "bi-merge" ("2B") refers to the average of two scores. Meanwhile, a single report framework ("1B") involves using just one credit score to represent a consumer's credit risk.

Since 2008, the GSEs have used Loan-Level Pricing Adjustments (LLPAs), risk-based fees that determine the cost of borrowing. LLPAs are calculated based on grids with rows reflecting representative credit scores in 20-point increments, from 640 to 779, and columns based on the mortgage LTV ranges. When the LTV ratio is greater than 80%, the consumer must generally purchase mortgage insurance, the cost of which is determined using a similar pricing grid based on credit scores and LTV. Thus, for consumers in the 640–779 score range, a 20-point difference in score guarantees a move to a different LLPA bin (and the differential pricing that results), while a 10-point score difference represents roughly a 50% chance of moving to a different LLPA bin, depending on the underlying score and whether the difference is higher or lower.

While the GSEs base loan qualification on all the data in a consumer's tri-merge credit report, 620 has been the minimum required representative score for GSE loans (in other words, 620 was necessary, but not sufficient, for approval). In mid-November 2025, Fannie Mae joined Freddie Mac in moving away from such a "hard" cutoff[2]; however, loans with scores below 620 presumably need compensating factors for approval, and individual mortgage companies may impose their own lender overlays. Thus, within this paper, we include examples relating to the traditional 620 threshold for conventional loans, with the understanding that the effective cutoff for an individual consumer and loan will vary depending on context and product. The upshot is that a decision must be made (and a qualifying loan must be priced) based on the information and score given, and this will vary depending on the NCRA data used for the score calculations. In addition, as we will see in section 5, consumers on the lower end of the credit score spectrum, where loan qualification is more likely to be an issue, have larger score variation between the NCRA data sets.

---

[2] Announcement SEL-2025-09: Selling Guide Updates," Fannie Mae, November 5, 2025, https://singlefamily.fanniemae.com/news-events/announcement-sel-2025-09-selling-guide-updates

4. Methodology

For this research, Equifax, Experian, and TransUnion provided AD&Co with data on an anonymized sample of consumers from October 2023 that included:
- VantageScore 4.0 credit scores
- Third party race/ethnicity data (when available)
- A flag indicating low-to-moderate income (LMI) level (when available)[3]

The sample was from the entire US population, not just consumers holding mortgages—and more specifically, not just consumers holding GSE-backed mortgages (those scores are higher on average, while Table 1 below shows that score variation increases on the lower end). The sample included all of the approximately 120 million consumers with VantageScore 4.0 scores less than 700 (102 million of which were scorable by data from all three NCRAs), and a 10% sample of the 144 million consumers with VantageScore 4.0 scores of 700 and above. AD&Co matched the resulting scores to each consumer record using a unique, anonymized ID, analyzed the observed score differences (see Section 5), and compared bi-merge and single score outcomes against the tri-merge standard (Section 6). Sections 7 and 8 discuss the implications for pricing and potential gaming, while Section 9 discusses the differential outcomes in the event a score cutoff such as 700 is used to limit the use of a single file approach. Note that the metrics and statistics in this paper are based on counts that were weighted to remove sampling bias; each consumer in the under-sampled population (with credit scores 700 and above) represented 10 consumers in the population at large.

Our study generally describes the absolute value of score differences, comparing scores derived using bi-merge and single report frameworks to tri-merge scores, which served as the baseline (Section 6). By using the absolute value score differences, we considered score variances in either direction—up or down. Regardless of whether 2B or 1B scores exceed or fall below tri-merge scores, material score differences matter. In either instance, some entity in the mortgage ecosystem suffers. When 2B or 1B scores fall below tri-merge scores, consumers may end up paying more for their mortgage. Consumers on the margins of approval could even have their application denied. Score differences in the other direction, where 2B or 1B scores exceed scores produced under a tri-merge standard, primarily, though not exclusively, harm the entity holding the credit risk—generally the loan guarantor and/or insurer, who may be undercompensated for the risk they've assumed. Along with credit risk holders, consumers may also be harmed by inaccurately high scores. If inflated 2B or 1B scores cause lenders to extend credit to consumers who are later unable to repay and face foreclosure, those consumers would likely be worse off than if they'd been denied credit at application. For these reasons, inaccurate scoring in either direction can cause incremental strain on credit risk holders, higher interest rates and payment stress for a portion of borrowers, and increased uncertainty across the system, ultimately having the potential to increase rates for all borrowers.

5. Raw Credit Score Differences

In this section we describe the underlying individual VantageScore 4.0 differences (between two given scores for the same consumer) that were observed after combining the datasets by anonymized ID. Such analysis gives a sense of possible score variation when just a single score is used as a consumer's representative credit score, or when the scores are combined in various ways.

---

[3] The low-to-moderate income designation is based on the consumer's family income compared to the median family income of the relevant metropolitan statistical area (MSA) or the non-MSA portion of a state. Low-to-moderate income consumers are those with family income that is below 80% of the area median family income.

