Authors' Response to American Hospital Association's review of
Desai and McWilliams, "Consequences of the 340B Drug Pricing Program"
***New England Journal of Medicine*. 2018; 378: 539-548**

We thank the AHA for its comments on our study[1] and applaud the AHA for requesting a review from an independent economist with expertise in the methods we employed. Below we provide a point by point response to the concerns raised by the AHA and the reviewer, Dr. Partha Deb, including summaries of additional analyses conducted to address them empirically. In summary, none of these analyses substantively affected our results, and our conclusions remain unchanged. The specific methodological points raised about our implementation of the regression discontinuity design are minor; none are of great consequence, let alone "negate" our findings. Like any study, ours has some limitations, which we described clearly in the limitations section of our paper. In particular, we acknowledged that the use of surplus resources generated from drug discounts by some hospitals may differ from the effects we estimated for hospitals near the 340B Program eligibility threshold. We also noted that our reliance on Medicare and HCRIS data may have missed some beneficial uses of surplus resources for low-income groups. Nevertheless, our study should not be dismissed, as it presents strong evidence of hospital responses to Program incentives that are inconsistent with Program goals, evidence that is sufficient to motivate Program reform and more generally inform the design of policies intended to enhance care for low-income populations.

We also note that our study is not the only one to question whether the 340B Program has efficiently subsidized efforts to meet the unmet needs of underserved populations. Prior studies include one that found 340B hospitals have affiliated disproportionately with practices in more affluent communities[2] and another that found Program expansions under the ACA have shifted cancer treatment away from community practices and into hospital-owned facilities, with overall increases in area-level total spending on cancer care.[3] Our paper differs from another prior paper[4] that found no hospital-physician consolidation associated with expansions of the 340B Program under the ACA to additional categories of hospitals, such as critical access hospitals, that may be less able to pursue mergers and acquisitions. Our paper focuses on general acute care hospitals, the hospital type that is most common, accounts for the majority of hospital spending, and was a category already exposed to the 340B Program before the ACA. The regression discontinuity design we employed also avoids having to make strong assumptions about how ongoing trends in hospital-physician consolidation would have evolved in the absence of the 340B Program.

The study uses a methodological approach called "regression-discontinuity" that does not support making broad claims about the 340B program

Generally speaking, any strong study design trades off external validity for internal validity. Like instrumental variables estimation methods, regression discontinuity designs identify causal effects of treatment on a marginal subject – in this case the effect on hospitals close to the 340B eligibility threshold. While marginal effects are not the average treatment effects among the full treated population, they are often of great policy relevance, if not greater relevance. For example, our results would inform consideration of policies to roll back expansions of the 340B Program to focus more exclusively on hospitals serving very high proportions of underserved patients (e.g., by increasing the DSH percentage for eligibility) or policies that would expand care for the underserved more directly, without risking inefficiencies on the margin.

We agree that our approach limits the generalizability of our findings with respect to hospitals with high DSH percentages and categories of hospitals we did not study (we studied only general acute

hospitals), which is why we included a clear description of this limitation in the limitations section of our paper:

"*Third, our regression-discontinuity approach supported inferences about hospitals just above the eligibility thresholds. Hospitals with higher DSH percentages could have responded differently to program discounts. Fourth, our conclusions may not apply to categories of eligible hospitals we did not study, such as critical access hospitals.*"

The limited external validity of our results, however, does not negate their importance. Instead, it raises questions about how other hospitals respond to the 340B Program that future research should address. In randomized control trials, too, some external validity is sacrificed to isolate the causal effects of treatment in a specific population. Yet when we learn that a drug does not benefit relatively healthy patients with a certain condition, for example, we do not dismiss that evidence when considering whether it might be beneficial for a sicker population with the same condition. The appropriate response is to exercise caution when using the drug and to try to learn more about its effects in other populations. Accordingly, we would encourage more research rigorously examining the effects of the 340B Program on hospitals with high DSH percentages.

In addition, the generalizability and relevance of a marginal effect depends on how atypical the marginally affected subject is. In our study, hospitals near the Program eligibility threshold were not atypical. For example, general acute care hospitals included in our study with DSH percentages within 10% of the eligibility threshold accounted for 56% of hospital inpatient spending in Medicare and 58% of hospital outpatient spending among public or non-profit general acute care hospitals in 2010, and public or non-profit general acute care hospitals accounted for 66% of spending among all hospital types.

