November 26, 2021

Ellen Montz
Director
Center for Consumer Information and Insurance Oversight
Centers for Medicare and Medicaid Services
Department of Health and Human Services
200 Independence Avenue SW
Washington, DC 20201

Re: HHS-Operated Risk Adjustment Technical Paper on Possible Model Changes

Dear Director Montz:

Thank you for the opportunity to comment on the "HHS-Operated Risk Adjustment Technical Paper on Possible Model Changes" that was published on October 26, 2021 by the Center for Consumer Information and Insurance Oversight (CCIIO). We appreciate CCIIO's decision to publish additional information about the risk adjustment changes it is currently considering, as well as the opportunity to comment on those potential changes.[1]

The first section of this letter comments on the changes to methods for calculating risk scores that are described in the main text of the technical paper. We make three main points:

- **The transfers generated by the current risk adjustment methodology are most likely too small, not too large:** Because health status is only imperfectly observed in risk adjustment, fully mitigating selection incentives is likely to require transferring more from insurers with low-risk enrollees to insurers with high-risk enrollees than is suggested by directly observable differences in risk. But it is doubtful that the existing mixture of under- and over-prediction documented in the technical paper is, on net, producing transfers that meet this standard. Indeed, the large role that narrow network, tightly managed plans currently play in the individual market is consistent with the view that risk adjustment transfers are too small and insurers retain strong incentives to pursue low-risk enrollees.

- **In this environment, changes that increase risk scores for low-risk enrollees are likely to worsen market outcomes by reducing plan quality and, potentially, competition:** All else equal, steps that increase the risk scores of low-risk enrollees (or reduce the risk scores of high-risk enrollees) will tend to reduce risk adjustment transfers. Since we believe that transfers are already too small under the existing risk adjustment methodology, we expect that such changes will tend to increase selection incentives, thereby reducing plan quality and, potentially, making it harder to sustain robust plan competition.

  Raising risk scores for low-risk enrollees is also unlikely to have the countervailing benefit of increasing enrollment, as posited by the technical paper. Since most enrollees receive the premium tax credit and the tax credit's value depends on the premiums of the lowest-cost silver plans, changes that reduce the premiums of existing low-cost plans or spur

[1] Please note that the views expressed in this letter are our own and do not necessarily reflect the views of the Brookings Institution, Harvard University, or anyone affiliated with either institution other than ourselves.

insurers to market lower-cost products aimed at low-risk enrollees would leave *net* premiums of the lowest-cost plans little changed. Thus, increased enrollment is unlikely; indeed, since the quality of the lowest-cost plans could fall, enrollment could actually fall.

In light of these concerns, we recommend that CCIIO not adopt its proposed two-stage procedure to estimate model coefficients, which the technical paper shows would increase predicted spending for low-risk enrollees and reduce it for high-risk enrollees, thereby exacerbating selection incentives. By contrast, the proposals related to enrollment duration factors appear beneficial along this dimension (and may have benefits along other dimensions). We are uncertain what to recommend with respect to the proposed severity-HCC-count interactions since this change would increase predicted spending for both the lowest- and highest-risk enrollees, and the net effect on selection incentives is thus unclear.

- ***Combining* steps to improve model fit with changes to the transfer formula could potentially reduce selection incentives and improve market outcomes:** The fact that the existing risk score model fits poorly on some dimensions does suggest that there may be opportunities to improve risk adjustment. In particular, improving model fit could allow the model to produce a more accurate *rank-ordering* of insurers' risk mixes, even if it also led to smaller, less appropriate transfers. Coupling changes to improve fit with changes to the transfer formula that magnified transfers could make it possible to achieve the benefits of a more accurate risk score model without shrinking already-too-small transfers.

  If CCIIO does elect to go down this road, we recommend that it give further consideration to improving model fit via a non-linear risk score model. Adopting a non-linear model has the potential to improve fit along some of the dimensions of interest without sacrificing fit along other dimensions (as the two-stage procedure would) and without creating new coding incentives (as the severity-HCC-count interactions approach would).

We close by commenting on the potential changes related to cost-sharing-reduction (CSR) plan variants that are discussed in the appendix of the technical paper. We make two main points:

- **Modifying risk adjustment to account for "silver loading" is appropriate, but CCIIO should refine its specific proposals:** Explicitly incorporating the cost of providing CSRs into risk scores and modifying the transfer formula rating term to account for the higher effective actuarial value (AV) of silver plans would mitigate incentives to avoid CSR enrollees (and higher-risk CSR enrollees in particular). However, rather than continuing to construct CSR-variant risk scores as a multiple of silver-tier risk scores, as the appendix proposes, CCIIO should explore creating separate risk score models for the CSR variants. Additionally, in modifying the silver AV used in the transfer formula rating term, CCIIO should use *state* average effective silver AVs, not the national average CCIIO is currently considering, as there are large cross-state differences in average effective silver AV arising from state Medicaid expansion decisions and various other factors.

