



OFFICE OF THE ACTUARY

DATE: August 21, 2026

FROM: Paul Spitalnic, ASA, MAAA
Chief Actuary, Centers for Medicare & Medicaid Services

SUBJECT: Certification of Comprehensive Care for Joint Replacement Expansion

Certification

Section 1115A of the Social Security Act established the Center for Medicare and Medicaid Innovation (CMMI) within the Centers for Medicare & Medicaid Services (CMS) to test innovative payment techniques and service delivery models. For successful models, the law states that “the Secretary may, through rulemaking, expand (including implementation on a nationwide basis) the duration and the scope of a model that is being tested...to the extent determined appropriate by the Secretary, if—

- (1) The Secretary determines that such expansion is expected to—
 - (A) reduce spending under the applicable title without reducing the quality of care; or
 - (B) improve the quality of patient care without increasing spending;
- (2) The Chief Actuary of the Centers for Medicare & Medicaid Services certifies that such expansion would reduce (or would not result in any increase in) net program spending under the applicable titles; and
- (3) The Secretary determines that such expansion would not deny or limit the coverage or provision of benefits under the applicable title for applicable individuals.”

We were asked to consider an expansion of the Comprehensive Care for Joint Replacement (CJR) model, a bundled payment model for lower extremity joint replacement (LEJR) episodes of care that occur in the hospital setting. The expanded model would include some specification changes from the CJR model, and would be called CJR-X.

CJR-X will target reductions in LEJR spending for the anchor stay or procedure and the 90 days post discharge. Benchmark spending for each episode of care will be calculated using a 3-year rolling baseline period of average spending, broken out by geography and Diagnosis-Related Group (DRG). The benchmark will be trended forward to the performance period using a prospective trend factor, which will be based on historical changes in average episode spending during the baseline period. Once the actual trends are known, a retrospective trend factor, capped at +/- 3 percent of the projected trend, will be applied to determine the final benchmark spending amount.

Target prices will be calculated by risk adjusting the benchmark spending using performance year adjustments to minimize the risk of patient mix differences between the historical period and performance period. After risk adjustment, the benchmark spending for each hospital will be adjusted by a discount, which will depend on the hospital's quality of care measure. The discounts will be 0 percent for hospitals with "excellent" quality of care, 1 percent for hospitals with "good" quality of care, and 2 percent for all other hospitals. After each performance year, reconciliation payments or repayments for each hospital will be determined based on the cumulative difference between target prices and actual episode spending during that year, subject to stop gain/loss limits and episode exclusions for episodes subject to extreme and uncontrollable circumstances. Stop-gain/loss limits will be set at 20 percent of the cumulative target price, except rural hospitals and hospitals with a high percentage of duals will receive a stop-loss limit of 5 percent. In addition, hospitals that are deemed to deliver quality care that is below acceptable will not receive reconciliation payments.

Based on historical evidence from evaluations of CJR and other bundled episode payment models, and analysis of recent episode spending data, **I certify that CJR-X would not result in an increase in program spending.** If the actual parameters of CJR-X differ from those specified above, or if there is additional evidence available that would cause us to update our methods or assumptions, this certification would not apply and the program would need to be reevaluated. The remainder of this memorandum summarizes the analysis supporting this certification.

Office of the Actuary Analysis

For our analysis, we first estimated the savings of the target pricing methodology for each of the policy specifications and combined the results. Then, we assessed the potential for the intervention to change provider behavior. The estimated behavioral impacts were combined with the savings from the target pricing methodology to determine the overall impact on program spending.

