Chesapeake Governance Study: Report of 2021 Decision-Maker Interview Results

D.G. Webster, Dartmouth College

Abstract: This report describes the aggregate results from a series of interviews conducted with decision makers involved in governance of the Chesapeake Watershed. Interviews began in June and ended in December of 2021. Information collected will be combined with other data to create and then test a computer model to predict likely policy changes under a range of future scenarios. It is part of a larger project funded by the National Science Foundation called Modeling the Dynamics of Human and Estuarine Systems with Regulatory Feedbacks (Award #2009248). Using the Chesapeake Bay as an example, this project will combine the policy model that we are designing with biophysical models to predict how social, economic and policy changes impact water quality, and how changes in water quality influence human behavior and decision-making.

Contents

Acknowledgements

This project was funded as part of the project CNH2-L: Modeling the dynamics of human and estuarine systems with regulatory feedbacks (Award Abstract # 2009248; https://www.nsf.gov/awardsearch/showAward?AWD_ID=2009248). Project colleagues, Gary Shenk, Peter Claggett, and Victoria Coles helped to identify respondents for these interviews. Gary also provided information on the phases of the Watershed Model. Patrick Bitterman provided advice and assistance with data analysis and reporting. Dartmouth undergraduates Ellen O'Rourke and Paridhi Kapadia assisted in scheduling interviews, researching respondent backgrounds, and taking notes during interviews. Maria Angelino checked coding in Atlas.ti and reviewed the final report for consistency. I'm also grateful to all 59 respondents who took the time to share their perspectives on the CBP policy process with me and the many who helped keep the snowball sample rolling.

List of Symbols

n	Number of blocks of text describing a particular component of the policy process
M or m	Number of "mentions" or times that a code/code group occurs in blocks of text
M*	Maximum number of mentions per respondent
$\bar{m}$	Average number of mentions per block of text about a particular topic
cc	Number of co-occurrences between two codes
ccf	Co-occurrence co-efficient = cc (1, 2) / (m1 + m2 – cc(1,2))

List of Acronyms

BMP	Best Management Practice
CAST	Chesapeake Assessment Scenario Tool
CBC	Chesapeake Bay Commission
CBP	Chesapeake Bay Program
CBW	Chesapeake Bay Watershed
CBPO	Chesapeake Bay Program Office
EPA	Environmental Protection Agency (Region 3 unless specified)
TMDL	Total Maximum Daily Load
WIP	Watershed Implementation Plan
USC	Upper Susquehanna Coalition
USDA	United States Department of Agriculture
CREP	USDA Conservation Reserve Enhancement Program
NEIEN	National Environmental Information Exchange Network
NFWF	National Fish and Wildlife Federation

See also Appendix F: List of Tables and Appendix G: List of Figures.

Executive Summary

This report summarizes data collected as part of the National Science Foundation-Funded Project CNH2-L: Modeling the dynamics of human and estuarine systems with regulatory feedbacks (Award Abstract # 2009248; https://www.nsf.gov/awardsearch/showAward?AWD_ID=2009248). Fifty-nine key participants in the Chesapeake policy process were interviewed between June-December 2021. The primary purpose of these interviews was to collect data to aid with the development of a multi-level computational model of the policy process for nutrient management in the Chesapeake Bay watershed. A secondary purpose was to help guide exploration of potential future scenarios once the policy model is coupled with models of other components of this social-ecological system.

Respondents provided the research team with a wealth of information about all aspects of the policy process for the Chesapeake watershed. In order to make the most of these data we used a systematic approach to code responses (see Appendix A: Detailed Methods). Three main policy-relevant findings emerged from the interviews:

- The CBP has helped to improve Bay water quality since 2010, but progress toward the 2025 TMDL has been stymied by three main factors:
    - Loading amplifiers (e.g., population growth, climate change)
    - Increasing marginal costs of reducing nutrient loads
    - Institutional constraints within and outside the CBP
- Water quality governance in the region is a highly contested process; this is a necessity given legal and sociopolitical constraints on CBP activities, but it has also undermines the legitimacy of the CBP and the computer models that it relies on to guide policy
- Reducing nutrient loading is likely to be more difficult in future, so to improve water quality governance the Chesapeake Bay Program (CBP), jurisdictional leaders, and others will need to address a number of challenges, but they can do this by taking advantage of internal and external opportunities.

The rest of this executive summary provides additional detail on these and other insights from our interview data. We start by describing how respondents perceive the effectiveness of water quality governance in the Chesapeake watershed and how those perceptions interact with perceptions of the legitimacy of the CBP. Then we describe the top challenges and opportunities identified by our respondents and highlight important synergies among key opportunities. Lastly, we review respondents' thoughts about the future of governance in the region, including their concerns about the viability of the CBP post-2025.

Effectiveness and Legitimacy

An overwhelming majority of respondents indicated that water quality governance in the Chesapeake watershed has been partially effective in terms of **goal attainment**. This is not surprising, given that current estimates suggest that the total maximum daily load (TMDL) will not be reached for the Bay by the 2025 deadline. On the other hand, over half of the 58 statements on goal attainment indicated that the CBP had improved water quality and/or that water quality was better than it would have been without the CBP. Evaluation depended heavily on the counterfactual considered by the respondent; those who compared current conditions to pre-2010 conditions tended to evaluate goal attainment

more favorably while those who compared current conditions to the goal for 2025 or related environmental criteria tended to evaluate goal attainment more negatively.

Many respondents also indicated that the CBP is partially effective or ineffective in its **institutional design** and most of these cited procedural failures as limits on goal attainment. Bureaucratic inefficiency, lack of engagement with local-level stakeholders, lack of resources, and politicization of multiple components of the policy process were most often described as limitations on institutional effectiveness. There was not agreement on the strengths and weaknesses of CBP institutions, however. Respondents were particularly divided over the role of science and models and the level of collaboration across the CBP. What is clear, however, is that the majority of respondents viewed CBP governance as a political exercise and they generally do not view politicization as a positive component of the process.

Interestingly, few respondents evaluated the effectiveness of the CBP in terms of **equity**, either among states or at the local level, but of those who did mention it, most evaluated it as partially effective or ineffective. Overall, equity was not mentioned very often in any context, and when it was, it was more likely to refer to concerns over allocation of loading goals among state-level jurisdictions than equity for stakeholders at the local level. This does not mean that equity—or diversity, equity, and inclusion more broadly—was not important to respondents, but rather that few of our respondents viewed equity as a central to the governance process. That is, they might view equity as important in its own right, but if it was mentioned, it was primarily described as a separate goal for the CBP; only a few respondents indicated that greater equity for local stakeholders was an opportunity to improve water quality (goal attainment) or CBP procedures (institutional design).

Concerns about effectiveness in all three guises appeared in our more detailed analysis of the policy process as well, mainly in conjunction with the WIP Design Process and the Modeling Process. In both cases, respondents expressed concerns about the legitimacy of these processes, and some linked factors that reduced legitimacy to factors that reduced effectiveness. Specifically, WIPs were described as "paper processes" or "good on paper but not in practice" 44 times and, of these, about half indicated that this undermined the legitimacy of the entire process. Some respondents indicated that designing WIPs that were "good on paper" was the best that could be done given legal, political, and resource constraints, but none put a more positive spin on this issue. The WIP Design process was defended for specific states 12 times, however, including Pennsylvania, West Virginia, and New York.

Questions about the Chesapeake Assessment Scenario Tool (CAST) or related "Bay Models" and the legitimacy of the modeling process also arose, with 29 of 31 statements about the model indicating that it was not accurate or acceptable and 30 of 31 statements about the legitimacy of the modeling process indicating that it is not considered legitimate. Positive statements about the models or modeling process were much less numerous (< 5), but we did not ask about the modeling process specifically, so the results may be biased towards respondents who are dissatisfied with this aspect of the CBP.

Respondents also indicated that the model was not accurately accounting for changes in nutrient loading, with 33 statements indicating that BMPs are under-credited in CAST, 2 indicating that credits from CAST were about right, and 4 indicating that BMPs are over-credited in CAST. In other words, many more respondents believed that the model is under-estimating the effects that WIPs have had on loads in the Chesapeake Bay (thereby underestimating goal attainment) than felt that the estimates were either accurate or overestimated (thereby overestimating goal attainment). As described below, these questions about accounting using the model were seen as a major challenge for the CBP, but we also

note that model development and specification was largely viewed as a political process, and that this, combined with lack of transparency, generated concerns about the legitimacy of the models and the CBP as a whole among a number of respondents.

Challenges and Opportunities

Respondents identified 42 types of **challenges**, or things that make governance of Chesapeake water quality more difficult, and 21 types of **opportunities**, or things that might make governance of Chesapeake water quality easier. Challenges and opportunities mentioned most often are summarized in Figure E. 1, which also shows how respondents often talked about top challenges and opportunities together. It was clear that most respondents saw water quality governance as a complex problem where challenges were inter-related to each other, meaning that some opportunities were viewed as methods to address multiple problems at the same time. Identifying these synergies is important both for understanding the history of the CBP and anticipating its future.

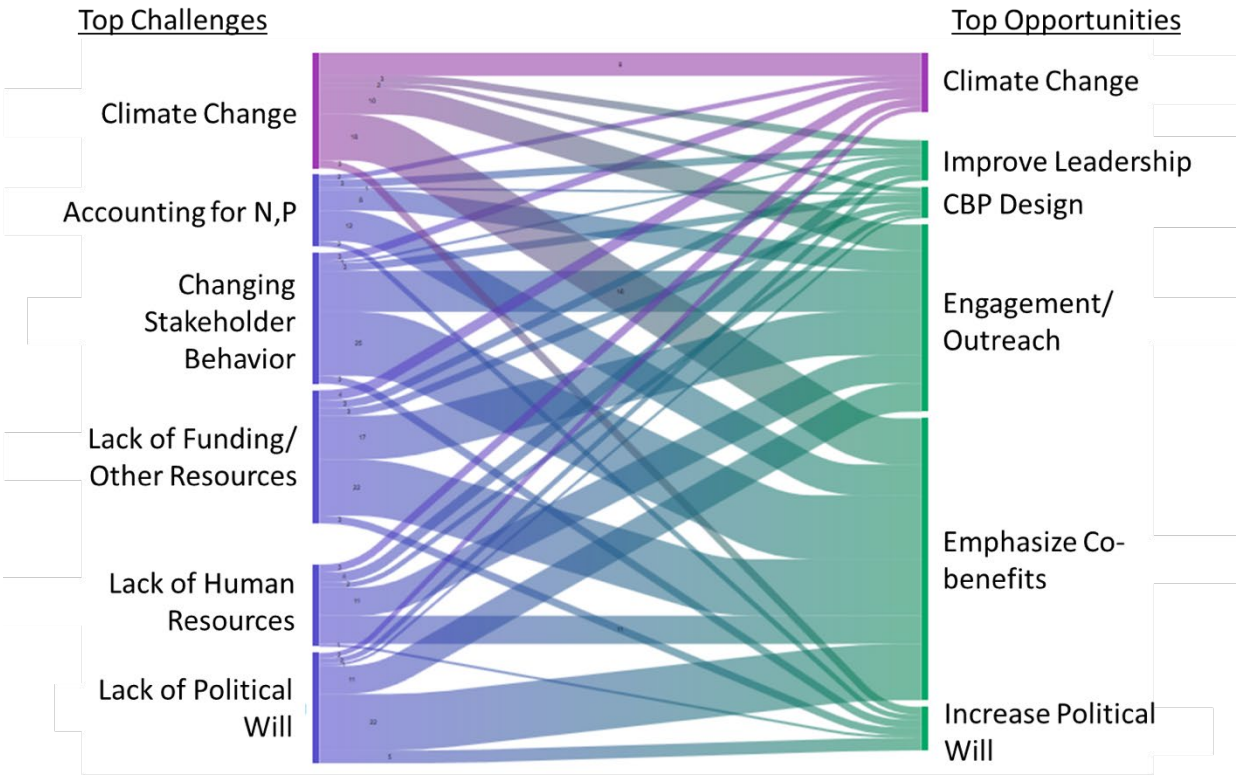


Figure E. 1 Sankey Diagram of the Most Frequently Mentioned Challenges and Opportunities. Width of lines linking Challenges on the right with Opportunities on the left indicates how often these topics were mentioned together, or co-occur, in a given interview. Co-occurrence does not always imply causal connection, but Engagement/Outreach and Emphasize Co-benefits were both frequently mentioned as methods for addressing all challenges listed here except Accounting for N,P.

Climate change was identified as a top challenge, but some respondents also saw it as an opportunity, since it could increase funding to implement BMPs that have climate-related co-benefits such as flood mitigation. Other challenges were related to respondent concerns about the models or modeling process describe above. Difficulties accounting for N and P were primarily associated with concerns about the design and parameterization of CAST and were frequently tied to other challenges or questions about the models used by the CBP. Difficulties of data collection and verification did not make this list of top challenges but were mentioned as a separate challenge 41 times (see Section 4 Challenges

and Opportunities for a more detailed look at these results). The remaining top challenges focused on barriers to BMP implementation, including difficulties changing stakeholder behavior and lack of political will, funding, or human resources. All of these challenges were seen as closely connected to each other and to other challenges that were not mentioned as frequently.

Of the top five opportunities shown in Figure E. 1, all were viewed as methods to address a wide array of challenges. Emphasis on co-benefits, in particular, was seen as way to bridge the power disconnect between people who live on the Bay and therefore benefit directly from improvements in Bay Water Quality and those who live inland and therefore have considerable control over the amount of nutrients that reach the Bay, but do not directly benefit from improvements in Bay water quality. We expected to see evidence of this disconnect at the state level but were surprised to find that, even within Bay States like Maryland and Virginia, there was a perception that people who did not either live or work on the Bay were not concerned with Bay water quality. This, in turn, contributed to the challenge of low political will, which many respondents viewed as the root of the other challenges they described.

Some of the less frequently mentioned opportunities are logical extensions of related challenges. For instance, increasing funding and human resources is an opportunity to address the challenges of lack of funding and lack of human resources. On the other hand, climate change came up as an opportunity, as well as a challenge. Often, the connection was made by the same respondent. Although climate change is generally expected to make it more difficult to achieve the TMDL due to increased precipitation and other environmental shifts, some respondents saw it as an opportunity to increase access to funding and other resources, to address the behavior change challenge, and build political will through climate-related co-benefits. Given the circularity of many of the relationships described between challenges and opportunities, there appears to be a bit of a chicken-and-egg problem for the CBP, or at least we can say that a number of respondents indicated that people (including decision makers, stakeholders, and the public) would need to address multiple challenges by taking advantage of multiple opportunities at the same time to improve CBP effectiveness.

Co-benefits were described as an opportunity to improve water quality governance 86 times throughout our interviews and were mentioned more than any other challenge or opportunity. Because they were so important to respondents, we coded all co-benefits mentioned and organized them into 15 categories. The two most frequently mentioned co-benefits were reducing harm from flooding and improving local water quality. Reduced flooding via floodplain restoration or similar stormwater practices was mentioned as an increasing opportunity because severe rainfall events and coastal inundation are expected to increase with climate change. Local water quality was also linked to many other co-benefits that were not mentioned as often, including restored ecosystems or habitats for key species, increased recreational and real estate values, green space and livability, and healthier commercial fisheries. A number of co-benefits were specific to agriculture, including improved soil health, improved animal health, and higher profits (including through cost-share or payments for ecosystem services).

Interestingly, when we asked respondents specifically about why best management practices are either included in states' Watershed Implementation Plans or implemented on the ground, a number of statements described what we could call "co-costs" as impediments to BMP uptake. While co-benefits are positive side-effects of nutrient reduction policies, co-costs are negative side effects. Co-costs were not mentioned as challenges per se, but some are related to the difficulty of convincing stakeholders to

implement voluntary BMPs. Co-costs could be agricultural (e.g., loss of crop land, increased shade), urban (e.g., less parking, reduced safety if streetlights are blocked by trees), or suburban (e.g., loss of aesthetic and reputational values associated with lawns). Potential regressive impacts of funding mechanisms like "flush taxes" were also mentioned as co-costs. That is, when a flat rate is charged to pay for things like improved wastewater treatment plants, poorer people are more affected because the fee takes up a larger proportion of their income.

Uncertainty and the Future of Chesapeake Water Quality Governance

Overall, a majority of respondents expected that water quality governance for the Chesapeake Bay Watershed will become more difficult in the future. **Uncertainties** were highlighted around climate change and other challenges, but many respondents were also unsure how the system will change once Phase III of the TMDL ends in 2025. Some even believed that the increasing difficulty of reaching the TMDL could be a threat the Partnership, while others expected that the deadline would simply be extended. Of those who were more pessimistic, responses were further divided between those who favored more stringent load reduction requirements to provide resilience against environmental stressors (and therefore a lower goal for the post-2025 TMDL) and those who believed that load reduction requirements needed to be less stringent (higher TMDL post-2025) to make the goal more achievable given existing resources and technologies. A third group of responses indicated that changes to the TMDL itself would not be possible due to legal constraints on the CBP.

In addition to load multipliers like climate change and development, concerns about the future arose because of the **increasing difficulty** of reducing nutrient loads. Although not mentioned often as broad challenges for the CBP as a whole, detailed responses about WIP Design and Implementation both indicated that much of the "low hanging fruit" has already been harvested and that marginal costs (or cost per additional unit) of nutrient reduction are now very high for some of the most effective technologies. As noted by a number of respondents, investment in wastewater treatment plants helped states like Maryland and Virginia make large reductions in their municipal loads in early phases of the TMDL, but there is little room for additional plants in these states and further improvements load reduction efficiency per plant will be expensive. Similarly, early adopters among farmers and other landowners have already implemented many of the more cost-effective BMPs, so the costs of expanding voluntary practices is also increasing.

This brings us back around to questions about the ability of the Partnership to overcome the challenges it will face in future, possibly by taking advantage of some of the opportunities described by our respondents. Environmental economists generally anticipate increasing marginal costs of pollution reduction and would argue that the pollution limits should be set where the marginal costs to society equal the marginal benefits to society. As our results show, determining how society values nutrient reduction is a techno-political process that does not yield a simple solution, particularly for a social-ecological system as large and complex as the Chesapeake Bay watershed. What is clear is that the CBP has helped to improve water quality in the Bay and the rest of the watershed, but that it will need to improve its own institutional design in order to continue to reduce loading in spite of increasing costs and environmental amplifiers. This, in turn, may require broader changes in the political and legal systems in which the CBP is embedded. For instance, increasing equity and social justice could narrow disconnects between those who benefit from water quality improvements and those who contribute to nutrient pollution, thereby improving goal attainment and design effectiveness.

1 Introduction

This report summarizes data collected as part of the National Science Foundation-Funded Project CNH2-L: Modeling the dynamics of human and estuarine systems with regulatory feedbacks (Award Abstract # 2009248; https://www.nsf.gov/awardsearch/showAward?AWD_ID=2009248). Fifty-nine participants in the Chesapeake policy process were interviewed between June-December 2021. The primary purpose of these interviews was to collect data to aid with the development of a multi-level computational model of the policy process for nutrient management in the Chesapeake Bay watershed. A secondary purpose was to help guide exploration of potential future scenarios once the policy model is coupled with models of other components of this social-ecological system. Data collected here are complimentary to other sources, including government archives, news reports, and data on WIP Design and Implementation. We were particularly interested in collecting information on components of the policy process that could not be easily observed through public data sources. In addition to information on behind-the-scenes components of the process, we wanted to document variations in perceptions across different groups of people involved in the policy process. This helps to provide a more holistic understanding of political components of the system and potential paths for the Chesapeake Bay Program after the 2025 deadline to reach the TMDL has passed.

As described in Appendix A: Detailed Methods, we used a stratified snowball sampling method to recruit a representative sample of decision-makers based on their jurisdiction and areas of expertise.[1] Interview results were then coded using the program Atlas.ti, so that we could evaluate how often a particular view or statement—referred to here as a code or code group—occurred throughout the sample (m = number of mentions) and compare across different groupings (e.g., by the component of the process being described, by respondent characteristics, or by the time period covered).[2] Rather than counting every single time the statement is mentioned, we count how often it is mentioned as part of a particular block of text (n = number of blocks of text). This helps to minimize over-representation of repetitive statements within an interview (see Section E.1 Block Identifiers). We also use co-occurrence calculations to see how often two ideas or statements appear together (cc = number of co-occurrences; ccf = co-occurrences standardized by number of mentions for each code). This can assist with analysis of correlation and can shed light on causation when causal statements are coded. Because these data are categorical and the sample size becomes very small once categorized, we did not undertake statistical analysis for this report.

The report is organized in two main sections. Part I covers results that relate to the policy process generally. It starts by describing how respondents frame their understanding of the policy process as a whole then breaks down the results by different components or sub-prosses (Section 2 Process Frames). The next section reviews respondent perceptions of the effectiveness of the governance process for the watershed (Section 3 Overall Effectiveness), and the last section in Part I covers responses about challenges and opportunities for the Chesapeake, both historically and in future (4 Challenges and Opportunities). Part II then delves into more detail on different components of the policy process

[1] The one exception is Maryland, which tends to be weighted toward Scientist general roles more than the other jurisdictions. We will take this into account when interpreting results.

[2] To learn more about how we created these groupings see Appendix D: Respondent Backgrounds (Aggregated) and Appendix E: Processes and other General Variables.

including: Section 5 Setting the TMDL and Allocating State-level Loading Targets, Section 6 WIP Design and Sub-State Allocation Processes, Section 7 Implementation, Section 8 Other Processes, and Section 9 System Process. Section 8 covers the modeling process, funding process, and a few ancillary processes. Section 9 brings us back around to system-wide findings, discussing respondents' expectations about the future and specific recommendations for improving the policy process as a whole.

Part I: General Results

The three sections in this part of the report describe general results from our interviews, mainly derived from variables created early in the coding process.

Section 2 Process Frames describes how groups of respondents use technocratic, participatory, or political frames to describe components of the policy process and the CBP system as a whole. In Section 3 Overall Effectiveness, we examine evaluations of the effectiveness of the CBP by analyzing comments about goal attainment, procedural or institutional design effectiveness, and equity. Section 4 Challenges and Opportunities rounds out our analysis of a priori codes by looking at the topics respondents identified as either challenges that limit CBP effectiveness or opportunities to improve CBP effectiveness. We also elaborate on an important set of opportunities, co-benefits or benefits from BMP implementation that are additional to nutrient reduction, and another important set of challenges, disconnects between people who have incentives to improve CBP effectiveness and those who have the power to do so.

2 Process Frames

For the most part, participants view water quality governance in the Chesapeake watershed using a political frame that focuses on contestation over the distribution of costs and benefits associated with load reduction. This includes subprocesses like Funding and computer Modeling. Only WIP Design was viewed to be more participatory than political, emphasizing engagement and cooperation rather than conflict, though the difference was small. Respondents in the Implementation Role were more likely to view processes as participatory. Respondents in the Tech/Modeler Role were more likely to use technocratic frames, but even they used political frames overwhelmingly when describing the system as a whole.

As described in Appendix A: Detailed Methods, three major frames were defined *a priori* and then coded for the main sections of the policy process: technocratic, participatory, and political. Frames are different ways of describing the same thing, so it is normal for there to be some overlap among them. Figure 1 shows the number of blocks tagged with each of these policy frames by the policy process described.[3] Technocratic frames describe processes primarily in terms of following existing rules and procedures. These rules are believed to be fixed and outcomes are deterministic. In other words, decisions are prescribed by rules and are not altered by who is participating or how power is distributed among participants. While technocratic frames are high in frequency for the TMDL/Allocation and WIP Design processes, they are low for all other process. In fact, there were no technocratic descriptions of the whole CBP system. Participatory frames highlight the importance of including diverse actors in the policy process, including building relationships among the CBP and state-level agencies and engaging stakeholders in a given process. Participatory frames were only high in frequency for the WIP Design

[3] See Table 1 at the end of this section for specific definitions of each frame as it applies to the different policy processes described below.

process, where they were only slightly lower than technocratic frames. In contrast, political frames were high frequency in all six process categories and were the most frequent frame used in all but WIP Design. Political frames view policy process as contested, where groups (e.g., stakeholders, NGOs, state agencies, political parties) with different levels and sources of power vie for benefits or to reduce costs they expect to pay. This frame also includes statements about the distributional effects of a process which may be "baked in" as artifacts of past power struggles/structural injustices.

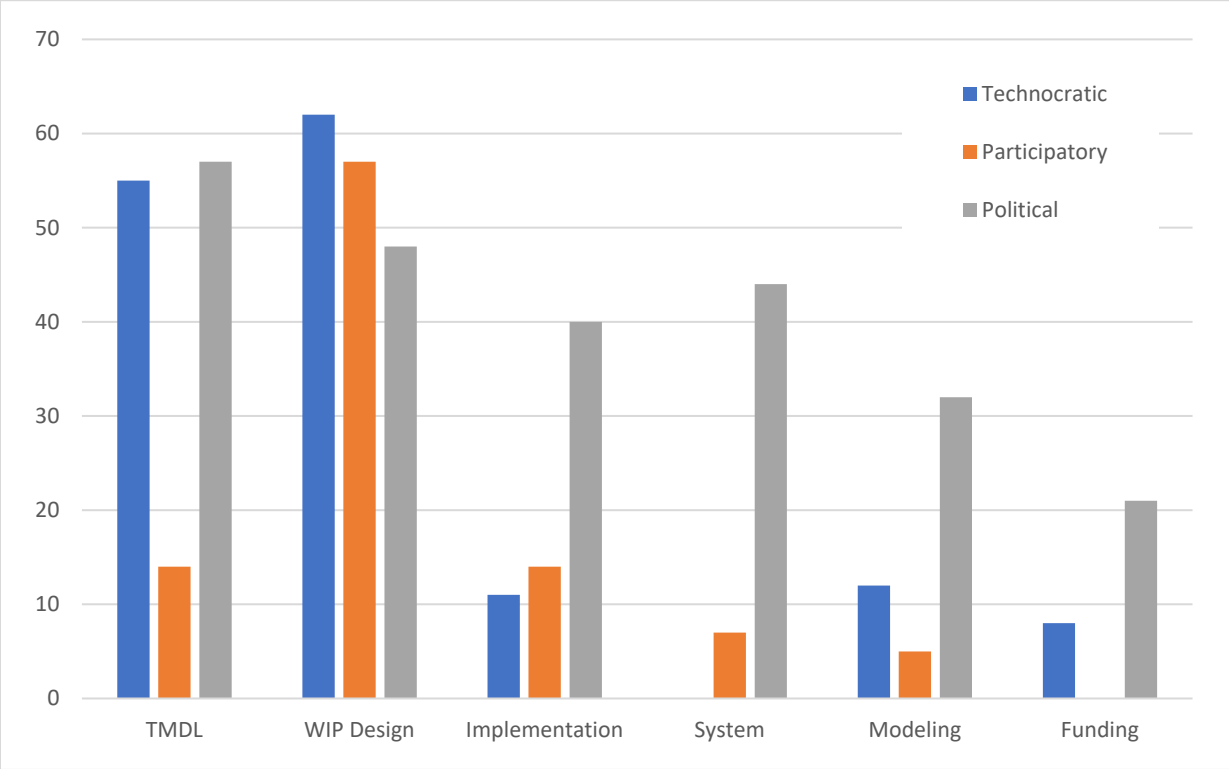


Figure 1: Policy Frames by Number of Blocks per Process. Political frames dominate in all but the TMDL and Design processes but are still high even there. The TMDL/Allocation process is also described frequently using a technocratic frame while the WIP Design process is most often described as technocratic, then participatory, then political. Data also illustrate the higher number of blocks of text for WIP Design/Allocation compared to other sub-processes.

Although they are distinct, these frames are not mutually exclusive, even within the same stage of the design process. For instance, the same respondent might express a technocratic frame when first discussing WIP Design but later switch to a political frame. In such cases, both frames would be included in the code for the relevant block of text. Nevertheless, co-occurrences between frames were low. In fact, the co-occurrence coefficient was 0.15 or less for all combinations of frames except the TMDL: Political and Modeling: Political frames (cc = 13, cf = 0.17), reflecting the importance of computer models in establishing the TMDL and allocating loading targets among state-level jurisdictions.[4]

It is particularly interesting that the modeling process was described much more often using a political frame than a technocratic frame. This may reflect the roles and expertise of our respondents, particularly the fact that only 11 of our 59 respondents identified their primary role as

[4] See Section A.2.2 Code Development for more on how these frames were coded and Section A.2.4 Annotations and Code Groups for details on the use and interpretation of blocks, mentions, and co-occurrence measures (cc and cf).

technical/modeling. Respondents in the Tech/Modeler role were somewhat more likely to describe the modeling process using a technocratic frame (50% of blocks) when compared to other roles (0-33%). They were also highly likely to use the technocratic frame when describing the process of setting the TMDL/Allocation (76.47%), but this frame was less dominant than the political frame for most other processes described by this group (see Figure 2).

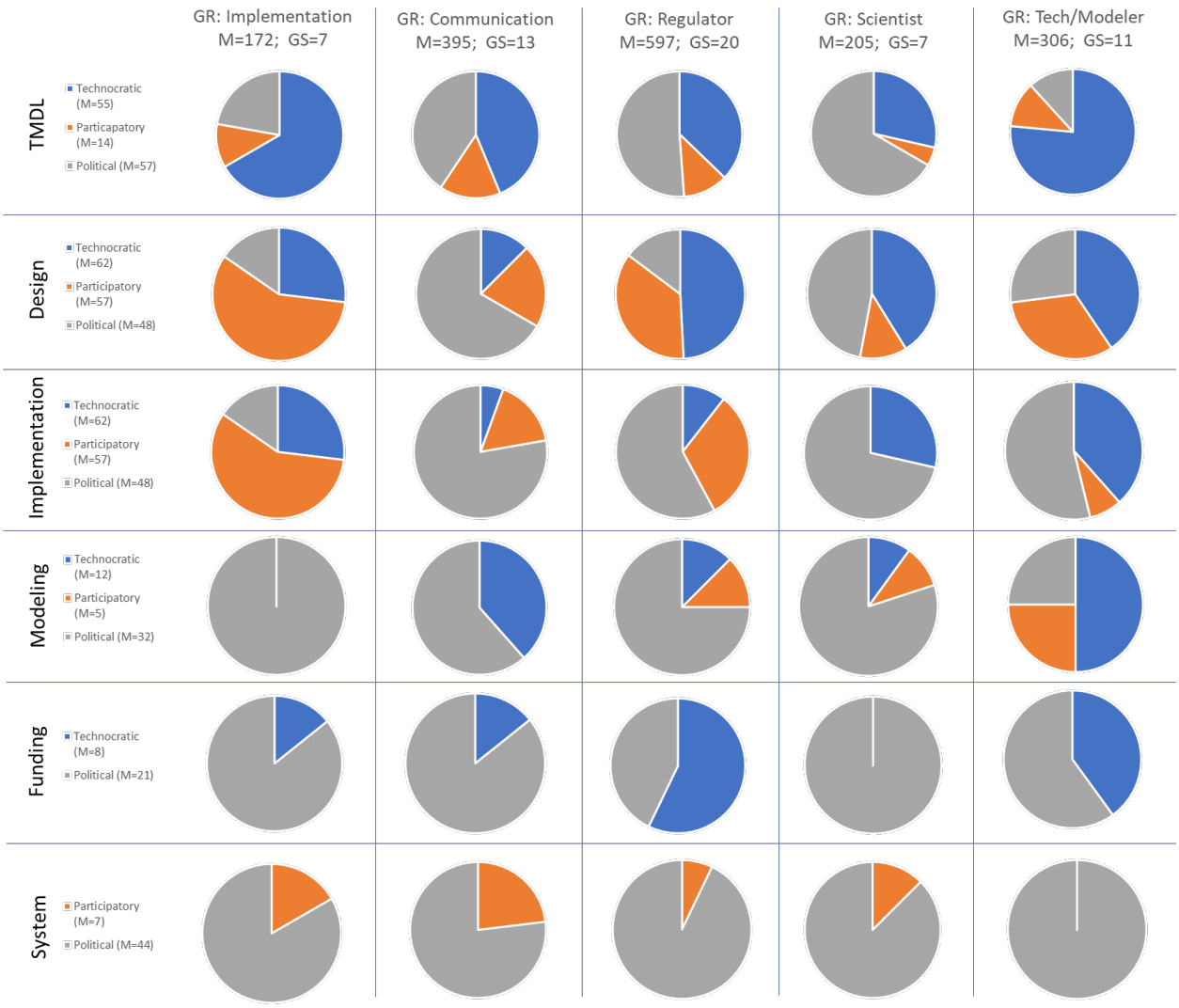


Figure 2: Policy Frames by Component of the Process and by Respondent's General Role. Reinforces the general result that the policy frame dominates across most components of the policy process. Respondent's general role only seems to have a systematic relationship with their framing for tech/modelers, except for Funding and System Processes. Implementors are also more likely to use participatory frames for WIP Design and Implementation but political frames for other processes.