Figure 1 and Table 1 display the distribution of (absolute value) VantageScore 4.0 score differences among consumers that were scored based on data from each of the three NCRAs, when two scores were chosen randomly for each consumer and compared against each other. The red dotted line displays the distribution over all consumers, while other lines (and table columns) display the differences for various subgroups. The White and Minority Groups 1 and 2 are based on the race/ethnicity field mentioned previously, while the LMI groups are based on the low-to-moderate income flag. Such segmentations allow us to see the potential differing impacts that altering the tri-merge standard could have on different subgroups and underserved populations. Meanwhile, for credit score bands, we focused on the groups of consumers between 600 and 639 (near the traditional 620 threshold for GSE loans), 640–779 (internal LLPA bins), and 780+.

As seen in column 1, the two credit scores were identical 42% of the time, but 27% of the time they differed by at least 10 points. For 14% of the score pairs, the difference was at least 20 points, and for 4% it was 50 or more points. Scores for LMI and Minority Groups varied more frequently than other groups, with 17% of LMI consumers and 16%–18% of consumers in Minority Groups having differences of 20 points or more between the two chosen scores.

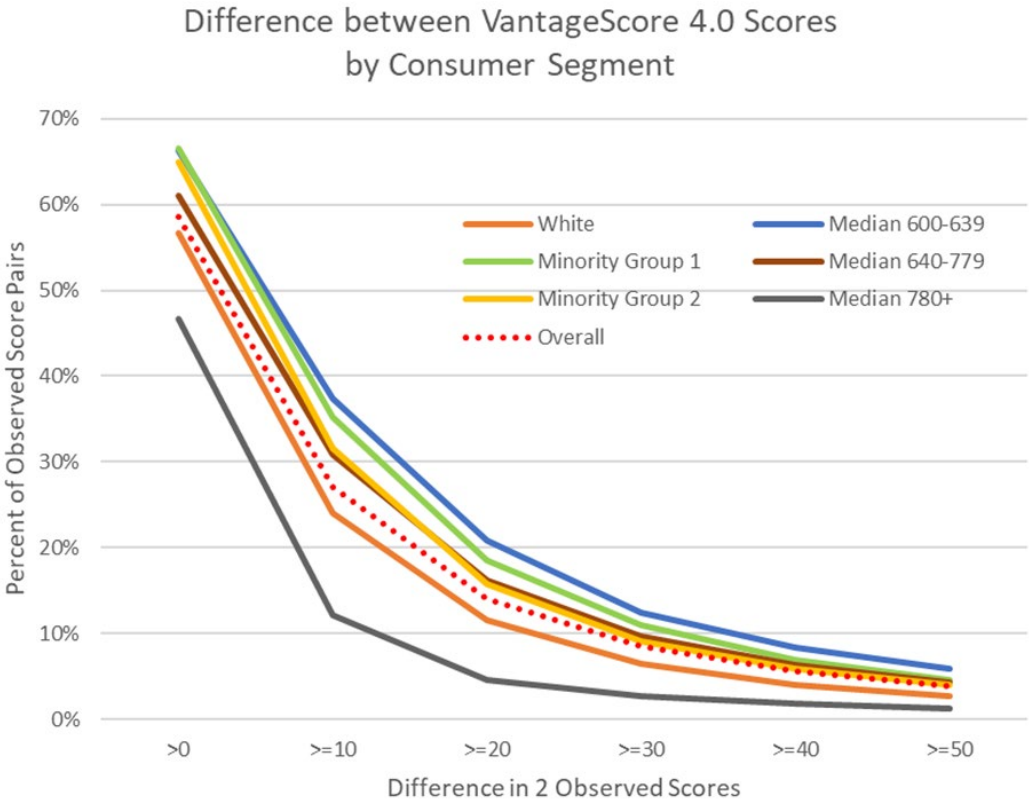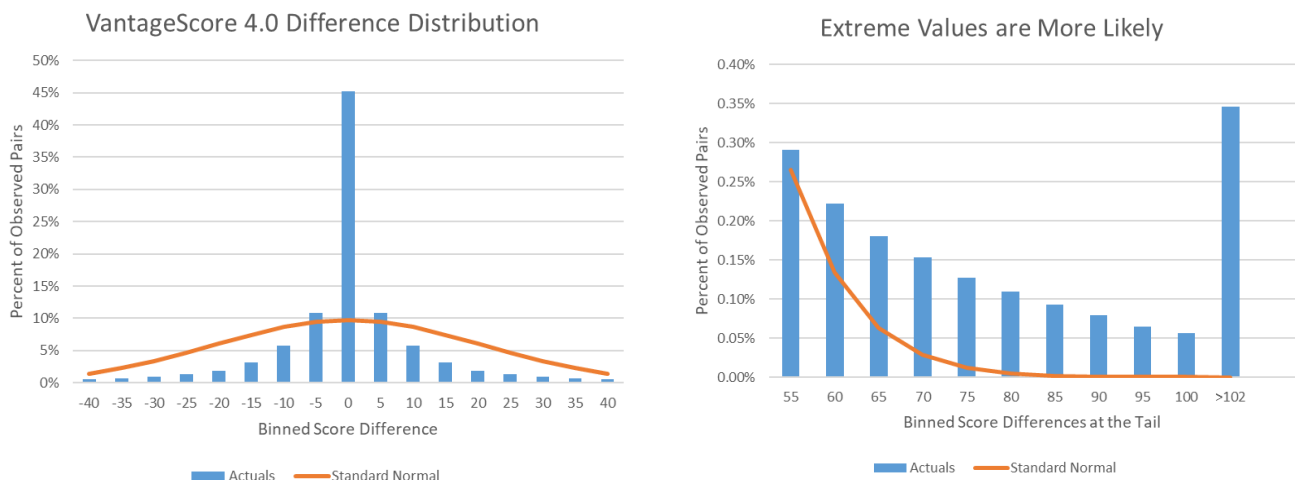**Figure 1. Distribution of VantageScore 4.0 Pair Differences**