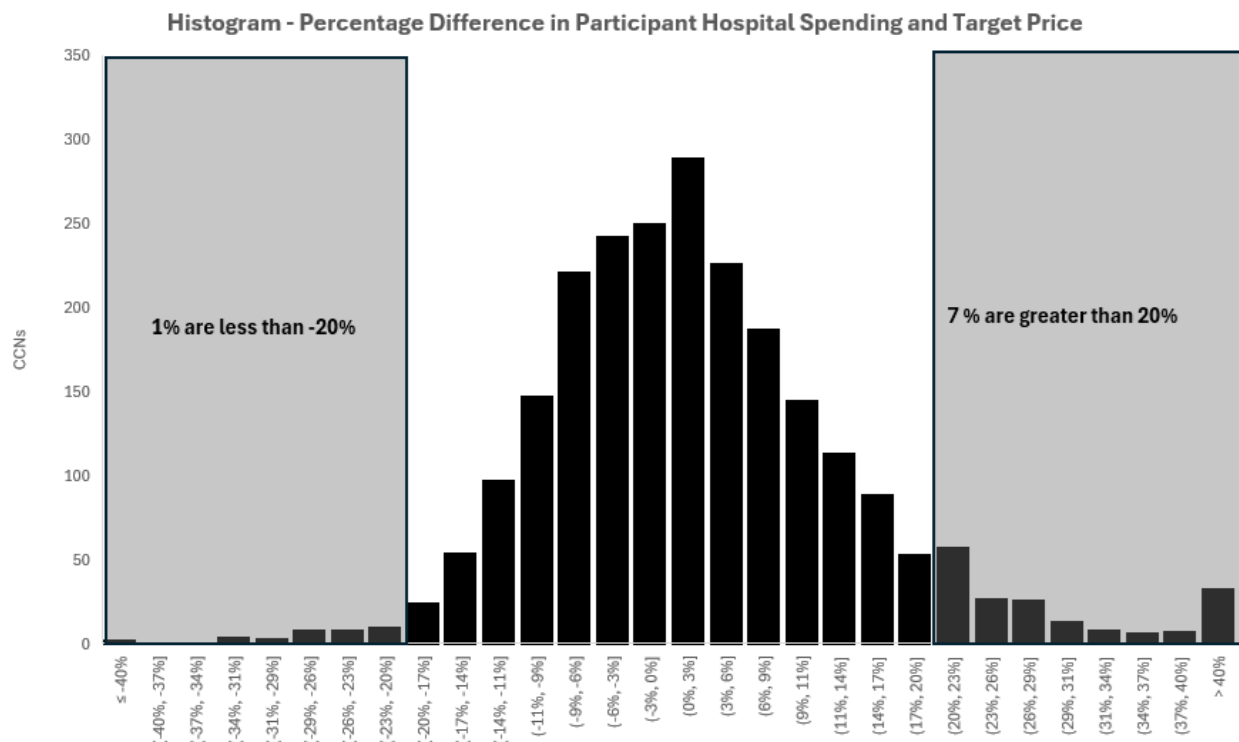
1. Financial Estimate of Target Pricing Methodology

Using episode spending data for episodes beginning in 2024, we performed a reconciliation calculation, using 2024 as the performance year and the baseline period. Using the same year for the baseline and performance period provided us the ability to value policy specifications individually and cumulatively, ignoring possible behavioral effects and trend error impacts.

The discount varies based on the quality adjustment. The discount is 2 percent of total episode spending for hospitals in the quality categories of "acceptable" and "below acceptable". However, the quality adjustment reduces the discount to 1 percent for hospitals rated as "good" and 0 percent for hospitals rated as "excellent", resulting in an average discount of about 1.3 percent. The following distribution of quality categories was provided by CMMI and the certification includes the assumption that this distribution will be constant in the future:

Quality Category	Percentage of hospitals	Discount
Below Acceptable	20%	2%
Acceptable	30%	2%
Good	30%	1%
Excellent	20%	0%

Applying stop-gain/loss limits of 20 percent to all hospitals reduces the average savings from the target pricing methodology by 0.4 percentage points, from 1.3 percent to 0.9 percent. This is due to skewed episode spending, as episode spending is more likely to exceed the target by a large margin rather than be less than the target by a large margin. The chart below shows the distribution of the percentage difference between hospital episode spending and their target price. Episode spending from approximately 7 percent of hospitals exceeds its target by 20 percent or more, while only 1 percent of hospitals have spending below their targets by more than 20 percent.



Applying a 5-percent stop-loss limit to high-dual and rural hospitals while maintaining the stop-gain at 20 percent reduces the projected savings from 0.9 percent to 0.6 percent (a 0.3 percentage point reduction).

Nonpayment of reconciliation to hospitals in the “below acceptable” quality category increases the estimated savings by 0.1 percentage point. As a result, we estimate that with no participant behavior, CJR-X would result in savings equal to 0.7 percent of LEJR episode spending.

As indicated earlier, using 2024 as both the baseline and the performance period ignores the potential for trend forecast errors that exceed the three percent retrospective trend factor cap. OACT estimated the distribution of financial outcomes due to trend errors using the Black-Scholes formula and the historical autocorrelation of episode spending growth. This simulation projects an additional .05 to .1 percent in reconciliation payments as a share of episode spending from the variation. Thus, with no behavior, OACT assumes that the target pricing methodology alone produces savings of 0.6 percent.

