

# The State of Open Government in the U.S.

A Comparative Empirical Analysis of U.S. Performance under NAP3 Relative to U.S.-NAP2  
and OECD Peers

*Jason I. McMann, PhD*

## Abstract

This paper uses the Open Government Partnership’s (OGP) assessment methodology and difference-in-means tests to conduct a comparative empirical analysis of U.S. performance under the federal government’s third National Action Plan (NAP3) relative to its own performance under the previous action plan (NAP2), and relative to a group of OECD peer countries. The paper yields four key findings. (1) On average, U.S. performance under NAP3 is statistically indistinguishable from its performance under NAP2. (2) On average, U.S. performance under NAP3 is statistically indistinguishable from that of its OECD peers, with the exception of minor differences in mid-term completion rates. (3) On average, the U.S. under NAP3 advanced commitments with low potential and actual impact that only marginally opened government. U.S. performance in this regard is nevertheless statistically indistinguishable from its own performance under NAP2 and that of its OECD peers, reflecting a high historical prevalence of low-impact action plans, both within the U.S. and globally. (4) For the first time in its assessment history, U.S. performance under NAP3 resulted in a closure of government, as reflected in a ‘Did-It-Open-Government’ score of ‘Worsened’ for two commitments. Collectively, these findings suggest that while the U.S. performs well on commitment completion (in both absolute and relative terms), it continues to advance low-impact commitments that only marginally shift the needle toward more open government. U.S. backsliding and the delayed release of NAP4 cast doubt on the U.S.’ commitment to creating a more open government.

**Author Bio:** *Dr. Jason I. McMann served as National Researcher for the United States at the Open Government Partnership’s Independent Reporting Mechanism (IRM) during the latter half of the 2015-2017 reporting cycle. He led the IRM’s end-of-term assessment of U.S. performance under the third National Action Plan (NAP3) and authored the corresponding [End-of-Term Report](#); he also co-authored the NAP3 [Mid-Term Assessment Report](#). Dr. McMann obtained his PhD from Princeton University’s Politics Department in 2016. He can be reached at [jason.ian.mcmann@gmail.com](mailto:jason.ian.mcmann@gmail.com)*

**Disclaimer:** *The views and opinions presented here reflect those of the author and do not necessarily reflect those of the Open Government Partnership or any agency of the U.S. government.*

## Introduction

In August 2018, the Open Government Partnership’s Independent Reporting Mechanism (OGP/IRM) released its end-of-term assessment of U.S. performance under the federal government’s third National Action Plan (NAP3), covering the July 2015 – June 2017 reporting period. As the OGP’s National Researcher for the United States, I led the assessment and authored the corresponding *End-of-Term Report (EOTR)*.

Shortly after the report’s release, the OGP made available an updated version of its global [Commitment Dataset](#). The dataset tracks progress made toward completing OGP commitments, as well as their potential and actual impact (“Did-it-Open-Government” Scores), across all participating OGP countries.

This paper draws upon the *EOTR* and the OGP’s recently updated Commitment Dataset to carry out a comparative empirical analysis of U.S. performance under NAP3 relative to its own performance under the previous action plan (NAP2), and relative to a group of OECD peers who have participated in the OGP assessment process.<sup>1</sup> The paper presupposes a familiarity with the [OGP assessment methodology](#).

To conduct the analysis, I use difference-in-means tests to measure the difference in average performance along key OGP reporting dimensions, and pair these tests with related descriptive analysis; I highlight comparisons between U.S. performance under NAP2 and the OECD when relevant. The analysis covers 1,113 commitments, distributed as follows: 52 from U.S.-NAP3, 26 from U.S.-NAP2, and 1,035 from the OECD peer group. Table 1 below provides a full summary of the results from the difference-in-means tests.<sup>2</sup>

The analysis that follows is structured around four key findings. The Methodology Note at the end of this paper provides additional information on test specifications, as well as links to replication materials.

**Table 1: Results of Difference-in-Means Tests.**

Quantity of Interest	U.S. NAP: Mean	OECD: Mean	Diff (U.S. - OECD)	P-Value
<b>U.S. Performance on NAP3 Relative to OECD</b>				
Completion Rate: Mid-Term	1.38	1.73	-0.35	0.0073***
Completion Rate: End-of-Term	1.96	2.12	-0.16	0.152
Potential Impact	1.62	1.56	0.06	0.657
"Did-It-Open Government" Score	0.85	0.87	-0.02	0.767
<b>U.S. Performance on NAP2 Relative to OECD</b>				
Completion Rate: Mid-Term	1.77	1.73	0.04	0.741
Completion Rate: End-of-Term	1.96	2.12	-0.16	0.165
Potential Impact	1.69	1.56	0.13	0.324
"Did-It-Open Government" Score	0.96	0.87	0.09	0.477
<b>U.S. Performance on NAP3 versus NAP2</b>				
Completion Rate: Mid-Term	1.38	1.77	-0.39	0.0209**
Completion Rate: End-of-Term	1.96	1.96	0.00	n/a
Potential Impact	1.62	1.69	-0.07	0.572
"Did-It-Open Government" Score	0.85	0.96	-0.11	0.408

*Note:* \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

<sup>1</sup>The OECD peer group includes: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, South Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the U.S. The U.S. is excluded from the OECD peer group in the empirical analysis.

<sup>2</sup>The figures presented throughout this paper denote statistical significance as described in the note accompanying Table 1.

## Empirical Analysis

**Key Finding 1: On average, U.S. performance under NAP3 is statistically indistinguishable from its performance under NAP2 along key OGP reporting dimensions.**

Comparing U.S. performance across NAP3 and NAP2 offers an opportunity to assess variation in U.S. performance over time, as well as across administrations, with President Trump assuming office partway through the NAP3 evaluation period.

With respect to average completion rates, potential impact scores, and “Did It Open Government” (DIOG) scores, U.S. performance is statistically indistinguishable across NAP3 and NAP2, reflecting strong continuity in performance across action plans and administrations.

Figures 1-3 provide a visual representation of these comparisons in the form of violin plots, with jittered points representing individual commitments and dashed horizontal lines indicating mean performance under NAP3 and NAP2. Across all three variables of interest – average completion levels, potential impact, and DIOG scores – the difference in means is statistically insignificant, suggesting that no meaningful difference in performance exists across action plans.

Figure 1: U.S. Completion Status at End-of-Term: NAP2 vs. NAP3.

