An Assessment of the Impact

Of the Premature Release of Subject-Matter Content on

the July 2019 California Bar Examination

Prepared By

Research Solutions Group

November 13, 2019

EXECUTIVE SUMMARY

A premature release of the written portion subject-matter areas of the July 2019 California Bar Examination (CBX) to deans from 16 California law schools led the State Bar of California ("State Bar") to inform all test-takers of the topics less than one week before the exam administration[1]. This had led to concerns that the results of the examination would, in some manner be impacted. In response, the State Bar requested that statistical analyses be conducted to address the concerns. Specifically, the analyses addressed three issues:

- To what degree, if any, did the performance on the July 2019 CBX change as a result of the premature content release of the examination?
- To what degree, if any, was the CBX passage rate systematically impacted by the release?
- Was the performance of the students from the 16 law schools whose deans were made aware of the content of the written portion of the examination, differentially impacted relative to the performance of all other applicants?

Employing historical CBX and national data published by the National Conference of Bar Examiners (NCBE), the analytic methodology used to examine these issues first established historical patterns and relationships and then compared them to results from the July 2019 CBX administration.

The results of the following analyses indicated that the premature release of the content had no statistically significant impact on the results of the July 2019 examination.

Based on a regression model documenting the strong and consistent relationship between national and California-specific Multistate Bar Examination (MBE) performance between 2004 and 2018, it was first established that California scores on the MBE in 2019 did not deviate from what would be predicted given results from the rest of the country. A second regression analysis indicated that the July 2019 CBX bar passage rate did not differ statistically from predicted rates given the current CBX MBE performance and the historical relationship between bar passage and MBE scores in California. Application of a third set of multiple regression models indicated that actual written score performance of graduates from the 16 law schools did not differ from expected results. The findings were true for students from all 16 schools as a group, and separately for each law school. Finally, we observed that the statistical relationship between written scores and MBE scores on the July 2019 CBX did not significantly differ from that of recent July examinations, lending further evidence that performance on the written section of the examination was not systematically impacted.

---

[1] We abbreviate this event as "early release" throughout the report.

I. BACKGROUND

After the California Bar Examination (CBX) is administered, a group of deans from a sample of California law schools are routinely sent e-mail invitations to attend the second calibration session, a meeting designed to give law schools insight into the methods used to grade the written portion of the CBX. The invitation contains the legal subject-matter areas (e.g., Torts, Constitutional Law) that were covered on the written portion of the examination, so that the deans can opt to send other representatives from their faculty to the workshop. For the July 2019 administration a similar invite was prepared, but inadvertently sent out to deans from 16 law schools at the end of the week *preceding the administration of the examination*.

Two days following the invitation the State Bar of California ("State Bar) discovered the mistake and made the decision to notify all applicants who would be taking the examination about the error and informing them of the specific subject-matter areas. While applicants sitting for a bar examination in California are aware of the 13 potential legal subject-areas on which they may be tested, this was the first time that they knew which of those areas the essay and performance task would cover.

Concerns were voiced to the State Bar by various individuals that students from the 16 law schools might have had an unfair advantage over all other test takers; arguing that those students possibly could have had additional focused study, resulting in higher performance on the CBX. Additional comments were made that the error could also lead to higher overall passage rates on the July 2019 CBX, or, conversely, harder grading, resulting in a lower passage rate.

In response, the State Bar requested that statistical analyses be conducted to empirically determine whether, and to what degree these issues and concerns were valid. The intent of this report is to document the statistical methods used to address these issues, including the data that was used, the results of the analyses and the subsequent conclusions that were drawn.

2. RESEARCH ISSUES AND ANALYTIC APPROACH

Three key issues were studied in this report.

- To what degree, if any, did the performance on the July 2019 CBX change as a result of the premature content release of the examination?

- To what degree, if any, was the CBX passage rate systematically impacted by the release?

- Was the performance of the students from the 16 law schools whose deans were made aware of the content of the written portion of the examination, differentially impacted relative to the performance of all other applicants.