2. Behavioral impacts of the model

The impact of the pricing methodology detailed above does not consider the potential effects of provider behavior on episode costs. If the model does not result in any provider practice pattern changes related to an LEJR episode, then we estimate that providers would owe the government, on average, 0.6 percent of LEJR episode spending, as detailed in the Financial Estimate of Target Pricing Methodology section of this memorandum. The estimated amounts owed would be the difference between the actual costs and the discounted targets. However, practice pattern changes could affect the savings. If practice pattern changes result in a reduction of costs, the model-related savings would be offset by reductions in amounts owed by providers. However, the lower costs from the changes in practice pattern would also result in lower benchmarks, creating additional savings to the Medicare program. If instead practice pattern changes result in increased costs, the opposite would be true. To conclude that the model would not result in increased costs, we must determine that the model is not expected to result in increased average costs for the procedures included in CJR-X.

a. Evaluation results for the CJR model

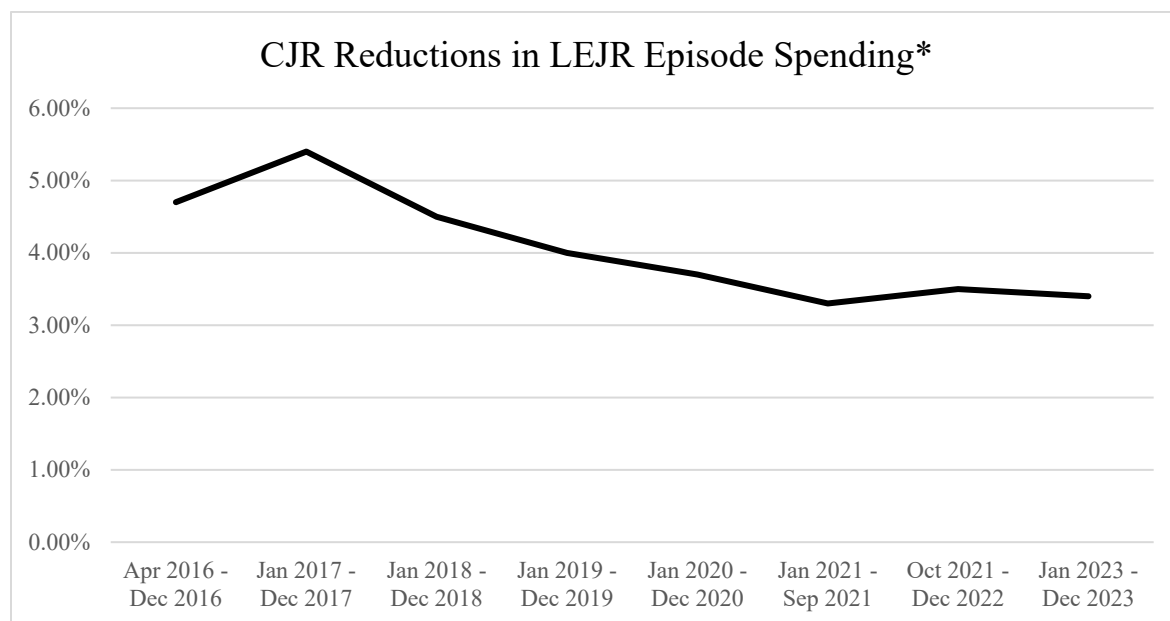
To assess the potential effect on actual costs for CJR-X, we began by examining the evaluation results for CJR. In 2016, CJR began as a fully mandatory bundled payment model. CMS stratified MSAs into eight strata based on historical episode payments and population size. Within each stratum, MSAs were randomly selected to participate in CJR. In total, 67 MSAs were chosen to participate. However, starting in 2018, only hospitals belonging to the 34 MSAs with the greatest historical episode spending were required to continue in the model.

The evaluation employed a difference-in-differences approach to estimate model savings. This approach compared differences in CJR participant spending during performance years to spending during the base period (2012 to 2014) with the differences in spending from comparison hospitals. Considering only the mandatory participating hospitals, the model has consistently shown indications of overall reductions in total episode spending. In the latest evaluation report¹, comparing spending from episodes ending in October 2021 through December 2023 with spending from the base period, CJR was estimated to reduce episode spending in Original Medicare (formerly described as fee-for-service Medicare) by 3.4 percent (before considering reconciliation) with a p-value of 0.12. Much of these savings can be

¹ <https://www.cms.gov/priorities/innovation/data-and-reports/2025/cjr-py7-annual-report>, page 23

attributed to reductions in post-acute care spending, as the evaluation indicated that CJR hospitals decreased the proportion of patients first discharged to an Inpatient Rehabilitation Facility (IRF). These savings were partially offset by increases in home health spending.