Table 1. Distribution of VantageScore 4.0 Pair Differences

	Percentage of VantageScore 4.0 Pairs Having Differences Above the Given Threshold								
Absolute Value Difference	Overall	White	Minority Group 1	Minority Group 2	LMI = 0	LMI = 1	Median 600-639	Median 640-779	Median 780+
=0	42%	43%	33%	35%	44%	38%	34%	39%	53%
>=10	27%	24%	35%	32%	23%	31%	37%	31%	12%
>=20	14%	11%	18%	16%	12%	17%	21%	16%	5%
>=30	9%	7%	11%	9%	7%	11%	12%	10%	3%
>=40	6%	4%	7%	6%	4%	7%	8%	6%	2%
>=50	4%	3%	5%	4%	3%	5%	6%	4%	1%

An important point about these score differences is that they do not follow a normal distribution: both small and large values appear with greater frequency than they would for a normal distribution having the same standard deviation. In general, the standard deviation of score differences is 20 points. Figure 2 displays the actual distribution as a histogram of five-point buckets centered at 0 (e.g., "0" represents the interval of differences from -2 to 2). Overlaid on this histogram is the equivalent weight for a (binned) normal distribution having the same standard deviation. As can be seen in the left-hand chart in Figure 2, the actual distribution is highly concentrated in the center bucket, where there is a large peak. The remaining bars appear to lie almost entirely below the normal distribution. However, the right-hand chart displays a magnified portion of the tail, where the score differences are above 50, illustrating the fact that the actual distribution has heavier tails than normal; extreme differences are much more likely to occur.

In summary, the differences between scores are not normally distributed. Score differences are often near zero but have more extreme highs and lows than a typical bell curve. Therefore, the usual metrics such as average score differences and standard deviations do not tell the whole story; binning the score differences (and other derived metrics) yields a more complete picture.

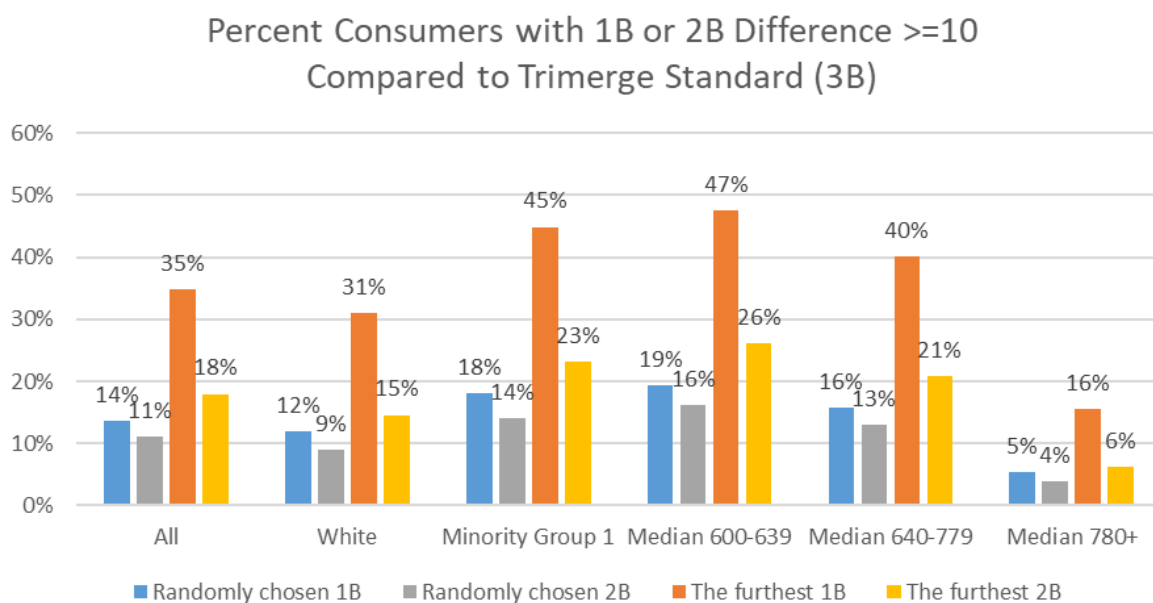## Figure 2. VantageScore 4.0 Difference Distribution and Close-Up of Tail

6. Single and Bi-Merge Scores vs. the Tri-Merge Standard

We now analyze the potential score differences that result when moving from the tri-merge standard to just one score or the average of two.