The study relies on fee-for-service Medicare data only to make claims about the impact of the 340B program on low-income individuals, ignoring that the vast majority of low-income people are not enrolled in Medicare. Low-income individuals are diverse: the largest subgroup are children, who make up approximately 35 percent of low-income individuals, followed by the working poor. Only 23 percent of low-income individuals are elderly or disabled and therefore potentially eligible for Medicare. However, the study authors chose to look at fee-for-service Medicare claims only to evaluate the impact of the 340B program on low-income individuals more broadly.

The total reliance on fee-for-service Medicare claims is an odd choice for another reason – that the study authors claimed to be interested in looking at improvements in access to care. Medicare beneficiaries receive a relatively comprehensive benefit package compared to other insured individuals and are the group most likely to have a medical home. The authors could instead have evaluated the impact on vulnerable populations such as the uninsured, the underinsured and Medicaid beneficiaries

That Medicare benefits are comprehensive is a popular misconception. Coverage in the traditional fee-for-service program leaves beneficiaries exposed to substantial cost-sharing and does not include stop loss provisions. Thus, beneficiaries without supplemental sources of coverage are at a high risk of financial ruin if they become ill. While most fee-for-service Medicare beneficiaries have a source of supplemental insurance, approximately one in five do not, and that proportion is higher for low-income groups, black beneficiaries, and beneficiaries with high medical needs.[5] For example, 36% of beneficiaries with less than a high school education and 26% of beneficiaries who

were admitted to a post-acute nursing facility reported no source of supplemental coverage in 2014, based on data from the Consumer Assessment of Healthcare Providers and Systems. An additional approximately 20% of Medicare beneficiaries have supplemental benefits from Medicaid, which generally reimburses providers at significantly lower rates for the portion of care not covered by Medicare, or from assistance through the Medicare Savings Program, which often provides minimal additional coverage and also reimburses at lower rates. These dual eligible enrollees account for a high proportion of Medicaid spending – for example, 36% of Medicaid spending in 2010.[6] Given the much greater burden of illness than in the rest of the population, on average, the Medicare population therefore includes a substantial proportion of patients with the greatest unmet needs. These patients would be natural targets of hospital efforts to enhance care for the underserved. Accordingly, we focused on subgroups of the Medicare population, including dual-eligible enrollees, beneficiaries residing in low-income areas, and beneficiaries served by safety-net providers (who also disproportionately serve Medicaid enrollees and the uninsured).

In addition, because these subgroups of Medicare beneficiaries live in the same communities or are served by the same providers as Medicaid enrollees and the uninsured, we would expect that major investments in care for Medicaid patients or the uninsured would spillover onto the care provided to low-income Medicare beneficiaries. For example, if hospitals are using savings from the 340B hospitals to provide "increased access to care" and "community outreach programs" (as stated here), we would expect those efforts, if substantial, would affect underserved and needy patients in the Medicare population, too.

Finally, we do not rely exclusively on Medicare data. We also used HCRIS data to examine effects on hospital investments in federally qualified community health centers (FQHCs), which would affect all patient populations served by FQHCs. We found no evidence from these data of greater hospital integration with FQHCs or increases in staffing or amount of care delivered in FQHCs already integrated with hospitals.

We acknowledge that the data available to us may have missed Program effects on low-income populations outside of Medicare and clearly stated our reliance on Medicare data as a limitation of the study (excerpted below). We would welcome the use of other data sources and rigorous designs to further study the extent to which hospitals provide expanded or improved care for other low-income populations as a result of the 340B Program. But the consistent lack of an effect on the measures we were able to examine suggests that any benefits have been small relative to the magnitude of the surplus generated from the Program. Thus, our findings should not be dismissed as they support greater scrutiny of the ways in which hospitals are using the financial gains from the 340B Program.