Interestingly, respondents coded as Scientists were more likely to use a political frame compared to other roles (65.57%; keeping in mind that this group includes social scientists), with only WIP Design standing out with more blocks using Technocratic and Participatory Frames combined. Communicators also used the political frame in more than half of their blocks of text (60.82%), but for them, the TMDL process was a bit more balanced between Political and Technocratic Frames. The distribution is a bit flatter for Regulators, though political frames are still dominant, followed by technocratic and then participatory frames. Regulators were more likely to describe WIP Design and Funding using

Technocratic frames. Respondents coded with the Implementation Role were the only group that was most likely to use participatory frames (37.04%), but this was very close to their use of political frames (35.19%). For them, Participatory frames were concentrated in the WIP Design and Implementation process blocks.

Overall, there was much less variation across respondent role types for the system as a whole, which all groups perceived to be primarily political in nature. Funding was also described mainly as political processes by all groups except for Regulators, who used technocratic frames more often for this process. Again, Modeling was described primarily as a political process by all groups except Tech/Modelers. Variation among groups increases as we move to the processes that were identified a priori and which were the focus of specific sections of the survey instrument. This may reflect differences between information volunteered (on System, Funding, and Modeling processes), vs. information requested (on TMDL, WIP Design, and Implementation processes).

Variation in perceptions of all the above processes may also reflect individual experience. In particular, use of participatory frames may be shaped by an individual's experience with each type of process. That is, respondents whose roles give them more opportunities to participate in a particular process may view it as more participatory. This can help to explain the dominance of participatory frames used by Implementors when describing both the WIP Design and Implementation processes. In contrast, respondents who do not have much input into a process may be more likely to describe it using political or technocratic frames. Looking at the frames in aggregate, we do not see much difference among groups when respondents are categorized by their length of experience, jurisdiction, or type and level of participation in the policy process. Respondents with 4-9 years of experience were slightly more likely to use technocratic frames (48%) than political frames (34%), but otherwise political frames were most frequent across all levels of experience (40-65% of blocks), with remaining blocks split fairly evenly between participatory and technocratic frames.

Jurisdictional variation in the use of frames across all policy processes are shown in Figure 3. Maryland stands out with a very high concentration of political frames, but, this likely reflects the larger than average representation of scientists in this group (an artifact of the sampling process, see Section A.1 Sampling). In contrast, the Multi/Independent group also used the political frame much more than either of the other two frames, but this likely represents their outsider perspective rather than sampling bias per se. Virginia, Delaware, Pennsylvania, and Federal respondents also used political frames more than others, though the difference is not as large as for Maryland or DC. Respondents from West Virginia and Washington, DC were somewhat more likely to used technocratic frames than political or participatory frames. New York was split fairly evenly among the three frames but, like Pennsylvania, representatives from New York did use participatory frames more than respondents from other jurisdictional groups. While these differences are interesting, they are also small for the most part, and we cannot rule out the effects of sampling bias or small-n bias. Indeed, we see this as generally supportive of a nuanced view of the policy processes, where all three frames provide insights.

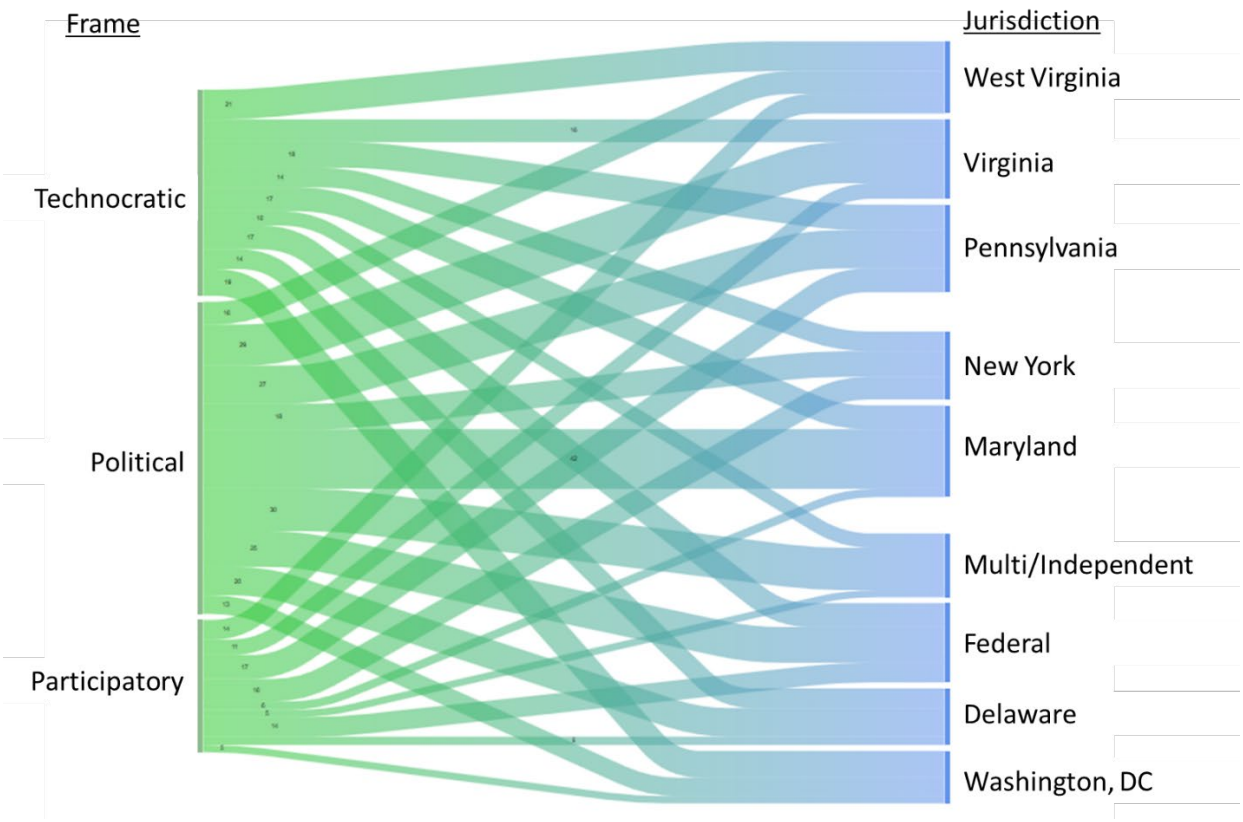


Figure 3: Aggregate Frames by Jurisdiction (Blocks of Text, not Respondent Characteristics). Shows no pattern of variation among frames used to describe jurisdictional processes (mainly WIP Design or Implementation). Maryland may be exceptional because of the high number of political frames, but this also reflects the higher number of Scientist general roles in this group.

Lastly, when we break down the results by process and respondent characteristics, the most obvious source of variation is in the number of blocks where any frame was used, as opposed to the distribution of frames per se. This can be seen in the radar charts shown in Figure 4. WIP Design and Modeling both show a "butterfly" pattern in at least one frame, where we see the highest number of blocks used by those with the high or medium rates of participation in each of the procedural categories described in Section D.3 Participation and Outside Engagement. Looking across all of the graphs, respondents in the Management: High category consistently used the political frame most frequently for all but WIP Design, where it was edged out only slightly by the technocratic frame. This contrasts to the results for the general role of Regulator, which was much more balanced between political and technocratic frames. Medium and High participation in Advisory Committees follows a similar pattern, which is more consistent with the frames observed for respondents in the general role of Scientist.

Looking at the two technical participation categories, high participation in Core Technical committees can be linked to dominant usage of political frames for volunteered descriptions of Modeling, Funding, and System processes, but political frames are used less frequently for implementation processes, are on par with technocratic frames for setting the TMDL/Allocation, and lower than both Participatory and Technocratic frames for WIP Design. Work group participation is most difficult to interpret, possibly because it is an intrinsically different measure than the other three. There is also likely to be considerable overlap between low workgroup participation and medium or high levels of participation in the other categories, simply due to the hierarchical nature of the CBPO. That said, it is interesting that

high participation in work groups is correlated with a more participatory view of implementation, or at least participatory frames are used as often as political frames for this group. Work group participation often involves area experts and implementors, so this is in line with the analysis above.

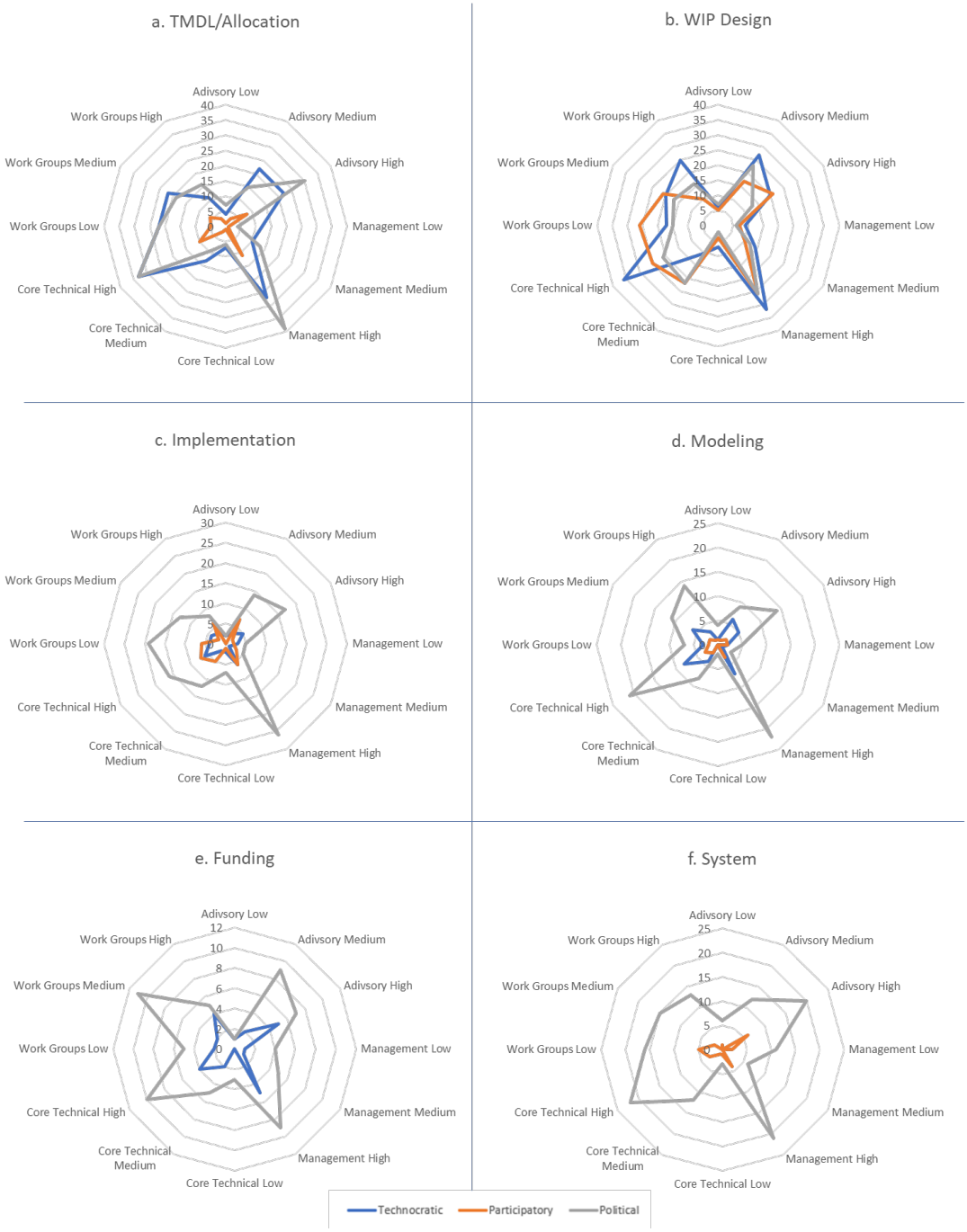


Figure 4: Frames by Process and Respondent Participation in CBP Structures

Methods Note: A few more notes on methods are important. First, as shown in Table 1, the definition of each frame differs slightly depending on which portion of the policy process it describes. Because additional processes were identified in the first round of coding, we added variations on these frames for modeling processes, the process of finding or distributing funding, and for the system as a whole.

Second, the difference between blocks/code and respondents/code may be important. Of these frames, the majority are only mentioned once per interview, but the design frames in particular were mentioned more frequently, particularly in interviews with the Federal and Multi/Independent group because blocks were coded by state as well as process and time period. Since respondents in these groups often described WIP Design for multiple states, they had more opportunities to use different frames, or at least their responses contain more blocks of text than other respondents in the Design section of the interview. In theory, this should also affect implementation frames, but frequency/respondent is not higher for this process. Across all processes, other than WIP Design, political frames are mentioned more often per respondent, but only TMLD1c: Political is mentioned more than 3 times per interview (4 times/respondent in 2 interviews).

Table 1 Definition of Policy Frame by Policy Process. Frames vary slightly from one policy process to another because of variation in the nature of each process.

PROCESS	TECHNOCRATIC	PARTICIPATORY	POLITICAL
SETTING/ REVISING THE TMDL AND ALLOCATION OF LOAD AMONG STATES [TMDL1]	Use model to estimate spatial and sectoral distribution of loads, then distribute load to states based on results; Decision = Top-down from EPA/CBP with minimal discussion of negotiation among states [TMDL1a]	Negotiation within partnership; highlights importance of engaging all member states and ensuring the equity (e.g., Equity Principles, Tiered Scenarios) and legitimacy of the process [TMDL1b]	Negotiation within partnership but may include EPA as important player; highlights power of different partners, strategic effects of decision rules (e.g., consensus), and the distribution of costs/benefits; terms like "under the table", "behind the scenes", etc. [TMDL1c]
WIP DESIGN [DES1]	Use model (CAST/Bay Models) combined with expert opinion to allocate loads across sectors and sub-state jurisdictions and to determine the best suite of management practices [DES1a]	Work with local partners to allocate load and decide on the best suite of management practices; focus on stakeholder engagement and transparency [DES1b]	Allocate load and decide on management practices based on the relative political power of relevant interest groups; focus on seeking to placate vested interests and avoid backlash [DES1c]
WIP IMPLEMENTATION [IMP1]	Focus on procedural aspects of implementation (e.g., providing technical support, monitoring progress, punishing non-compliance as per existing regulations, etc.) [IMP1a]	Focus on engagement with stakeholders and providing information to persuade them to effectively implement management practices [IMP1b]	Focus on role of power, usually through control of funding, threat of punishment, or influence through social networks (e.g., farmers influencing other farmers). May discuss effects of politics at any level as long as it relates to implementation [IMP1c]

MODELING [MOD1]	Focus on the technical or procedural aspects of model development and parameterization [MOD1a]	Focus on the role of stakeholders in model development or parameterization [MOD1b]	Focus on the political aspects of model development and parameterization, particularly how the model affects the allocation of loading targets and accounting for BMPs [MOD1c]
FUNDING [FND1]	Focus on the procedural aspects of funding, such as how to calculate funds needed, what bureaucratic procedures are used to find or allocate funds [FND1a]	NA. No blocks included text about stakeholder input in the funding process.	Focus on political aspects of funding decisions, including why some groups/communities might receive more funds than others [FND1c]
SYSTEM [SYS1]	NA. No blocks described the system as a whole as a technocratic process.	Focus on the entire system as a participatory process, including among the states and CBP and/or with local communities and stakeholders [SYS1b]	Focus on entire system as a political process. Includes exogenous political influences, contestation among the state- and federal-level government agencies, and political implications of CBP policies for state and local governments or stakeholders [SYS1c]

3 Overall Effectiveness

The vast majority of statements about effectiveness indicated that water quality governance is only partially effective or ineffective. Most evaluated effectiveness in terms of goal attainment (reaching the 2025 TMDL or related environmental improvements) but institutional design effectiveness was also evaluated frequently. Inefficiency and bureaucratic inertia were two common indications of poor design effectiveness, especially at the CBP. Very few respondents mentioned equity as a criteria for effectiveness and those who did were more likely to discuss equity of allocation among state-level jurisdictions than diversity equity and inclusion for stakeholders.

Respondents were asked to evaluate the effectiveness of water quality governance in the final section of the interview instrument.[5] When coding the documents, we broke overall effectiveness down into three main categories, following typical definitions of governance effectiveness in political science: goal attainment, institutional design, and equity (see Figure 5). **Goal attainment**—or meeting the

[5] See Part II In-Depth Descriptions of Processes for a more detailed break-down of effectiveness by policy process, jurisdiction, etc.

environmental goals set out by the law—is the most straightforward measure of effectiveness in some ways, although there is an extensive literature on the difficulties of attributing environmental change to any one policy or set of policies. Most respondents talked about effectiveness as goal attainment (m = 58), though many also commented on institutional design (m = 46), and some also commented on effectiveness as either equity among states (m = 8) or equity for stakeholders and the public (m = 6). We also added annotations with an ECF prefix whenever respondents provided rationale for their evaluations of effectiveness (m = 79).

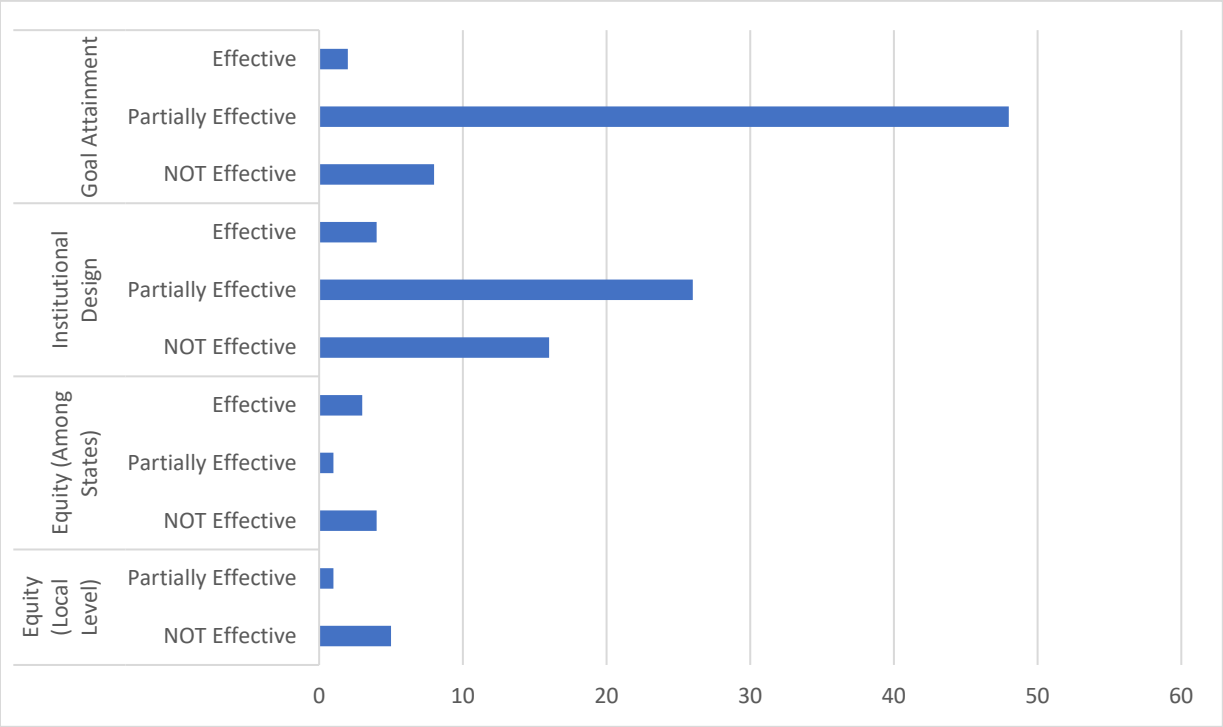


Figure 5: Mentions of Effectiveness by Type. Most statements described effectiveness in terms of goal attainment. Of these, the vast majority indicated that Chesapeake water quality governance was Partially Effective. Respondents also evaluated design effectiveness as partially effective for the most part, though CBP procedures were deemed ineffective in a larger number of blocks of text. Equity was used least often as a measure of effectiveness and was divided between concerns over the equity of loading allocation among states and concerns about equity for participants at the local level. Respondents were split on their evaluation of state-level equity but statements about local-level equity indicated that the system is not effective.

In general, respondents had a complex understanding of effectiveness which encompassed more than one of the categories described here, so multiple effectiveness codes per document and even per block was common. Although we asked about the effectiveness of "governance" generally, all respondents tended to focus on the effectiveness of the CBP and/or the EPA as the administrator of the CBP. Some respondents did consider how outside factors might affect the effectiveness of the CBP, however. We provide detailed information on all of these facets of effectiveness in the rest of this section, starting with Section 3.1 Goal Attainment, then Section 3.2 Institutional Design, and finally Section 3.3 Equity. Specific definitions of institutional design effectiveness and equity effectiveness are covered in the relevant subsection.

Methods Note: In a few cases, we ran out of time to ask about effectiveness specifically, but many respondents volunteered information about their perceptions of effectiveness in other parts of the interview. Also note that in some cases, descriptions of effectiveness might have been limited by time

and respondent characteristics (some were more likely to provide detailed answers than others). Because we are reporting in aggregate and codes used are too generic to identify any one individual respondent, we will report codes mentioned fewer than 5 times in this section. We are still only reporting on the number of blocks in which a code occurs, not the number of respondents who made a particular type of comment. This means that we may have higher frequencies, particularly in the tails of the distribution. That is, people who really like the CBP may make more statements that it is effective and people who do not are likely to make more statements that the CBP is ineffective.

3.1 Goal Attainment

The vast majority of responses indicated that Chesapeake water quality governance has only been partially effective on goal attainment. More positive evaluations of this type of effectiveness compared current water quality to the counterfactuals of either pre-CBP water quality or beliefs about water quality in the absence of the CBP. More negative evaluations compared current water quality to the goal set with the TMDL or related environmental outcomes. Many also included the caveat that CBP institutions could be improved even if they considered goal attainment to be effective or partially effective.

As shown in Figure 5 above, of the 58 mentions of goal attainment, the majority (m = 48) indicated that the CPB was at least partially effective. Only two blocks contained unqualified statements that governance was effective overall, and 8 statements said unequivocally that governance was ineffective at meeting the TMDL. The low number of Effective codes is not surprising, given that recent estimates show that it is unlikely that all jurisdictions will meet their loading goals by 2025 (Garner 2022). Explanations that respondents volunteered for their evaluation of goal attainment by the CBP are more interesting (see also Section 4 Challenges and Opportunities). Figure 6 shows the breakdown of statements about goal attainment by related rationale (ECF annotation code groups). Note that respondents might provide more than one reason for a single evaluation, and that these topics may have been raised in other contexts, therefore may be covered more fully in other sections of this report (e.g., Sections 4.1 Challenges, 6.1 WIP Design Aggregated, and 8.1.3 Transparency and Legitimacy in the Modeling Process).

It is interesting here that rationale for statements that the CBP has been effective at attaining its goals overlap with explanations for statements of partial goal attainment. One statement said that the CBP is effect, and then went on to say that the TMDL-based system currently in place is more effective than pre-2010 governance (“> Pre-CBP Attainment”). Another statement about goal attainment said the CBP was effective, but that the process could be improved with different institutions (“Could Improve Institutions”; could be any of: CBP rules and norms, formal laws and regulations at all levels, and broader social institutions such as political will, policy entrepreneurship/leadership, and stakeholder buy-in).

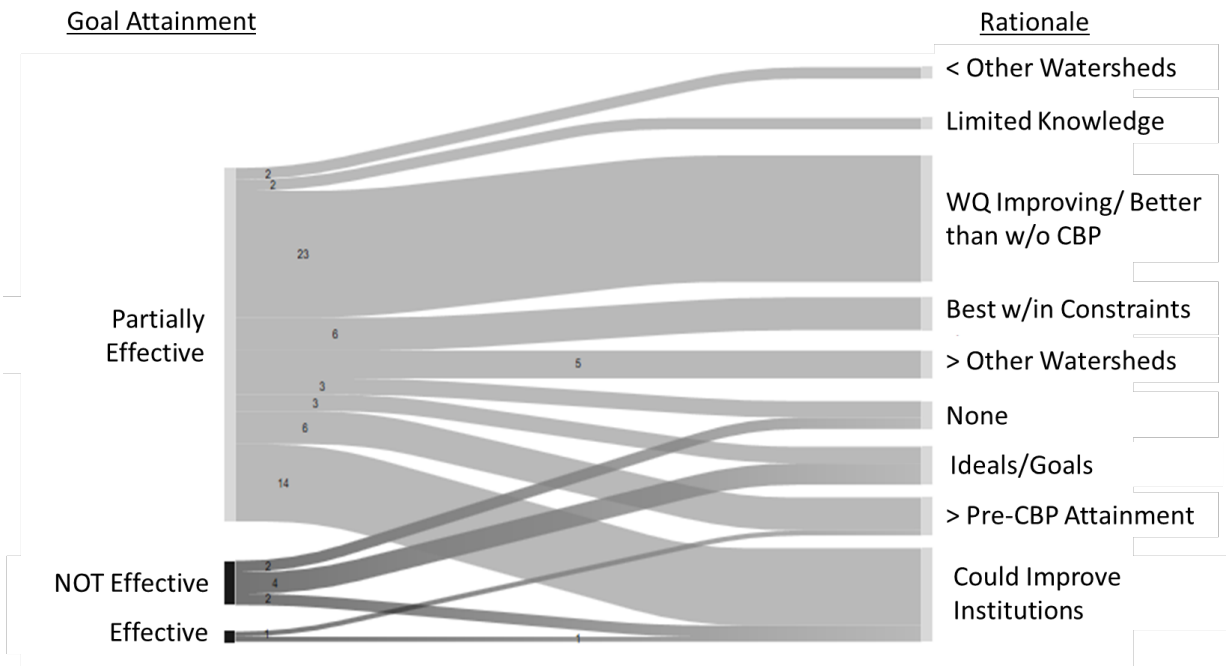


Figure 6: Sankey Diagram of Rationale for Statements about Goal Attainment. Where rationale was provided, comparison to the historical counterfactual (water quality is improving/better than without the CBP) was most commonly used to support the evaluation of partial goal attainment. For all three evaluations of goal attainment, caveats that CBP institutions could be improved were also common. Most evaluations of ineffectiveness on goal attainment compared the attained level of water quality to the 2025 goal set by the TMDL or related improvements in ecosystem function.

Suggestions for improving institutions were also mentioned in conjunction with two statements that CBP has not been effective in meeting its goals. However, in this group it was more common to either have no supporting statements (None) or for respondents to list the ways in which the CBP has not provided sufficient improvement in water quality and/or met its own goals for the TMDL (Ideals/Goals). Here again, however, these rationales were not unique to "goal attainment = ineffective" statements. In fact, statements that the CBP is partially effective on goal attainment were also linked to these three rationale, suggesting that differences in the basic statements about goal attainment (effective or NOT effective) may reflect individual personalities or general attitudes toward the CBP more than actual goal attainment by the CBP. Certainly, this is a subjective question and the answer depended heavily on respondent values, as well as the information available to them and their overall willingness to explain their response.

Keeping the above in mind, it is still useful to look at the remaining rationale summarized in Figure 6. Statements that water quality in the Bay has improved or is better than it would have been without the CBP were the most common (m = 23). Political scientists call the latter rationale a **counterfactual**. In the absence of direct causal linkages between policy and changes in environmental variables, the counterfactual (**what would have happened in the absence of said policy**) is a useful tool for assessing effectiveness in its own right. Continuing in descending order of frequency, fourteen statements of partial goal attainment were followed by recommendations for improving related institutions, suggesting that the respondent believed there was a causal link between lack of design effectiveness and lack of goal attainment. A wide range of specific suggestions for improvements were made by respondents, but these will be combined with recommendations provided in other contexts and summarized in Section 9.3 Future Needs and Expectations.

Six blocks in the partial goal attainment category showed that respondents believed that the CBP was doing the best job possible within the constraints of existing laws or other political-economic realities (outside of CBP institutions but still another causal link to Design Effectiveness). Other supporting statements for partial goal attainment included comparisons against other watersheds (mostly that CBP was doing better (> Other Watersheds), but in two blocks that CBP was doing worse (< Other Watersheds)) or against goal attainment in the Chesapeake watershed prior to the 2010 TMDL system (>Pre-CBP Attainment). Lastly, two blocks are a poor fit for this category because the respondents did not say that attainment was partial per se, but rather attainment was effective w/in their own sector or region of expertise, but they could not comment on other sectors or regions. We chose to include these here for the sake of transparency.

3.2 Institutional Design

Evaluations of Institutional Design effectiveness are generally more negative than goal attainment, but a majority still rate the CBP as partially effective. Inefficiency was the most prominent rationale for evaluations of partial effectiveness or ineffectiveness. Politicization, failure to listen to local-level stakeholders, lack of resources, and over-reliance on computer models were also common rationale for negative evaluations of design effectiveness. Some effective or partially effective evaluations recognized that the CBP was designed as well as it could be given external constraints like federal and state laws, existing political will, and the large scale of the watershed.

Statements about institutional design of the CBP proved more difficult to parse. **Institutional design** refers to the formal rules and informal norms that circumscribe the operations of governance in the watershed and which the CBP, EPA follow as they seek to carry out their responsibilities under the law. These include the various official agreements among the EPA and member jurisdictions, the committee structure of the CBP and related decision processes (i.e., the "consensus continuum"), and policies such as the "Equity Principles", "2-Year Milestones", "Backstop Measures", the CBP Office (CBPO) which facilitates Partnership activities, and, indeed, the TMDL and WIPs themselves. In theory, state, federal, or local laws that circumscribe or facilitate CBP activities could be included, along with state agency and local government norms associated with the CBP. However, respondents tended to focus on the effectiveness of the CBP/EPA rather than the broader "governance" system, so we mainly discuss internal design effectiveness of the CBP writ large. Some respondents recognized the importance of external components of institutional design, so these will be described as well.