Figure 3 displays a selection of metrics describing the frequency with which single and dual scores (1B and 2B) differed from the tri-merge (3B) by at least 10 points. Note that while a 10-point minimum difference may seem negligible, analysis of the data showed that within the 640–779 score range, these 10+ point differences resulted in a shift to a different LLPA bin about 80%–83% of the time (for 2B and 1B, respectively). The results here are broken down by six consumer segments, with the leftmost set of bars describing the set of all scored consumers represented by the study dataset. First, the blue and gray bars display the percentage of single (blue) and bi-merge (grey) scores that differed from the tri-merge standard by at least 10 points when the scores were selected randomly for each consumer. This is what we would expect to see, on average, in the absence of any gaming or adverse selection. For example, over all consumers, the randomly chosen single score differed from the tri-merge standard (either higher or lower) by 10 or more points about 14% of the time. This frequency increased to 18% for Minority Group 1 and to 16% for consumers in the 640–779 score range. For a random choice of bi-merge, the frequency of 10-plus point difference compared against the tri-merge score fell to about 11% overall and 13% for those with standard tri-merge scores between 640 and 779.

Meanwhile, the orange and yellow bars display the percent of consumers for which the single (orange) or bi-merge (yellow) score that lay furthest from the tri-merge standard differed from it by at least 10 points. In other words, for these consumers there was at least one choice of a single-score or bi-merge score that differed from the tri-merge by at least 10 points (the selection of such a score could merely be due to random chance or, as discussed in Section 8, could be purposefully chosen, or "shopped"). About 35% of all consumers had at least one single score 10 or more points from the tri-merge. However, for consumers in Minority Group 1 or in the 600–639 range, the potential for such a score difference rose to about 45%–47%.

Figure 3. Single Report and Bi-Merge 10+ Differences vs. Tri-Merge Standard

Percent Consumers with 1B or 2B Difference >=10 Compared to Trimerge Standard (3B)

Legend: ■ Randomly chosen 1B ■ Randomly chosen 2B ■ The furthest 1B ■ The furthest 2B

Figures 4 and 5 display the same set of metrics, but with a focus on 20- and 40-point differences. Variances of this magnitude guarantee a move of at least one or two 20-point bins, respectively. Within the 640–779 range, 21% of consumers had at least one single score that would result in a minimum 20 point shift, and a random selection of one score resulted in such a shift about 8% of the time. In addition, 8% of consumers in this range had at least one single score that differed a full 40 points or more from the tri-merge standard; a random selection of one score resulted in a minimum 40 point shift about 3% of the time.