While evaluation reports for CJR show savings, reductions in episode spending have leveled off in recent years. The following graph shows estimated reductions in episode spending for CJR hospitals over time:



*Data comes from multiple sources^{2,3,4}

The greatest estimated reductions in episode spending occurred in the early years of the model, with a greater than 5 percent reduction in spending in 2017. Since 2019, average estimated reductions in episode spending have been between 3.3 and 4.0 percent, maintaining most, but not all, of the savings achieved in earlier years. Based on these results, it may be challenging for CJR participants to generate additional reductions in episode spending going forward, but the achieved savings are likely to persist.

Reductions in episode costs in the initial years of CJR have resulted in lower benchmarks in more recent years. Since the savings have leveled off, participants have owed payments to CMS, on average, which has resulted in continued federal savings.

Due to the methods used for selecting the mandatory participants, the hospitals participating in CJR were not representative of a national sample. For the original 67 MSAs chosen to participate

² Apr 2016 – Sep 2021: <https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report>, page 16

³ Oct 2021 – Dec 2022: <https://www.cms.gov/priorities/innovation/data-and-reports/2024/cjr-py6-annual-report>, page 18

⁴ Jan 2023 – Dec 2023: <https://www.cms.gov/priorities/innovation/data-and-reports/2025/cjr-py7-annual-report>, page 135

in CJR, the model oversampled MSAs with greater episode spending and larger population size. The following table provides details regarding CMS’ original stratified random sample of MSAs included in the latest evaluation report⁵:

MSA Population	MSA sampling stratum	MSA average episode payment	# MSAs eligible for sampling	# CJR MSAs	Percent of MSAs selected for CJR (%)
Less than median population	1	Lowest quartile	25	8	32.0
	2	2 nd lowest quartile	18	6	33.3
	3	3rd lowest quartile	19	8	42.1
	4	Highest quartile	22	11	50.0
More than median population	5	Lowest quartile	15	5	33.3
	6	2 nd lowest quartile	28	10	35.7
	7	3rd lowest quartile	22	9	40.9
	8	Highest quartile	22	10	45.5
Total				67	

From the table above, 38 out of the 67 MSAs are MSAs with spending levels in higher quartiles (strata 3,4,7 and 8). In 2018, the 34 MSAs with the greatest historical spending were selected to remain mandatory participants in the model. Baseline (during years 2012 to 2014) per episode spending for the 34 mandatory MSAs was 16 percent greater than baseline spending for other hospitals (\$32,516 compared to \$28,005)⁶. This bias in MSA selection puts into question whether non-CJR hospitals would have similar behavioral outcomes in a model expansion.

b. OACT Simulation of CJR Participant Selection Bias

Due to the oversampling of high cost MSA in the selection procedures for mandatory CJR participants, OACT simulated an intervention effect for a nationally representative set of eligible hospitals. OACT utilized nearest neighbor matching and a difference-in-difference analysis between CJR participant and non-participant hospitals residing within MSAs initially eligible for CJR. Overall hospital and LEJR-specific billing patterns were estimated and matched between both cohorts of hospitals from a pool of 2012-2014 baseline experience. Geographic payment standardization and LEJR-specific risk adjustment were estimated and applied to all experience. Lastly, hospital-specific study weights were estimated to form a nationally representative comparison of treatment period experience during 2023-2024.

The estimated impact on episode spending from the evaluation results for the CJR model is a 3.4 percent savings with a p-value of 0.12. The estimated reduction in episode spending from the

⁵ <https://www.cms.gov/priorities/innovation/data-and-reports/2025/cjr-py7-annual-report>, page 82

⁶ <https://www.cms.gov/priorities/innovation/data-and-reports/2020/cjr-thirdannrpt>, page 26

nationally-representative sample of CJR participants is 1.6 percent with a p-value of 0.47. The large p-value is due to the small sample of participants with lower-than-average spending.

c. Review of Similar Bundled Payment Models

To assess the potential risk of CJR-X resulting in increases in episode spending, OACT reviewed evaluations of other CMMI bundled payment models, namely Models 2 and 3 from the Bundled Payments for Care Improvement (BPCI) initiative and BPCI Advanced.