Analyses of each of these issues were conducted in two stages. In the first stage, available historical data was examined to establish a baseline profile of the relevant CBX performance outcome (e.g., raw scores, pass/fail status) and to assess that profile to that of a relevant comparison group. Once the profile was established, performance on the same July 2019 CBX outcome was then calculated and matched to the performance of the comparison group's 2019 outcomes. It was reasoned that if the historical relationships that were observed in the years preceding the July 2019 CBX did not change during the July 2019 CBX, the findings would suggest that the early subject-topic release did not have an impact. If the July 2019 CBX relationships did change, then it might suggest that the early release might have had an effect.

The analytic approach was influenced by several factors. First, raw score performance on the Multistate Bar Examination (MBE), the multiple-choice portion of the CBX, is "equated" annually such that scaled scores on this portion have comparable meaning across years, i.e., identical scaled MBE scores on two separate administrations, imply identical skill/knowledge levels. Thus, MBE performance served as a valuable cross-year anchor for the analyses. Secondly, the subject matter areas on the MBE are consistent across years and known to everyone in advance of the examination. Therefore, we would assume no particular group would have an advantage over another. Concerns were reportedly raised to the State Bar by some that the early release raised test-anxiety levels that actually could have bled over to the MBE. Given the importance of the MBE, as an anchor in our analyses. we evaluated the possibility that MBE performance might be affected as well as the written sections. We also note that analyses utilizing historical data focused on July administrations only because of the known differences in the composition of applicants sitting for the July and February exams.

3. RESULTS

Issue 1. To address the first issue related to the impact of the early topic release on the performance on the 2019 CBX, we first examined performance on the MBE by studying the relationship between historic California and United States-wide trends and then using those results to compare actual to expected results on the 2019 administration.

Approach. In this analysis, we first collected 15 years (2004-2018) of July examination MBE performance.   U.S. data was obtained from the published National Conference of Bar Examiners (NCBE) archives (www.ncbex.org).   For each administration, the average MBE performance and number of test-takers was recorded. For those same examinations, the average MBE performance and number of test-takers were calculated for California, using data available from the archive files of the State Bar. To isolate the non-California results from the U.S. Data, for each administration, we subtracted the number of California test-takers and the sum of their MBE performance, and then recalculated the U.S. average MBE performance. Using Ordinary Least Square (OLS) regression methods, we then compared actual to expected results.

Findings. The original and adjusted data can be found in Table 1 on the following page. Table 1 also contains a column for the difference between the Adjusted U.S. and California MBE performance for each administration. As can be seen in Table 1, the average performance on the MBE is consistently higher than the rest of the country. The average difference across the 15 administrations was 24 points (Sd. = 4.29 points).

We next estimated the correlation between the historic California and adjusted U.S. MBE performance to determine if there was a meaningful historical relationship using the annual performance data (N=15 administrations). The resulting correlation was quite high[2] (r=.97) and statistically different from zero (p< .0001). Given this strong relationship, we then calculated the regression equation on that data and used the equation to estimate what the California performance on the MBE would be based upon U.S. performance in a given year. The resulting equation and related statistics are presented in Table 2[3].

The results in Table 2 indicate that the overall regression model was very accurate in explaining the historic relationship between California and U.S. MBE performance as evidenced by the fact that 95% of the variability in California MBE performance could be explained by the corresponding variability in U.S. performance.

---

[2] Correlations range from -1.00 to +1.00
[3] A simple OLS regression equation consists of an "intercept" which anchors the equation (i.e., the expected outcome when the value" of the predictor is zero) and a regression "weight" that serves as a multiplier. The "expected value equals the intercept plus the product of the weight times the new value (in this case, the 2019 adjusted U.S. average MBE)

The same MBE data used to establish the historic relationships were then collected for the current July 2019 administration. A similar set of calculations were made to adjust the 2019 NCBE-based U.S. data to exclude California. The adjusted-U.S. data was then entered into the regression equation referenced above to *estimate* what the California-specific performance would be under the assumption of et ceteris paribus (i.e., all things remained equal). If the estimated California MBE performance was similar (i.e., not statistically different) to the actual performance, then there would be supporting evidence that the early release of topics for the written portion of the examination did not have a discernible effect on the overall MBE performance in 2019[4].