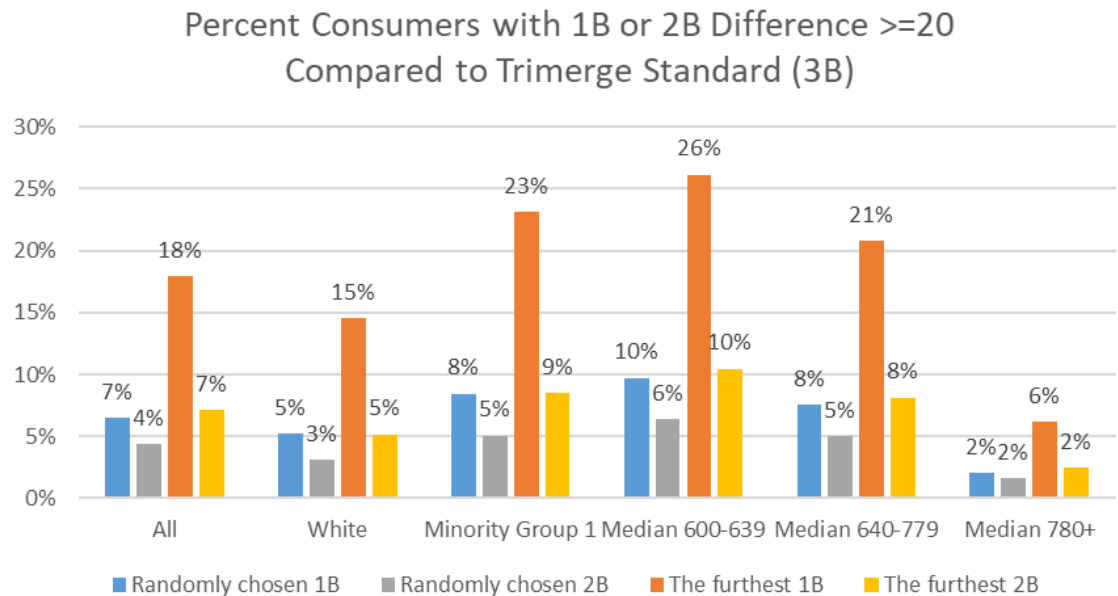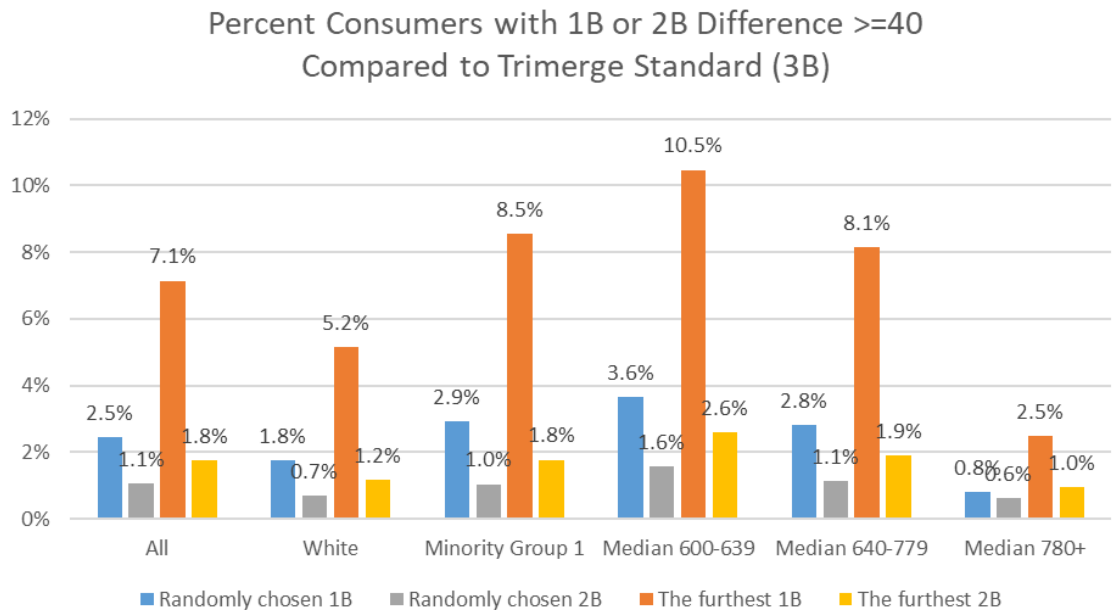## Figure 4. Single Report and Bi-Merge 20+ Differences vs. Tri-Merge Standard

Percent Consumers with 1B or 2B Difference >=20
Compared to Trimerge Standard (3B)

Legend: Randomly chosen 1B, Randomly chosen 2B, The furthest 1B, The furthest 2B

Figure 5. Single Report and Bi-Merge 40+ Differences vs. Tri-Merge Standard

Percent Consumers with 1B or 2B Difference >=40
Compared to Trimerge Standard (3B)

Legend: Randomly chosen 1B, Randomly chosen 2B, The furthest 1B, The furthest 2B
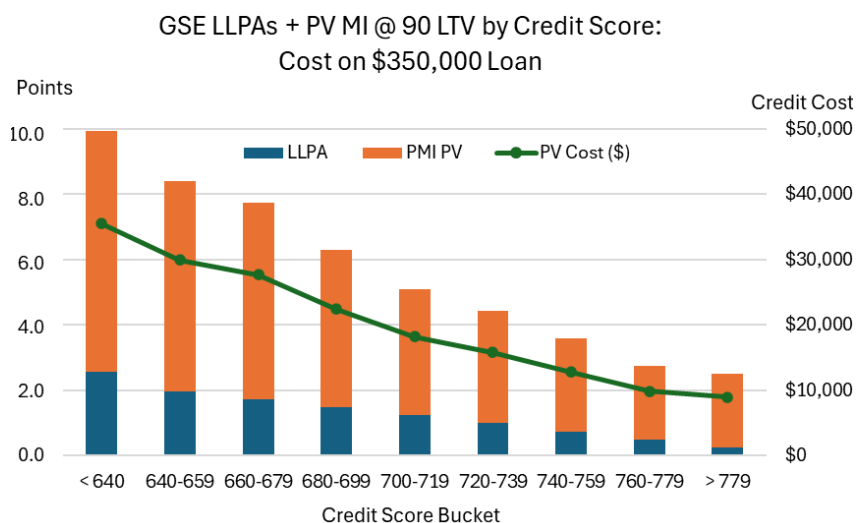
7. The Impact of Increased Uncertainty on Pricing and Qualification

As shown above, for about 26% of consumers in the 600–639 range and 21% of those in the 640–779 range, it was possible to select a single score that differed from the tri-merge median by 20 or more points. However, we found that even randomly selecting a single score in place of the tri-merge standard would move the 600–779 consumers into a different 20-point bin about 19% of the time (note that differences of less than 20 may result in a move between bins). Even if unbiased, such moves increase uncertainty for credit risk takers and may lead to pricing increases for all consumers (the GSEs currently guarantee about 28 million loans).