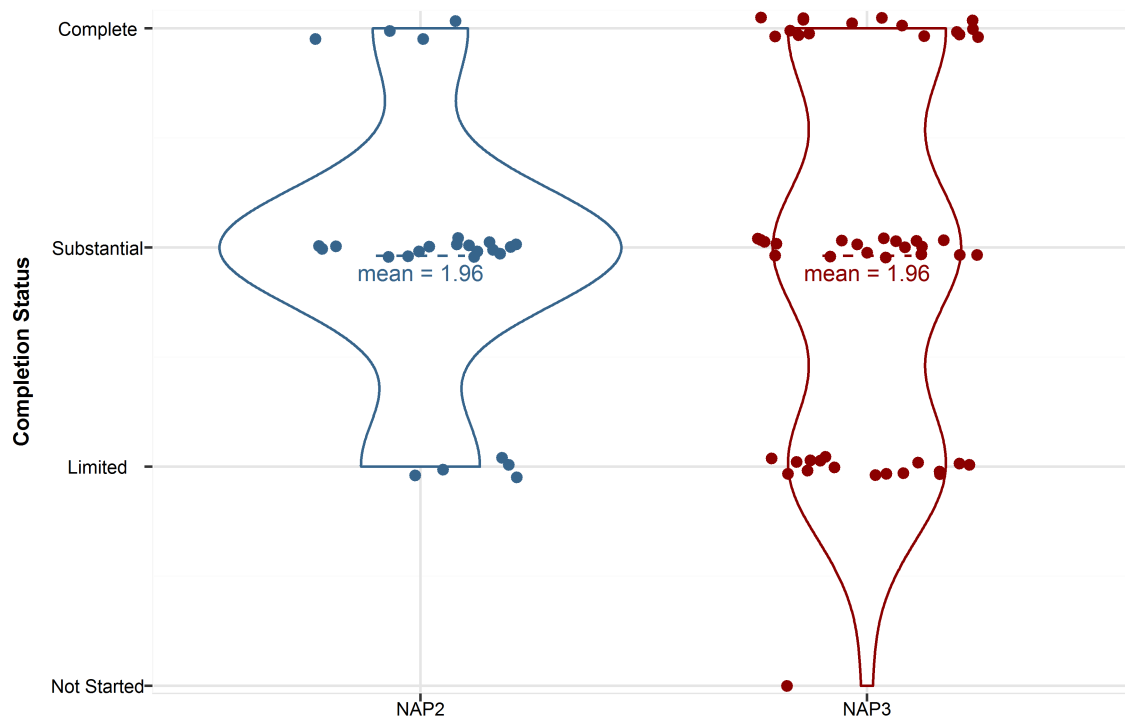


Figure 2: U.S. Potential Impact Scores: NAP2 vs. NAP3.

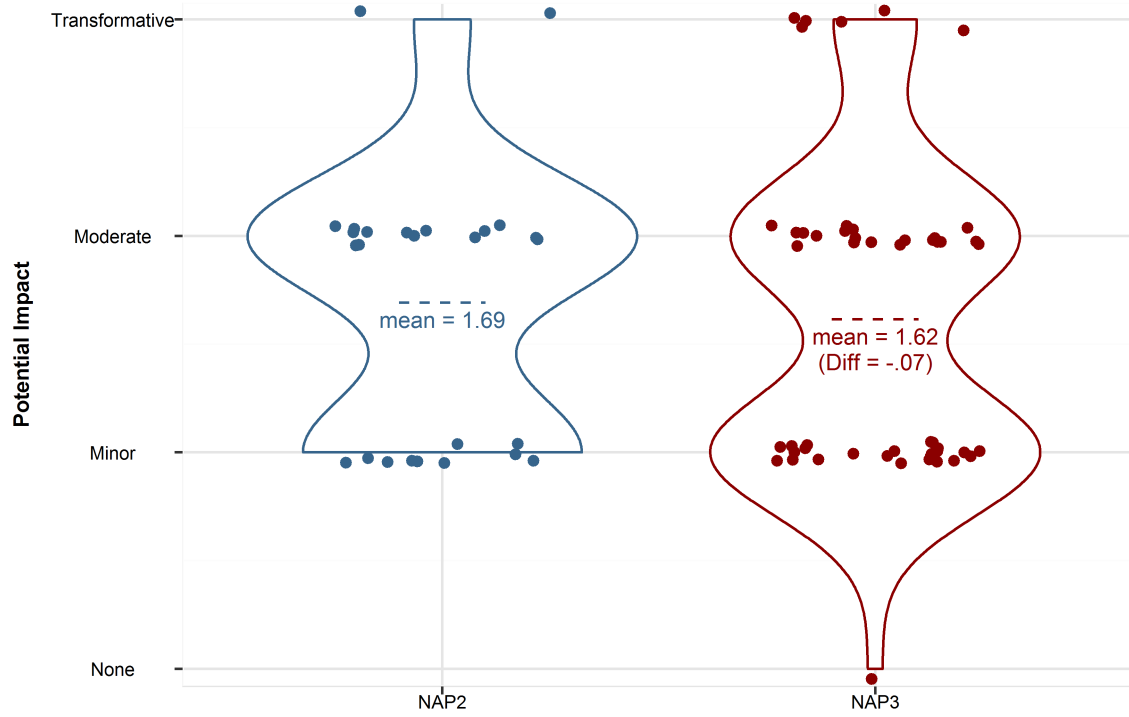
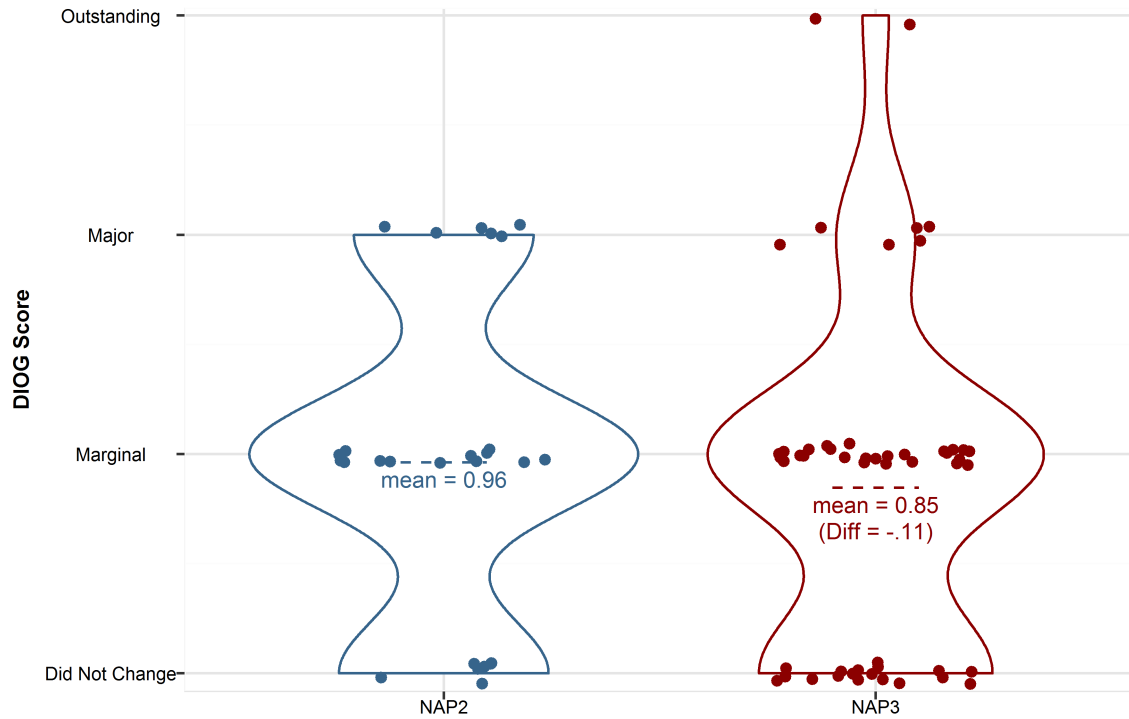