> *"Our study had several limitations. First, we relied largely on Medicare data. We would expect, however, that major investments in clinical resources for low-income groups outside of Medicare, such as the uninsured, would also affect care for low-income Medicare beneficiaries. Moreover, we found no evidence of enhanced care for a subgroup of Medicare patients with supplemental insurance that is less generous or reimburses hospitals at lower rates than private supplemental insurance. In addition, measures of hospital investments in FQHCs were not specific to Medicare."*

The study authors applied their own limited interpretation of the intent of the program when selecting measures to evaluate 340B hospitals. According to the authors, "…the program may not elicit the intended responses from hospitals – such as providing more care to low-income communities, investing in safety-net providers, or reducing health

disparities….” While these are all laudable goals, they reflect the study authors’ own beliefs of how the 340B program should work, not Congress’s. Congress intended for the program to “stretch scarce Federal resources as far as possible, reaching more eligible patients and providing more comprehensive services.”[3] While many hospitals may use 340B savings to invest in safety-net providers or fund programs intended to reduce health disparities, there are many ways in which 340B entities may meet the true Congressional intent of the program of expanding access to care and more comprehensive services. The study authors cannot make broad claims about program performance when looking at such limited performance measures.

We believe our characterization of how hospitals would implement Congress’s stated intention for the Program is reasonable, broad, and consistent with the ways the AHA reports hospitals are using 340B Program savings. We are unsure how the language in our paper departs significantly from the Program’s stated goal of reaching more low-income patients and providing them with more services, the original focus of the program on safety net providers when created in 1992, or the AHA’s reports of hospitals providing “increased access to care,” “community outreach programs,” and “free vaccinations.” The full text of from our paper is as follows:

> “*The program is explicitly intended to encourage hospitals to dedicate resources generated from the discounts to expanding or improving care for vulnerable populations, particularly those served by safety-net providers… Thus, the program may not elicit the intended responses from hospitals — such as providing more care to low-income communities, investing in safety-net providers, or reducing health disparities — and may even have unintended consequences*.”

The study also fails to account for changes in coding of physician practices during the study period. The authors note that “…increases in hospital ownership of physician practices could have been over stated if practices owned by hospitals merely changed place-of-service codes….” However, they choose to ignore a change in federal reporting requirements that did just this. Beginning in 2011, the Health Resources and Services Administration (HRSA) required that all outpatient and other community-based sites of care that intended to use 340B drugs for their patients register separately for the 340B program, along with other requirements. By ignoring this HRSA reporting change, the study authors fail to acknowledge that the increase in the registration of hospital-owned outpatient clinics and services in the 340B program may simply be a matter of changes in reporting.

We thank the AHA for bringing this regulatory change to our attention, though we were unable to find further detail that would suggest that the change in reporting requirements should have caused a change in billing among hospital-owned outpatient facilities and practices from independent office settings to the hospital-owned setting (which was the basis for our assessment of hospital-physician consolidation). Nevertheless, to address this concern, we repeated our analyses after restricting the study years to those before 2011, and our results remained unchanged. In fact, estimates of hospital-physician consolidation associated with the Program eligibility threshold became larger, and the estimate for ophthalmologists became statistically significant. Specifically, the adjusted discontinuities at the 340B eligibility threshold were 2.7 additional hematologist-oncologists/hospital (p=0.01) and 1.6 additional ophthalmologists/hospital (p=0.03). Thus, the regulatory change in 2011 could not have explained our results.

As noted in our limitations section, changes in place of service coding from office to hospital-owned in practices already owned by hospitals could have contributed to our results. But hospitals already had strong incentives to implement such coding changes in the absence of the 340B

Program because outpatient care is reimbursed by Medicare (and commercial payers) at higher total rates when delivered in settings coded as hospital-owned than in the independent office setting. Moreover, any contributions from changes in coding to our estimates would not change our overall conclusion that there is evidence of hospitals responding to Program incentives by maximizing financial gains without clear accompanying evidence of expanded care for low-income patients.

Additional comments on analysis from Dr. Partha Deb:

Unclear text on whether the DSH percentage variable is centered around the 340B eligibility threshold of 11.75%.

We thank Dr. Deb for picking up on this inconsistency. We confirm that in the analysis the hospitals' DSH percentage variable was centered around the threshold for 340B eligibility and therefore defined as the difference between the hospital's DSH percentage in the previous year and 11.75%. As such, the coefficients can indeed be interpreted as the effect of eligibility for a hypothetical hospital with a DSH percentage of 11.75%. In an earlier version of our manuscript, we had made this clear in the equation (DSH-11.75), but in our efforts to condense the equation for inclusion in the main body of a published version, we removed the "-11.75" and instead added clarifying text in the description of the variables to note that "DSH" was centered around the threshold value. Over the course of multiple drafts, unfortunately this clarifying text was inadvertently omitted. We have corrected the equations in the Appendix of our paper and have notified the journal of this textual omission so that we can post the updated Appendix. That our estimates are derived from the correct equation is evident from the visual discontinuities in Figure 1 (and figures in the supplementary appendix), which are consistent with estimated discontinuities from our models, as well as from the consistency of our results from a model with a single constrained slope, in which centering "DSH" around the threshold is unnecessary.