Returning to Figure 5, statements that the CBP is partially effective were most common (m = 26). Unqualified statements that the CBP was poorly designed (m = 16) were also fairly high, while unqualified statements that the CBP was well designed were much lower (m = 4). Unsurprisingly, respondents from the federal level were more likely to evaluate the CBP design positively, but only by a few blocks. Respondents from the multi/independent category were also more likely to evaluate the CBP design negatively (5 of the 16 negative evaluations; 5 of the 7 design effectiveness blocks from multi/independent respondents), though there was also one positive evaluation from this group. Respondents with expertise in Policy/Law/Social Science were also more critical of CBP design (17 of 22 evaluations were negative).

The rationale provided for statements about the design of the CBP were more varied than those provided for goal attainment (see Figure 7). Here again, many of these same topics are covered in Part II

In-Depth Descriptions of Processes, so we will not dwell on these results. However, there are some clear areas of agreement and disagreement among blocks, which also reflect common themes throughout the interviews. One area of agreement is that the CBP process is not efficient in terms of time spent in meetings, red tape, and other bureaucratic processes (m = 14). With the highest number of mentions in the context of design effectiveness, this rationale was used to explain 5 assessments of "Not Effective" and 8 assessments that the CBP was "Partially Effective". One block also evaluated the CBP as "Effective" on design because of its inefficiencies. That is, in this block of text the respondent said that the CBP was effective in forming collaborations across such a large area BECAUSE of design elements that created inefficiencies (e.g., extensive committee structure, consensus decision process, etc.). No blocks of text contained language indicating that the CBP was procedurally efficient.

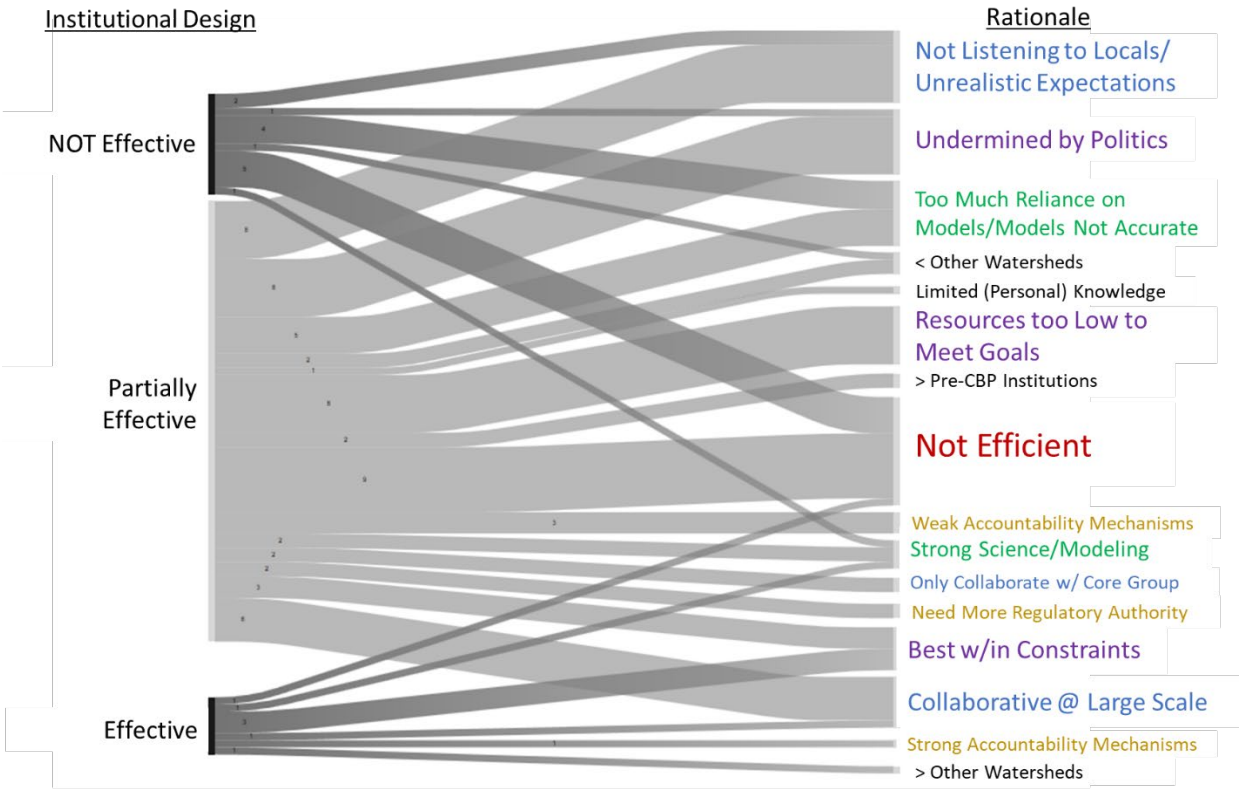


Figure 7: Sankey Diagram of Institutional Design Effectiveness by Rationale. Shows that Inefficiency is the most common rationale for negative evaluations of institutional design and is even mentioned for one positive evaluations. Lack of concern for local-level constraints, politicization, lack of resources, and over-reliance on models were also mentioned frequently to explain ineffective or partially effective evaluations of design. Most positive and some partially effective evaluations indicated that the CBP was doing the best it could within external constraints or given the large scale of the watershed.

Similar tradeoffs were mentioned elsewhere. For instance, 4 of the 7 blocks that mentioned the CBP fostered collaboration at a large scale also mentioned that the CBP was not efficient. However, inefficiency was also mentioned in 1 of the 2 blocks indicating that the CBP was collaborative, but only for a core group of decision-makers close to the CBPO, and 5 of the 9 blocks indicated that the CBP might be collaborative at the state and federal level, but out of touch with stakeholders at the local level. Thus, while there is certainly disagreement about the effectiveness of CBP design in fostering collaboration and engagement throughout the watershed, respondents with different views on the effectiveness of said collaboration agreed that the governance process is not procedurally efficient.

Other tradeoffs that were seen to limit CBP effectiveness included those associated with political, legal, social-ecological, and resource constraints. For blocks which evaluated CBP design as either "Effective" or "Partially Effective", some respondents indicated that the design was the best it could be within one or more of those constraints ("Best w/in Constraints" in Figure 7). Others were more critical, asserting that design could have been improved in order to minimize the effects of some constraints. Failure to supply sufficient resources to meet the goals set by the program was generally seen as a design flaw, although in one block there was a co-occurrence with the "Best w/in Constraints" code. Similarly, sensitivity to political concerns at the state level was usually seen as a design flaw, though in one block it was viewed as a constraint on CBP design. Human behavior was also seen as a constraint in a few blocks, particularly given the difficulties of monitoring or enforcing regulations on non-point-source pollution. We will delve into this in greater detail in future sections, but it is interesting to see that at least one respondent believes that the CBP has strong accountability mechanisms, which makes its design effective, while in a few other blocks statements that design is only partially effective were supported by criticisms of lack of accountability or lack of regulatory authority.

A last important area of disagreement was on the use of models and science. In 4 evaluations of design effectiveness, respondents mentioned that the CBP has fostered important advances in watershed modeling and science. Of these, 1 block said that the CBP was effectively designed because of its use of models and science, 2 indicated that production of good models and science was evidence that the design was partially effective, and 1 evaluated the design as ineffective in spite of being good on models and science. In contrast, 9 blocks mentioned that over-reliance on models and/or the inaccuracy of the CBP models reduced design effectiveness to either partially effective (m = 5) or not effective (m = 4). Although the sample size here is too small to evaluate relationships statistically, visual inspection of the data suggests that these results are not based in respondent biases; there is no clear pattern of relationship between respondent backgrounds/jurisdiction and their evaluation of the effective use of modeling/science in the CBP policy process. Importantly, modelers and scientists are just as likely to evaluate the role of models positively or negatively when compared to regulators, communicators, and implementors.

3.3 Equity

Equity was mentioned much less often as a criteria for effectiveness and when it was, it could either refer to equity for local stakeholders or equity among state-level jurisdictions. Concerns about the use of models for allocation of loading goals was a primary rationale for negative evaluations of equity among states. Ignoring local-level concerns was most often mentioned as rationale for negative evaluations of local-level equity. Provision of resources was the most common rationale for positive evaluations of equity effectiveness at both the state and local levels.

When we first decided to code for **Equity** effectiveness, we were largely thinking of diversity, equity, and inclusion for the stakeholders in the watershed. However, respondents also discussed equity in terms of allocation of loading targets and funding/resources among the state-level jurisdictions, so we created two separate codes for equity. As shown in Figure 5, equity in general was not mentioned as often as a criteria for effectiveness as either goal attainment or institutional design, with 14 blocks total, of which 6 discussed equity for local stakeholders and 8 discussed equity among states. Rationale for these assessments of effectiveness in attaining equity were also less diverse than explanations for other types of effectiveness.

Figure 8 shows the break down in evaluations of equity effectiveness among the state-level jurisdictions that are partners in the CBP. Here again, the sample size is very small, but these issues will arise in other sections of this report as well (see, i.e., Sections 8.1.3 Transparency and Legitimacy in the Modeling Process, 8.2.2 Influence and Options in the Funding Process, and 9.1 Factors and Actors). For now, lack of equity among states was attributed once to funding asymmetries (some states receive more funding than others), but problems with the way the models are used to distribute load were mentioned more often (m = 3). This rationale was also provided to support the one evaluation of partial equity among states. In contrast, increases in state funding through the CBP were used to explain evaluations of equity among states once, and in two blocks of text the CBP procedures, particularly the Equity Principles, Hockey Stick, and related processes for assigning loading targets to states were viewed as equitable in and of themselves. Both of these blocks came from respondents in the Federal grouping.

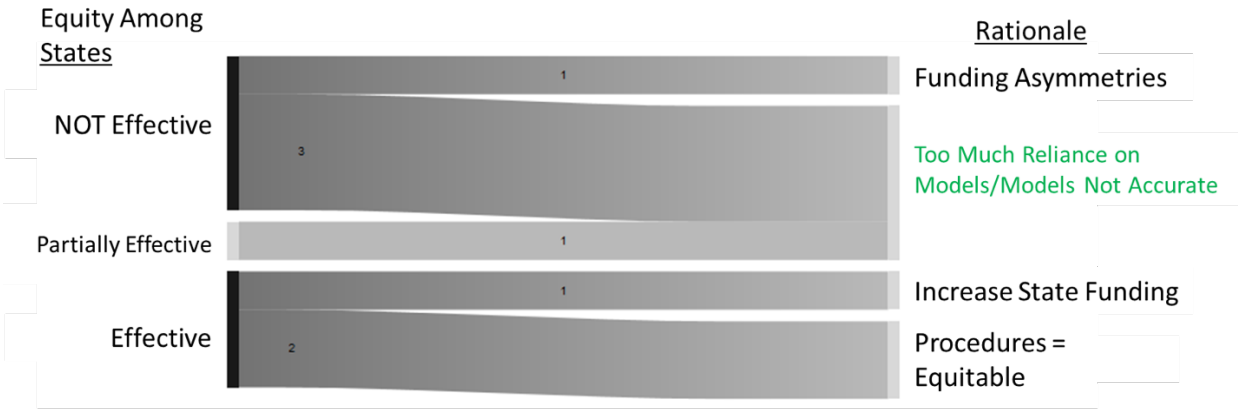


Figure 8: Sankey Diagram of Equity Effectiveness among States by Rationale. Over-reliance on models was the main rationale for negative evaluations of equity among states.

None of the respondents indicated that the CBP was effective at establishing equity at the local level (see Figure 9). Of those who said that the CBP does not foster equity at the local level, three attributed this to failure to listen to stakeholders and local governments and 2 mentioned funding asymmetries among local-level jurisdictions. That is, some jurisdictions are better at garnering funds than others. This topic came up in other contexts, so we will return to it in later settings (see Section 9 System Process). For the block of text saying that the CBP is partially effective for local-level equity, the reason provided was that the program did increase access to funding for some marginalized communities.

Figure 9: Equity for Local Stakeholders by Rationale. Though the number of mentions is low, it is telling that there are no positive evaluations of equity for local stakeholders. Main rationale for negative evaluations include funding asymmetries and failure to take local-level concerns into account in water quality governance.

Again, given the small sample size, the above analysis should be taken with a large grain of salt. It is likely that the distribution of responses across the three types of effectiveness can be attributed to how respondents define effectiveness, which in turn is shaped by the signals of effectiveness highlighted by the CBP and other organizations. Although equity and inclusion are increasingly important at the CBP, ecological goals are still the legal foundation of the TMDL process. Meeting the loading requirements—or improving water quality generally—is a large part of the job of many of the respondents to this interview, so it is understandable that this is their main focus when asked about CBP effectiveness. Similarly, the CBP has received a number of accolades and much attention in the academic literature for its institutional design, so participants are aware of design effectiveness, even if they may not use that term. In addition, many respondents experience the effects of CBP design in their regular working lives (e.g., attending meetings, grappling with data systems, etc.), and this firsthand experience makes it more likely that these elements will come to mind when they are asked about effectiveness in general. It is also telling that all of the respondents who talked about equity for local-level stakeholders were either implementation specialists or worked with stakeholders in some important aspect of their job.

4 Challenges and Opportunities

Respondents identified 42 types of challenges and 21 opportunities that could reduce or increase governance effectiveness, respectively. Respondents generally showed a sophisticated understanding of both challenges and opportunities, with most describing a tightly connected system of challenges and identifying opportunities to address multiple challenges by fostering systemic change. Of these, emphasizing the co-benefits of nutrient reduction and increasing engagement and outreach stood out. These approaches were viewed as opportunities to narrow power disconnects, or misalignments between incentives, resources, and understanding (see Section 4.4 Power Disconnects).

The next set of variables covers **challenges**, or things that make governance of Chesapeake water quality more difficult, and **opportunities**, or things that might make governance of Chesapeake water quality easier. Figure 10 shows the distribution of Challenge and Opportunity codes across the process codes that were used to delineate different blocks of text. Note that challenges were mentioned much more often than opportunities throughout the survey instrument, but otherwise the distribution loosely matches the break-down of codes by process found in Figure 101 (see Section E.1.1 Process Codes). One big difference is that system-wide challenges were mentioned most, even though the Process: System code is attached to only 20% of the blocks of texts. In fact, the average number of challenges mentioned/block is highest for Nutrient Trading ($\overline{m}$ = 3.5, but n is very small), System ($\overline{m}$ = 2.57), and Implementation ($\overline{m}$ = 2.23) Processes. Challenges are also fairly high for the modeling process ($\overline{m}$ = 1.89), but are greater than one for all processes except TMDL Only ($\overline{m}$ = 0.33, but n is small). In contrast, the average number of opportunities mentioned is less than one for all types of processes, with the highest being in Implementation ($\overline{m}$ = 0.93), System ($\overline{m}$ = 0.78) and Funding ($\overline{m}$ = 0.69) Processes.

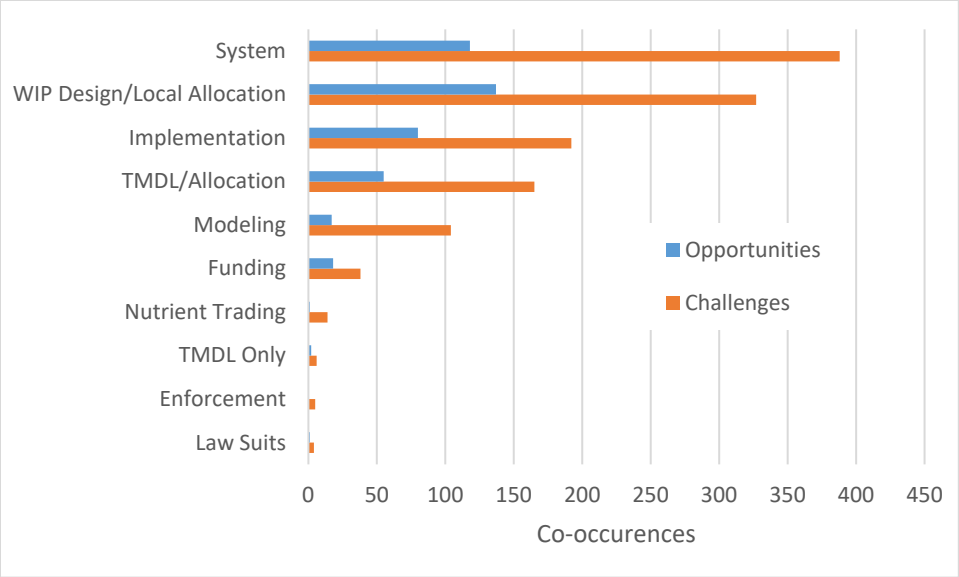


Figure 10: Challenges and Opportunities as they Co-Occur with Process Codes. Challenges outweigh opportunities for all processes and are proportionately highest for the System as a whole. TMDL/Allocation, and Modeling processes also have higher mentions of challenges relative to opportunities.

Methods Note: Here, we describe challenges for the system as a whole. Challenges and opportunities for specific jurisdictions or components of the policy process will be covered in Part II In-Depth Descriptions of Processes. There may be some overlap between these results and the rationale for effectiveness described above. That is, when a rationale for effectiveness described either a challenge or an opportunity, the block of text would have been coded for both effectiveness and challenge/opportunity. A total of 42 types of challenges were mentioned in 1,035 blocks of text. This compares to 21 types of opportunities mentioned in 378 blocks of text. Some things, like climate change, were described as both challenges and opportunities, sometimes by the same respondent.

4.1 Challenges

Respondents described many challenges, or factors that prevented effective loading reduction in the past or were expected to prevent effectiveness in the future. Of these, the majority focused on barriers to the effective implementation of BMPs, including lack of funding, lack of human resources, and the difficulty of convincing people to change their behavior. These implementation barriers were also linked to problem multipliers that were expected to make loading reduction more difficult in future (e.g., Climate Change, Conowingo Dam), problems within the Chesapeake Bay Partnership (e.g., inefficiency, negotiations among the states, and over-reliance on computer models), and other political factors (e.g., lack of political will, partisan politics, and local backlash). Although the number of mentions was lower, a wider variety of political challenges was identified compared to other types. Environmental Justice was mentioned as a challenge but was more often seen as a difficult goal for the CBP to attain than a factor that might make attainment more challenging.

From 114 detailed codes describing past or current challenges mentioned by respondents, we created 42 type codes. Table 2 (end of this section) describes the types of challenges identified by respondents in descending order by number of blocks in which each type was mentioned. To facilitate analysis, we further grouped these types of challenges into four major categories, as shown in Figure 11. Barriers to

implementation of BMPs were mentioned most often, but problem multipliers, weaknesses in Partnership functions, and other political challenges were also mentioned frequently.

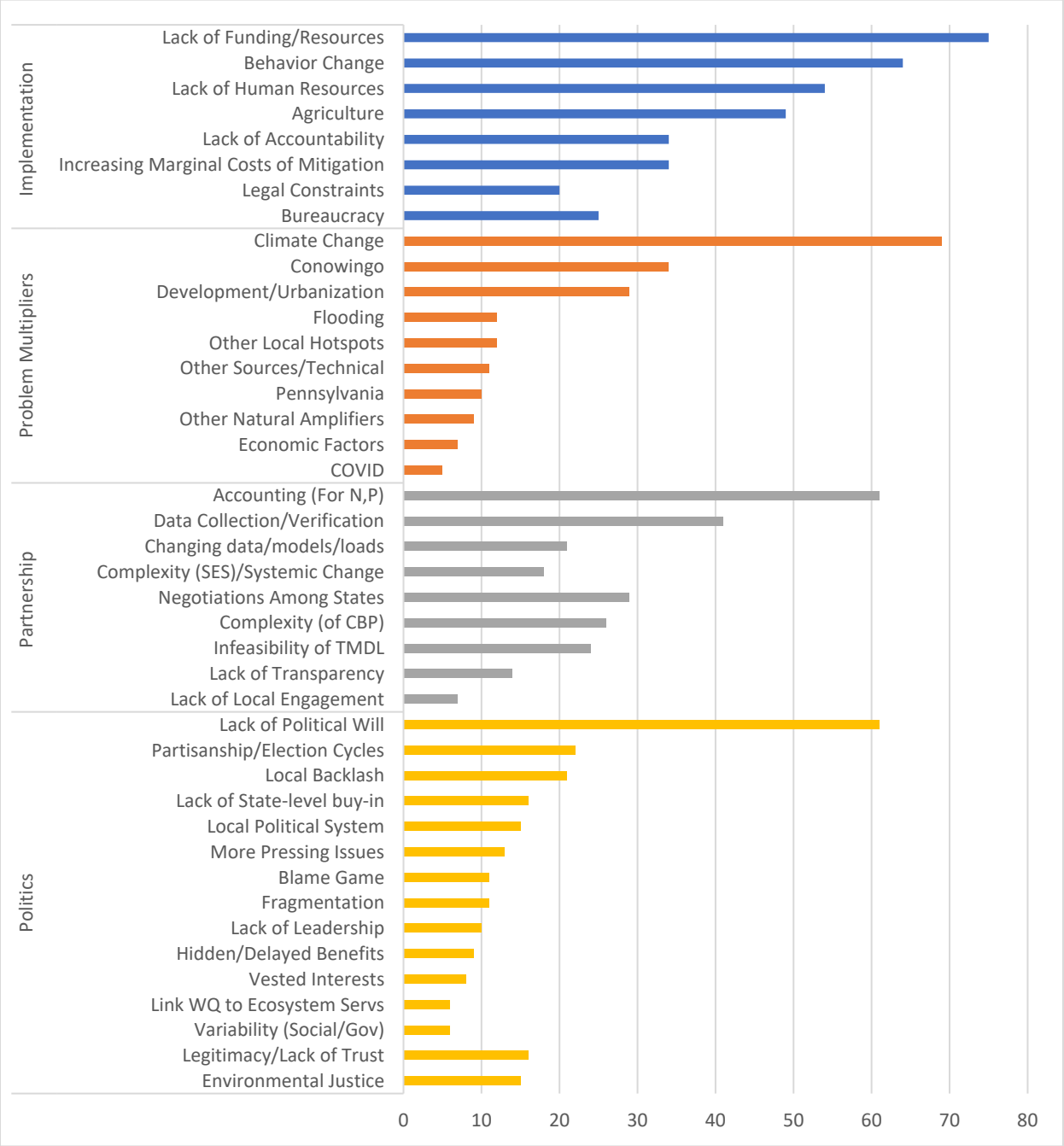


Figure 11: Challenges by Major Grouping. Most of the challenges mentioned were things that made the implementation of BMPs more difficult. Other challenges could by grouped as problem multipliers that could increase loading rates in future, partnership problems, which hindered effective functioning of the CBP, and other political barriers to effective water quality governance.

Partnership challenges, in particular, reinforce our analysis of institutional design effectiveness above. Although useful for reporting purposes, these categories are not completely separate, and so analytical interpretation should be undertaken with caution. A number of the challenges could easily have been

placed in more than one of these major categories. For instance, Economic Factors were described as problem multipliers because these blocks of text most often referred to things like agricultural prices, which could increase fertilizer-intensive agriculture, but some also described economic factors as limit on state or federal funding for BMP Implementation. Keeping this caveat in mind, the rest of this section describes the types of challenges in each of these major groupings and explains how respondents viewed challenges in relation to each other.

4.1.1 Implementation Challenges

First, lack of funding or other physical resources was mentioned most often, usually for the implementation of WIPs/getting BMPs on the ground. Other factors that were frequently mentioned as challenges for implementation included stakeholder's unwillingness to change their behavior (Behavior Change; m = 69), lack of human resources for outreach, education, and technical support (Human Resources; m = 54), and the increasing marginal costs of/diminishing returns to mitigation measures (Increasing MC; m = 34). Lack of Accountability (m = 34) at both the state or local level was also mentioned as a challenge because it reduces incentives to implement the WIPs (state-level) or to implement specific BMPs (local level). Co-occurrences between Funding and Human Resources were particularly high (cc = 31, cf = 0.32), with somewhat lower co-occurrences between Behavior Change and Funding (cc = 17, cf = 0.14) or Human Resources (cc = 11, cf = 0.10) and between Increasing MC and Funding (cc = 12, cf = 0.12) and Human Resources (cc = 9, cf = 0.11). Additionally, Bureaucracy in the form of red tape and cumbersome procedures was mentioned 25 times as a challenge that hindered implementation of BMPs (see Figure 12).

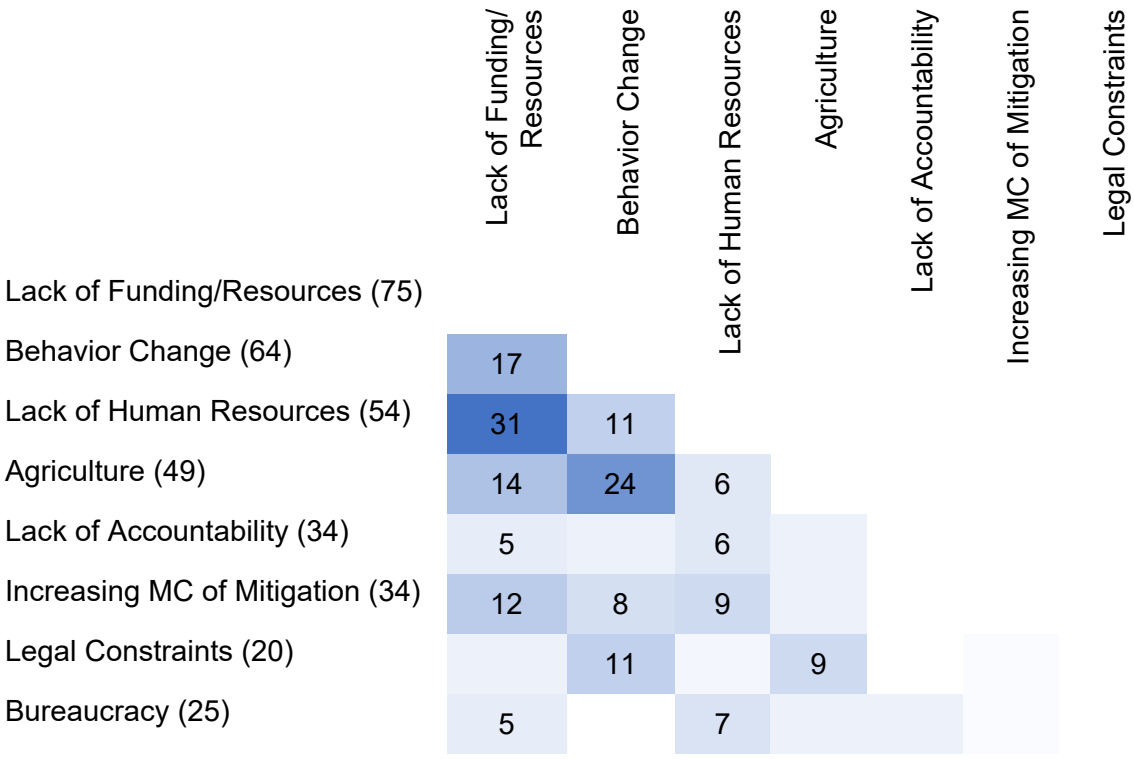


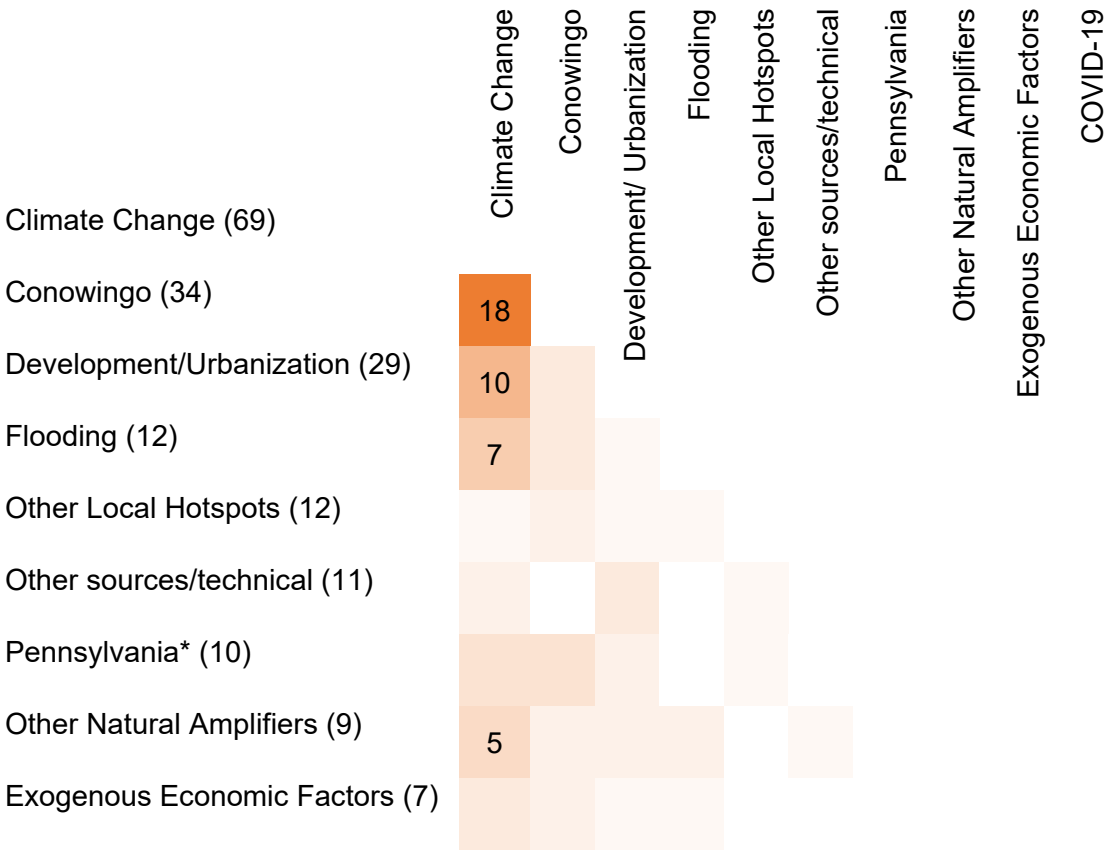
Figure 12: Heatmap of Co-occurrences between Implementation-related Challenges. There is a strong correlation between Lack of Funding and Lack of Human Resources, and causal relationships between Lack of Funding, Lack of Human Resources, and most other barriers to implementation. A causal relationship is indicated between Agriculture and Behavior Change because Agricultural BMPs are primarily voluntary. Increasing marginal costs of implementation also co-occur with Lack of Funding, Lack

of Human Resources, and Behavior change because higher costs are expected to use up more resources and make behavior change more difficult. Number of Mentions (m) in parentheses, all CC > 5 reported.

A special case of implementation challenges, Agriculture, was mentioned as a sectoral hotspot in 49 blocks of text, largely due to difficulties with implementation of BMPs. There was also a high co-occurrence between Agriculture and Behavior Change (cc = 24, cf = 0.27), reflecting the fact that most policies aimed at nutrient reduction in the agricultural sector are voluntary programs. This may also explain moderate co-occurrence between Agriculture and Funding (12, cf = 0.13), Human Resources (11, cf = 0.06), and Legal Constraints (cc = 9 of m = 20, cf = 0.15), particularly when compared to Development/Urbanization (cc = 2, 1, and 0 respectively).