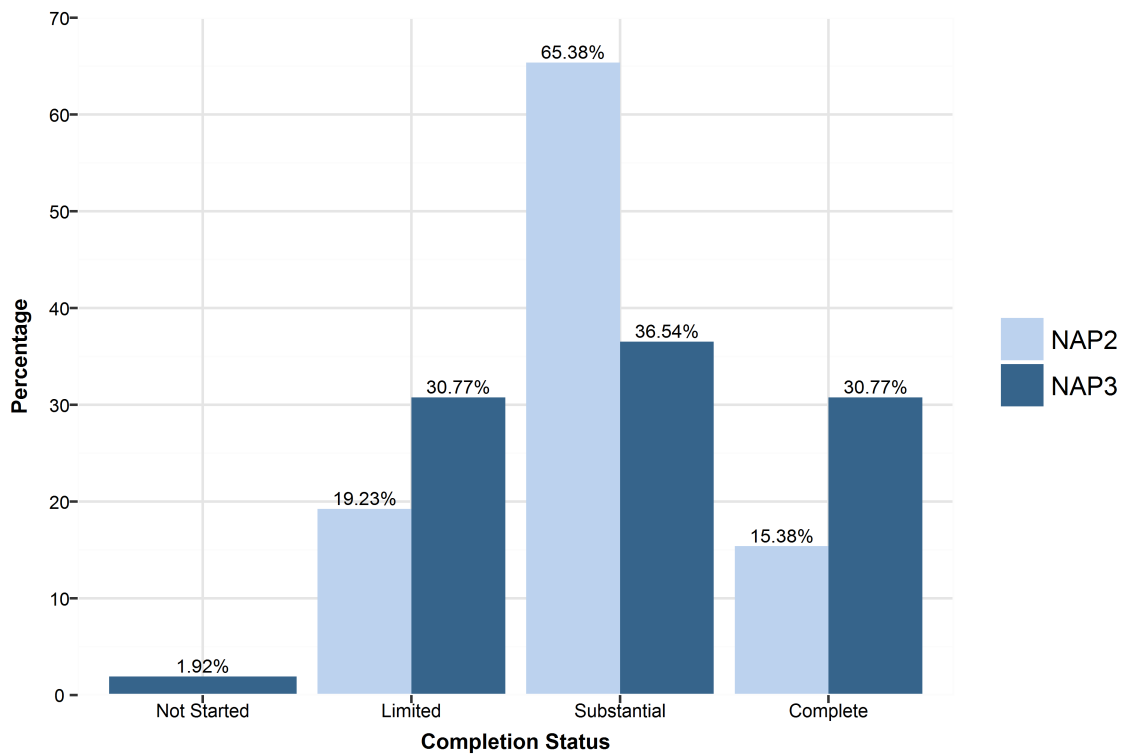
Figure 3: U.S. Did-It-Open-Government Scores: NAP2 vs. NAP3.



Along all three dimensions, the lack of a statistically meaningful difference at the mean nevertheless masks distributional variation in performance that is worth briefly reviewing here.

To that end, Figure 4 depicts distributional variation in end-of-term completion rates across NAP3 and NAP2. As can be seen in the figure, U.S. performance under NAP2 reflects a far higher percentage of substantially complete commitments relative to NAP3 (65.38% versus 36.54%, respectively). Under the latter plan, the U.S. has effectively split the difference: roughly half of the relative decline in the percentage of substantially complete commitments across NAP2 and NAP3 is reflected in a higher share of complete commitments under NAP3 (30.77% under NAP3 relative to 15.38% under NAP2); the other half is reflected in a higher share of commitments with limited completion (30.77%), as well as one commitment on which no progress was made by the end-of-term (Commitment #6).

Figure 4: U.S. Completion Status at End-of-Term: NAP2 vs. NAP3.



As Figures 5 and 6 makes clear, virtually identical patterns in U.S. performance are visible across NAP3 and NAP2 with respect to commitments’ potential impact and DIOG scores. With respect to potential impact, NAP3 reflects a simultaneous distributional shift away from commitments with moderate potential impact and toward commitments with more transformative potential impact (on the positive side of the ledger), but also toward commitments with minor and no potential impact (on the negative side of the ledger). With respect to DIOG scores, NAP3 exhibits a similar simultaneous shift away from commitments that opened government in a major way, and toward commitments with both outstanding impact and no clear impact. In two particularly troubling cases, U.S. actions taken under NAP3 resulted in government closure, as reflected in a DIOG score of “Worsened” for two commitments and discussed in further detail on the following pages.

Figure 5: U.S. Potential Impact Scores: NAP2 vs. NAP3.