- **It is unclear whether risk scores should incorporate a special adjustment to reflect the distinctive spending patterns of CSR enrollees:** The technical paper presents persuasive evidence that CSR enrollees spend less than would be expected based on their observed health status and cost-sharing, and it suggests that it may be appropriate to align CSR enrollees' risk scores with their observed spending. However, CCIIO's explanation

for CSR enrollees' unexpectedly low spending—that lower cost-sharing does not cause higher utilization in this population—is implausible. And, importantly, some plausible alternative explanations for this pattern, like that plans are currently doing a poor job serving their low-income enrollees or that low-income enrollees opt for plans with narrower networks or tighter utilization controls, would argue against making a special adjustment to align CSR enrollees' risk scores with their observed spending.

While we do not currently have a firm view as to whether CCIIO should or should not make such an adjustment, we do believe that determining the right policy response requires better understanding *why* CSR enrollees exhibit unexpectedly low spending. We note that a final decision would also need to weigh how adjusting risk scores for CSR enrollees would affect premiums of silver plans relative to plans in other metal tiers and the effects those shifts would have on net-of-subsidy premiums and federal outlays.

The remainder of this letter discusses these points in greater detail.

Changes Related to the Calculation of Risk Scores

The technical paper describes three potential changes to risk score calculations motivated by concerns that the current approach mispredicts plan liability for certain groups. First, it describes a two-stage estimation procedure that would assign a higher weight to enrollees with low predicted spending (as estimated in the first stage) in the second-stage regression used to estimate the final risk score model coefficients. Second, it describes a proposal to interact the existing enrollment duration factors with an indicator for the presence of one or more HCCs. Third, it describes a proposal to replace the current severity-HCC interactions in the model with a set of interactions between a severity indicator and several HCC count indicator variables.

In this section of the letter, we first lay out a general framework for how policymakers should evaluate these changes to risk adjustment, concluding that risk adjustment transfers are likely already too small and that policymakers should avoid steps that reduce the transfers paid by plans with lower-risk enrollees. We then use that framework to evaluate these three specific proposals, concluding that the two-stage estimation procedure is unappealing, while changes to the enrollment duration factors are appealing and the severity-HCC interactions may be either good or bad. Finally, we describe another path forward that would combine changes to improve model fit with changes to the transfer formula aimed at magnifying the overall size of risk adjustment transfers; this approach could realize the benefits of better fit without the downsides.[2]

Conceptual Framework for Considering Risk Adjustment Changes

The underlying goal of risk adjustment is to eliminate (or at least weaken) the link between what types of enrollees an insurer attracts and the costs the insurer incurs. Severing this link eliminates insurers' incentive to make plan design and pricing decisions based on how those decisions affect their enrollee mixes. That, in turn, can help avoid two types of problems:

- *Inadequate plan quality:* High-risk enrollees typically have a higher willingness to pay for more generous coverage (e.g., broader networks, lower cost-sharing, and laxer utilization

[2] Throughout this section our main focus is on how these changes would affect the individual market. Many of our qualitative conclusions likely apply to the small group market as well, although their quantitative importance is likely smaller since adverse selection is less intense in the small group market.

controls).[3] Thus, without effective risk adjustment, insurers are likely to charge high premiums for plans that offer those features or decline to offer such plans entirely. The result is that enrollees will be driven into overly stingy plans (e.g., plans with excessive cost-sharing, overly narrow networks, or overly stringent utilization controls).

- *Inadequate competition:* Inadequate risk adjustment may also make it hard to support robust competition among insurers. This is because low-risk enrollees may be particularly sensitive to premiums differences across plans. Consider, for example, a situation where most plans' networks have similar breadth overall but there is some variation in exactly which hospitals and physicians are included in each network, a relatively common pattern in the individual market. In this case, low-risk enrollees (who care a lot about premiums) may be much more likely to leave a plan in response to a premium increase than high-risk enrollees (who care a lot about which specific providers are included in a plan's network).