Figure 6 shows consumer cost for GSE LLPA plus present value (PV) of mortgage insurance by credit score, assuming a $350,000 loan with a 90% LTV ratio. (Note that the corresponding LLPA bucket for LTV runs from >85 to <= 90; we use "90" as shorthand for loans that fall in this bucket.) We consider these values generally representative, as 30% of GSE purchase money loans since 2023 have LTVs of 90% or higher, and 40% of first-time buyer loans during this period have LTVs of 90% or higher. Thus, the 90% LTV bucket is economically significant. For these loans, moving between credit score LLPA bins can cost borrowers or insurers between $3,000 and $5,000 in PV, depending on credit score bucket with the impact generally increasing at lower score levels. The costs for 95% LTV loans would be even higher. About 7%, or 1.9 million, current GSE borrowers have original LTVs of 90, with another 16% at 95 or more, representing about 6 million loans total.[4] So, nearly one-quarter of GSE borrowers have original LTVs of 90 or more and therefore have a meaningful chance of a several thousand dollar change in the cost of finance by using one random score out of three.

Importantly, at the lower end of credit scores, some borrowers could fall below the credit score threshold for qualification. As seen in Table 1, consumers in the 600–639 credit score range had a greater amount of score variation than the population at large. A randomly selected single score from this group differed by at least 10 points from the tri-merge score about 19% of the time (Figure 3). Furthermore, using the lowest of the three scores, about 30% of the 13 million scored consumers in the 620–639 range (or the 240,000 GSE loans with VantageScore 4.0 in this range) could fall below 620 and potentially be denied a GSE mortgage loan. (The opposite problem of lenders specifically selecting the scores that would qualify consumers is discussed in the next section.)

Figure 6. Loan-Level Pricing Adjustment by Credit Score for 90% LTV

GSE LLPAs + PV MI @ 90 LTV by Credit Score: Cost on $350,000 Loan

---

[4] Query of GSE mortgages as of December 2025, using Recursion.

8. The Consequences of Distorted Incentives: Loopholes and Score Shopping

A move away from the standard tri-merge credit score creates an information gap. Because of the tri-merge standard, this information asymmetry doesn't currently exist, but would be introduced by going to a bi-merge or single score requirement. Lenders would suddenly have more credit data on a consumer than everyone else in the market. This creates the opportunity for score shopping—where lenders can pick the consumer's best credit score(s) to make a loan look as attractive as possible. We have quantified this incentive at $3,000–$5,000 in PV for an LTV of 90, which accounts for almost a quarter of first-time buyer loans. As this potential gain represents a substantial share of originator economics, the incentive to engage in score shopping would be significant. Consumers who comparison shop for the best deal would also tend to choose a lender that uses their highest score, regardless of whether lenders themselves score shop. Based on this study, for consumers in the 640–779 range, under a single-score requirement, score shopping (or comparison shopping) to select the highest score would result in a lower priced bucket about 26% of the time.

It is important to consider how the GSEs, mortgage insurers, and other credit investors may react to this potential incentive and resulting loss in credit risk fees. From the perspective of these stakeholders, it would be logical, indeed prudent, to assume score shopping occurs frequently and, based on this assumption, raise loan pricing across the board to offset the risk—which would penalize all borrowers. Increasing loan pricing could reinforce the benefits of score shopping, which could increase its frequency and lock in less precise credit pricing.
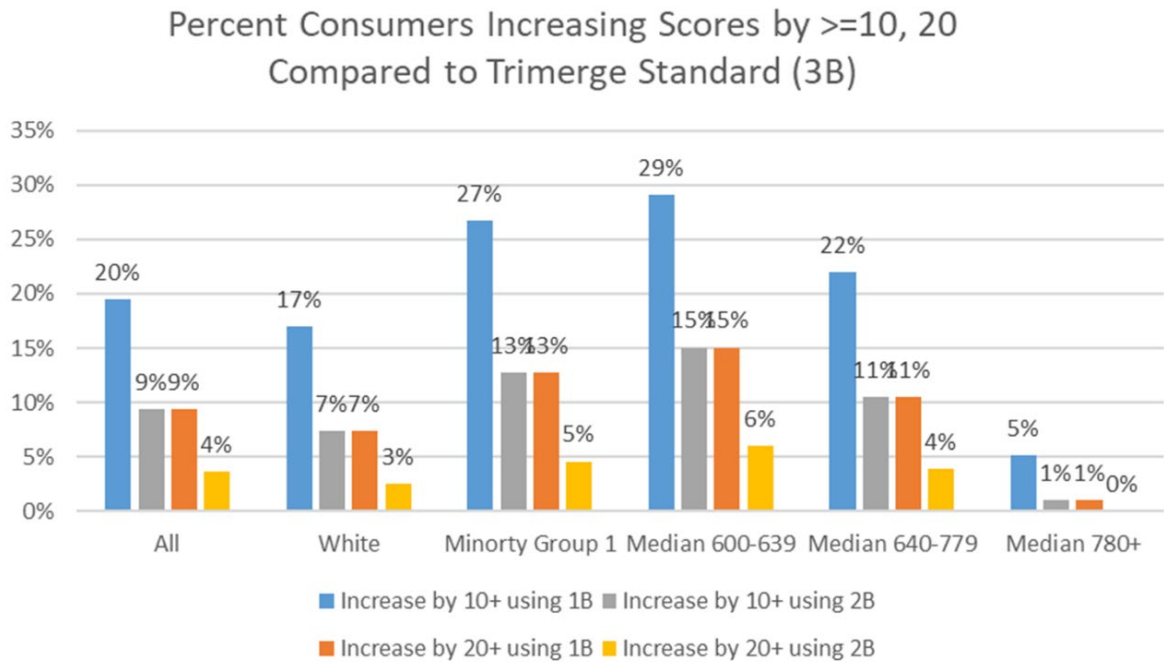
Concerns related to score shopping are particularly acute when considering consumers with scores on the margin of approval. More than 10 million scored consumers in the research dataset had a median score between 600 and 619. For consumers in this range, randomly selecting just one of the three scores would result in the consumer exceeding the traditional threshold of 620 about 12% of the time. Compared with tri-merge-based decisions, these could be viewed as errors or underpricing. However, through score shopping, 36% of the consumers in this group would exceed the 620 threshold based on their single highest score. A consumer who should not have qualified and who ultimately defaults is potentially the biggest loser in this case.