The authors drop all hospitals with less than 50 beds, which they state include critical access hospitals and sole community hospitals. The argument that critical access hospitals and sole community hospitals are different from other short term general hospitals is a reasonable one. But there are plenty of small short term general hospitals and it is not clear why they should be dropped from the analysis. Admittedly, such hospitals would get small weights in the beds weighted regressions, but they do constitute a substantial fraction of hospitals. Their inclusion could change results substantively.

As Dr. Deb suspects, the inclusion of hospitals dropped because they have fewer than 50 beds continues to support conclusions that the 340B Program contributed to consolidation in hematology-oncology (2.4 additional hematologist-oncologists/hospital, $p = 0.02$) and increased drug administration by hospitals in hematology-oncology (220 more drug claims billed/hospital, $p=0.002$) and ophthalmology (76 more claims/hospital, $p=0.03$).

Hospitals with fewer than 50 beds were dropped from the main analyses to ensure all critical access hospitals and sole community hospitals were removed. These two hospital types have fewer than 50 beds and are eligible for 340B through alternate criteria and can be reimbursed differently by Medicare. Specifically, following the expansion of the 340B Program under the Affordable Care Act, critical access hospitals are 340B eligible regardless of their DSH percentage and sole community hospitals with a DSH percentage exceeding 8% are eligible. Inclusion of hospitals eligible through alternate criteria would increase imprecision of our estimates. Our study

hypotheses are appropriate to general acute hospitals rather than special designation hospital types that might have different incentives to maintain their special status.

Moreover, hospitals dropped because they have fewer than 50 beds constitute a small share of care provided to Medicare beneficiaries - 2% of Medicare hospital inpatient and 3% of hospital outpatient spending in 2010 the midpoint of our study period. Further, we demonstrate that study conclusions remain unchanged after inclusion of these hospitals (hospitals with fewer than 50 beds but not explicitly categorized as special designation hospitals) in analyses that weight hospitals by number of beds as well as analyses that weight all hospitals equally.

The authors also drop hospitals with DSH percentages within 1 percentage point of the threshold, leading to omission of a non-trivial number of hospitals. Omission of these hospitals quite likely has substantive implications for the results. The authors argue for dropping hospitals within 1 percentage point as being due to misclassification, which is an issue but a common one in studies that use regression discontinuity designs. It can be dealt with, as the authors do, using instrumental variables methods. But for the main analysis, the authors drop these observations and cite a paper that deals with heaping, which is not the same thing as misclassification. More substantively, the context of the cited paper is not relevant here. That paper states, in the abstract: This study uses Monte Carlo simulations to demonstrate that regression-discontinuity designs arrive at biased estimates *when attributes related to outcomes predict heaping in the running variable*. (emphasis added)

While the primary reason to exclude hospitals within 1% of the eligibility threshold was to reduce measurement error introduced by misclassification of hospital eligibility for hospitals close to the eligibility threshold, an additional reason was to exclude hospitals that may have manipulated their DSH percentage to select into eligibility. Misclassification, which largely results from differences in HCRIS reporting periods and/or 340B participation that can begin mid-year, was vastly greater within 1% of the eligibility threshold relative to hospitals outside of that range. Specifically, 41% of observations with a DSH percentage in the previous year between 10.75% and 11.75% were defined in the data as 340B participants. In contrast, only 3.8% of observations with a DSH percentage between 1.75% and 10.75% in the previous year were 340B participants suggesting there is substantially greater misclassification among observations within 1% of the threshold.

While a number of analyses detailed in the Supplementary Appendix ruled out evidence of meaningful hospital manipulation of DSH percentage by hospitals to select into 340B eligibility, the secondary rationale for dropping these observations was a precautionary measure to exclude hospitals that were most likely to engage in such manipulation (those closest to the eligibility threshold). The cited paper[7] addresses such manipulation in RD studies and discusses the donut-RD approach as yielding unbiased estimates of the treatment effect. See Barreca, Lindo, and Waddell (2011) for further discussion of this approach and an implementation.[8] This working paper version of the study focused more on the issue of treatment manipulation near the threshold, but we cited the published version per journal style.