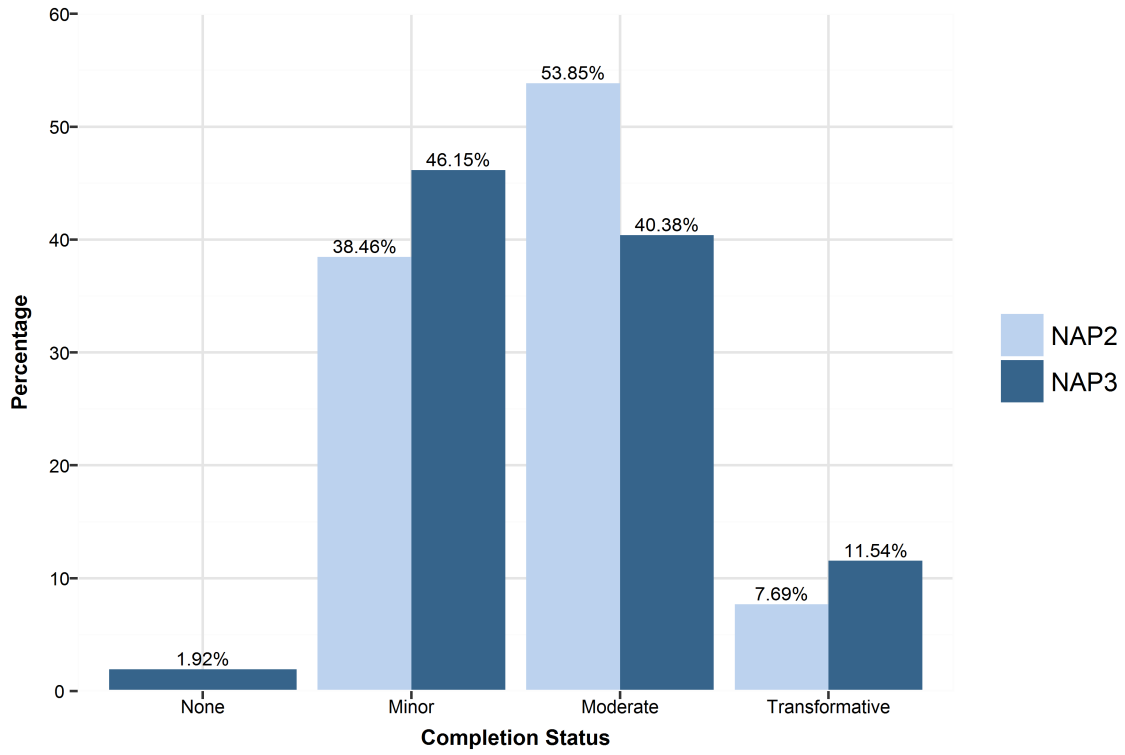
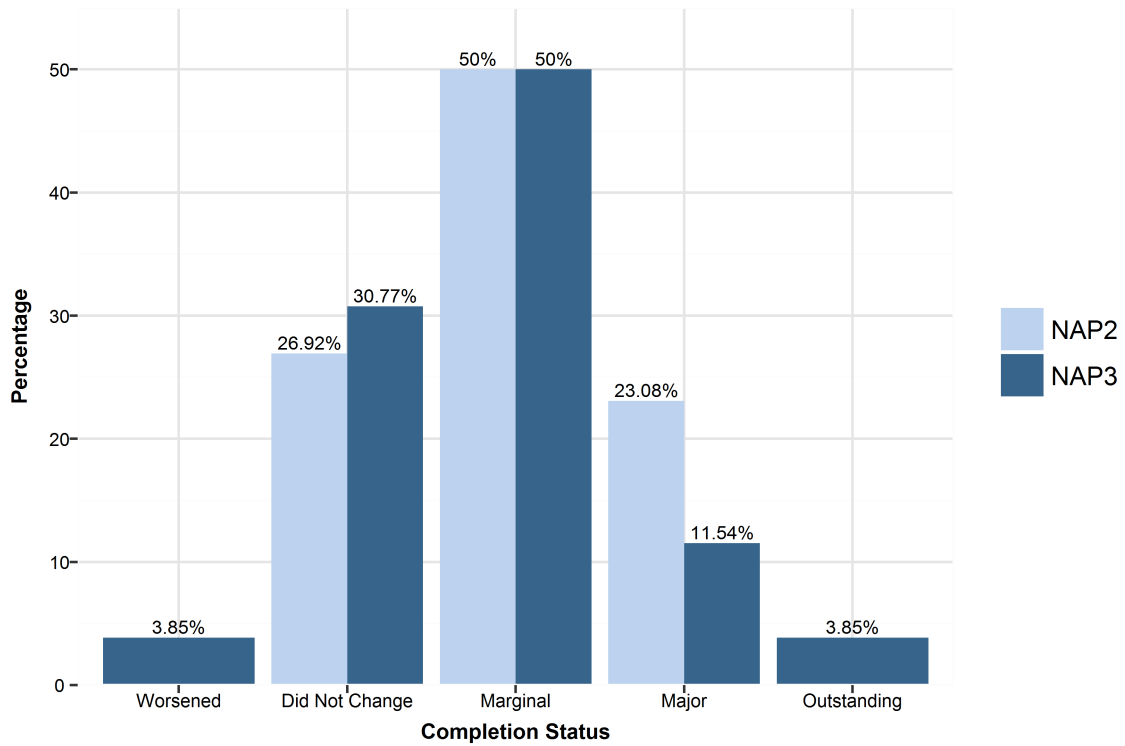
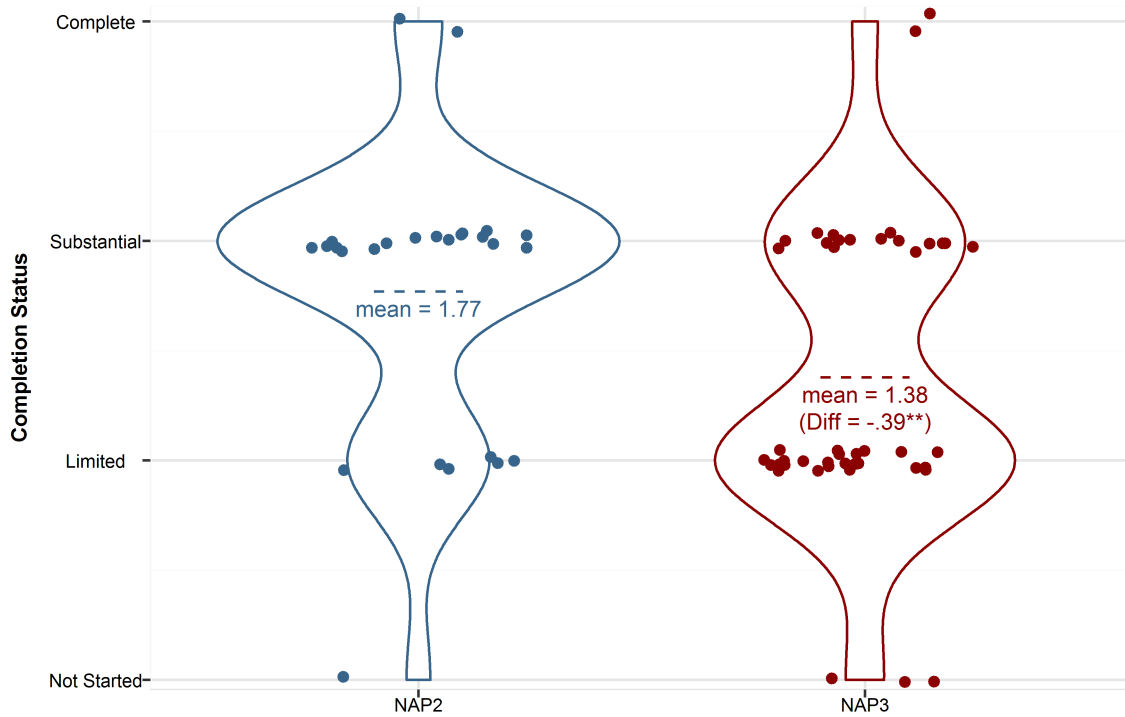


Figure 6: U.S. Did-It-Open-Government Scores: NAP2 vs. NAP3.



Importantly, the U.S. presidential transition – which occurred partway through the NAP3 evaluation period – did not adversely affect the average commitment completion rate relative to NAP2, as reflected in identical average completion rates at the end-of-term across action plans. On the contrary, as can be seen by comparing average completion rates at mid-term (Figure 7 below) and at the end-of-term (Figure 1 above), the mean completion rate for NAP3 at mid-term (halfway through the reporting period in June 2016) was in fact lower than the mean mid-term completion rate for NAP2, with statistically distinguishable means of 1.38 and 1.77, respectively. Given that we observe identical mean completion rates at the end-of-term, U.S. completion progress under NAP3 actually accelerated post-midterm. While some progress was made during the tail months of the Obama administration, progress continued following President Trump’s inauguration in January 2017. Though narrow in scope, this finding stands in partial contrast to [recent work](#) (most notably by the [Sunlight Foundation](#)) which highlights declining enthusiasm for open government initiatives under the Trump Administration.

Figure 7: U.S. Completion Status at Mid-Term: NAP2 vs. NAP3.



**Key Finding 2:** On average, U.S. performance under NAP3 is statistically indistinguishable from that of its OECD peers along key OGP reporting dimensions, with the exception of minor differences in mid-term completion rates. The distribution of U.S. and OECD performance across response categories is also similar. These same findings roughly hold for comparisons of U.S. performance under NAP2 to that of its OECD peers, reflecting relative continuity in average U.S. performance.