  Without effective risk adjustment, this pattern gives plans a strong incentive to undercut their competitors on price and capture the low-risk enrollees. While incentives to compete on price are often socially useful, unchecked selection can make these incentives so strong that all insurers lose money, forcing some to leave the market. At that point, consumers lose valued choices (e.g., access to a hospital not included in a surviving plan), and premiums can rise from reduced competition, burdening consumers or the government. Consistent with this concern, research by one of us (Layton) finds that Massachusetts' individual market would devolve to a monopoly without strong risk adjustment.[4]

The technical paper implicitly broaches the possibility that strong risk adjustment may also have an adverse side effect: discouraging enrollment by low-risk consumers. The logic is that stronger risk adjustment is likely to raise the premiums charged by the lowest-cost plans, both by directly increasing the risk adjustment transfers these plans pay and by encouraging insurers to offer more generous coverage, and thereby reduce enrollment. However, this argument neglects the fact that the large majority of individual market enrollees receive the premium tax credit, the value of which is based on the premium of the second-lowest-cost silver plan. Thus, if more aggressive risk adjustment does raise the premiums of the lowest-cost plans, it is likely to raise the value of the tax credit commensurately, leaving the *net* premium of the lowest-cost plans little changed. This implies that more aggressive risk adjustment is unlikely to reduce enrollment. Indeed, it may actually *increase* enrollment by improving the quality of the lowest-cost plans, thereby allowing potential enrollees to purchase a better plan for the same net premium.

Furthermore, it is not a given that weaker risk adjustment will reduce even pre-subsidy premiums for the lowest-cost plans. As risk adjustment weakens, it is indeed the case that transfers from low-risk plans to high-risk plans decrease, lowering the costs of the low-risk plans and raising the costs of the high-risk plans. However, as the high-risk plans raise their premiums in response to this shift, some consumers will choose the low-risk plans instead. Importantly, in a market with adverse selection, those "marginal" consumers will often be higher risk than current enrollees in the low-risk plan, driving the costs of the low-risk plan back up again, possibly offsetting the reduced

[3] See, for example, Michael Geruso et al., "The Two Margin Problem in Insurance Markets," *The Review of Economics and Statistics*, July 9, 2021, 1–46, https://doi.org/10.1162/rest_a_01070.

[4] Edward Kong, Timothy Layton, and Mark Shepard. "Adverse Selection and Insurer Participation in Health Insurance Markets." Work in progress.

transfers and causing its premiums to remain steady. Indeed, in simulations in work co-authored by one of us (Layton), this is precisely what occurs: weakening risk adjustment rarely actually *lowers* premiums for the lower-risk plans, and it always *raises* premiums for the higher-risk plans.[5]

In light of the potential harms from risk selection, an important question is how risk adjustment transfers should be calculated to minimize insurers' incentives to engage in it. A conventional approach, which is embodied in the current risk adjustment methodology, is to develop a model that predicts enrollee-level spending as well as possible based on a set of observed enrollee characteristics (in this case age, sex, and HCCs) and then calculate risk adjustment transfers based on the difference in average predicted spending across insurers. Under this approach, it is natural to choose risk adjustment coefficients to maximize measures of model "fit," like R^2.

But economic research on the design of an optimal risk adjustment system suggests that the transfers generated by this approach will generally be too small to fully mitigate selection incentives.[6] This is because many aspects of health status are not observed for risk adjustment purposes, so there is substantial variation in health care needs even among enrollees that look identical to the risk adjustment system. Furthermore, it is generally reasonable to expect that a plan that disproportionately attracts enrollees who appear high-cost on the characteristics that *are* captured in risk adjustment will also attract costlier-than-average enrollees within any given observable group.[7] (For example, a plan that attracts a disproportionate share of diabetics is likely to also attract diabetics who cost more, on average, than diabetics overall.)

Consequently, fully mitigating selection incentives is likely to require "underpaying" for enrollees with characteristics that predict low spending and, similarly, "overpaying" for enrollees with characteristics that predict high spending. In principle, this can be achieved via the transfer formula by scaling up transfers beyond the level suggested by risk score differences. But it can also be achieved by calculating risk scores in a way that *underpredicts* spending for enrollees with low predicted spending and *overpredicts* spending for enrollees with high predicted spending.

The preceding discussion supports two conclusions that will be useful in assessing the proposals in the technical paper. First, the fact that the current methodology underpredicts spending for enrollees with low predicted spending is likely a blessing in disguise. Thus, all else equal, efforts to remedy this "problem" are likely to actually exacerbate selection incentives, with deleterious downstream effects on plan quality and, potentially, competition. On the other hand, the analysis

---

[5] For a fuller analysis, see Geruso et al., "The Two Margin Problem in Insurance Markets," July 9, 2021.