The reviewer is correct that an instrumental variables approach can be used to address measurement error. We reported instrumental variables (IV) analyses results in the Supplementary Appendix to provide estimates of the effects of program participation, which are greater than estimated effects of Program eligibility because not all eligible hospitals participate in the Program and this approach could be extended to include observations in the donut. However, misclassification would be expected to contribute to a weaker instrument and in turn less precise estimates. Indeed, when we conduct a post hoc analysis including observations in the donut, our IV

estimates are similar to those from our main IV analysis (enforcing the donut) but slightly less precise. For the reasons described above and in our manuscript, including evidence that the misclassification was overwhelmingly concentrated within 1% of the eligibility threshold, we made an a priori decision to use the donut-RD as our primary approach to address the misclassification and potential manipulation around the threshold in a clean fashion.

The authors use number of beds as importance weights in the hospital-level regressions. If number of beds are important as weights, they may well be important as regression controls. In fact, one would expect hospital size to be significantly related to the number of physicians employed. To be precise, the authors show that there is no evidence of a discontinuity in number of beds across the threshold, suggesting that number of beds as a regression control would not change the results substantively, but that remains an untested proposition. Given the a priori expectation of substantial associations between hospital size and the outcomes, it should be included as a covariate, not just as a weight, in the regressions.

As Dr. Deb suspected, controlling for beds does not change the study conclusions. Results continue to support that the 340B Program contributed to hospital consolidation in hematology-oncology (2.1 more hematologist-oncologists/hospital, p=0.03) and greater parenteral drug administration in hospital outpatient departments in hematology-oncology (192 more drug claims billed/hospital, p =0.001) and ophthalmology (64 more drug claims billed/hospital, p=0.03).

The number of beds was used as a proxy for hospital size to weight the analyses, since we (as well as the editors and referees) believed that hospitals constituting a greater portion of care should be given greater importance/weight in the analyses. Our main analyses do not control for hospital size (e.g. number of beds) because conceptually, size could be affected by the 340B Program. For example, hospitals could use the financial gains from the 340B drug discounts and accompanied consolidation to increase inpatient capacity. Including the number of beds as an independent variable could theoretically "control away" some of the effect we aim to measure. It was for this reason that we tested whether number of beds trended continuously across the eligibility threshold. Thus, as expected, the study conclusions remain unchanged after including controls for the number of beds in the specification.

Moreover, in sensitivity analyses, we conducted analyses without weights (Table S10) and analyses fixing weights to the number of beds in a single year to address potential bias in hospital response to 340B via the number of beds. In both, study conclusions remain unchanged.

The authors use census regions as geographic controls in their primary analyses and states in a supplementary analysis. Both might be considered inadequate. A much sharper quasi experiment would compare hospitals on either side of the DSH threshold in the same market (e.g., hospital referral region (HRR)).

To be clear, the RD design should control for geography because we would not expect a discontinuous change in the geographic distribution of hospitals at the eligibility threshold (by design, hospitals just above and below the threshold are quasi-randomized to program exposure). We demonstrated that empirically with respect to census region. For this reason, estimating discontinuities among hospitals in the same HRR would not be expected to produce a "sharper quasi-experiment." Even if within-HRR comparisons were desirable, there are generally too few hospitals per HRR to establish a counterfactual based on the relationship between DSH percentage and outcomes. Such a within-HRR analysis would not just include HRR fixed effects in the model but

would also include interactions between HRR and the slopes with respect to DSH to estimate a within-HRR effect. The inclusion of geographic fixed effects only generates estimates that reflect discontinuities in the deviations between hospitals and their area averages, not strictly within-area effects. Moreover, conceptually it is not clear that within-area effects would be the most informative, since responses by 340B hospitals may have triggered competitive responses by ineligible hospitals that could obscure some of the effects of the Program.

Including HRR fixed effects instead of state or region effects yielded discontinuity estimates that are overall similar in magnitude and more precise (greater "statistical significance" presumably from the variance reduction achieved with smaller geographic units as predictors). For the analysis of hospital-physician consolidation, adjusted discontinuity estimates are 2.1 more hematologist-oncologists/hospital (p=0.02) and 1.1 more ophthalmologists/hospital (p=0.01).