Comparing U.S. performance under NAP3 to performance among a group of OECD peer countries offers another lens through which to assess U.S. performance. While the OGP has many non-OECD countries as members, I focus here on the U.S.-OECD comparison based on the assumption that OECD countries have similar capacity to implement open government initiatives, and therefore comprise the most relevant peer comparison group.

Figure 8: Completion Status at End-of-Term: U.S. NAP3 vs. OECD Average.

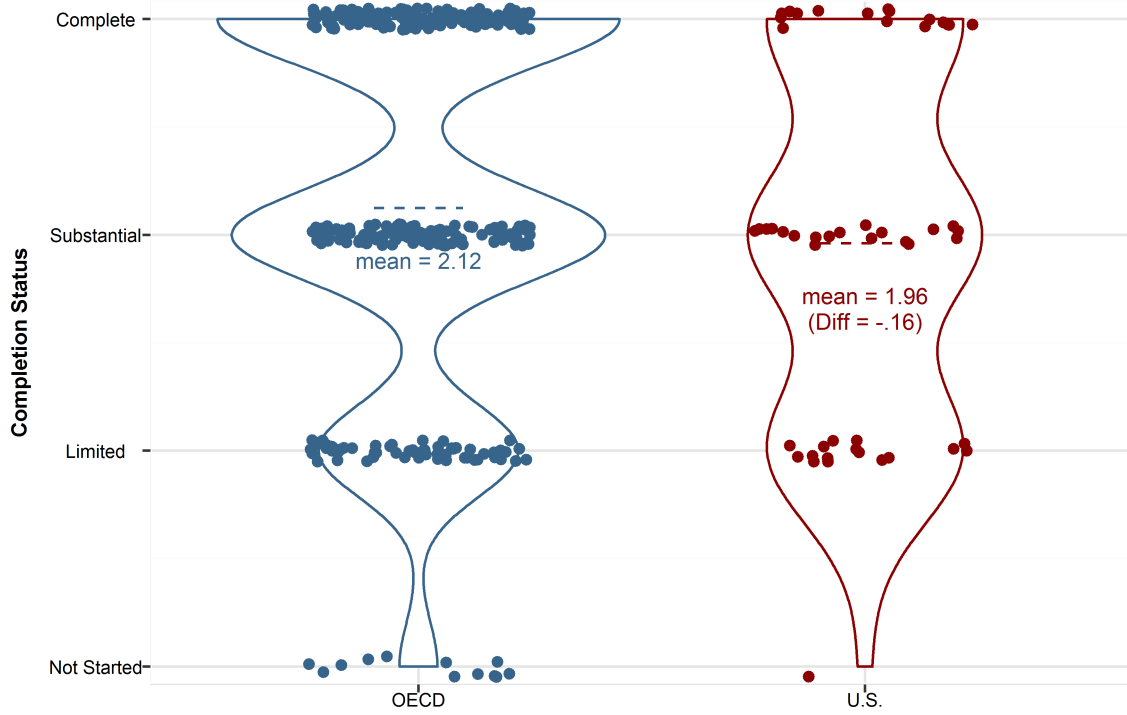


Figure 9: Potential Impact Scores: U.S. NAP3 vs. OECD Average.

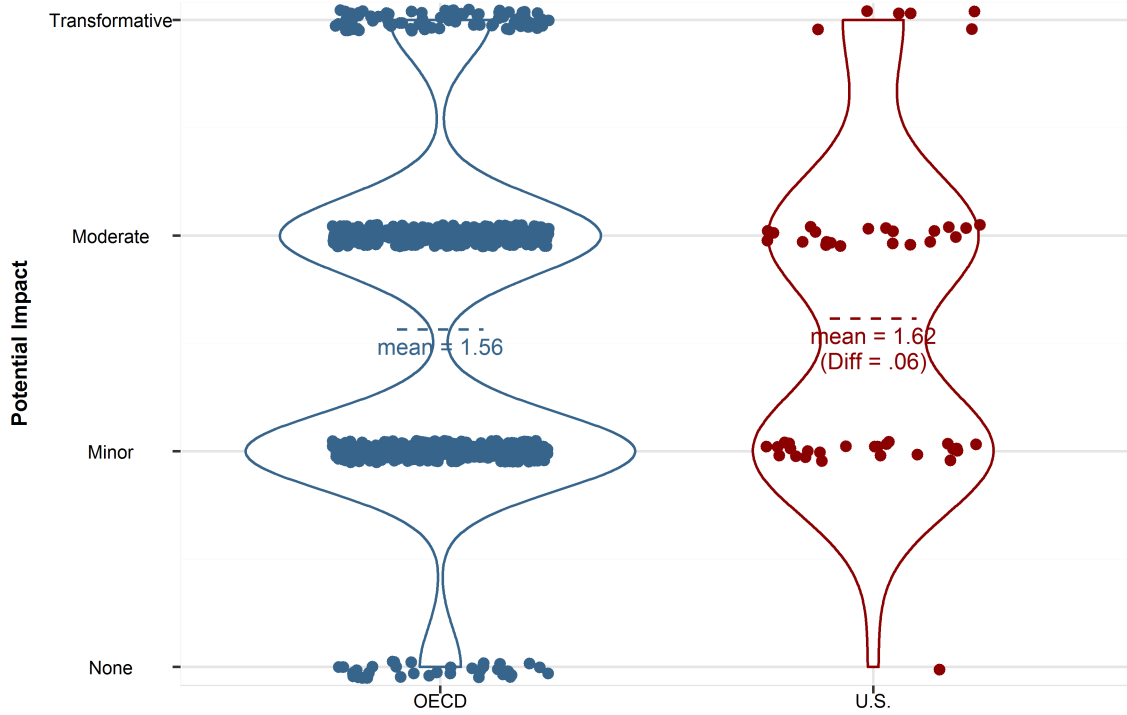
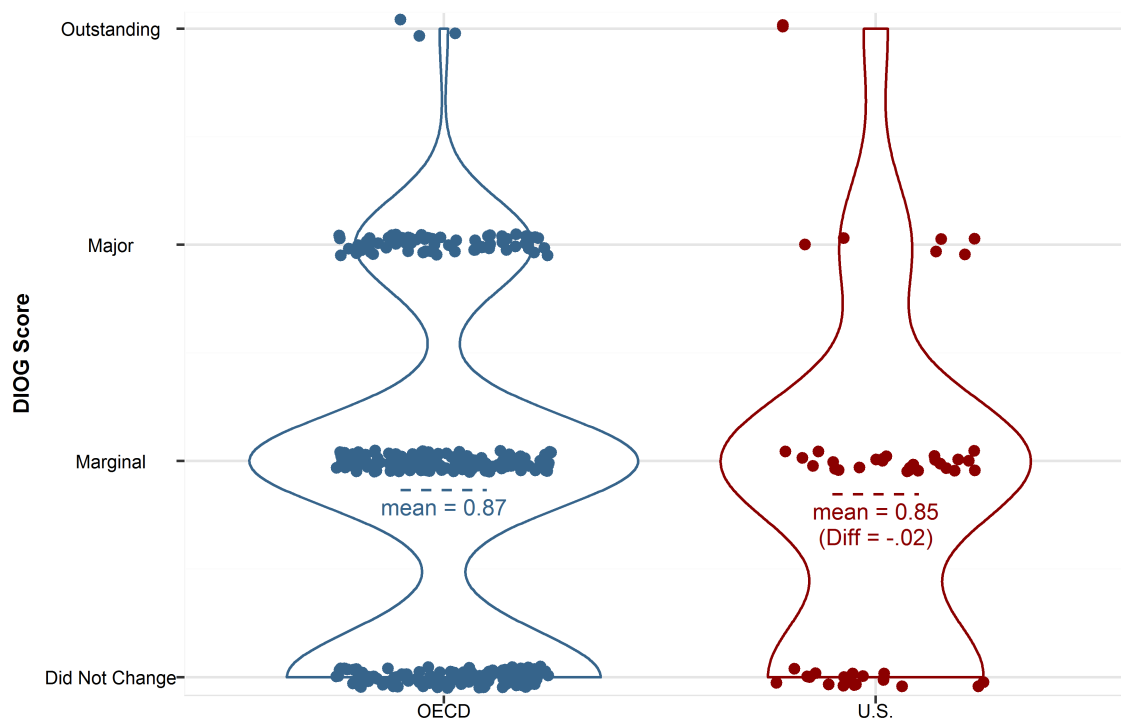