[6] See Jacob Glazer and Thomas G. McGuire, "Optimal Risk Adjustment in Markets with Adverse Selection: An Application to Managed Care," *The American Economic Review* 90, no. 4 (2000): 1055–71; Jacob Glazer and Thomas G. McGuire, "Setting Health Plan Premiums to Ensure Efficient Quality in Health Care: Minimum Variance Optimal Risk Adjustment," *Journal of Public Economics*, ISPE Special Issue, 84, no. 2 (May 1, 2002): 153–73, https://doi.org/10.1016/S0047-2727(01)00123-2. For a recent empirical illustration of this finding, see Timothy J. Layton, Thomas G. McGuire, and Richard C. van Kleef, "Deriving Risk Adjustment Payment Weights to Maximize Efficiency of Health Insurance Markets," *Journal of Health Economics* 61 (September 1, 2018): 93–110, https://doi.org/10.1016/j.jhealeco.2018.07.001.

[7] For empirical evidence on this point, see: Michael Geruso et al., "The Two Margin Problem in Insurance Markets," SSRN Scholarly Paper (Social Science Research Network, May 3, 2019), https://papers.ssrn.com/abstract=3385492; Vilsa Curto et al., "Health Care Spending and Utilization in Public and Private Medicare," *American Economic Journal: Applied Economics* 11, no. 2 (April 2019): 302–32, https://doi.org/10.1257/app.20170295.

above suggests that underpredicting spending for enrollees with very high predicted spending is indeed pernicious and that fixing this problem would reduce selection incentives.[8]

We note that this conclusion would not hold if current patterns of over- and under-prediction were generating excessive risk adjustment transfers on net. But actual experience offers little reason to believe that this is the case. Indeed, it is striking that the individual market has become increasingly dominated by plans with relatively narrow networks and tight utilization controls.[9] While this could, in principle, be merely a response to consumer preferences for low-cost plans, it is consistent with substantial selection incentives remaining even after risk adjustment. At a minimum, this pattern suggests there is little reason to worry that current risk adjustment transfers are substantially too large since excessive transfers would likely drive these plans from the market or at least cause them to have relatively low market shares, which is clearly not the case.

Second, while maximizing model fit is the wrong overarching goal for risk adjustment, it is still the case that improving model fit by adding new predictor variables will often reduce selection incentives. As described above, the conventional approach of choosing risk adjustment coefficients to maximize the model R^2 goes awry because it fails to account for variation in spending associated with aspects of health status that are not captured in the available predictor variables. Thus, adding predictor variables that allow the model to better distinguish between higher- and lower-cost enrollees and, correspondingly, reduce the amount of unexplained spending variation can ameliorate the disadvantages of choosing model coefficients to maximize R^2.

Implementation of a Two-Stage Estimation Procedure

The framework outlined above suggests that CCIIO's proposal to adopt a two-stage estimation procedure is likely to increase selection incentives and worsen market outcomes. The two-stage approach prioritizes finding model coefficients that accurately predict spending for enrollees who are predicted to have low spending in the first-stage regression, while deprioritizing accuracy in other groups. Unsurprisingly, CCIIIO's results show that this approach both increases risk scores for enrollees with low predicted spending *and* reduces risk scores for enrollees with high predicted spending.[10] Consistent with the lessons from literature on optimal risk adjustment discussed above,

[8] The scope to *improve* competition by avoiding a race to the bottom in premium setting may be limited in light of the recent recovery in plan participation in the individual market. See Daniel McDermott and Cynthia Cox, "Insurer Participation on the ACA Marketplaces, 2014-2021," November 23, 2020, https://www.kff.org/private-insurance/issue-brief/insurer-participation-on-the-aca-marketplaces-2014-2021/.

[9] See, for example: John A. Graves et al., "Breadth and Exclusivity of Hospital and Physician Networks in US Insurance Markets," *JAMA Network Open* 3, no. 12 (01 2020): e2029419, https://doi.org/10.1001/jamanetworkopen.2020.29419; Erica Coe, Jessica Lamb, and Suzanne Rivera, "Hospital Networks: Perspective from Four Years of the Individual Market Exchanges" (McKinsey and Company, May 2017), https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/hospital-networks-perspective-from-four-years-of-the-individual-market-exchanges; Erica Coe, Alex Luterek, and Jim Oatman, "Insights into the 2019 Individual Exchange Market" (McKinsey and Company, December 2018), https://healthcare.mckinsey.com/insights-2019-individual-exchange-market.