The confidence intervals shown in Figure 1 are probably not correct and paint a picture with more statistical precision than is likely true. The authors report 95% confidence intervals, yet the data points (which are group means) are frequently outside the intervals. That strikes me as being incorrect.

The plotted 95% confidence intervals in Figure 1 were presented for illustrative purposes and were calculated for the hospital-level means (not the bin means which are displayed). Because the number of hospitals varies by bin, bins with fewer hospitals contribute less to the mean and so may sometimes lay outside the confidence intervals. As we note in our manuscript, the confidence intervals for our results were estimated appropriately, using robust variance estimation. Thus, the appearance of the figure should not be interpreted as evidence that we underestimated the uncertainty of our estimates.

For patient-level analysis, zip code is not a reasonable definition of a market area. Most researchers would use areas defined by Dartmouth (hospital service area (HSA) or HRR). When local policy may have impacts, researchers might use counties or statistical areas as market areas. Zip codes are almost meaningless in any of those contexts.

The purpose of the patient-level analysis was not to define a hospital market but to identify an area around a hospital with substantial exposure to the hospital for the purposes of measuring the effects of potential community investments hospitals have made in response to the 340B Program. We show in Table 2 that inpatient and outpatient care for beneficiaries residing in a hospital's ZIP code are in fact attributable to the local hospital, providing clear evidence that ZIP codes constitute meaningful geographic designations for this study objective. Specifically, residing in a ZIP code with a 340B hospital is associated with a 48 percentage point increase in the portion of inpatient admissions and hospital outpatient spending attributable to a 340B hospital.

ZIP codes are preferable to hospital service areas or hospital referral regions, because hospital service areas and hospital referral regions are based on admission flows which could be affected by the 340B Program and therefore could be endogenous. ZIP codes are superior to counties and metropolitan statistical areas for our purposes, because the latter two designations more frequently include multiple study hospitals given their larger size and would in turn require us to drop more hospitals, a concern Dr. Deb states in his subsequent point. Finally, since patient exposure to hospitals is greater the closer they live, and hospitals that disproportionately serve low-income populations are often located in low-income neighborhoods, the immediate ZIP code around a hospital should be more sensitive than larger geographic markets to hospital efforts to expand access and engage in community outreach.

The authors restrict the sample to hospitals that are unique within zip code. They state that this restriction still covers 75% of hospitals. But it means that they drop 25% of hospitals in what is already a restricted sample. Hospitals with nearby competitors might behave quite differently than those without. If that is true, their inclusion could easily change results substantively.

Because the analysis requires us to assign each area to a single hospital and DSH percentage, we cannot cleanly examine the 25% of hospitals with another study hospital in the same ZIP code. However, economic theory and past literature[9] suggests that greater competition elicits less provision of community benefit (measured in the form of charity care) by hospitals, not more. Therefore, we believe inclusion of the 25% of hospitals facing greater local competition would strengthen the conclusion that the 340B Program is not leading hospitals to reinvest 340B surplus in such a way to improve mortality rates among Medicare beneficiaries in their local service area.

Using Medicare claims, I do not understand how one can show impact on low-income patients. Those patients would be enrolled in Medicaid or would be uninsured. This patient level analysis cannot demonstrate whether their care changed in any way. The authors acknowledge this "limitation" in the discussion but it is buried deep at the end of the discussion section.

Please see our response to this comment above. The use of Medicare data did not prohibit us from examining effects on low-income populations. And far from "buried", discussion of this issue is at the outset of our limitations section, with an entire paragraph devoted to it:

> *"Our study had several limitations. First, we relied largely on Medicare data. We would expect, however, that major investments in clinical resources for low-income groups outside of Medicare, such as the uninsured, would also affect care for low-income Medicare beneficiaries. Moreover, we found no evidence of enhanced care for a subgroup of Medicare patients with supplemental insurance that is less generous or reimburses hospitals at lower rates than private supplemental insurance. In addition, measures of hospital investments in FQHCs were not specific to Medicare."*

The authors begin their discussion of findings by stating that "… hospitals that are eligible for the 340B Drug Pricing Program have responded to program incentives by …". This is false, in a strict sense, and more broadly misleading. In fact, their analysis leaves out acute care hospitals with fewer than 50 beds, those with DSH percentages greater than 21.75%, hospitals that are very close to the threshold and several other categories of hospitals like critical access hospitals, sole community hospitals, rural referral centers, pediatric hospitals, and free-standing cancer centers. So, clearly the scope of the interpretation should be limited to that group of hospitals. In fact, to be more precise, because a regression discontinuity design is used, interpretation should be focused on hospitals that are close to the threshold, e.g., "short-term general hospitals close to the threshold have responded to program incentives by …" To be fair, the authors discuss this and other limitations (limitations paragraphs on p. 9) but these are more than just technical limitations. These speak to policy implications that might be drawn from the work.