Table 1

California and U.S. MBE Performance

2004-2018

July Administrations

	California		U.S.		
Year	Examinees	MBE Ave	Examinees	MBE Ave	Adj. MBE Ave.
2004	8,020	1,434	47,433	1,412	1,408
2005	8,310	1,437	49,998	1,416	1,412
2006	8,858	1,452	51,176	1,433	1,429
2007	8,115	1,459	50,181	1,437	1,433
2008	8,590	1,476	50,011	1,456	1,452
2009	8,607	1,463	50,385	1,445	1,441
2010	8,521	1,454	50,114	1,436	1,432
2011	8,412	1,458	49,933	1,438	1,434
2012	8,664	1,460	52,337	1,434	1,429
2013	8,822	1,461	53,706	1,443	1,439
2014	8,428	1,436	51,005	1,415	1,411
2015	8,236	1,426	48,384	1,399	1,393
2016	7,648	1,423	46,518	1,403	1,399
2017	8,546	1,432	46,627	1,417	1,414
2018	7,943	1,408	45,274	1,395	1,392

---

[4] This analysis could not be used to determine if there might have been an effect on individual test-takers.

Table 2

Results of Regression Analysis using Adjusted Historical U.S MBE

Performance Data to Predict MBE Performance

Within California[5]

1. Model Evaluation					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	4656.51	4656.51	239.81	<.0001
Error	13	252.42	19.42		
Corrected Total	14	4908.93			
2. Parameter Estimates					
Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	
Intercept	38.77	90.83	0.43	0.6765	
Adjusted U.S. Ave. MBE	0.99	0.06	15.49	<.0001	
R-square	0.95				
Mean Square Error (MSE)	4.41				

Publication of U.S.-wide results from the NCBE indicated that 45,334 applicants sat for the July 2019 MBE with an average score of 1,411. Their reports also indicated that 7,768 applicants took the test in California and achieved an average score of 1,427. Using the same adjustment methods described above, we estimated that 37,566 took the MBE outside of California and that the estimated Non-California mean MBE score was a 1,408.

Using the regression equation and the U.S. adjusted mean of 1,408, we would have estimated the California mean MBE to be 1,432. The difference between the actual California result (1,427) and the estimate result (1,432) was approximately 5 MBE points. Since the average error in the estimate was 4.41 points, a 95% confidence interval was +/- 8.6 points. The 95% confidence interval represents the range (+ or -) that the estimated result would be expected to fall 95 times out of 100. The 5-MBE point difference is well within the 95% confidence interval of the expected value.

Conclusion. The findings suggest that given the increase in MBE scores, the overall knowledge and skills of the applicants sitting for the July 2019 CBX have increased to approximately the

---

[5] Many of the statistics presented in Table 3 (e.g., Df, F-value) and other tables containing regression results are provided for the sake of completeness and directed to the technical reader who may be reviewing this report.

same levels seen in 2015. The increase is totally consistent with the historic pattern of relationship between California and the rest of the U.S. *We conclude that that any impact that the early topic release had on the population of test-takers was not statistically large enough to be distinguishable from chance differences alone*.

Issue 2. We next examined whether the early content release positively or negatively impacted the passage rates on the July 2019 CBX in some systematic manner.

Approach. To address this issue, we again looked to historic relationships, as well as to the procedures used to both score and make pass/fail decisions on the examination. We first examined the same 15 years (2004-2018) of July examination CBX MBE performance and eventual bar passage[6]. For each administration, the average MBE performance was calculated along with bar passage rate (i.e., the number passing the examination/total taking the examination). Similar to the first analysis, we then used regression methods, to compare the actual bar passage rate on the current examination with the expected rate to determine how much the two rates deviated, if any.

We next looked to the published protocols for scoring the CBX. These protocols dictate that the raw scores (i.e., the grades assigned by the trained readers) on the written section of the exam be summed together and "scaled" to the distribution of the MBE scores earned during the same examination. This procedure is used in all but a few states in the U.S. administering a bar examination and is intended to control or adjust for differences in the inherent difficulty of written questions from administration to administration, as well as for differences in the leniency or stringency of graders standards. If the early topic release led to systematically higher written raw scores, then the process of scaling would anchor them to the same distribution of MBE scores. *Thus, the resulting distribution of "scaled written" scores would subsequently be the same as if there was no artificial improvement due to an early release*.