4.1.2 Problem Multipliers

Second, factors that were likely to make loading worse in the future, thereby making the job of meeting loading goals more difficult, were also brought up as challenges. Climate Change (m = 69) was mentioned most frequently in this category, though it should be noted that we did prompt respondents to consider climate change as challenge or opportunity if they did not volunteer it themselves (n < 5). Interestingly, Climate Change also had the most co-occurrences with other challenges (cc = 203) and, when described as a challenge was seen as a competitor for funding and political will as well as an amplifier of nutrient pollution (see Figure 13). Flooding was also mentioned as a potential problem multiplier 12 times, as were economic factors such as prices for agricultural inputs and outputs (m = 7; could also affect implementation), other sources or technical changes that might make loading worse (m = 11), and COVID-19 (m = 5). More detailed results on climate change can be found in Section 8.3 Climate Change, Conowingo, etc.



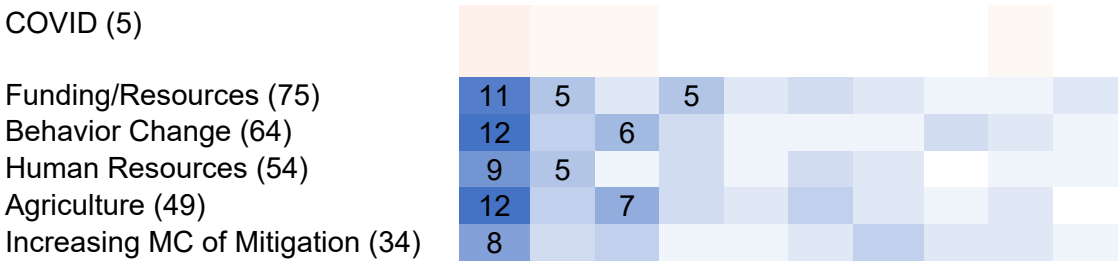


Figure 13: Heatmap of Co-occurrences between Problem Amplifying Challenges and Selected Implementation Challenges (cc > 7). Most co-occurrences are with Climate Change, including the Conowingo Dam Challenge, which is sometimes expected to get worse because of Climate Change. Respondents also expected Flooding and other Natural Amplifiers to become more common because of Climate Change. High co-occurrence with Implementation Challenges may reflect high numbers of mentions generally but in a number of blocks respondents indicated that these barriers to implementation would have greater impacts because of Climate Change. Number of Mentions (m) in parentheses, all CC > 5 reported.

The increasing frequency and magnitude of the periodic discharges of nutrients from the area behind the Conowingo Dam in Maryland was also mentioned as an important challenge (m = 34), as was increasing land Development/Urbanization (m = 29), which could increase runoff from impervious surfaces. Some respondents treated these as similar or related issues, with 18 co-occurrences (cf = 0.21) between Climate Change and Conowingo and 10 co-occurrences (cf = 0.11) between Climate Change and Development/Urbanization. Other geographic hot spots were mentioned less frequently and so were aggregated into a single group (m = 12), though Pennsylvania was mentioned specifically as a challenge to goal attainment in 10 blocks. Respondents differed in their rationale for mentioning PA as a challenge, but its large contribution to both freshwater and nutrient loads in the Bay was an area of agreement. We will look more closely at the role of Pennsylvania and other state-level jurisdictions in Part II In-Depth Descriptions of Processes.

4.1.3 Partnership Challenges

Third, a number of concerns about the functioning of the CBP were raised. Within this group, use of models, data collection, and related technical issues were mentioned frequently as challenges. Problems with the way that models were used to account for changes in loads of nitrogen (N) and phosphorus (P) in response to policy measures were mentioned most frequently (Accounting for N,P; m = 61). Interestingly, while some respondents felt that the current system of accounting for the effects of best management practices overestimated the reduction in load delivered to the Bay, more respondents indicated that the CBP system is underestimating the effects of BMPs or not counting all BMPs, and therefore underestimating load reductions. The Accounting for N,P challenge co-occurred 16 times (cf = 0.19) with the related challenge of collecting or reporting data on BMPs and verifying BMP implementation (Data Collection/Verification; m = 41) (see Figure 14). Both of these processes were seen to be resource intensive on the ground but also to be difficult because of the number of different types of BMPs available and natural variations in their effectiveness around the watershed. Frequent changes in the design or parameterization of CAST and related Bay Models were also mentioned as a challenge when results increased required load reductions or reduced credits for BMPs (Changing data/models/loads, m = 21). Difficulties accounting for N,P were also considered to contribute to the challenge created by the Complexity of the CBP itself (m = 26, cc = 7, cf = 0.09), but in some cases respondents argued that it reflected the challenge of the Complexity of the CBW as a whole (Systemic Complexity; m = 18, cc = 6, cf = 0.08).

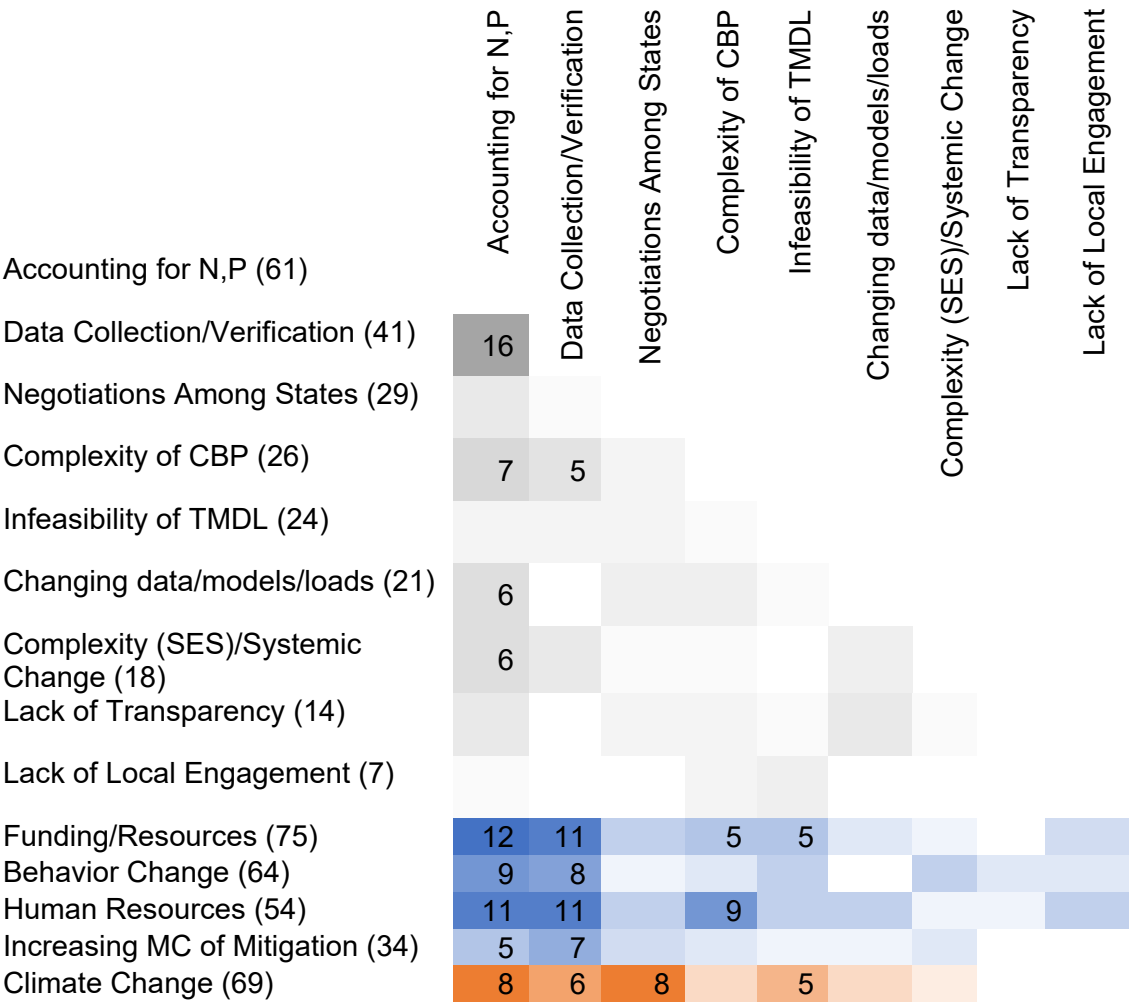


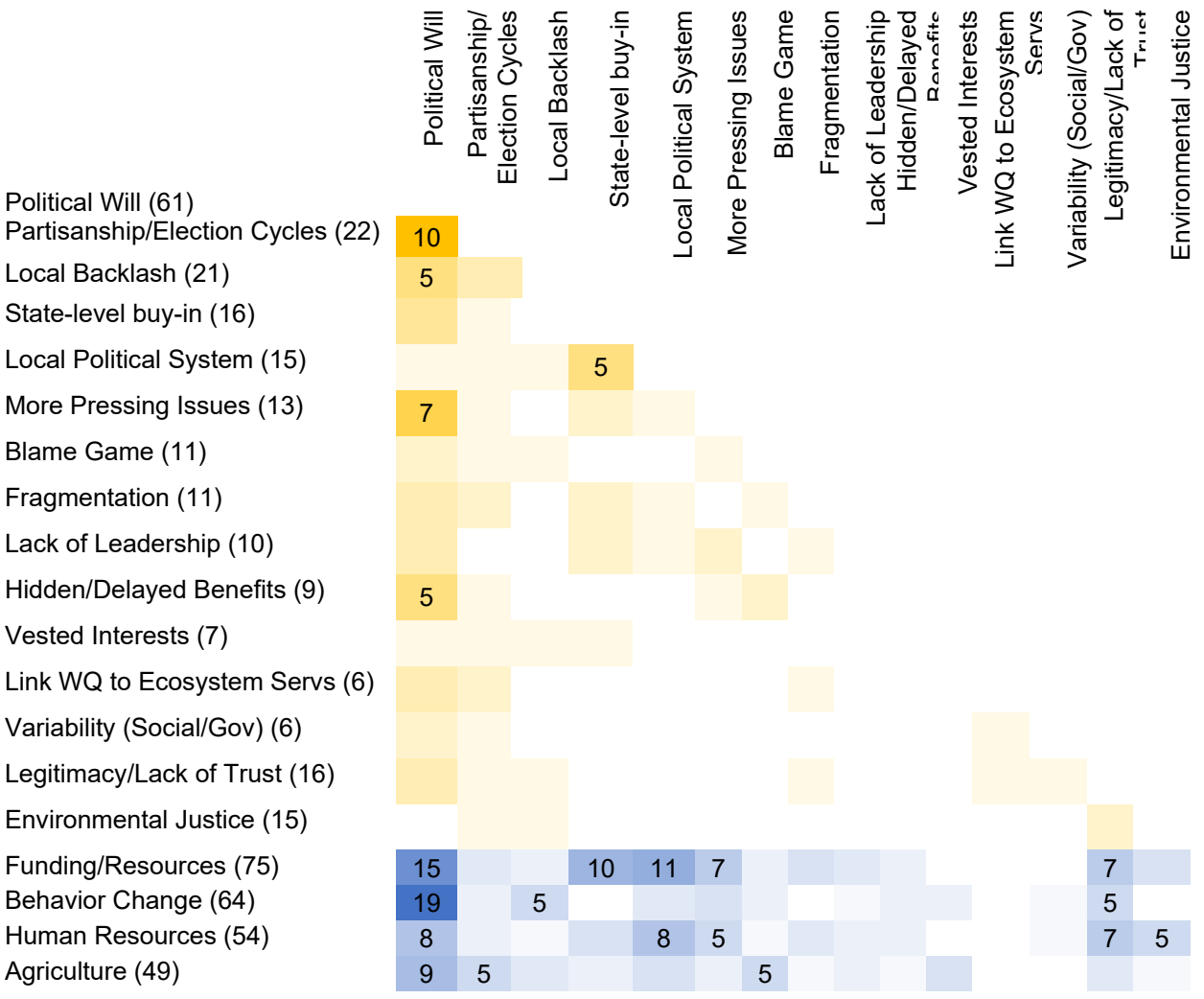
Figure 14: Heatmap of Co-Occurrences Between Partnership Challenges and Selected Other Challenges (m > 7). Shows connections between modeling challenges, with highest co-occurrence between Accounting for N,P and Data Collection/Verification. SES Complexity/Systemic Change and the Complexity of the CBP were both seen to make Modeling Challenges more difficult, but the relationship was reciprocal for Complexity of the CBP. Number of Mentions (m) in parentheses, all CC > 5 reported.

Some Partnership Challenges related to other aspects of the CBP. Of these, negotiations among state-level jurisdictions were mentioned most frequently (Negotiations among States; m = 29), usually as a process that was just difficult in itself but in some cases as a process that undermined the effectiveness of CBP policies. As noted above, 26 blocks stated that the Complexity of the CBP structure made it more difficult to navigate the committee process, advocate for the respondent's/organization's goals, understand how decisions were made, get access to needed resources, or evaluate the effectiveness of CBP policies. This challenge echoes respondent concerns about institutional design effectiveness (see Section 3.2 Institutional Design). The Infeasibility of the TMDL was also mentioned as a challenge in 24 blocks, with some simply stating that the goal could not be reached with others citing reasons it was not feasible, including referencing various other challenges already described above. Other challenges associated with the CBP were a Lack of Transparency (m = 14) and Lack of Local Engagement (m = 7), which is a reiteration of the findings on equity and effectiveness described in Section 3.3 Equity. Although it was mainly described as a challenge at the state-level, Fragmentation—or the splitting of up responsibility among multiple organizations that are not well-coordinated—was mentioned 11 times

and co-occurred with Complexity of the CBP in 6 of those blocks (cf = 0.19), indicating that these challenges are interconnected at least in the minds of some respondents.

4.1.4 Other Political Challenges

Fourth, 14 of the 42 challenges could be considered primarily political in nature, and that is exclusive of CBP-specific challenges that have a political component (i.e., Negotiations among States, Accounting for N,P) or structures that were historically determined through political processes (i.e., Complexity of the CBP, Legal Constraints, Fragmentation). Of the 14, lack of Political Will (m = 61) is mentioned most frequently. It co-occurs most often with Behavior Change (cc = 19, cf = 0.18), Climate Change (cc = 16, cf = 0.14), Conowingo (cc = 10, cf = 0.12), Funding/Resources (cc = 15, cf = 0.12), and a number of the other political codes described below. Related challenges included the variation in funding and policy support associated with Election Cycles and Partisan Politics (m = 22, cc = 10, cf = 0.14 w/Political Will), More Pressing Problems which compete with nutrient pollution for funding and public attention (m = 13, cc = 7, cf = 0.10), and Hidden or Delayed Benefits from nutrient reductions (m = 9, cc = 5, cf = 0.08) (see Figure 15). A more specific political process, the "Blame Game" (m = 11, cf = 0.03) which may be used by states or stakeholders to shift responsibility for load reductions onto other groups co-occurred most often with Agriculture (cc = 5, cf = 0.09) and Conowingo (cc = 3, cf = 0.07).



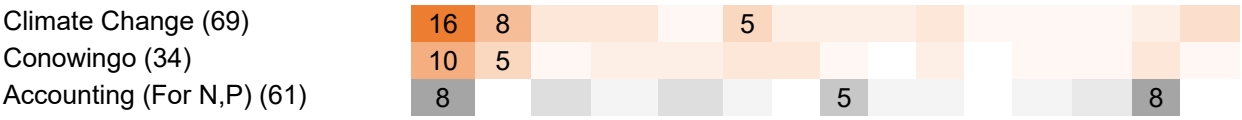


Figure 15: Heatmap of Co-occurrences between Political Challenges and Select Other Challenges (m > 7). Political Will co-occurs with most other challenges, particularly those that may affect political will such as Partisanship/Election Cycles, Local Backlash, More Pressing Issues, and Hidden/Delayed Benefits. Political Will also co-occurs frequently with other types of challenges but the direction of the relationship varies (see also Increasing Political Will in Section 4.2 Opportunities). Environmental Justice and Equity had few mentions but the former co-occurred frequently with non-political challenges, including Accounting for N, P. Number of Mentions (m) in parentheses, all CC > 5 reported.

Local Backlash against CBP policies (m = 21) was also mentioned fairly often as a challenge in its own right, a drain on Political Will (cc = 5), and a barrier to Behavior Change (cc = 5). Although it was mentioned less often than Local Backlash, challenges associated with the structural limitations of Local Political Systems (m = 15) co-occurred much more often with other challenges, particularly with lack of Funding/Resources (cc = 11), Human Resources (cc = 8), and State-Level Buy-In (cc = 5). That last challenge was mentioned 16 times and co-occurred frequently with Funding/Resources (cc = 10) and moderately with Human Resources (cc = 4) and Political Will (cc = 4). Lack of leadership at the state or federal levels was also mentioned as a challenge in 10 blocks, co-occurring most often as a cited result of low Political Will (cc = 3) and a cause of low levels of Funding (cc = 3) and lack of Accountability (cc = 3). However, lack of Leadership co-occurred most often with Complexity of CBP (cc =4), without consistent indication of any type of relationship between the two challenges.

A last sub-set of the political challenges are related to trust, legitimacy, and equity, though these were mentioned less often than most of the other types of challenges. Lack of Trust in or belief in the Legitimacy of the CBP, state level agencies, or the government generally was mentioned 16 times. Like many other challenges, Trust/Legitimacy co-occurs often with Funding/Resources (cc = 7, cf = 0.08) and Human Resources (cc = 7, cf = 0.11), but it occurs most frequently when described as a result of concerns over Accounting for N,P (cc = 8, cf = 0.12). Other moderately-high co-occurrences are with Data Collection/Verification (cc = 4, cf = 0.08), Changing Data/Models/Loads (cc = 4, cf = 0.12), and lack of Transparency (cc = 4, cf = 0.15), all of which are more closely related to equity among state-level jurisdictions than for local stakeholders, as explained in Section 3.3 Equity. In contrast, Environmental Justice was mentioned as a challenge in 14 blocks but was more often seen as a difficult goal for the CBP to attain than a factor that might make attainment more challenging. Co-occurrence was highest here with Human Resources (cc = 5, cf = 0.08), Funding/Resources (cc = 4, cf = 0.05), Climate Change (cc = 4, cf = 0.05), and Flooding (cc = 3, cf = 0.13). Other factors that one might expect to be highly correlated with Environmental Justice did not co-occur as often, including lack of Trust/Legitimacy (cc = 2, cf = 0.07), lack of Local Engagement (cc = 1, cf = 0.05), lack of Accountability (cc = 0, cf = 0), Local Backlash (cc = 1, cf = 0.03), and lack of Transparency (cc = 0, cf = 0).

Table 2: Challenges in Descending order by Number of Mentions

Type of Challenge	#M	#CC	M*	Description
Funding/Resources (Lack of)	75	225	4	Statement indicated that lack and/or inconsistency of funding and other physical resources made design and implementation of WIPs difficult. Does not include Human Resources.
Climate Change	69	200	2	Statement indicated that climate change would pose a major challenge for the Partnership in future. Usually because it is likely to increase

				loading/reduce water quality through higher rainfall and temperatures, which will increase conflict among the state-level jurisdictions. In some blocks, reductions in the efficacy of BMPs was also mentioned.
Behavior Change*	64	188	4	Statement indicated that changing individual behavior (farmers, landowners, etc.) is an important but difficult requirement for goal attainment. Includes factors like: pro-lawn norms, anti-wetland norms, lack of knowledge/know-how, dueling sources of information, lack of engagement, free riding/gaming the system, and general lack of trust in government
Accounting for N,P (Inaccuracies in)	61	160	5	Statement indicated that the way that BMPs and other practices are accounted for in CAST prevents them from selecting more effective BMPs (WIP Design) or getting credit for all BMPs in place (WIP Implementation). Includes skepticism about how the model was parameterized and how it translates (only) verified practices into changes in load at the edge of tide.
Political Will (Lack of)*	61	190	4	Statement indicated that lack of political will prevents the government from taking sufficient actions/providing sufficient funding for the CBP or related programs (state and federal levels). Reasons for low political will mentioned included: temporal myopia, unclear problem signals, fear of negative consequences of regulation, lack of federal engagement, lack of local government buy-in.
Human Resources (Lack of)*	54	193	4	Statement indicated that lack and/or inconsistency of human resources made design and implementation of WIPs difficult. Includes high turn-over in government agencies at different levels, lack of institutional memory, and insufficient “boots on the ground” for implementation.
Agriculture*	49	161	4	Catch-all for general frustration with the difficulty of reducing nutrient loading from agriculture. May be associated with political pressures that limit effectiveness in different components of the policy process. Includes general mentions of agriculture and mentions of poultry litter specifically.
Data Collection/ Verification (Difficulty of)*	41	115	3	Statement indicated that data collection and verification is too difficult or that these create inaccuracies in model results. Similar to CHLNG: Accounting for N/P but focuses more on problems with the data used in the models than the model itself.
Accountability (Lack of)	34	83	4	Statement indicated that lack of accountability at

				either the state or local level prevents effective design and/or implementation of WIPs.
Conowingo	34	114	2	Statement indicated that the Conowingo Dam and its WIP created conflict among the partnership jurisdictions, usually overload allocation and funding.
Increasing Marginal Costs/ Diminishing Returns*	34	110	3	Statement indicated that increasing marginal cost of mitigation and compliance and/or diminishing returns to mitigation technologies decreased ability to further reduce loads or implement additional BMPs (i.e., we've already collected the "low hanging fruit").
Development/Urbanization	29	85	2	Statement indicated that increasing development and urbanization were factors that increased water pollution and made load reductions hard to manage accurately. Population growth was also mentioned in this context.
Negotiations among States	29	82	5	Statement indicated that negotiations among the state-level jurisdictions inhibited or slowed progress on reaching the water quality goals, usually through allocation of state-level targets, but sometimes through impact on modeling process.
Complexity (of CBP)*	26	83	3	Statement indicated that complexity of the CBP structure made design and/or implementation of WIPs difficult. Includes statements that communication is poor within the CBP and that the committee structure creates considerable inertia/inefficiency.
Bureaucracy*	25	72	3	Statement indicated that federal, state, or local red tape, transaction costs, or similar components of bureaucratic inertia make it more difficult to design or implement WIPs. Similar to Complexity of the CBP ,but more focused on paperwork required for grants, reporting requirements, etc.
Infeasibility of TMDL*	24	62	3	Statement indicated that the TMDL targets were unrealistic and not feasible.
Partisanship/Election Cycles	22	75	3	Statement indicated that the partisanship and election cycles slowed progress at important points in the CBP process (e.g., negotiating Phase III targets in 2017).
Changing data/models/loads	21	48	2	Statement indicated that changes in data and models reduced transparency of the system, made it hard for individuals to plan for the future, decreased legitimacy/increased backlash, or otherwise made design or implementation of WIPs more difficult.
Local Backlash	21	43	2	Statement indicated that local backlash on TMDL allocations made it more difficult to meet local

				planning targets.
Legal Constraints*	20	57	2	Statement indicated that legal constraints posed a challenge, especially in regard to ensuring compliance and incentivizing voluntary action. Also includes a few mentions of the role of lawsuits in polarizing negotiations among partners.
Systemic Complexity (SES)*	18	40	2	Statement indicated that the complexity of the CBW as a social ecological system made it difficult to regulate at the state and watershed levels. Includes statements about accommodating geographic variation in WIPs and in CAST as well as statements about difficulties understanding the effects of a change in one time/location on the system as a whole.
Legitimacy/Trust (Lack of)	16	79	2	Statement indicated that lack of legitimacy or trust in the CBP process reduced state or local willingness to participate fully. Includes statements about lack of transparency in the allocation of state-level loading targets, distribution of funding, and CAST/model design.
State-level buy-in (Lack of)	16	54	3	Statement indicated that lack of state-level buy-in prevented effective WIP design or implementation. In some cases, state-level agencies were indicated, but state legislatures and governors' offices were cited more often for not providing sufficient funding or leadership to fully implement well-designed WIPs.
Environmental Justice (Lack of)	15	34	2	Statement indicated that the inequitable distribution of resources or failure to engage marginalized populations prevented effective implementation of WIPs.
Local Political System	15	56	2	Statement indicated that the varying local political structures and local politics made design and implementation of WIPs more difficult. Includes conflicts among local jurisdictions which might prevent collective action on large-scale projects like wastewater treatment plants.
Transparency (Lack of)	14	47	4	Statement indicated that a lack of transparency in the CBP process (especially TMDL/Allocation and Modeling) made governance more challenging for at least one of several reasons: reduced trust/legitimacy, reduced accountability, increased the transaction costs of interacting with the CBP.
More pressing issues	13	59	2	Statement indicated that competition with other issues reduced access to funding, political will, and other resources needed for WIP design and implementation.

Flooding	12	59	1	Statement indicated that increasing flooding (often due to climate change) would pose challenges for WIP implementation, both through increased competition for resources and decreased efficacy/durability of BMPs.
Other Local Hotspots*	12	39	2	Statement mentioned a specific area where high loading was a challenge for their jurisdiction. Usually a creak, river, or other body of water with high load levels but sometimes a county or other local-level jurisdiction.
"Blame Game"	11	35	3	Statement indicated that there was push-back against further load reductions/introduction of more BMPs due to stakeholders perceiving that they had already achieved a "fair" level of load reduction and that additional reductions should be borne by others. Usually at the state level, sometimes among sectors.
Fragmentation	11	52	2	Statement indicated that fragmentation of governance structure (e.g., different sectors managed by different agencies, even within the same state, some managed locally, other by state, others federal, etc.) made design and implementation of WIPs difficult
Other Sources/Technical*	11	33	2	Statement mentioned a source of nutrient pollution other than agriculture or development which they viewed as a challenge for the CBP. Includes lawn fertilizer, loss of forests, stormwater, mining, and greenwashing or false claims about the N,P levels of products or ability of a BMP to reduce N,P.
Leadership (Lack of)	10	34	2	Statement indicated that a lack of leadership or poor leadership either by state governors or within the CBPO resulted in poor WIP design or implementation.
Pennsylvania**	10	38	1	*When specifically asked about challenges faced by the CBP*: Statement indicated that "Pennsylvania" was a challenge, usually because of its high contribution to nutrient loads in the Bay and the difficulty of reducing that load given that most of it comes from non-point sources and/or because state and local politics are not favorable in PA.
Hidden/Delayed Benefits	9	28	1	Statement indicated that hidden or delayed benefits from nutrient reduction made it more difficult to convince stakeholders to implement BMPs voluntarily or to raise sufficient political will to ensure effective WIP design or implementation.
Other Natural Amplifiers*	9	32	2	Statement mentioned natural factors other than flooding and climate change that are likely to make governing nutrient pollution more challenging.

				Some are related to climate change but were mentioned separately (e.g., hurricanes, sea level rise, and high precipitation) others will amplify its effects locally (e.g., high groundwater table).
Vested Interests	7	30	1	Statement indicated that at least one powerful interest group would resist additional regulations or implementation of BMPs because of perceived high costs to the group.
Exogenous Economic Factors	7	23	2	Statement indicated that exogenous economic factors made prioritizing water quality difficult. Examples include crop prices and economic growth.
Local Engagement (Lack of)	7	12	2	Statement indicated that lack of local engagement resulted in impractical policies and/or limited stakeholder buy-in.
Link N,P to Water Quality to Ecosystem Services	6	19	1	Statement described difficulties of translating ecosystem-wide goals into water quality metrics and then to N,P loading requirements. Usually refers to modeling difficulties but can refer to regulatory priorities. Feedbacks to implementation and perceptions of fairness were also mentioned.
Variability (Social/Gov)*	6	23	1	Statement indicated that variability in social values and government priorities resulted in changes in funding opportunities and loading requirements that made implementation more difficult and created il-will at the state or local level.
COVID	5	18	1	Statement indicated that the Covid-19 pandemic and its repercussions had created challenges for implementation, both in terms of limiting in-person interactions and in terms of drawing resources away from water quality governance. A few statements also mentioned potential changes in housing and transportation decisions that could affect patterns of development.

#M = Number of mentions, #CC = Total Co-occurrences with other Challenges, M = Maximum mentions/interview*
** Heterogenous grouping of similar challenges after coding was completed. Merged to aid analysis and to produce groups > 5 blocks each.*
*** Does not include challenges associated with PA mentioned outside of the last question in Section 5*

Methods Note: As noted in the introduction to this section, all of the above analysis should be interpreted cautiously. Where possible I have indicated causal inferences provided by respondents, but unless otherwise stated, these co-occurrences do not indicate correlation or causation. Furthermore, these are blocks of texts, not frequency per respondent, so one particularly adamant respondent who mentioned the same challenge multiple times will increase the number of mentions for a particular challenge or co-occurrence. Figure 16 shows that this problem is minimal for the majority of challenges, as most were only mentioned once or twice per interview. However, as shown in Table 4, the maximum number of mentions of a particular challenge per respondent (M*) did reach as high as 5 for two of the challenges, 4 for 7 of the challenges, and 3 for 8 of the challenges. This will only have a large impact on the analysis for challenges with a relatively low number of mentions (e.g., Transparency and "Blame

Game"). Unfortunately, we cannot easily analyze co-occurrences by respondent, but interpretation of challenges diodes with low co-occurrences and at least one node with a high M* should be undertaken with caution. We can say, for instance, that co-occurrences between Transparency and other challenges like Trust/Legitimacy and Political Will would be lower by 1-3 blocks if we reported results by respondent. Whether this is a "bias," or it simply reflects the importance of these challenges (i.e., if the challenge is mentioned more often by a respondent it is more important to them and this should be counted), the fact remains that reporting by block is a better method for protecting respondent confidentiality.

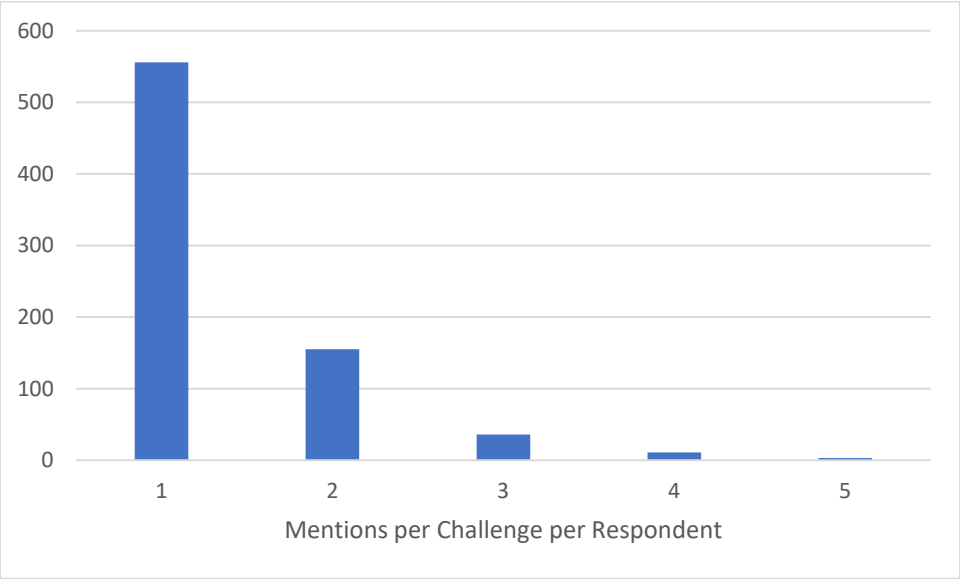


Figure 16: Mentions per Challenge per Respondent. Most respondents only mentioned a given challenge once per interview. A few mentioned the same challenge twice in their interview and even fewer mentioned the same challenge more than 3 times/interview.