[10] At first glance, Figure 2.2 of the technical paper suggests that predicted spending is roughly unchanged at the top of the predicted spending distribution. However, footnote 63 of the technical paper indicates that the mean risk score under the two-stage approach is 1.036 and that CCIIO has not renormalized the risk scores to have mean one. It is the normalized risk scores that would ultimately determine transfers, and it is clear from the results reported in Figure 2.2 that the *normalized* risk scores of enrollees with high predicted spending would fall.

changes like these are likely to increase insurers' incentives to engage in risk selection, raising the risk of a race to the bottom that reduces plan quality and competition.

We note that CCIIO finds that the proposed change also worsens overall model fit, at least as measured by R^2.[11] As we have argued above, there are strong rationales for deviating from the goal of maximizing fit, but the main arguments for doing so *do not* support prioritizing fit at the bottom of the spending distribution; indeed, they argue for doing precisely the opposite. Regardless, to the extent that CCIIO's objective is improving fit, this fact should the agency pause.

Recommendation: CCIIO should not adopt the two-stage estimation procedure.

Implementation of HCC-Duration-Factor Interactions

By contrast, CCIIO's proposal to interact the enrollment duration factors with an indicator for whether the enrollee has one or more HCCs appears likely to reduce selection incentives and thereby improve market outcomes. The results reported in Figure 3.2 of the technical paper suggest that, on average, the changes will reduce risk scores among enrollees with no HCCs and increase risk scores among enrollees with one or more HCCs. Combined with the fact that this change would improve model fit along other dimensions, the framework we laid out above suggests that these changes would reduce insurers' incentives to engage in risk selection.

Recommendation: CCIIO should adopt the HCC-interacted enrollment duration factors.

Implementation of Severity-HCC-Count Interactions

The effects of CMS' proposal to add severity-HCC-count interactions are more ambiguous. While CCIIO does not present estimates of the effect of this proposal on its own, the results presented in Figures 4.2 and 4.3 suggest that the effects of this change would vary over the risk score distribution. Namely, it would increase risk scores for the lowest-risk enrollees, reduce them for enrollees in an intermediate range, and raise them at the top of the risk distribution. It is unclear how this combination of changes would, on net, affect the transfers paid by insurers who tend to attract healthier enrollees and, thus, overall incentives for risk selection. Given this ambiguity, additional analysis of these changes' net effect on transfers would be valuable.

An additional potential concern with this change is that including HCC count variables may encourage insurers to invest additional resources in diagnosis coding. It is unclear, however, how large the incremental coding incentives under this approach actually are, so those costs might be worth bearing if this proposal increased the transfers paid by insurers with lower-risk enrollees.

Recommendation: CMS should analyze how this proposed change would affect the transfers paid and received by different types of insurers on net. If it would tend to reduce the risk adjustment transfers paid by insurers with lower average risk scores, then CCIIO should not adopt this change. If it would tend to increase the transfers paid by these insurers or leave those transfers roughly unchanged, then CCIIO should adopt this change.

[11] This is unsurprising since ordinary least squares regression maximizes R^2. Thus, any other estimation procedure—including the proposed two-stage procedures—must necessarily achieve a lower R^2. Also, we note that the reduction in model fit may be larger than CCIIO reports; footnote 63 of the technical paper states that the mean risk score under the two-stage procedure exceeds one. After the risk scores are appropriately normalized, fit could worsen.

An Alternative Path Forward

We argued above that a narrow focus on improving model fit (particularly a focus on improving specific dimensions of fit rather than overall model fit) is likely to reduce the effectiveness of risk adjustment. However, the fact that the current risk score model fits poorly along some dimensions does suggest that there may be opportunities to improve the performance of risk adjustment. In particular, we believe that changes that improved risk score model fit might produce a more accurate *ranking* of insurers with respect to risk mix (and a more accurate measure of the *relative magnitude* of differences in risk mix across insurers), even if they would also produce less accurate risk adjustment *transfers* when entered into the current transfer formula.

Thus, we believe that it would be worthwhile to explore an approach that combined two types of changes: (1) changes to the risk score model that improved overall model fit, including improving fit along some or all of the dimensions discussed in the technical paper; and (2) changes to the transfer formula to ensure the transfers owed by insurers with below-average risk did not fall and, ideally, increased. Below, we discuss what we would propose under each heading.