The BPCI initiative included four models. Participation in these models was voluntary. Models 1 and 4 defined an episode of care as an inpatient stay in an acute hospital, neither model bundling post-acute care spending. Since post-acute care was not bundled in these models, OACT focused more on the other BPCI models.

BPCI Model 2 bundled spending for the acute care stay and various post-acute care periods. Participants could select up to 48 episodes of care to be bundled, including LEJR episodes. The seventh evaluation report of BPCI Model 2 estimated reductions in episode spending for LEJR episodes of 4.5 percent for hospital-initiated episodes.⁷ Furthermore, estimated impacts in the 95 percent confidence interval all indicated savings. Episodes initiated by physician group practices (PGPs) experienced even larger reductions in spending.⁸ However, due to the model design (for example, waiving downside risk during reconciliation), reconciliation payments exceeded those savings and BPCI Model 2 was estimated to result in a net cost to Medicare.

BPCI Model 3 bundled only post-acute care spending for periods ending 30, 60 or 90 days after discharge, and also had a selection of 48 different episodes. The seventh evaluation report of BPCI Model 3 estimated LEJR episode spending reductions for post-acute care of 7.0 percent.⁹ Similar to Model 2, estimated impacts within the 95 percent confidence interval indicated savings for this piece of the model. Also similar to Model 2, reconciliation payments were greater than the estimated reductions in episode spending, resulting in an overall model cost.

BPCI Advanced was a voluntary bundled payment model for hospitals or physician group practices (PGPs). Episodes began with a medical or surgical hospital stay or an outpatient procedure initiated by a participating hospital or PGP, and ended 90 days after the patient was discharged from the hospital. BPCI Advanced covered episodes from 8 clinical episode lines, including orthopedic episodes. The latest evaluation report for BPCI Advanced showed 5.7 percent savings to orthopedic episodes when comparing episode spending in 2022 with spending in the base period (2015 to 2018), with estimated impacts within the 95 percent confidence interval indicating savings for this part of the model.¹⁰ Prospective target pricing in BPCI Advanced was used in early model years, resulting in large payouts to participants. These payouts exceeded the savings from reductions in episode payments. After retrospective target pricing was implemented in 2021, participation in the model swiftly declined. About a third of

⁷ <https://www.cms.gov/priorities/innovation/data-and-reports/2021/bpci-models2-4-yr7evalrpt-app>, page E-18

⁸ <https://www.cms.gov/priorities/innovation/data-and-reports/2021/bpci-models2-4-yr7evalrpt-app>, page F-11

⁹ <https://www.cms.gov/priorities/innovation/data-and-reports/2021/bpci-models2-4-yr7evalrpt-app>, page H-3

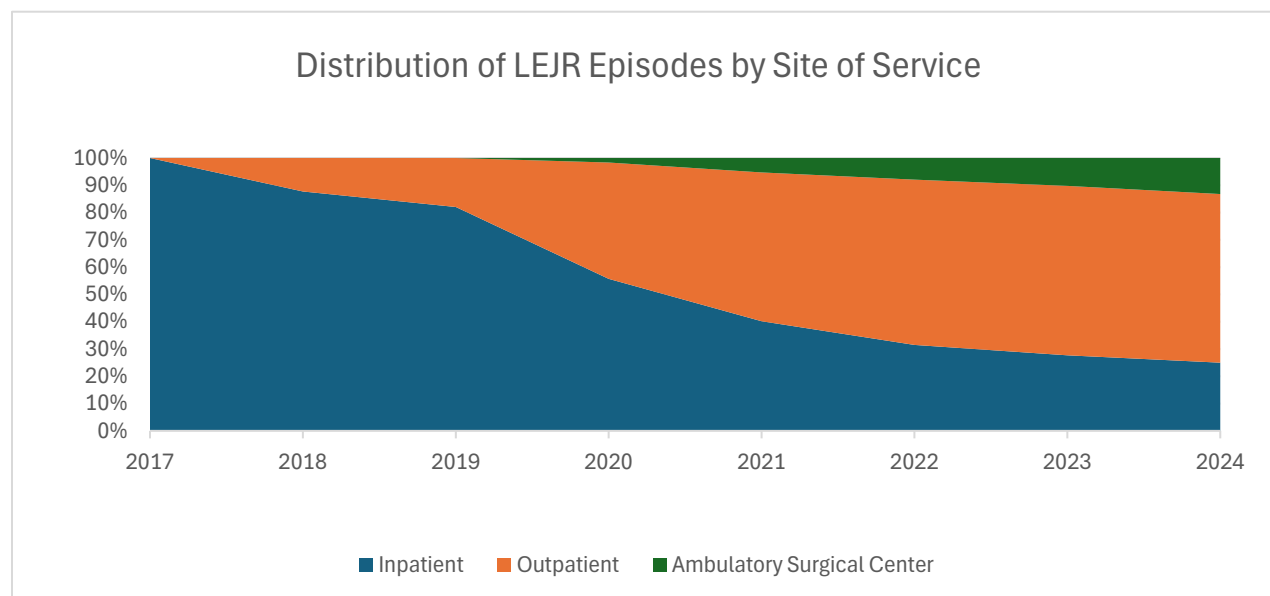
¹⁰ <https://www.cms.gov/priorities/innovation/data-and-reports/2025/bpci-adv-ar6-appendices>, page H-6