Findings.

Table 3 contains the historic July CBX average MBE performance and bar passage rates for the July administrations of the CBX in years 2004 through 2018. The table indicates that bar passage rates during these years have ranged from a high of 62% to a low of 41% with corresponding high and low average MBE scores of 1,476 and 1,408. As expected, the correlation between the average MBE scores and CBX passage rates during that period was r=.98, which suggests that the passage rate tracks almost perfectly with performance on the MBE.[7]

---

[6] Historical data was available only applicants sitting for the MBE during those examinations and completing the written section are included in Table 3.
[7] Note that the weighting of the MBE in terms of its contribution to the total score increased to 50% in 2017 from 35% on previous exams.

This high degree of correlation suggested that the average MBE score could serve as a predictor of what the bar passage rate might be on future administrations. We developed on OLS regression equation based on the 2004 to 2018 results[8]. The results, presented in Table 4 indicate that the overall regression model was very accurate in predicting the passage rates with an average error of only 1.1%

Table 3

Average MBE Performance and

Bar Passage Rates on the July California Bar Examination

2004-2018

Year	Examinees	MBE Ave	Pass Rate
2004	8,020	1,434	49
2005	8,310	1,437	49
2006	8,858	1,452	52
2007	8,115	1,459	56
2008	8,590	1,476	62
2009	8,607	1,463	59
2010	8,521	1,454	55
2011	8,412	1,458	55
2012	8,664	1,460	56
2013	8,822	1,461	56
2014	8,428	1,436	49
2015	8,236	1,426	47
2016	7,648	1,423	44
2017	8,546	1,432	50
2018	7,943	1,408	41

---

[8] To account for the fact that weighting of the MBE changed over the period, we weighted the results from 2017 and 2018 by a factor of 2.

Table 4

Results of Regression Analysis using California Average MBE

Performance Data to Predict CBX Passage Rates

1. Model Evaluation					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	551.7116	551.7116	370.71	<.0001
Error	13	19.34719	1.4883		
Corrected Total	14	571.0588			
2. Parameter Estimates					
Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	
Intercept	-374.7849	22.1284	-16.94	<.0001	
Adjusted U.S. Ave. MBE	.2994	.0153	19.25	<.0001	
R-square	.97				
Mean Square Error (MSE)	1.22				

Utilizing the parameters from Table 4, we then estimated what the CBX passage rate on the July 2019 administration would be based on an average MBE score of 1,431[9]. The estimated passage rate was 47.7% with a 95% confidence interval of +/-2.4% (i.e., 45.3% to 50.1%). The actual bar passage rate for these applicants was 50%, a rate within expected range. The actual 2019 CBX mean MBE and bar passage rate is virtually identical to that observed in 2017.

Conclusion. We first conclude that procedural protocols requiring raw written scores be scaled to the distribution of MBE performance on the current administration would preclude *that any systematic effect the early topic release might have had on* <u>*raw written score*</u> *performance could (if any) not have translated into a subsequent and systematic statistically significant impact on the July 2019 CBX passage rate.* Secondly, the marked similarity of the actual and predicted bar passage rates suggest that any increase in the passage rate on the July 2019 examination relative to 2018, is due to the increase in the ability/skills level of the applicants as measured by their MBE scores, and is in line with expectations based on the MBE performance.

---

[9] The average MBE for applicants taking and completing the entire examination was used for this analysis to allow for an apples-to-apples comparison to the available historical data

Issue 3. The preceding analyses demonstrated that any systematic effects of the early release of the written section's subject-matter coverage did not appear to have any impact on the MBE performance or the overall bar passage rate on the examination. However, the question remained as to whether the early release may have had an impact on a specific group of applicants within the general population. A specific subgroup of possible concern was the examinees from the 16 law schools whose deans were made aware of the content of the written portion of the examination, earlier than all other test-takers.[10] Thus, analysis of these groups offer insight into whether there might have been some impact on the written scores themselves.