Figure 10: Did-It-Open Government Scores: U.S. NAP3 vs. OECD Average.



As can be seen in Figures 8-10 (which mirror the analysis in the previous section), U.S. performance under NAP3 is statistically indistinguishable from that of its OECD peers with respect to average commitment completion rates, potential impact, and DIOG scores at the end-of-term. The U.S. under NAP3 and its OECD peers also exhibit similar end-of-term performance in a distributional sense: as the violin plots above make clear, there is relatively minor variation in the distribution of completion rates, potential impact, and DIOG scores around the mean.

Per several additional figures included in the Appendix (omitted here in the interest of space), these findings similarly hold for NAP2, with a few minor exceptions: (1) the OECD group has several commitments that remain un-started at the end-of-term, whereas the U.S. under NAP2 has none; (2) the OECD group has several commitments that exhibit no potential impact, whereas the U.S. under NAP2 has none in that category; and (3) several OECD commitments have “Outstanding” DIOG scores, whereas the U.S. under NAP2 has no commitments in that category.

The main area of divergent performance across NAP3 and NAP2 relative to the OECD peer group – as made clear by Figure 11 – is a difference in mean completion rates at mid-term under NAP3, reflecting U.S. and OECD means of 1.38 and 1.73, respectively. By contrast, as Figure 12 makes clear, no such difference existed for NAP2. With respect to NAP3, the fact that no such gap persists at the end-of-term indicates an acceleration of U.S. performance post-midterm, as highlighted in the previous section.

Figure 11: Completion Status at Mid-Term: U.S. NAP3 vs. OECD Average

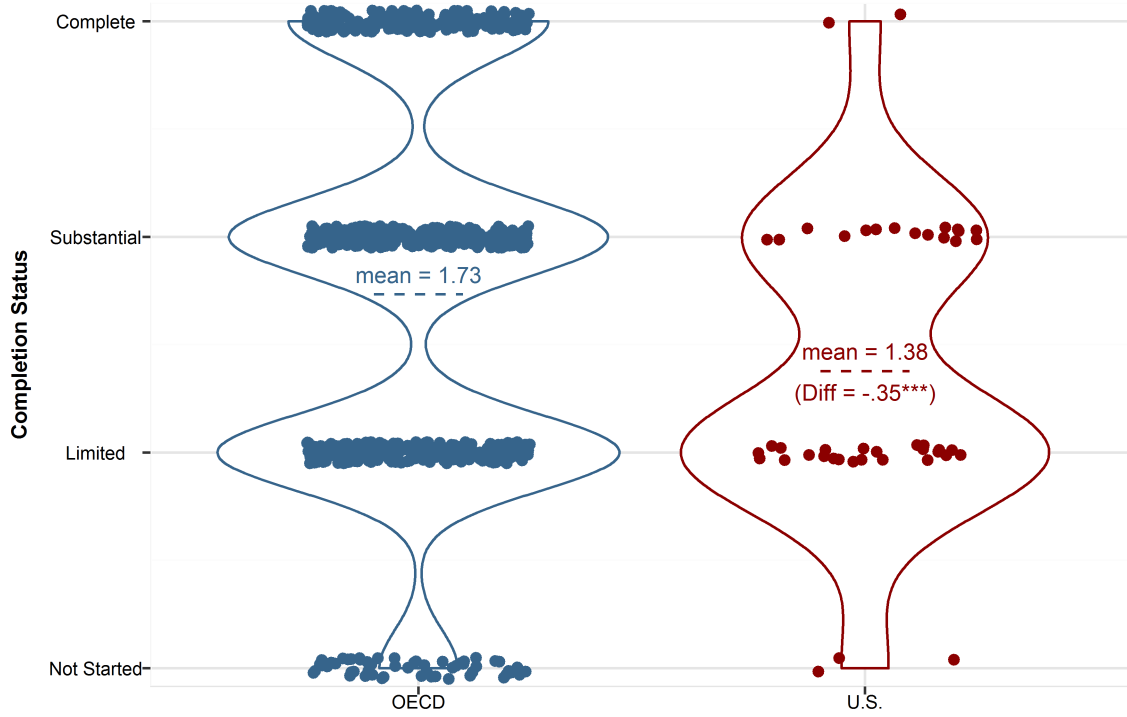
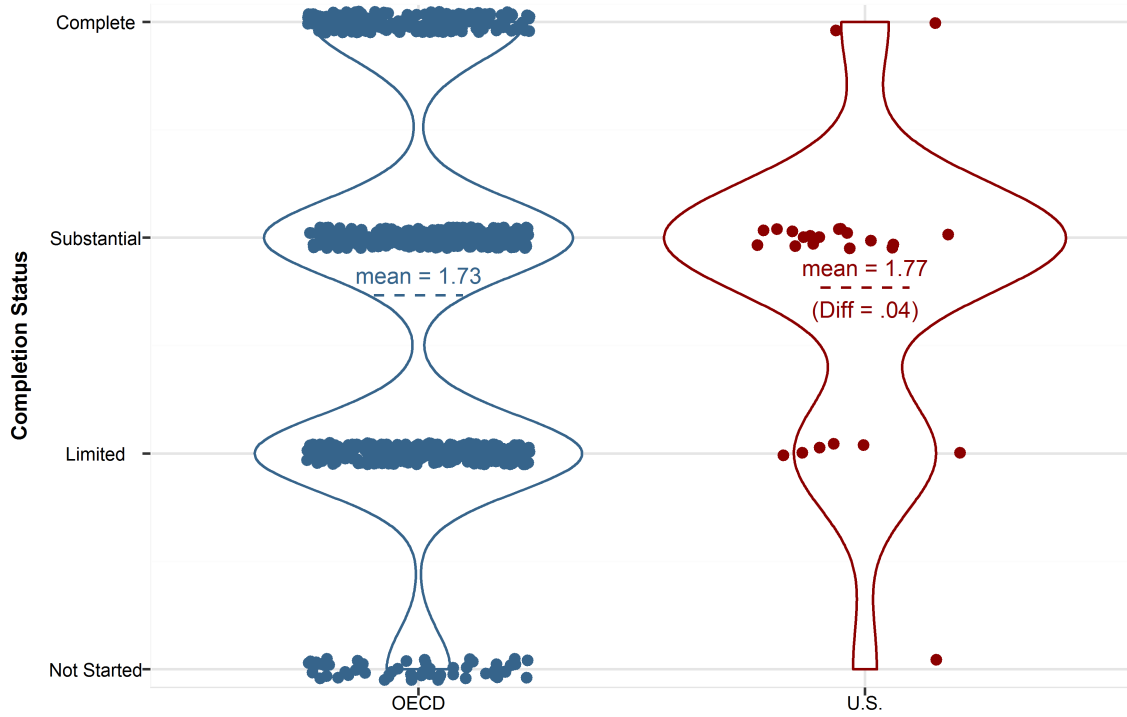