Changes to Improve Model Fit

While one prong of our proposed approach would be to improve model fit, paralleling the main goals of the technical paper, our vision for doing so differs from CCIIO's. In particular, since CCIIO's proposed two-stage estimation procedure actually worsens overall model fit (at least as measured by R^2), we suggest giving further consideration to non-linear models, which we suspect could address some of the problems the two-stage procedure is attended to address without worsening fit overall. (A non-linear specification could—and likely should—incorporate the modified enrollment duration factors in the technical paper. In principle, a non-linear model could also incorporate the proposed severity-HCC-count interactions, although we suspect that a non-linear model might obviate the need for these interaction terms in practice.)

We recognize that CCIIO sees some challenges in adopting a non-linear model. One significant barrier is that CCIIO reports encountering convergence problems when trying to estimate the parameters of a non-linear model. While troubleshooting these convergence problems is challenging without more information, we suspect that they are surmountable.

There are two specific approaches that we encourage CCIIO to explore if it has not already. To describe those approaches, it is useful to establish some notation. Specifically, CCIIO has indicated that its work with non-linear models has focused on specifications of the following form:[12]

$$Y = X\alpha + \mathrm{HCC}\beta + (\mathrm{HCC}\beta)^{\gamma}, \qquad (1)$$

where Y is plan liability, X is the vector of non-HCC variables included in the risk score model, and HCC is the vector of HCC indicators included in the risk score model.

Our first suggestion is that CCIIO consider a different algorithm for obtaining the least-squares estimates of the parameters in equation (1). We suspect that CCIIO is currently using an algorithm that iterates over all of these parameters simultaneously and that a "nested" algorithm might produce better results.[13] Concretely, we envision conducting a "top-level" search over the

[12] See 85 FR 29188.

[13] This algorithm is similar in some respects to the family of "alternating optimization" algorithms, under which a function is alternatively optimized with respect to one parameter and then the other until convergence is reached.

coefficient γ alone; for each value of γ, the algorithm would then iteratively solve for the values of α and β that maximized fit for the current value of γ. Our conjecture is that solving for α and β for a fixed γ would not present convergence problems. Similarly, the "top-level" search would be over a single parameter, which should limit convergence problems at that stage as well.

Our second suggestion is that CCIIO experiment with the alternative functional forms:

$$Y = (X\alpha + \mathrm{HCC}\beta)^{\gamma}. \quad (2)$$

Equation (2) differs from CCIIO's existing specification in that it applies the exponent symmetrically to the HCC and non-HCC terms. We suspect that this functional form might have somewhat better numerical properties. This might be particularly true in combination with the "nested" approach described above since it would make it possible to obtain starting values for α and β for each fixed value of γ from a regression of $Y^{1/\gamma}$ on the vectors X and HCC. We also see it as appealing that the specification in equation (2) treats the HCC and non-HCC risk adjusters symmetrically, as we see no clear rationale for the asymmetry reflected in equation (1).

Changes to the Transfer Formula

The main potential downside of these steps to improve model fit, similar to the steps that CCIIO advances in the technical paper, is that they could reduce the transfers paid by insurers that disproportionately attract low-risk enrollees. For the reasons discussed above, we believe that reducing these transfers would increase incentives for risk selection and worsen market outcomes.

Those consequences could, however, be avoided with suitable modifications to the transfer formula. While there are a variety of modifications to the transfer formula that would achieve this objective, one straightforward approach would be to simply apply a multiplier to the transfers determined under the current formula. Conceptually, this multiplier would be justified by the theory and evidence described earlier in this letter that insurers who attract enrollees who appear to be low-risk on dimensions of health status that *are* captured in risk adjustment are likely to also attract enrollees who are low-risk on unmeasured dimensions of health status.

Ideally, this multiplier would ultimately be set using evidence on how much of the difference in spending across plans is attributable to enrollee characteristics that are not captured in risk adjustment. However, it could also be set (at least on an interim basis) to achieve a specified objective, like ensuring the total value of transfers paid by insurers with average risk in the bottom quartile of the risk distribution do not pay less under the revised risk score model and transfer formula than they do with the current risk score model and transfer formula. But overall, we would strongly suggest adjustments to the transfer formula that result in larger transfers rather than maintaining the status quo, as we believe the costs of doing this to be minimal while the potential benefits for individuals in this market could be substantial.

Changes Related to CSRs

In closing, we comment on the changes related to CSR variants discussed in the appendix of the technical paper. We view the appendix as raising two conceptually distinct issues: (1) how the lack of CSR payments and the transition to silver loading should be accounted for in risk adjustment; and (2) how to construct risk scores for CSR enrollees in light of evidence that they consume less care than expected based on their observed health status and cost-sharing. We address each in turn.