4.2 Opportunities

When discussing opportunities, respondents mainly focused on factors that they believed would help to address multiple challenges at the same time. Emphasizing Co-Benefits and Increasing Engagement/ Outreach were mentioned most often because they were perceived as methods for increasing access to funding, encouraging behavior change, building political will, and, ultimately, making fundamental improvements in system function. Other opportunities identified are either more specific versions of these top two opportunities or they provide more specific approaches to tackling Implementation or Partnership challenges. Increasing Environmental Justice is mentioned as an opportunity to improve governance in the watershed, though not as often as many other factors.

Opportunities, or factors that respondents believed either improved governance in the past or could improve governance in the future, were mentioned fewer times than challenges. Table 3 (end of this section) presents the types of opportunities identified by respondents in descending order by number of blocks in which each type was mentioned. As shown in Figure 17, the opportunities mentioned most often tended to address multiple challenges. Like the major groups for challenges, opportunity groups are fuzzy sets, so opportunities like Increasing Political Will, which is clearly political, could also have

been placed in the Multiple Challenges group because most respondents who mentioned this opportunity viewed overcoming Political Challenges as a way to overcome Implementation, Partnership, and Problem Multiplier challenges as well. Because there are fewer types of opportunities than challenges, in this section we review them together, rather than breaking them down further by major category.

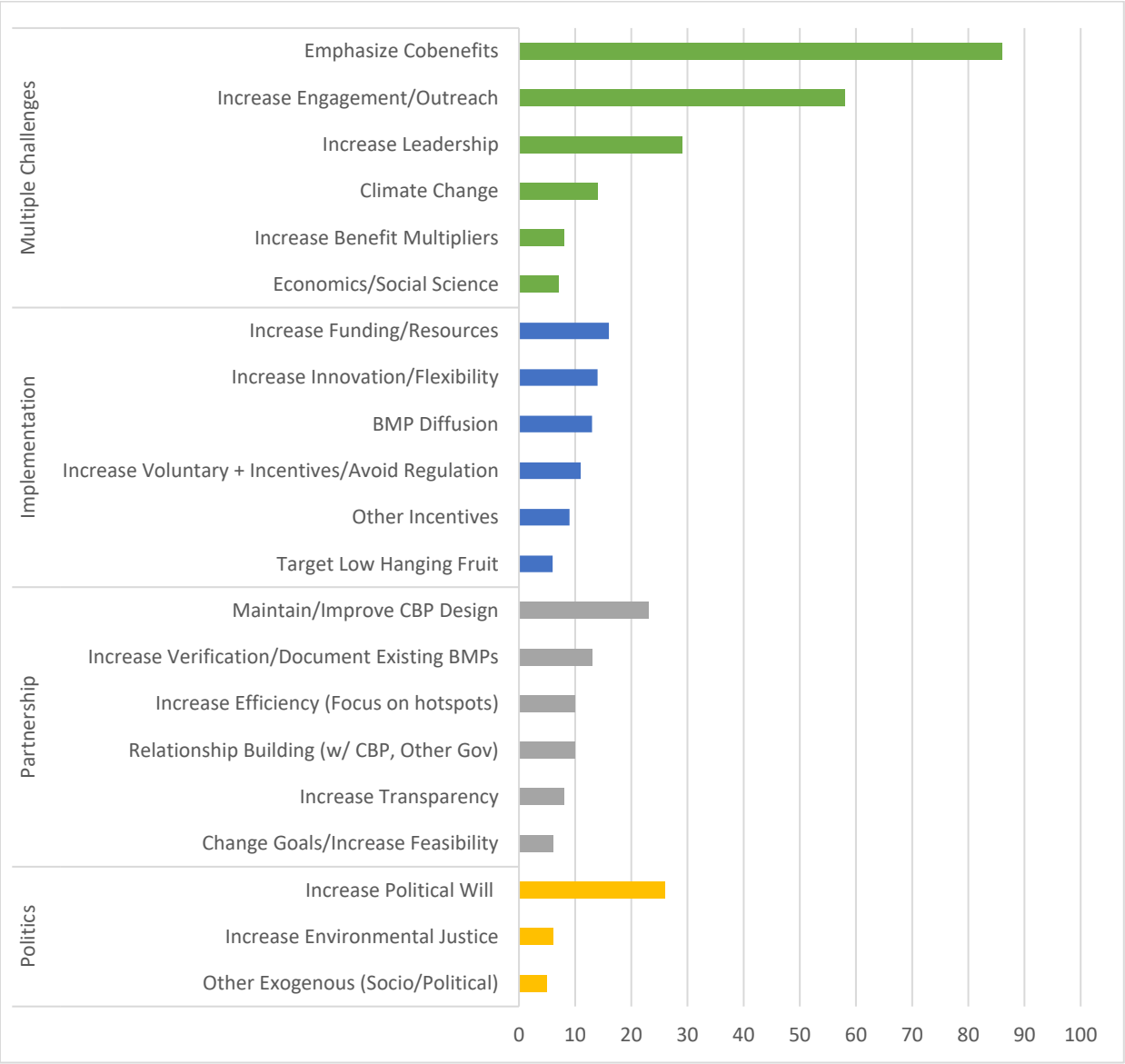


Figure 17: Opportunities by Major Category (Number of Mentions). Respondents mentioned opportunities that were thought to address multiple challenges most often. These included emphasizing co-benefits, which was mentioned more often than any other opportunity or, indeed, more than any challenge.

Co-benefits (m = 86), or benefits that accrue to stakeholders as a result of nutrient reduction measures, were most frequently mentioned as opportunities to improve WIP design and implementation through increased political will, stakeholder willingness to implement BMPs, and wider funding opportunities. This is a very diverse group, including habitat restoration and protection of charismatic fauna, improvements in human health, living conditions and recreational opportunities, and increased

economic efficiency/profitability through implementation of BMPs. Because this type of opportunity was mentioned so often and in a number of different contexts, we coded Co-benefits separately and will review those results in the next section.

Engagement and/or Outreach were the next most mentioned opportunity for improvement (m = 58), followed by more Effective Leadership (m = 29), Increased Political Will (m = 26), and Maintaining or Improving CBP Design (m = 23; another echo of concerns about institutional design effectiveness). Co-occurrences were highest between Co-benefits and Engagement/Outreach (cc = 23, cf = 0.19) but this is not surprising given the higher number of mentions for these two groups (see Figure 18). It does, however, reflect the shared perception that both opportunities could help change stakeholder incentives, which in turn, was expected to increase voluntary implementation of certain BMPs (Behavior Change Challenge, see Figure 19). Co-occurrences between Co-benefits and Leadership (cc = 6, cf = 0.06) and Political Will (cc = 7, cf = 0.07) are about the same as between Engagement/Outreach and Leadership (cc = 7, cf = 0.09) and Political Will (cc = 5, cf = 0.06). Co-occurrences with CBP Design were less than three (cf < 0.04) for all of these opportunities except Leadership (cc = 5, cf = 0.11). These co-occurring responses generally indicated that strong federal or state leadership either has led to or could lead to improved CBP Design or WIP implementation. Otherwise, both Co-Benefits and Engagement/Outreach were sometimes described as means of increasing Political Will, whereas Political Will was mentioned as a source of pressure for improved leadership in a few blocks.

	Co-benefits	Engagement/ Outreach	Leadership	Political Will	CBP Design	Funding/ Resources
Co-benefits (86)						
Engagement/Outreach (58)	23					
Leadership (29)	6	7				
Political Will (26)	7	5				
Maintain/Improve CBP Design (23)						
Increase Funding/Resources (16)		6				
Climate Change (14)	6					6
Innovation/Flexibility (14)	6					
BMP Diffusion (13)	8	5				
Increase Verification/Document BMPs (13)	6					
Voluntary/Avoid Regulation (11)						
Efficiency (Focus on hotspots) (10)	5					
Relationship Building (Gov) (10)						
Other Incentives (9)						
Benefit Multipliers (8)	5					
Increase Transparency (8)						
Economics/Social Science (7)						
Change Goals/Increase Feasibility (6)						

Figure 18: Heatmap for Co-occurrences between Six Opportunities with Highest Number of Mentions and All Other Opportunities. Majority of opportunities co-occur most frequently with Co-benefits, in part because Co-benefits are mentioned so often. Engagement/Outreach also co-occurs with several multi-challenge and implementation opportunities. Climate Change co-occurs frequently with Increasing Funding/Resources because many respondents believed that funding for climate change adaptation could be used to augment funding for certain BMPs like floodplain restoration. Number of Mentions (m) in parentheses, all CC > 5 reported; CC < 5 for all opportunity pairs omitted from this figure except for those with Co-benefits, which will be covered in the next section.

Given the high frequency of lack of Funding/Resources as a challenge (see previous section), it is somewhat surprising that increasing Funding/Resources was not mentioned as an opportunity more often (m = 16). As shown in Figure 19, the opportunity of Increased Funding/Resources does not co-occur with many challenges but emphasizing Co-benefits and increasing Engagement/Outreach do co-occur with Lack of Funding/Resources (cc = 22, cf = 0.16) and multiple other challenges. This further emphasizes the finding that respondents tended to focus on opportunities as fundamental methods to address multiple challenges. For instance, Climate Change (m = 14) and Benefit Multipliers (m = 8) were mentioned most often as opportunities to leverage additional (non-CBP) funding and, while the Other Exogenous Socio/Political (m = 5) category mainly contains references to Obama's 2009 Executive Order some refer to factors that might increase funding for CBP programs. In contrast, Human Resources were mentioned frequently as a challenge, but came up fewer than 5 times as a source of opportunity and so were folded in with all types of Funding/Resources in this set of codes. Similarly, while concerns about modeling, data, and other technical aspects of the CBP policy process were mentioned frequently as challenges, changing these systems was mentioned fewer than 5 times as opportunities and so were folded into the broader categories of CBP Design (m = 23) or increasing Transparency (m = 8), depending on the focus of the comment (general improvement or specifically mentioning transparency).

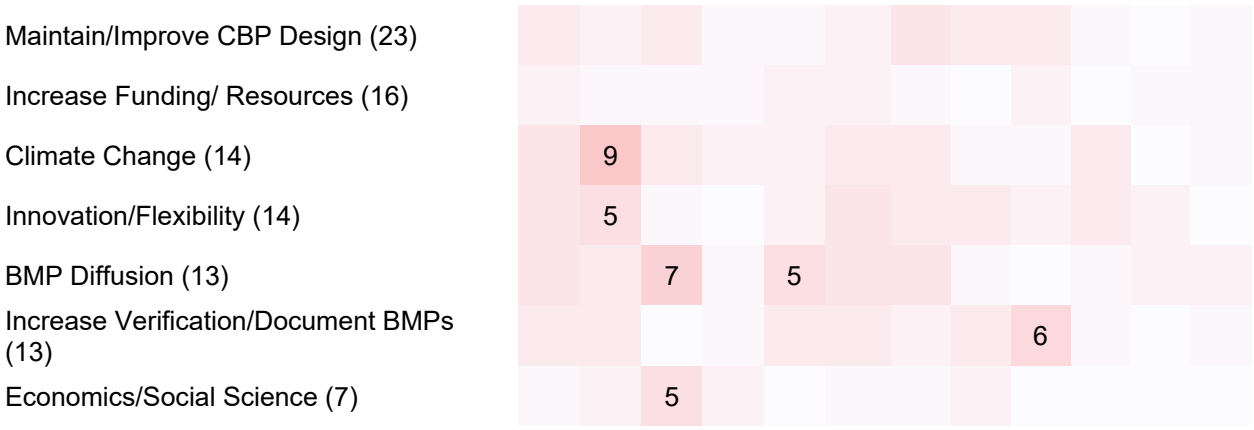


Figure 19: Heatmap of Co-occurrences between Opportunities and Select Challenges (cc > 5 for ANY except Co-benefits). Relationships between Co-benefits and Funding/Resources, Climate Change, Behavior Change, and Political Will stand out. Engagement/Outreach is viewed as a way to increase Funding/Resources and encourage Behavior Change. Climate Change is frequently mentioned as a challenge and an opportunity in the same block of text. BMP Diffusion and increasing reliance on economics/social science were frequently posited as methods to address Behavior Change. Number of Mentions (m) in parentheses, all CC > 5 reported; CC < 5 for all opportunity/challenges pairs omitted from this figure except for those with Co-benefits, which will be covered in the next section.

In fact, most of the opportunities that were mentioned less frequently can be thought of as special cases of improving CBP Design, building Political Will, Changing Behavior, or overcoming some combination of these three challenges. In addition to increasing Transparency (m = 8), opportunities associated with improving CBP Design included increasing the Verification and/or Documentation of Existing BMPs (m = 13; either to increase accountability or because the respondent indicated that existing BMPs are undercounted in the models), Increase Efficiency (m = 10) for goal attainment by focusing more on hot spots with high levels of nutrient pollution, and increasing Relationship Building among the CBP and jurisdictional agencies (m = 10; includes state-level agencies other than the main environmental agency). Changing CBP design to encourage more innovation and provide greater flexibility for stakeholders and state/local governments was also mentioned and grouped with a few comments indicating that technical innovations could be opportunities for improvement (without policy change) (Innovation/Flexibility; m = 14). Last for CBP design, 6 blocks mention that changing CBP goals, including instituting a less stringent TMDL or expanding the range of water quality indicators considered could be an opportunity to improve governance. This is not as simple as moving the goalpost, however, as in most of these blocks the intention was to bring WIP Design into alignment with Implementation (see Sections 3 Overall Effectiveness and 6 WIP Design and Sub-State Allocation Processes).

Most of the less frequently mentioned opportunities in the changing Incentives/Outreach category focus on stakeholder behavior, but there is an overlap with increasing Political Will both in the public sphere and within government bodies. BMP Diffusion (m = 13) blocks are focused on stakeholder behavior, indicating that if influential stakeholders can demonstrate real benefits from BMP implementation it tends to increase the willingness of other stakeholders to implement said BMPs. Demonstrating Co-benefits was mentioned frequently in tandem with BMP Diffusion (m = 8), at least relative to the mentions in other low-frequency categories. Increasing incentives in voluntary programs and/or using the threat of increased regulation were also mentioned as past/future opportunities to improve WIP Implementation in 11 blocks of text (Voluntary+Incentives/Avoid Regulation). This group over-laps CBP Design somewhat, as a few of the blocks call for increased regulatory capacity or changes in the way that the EPA implements backstop procedures. Most, however, assert that the current voluntary

approach is more feasible than increasing regulatory requirements. More broadly, opportunities that leveraged exiting incentives or created new ones also included: collecting remaining "Low Hanging Fruit" (m = 6), using Insights from Economics and other Social Sciences (m = 7) to change stakeholder incentives or build political will was also mentioned as an opportunity, as were various "Other Incentives" (m = 9) which were not described often enough to report separately.

Finally, Environmental Justice (m = 6) was mentioned as an opportunity in a handful of blocks. That is, respondents indicated that increasing the representation of marginalized groups in the decision process and/or distributing resources more equitably throughout the watershed could improve WIP design and implementation. This highlights the differences in perceptions of environmental justices as a challenge for the CBP to overcome vs. an opportunity to improve CBP governance. Moreover, there is only on co-occurrence between Environmental Justice as a challenge and as an opportunity, which further suggests that even respondents who think about EJ in this context do not perceive the duality of this topic (document co-occurrence = 3). In contrast, 9 of the 14 mentions of Climate Change as an opportunity co-occur with mentions of Climate Change as a challenge (document co-occurrence = 12). Again, we did prompt respondents to discuss climate change as a challenge/opportunity, so the divergence from EJ may be due to project design, but this is, at least, is evidence that further research could be useful.

Table 3: Opportunities by Number of Blocks in which they were Mentioned

Opportunity	M	CC	M*	Description
Co-benefits*	86	94	5	Respondent indicated that highlighting the co-benefits of nutrient reduction could increase political will, funding opportunities, willingness to implement voluntary BMPs, or otherwise make WIP design and implementation more effective. A breakdown of these co-benefits will be covered in the next section.
Engagement/Outreach*	58	71	4	Respondent indicated that more engagement with stakeholders and public outreach has or could increase political will, funding opportunities, willingness to implement voluntary BMPs, or otherwise make WIP implementation more effective.
Effective Leadership	29	36	5	Respondent indicated that strong leadership at the state or federal level either had made the CBP more effective in the past or could make the CBP more effective in the future.
Political Will*	26	28	2	Respondent indicated that higher levels of political will had or could increase allocation of funding and human resources, increase accountability at state or federal levels, or otherwise make WIP implementation more effective.
Maintain/Improve CBP Design*	23	15	4	Respondent indicated that some element of the CBP design had or could improve some component of the policy process. Examples include maintaining the Partnership Principles, increasing the regulatory authority of the

				EPA/CBP, streamlining the WIP Design process, and improving data collection/modeling processes.
Increase Funding/Resources	16	17	2	Respondent indicated that increased access to funding or other resources, **including human resources**, could improve WIP implementation through increased outreach/engagement, increased cost share or other incentive programs, and/or increased public spending on infrastructure.
Climate Change	14	17	2	Respondent indicated that climate change could provide opportunities to access new sources of funding and might provide additional political will to improve WIP design and implementation. In a few instances, it was also seen as an opportunity to increase cooperation among state-level jurisdictions.
Innovation/Flexibility*	14	25	2	Respondent indicated that technological and policy innovation could be an opportunity to improve the efficiency of nutrient reduction. Includes new policies that would allow stakeholders to take advantage of or develop innovative technologies.
BMP Diffusion	13	25	1	Respondent indicated that implementation of specific BMPs either has increased or will increase in future if stakeholders are shown a proof of concept, usually by influential members of their community.
Increase Verification/ Document Existing BMPs	13	22	2	Respondent indicated that increased verification and documentation of existing practices to reduce nutrient loading would show that the CBP has been more successful than currently estimated. In some cases, also mentioned that increasing accountability could incentivize more effective and longer-term BMPs.
Voluntary + Incentives/Avoid Regulation*	11	10	1	Respondent indicated that implementation of voluntary BMPs either has increased or will increase because stakeholders want to avoid the imposition of backstops by the EPA and/or to avoid transition to a more stringent regulatory system due to either legislative or judicial action.
Efficiency (Focus on hotspots)	10	13	2	Respondent indicated that focusing funding and other resources on sectors or regions that are highly polluting would improve goal attainment in the Bay.
Relationship Building (w/ CBP, Other Gov)	10	9	2	Respondent indicated that building relationships among the CBP and state agencies either has or could improve collaboration and thereby may

				improve WIP design and implementation.
Other Incentives*	9	19	1	Respondent mentioned at least one of several methods to increase incentives for stakeholders to implement BMPs, including nutrient trading and local organizing.
Benefit Multipliers	8	16	2	Respondent indicated that government agencies were able to leverage CBP resources to improve governance in other geographic areas or to meet other management goals. Similar to Co-benefits, but focus is on incentives at the level of state agencies.
Increase Transparency	8	7	2	Respondent indicated that increasing the transparency of the CBP policy process could increase political will, willingness to implement voluntary BMPs, or otherwise make WIP implementation more effective. Transparency in the modeling process and improvements in transparency associated with the use of CAST during the Phase III WIP are included here.
Economics/Social Science*	7	12	1	Respondent indicated that better understanding of human behavior through insights from economics or other social sciences could increase political will, willingness to implement voluntary BMPs, or otherwise make WIP implementation more effective.
Change Goals/Increase Feasibility	6	2	1	Respondent indicated that changing the water quality goals or reducing the amount of load reduction required by the TMDL would be an opportunity to improve goal attainment/make goals more attainable. Includes considering costs as well as benefits and establishing more holistic metrics of good water quality.
Environmental Justice	6	6	1	Respondent indicated that increased representation of diverse groups of stakeholders could provide opportunities to improve decision making processes, increase political will and willingness to implement voluntary BMPs, or otherwise make WIP design and implementation more effective.
Low Hanging Fruit (Remaining)	6	3	1	Respondent indicated that there is still capacity for large reductions in nutrient loading through high-impact/low-cost BMPs (i.e., low hanging fruit).
Other Exogenous (Socio/Political)	5	5	1	Respondent indicated that changes in political or economic systems that are not directly caused by CBP processes could provide opportunities for improvement in CBP governance. Examples

				include changes in the dominant political party in state or federal government and changes in the geographic distribution of urban/suburban development.

M = Number of mentions, CC = Total Co-occurrences with other Opportunities, M = Maximum mentions/interview*
** Heterogenous grouping of similar challenges after coding was completed. Merged to aid analysis and to produce groups > 5 blocks each.*

Methods Note: We did specifically ask about future challenges in the last section of the interview protocol, following up on opportunities only if time permitted, but opportunities also came up less often in other sections of the report (see Figure 10). From the 77 detailed opportunities to improve the functioning of the CBP, we created 25 more general types of opportunities. Like the challenge codes, co-occurrences among the opportunity codes are highly correlated with the number of mentions for a given code ($b = 1.11$, $p = 0.00$, $r^2 = 0.94$) and "mentions" indicates the number of blocks of text in which an opportunity was mentioned, without controlling for which respondents made the statements. Most opportunities were only mentioned once or twice by each respondent but some of the most frequently mentioned opportunities occurred more often in a few interviews (3-5 blocks; M* in Table 3, see Figure 26), so we still need to take care with the co-occurrence analysis, especially when co-occurrences with the top four opportunities are low.

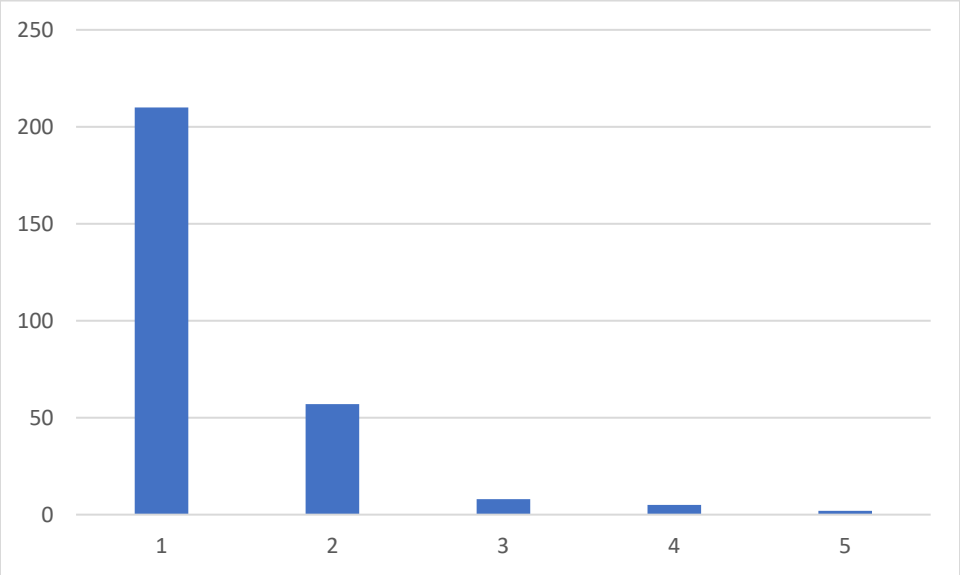


Figure 20: Mentions per Opportunity per Respondent (M). Most respondents only mentioned a given Opportunity once per interview. A few mentioned the same opportunity twice in their interview and even fewer mentioned the same opportunity more than 3 times/interview.*

4.3 Co-benefits

Respondents identified 15 types of co-benefits to load-reducing policies. Multi-sector co-benefits were mentioned most often, particularly reducing flooding and improving local water quality. While most co-benefits were viewed as means to address the challenges of Behavior Change, lack of Funding/Resources, and lack of Political Will, a few highlighted management-related benefits such as data sharing within the region and the diffusion of policies outside the region.

Co-benefits emerged as an unexpected, but important, factor as we conducted the interviews, and so we decided to code them separately, whether they were mentioned specifically as an opportunity to improve CBP governance or in any other way. We identified 46 specific co-benefits and grouped them into 15 types, as summarized in Table 4 at the end of this section. As shown in Figure 21, most frequently mentioned co-benefits can provide value across multiple sectors, including both urban and rural communities. Agricultural co-benefits were also mentioned frequently, along with co-benefits for urban and recreational sectors (Other Sectors). Environmental co-benefits included protection of specific species and of ecosystems/habitat generally. Interestingly, some co-benefits occurred outside the watershed through sharing data or policy diffusion.

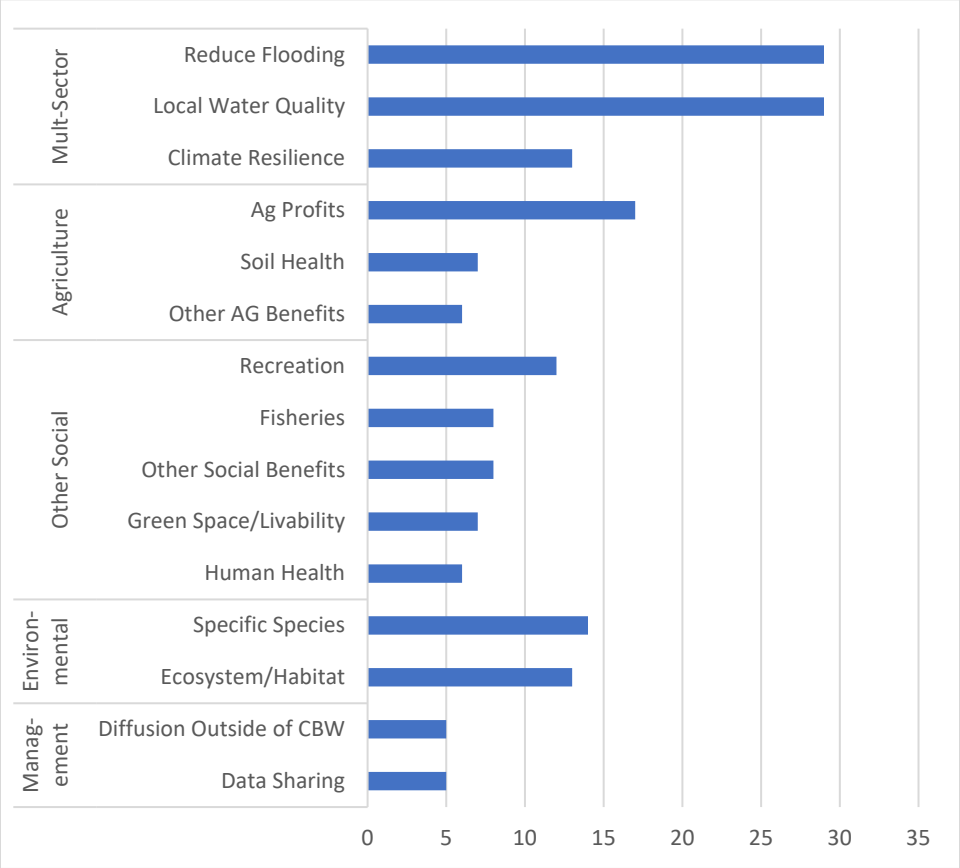


Figure 21: Co-benefits by Category and Number of Mentions. Most frequently mentioned co-benefits can provide value across multiple sectors, including both urban and rural communities. Agricultural co-benefits were also mentioned frequently, along with co-benefits for urban and recreational sectors. Environmental co-benefits included protection of specific species and of ecosystems/habitat generally. Interestingly, some managerial co-benefits were also mentioned.

The two most frequently mentioned co-benefits were reducing harm from flooding and improving local water quality. The former was associated with specific BMPs (see Sections 6 WIP Design and Sub-State Allocation Processes and 7 Implementation) while the latter was more often described as a potential effect of the CBP process as a whole. Although they do not appear to be directly related in any causal way, these two top co-benefits co-occurred more than any others (cc = 10, cf = 0.21). Several respondents mentioned a direct connection between Reduced Flooding and another frequently mentioned co-benefit, Climate Resilience (m = 13, cc = 5, cf = 0.14), but co-occurrence was the same with several other co-benefits and higher with Ecosystem/Habitat restoration (m = 13, cc = 8, cf = 0.24).

A much more common causal link was made between improved Local Water Quality and other ecological or economic co-benefits, though the connection was not always made explicit by the respondent. Ecological co-benefits included increased populations of charismatic species like black ducks, blue crabs, or the hellbender lizard (Specific Species, m = 14) and the restoration of ecosystems and habitat for wild animals (Ecosystem/Habitat, m = 13). Two economic co-benefits were also related to ecological benefits and improved local water quality: improved recreational opportunities (e.g., fishing, swimming, m = 12) and replenished commercial fisheries (e.g., blue crabs, oysters). Improved Human Health (m = 6) was also linked to improved water quality by some respondents.

Several other co-benefits were concentrated in either rural or urban areas. Increasing Profits in Agriculture was mentioned with moderate frequency (m = 17) as an added benefit from implementation of some BMPs, though respondents did not usually use the term "co-benefit" in this context. Improved Soil Health (m = 7) was also mentioned as an agricultural co-benefit of certain BMPs, along with various Other Agricultural Benefits (m = 6). On the urban/suburban side, Green Space and Livability (m = 7) was mentioned and sometimes was linked to improved Human Health (cc = 4, cf = 0.44). A range of other social benefits were also mentioned once or twice and so were aggregated into the category Other Social Benefits (m = 8).

The last set of co-benefits were associated with the institutions and design of the CBP. Five blocks mentioned improved data sharing within and among jurisdictions as a co-benefit of the program and five described the diffusion of CBP policy design to other watersheds as a co-benefit. There was not much overlap between these co-benefits and the others described above, but Diffusion did co-occur with Local Water Quality 3 times (cf = 0.10) and twice with Soil Health (cf = 0.20).