Figure 12: Completion Status at Mid-Term: U.S. NAP2 vs. OECD Average



Collectively, these findings indicate that U.S. performance continues to remain in line with that of its OECD peers. As the figures above make clear, while the U.S. and its OECD peers perform quite well on commitment completion, reflecting an average completion rate of “Substantial,” they nevertheless perform poorly along other dimensions. I elaborate on this point below.

***Key Finding 3:*** On average, the U.S. under NAP3 advanced commitments with low potential and actual impact that only marginally opened government. With respect to both potential and actual impact (proxied by “Did-it-Open-Government” scores), U.S. performance under NAP3 is nevertheless statistically indistinguishable from its own performance under NAP2 and that of its OECD peers, reflecting a high historical prevalence of low-impact action plans, both within the U.S. and globally.

High average completion rates are only laudable to the extent that action plans exhibit high average levels of impact, with governments translating potential into actual impact both during and beyond the end-of-term. As alluded to above, the OGP’s “Did-it-Open-Government” scores most closely approximate performance along the latter dimension.<sup>3</sup>

While the U.S. obtains an average score of roughly “Substantial” for commitment completion at the end-of-term under NAP3, its average DIOG score hovers just below “Marginal,” while its average Potential Impact score hovers between “Minor” and “Moderate”, as shown in Figures 9 and 10 above (respectively). Taken together, these figures suggest that under NAP3, the U.S. has advanced a low-impact action plan whose actual impact on open government at the end-of-term is lower than its assessed potential.

Comparisons of U.S. performance under NAP3 relative to both NAP2 and its OECD peers offer additional insight into U.S. performance, with both positive and negative connotations.

In particular, with respect to both average Potential Impact and DIOG scores, U.S. performance under NAP3 is statistically indistinguishable from its own performance under NAP2 and that of its OECD peers (see comparison in Figures 9 and 10 above). On the positive side of the ledger, while consistency in the U.S.’ own performance over time provides no indication of substantially improved performance, nor does it provide any indication of substantially worse performance on average. Moreover, the lack of a meaningful difference in performance across the U.S. and its OECD peers confirms that the U.S. is not an outlier among countries that are presumed to have similar implementation capacity. The same comparison holds for U.S. performance under NAP2 (see corresponding figures in the Appendix), further demonstrating the consistency of U.S. performance relative to its OECD peers.

On the negative side of the ledger, relatively poor performance by both the U.S. and its OECD peers with respect to potential and actual impact reflects a relatively high historical prevalence of low-impact action plans, both within the U.S. and globally. This finding calls into question the value of the OGP assessment process as a means of incentivizing governments to pursue high-impact commitments in the direction of more open government. While low-impact commitments may ultimately shift the needle in this direction, their prevalence highlights the challenges that both the OGP and its members face in advancing down this path more quickly.

***Key Finding 4:*** For the first time in its assessment history, U.S. performance under NAP3 resulted in a closure of government, reflected in a “Did-It-Open-Government” score of “Worsened” for two commitments.

In addition to the high prevalence of low-impact commitments, actions taken under the Trump administration resulted in a DIOG score of “Worsened” for two NAP3 commitments, indicating a closure of government relative to the status quo. This represents the first time the U.S. has obtained such an assessment.

---

<sup>3</sup>While the OGP is cagey about performing a direct assessment of actual impact at the end-of-term, there is little meaningful distinction between DIOG scores and impact, hence its use here.

The first such commitment (#24, Milestone 1, “Improve Public Participation”) was intended to “increase responsiveness and encourage re-use of *We the People*, an online government-facing petition platform which mandates that the White House respond within 60 days to public petitions that meet a prescribed 100,000 signature threshold. During the latter half of the OGP evaluation period (specifically from January-June 2017), nine petitions met the signature threshold. None received a government response within the prescribed window. This contrasts markedly with an improvement in the government’s average response time from 45 to 34 days during the first half of the evaluation period (spanning July 2015 - June 2016), reflecting a clear case of backsliding. The *EOTR* (see pps.98-101) contains additional details on the scope of backsliding.

The second such commitment (#31, “Transparency of Extractive Industries”) intended for the U.S. to deepen its engagement with the *Extractive Industries Transparency Initiative* (EITI), a prominent international initiative that aims to improve global transparency and accountability in the extractive sector, with a particular emphasis on payments made by extractive companies to governments, and government revenue received from extractive sector activities. Instead, on 2 November 2017 (shortly after the conclusion of the end-of-term evaluation period) the U.S. Department of the Interior published a [letter](#) withdrawing the U.S. from the EITI. As described more fully in the *EOTR* (see pps.123-128), the U.S.’ withdrawal eliminates its obligation to publish EITI reports on sector-specific payments and revenue, as well as its commitment to validate U.S. extractive sector policies and practices against the *EITI Standard*. The U.S.’ decision to withdraw from the EITI was widely panned by civil society organizations engaged in research and advocacy on the extractive sector.

While these commitments represent isolated cases of government backsliding relative to overall U.S. performance across NAP3’s 52 total commitments, they definitively shift U.S. policies and practices toward government closure in direct contravention of the OGP’s foundational *Open Government Declaration*. With respect to timing, both cases of U.S. backsliding occurred under the Trump Administration, which has also substantially delayed the release of NAP4, as reported by [FedScoop](#) in September 2018. Whether this backsliding signals an isolated case of worsening performance or the beginning of a longer-term trend remains to be seen.

## Conclusion

The findings presented above suggest that while the U.S. continues to perform well when it comes to completing commitments in a timely manner – both relative to its own performance under NAP2 and that of its OECD peers – it continues to advance low-impact commitments that only marginally open government. The latter trends are not limited to U.S. performance under NAP3; rather, they coincide with a broader global prevalence of low-impact action plans among the OGP’s OECD member-countries. U.S. backsliding on two commitments under NAP3 and the delayed release of NAP4 (which remains pending at the time of writing) jointly cast forward-looking doubt on the U.S.’ commitment to creating a more open government.