Approach A.

To address this question, a series of additional analyses were conducted; again, using both historical performance data and performance on the current administration.

First, examination performance (Pre-Regrade Written {WS} and MBE Scores), repeater status, and law school attended was extracted from the databases of the four most recent July administrations (i.e., 2015, 2016, 2017 and 2018) of the CBX[11]. Similar data was gathered from the current July 2019 administration. Applicants graduating from the 16 institutions whose dean was emailed the early notification were classified as "Early Release (ER)" while all other applicants were classified as "No Early Release (NER)".

Two sets of analyses were then conducted. The first set of analyses examined the relative examination performance (WS and MBE) for any differences that occurred between the groups before the breach and to evaluate the consistency/stability of those differences. We next analyzed salient applicant characteristics from both groups (including percentage of repeaters and type of school attended), as those variables have been known to be related to performance on the CBX. We reasoned that the greater the similarity of results in the two groups on the July 2019 CBX performance relative to historic findings, the stronger the evidence that the early disclosure had no differential effect on the students from the 16 law schools.

The second set of analyses focused on the July 2019 results only. In these analyses, MBE scores, repeater status, and type of school attended *of the NER applicants only,* was used to "predict" their WS (via OLS Multiple Regression). If the set of indicators were found to be highly correlated, then it was reasoned that the equation could then be subsequently applied to

---

[10] This concern was raised, and analyzed here, although the State Bar informs me that no evidence has been presented to date that students from these 16 schools actually received the subject matter coverage any earlier than all other test takers.

[11] As of the July 2017 administration, the CBX was changed from a 3-day administration to a 2-Day administration. The number of essay questions was reduced from 6 to 5, and the number of performance test questions was reduced to 1 from two. As a result, the possible range of raw WS was reduced from 0-1000 to 0-700.

the applicants in the ER group to predict their WS as well. If (1) the relationships between the WS and MBE was not disrupted or (2) there were little to difference be found between the actual WS and the predicted WS, then it would lend evidence to the conclusion that early release did little to systematically lower or raise the WS performance of the ER group.

Where appropriate, statistical tests of significance were applied to the analysis with α-levels set at .05. Results from the tests where p < .05 were considered statistically meaningful and not occurring by chance.

Findings A.

Table 5 presents statistical data on the performance measures and classification variables for both the ER and NER groups for each of the four historical examinations. Results from Table 5 indicate that applicants from the 16 ER law schools consistently made up about 13% of the overall test-taking population and consistently performed lower on both the MBE (63 scale points) and Written tests (15 raw points). These applicants, as a group, were also disproportionately repeaters (51% vs 32% on average) and greater proportions came from California Accredited and Unaccredited schools (37% and 13% vs 6% and 2% from the NER group). We also noted that the relationship (i.e., correlations) between the MBE and WS tended to be consistent across the exams within each group, though slightly higher in the NER group, due most likely to the fact that applicants in the NER groups were more homogenous (see Sd's in each group).[12]

Then, to more clearly illustrate the consistency of the difference between the groups on each of the performance measures, we calculated the standardized difference between the applicants in the NER and ER groups. This calculation simply involved subtracting the means in each group, and then standardizing that difference by dividing by Sd. of the scores in the NER group. Calculations were conducted for both the WS and MBE for each year. The results are presented in Table 6.

The statistics in Table 6 illustrate several findings. First, they show that the relative size of the difference in performance between the ER and NER groups are quite similar for both the MBE and WS sections and that pattern seems to hold for each of the examinations. Across the four administrations, the average difference is only .03 standard units (.32 vs .29). We did see slightly larger differences when the examination changed to a 2-Day format, though these differences are still quite small.

---

[12] A separate analysis compared the relationship between <u>all</u> test taker's MBE and WS scores in 2018 (r=.709) and 2019 (r=.712). No statistical difference was found (p=.379).

The implication of these analyses is that if any additional study time potentially afforded to students from the ER schools did have an impact, we would expect to see an increase in the standardized difference of the July 2019 WS, relative to the difference in the July 2019 MBE.   If not, we should see results within the ranges observed in Table 6.