Table 4: Types of Co-benefits

Code	M	CC	M*	Description
Reduce Flooding	29	61	3	Respondent indicated reduced flooding was an important co-benefit of BMPs like stormwater management, wetlands and flood plain restoration
Local Water Quality	29	53	3	Respondent indicated that improved local water quality was an important co-benefit of CBP policies
Agricultural Profits	17	21	3	Respondent indicated that some BMPs provided co-benefits in the agricultural sector through reduced costs of production or direct transfers (i.e., payments for BMP implementation)
Specific Species*	14	39	2	Respondent indicated that increased populations of certain species was a co-benefit of CBP policies (through improved local water quality). Examples of

				species include: black ducks, blue crabs, brook trout, brown trout, elk, and the hellbender lizard.
Climate Resilience	13	23	1	Respondent indicated that improved resilience to climate change was an important co-benefit of some BMPs. Includes some cross over with flood mitigation plus more efficient use of freshwater, protection of groundwater, and other ecosystem services.
Ecosystem/Habitat*	13	37	2	Respondent indicated that restoration of ecosystems/ provision of habitat was a co-benefit of CBP policies, mainly through improved water quality (especially for seagrasses) but also as a direct results of some BMPs (i.e., wetlands restoration, pollinator gardens).
Recreation/Economic Value*	12	36	3	Respondent indicated that improvements in recreational opportunities or other economic values were a co-benefit of CBP policies, mainly through improved water quality. Types of recreation mentioned included: recreational fishing, swimming, hunting, enjoying natural beauty, and tourism generally. Improved real estate values are also included in this type of co-benefit.
Fisheries	8	26	1	Respondent indicated that more productive commercial fisheries were a co-benefit of CBP policies, mainly through improved water quality, habitat restoration, and related increases in populations of commercially valuable species like oysters and blue crabs.
Other Social Benefits*	8	26	1	Respondent indicated that CBP policies provide some other type of social benefit, including reducing harmful algal blooms, reducing other types of water pollution, providing employment, and providing infrastructure security.
Green Space/Livability*	7	28	2	Respondent indicated that certain BMPs can provide co-benefits by increasing access to green spaces and improving livability in urban or suburban areas. Includes reducing urban heat effects.
Soil Health	7	20	2	Respondent indicated that some BMPs provide additional benefits to agriculture by improving soil health.
Human Health	6	26	2	Respondent indicated that CBP policies could improve human health, usually through improved water quality but also through green space/livability or various other social co-benefits.
Other AG Benefits*	6	24	1	Respondent indicated that some BMPs could provide additional co-benefits to agriculture, including diversifying production or shifting to higher-value-added products (i.e., grass fed beef, fruits, nuts, flowers), improving cattle health, improving farm organization, and providing other benefits such as

				wind breaks.
Diffusion Outside of CBW	5	0	1	Respondent indicated that the science and modeling undertaken by the CBP provided co-benefits through diffusion outside of the Chesapeake Bay Watershed.
Data Sharing	5	10	1	Respondent indicated that improved data sharing within and among states and federal-level agencies was an added benefit of the CBP policy process.

#M = Number of mentions, #CC = Total Co-occurrences with other Challenges, M = Maximum mentions/interview*
** Heterogenous grouping of similar challenges after coding was completed. Merged to aid analysis and to produce groups > 5 blocks each.*

Methods Note: As with challenges and opportunities, most mentions of a given type of co-benefit only occur in one block per interview, but the same co-benefit is mentioned in two blocks/interview 13 times, and 3 blocks/interview 7 times.

4.4 Power Disconnects

Although not identified by respondents as such, power disconnects—or misalignments between incentives, understanding, and resources—were described by some as underlying challenges for governance in the watershed. We expected that disconnects between Bay states and non-Bay states would be important, but respondents also described disconnects between people who live on the Bay and those who do not (even within Bay states), rich communities and poor communities, and state and federal government agencies. We find it particularly interesting that many of the challenges described above stem from these power disconnects (e.g., Behavior Change, Lack of Funding/Resources, "Blame Game") and many of the opportunities described by respondents were thought to address challenges by narrowing disconnects (e.g., emphasizing Co-benefits, Engagement/Outreach).

The last group of codes described in this section was designed to evaluate perceptions of power disconnects in the Chesapeake Bay Watershed. A **power disconnect** occurs when the people who have incentives to conserve the environment lack the understanding and/or resources to do so and vice versa. This is different from inequalities, as a highly unequal system may still have a narrow power disconnect if well-off residents also have strong incentives to protect the environment. When we wrote the proposal for this project, we anticipated that there would be a transboundary power disconnect between upstream and downstream states. That is, while downstream states are most affected by nutrient pollution in the Bay, upstream states are important contributors to the amount of pollution entering the Bay. Indeed, Pennsylvania's section of the watershed contributes approximately 50% of the freshwater inflow to the Bay (via the Susquehanna River), even though PA does not boarder on the Bay. As shown in Figure 22, 24 blocks did mention this disconnect between Downstream (Virginia/Maryland) and Upstream (All Other) jurisdictions. However, respondents also described three other power disconnects, which we explain in more detail below.

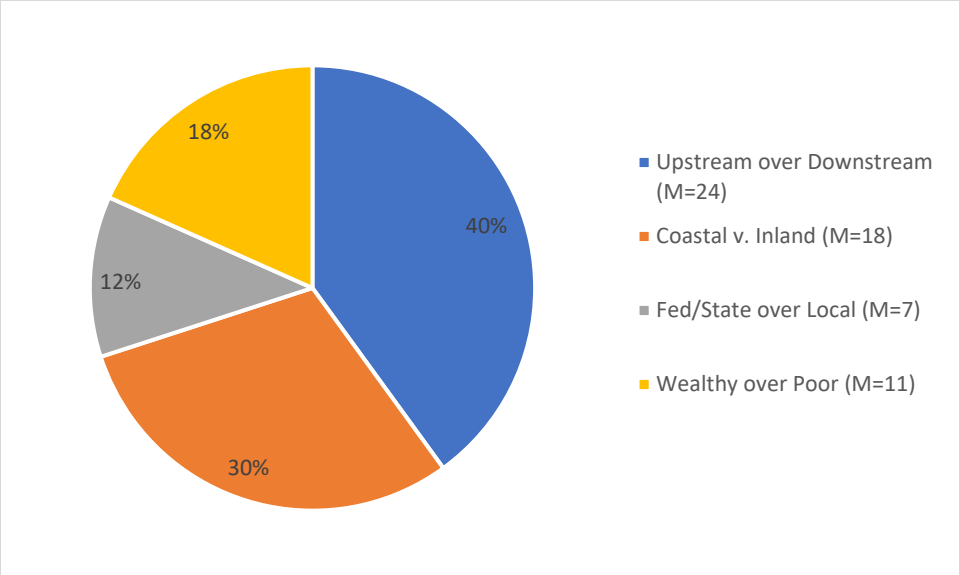


Figure 22: Power Disconnects by Number of Mentions

In an interesting twist on our expectations, 18 blocks mentioned power disconnects between people who live right on the Bay, and therefore have incentives to protect it, and those who live inland from the Bay, and therefore make the majority of decisions affecting nutrient pollution in the Bay. This disconnect was described within both Bay and Non-Bay states and was mentioned by respondents from 7 of the 9 jurisdictional groups. Interestingly, 50% of these Coastal-Inland blocks co-occur with the code for Co-benefit: Local Water Quality (cc = 9, cf = 0.24). In contrast, the Upstream-Downstream Disconnect only co-occurs with Local Water Quality in 3 blocks (cf = 0.06). The only other moderately high co-occurrences between Disconnects and Co-benefits also occur between the Coastal-Inland Disconnect and Commercial Fisheries (cc = 4, cf = 0.18), Human Health (cc = 4, cf = 0.20), and Recreation (cc = 4, cf = 0.20). Co-occurrences between Disconnects and other types of Opportunities are low except for between Coastal-Inland and Engagement/Outreach (cc = 7, cf = 0.10). Co-occurrences with Challenges are also low, except between Coastal-Inland and Challenge: Political Will (cc = 11, cf = 0.16).

This reflects a common theme in a number of responses: shifting goals to include local water-quality related co-benefits could narrow power disconnects and thereby build political will and encourage behavior change needed to improve CBP governance.

In many ways, the Federal/State Government—Local Government/Stakeholder Disconnect also supports the general theme that locals have ample incentives to try to improve local water quality but are denied resources to do this because state or federal authorities are focused on Bay water quality. With only 7 mentions, however, co-occurrences are quite low with most other variables.

The last disconnect is based on wealth and is somewhat different from the others in that, wealthy groups are described as having both the incentives and the resources to protect Bay water quality, but they displace the costs of that protection onto poorer communities in urban or rural areas. These blocks of text describe policies like Flush Taxes as regressive and may also mention disproportionate costs of nutrient management BMPs to small farmers vs. large industrial farming operations. In short, this disconnect is less about protecting the environment and more about determining who pays the cost for

protection. We will go into this in greater depth in Sections 8.2 Funding Process and 9.1 Factors and Actors.

Part II In-Depth Descriptions of Processes

The four sections in Part II of the report provide more detailed information on the setting of the TMDL and allocation of loading goals among state-level jurisdictions, the WIP Design and local-allocation process, the implementation of BMPs, and other processes like modeling and funding. It wraps up with a look at respondent perceptions of the future and suggestions for improving the system as a whole.

This part of the report provides more in-depth information on respondent perceptions of the various processes that make up the Chesapeake Bay water quality governance system, as well as some descriptions of the policy process as a whole. The first three sections analyze annotation code groups paralleling sections 2-4 of the interview instrument (see Appendix C: Annotated Instrument). We start by examining the process of setting the TMDL and related loading targets (TMDL/Allocation), with particular attention to sectoral and state-level aspects of the process. Then, Section 6 WIP Design and Sub-State Allocation Processes goes into detail about the WIP Design and sub-state allocation processes (WIP Design/Allocation), including comparisons across states, sectors, and types of BMPs. Respondents provided more information about this process than any of the others, so the analysis is quite lengthy. Next, Section 7 Implementation describes the process of implementing WIPs generally and different types of BMPs specifically.

The last two subsections analyze volunteered statements about other subprocesses. Since we did not ask about these processes specifically, we are cautious about interpreting results, but they do provide interesting areas for further research. Section 8 Other Processes covers the modeling process, funding process, lawsuits, nutrient trading, climate change, and the Conowingo Dam. The last four processes are associated with very few blocks of text, which increases our concerns about the representativeness of this information. Lastly, Section 9 System Process, analyses descriptive statements about the system as a whole that were not covered in Section 2 . This includes the types of factors that affect the system as a whole, actors with influence on the system as a whole, descriptions of the evolution of the system, and specific prescriptions for changes to the system (mainly to improve effectiveness).

5 Setting the TMDL and Allocating State-level Loading Targets

This section covers those statements made by respondents regarding the TMDL and Allocation processes in general, as well as state-specific roles in these processes.

As noted in Section E.1.1 Process Codes, most respondents treated the setting of the watershed loading target (TMDL) and allocation of loading targets among state-level jurisdictions (Allocation) as two parts of the same, iterative process. In fact, many respondents also saw WIP Design as an integral part of the TMDL/Allocation process, but we can still break down responses into meaningful variables based on descriptions of different portions of this recursive system-level process. This section covers those statements made by respondents regarding the TMDL and Allocation processes in general, as well as state-specific roles in these processes. As per our general confidentiality guidelines, in cases where respondents might be identified, we only report responses which were observed in 5 or more blocks of text. Because our colleagues from the CBPO already have extensive expertise on the technical aspects of

setting the TMDL and loading targets, we did not ask respondents for detailed descriptions of the use of the models but instead asked if they could provide a brief overview of the process and highlight any parts of the process that would not be easily observed through the CBP website and related documentation.

5.1 Setting/Revising TMDL Loading Goals

In general, respondents described this as an iterative process between setting loading goals for the watershed and allocating goals to state-level jurisdictions, but many were confused or indicated that they did not fully understand how the process works. Responses were also divided between technocratic depictions of the process which focused on following past precedents and political depictions of the process which focused on the effects of negotiations over allocation of state-level loading goals. Respondents also saw the process as top-down and were mainly critical, though some mentioned that the process was improving as of Phase III.

The total maximum daily load (TMDL) for the Chesapeake Bay was set by the EPA in 2010. Specifically, "the TMDL set Bay watershed limits of 185.9 million pounds of nitrogen, 12.5 million pounds of phosphorus, and 6.45 billion pounds of sediment per year. This equates to a 25 percent reduction in nitrogen, 24 percent reduction in phosphorus, and 20 percent reduction in sediment from the base year of 2009" (US EPA 2015). This was part of a comprehensive agreement among the EPA and state-level jurisdictions which established the current system of watershed implementation plans which was designed to achieve the TMDL target by 2025.

The TMDL target has not changed since 2010, but the resulting goals for external loading into the Bay—that is, how much the Partnership is expected to reduce nutrients moving from the land into the water to meet that overall goal in the Bay—can change for a number of reasons. First, loading goals may change if exogenous factors alter internal loading dynamics or increase the rate of external loading. Climate Change, Conowingo Dam, and several other problem multipliers described in Section 4.1 Challenges are examples of this type of factor. Second, if the loading goals are not reached in earlier phases (target = 60% of TMDL by 2017), then remaining amounts will be added to loading reduction required in subsequent phases (III, post 2025?). Third, because loading targets are determined by models, changes to the CBP modeling system can alter the external loading goals for the watershed (and therefore the goals for jurisdictions, as described in Section 6 WIP Design and Sub-State Allocation Processes). In fact, the first two factors will only affect loading goals once they are incorporated into the models, but we distinguish between changes on the ground that are reflected in the model vs. changes in data, methods, parameters or other factors that alter model outcomes without changes in the natural system.

Although we try to maintain a clear distinction here, we note that, among CBP participants, the term "TMDL" is used by some respondents to describe the 2025 target for the Bay, but others use it to describe the watershed-wide loading goals that are used as steppingstones to that target.[6] Further complicating matters, the EPA defines the TMDL as a "comprehensive pollution diet" and uses the term to describe the entire system of loading goals, WIPs, Milestones, etc. Admittedly, we also did not recognize the distinction between the very specific interpretation of the TMDL as the final target for

[6] Some respondents expressed a preference for the term "loading goal" over "loading target", especially in the contexts of allocation among states or among local-level jurisdictions.

2025 and the broader process of setting watershed-wide goals to reach that target when designing this survey. It was only through respondent reactions that we understood the term meant different things to different people. Thus, although we started by only asking about the setting of the TMDL, after the first 6 interviews, we revised the instrument to ask first about the setting of the TMDL in 2010 and then about the revision of the loading targets when Phase III WIPs were implemented in 2017. Where possible, we also asked about revisions to loading goals to account for Climate Change and the Conowingo Dam. However, since there were so few blocks for these periods, we generally report results in aggregate across all watershed-wide TMDL-related decisions (see Section E.1.2 Timeline Codes).

When asked to briefly describe how the TMDL was set or how those external loading targets for the watershed were revised, respondents did not have a shared understanding. As shown in Figure 23: Describing the timing of steps in the process of setting the TMDL/Adjusting loading targets for the entire watershed. Figure 23, 23% of blocks indicated that they did not know enough about these processes to answer the question. 34% described a process that was similar to the description provided on the CBP website, where the TMDL goals were designed to achieve a certain level of water quality in the Bay, then nutrient loading targets are set in order to meet those goals, and then state-level targets are set to meet that watershed-wide target. However, 18% indicated that state-level targets were set first, then states went through the WIP design process, and finally the total loading target was set for the entire watershed. The last group may provide some way to understand these divergent perspectives; 25% of blocks said that the process was an iterative and non-linear process of adjustment among the committees work on the Bay models, watershed-wide targets, and state-level targets/WIP design processes. If the process is indeed iterative, then the Bay First and State First responses may simply reflect different sides of a policy cycle.

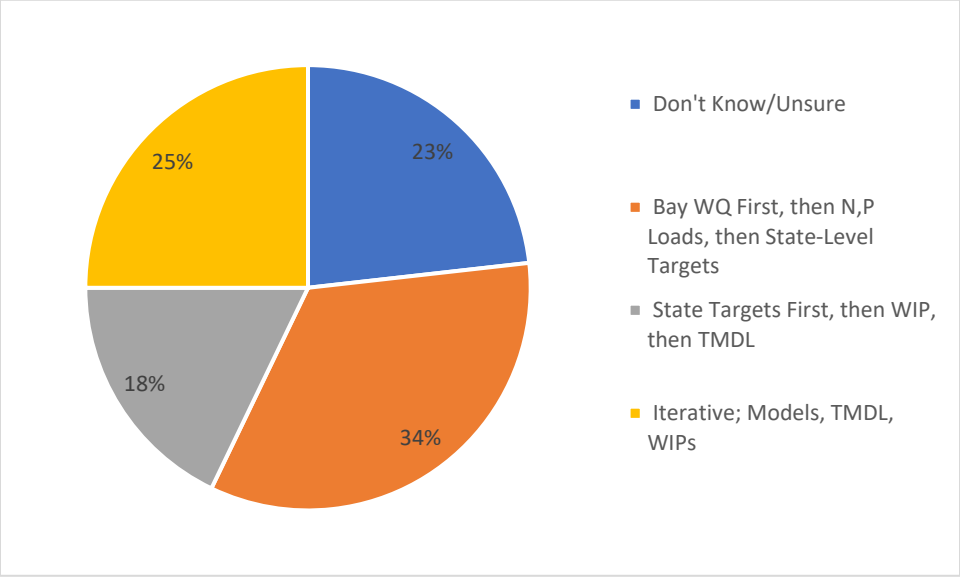


Figure 23: Describing the timing of steps in the process of setting the TMDL/Adjusting loading targets for the entire watershed. Since the TMDL/Allocation process is iterative, discrepancies between Bay Water Quality First… and State Targets First… may depend on the segment of the policy process the respondent was considering at the time. Nevertheless, results suggest there is considerable confusion about the process and a large number of respondents indicated that they did not understand it.

Looking at the co-occurrences, TMDL-Only process codes only co-occurred with “Don’t Know” (cc = 6) and “Bay First” (cc =5) blocks. In contrast, TMDL-Allocation process codes co-occurred with all of the

process timing variables shown in Figure 23, with 10 or greater co-occurrences for all variables except "don't know/unsure". This suggests to us that at least some of the variation in responses may be due to respondent's interpretation of the terms "TMDL" and "watershed-wide loading targets", with those who answered while thinking of the setting of the 2025 TMDL Goals more likely to answer that the process started with "Water quality in the Bay", while those who were thinking more about setting loading targets to reach that goal were more focused either on the distributional aspects of loading targets (i.e., "State targets first") or described it as an "Iterative" process. The last assertion is supported by high co-occurrence between "State targets first" blocks and blocks that identified Conflict over Allocation (among states) as an important factor in the TMDL Process (cc = 10).

Given the analysis above, it is odd that there is no clear pattern of co-occurrence among the processes timing blocks and timeline codes (especially Phase I vs. other Phases), and there are few strong patterns across respondent characteristics, either. There is marginal support for the hypothesis that longer time spent involved in CBP governance increases respondent confidence in their understanding of the TMDL process, as smaller percentages of respondents responded "I don't know" in the categories with 20 or more years of experience. Respondents with more than 30 years of experience were more likely to express the "Bay water quality first" opinion (41.67% of respondents in this EXTEMP category), whereas respondents with 10-30 years of experience were more likely to describe the process as "Iterative" (43.75% for 10-15 years, 33.33% for 15-20 and 20-25 years, 60% for 25-30 years). Compared to other groups, respondents with general roles of Tech/Modeler (33.33%) and Implementation (42.86%) were somewhat more likely to say that they "Don't know", but the only clear difference by role was that 48.72% of Regulators and 42.72% of Scientists indicated that the process was "Iterative". Regulators were also more likely than other groups to respond that the process started with "State Targets". Uncertainty was also higher among non-Bay jurisdictions (21.05% vs. 14.81% of responses in each category), but the biggest differences were on the timing variables, as show in Figure 24.

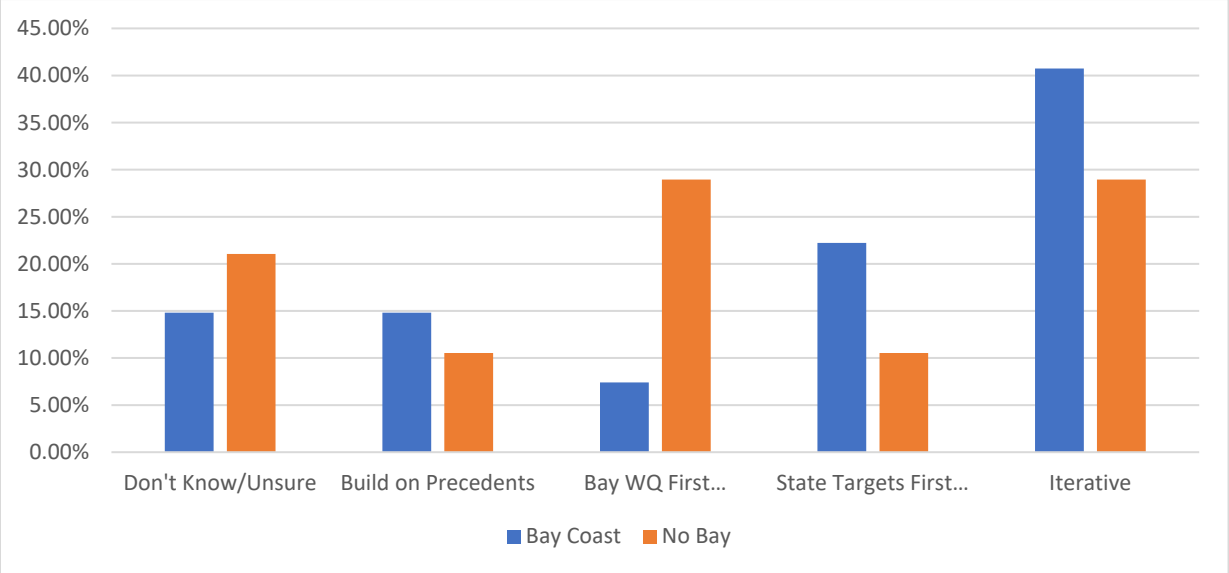


Figure 24 Descriptions of the TMDL Process by States with Coastal access to the Bay and those without. It is interesting that a larger percentage of blocks from Non-Bay states describe the process as starting with Bay Water Quality… while a larger percentage of blocks from Bay Coast states see it as starting with state loading goals. This may reflect transboundary disconnects between Bay Coast and Non-Bay states, but the number of observations is too small to draw strong conclusions.

In addition to these descriptions of the process, respondents identified two factors as important endogenous determinants of setting the TMDL/loading targets for the watershed as a whole: building on past precedent (m = 11) and conflict over allocation among states (m = 16). In this context, "past precedents" include the 2000 Bay Agreement and other legal precedents, various (failed) voluntary commitments to load reductions pre-2010 which triggered the TMDL, and negotiations over allocation criteria (the Equity Principles and E3 Scenarios) which concluded prior to the creation of the partnership in 2010. These annotations co-occurred most often with the technocratic frame for the TMDL Process (cc = 5, cf = 0.12) and mostly reflected a perception that many of the "big decisions" were already made in the past, and now setting the target loads is largely a matter of following those rules and "precedents". As might be expected, blocks in the Conflict over Allocation (among states) code group co-occurred more often with the political frame for the TMDL process (cc = 11, cf = 0.17), though co-occurrence with technocratic frames was also high (cc = 8, cf = 0.12). This group was more diverse and included both broad statements that allocation affected the setting of the TMDL and more specific statements about the ways that this occurred (often through voting procedures and modeling decisions). This will be discussed further in the next section.

Overall, setting the TMDL/loading targets for the watershed was seen as the most top-down component of the entire policy process. 21 blocks of text indicated that the process was largely top-down, including statements that engagement was performative and that there was minimal local input into the process. There were no other responses in this category, meaning that no blocks of text indicated that the TMDL process was even moderately bottom-up. The majority of blocks that mentioned the use of power in setting the TMDL ascribed it to the EPA (m = 14), with lower number of mentions for other actors like state-level jurisdictions/the Partnership (m = 8) and non-governmental organizations (m = 2). Some descriptions of the EPA's role describe federal overreach, but others describe them as primarily acting as facilitators for state-level agencies. Interestingly, technocrats were specifically mentioned as powerful actors in 7 blocks of text associated with setting of the TMDL.

Most value-based descriptions of the TMDL process (not including allocation) were critical (m = 10) and of these, 7 referred to problems with the use of CAST or other Bay Models. However, in 5 blocks of text, respondents indicated that they thought the TMDL process was improving, especially with Phase III, and that fairness was therefore increasing. Here again, it is difficult to disentangle these statements about setting the TMDL/loading targets for the watershed from the allocation of state-level loading targets which is described in the next section. Responses on the allocation process will provide more information on how this can be achieved.

5.2 Allocating State-Level Loading Goals

Statements specifically describing the process of allocating loading goals among state-level jurisdictions parallel findings in the previous section, insofar as most focus either on the use of CBP models or negotiations among states and the EPA. About 2/3 of statements about the legitimacy of the allocation process indicated that it was questionable, either because of concerns about the model or about negotiations among states. Responses from Non-Bay States were much more likely to question the legitimacy of the process compared to responses from Bay States.

As shown in Figure 25, on the surface, respondents were generally split about the process of load allocation among jurisdictions, though it should be noted that the two main perspectives were not

mutually exclusive (cc = 9, cf = 0.16). Just over half of the 68 blocks describing jurisdiction-level loading goals focused on the use of CAST or other Bay Models to estimate allocations. A few blocks mentioned that loading goals were first allocated based on sector or watershed (m = 4), but most were more focused on the effects that different modeling options had on the distribution of load. Slightly fewer blocks in this category (38%) described allocation as a process of negotiation among state-level jurisdictions and/or the EPA. Technically, both of these responses are accurate to a certain extent. Initial distributions of loads are carried out using CAST (Phase III) or other versions of the Watershed Model (Phases I and II), but the Principals' Staff Committee (PSC) and Management Board (MB) guide the modeling process, and the Executive Council (EC) has the final say on both watershed-wide and jurisdiction-specific loading targets. Furthermore, state-level jurisdictions have input at most stages of the modeling process, via the Water Quality Goal Implementation Team (WQGIT), its subsidiary work groups (i.e., Modeling Work Group), various technical subcommittees, and other committees like the Scientific, Technical Assessment and Reporting (STAR) team (see Section 8.1 Modeling Process).

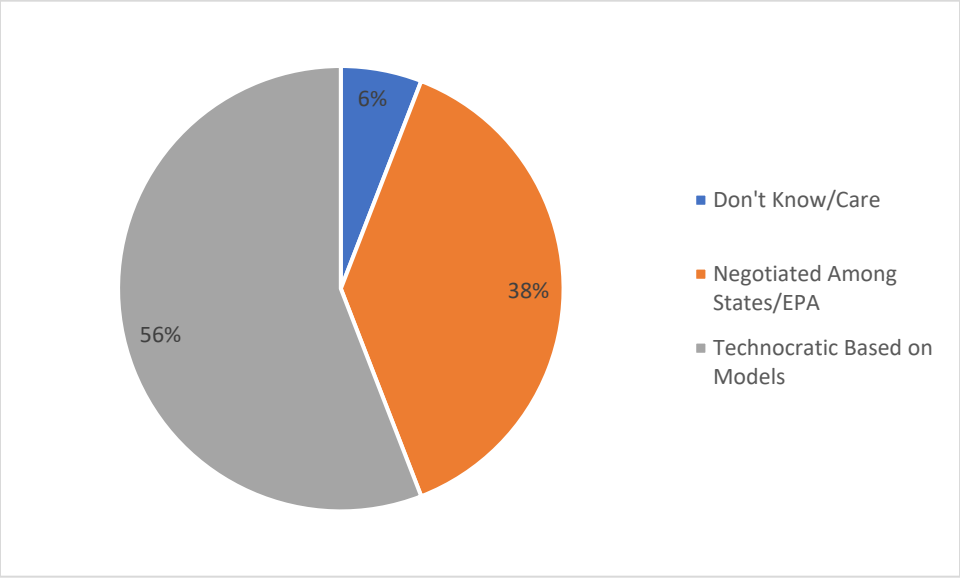


Figure 25: Descriptions of Allocation Process by Type. Descriptions largely fell into two overlapping categories. Many said that the allocation process is technocratic, based largely on results from the CBP Models but a large number also emphasized negotiations among states and the EPA. Some indicated that the process starts with CBP Models and then is finalized through negotiations. The lower percentage of Don't Know/Care answers reflects attrition, since those who answered that they did not know about the TMDL process as a whole were not asked this question.

In fact, the most common specific assertion in the TMDL/Allocation section of the interviews was that the details of the models can make a big difference in the allocation of loading targets among states (m = 34). Only a few respondents made direct links between the importance of the models and the legitimacy of the allocation process. These were split 50/50 between opposing beliefs; half said that more detailed models improved legitimacy and half said that more detailed models reduced legitimacy. However, there was agreement that improved transparency in the modeling process increased the legitimacy of the allocation process. Again, statements specific to this portion of the policy process were few, so we return to this issue for the entire system in Section 8.1 Modeling Process.

Overall, 31.58% (m = 12) of responses that addressed the question of the legitimacy of state-level allocations indicated that the process was equitable among jurisdictions while 68.42% (m = 26) indicated the process was questionable. Note that these responses were volunteered as the respondent described

the allocation process and so reflect the perceptions of those who were concerned enough about the issue of legitimacy to mention it to us without being asked. These numbers might change if we had asked every respondent to evaluate the legitimacy of the process. However, this clearly indicates that future research into perceptions of legitimacy could be interesting. Negotiations among states were mentioned most often as rationale to question the allocation process. Some respondents also mentioned factors like the modeling issues described above while others questioned the E3 Criteria themselves. For the latter, failure to incorporate the costs of nutrient reductions, the full spectrum of activities that could reduce nutrient pollution (i.e., climate mitigation), and resources available for nutrient reduction as criteria were all mentioned as factors that reduced the legitimacy of the allocation process. As shown in Figure 26, there is a clear delineation in perceptions of legitimacy between respondents from jurisdictions that are right on the Bay (Bay Coast) and those that are not (No Bay), with the latter less convinced that the process is equitable.

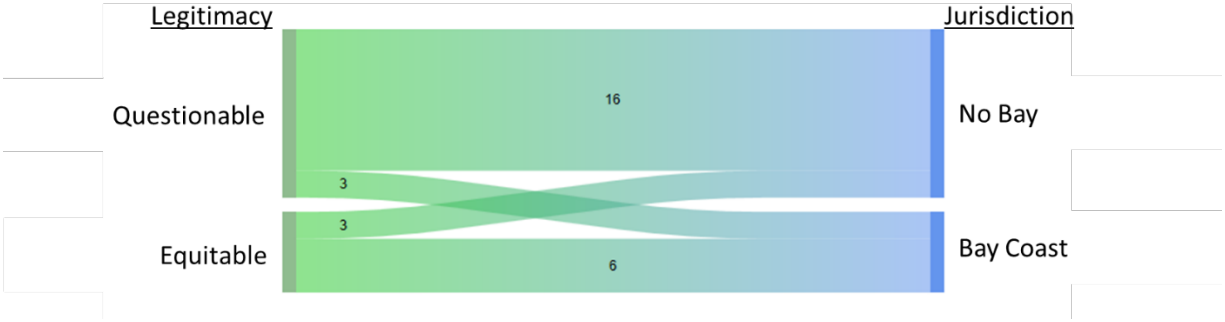


Figure 26: Perceptions of the Legitimacy of the State-Level Allocation Process by Bay/No Bay. Statements from Bay State respondents were more likely to describe allocation of loading goals among states as equitable, while statements from Non-Bay states were more likely to describe the process as questionable or not legitimate.

5.3 Influence of States and Other Actors on the TMDL/Allocation Process

Results in this section primarily describe the perceived behavior of state-level jurisdictions because there were very few statements about other actors in the TMDL/Allocation Process. Bay States (Maryland and Virginia) were generally perceived to be leaders in the TMDL/Allocation Process while Non-Bay States (especially Pennsylvania, Maryland, and West Virginia) were described more often as laggards. That said, all states asserted that they are already paying their "fair share" for water quality governance in the Bay, which suggests that the CBP may be approaching an impasse due not to transboundary power disconnects per se, but rather lack of political will in both Bay and Non-Bay states. Responses from multiple states acknowledged that Pennsylvania faced greater difficulties meeting its loading goals compared to other states, but most did not view this as rationale to change allocation procedures. West Virginia and New York have received concessions on their loading in the past and will continue to defend those concessions in future.