---

## Methodology Note

The econometric analysis for this paper was performed using R in RStudio. The analysis covers 1,113 total commitments, distributed as follows: 52 from U.S.-NAP3, 26 from U.S.-NAP2, and 1,035 from the OECD peer group. The paper does not assess U.S. performance under NAP1 due to a differential coding scheme used during the assessment process at that point in time. Difference-in-means tests were performed using specifications for two-sample comparisons from independent groups. Test specifications were adjusted as necessary to account for non-normal distribution of variables and equal/unequal variances across groups. The p-values and significance levels reported in the tables and figures that accompany this paper correspond to the most appropriately specified test with respect to assessments of normality and variance. Multiple specifications were employed as robustness tests in cases of ambiguity. All U.S.-OECD comparisons exclude the U.S. from the OECD peer group. For the purposes of data visualization, violin plots are used due to the ease of simultaneously displaying grouped variable distributions, performance values for individual commitments,

and group means. Jittered points represent individual commitments. Dashed horizontal lines indicate group means. With respect to coding, all variables assessed in this paper have been rescaled on the [0-3] interval to facilitate cross-indicator comparisons, as described further in the replication code. Commitments that received a DIOG score of “Worsened” are coded as 0 along that dimension, and are therefore subsumed by the “Did Not Change” category in violin plots indicating the distribution of DIOG scores. These cases are coded as 0 because an OGP assessment of “Worsened” does not include a corresponding assessment of degree, and therefore cannot be used to meaningfully assign a negative DIOG score. Fully annotated replication code is available [here](#). The version of the OGP Commitment dataset used in this analysis is available [here](#).

# Appendix

Figure A-1: Completion Status at End-of-Term: U.S. NAP2 vs. OECD Average.

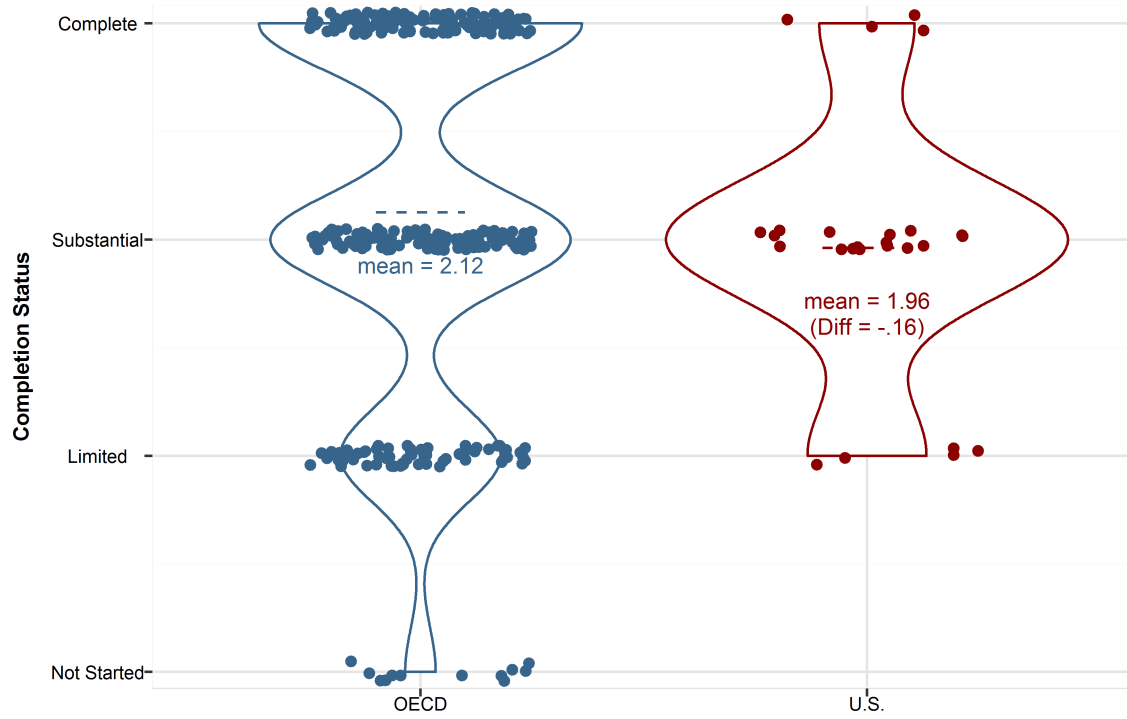


Figure A-2: Potential Impact Scores: U.S. NAP2 vs. OECD Average.

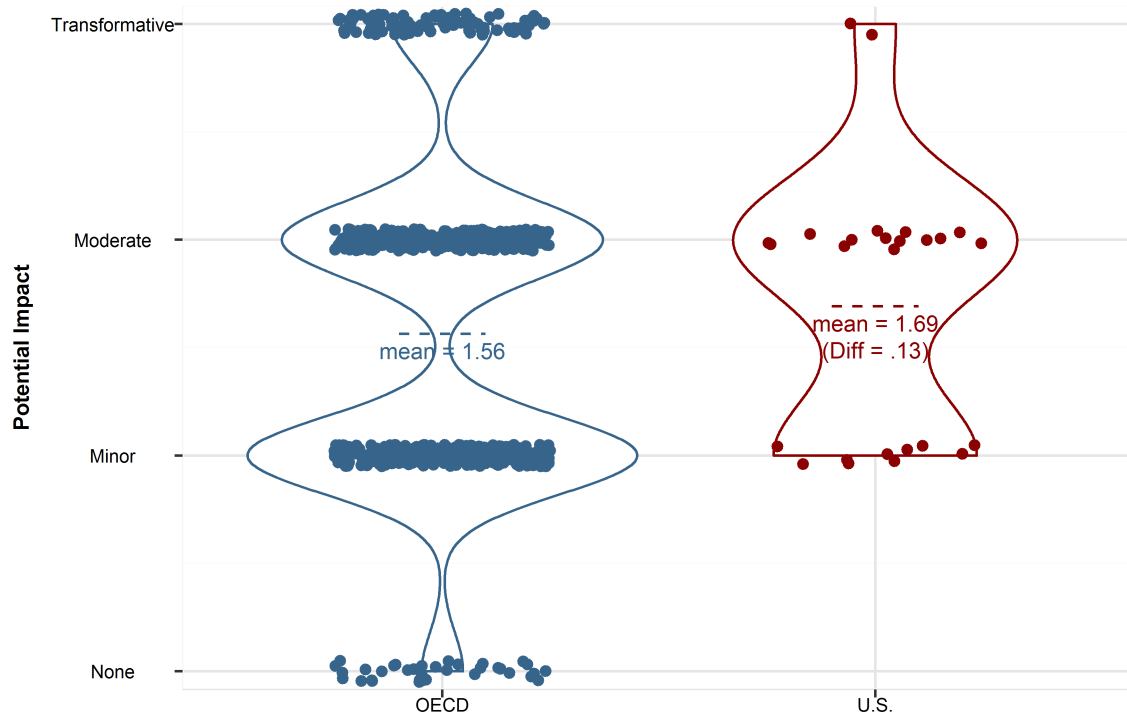


Figure A-3: Did-It-Open-Government Scores: U.S. NAP2 vs. OECD Average.