Table 5

Descriptive Statistics for July GBX Performance Measures and Classification Variables

By "Early Release" vs. "Non-Early Release" Applicants

	3-Day Format				2-Day Format				4 Examination	
	2015		2016		2017		2018		Average	
	NER	ER	NER	ER	NER	ER	NER	ER	NER*	ER*
Metric/ Indicator	(N=7,219)	(N=1,017)	(N=6,665)	(N=983)	(N=7,471)	(N=1,075)	(N=6,888)	(N=1,055)	(N=6,888)	(N=1,055)
	88%	12%	87%	13%	87%	13%	87%	13%	87%	13%
MBE										
Mean	1,433	1,383	1,431	1,372	1,439	1,383	1,415	1,365	1,429	1,376
Sd.	159	145	168	148	169	146	165	143	165	145
% Repeaters	27%	46%	31%	49%	34%	55%	34%	53%	32%	51%
Type of School										
ABA	76%	50%	75%	51%	71%	46%	71%	50%	73%	49%
CA Accreditted	6%	37%	6%	37%	6%	40%	6%	35%	6%	37%
Non-Accredited	2%	12%	2%	12%	3%	14%	2%	14%	2%	13%
Written Score										
Mean	614	598	599	582	431	420	421	412	_607_	_592_
Sd.	52	46	48	40	39	31	38	32		
Correlation										
Written w. MBE	.71	.63	.73	.61	.73	.57	.72	.61	.72	.61

* Average Written Score based on pro-rated 2017 and 2018 8-question tests.

Table 6

Standardized Differences Between "Early Release" vs. "Non-Early Release" Applicants

On the MBE and Written Section of the GBX

July 2015 through July 2018

Metric	3-Day Format		2-Day Format		4 Exam
	2015	2016	2016	2016	Average
MBE	0.31	0.35	0.33	0.30	0.32
Written	0.30	0.36	0.28	0.23	0.29

With these historical similarities and differences as a reference, we replicated the analyses using the results from the current July 2019 CBX (providing comparisons to the recent two July administrations). Table 7 provides the descriptive statistics, while Table 8 presents the standardized differences.

The statistics in Table 7 point to striking similarities in (a) the 2019 performance and characteristics of the examinees in the respective two groups relative to the two previous examinations, as well as (b) the magnitude of the differences in performance between them. For example, we see that in the NER group, both the percentage of repeaters and proportion attending each of the types of school are identical, while in the ER group, the proportions vary only slightly. Reflecting the overall change in MBE performance, we see applicants in both groups have increased slightly. Perhaps more importantly, we see that the mean raw WS have changed by only 1 and 2 points in the NER and ER groups, respectively, with comparable score spreads (i.e., Sd.). Further, the relationship of the WS performance to MBE performance has not been significantly impacted in either group, relative to historic values.

The consistency of the size of the differences is illustrated in Table 8. We note that the size of the difference between ER and NER on the MBE is identical to the average of the preceding two examinations (.32), while difference on the WS is only .02 Sd units. *This finding lends strong evidence to the issue that even if examinees from the 16 law schools received the topics earlier than examinees from other schools (of which there is currently no evidence that they did), they received no statistically significant added benefit.*

Table 7

Descriptive Statistics for July GBX Performance Measures and Classification Variables

By "Early Release" vs. "Non-Early Release" Applicants

Historic 2-Day Exams Compared to July 2019

	2017		2018		Average 2017-2018		2019	
	NER	ER	NER	ER	NER*	ER*	NER	ER
Metric/Indicator	(N=7,471)	(N=1,075)	(N=6,888)	(N=1,055)	(N=7,180)	(N=1,065)	(N=6,625)	(N=1,055)
	87%	13%	87%	13%	87%	13%	86%	14%
MBE								
Mean	1,439	1,383	1,415	1,365	1,427	1,374	1,437	1,389
Sd.	169	146	165	143	167	144	167	155
% Repeaters	34%	55%	34%	53%	34%	54%	34%	50%
Type of School								
ABA	71%	46%	71%	50%	71%	48%	71%	53%
CA Accredited	6%	40%	6%	35%	6%	38%	6%	34%
Non-Accredited	3%	14%	2%	14%	3%	14%	2%	12%
Written Score								
Mean	431	420	421	412	426	416	427	418
Sd.	39	31	38	32	39	31	37	32
Correlation								
Written w. MBE	.73	.57	.72	.61	.73	.59	.72	.64