Reflecting Silver Loading in Risk Adjustment

In general, we believe that risk adjustment should reflect the liabilities plans actually incur and the revenues they actually collect under silver loading. Concretely, the risk scores used to calculate the risk term of the transfer formula should reflect the fact that delivering coverage with CSRs increases plan liability relative to delivering equivalent coverage without CSRs. And the rating term of the transfer formula should reflect the fact that silver plans are now priced in accordance with an actuarial value (AV) higher than the nominal 70% AV of a silver plan.

On a conceptual level, this approach has two appealing features. First, it mitigates plan incentives to avoid CSR enrollees stemming from those enrollees being more costly due to the more generous coverage provided to them. Second, it ensures that plans do not have particular incentives to avoid high-risk CSR enrollees by aligning the compensation plans receive for those enrollees with the actual generosity of their coverage rather than the nominal 70% AV of a silver plan.

We note that this general conclusion is predicated on the assumption that silver loading should continue. If policymakers viewed "broad loading" as preferable, they could achieve that by reflecting the lack of CSR payments in calculating risk scores, but not reflecting the higher effective AV of silver plans in the rating term. This approach would generate larger transfers into the silver tier that would, in effect, remove the incentive to silver load. The transition to broad loading would, however, raise premiums for non-silver plans (both subsidized and unsubsidized), while reducing premiums for relatively few Marketplace enrollees with unsubsidized silver plans and reducing federal spending on premium tax credits. We suspect that this is not what CCIIO is aiming to achieve. It is also, in our view, likely undesirable on substantive grounds.

In light of this discussion, we make two specific recommendations about how CCIIO should modify risk adjustment to account for silver loading:

- *CSR enrollee risk scores:* Risk scores for CSR enrollees should be calculated using separate models that reflect plan liability in CSR variants, similar to the approach used to calculate risk scores for different metal tiers. Similarly, the induced utilization factors for each CSR tier should be set based on existing evidence on the relationship between cost-sharing and utilization. (Below, we assess the question of whether a special CSR-variant adjustment is warranted to reflect actual spending patterns in the CSR population.)

  The appendix suggests that creating variant-specific models would add substantial complexity and that it would be preferable to continue calculating risk scores for CSR enrollees by applying a multiplier to risk scores from the silver model. It is unclear to us why the added complexity is so large (and why that complexity is worth bearing with respect to different metal tiers but not the CSR-variants). Regardless, to weigh complexity concerns, it would be useful to better understand how risk scores derived from CSR-variant-specific models would compare to risk scores derived from the silver-tier model at different risk levels. Because of the differences in cost-sharing structure between the CSR variants and base silver plans, we suspect that these differences could be meaningful.

- *Silver plan AV in the rating term:* We generally agree with CCIIO that the rating term should incorporate the average effective AV of silver plans, inclusive of the value of CSRs, rather than the nominal AV of 70%. However, we think that CCIIO's proposal to use a national average effective silver AV in the rating term is ill-advised. There are large differences in effective silver AV across states, reflecting a variety of factors, most

importantly whether or not states have expanded Medicaid or operate Basic Health Programs.[14] Using a state-specific average silver AV would ensure that risk adjustment appropriately reflects these differences across states.

Constructing Risk Scores for CSR Enrollees

The technical paper presents evidence from EDGE data indicating that CSR enrollees spend less than would be expected given their observed health status and the actuarial value of the coverage they hold. Specifically, Table A.4 suggests that the *current* risk score model and induced utilization factors either slightly underpredict or modestly overpredict plan liability in the largest CSR tiers. This is surprising since the current risk scores are calculated to reflect the *nominal* 70% AV of silver plans, rather than the higher AVs of the CSR variants, so one would expect plan liability for these enrollees to be underpredicted if the model was otherwise predicting accurately.

The technical paper suggests that the reason for this surprising finding is that, contrary to expectations, the lower cost-sharing faced by CSR enrollees is not causing them to spend more than they would in the absence of CSRs. In support of this view, Table A.3 presents results from a cross-sectional regression showing that CSR recipients and non-recipients have similar spending after controlling for the individual characteristics observed for risk adjustment purposes.