A related concern is credit washing, where consumers or third parties attempt to remove legitimate derogatory tradelines from a single bureau file through false disputes or identity-theft claims. Today, the tri-merge standard effectively prevents this behavior, as it is nearly impossible to "wash" all three bureau files at once. Moving to a single- or bi-merge requirement could increase both the opportunity and incentive for credit washing, allowing manipulated files to mask underlying risk. Existing score disparities illustrate this vulnerability: 4% of consumers with a maximum score between 700 and 759 had a median score at least 40 points lower, 2% had gaps of 60 points, and 1% had gaps of 80 points or more. Such single-bureau anomalies could become more consequential without the tri-merge safeguard.

The GSEs generally do not use credit scores for risk assessment, but rather the full in-file tri-merge data set. Scores are generally used only for pricing by the GSEs. However, if tri-merge data is no longer pulled at origination, the GSEs would be at the same information disadvantage unless they pull the tri-merge data set themselves. Adopting this approach and pulling tri-merge data themselves would create new duplicative infrastructure and matching expenses, while potentially introducing operational complexity and data mismatches based on timing of credit pulls. If mortgage insurers, re-insurers, and credit investors don't independently pull consumer credit data, they will be at a disadvantage as well.

Figure 7 shows how the potential benefits of score shopping vary by consumer demographics and creditworthiness, based on the study dataset. Note that under a single score framework, 22% of consumers in the 640–779 range could, by using their highest score, improve on their standard tri-merge score by at least 10 points, and 11% could improve by at least 20 points. In fact, even improvements of 40 or more points (not shown here) would not be rare: 6% of consumers in the 600–639 range and 4% of those in the 640–779 range could improve by 40 or more points by using the highest of three scores.

Figure 7. Differences When Choosing the Highest Single Report and Highest Bi-Merge Report vs. Tri-Merge Standard

9. Basing the Tri-Merge Standard on a 700 Cutoff

Recently, some in the industry proposed a hybrid approach based on an initial score threshold of 700. Under this proposal, if the first-pulled score is 700 or higher, it serves as the consumer's representative score. If the initial score falls below 700, the standard tri-merge process applies, and the median score is used.