participants dropped out of the model the next year. By 2025, less than 10 percent of participants remained in the model.

Other bundled payment models showed reductions in Original Medicare episode spending by participants, however due to target pricing methodology issues in the models or the voluntary nature of the models, these bundled payment models did not result in long-term net savings to Medicare. CJR-X is mandatory and uses a target pricing methodology that is similar to the methodology used in more recent CJR years. That target pricing methodology has produced lasting net savings regardless of the magnitude of the Original Medicare episode spending reductions that occur.

d. Other Considerations

When CJR began in 2016, LEJR procedures were only delivered in the in-patient setting. In 2018, total knee replacements were removed from Medicare’s inpatient only list, and total hip replacements were removed from that list in 2020. Total knee and hip replacements are now often performed in ambulatory surgical centers (ASCs). The graph below shows how the utilization of different sites of service for these procedures has evolved over time:



In 2024, 86 percent of knee procedures were either outpatient procedures or performed in an ASC. In addition, 61 percent hip procedures were also performed outside of the inpatient setting in 2024.

As care delivery of LEJRs has become more efficient over time, episode spending has also declined. As more procedures are performed on an outpatient basis or at an ASC, fewer patients are eligible for SNF or IRF. Both the treatment and comparison groups in CJR experienced significant cumulative reductions in average episode spending during CJR (10.8 percent in the

treatment group versus 8.3 percent in the control group).¹¹ By 2023, average episode spending for the treatment group was in line with the comparison group. It is possible that much of the potential savings resulting from reductions in episode spending have already occurred and further reductions could be far less likely.

e. Overall assessment of the impact on actual spending

Evidence from the evaluation of CJR, combined with the evaluations of other bundled payment models, indicates that bundled payment initiatives on LEJR episodes have consistently shown reductions in the use of post-acute services. Earlier versions of the model were voluntary and were shown to produce greater reconciliation payments, likely resulting in increases in net program costs. The mandatory CJR model has maintained the gross cost reductions from the early model years, and the changes in the payment methodology have led to net Medicare savings. While savings from reductions in Original Medicare spending were achieved under the model, they have not generated additional savings in recent years as LEJR episodes both within and outside of the model have become more efficient. However, there is a low likelihood that CJR-X would result in increases in episode spending. Overall, OACT estimates that future CJR-X behavioral impacts are expected to be relatively small in magnitude and are more likely to produce additional savings.

Conclusion

The discounts applied to LEJR episodes under the CJR-X model are expected to reduce program costs by 1.3 percent. After accounting for the effects of the stop-gain/loss limits, the nonpayment of bonuses to “below acceptable” providers, and the impact of the cap on the retrospective trend adjustments, the impact on program costs is a savings of 0.6 percent. This impact is before accounting for behavioral effects from CJR-X participants

The evaluation of the CJR model as well as other similar models indicate that bundle payments for LEJR episodes have consistently shown reductions in actual costs. However, recent experience indicates that care efficiency improvements and reductions in spending have leveled off and suggests that the future potential for savings driven by participant behavior is minimal. The analyses also indicates that an increase in episode spending is unlikely. As a result, CJR-X is expected to lead to an overall reduction in Medicare spending.

¹¹ <https://www.cms.gov/priorities/innovation/data-and-reports/2025/cjr-py7-annual-report>, page 132