Since we are trying to model the policy process, it is important understand the roles played by different types of actors. As shown in Figure 27, Non-Bay states, particularly Pennsylvania, New York, West Virginia, and the generic "Upstream/Headwater" states were mentioned more frequently than Bay Coastal states, including Virginia, Maryland, and "Downstream/Bay" States. Delaware and Washington, DC were mentioned much less often than other jurisdictions, and, in spite of its central role, the EPA/CBP was also not mentioned very often. Of the state-specific comments, 7 asserted that coastal Bay States (Maryland and Virginia) had much more to gain from the CBP than Non-Bay States, mainly due to

the CBP focus on Bay water quality instead of local water quality. This is a common theme that will come up again in other sections. The rest of this section focuses on the perceived behavior of these actors in setting and allocation of loading goals.

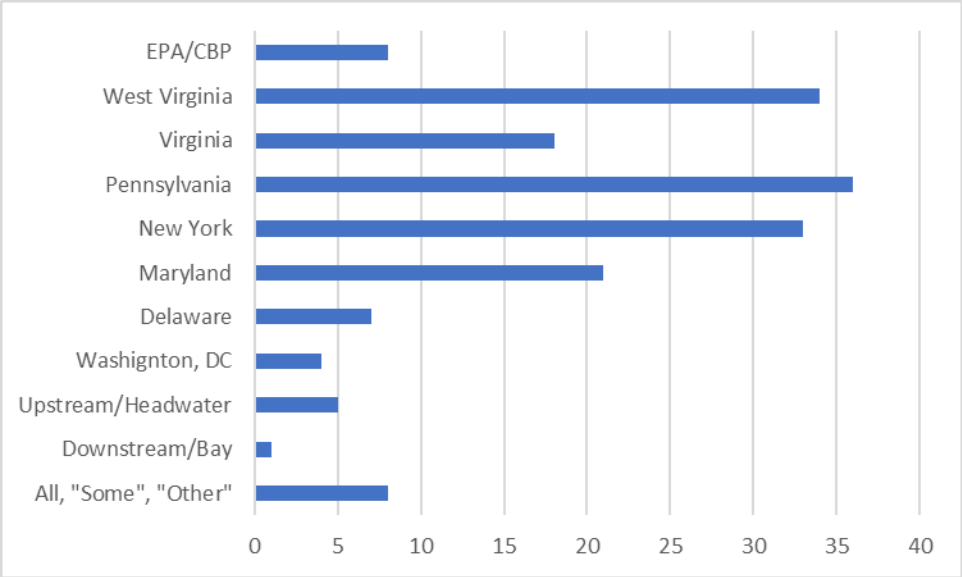


Figure 27: Number of times specific actors were mentioned in the TMDL/Allocation section of the interview. Non-Bay states, particularly Pennsylvania, New York, West Virginia, and the generic "Upstream/Headwater" states were mentioned more frequently than Bay Coastal states, including Virginia, Maryland, and "Downstream/Bay" States. Delaware and Washington, DC were mentioned much less often than other jurisdictions, and, in spite of its central role, the EPA/CBP was also not mentioned very often.

Respondents frequently described the roles of states in the negotiating process and, in some cases, tried to explain why certain states behaved in certain ways. Of the 46 statements about state roles, 21 blocks indicated that at least one state was acting as a leader while 23 indicated that at least one state was acting as a laggard and 1 indicated that one or more states were between a leader and a laggard. The terms "leader" and "laggard" were not necessarily used by respondents (especially laggard). These are terms that are used frequently when analyzing international negotiations. Leader states can be defined as states that push for more stringent environmental protection. Laggard states resist more stringent environmental protection, often because of concerns about the allocation of costs and benefits of an agreement rather than lack of concern for the environment per se. From the discussion below it is clear that respondents had somewhat different perceptions of "leading" and "lagging", with some statements asserting that leadership entailed goal attainment (meeting the state's load reduction goals). Therefore, we differentiate between leadership in the negotiation process and leadership on goal attainment in the analysis below.

As might be expected based on their perceived incentives, coastal Bay States were described as leaders more often than laggards, with Maryland described as a leader in 11 blocks and Virginia described as a leader in 7 (see Figure 28). West Virginia was also mentioned as a leader in 6 blocks, but these referred to leadership on goal attainment rather than within TMDL negotiations. New York and Washington, DC were also mentioned as leaders on goal attainment in fewer than 5 blocks. Pennsylvania was described as a leader once, particularly in that they "push downstream states to do more". In contrast, upstream/headwater states were described as laggards much more often with Pennsylvania (m = 9), New York (m = 11), and West Virginia (m = 7) mentioned more often than Delaware or DC. In one block,

Delaware was described as being “in the middle”, leading on some areas and lagging on others. Nevertheless, Maryland and Virginia were described as laggards on goal attainment (m = 3 and m = 1 respectively). This issue will come up again in Section 6.2 WIP Design State Variation.

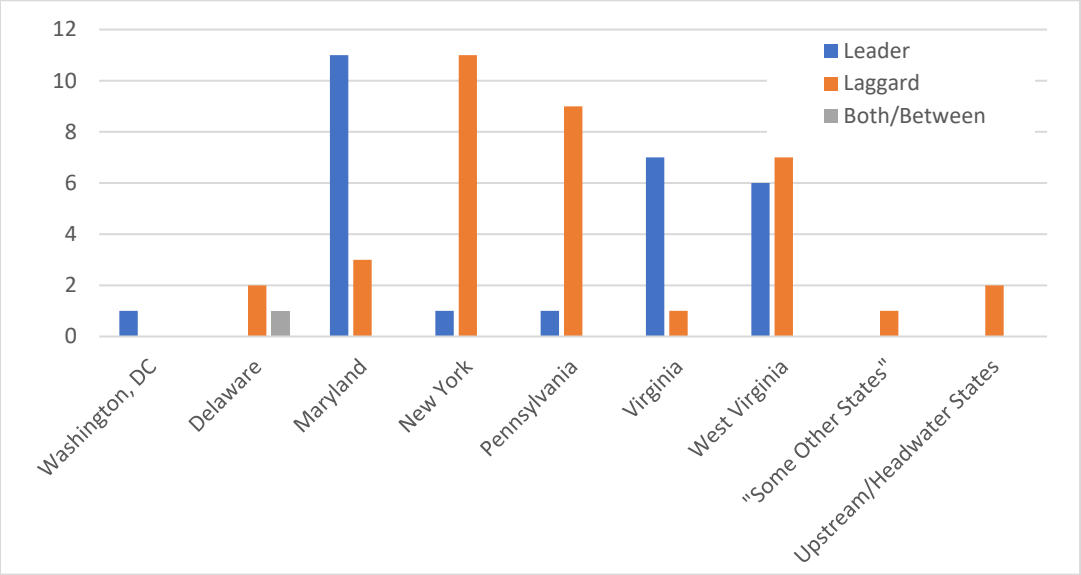


Figure 28: Role of States in TMDL/Allocation Process (# Mentions). Bay States Maryland and Virginia were described as leaders in the TMDL process, but a few statements indicated that they are laggards on goal attainment. New York and Pennsylvania were most often described as laggards in the TMDL process, but New York was also described as a leader on goal attainment and PA was described as a leader on equity among states. West Virginia was frequently described as a leader on goal attainment but as a laggard in the TMDL process. Mentions are low for other states, but it is interesting that Delaware was described as sometimes leading and sometimes lagging.

Although upstream states were only described as laggards on goal attainment in a few blocks, a number of statements indicated that they were not willing to pay more than their “fair share” or to pay for other states’ pollution (m = 7). On the other hand, 6 blocks indicated that Virginia and Maryland had already made concessions to non-Bay states to get agreement on the TMDL.

As might be expected, respondents from a given state tended to evaluate their state’s role in the TMDL/Allocation process more favorably. However, all agreed that Bay states were not likely to take on higher loads or otherwise make further concessions to non-Bay states in future. The only difference among respondents by state was their interpretation of the reasonableness of their state’s approach, which linked back to perceptions of questionable legitimacy in the distribution of load among states.

Another area of agreement was that, of the 7 jurisdictions, West Virginia and New York fought for (m = 10) and/or received (m = 21) concessions on their load allocations for each Phase. Responses varied as to whether these concessions were made by Bay states or by the EPA. Most of the blocks of text relating to concessions for New York and West Virginia (m = 18) did not express any thoughts about the fairness or legitimacy of these load reductions. Three blocks indicated that these concessions were fair, while two indicated that the concessions were not fair. All agreed that concessions to New York and West Virginia were small relative to the overall TMDL and that this made it easier for the EPA and other states to accept them. A few detailed explanations revealed that concessions were made early in the TMDL process and that representatives from each state defended these initial precedents in subsequent phases. Like the rest of the process, concessions to these states were determined in part through

technical decisions about how the Bay Models would be applied, including which baseline should be used and what should/should not be counted against a state's loading goals. There were also indications that the EPA was able to make up for these relatively small concessions by "banking" unexpectedly large reductions in atmospheric loading due to changes in the Clean Air Act and related legislation.

In contrast, Pennsylvania has a large proportion of the load, but responses indicated that this state did not receive concessions, at least in terms of their loading targets. This is interesting because theories on transboundary pollution would suggest that Pennsylvania would have considerable bargaining power as an upstream state whose collaboration is necessary for the success of the TMDL. It may be that this power was mediated by federal regulations (i.e., the Clean Water Act and the mandate provided by Barak Obama's 2009 Executive Order which set the TMDL in motion), but political limitations on Bay-State willingness to pay also play a role. As noted above, Coastal Bay States are not willing to pay more than their "fair share" but, in addition, side payments or other transfers of resources from one state to another (i.e., through nutrient trading) were considered politically unpopular in five blocks of text.

Interestingly, respondents from multiple states and the federal and independent groups recognized that Pennsylvania has a more difficult job meeting its loading goals (m = 16) when compared to other states. Reasons for this difficulty included the large size of the load, the fact that more of the load comes from agriculture, lack of political will and leadership within the state, political structures like townships which fragment decision making and resources, and exogenous party politics (i.e., "small government" ideologies). Thus, there was widespread sympathy for the difficulties faced by Pennsylvania, but this did not translate into motivation to change the state's loading targets/make concessions except among respondents from PA and some from the Multi/Independent category. Although loading reduction is difficult for all jurisdictions, difficulty was only mentioned as a reason for not meeting loading goals for Pennsylvania and Delaware (m = 2). That said, failure to consider costs in the E3 Scenarios and related technical mechanisms for loading distribution was mentioned as a problem for the entire CBP system (e.g., cc = 3 w/ Attainment = Not Effective), so we will return to this issue in Section 9 System Process.

Last but not least, 9 blocks mentioned that state-level jurisdictions played a watchdog role on the model in order to ensure fairness in the process of load allocation among states. This includes identifying errors in the model design and specifications, but also keeping an eye out for manipulations of the models that might benefit some states to the detriment of others. We will delve into this more in Section 8.1 Modeling Process.

Methods Note: When coding, we divided actor-specific annotations into a number of code groups to assess the meanings of each statement. Most of these groups were too small to report individually and could not be broken down by actor while also maintaining confidentiality. Thus, the results above describe only a sub-set of comments about specific actors in the TMDL/Allocation process.

6 WIP Design and Sub-State Allocation Processes

This section covers those statements made by respondents regarding the WIP Design and Local-Level Allocation processes in general, as well as state-specific processes. Note that respondents did not always view WIP Design and TMDL/Allocation as separate processes, because states must go through the WIP Design process to estimate the costs of a given load allocation, which then may affect their stance in negotiations. Thus, this division is somewhat artificial and, in reality, there is iteration between the TMDL/Allocation process and the WIP Design/Allocation process.

Like the TMDL and state-level allocation, respondents rarely differentiated between the process of designing Watershed Implementation Plans and the allocation of loading goals among sub-state jurisdictions. At this level, too, design and allocation are generally viewed as a single if iterated or recursive process. However, since we asked about these processes separately, we are able to pull apart factors affecting each here. While the TMDL and state-level allocation processes are primarily viewed as interactions among states and the federal government, WIP Design/Allocation more intensively involves state governments, local governing entities, and stakeholders. Because WIP Design/Allocation occurs differently across state-level jurisdictions, respondents usually reported only their home-jurisdiction for this section. Some independent and federal respondents had knowledge of WIP Design/Allocation across multiple states and so we coded their description of each state separately.

We report on aggregate results for all states first and then provide breakdowns by state where possible given confidentiality restrictions. A few respondents also had little or no knowledge of the WIP design process, usually because their main area of expertise was in the setting of the TMDL and state-level loading requirements. Overall, however, respondents were able to provide much more detailed information about the WIP Design/Allocation Process than they were for the TMDL/Allocation process. Unless stated otherwise, all information in this section applies to the Phase III WIPS.

6.1 WIP Design Aggregated

Although there are parallels between responses about the TMDL/Allocation Process and WIP Design/Allocation Process (i.e., importance of Building on the Past, Details of the Models), many other factors were described, most of which touched on some aspect of the costs, feasibility, and popularity of plan options. The WIP Design process was also viewed as more bottom-up/less top-down, which supports findings in Section 2 Process Frames, though more statements evaluated stakeholder engagement as performative rather than reflexive. Both politicization and critiques of the modeling process were also raised as impediments to effective WIP Design. ***Most importantly, a large number of statements indicated that WIP Design produced plans that looked like they would achieve states' loading goals based on CAST, but which were not or could not be implemented effectively.*** *This problem of "paper WIPs" negatively impacted perceptions of the legitimacy of watershed governance generally.*

Looking at the factors that respondents identified as important in the WIP Design Process aggregated across states, the first thing to note is that respondents described a much wider range of factors for this stage of the process than they did for TMDL/Allocation. Produced by each state-level jurisdiction in each phase of the TMDL, WIPs generally consists of plans to implement best management practices (BMPs) which are measures that reduce nutrient pollution and therefore should reduce the state's load into the Bay. BMPs usually vary by sector, location, and similar factors. For instance, planting cover crops to absorb nutrients and prevent erosion in fallow periods is an important BMP for the agriculture sector,

while building or upgrading treatment plants can be considered a BMP for the wastewater sector. Planting and maintaining forest buffers can be BMPs for multiple sectors, but these practices are limited to areas with high runoff potential. There is some fuzziness around the idea of a BMP and, particularly, whether regulated reduction mechanisms, like MS4 permits for stormwater count as BMPs. In general, we use the term BMP in its broadest sense, though, again, respondents tended to vary in their own definitions.

When designing their WIPs, states create "decks" of BMPs, which are then run through scenario tools (Phase II) or CAST (Phase III) to provide estimates of the effects on the state's loading relative to their overall target. These decks can only include BMPs that are recognized by the CBP accounting system. We will go into why specific BMPs are selected for a WIP and how the models are used for WIP Design in Sections 6.5 BMP Selection Aggregated and 6.6 BMP Selection by Type of BMP. Here, we review the general factors affecting the WIP Design process and then break them down by state-level jurisdiction.

As shown in Figure 29, Building on Past Precedent is by far the most important set of factors affecting WIP Design, whether blocks indicated that current WIPs Built on Previous WIPs (m = 46), were based on experience of what did and did not work in the past (What Works/Doesn't (Historically); m = 38), or started with their existing set of BMPs and related programs (Start with Existing; m = 11). This focus on learning from the past is also highlighted in the 9 blocks of text which indicated that Novelty and Innovation were NOT important or, indeed, desirable when designing WIPs; respondents preferred proven over unproven BMPs. Most other factors with more than 5 mentions captured some aspect of either effectiveness or feasibility or both. The Costs or Cost Effectiveness of BMPs was described as important in 15 blocks, although a few blocks indicated that costs were not important. Effectiveness in reducing load—specifically in terms of the amount of load reduction that would be accredited in the Bay Models—was also mentioned separately from cost effectiveness (Effectiveness/Meet Loading Goals; m = 12).

Various aspects of Feasibility (Tech/Econ/Pol) were mentioned as a factor in their own right (m = 14), but many other aspects of feasibility were described as well. Technocratic components of feasibility included meeting Legal Requirements (m = 5) and the Availability of Funding/Resources (m = 6). Political factors rested on Stakeholder Buy-in (m = 7), Popularity (m = 4), and Local Input (m = 15), which may seem similar but differ in that decision makers could have their own perceptions of stakeholder acceptance without direct local input and that local input itself may not lead to stakeholder acceptance. Ease of Implementation (m = 5) is another important component of feasibility that is related to costs, and which may increase stakeholder acceptance. Similarly, the presence of Co-benefits (m = 9), or benefits that are additional to the reduction of nutrient pollution, were described as desirable because they may help to reduce the net costs of BMPs or may make BMPs more acceptable to stakeholders, thereby increasing their political feasibility. A number of other factors were mentioned fewer than 5 times, including the willingness of the public to accept BMPs (as opposed to specific stakeholder groups), how easy it is to verify the BMPs, and remaining capacity to implement the BMPs. Equity and environmental justice were only mentioned as a factor in one block of text.

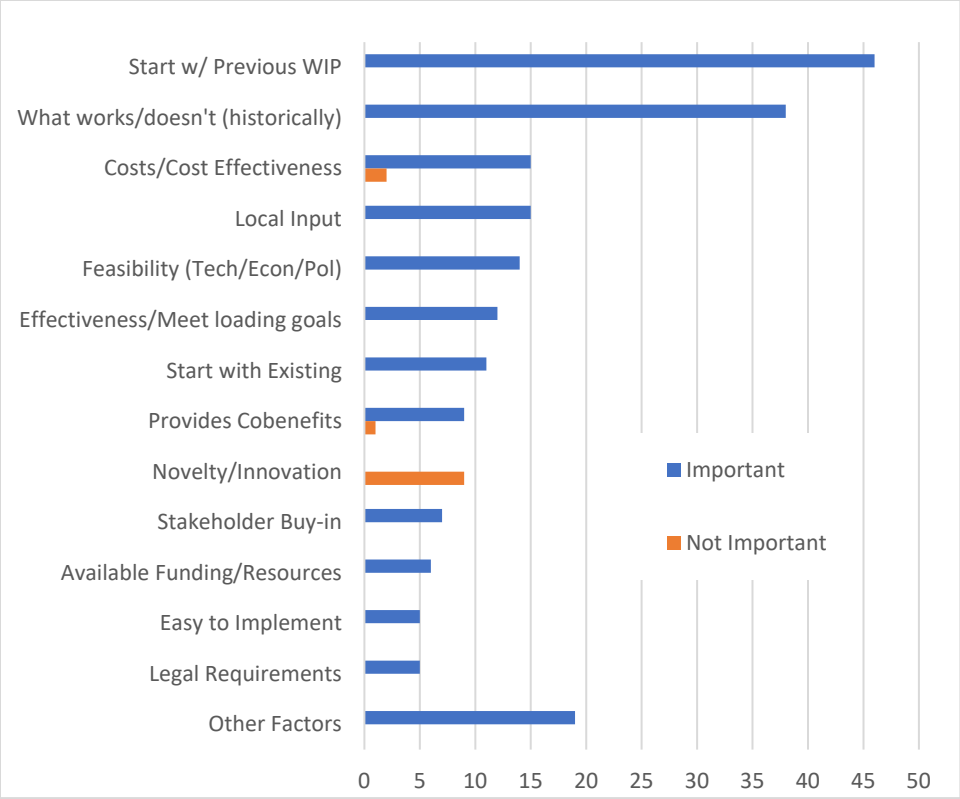


Figure 29: Factors Influencing WIP Design (# Mentions). Learning from the past is clearly important in a number of these factors, as are various aspects of feasibility, costs, effectiveness, and popularity.

Overall, the WIP design process was seen as less top-down/more bottom-up than the TMDL process, with "top" usually indicating state-level government and "bottom" indicating local governments or stakeholders. As shown in Figure 30, mentions that the process was Top Down/More Top Down (m = 26) were fewer than mentions that WIP design was Bottom Up/More Bottom Up (m = 31), with a small subset that indicated the process was both Bottom Up and Top Down (m = 9). The latter usually meant that the respondent saw a degree of iteration or back and forth between state and local actors. Note that some blocks of text mentioned that the design process was Top Down with Local Input (m = 17) and Bottom Up w/ State Override (m = 4), further indicating some back and forth between levels, though in both the state-level government was thought to have the final say.

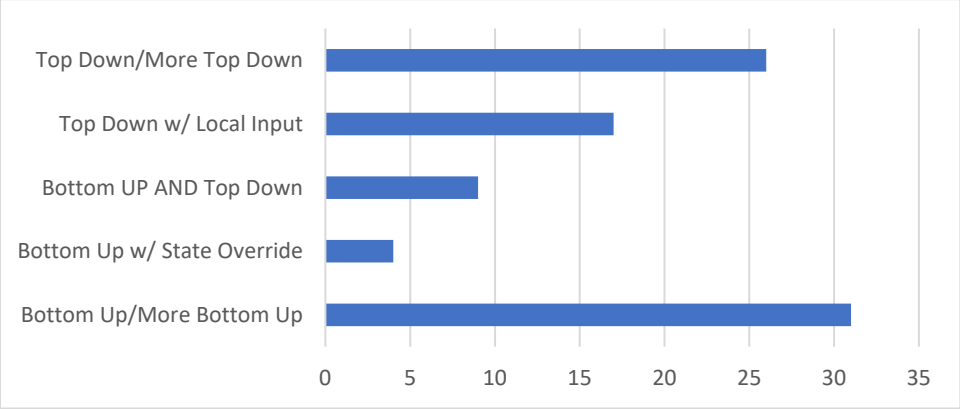


Figure 30: Hierarchy in WIP Design Process (# Mentions). Although there are still many statements describing the WIP Design Process as Top Down, there are also many that describe it as Bottom Up or some mix of Top Down and Bottom Up.

Percieved levels of stakeholder engagement were generally high, with engagement mentioned generically in 24 blocks (see Figure 31). Some responses were more specific, however, indicating that engagement was performative (stakeholder input did not change state decisions; m = 11) or reflexive (stakeholder input was taken into account; m = 5). Others indicating that the level of engagement was increasing (m = 8) or decreasing (m = 4) in the state whose WIP Design was being described.

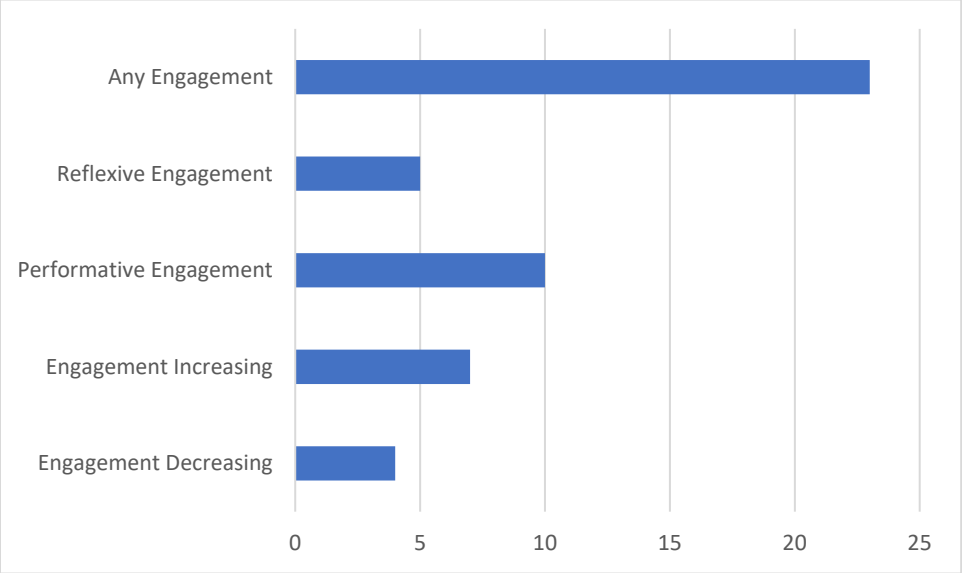


Figure 31: Stakeholder Engagement in WIP Design (# Mentions). Engagement was mentioned fairly often, but in 10 statements it was described as performative (stakeholders were invited to provide input but decision makers did not act on their suggestions). Statements that engagement was "increasing" were slightly higher than statements that engagement was "decreasing", but only 5 statements described reflexive engagement, where decision-makers actually act on stakeholder input.

Although seen as less top-down, and more about interactions between the state and local-levels, 12 blocks did mention that the federal government (mostly the EPA/CBP) had some influence over the WIP Design process. This influence was represented as facilitation, negotiation, or providing incentives to states. In 5 blocks, respondents mentioned that the federal government tried to influence WIP design but was ineffective. Other groups described as having influence over the WIP Design process are shown in Figure 32. Note that farmers were most frequently mentioned in this category (m = 14), followed by the federal government, experts, NGOs, other states, and generic stakeholders. The prominent role of farmers is not surprising, given that agriculture is a major contributor to nutrient pollution in the water shed and that currently agricultural BMPs can only be implemented voluntarily by farmers. In fact, 9 blocks in the WIP Design section mentioned that fear of a transition to regulation requiring mandatory BMP implementation in agriculture was an important incentive for farmers to participate.

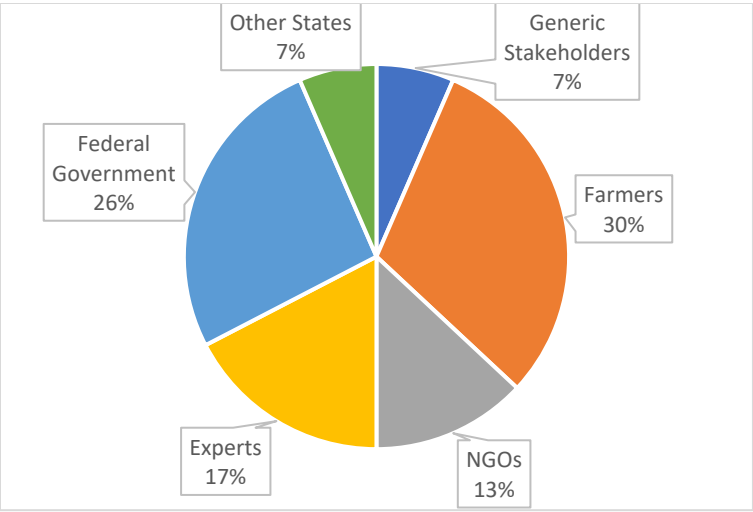


Figure 32: Groups that Influence WIP Design. Farmers and the federal government are mentioned most often as having influence over WIP Design, followed by technical experts, NGOs, generic stakeholders, and other states. It is interesting that state agencies were not mentioned at all, but this is likely because their central role was taken for granted by respondents.

A few other political factors were mentioned when respondents described the WIP Design process. These were generally considered roadblocks to effective WIP Design. First, 6 blocks of text mentioned that political contestation and lack of trust in government hindered the WIP Design process, usually by disrupting or undermining attempts at stakeholder engagement. In other words, stakeholders' political ideologies or lack of trust in government might prevent them from participating, or groups without a stake might try to disrupt the engagement process due to anti-government ideologies. Second, general contestation over loading goals or the division of resources within the state was mentioned as a roadblock to WIP Design in 6 blocks. These included concerns that resources were only going to those local government entities that were already well-endowed, minimizing the ability of other entities to influence WIP Design. Third, 5 blocks of text mentioned that WIP Design became more difficult from one Phase to another, with either increases in load reduction requirements or decreases in the amount of load reduction accredited to important BMPs. These changes were viewed as undermining stakeholder buy-in and reducing local governments' willingness to cooperate in WIP Design.

When asked specifically about the use of CAST or previous versions of the Bay Models, a majority of responses (m = 29) indicated that CAST was central to the WIP Design process. Only three blocks of text indicated the state used CAST, but it was not a central component in the design process. In addition, 8 blocks indicated that stakeholders (NGOs, local governments) also used CAST during the design process, with technical support from the state agency. A similar number of blocks (m = 9) indicated that stakeholders did not use CAST as part of the design process and a few of these stated that stakeholders did not understand CAST or how it was used for WIP Design. A number of blocks of text were critical of CAST, including problems with general specifications, parameterization, and transparency (m = 17) and parameterization of costs specifically (m = 8). On the other hand, 5 blocks of text indicated that CAST is better for WIP Design than previous models.

We wrap up this section with a summary of statements related to the legitimacy of the WIP Design process. **In particular, 44 blocks of text indicated that WIP Design was a "paper process", indicating that states were simply trying to design a plan that would allow them to meet their loading goals in CAST, but which could or would not be implemented effectively (see Figure 33). Of these, 41%**

indicated that the WIP Design process was not legitimate, 9% indicated that the process was flawed but the "best the state can do", and 29% did not comment on legitimacy as it related to "paper WIPs". Only 12 blocks indicated that the WIP Design was not a paper process, with a several of these stating that this increased legitimacy even though the end result was a WIP that would not allow the state to reach its loading goals. This reflects a larger conflict among state-level jurisdictions which will be discussed in later sections. It could also provide a testable hypothesis, given data on state-level implementation of their WIPs.

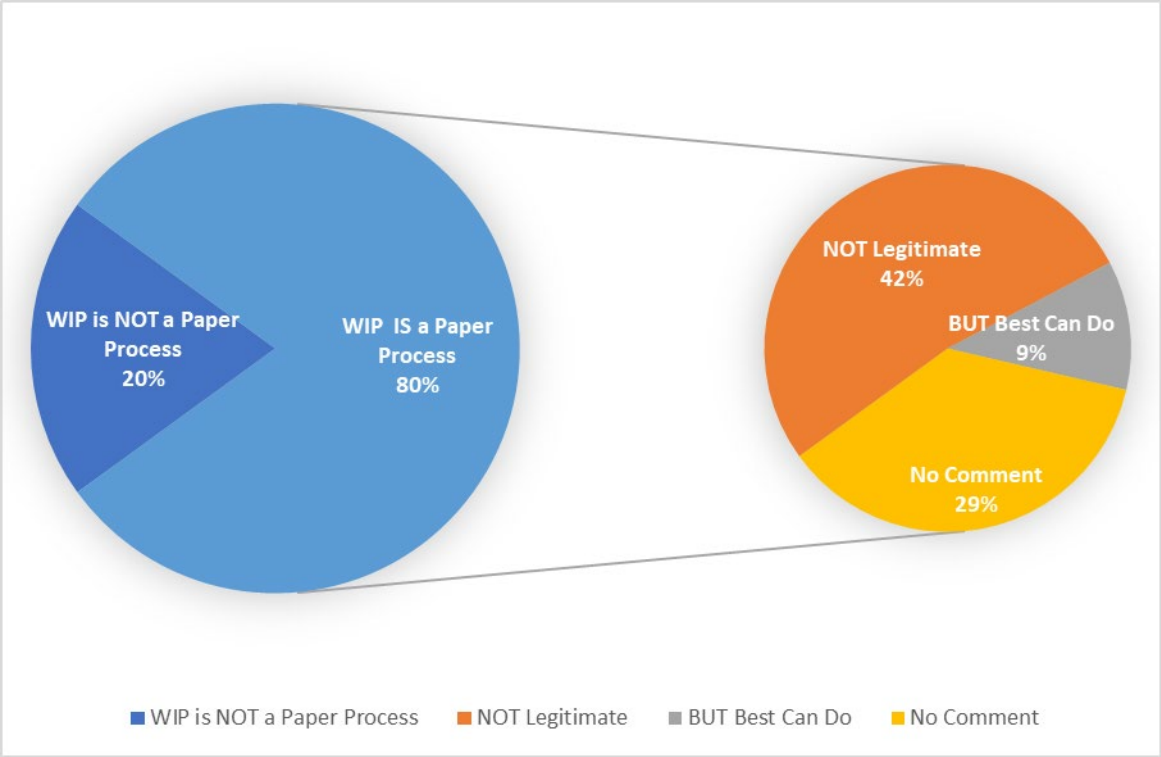


Figure 33: Break down of responses discussing whether or not the WIP Design is a Paper Process and how that affects legitimacy. 44 statements indicated that WIPs were designed to meet loading goals in CAST but could not or would not actually be implemented on the ground (Paper Process; 80% of all statements on this topic). Of these, 42% indicated that this negated the Legitimacy of the CBP, 9% indicated that this was the Best We Can Do given constraints on implementation, and 29% made no comment relating paper WIPs to CBP legitimacy.