We view this explanation as unlikely. A large literature convincingly establishes that lower cost-sharing increases utilization, including among low-income people.[15] More importantly, there are a variety of other explanations for the findings in Table A.3. Having a lower income may directly reduce demand for health care, even holding coverage characteristics constant. Alternatively, lower-income enrollees could be opting for lower-priced plans with non-cost-sharing features that tend to reduce utilization or provider prices (e.g., narrower networks or tighter utilization controls), or plans could be doing a poor job of making care accessible to their low-income enrollees, reducing utilization. It is also possible that CSR enrollees are healthier along dimensions that are not captured in risk adjustment; for example, healthier enrollees who would opt for bronze plans at higher income levels may opt for silver plans when CSRs are available.[16]

---

[14] See Matthew Fiedler, "The Case for Replacing 'Silver Loading'" (Brookings Institution, May 20, 2021), https://www.brookings.edu/essay/the-case-for-replacing-silver-loading/.

[15] For a few notable examples, see Willard G. Manning et al., "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *The American Economic Review* 77, no. 3 (1987): 251–77; Amitabh Chandra, Jonathan Gruber, and Robin McKnight, "The Impact of Patient Cost-Sharing on Low-Income Populations: Evidence from Massachusetts," *Journal of Health Economics* 33 (January 1, 2014): 57–66, https://doi.org/10.1016/j.jhealeco.2013.10.008; Zarek C. Brot-Goldberg et al., "What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics," *The Quarterly Journal of Economics* 132, no. 3 (August 1, 2017): 1261–1318, https://doi.org/10.1093/qje/qjx013; Liran Einav and Amy Finkelstein, "Moral Hazard in Health Insurance: What We Know and How We Know It," *Journal of the European Economic Association* 16, no. 4 (August 1, 2018): 957–82, https://doi.org/10.1093/jeea/jvy017.

[16] There are other reasons one might actually expect CSR enrollees to be sicker, namely that health status generally declines with income. See, for example, Raj Chetty et al., "The Association Between Income and Life Expectancy in the United States, 2001-2014," *JAMA* 315, no. 16 (April 26, 2016): 1750–66, https://doi.org/10.1001/jama.2016.4226. The relationship between income and health status could differ among people who actually enroll in coverage if the more generous premium subsidies and the CSRs themselves encourage higher take-up at lower income levels. It is unclear, however, that take-up is actually higher at lower-income levels. See Matthew Fiedler, "Enrollment in Nongroup Health Insurance by Income Group" (Brookings Institution, March 25, 2021), https://www.brookings.edu/research/enrollment-in-nongroup-health-insurance-by-income-group/.

We dwell on the question of *why* spending is lower because it is relevant to deciding whether and how to reflect these spending patterns in risk scores. If CSR enrollees have relatively low spending because plans are doing a poor job of serving them, then incorporating the lower spending into their risk scores may not be desirable, as this would encourage plans to do an even worse job of serving these enrollees. (As an analogy, if plans were doing a poor job of serving people with a specific health condition, resulting in unexpectedly low utilization in that group, reducing the group's risk scores is unlikely to be the right response.[17]) It may also be inappropriate to reduce the risk scores of CSR enrollees if their lower spending arises because they tend to choose plans that are less generous on non-cost-sharing dimensions. We do not, at present, have a strong view on the correct explanation and, thus, the correct policy response, but we do believe that this is a key question that must be answered to chart a path forward.

In closing, we note that choices about how to construct risk scores for CSR enrollees would also need to consider the effects on relative premiums of plans in different metal tiers. Reducing risk scores for CSR enrollees would tend to reduce transfers into the silver tier, increasing silver plan premiums and reducing premiums in other metal tiers. Like silver loading, this shift would tend to reduce premiums for people who purchase non-silver plans, while increasing premiums for the small number of unsubsidized enrollees who purchase Marketplace silver plans and increasing federal spending. Again, we do not take a position on how to weigh these effects here, but not that these effects should enter into CCIIO's ultimate decisions in this area.

Thank you for the opportunity to comment on these potential changes. We hope this information is helpful to you. If we can provide any additional information, we would be happy to do so.

Sincerely,

Matthew Fiedler
Fellow, USC-Brookings Schaeffer Initiative for Health Policy
Economic Studies Program
Brookings Institution

Timothy Layton
30th Anniversary Associate Professor of Health Care Policy
Harvard Medical School
and
Faculty Research Fellow
National Bureau of Economic Research

[17] For more on this point, see Savannah L. Bergquist et al., "Intervening on the Data to Improve the Performance of Health Plan Payment Methods," Working Paper (National Bureau of Economic Research, April 2018), https://doi.org/10.3386/w24491.