Table 8

Standardized Differences Between "Early Release" vs. "Non-Early Release" Applicants

On the MBE and Written Section of the GBX

Historic 2-Day Exams Compared to July 2019

Metric	2017	2018	Average 2017-2018	2019
MBE	0.33	0.30	0.32	0.32
Written	0.28	0.23	0.26	0.24

Approach B. Regression Analysis. Another set of more refined analyses were conducted to determine whether the WS of the ER schools were positively impacted. In this approach, a series of multiple OLS regressions were conducted using performance on the MBE and the available applicant characteristics to predict raw WS. However, these models were only conducted on the NER examinees. The purpose of these models was to see how accurately these factors could predict raw WS, and if both successful and consistent, then they could be used to predict what should have occurred in the ER examinees, all things held equal. If there was no difference between the historical actual and predicted WS, (in situations where no aberrant occurrences were reported before or after the examinations) for students from the 16 law schools, then the approach could be similarly applied to the July 2019 examination with a set of similarly expected outcomes.   A separate regression was run for each of the exams preceding the July 2019 administration.   Table 9 reports on the results of the models[13]

Table 9

Multiple OLS Regression Models

Predicting Written Scores

On Applicants from "Non-Early Release Schools"

July 2015 through July 2018

	3-Day Format		**2-Day Format**	
Predictor	**2015**	**2016**	**2017**	**2018**
Intercept	319.36	328.64	209.62	204.83
Scale MBE	0.20	0.19	0.15	0.15
Repeater	-19.23	-15.97	-12.01	-10.07
School Type				
ABA	16.43	12.69	16.12	15.13
Accredited	4.24	1.35	11.59	8.98
Unaccredited	-6.49	-12.90	1.01	2.30
Equation R²	0.54	0.58	0.58	0.56
Multiple R	0.73	0.76	0.76	0.75

---

[13] Recall that the models were estimated based solely on applicants from the NER applicants

Each of the models in the respective years were statistically accurate and quite consistent in predicting WS performance. The overall model in each year was statistically significant (p <.0001) and accounted for between 54% and 58% of the variability in WS scores (correlations ranging from .73 to .76 values that are quite high values by social science and testing standards). The "predictors" in Table 9 represent the weights that would be applied to the respective variable in estimating the WS for an applicant. All are statistically significant (p< .001). They differ in absolute size because they are on different scales of measurement (MBE on a 0 to 2000 scale; the others on a 0 or 1 scale). Differences in the size of the parameters between the 2-Day exams and 3-Day exams are also due to the fact that WS is on a different scale. The fact that the quality of the equation results (R^2 and R) remained consistent for both examination formats (2 and 3-day), add further evidence to the stability of these findings.

Based on the success of these models, a similar equation was developed for the July 2019 CBX using the same set of variables. The resulting equation yielded similar results, with an overall R^2 value of .54 and equation weights that not statistically different than the historic exams.

Given the relative strength of each of the models, we first applied them to the historic data of applicants from the ER schools to get a sense of what might have been expected under the condition of "no effect" from the early release. Table 10 shows the average actual and "expected" WS performance of the applicants from both the ER and NER schools for each examination, along with the differences and the results of a paired (correlated) t-test evaluating whether the differences are statistically different from 0[14].

The results in Table 10 conclusively showed that the models derived from the NER groups were very effective in predicting the WS performance in the ER groups for each of the historic exams. Average actual vs. predicted performance differed by more than 1 WS point on only the 2015 examination and in no case were the differences found to be statistically significant than zero (as evidence by the fact that none of the probability -p- values were less than .05. Thus, our baseline historical findings consistently showed that after accounting for the MBE performance, and the relative proportion of repeaters and school attendance, we would expect models for the July 2019 examination to yield similar findings in the absence of any impact to the written scores.