In assessing this approach, our analysis centered on two different consumer groups: those with a traditional tri-merge median score below 700, and those in the 700–779 range. Table 2 shows that nearly 6% of the 100 million scored consumers in the first group—those with a tri-merge median score below 700—had a maximum score of at least 700; this maximum score, if pulled, would be used as the representative credit score for the consumer. The frequency varied by credit score band; thus, for example, 4% of consumers with a median between 640 and 659 had a maximum score exceeding 700, with this percentage increasing to nearly 8% for consumers in the 660–679 score band and 29% for those in the 680–699 band.
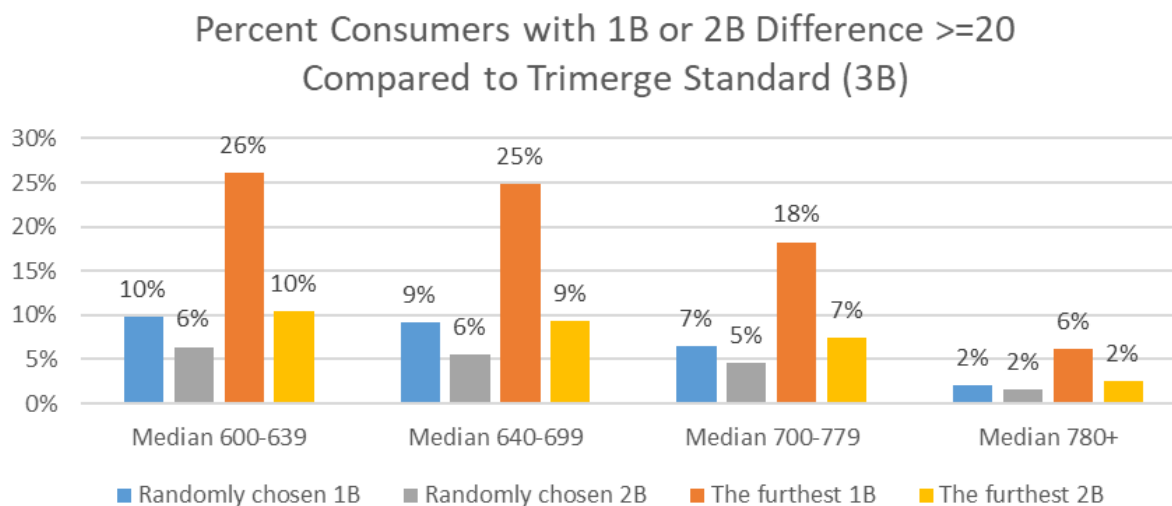
Since not all consumers in each band are looking to attain a mortgage, Table 2 also includes a column that indicates the number of existing GSE mortgages in each credit band, to give a sense of the actual market. The right-most column then uses the sample loan from Section 7 to show the approximate potential shortfall in credit pricing (via LLPA fees and private mortgage insurance) for a $350,000 90 LTV loan that is priced based on a credit score of at least 700. It is evident that score shopping would be a large risk for consumers having standard credit scores below 700.

Table 2. Consumers With Tri-Merge Score Below 700

Credit Score (median)	Approx. number of GSE loans (Dec 2025)[4]	Consumers represented in study dataset	Consumers in study having max score 700+	Percent of consumers with max score 700+	Approx. LLPA+PMI PV cost differential (vs. 700) for one $350K loan having LTV of 90
<600	61K	41,698,106	112,410	0.27%	
600-619	47K	10,560,522	128,530	1.22%	$17,200
620-639	360K	12,920,296	312,579	2.42%	$17,200
640-659	650K	12,022,850	489,894	4.07%	$11,700
660-679	1 million	11,338,124	886,999	7.82%	$9,400
680-699	1.8 million	13,141,827	3,823,957	29.10%	$4,200
Grand Total	**4 million**	**101,681,725**	**5,754,369**	**5.66%**	

As for consumers in the 700–779 credit score band, under the 700-cutoff proposal most of these would have just a single score as their representative score. For these consumers, the question again is how much the single score could differ from the tri-merge standard. Figure 8 repeats some of the metrics from Figure 4, focusing on score differences of at least 20 points but splitting the 640–779 range into two segments above and below 700. As might be expected based on the prior metrics, score variation decreases over the interval of 700–779, however, 18% of these consumers still had at least one single score that differed by at least 20 points from the tri-merge standard (compared with 21% of consumers in the full range 640–779).

Figure 8. Single Report and Bi-Merge 20+ Differences vs. Tri-Merge Standard

A 20-point difference guarantees that the consumer will move LLPA bins; however, smaller differences can also result in a move between bins. A randomly chosen single score resulted in consumers within the 770–779 range moving to a different LLPA bin (either higher or lower than the tri-merge standard) about 16% of the time—or 13%, if random outcomes below 700 were assumed to result in a tri-merge. Thus, even ignoring any incentive to score shop, risk could be mispriced for these consumers a meaningful amount of the time.

Conclusion

A credit score predicts a consumer's credit risk, and the score may vary based on the data from the three NCRAs; therefore, using the tri-merge score captures the more complete picture of a consumer's risk. This study shows that moving to a single score or to a bi-merge approach increases the uncertainty in assessing borrower risk, with direct implications for loan pricing and underwriting outcomes; this uncertainty is greater for minority and lower-scoring borrowers. Moreover, large differences in scores occur much more often than would be seen in a normal distribution. Single bureau and bi-merge scores often produce large discrepancies compared to tri-merge results; for example, 18% of consumers had a single score that differed from the tri-merge standard by at least 20 points, and 7% had a score that differed from it by 40 or more points. This can cost higher LTV/lower-score borrowers (or investors in such mortgages) thousands of dollars in mispriced fees and risk. Similarly, we have shown that the potential gain from score shopping (i.e., selectively choosing which NCRA data is used/reported) is several thousand dollars, which is large compared to the costs of origination. In short, this study showed that the call to move away from the tri-merge standard could have a meaningful negative impact and may not result in the most optimal outcome in terms of risk and price assessment for consumers or investors.