Please see our response to this comment above. In particular, the marginal effects we estimate are of significant policy interest because of the large number of hospitals at that margin and because the eligibility threshold is a policy lever that could be used in 340B reform efforts.

It is the style of medical journals to describe the limitations in a limitations section of the Discussion. As noted above and by Dr. Deb, we did that. If that style is not agreeable to some readers, there is little we can do about that as authors, and we would rather focus on the substance of our study.

I found the percentage change interpretations in the paper to be very misleading. For example, going from 1 hematologist-oncologist per hospital to 2.3 hematologist-oncologists per hospital is an increase of 1.3 hematologists-oncologists per hospital. Although it can also be framed as a 230% increase, it clearly makes it feel like it is enormously large and with potentially dire consequences.

The characterization of the increases by Dr. Deb is not quite right. The regression coefficient should not be subtracted by the mean to yield the marginal effect as suggested. Instead, the discontinuity estimate of 2.3 additional hematologist-oncologists was correctly interpreted as "340B eligibility was associated with 2.3 *additional* hematologist-oncologists practicing in a hospital's owned outpatient facilities, or 230% *more* than expected in the absence of the Program". The coefficient (2.3) was divided by the expected mean of the outcome (the counterfactual equal to 1.0) and multiplied by 100 to yield a relative effect in percentage term (230%). We presented both absolute and relative percentage effects because both are valuable for interpretation. In particular, because our analysis is conducted on a 20% sample of Medicare FFS claims, and because many hospitals lacked any physicians practicing in a given specialty (as noted in the table notes), the absolute effects are misleadingly small. Thus, presenting relative effect sizes in addition to absolute effect sizes provides greater balance, allowing readers to appreciate how ostensibly small effects are actually substantial shifts from the status quo. For outcomes measured in units or dollars, an approximation of the full absolute effect for the hospital and expected means for outcomes can be calculated by multiplying coefficients and expected means by 5.

References

1. Desai S, McWilliams JM. Consequences of the 340B Drug Pricing Program. *N Engl J Med*. 2018;378:539-548. doi:10.1056/NEJMsa1706475.
2. Conti RM, Bach PB. The 340B Drug Discount Program: Hospitals Generate Profits By Expanding To Reach More Affluent Communities. *Health Aff*. 2014;33(10):1786-1792. doi:10.1377/hlthaff.2014.0540.
3. Jung J, Xu WY, Kalidindi Y. Impact of the 340B Drug Pricing Program on Cancer Care Site and Spending in Medicare. *Health Serv Res*. 2018:1-21. doi:10.1111/1475-6773.12823.
4. Alpert A, Hsi H, Jacobson M. Evaluating the role of payment policy in driving vertical integration in the oncology market. *Health Aff*. 2017;36(4):680-688. doi:10.1377/hlthaff.2016.0830.
5. Potetz, Lisa; Cubanski, Juliette; Neuman T. A Primer on Medicare. *Henry J Kaiser Fam Found*. 2015;11(August):11-16. http://kff.org/health-reform/issue-brief/a-primer-on-medicare-financing/.
6. Young K, Garfield R, Musumeci M, Clemans-Cope L, Lawton E. Medicaid's Role for Dual-Eligible Beneficiaries. *Henry J Kaiser Fam Found*. 2013;(August):1-19. http://kff.org/medicaid/issue-brief/medicaids-role-for-dual-eligible-beneficiaries/.
7. Barreca AI, Lindo JM, Waddell GR. Heaping-induced bias in regression-discontinuity designs. *Econ Inq*. 2016;54(1):268-293. doi:10.1111/ecin.12225.
8. Barreca AI, Lindo JM, Waddell GR. *Heaping-Induced Bias in Regression-Discontinuity Designs.*; 2011.
9. Sloan FA. Chapter 21 Not-for-profit ownership and hospital behavior. *Handb Heal Econ*. 2000;1(PART B):1141-1174. doi:10.1016/S1574-0064(00)80034-7.