---

[14] A paired t-test is used to evaluate whether two sets of scores for the same individuals are significantly different, statistically speaking

Table 10

Average Actual vs. Predicted Written Scores

In "Early Release" vs. "Non-Early Release" Applicants

Based Upon Models Calculated in the Non-Early Release Groups

July 2015 through July 2018

Metric	3-Day Format				2-Day Format			
	2015		**2016**		**2017**		**2018**	
	NER (N=7,219)	**ER** (N=1,017)	**NER** (N=6,665)	**ER** (N=983)	**NER** (N=7,471)	**ER** (N=1,075)	**NER** (N=6,888)	**ER** (N=1,055)
Actual	613.93	598.42	598.65	581.50	430.81	419.82	420.78	411.81
Predicted	613.93	596.82	598.65	581.04	430.81	420.06	420.78	411.32
Actual - Predicted		*1.60*		*0.46*		*-0.24*		*0.49*
Significance		p > .136		p > .638		p > .749		p > .521

Anchored by these results, we then applied the model for the July 2019 NER examinees to the examinees from the July 2019 ER schools. The summary of the findings is presented in Table 11.

Table 11

Average Actual vs. Predicted Written Scores

In "Early Release" vs. "Non-Early Release" Applicants

Based Upon Models Calculated in the Non-Early Release Groups

July 2019

Written Scores	2019	
	NER (N=6,888)	**ER** (N=1,055)
Actual	427.20	417.65
Predicted	427.20	418.37
Actual - Predicted		*-0.72*
Significance		p > .330

We can see from Table 11 that the actual average written scores for examinees from the ER law schools was 417.65. The expected, or estimated average, based on the regression model was 418.37, only a .72 difference (on a scale of 0 to 700). The difference is statistically indistinguishable from zero, indicating that the results are entirely consistent with historical findings.

While there was no apparent differential impact on the set of students from the ER law schools as a whole, the question remains as to whether there were individual schools within that group that experienced a boost in their WS on the July 2019 CBX. A separate analysis was then conducted to compare the actual average WS performance to the predicted performance *for each school individually*, where the expected performance was based on the same equation as used in Table 11. Paired t-tests were used to compare the actual and expected means at each school. The individual school findings are presented in Table 12.

Table 12

Actual vs. Predicted Average Written Scores

In Individual "Early Release" Law Schools

Based Upon Models Calculated in the Non-Early Release Groups

July 2019

Law School	Actual Written	Predicted Written	Difference	P-Value
1	398.2	398.8	-0.7	0.670
2	408.1	417.5	-9.5	0.252
3	386.5	390.7	-4.2	0.638
4	413.5	401.9	11.6	0.086
5	406.6	401.3	5.3	0.132
6	426.1	427.5	-1.4	0.558
7	413.4	419.7	-6.3	0.013
8	*	*	*	0.223
9	401.4	398.2	3.1	0.201
10	410.7	403.2	7.5	0.246
11	389.5	392.6	-3.1	0.279
12	389.6	391.5	-1.8	0.717
13	408.5	403.8	4.7	0.074
14	437.3	438.1	-0.7	0.587
15	398.6	400.0	-1.4	0.685
16	415.0	419.4	-4.4	0.321

The schools in Table 12 are presented in random order (The school with an "*" had less than 11 examinees and State Bar policy precludes presenting score data on a cohort of this size to ensure that individual student is not identifiable). An inspection of the actual average written scores range from as low as 386.5 to 437.3, a difference of over 50 points and well over a full Standard Deviation. Despite this broad range, in no law school was the difference between the actual and predicted written score more than 11.6 points, and in none of the comparisons were the differences found to be statistically different from zero. In the one school in which the difference approached statistical significance (#7), the actual mean was actually lower than predicted.

Conclusions. The findings from these final set of analyses conclusively show that any additional study time that might have been afforded to the examinees from the schools in which the deans were sent the topics for the written section of the examination, had no positive, statistically significant impact on the written scores of their students. *The lack of impact was established for the group of 16 schools as a whole as well as for the students in each of the individual law schools.*