Methods Note: As with other sections, results here may be affected by respondents who make the same statement in multiple blocks of text. Most respondents who mentioned that WIPs either were or were not a Paper Process did so only once per interview. However, respondents were more likely to mention that a WIP was a Paper Process more than once per interview (M* = 5 compared to M* = 3), which may skew results somewhat. Still, 25 respondents indicated that a WIP was a Paper Process at least once compared to 9 who indicated that a WIP was not a Paper Process at least once.

6.2 WIP Design State Variation

Some aggregate patterns hold across a majority of states. CAST remained an important or useful tool for WIP Design in all States, though also one that was criticized. Learning from the past also remained important across the board, though statements about Pennsylvania indicated that they were less influenced by past WIPs and more influenced by recent implementation. This reinforces a key finding that there is path dependence in the WIP Design process and so we anticipate problems such as autocorrelation and heteroskedasticity in any statistical analyses of WIP content. Respondents described Maryland, DC, and Virginia WIP Design as more Top-Down with less Reflexive Engagement, while New York, Pennsylvania, and West Virginia processes were described as more Bottom Up with higher levels of engagement, including Reflexive Engagement in New York and PA. Delaware was described as both top-down and bottom-up but joined Virginia and Maryland in that most evaluations indicated that its WIP Design Process was good on paper (i.e., in CAST) but not in practice. New York's WIP Design was uniformly described as NOT a Paper Process. Evaluations for Pennsylvania and West Virginia were split between Paper Process and NOT Paper Process.

Rather than break all of the variables down by state numerically, it is easier to maintain confidentiality if we analyze patterns of variation among states for the variables described in aggregate above. Keep in mind that these are blocks of text describing the state's WIP design process. Although the majority of these blocks come from a respondent who is from the given state, some will also come from independent or federal respondents or respondents from other states (see Section E.1 Block Identifiers). First, Figure 34 shows a heatmap of the average number of mentions per state for each of the major factors influencing WIP design reported in aggregate in Figure 29. "Building on Previous WIPs" and "Past Experience" remain the most mentioned factors for all states except Pennsylvania, though "Start with Existing" was mentioned several times for PA. With the exception of Pennsylvania, these responses reinforce our expectation that there is path dependence in the WIP Design process and so we anticipate problems such as autocorrelation and heteroskedasticity in any statistical analyses of WIP content.

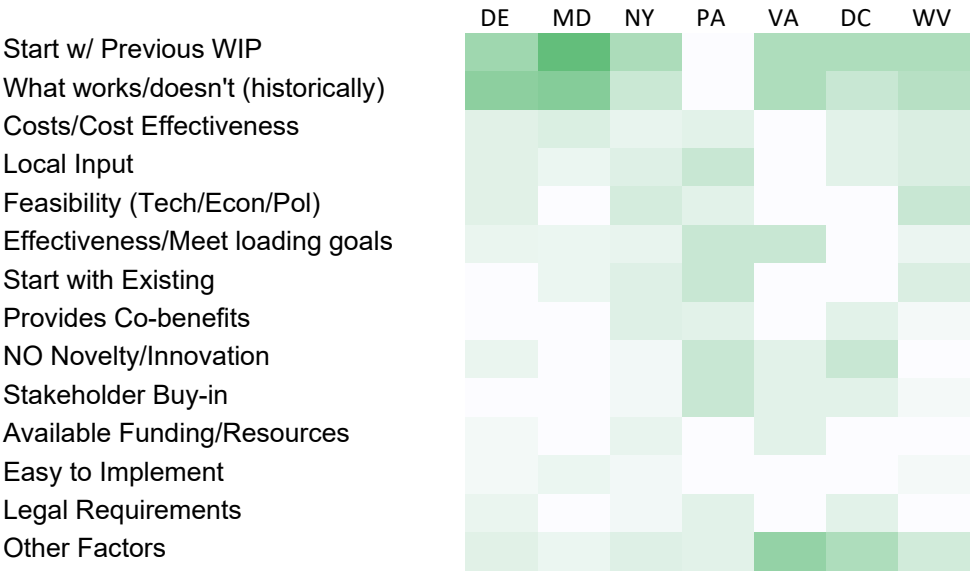


Figure 34: Heatmap of Average Mentions per State for each Major Factor Influencing WIP Design. Variables that indicate learning from past experience is the most important factor influencing WIP design remain dominant for all states except Pennsylvania.

"Feasibility of BMPs Generally" also received 6 mentions in blocks describing the WIP Design process for West Virginia. Almost all of the other factors were mentioned fewer than 5 times for each state. Because of the similarities between some categories (e.g., "Cost Effectiveness" and "Effectiveness", "Feasibility", and "Available Funding", etc.) and the small number of blocks/state, we do not think this information is definitive on its own and point the reader to Sections 6.5 BMP Selection Aggregated and 6.6 BMP Selection by Type of BMP. Surveys to collect data on each of these factors specifically might also be useful.

Patterns of responses by state for hierarchy and engagement are fairly consistent. but there are some areas that do not match up (see Figure 35). Maryland, Virginia, and Washington, DC were described as Top Down more often than Bottom Up and mentions of Any Engagement with stakeholders were also relatively low for these states. Performative Engagement, which involves consulting stakeholders without incorporating their feedback was mentioned most often for Maryland and Virginia, which is consistent with a Top Down approach to WIP Design. Note, however, that while all respondents were asked if the process was bottom up or top down, descriptions of engagement were not requested specifically and so reflect the perceptions of respondents who made the connection to engagement themselves. New York, Pennsylvania, and West Virginia had a larger proportion of mentions describing their WIP Design Process as "Bottom Up", which corresponds with higher levels of "Any Engagement" and "Reflexive Engagement", which occurs when stakeholder feedback is incorporated into decision making. Delaware was split between "Bottom Up" and "Both Bottom Up and Top Down", but it did have a high proportion of mentions of "Any Engagement".

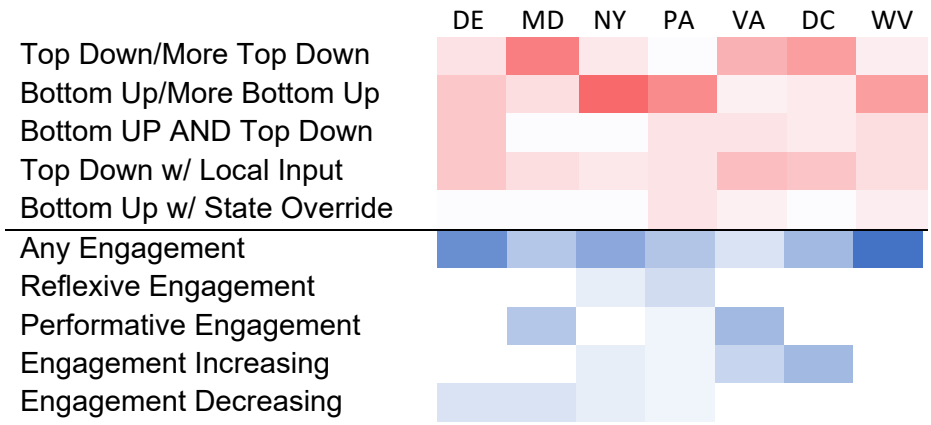


Figure 35: Heatmap of Average Mentions per State for Decision Hierarchy and Engagement Variables. Respondents described Maryland, DC, and Virginia WIP Design as more Top Down with less Reflexive engagement, while New York, Pennsylvania, and West Virginia processes were described as more Bottom Up with higher levels of engagement, including reflexive engagement in New York and PA.

There are some inconsistencies here as well, however. For instance, engagement was described as increasing in a larger proportion of blocks for Virginia and DC but decreasing in a small proportion of blocks for Delaware and Maryland. Responses were split between increasing and decreasing for New York and Pennsylvania. Some of these divergences may result from the use of "Engagement Increasing" as a diplomatic way to say that Engagement is relatively low vs. using it with is actual meaning. There is also some disagreement over whether or not engagement is reflexive or performative in Pennsylvania. All in all, these inconsistencies occur in only a few blocks of text, but they still signal different perceptions of the process among responses for each state.

The heatmap for the proportion of stakeholder groups with influence over the WIP design process is shown in Figure 36. Washington, DC stands out because the Federal Government is the only group mentioned with influence. This is not surprising, given that so much of the land—and therefore the nutrient reduction—in DC is controlled by federal agencies. In other words, federal agencies are major stakeholders in DC's WIP Design Process. Farmers were mentioned as an influential in all other jurisdictions, which is also commensurate with the fact that DC is an entirely urban area while other states all have some level of agriculture. Given that the number of blocks is quite low in this category, we do no read much into the rest of the state-level break-down. Looking more closely at "Federal Government", it is interesting that at least one block of text described the Federal (mainly EPA/CBP outside of DC) role for all states except Maryland. West Virginia and Delaware had more blocks that were critical of the Federal government's role, but again, due to low sample size/state these observations should not be considered generalizable to the entire state.

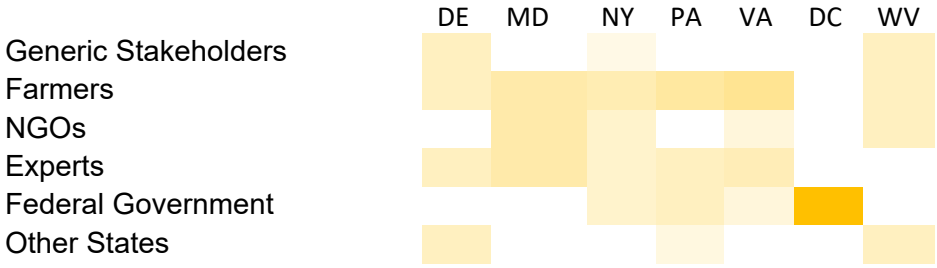


Figure 36: Heatmap of Average number of Mentions by State of Stakeholder Groups with Influence over the WIP Design Process. Federal government is most important in Washington, DC because a much larger proportion of their land area is owned or managed by federal agencies.

For use of CAST in the WIP Design process, the aggregate findings are largely reinforced, though there is some variation among states. CAST was designated as Central to the WIP Design process for a majority of blocks describing all states except Maryland and Virginia (see Figure 37). No blocks contained statements about the importance of CAST in the Maryland WIP Design process, though a few mentioned that it was better than previous models while others described difficulties using it. This finding may be a result of the high concentration of respondents from Maryland who identified their general role as Scientist rather than Regulator or Tech/Modeler. Virginia's blocks were split evenly between "CAST is Central" and "CAST is Useful, NOT Central". New York and West Virginia also had a few blocks in the latter category. All blocks for Delaware, Maryland, and West Virginia indicated that stakeholders did not use CAST themselves during the WIP design process, while blocks for other jurisdictions were split between the two variables in this category. Again, the number of blocks is low, though, so additional data collection on when and where stakeholders use CAST would be needed to make any definitive statements.

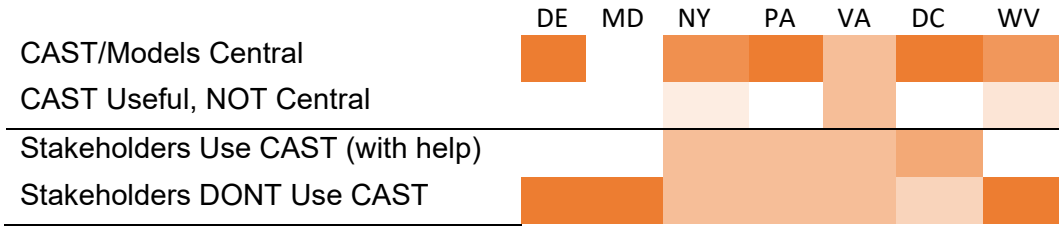


Figure 37: Heatmap of the Average Mentions by State for Use of Models in the Design Process. CAST was described as important or useful in the WIP Design process for all states except Maryland, though this may be a result of sampling bias. Stakeholders were more likely to use CAST (with state-level assistance) in New York, Pennsylvania, Virginia, and Washington, DC.

Looking at statements about whether or not WIP design is a Paper Process or some variation on that phrasing, a majority of blocks indicated that Delaware, Maryland, and Virginia's WIPs were good on paper but not in practice (see Figure 38). All blocks of text describing New York's WIP Design process indicated that it was NOT just a paper process, but that it was designed to be implemented effectively. Blocks describing Pennsylvania and West Virginia's WIP Design processes were split between these two descriptions. Qualifiers on statements that WIP Design was a paper process were generally split between "this makes the process less legitimate" and "BUT it's the best we can do", except for Virginia, where legitimacy of the process was questioned, but limitations were not mentioned.

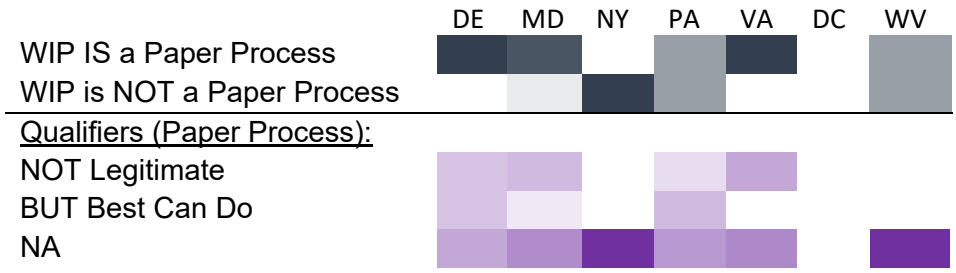


Figure 38: Heatmap of Average Mentions per State for "Paper Process" Variables. A majority of statements described WIP design as a paper process for Delaware, Maryland, and Virginia. All statements in this category for New York described their WIP Design as NOT a Paper Process. Responses were split for Pennsylvania and West Virginia. Respondents who indicated that WIP Design was a Paper Process indicated that this reduced the legitimacy of watershed governance for Delaware, Maryland, Pennsylvania, and Virginia, with at least some responses indicating that Paper Process is the Best We Can Do for Delaware, Maryland and Pennsylvania.

This last set of variables is particularly interesting in light of the description of state behavior in the TMDL/Allocation section of the survey instrument. It suggests that leadership by Maryland and Virginia at the CBP-level may not be matched by leadership in WIP Design or implementation. Furthermore, leadership on the TMDL may be more politically feasible for these states if they are not paying the political costs of fully implementing their WIPs. That said, it is possible that some respondents over-represented the lack of realism in a state's WIP Design process because of ideological leanings or inter-state animosities. In light of this possibility, it is important to note that, while about 42% of these critical blocks came from respondents in the Multi/Independent or Federal categories, the remainder were made by people from a given state about their own state's process. Indeed, state-level respondents were not likely to be critical of other state's processes and in some cases were more charitable toward other states. Furthermore, while 47% of blocks in the "Paper Process" category came from respondents in the General Role: Communication category, 25% came from respondents in the Regulator Role and 18% came from the Tech/Modeler Role. Sampling may have skewed results for Virginia, which has a higher proportion of respondents in the Communication role than other states (57% vs. 0-28%).

Methods Note: In most heatmaps we report number of mentions but when breaking down data by state-level jurisdictions we report average mentions/state in order to protect respondent confidentiality by obscuring the number of respondents who made specific statements about a given state. Relying on averages can skew the perceived importance of a given result, however, so we note that there were at least 7 mentions for each state whose WIP Design was primarily described as a Paper Process above. Results are also relatively robust for evaluations of Top Down/Bottom Up variables except for Delaware but are not robust for most of the engagement variables (except Any Engagement). Furthermore, although respondents in the Multi/Independent category were more likely to describe WIP Design as a Paper Process across all states, with the exception of New York, respondents from other jurisdictions

also frequently evaluated WIP Design as a Paper Process. Throughout, respondents were most likely to discuss their own state's WIP Design Process. Few blocks recorded a respondent from one state criticizing the WIP Design Process of another state (see Section E.1.3 Jurisdiction Codes). Nevertheless, the sample size per state is small, so we would recommend additional data collection to validate these results.

6.3 Sub-state Allocation

Allocation by sector was the most common approach across all states, which generally reflects differences in regulation outside of the CBP (e.g., the MS4 Permitting Process). Most responses indicated that states also allocated by local level jurisdictions, but some indicated that they preferred not to do so. Responses were not consistent across states, so additional data collection is needed. Pertinent local-level jurisdictions vary substantially by state, so this needs to be taken into account in the social science model. Political aspects of allocation, particularly local backlash, were also mentioned, though the number of blocks is too small to analyze these results in greater depth.

After asking respondents to describe the WIP Design process in general, we then asked them to explain how their state distributed loading goals to sub-state-level entities. Since there are fewer blocks and fewer variables to summarize, we combine aggregate and state-specific observations in this section.

When asked specifically to describe how their state allocated loading goals among sub-state actors during the WIP Design Process, respondents identified 4 main categories of sub-state actors. As shown in Figure 39, allocation by sector was the most common response in this category. This makes sense because stormwater and wastewater are regulated under separate rules from agriculture or forestry sectors. Allocation by basin or sub-watershed was mentioned in 6 blocks but was less common than other responses. Allocation at the local level, whether to counties (m = 14) or other entities (m = 17) such as municipalities, soil and water districts, or, in the case of Washington, DC, ownership by government agencies. Interestingly, 9 blocks of text indicated that the state did not undertake local level allocation at all. These answers were not consistent within states—that is, all states with blocks saying that they did not allocate at the local level ("NOT by Local") also had blocks indicating "County" or "Other Local" level allocations. It is likely that this is a result of differences in the sectoral expertise of respondents. For instance, MS4 permits are allocated at the local level across all states, but some states allocate agricultural loading goals to the local level while others do not.

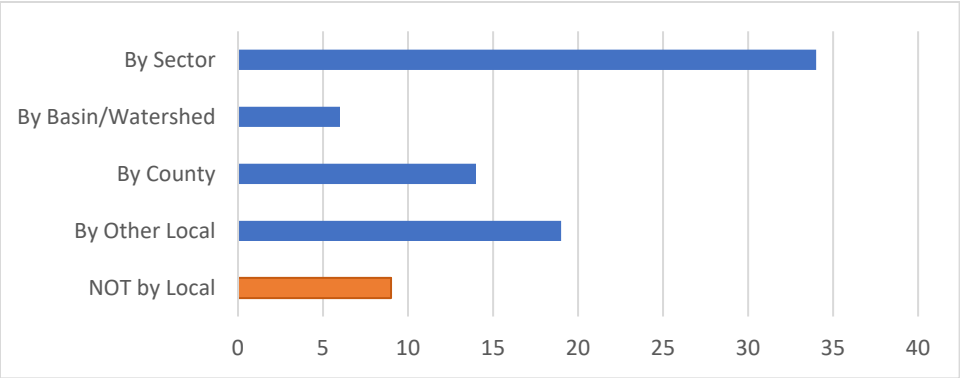


Figure 39: Sub-State Actors Considered in Allocation Process (# Mentions). Allocation by sector was most prevalent across all jurisdictions, reflecting the MS4 Permitting process among other regulatory requirements. Other responses were not consistent across state-level jurisdictions, so additional data collection is required.

There are fewer blocks of text describing the hierarchy of decision making for the sub-state allocation process than for the WIP Design process as a whole, and no statements regarding the level of stakeholder engagement. This is likely because the respondents already provided information on these topics on WIP Design broadly and were not asked specifically about them in the question on sub-state allocation. As can be seen by comparing Figure 30 and Figure 40, responses regarding hierarchy in the sub-state allocation are more evenly distributed than those for WIP Design as a whole. One interesting finding here is that blocks describing the process as both bottom-up and top-down distinguished between sectors; wastewater/stormwater were described as "More Top Down" while agriculture was described as "More Bottom Up". This is another reflection of the regulatory mechanisms for waste and stormwater compared to voluntary measures for agriculture (see Section 6.4 Sectoral Aspects of WIP Design/Allocation).

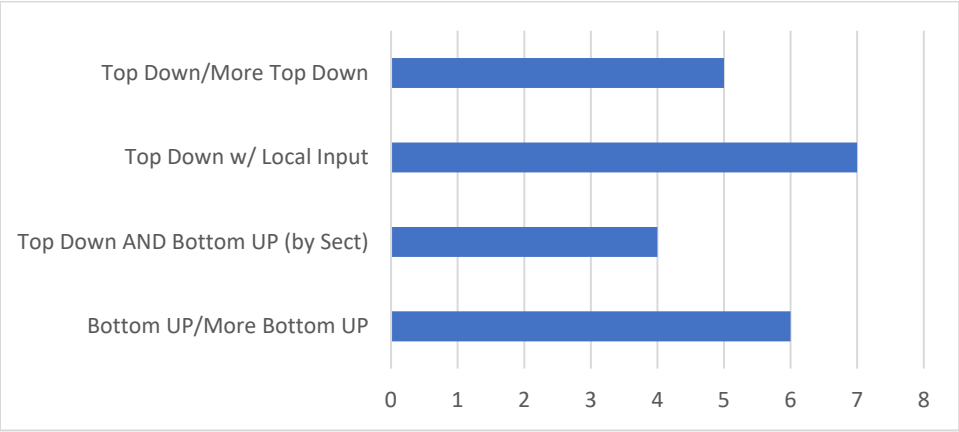


Figure 40: Hierarchy of Decision-Hierarchy Responses for Sub-State Allocation Process. Number of observations is small, but responses are generally split between categories.

Hierarchy responses differed by state, though we take these results as indications for future work with the low number of blocks per state. Blocks describing the sub-state allocation process for Maryland, Virginia, and Washington, DC were all either in the Top Down/More Top Down or Top Down w/ Local Input category. Other states were split among multiple categories, though Delaware, New York, and Pennsylvania were the only states with blocks of text describing their sub-state allocation process as Bottom Up/More Bottom Up. Responses are consistent across states, insofar as no state was coded with blocks of text from both the Top Down/More Top Down and Bottom Up/More Bottom Up categories.

Most of the other variables in this category were only observed in 2-3 blocks of text and much less information was provided than for WIP Design as a whole. Stakeholder influence was also not addressed by respondents who answered this question, though there was some mention of federal pressure to allocate loading goals down to the local level for states that were not meeting their loading goals. A few mentions were made of using CAST to allocate sub-state loadings and there were a few mentions of the difficulties of using CAST for this purpose. Most interesting, 11 blocks of text addressed political issues with sub-state allocation, all of which bore some relation to conflicts over allocation or concerns about the equity of allocation at this level. Concerns were expressed based on sectoral and geographic boundaries. There was also some mention of tradeoffs between the equity and cost-effectiveness of load distribution.

6.4 Sectoral Aspects of WIP Design/Allocation

Respondents focused mainly on Wastewater, Stormwater, and Agriculture when discussing WIP Design generally, followed by Point and Non-Point sources. Forest, Atmosphere, and Urban were mentioned much less often in the context of WIP Design/Allocation. Agriculture was generally described as the sector that had the greatest capacity for cost-effective load reduction, but also where load reduction was more difficult because BMP adoption is voluntary. A number of statements indicated that farmers or decision makers had a strong preference to avoid mandatory regulation, and that this might spur higher levels of voluntary implementation. The stormwater sector was overwhelmingly viewed as too difficult or expensive to include as a major source of load reduction in WIPs, though about a third of statements indicated that it was still important to include stormwater-related BMPs in WIPs because reductions in the sector were needed to meet loading goals. Wastewater was viewed as a success early phases of the TMDL, when large reductions in load were attained through investment in major wastewater treatment plants but most statements recognized that there was limited capacity or high marginal costs to further reduce loading in the sector. There was also a perception among some respondents that load reduction was easer or more effective in highly regulated/urban sectors because BMP implementation was mandatory rather than voluntary. This section highlights the importance of accounting for temporal dynamics in the social science model, especially since there appears to be path dependence in the capacity of load reduction by sector. In other words, we cannot expect that the distribution of load reductions in future WIPs will simply be a trend based on load reductions in past WIPs.

Although we did not ask about sectors specifically, respondents volunteered information about the roles of different sectors in the WIP Design/Allocation process. Figure 41 shows how many blocks of text contained information about WIP Design/Allocation by Sector. One of the first things to note is that respondents differed in how they categorized sources, with some breaking them down by sector, but others using broader groups like Point vs. Non-Point. Among the traditional sectors, there were many more comments about wastewater, stormwater, and agriculture than forests or atmospheric deposition. Some blocks also talked about WIP Design for Urban sectors, which generally encompassed wastewater, stormwater, and urban forestry or similar initiatives in urban areas specifically. There were no statements about the rural counterpart outside of agriculture, even though respondents did describe other types of rural issues when providing more detail on the selection of BMPs (e.g., upgrading septic systems). As shown in Section 6.6 BMP Selection by Type of BMP, the forest sector—especially forest buffers and urban forestry—were discussed much more when respondents described criteria for including specific BMPs in WIPs than when they discussed sectors generally. Atmospheric deposition was not mentioned often, as it is not a component of WIP Design, but it did come up in a few blocks of text in this section.

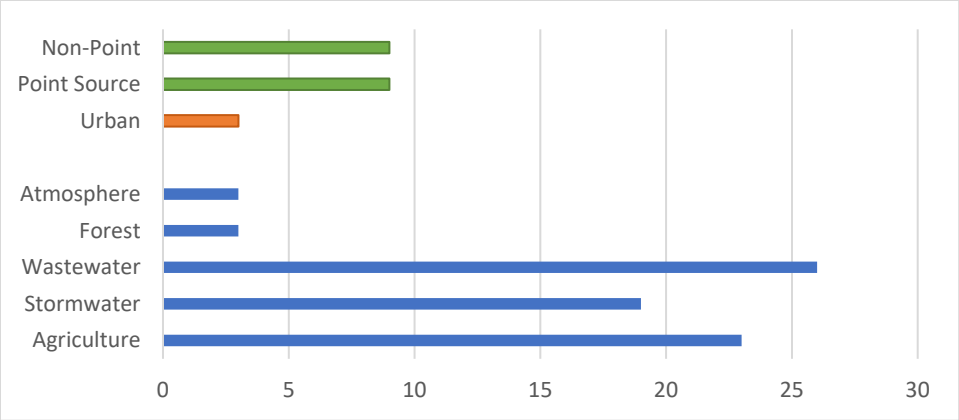


Figure 41: Comments about WIP Design/Allocation by Sector (#Mentions). Wastewater, Stormwater, and Agriculture received the highest level of attention, followed by Point and Non-Point sources. Forest, Atmosphere, and Urban were mentioned much less often in the context of WIP Design/Allocation.

State-level breakdowns of these variables suggest that they may reflect institutions rather than the importance of the sector in the state (see Figure 42). For instance, looking at only the top three sectors, blocks describing the WIP Design/Allocation process in Pennsylvania are weighted toward their largest sector (Agriculture), but Agriculture was not mentioned in blocks for either New York or West Virginia, even though both states' loads also come mainly from this sector. In contrast, Agriculture is even more predominant in blocks of text describing the WIP Design/Allocation process for Delaware, even though the state's load is more balanced among sectors. Maryland, Virginia, and Washington, DC are a better match between sectoral loading and mention of sectors, with DC's urban geography reflected in the lack of any statements about Agriculture. Interestingly, a number of blocks of text described sectoral components of WIP design for all state-level jurisdictions. This set of blocks leaned more toward Wastewater and Stormwater than Agriculture.

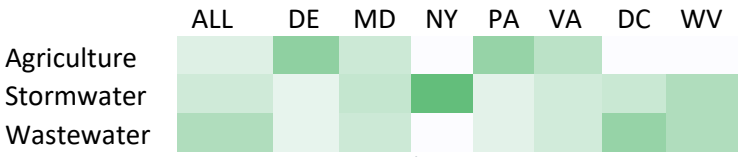


Figure 42: Heatmap of Average Mentions/State for Top 3 Sectors. There is no discernable relationship between how often a sector was mentioned as part of a state's WIP Design/Allocation process and how much that sector contributes to the state's load.

6.4.1 Agricultural Sector in WIP Design/Allocation

Breaking the data down by sector, annotations for Agriculture fell into the largest number of categories. However, only three of these detailed categories had more than 5 mentions, so we grouped them further into the four sets shown in Figure 43. The largest number of statements in this category stated that loading was more difficult to reduce for agriculture than for wastewater or stormwater. Reasons for this difficulty included: 1) the voluntary nature of agricultural programs (as opposed to mandatory programs in other sectors), 2) prevailing norms against conversion of agricultural land to urban development (which should reduce nutrient loading because new development is strictly regulated for stormwater), 3) lower visibility of agricultural runoff to the public, and 4) high variation within the agricultural sector (e.g., farm size, products grown/raised, technologies used, and social and financial capital availability, etc.). Under-crediting of agriculture BMPs in CAST was also mentioned as a factor

that made reducing agricultural loading more difficult—or at least made it more difficult to reduce the state's official loads and therefore to meet their loading goals.

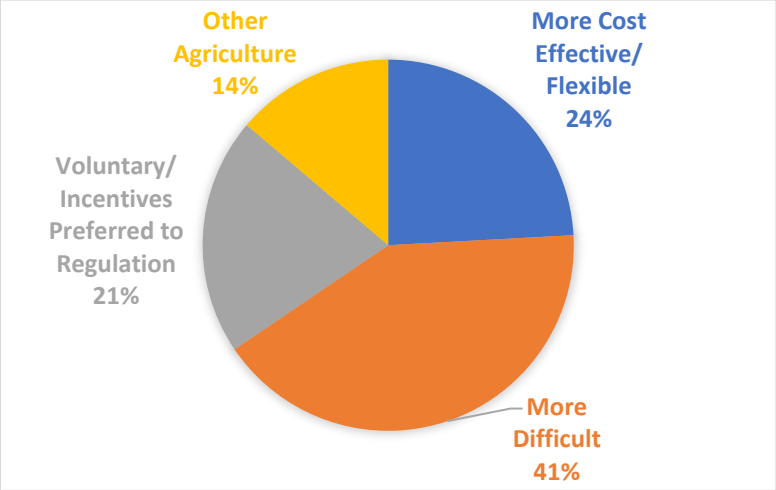


Figure 43: Statements about Agriculture as a Sector in the WIP Design/Allocation Process (% Mentions). The largest group of statements about the agricultural sector in WIP Design/Allocation indicated that it is More Difficult to reduce loading in the sector, though about a quarter of statements indicated that reductions in agricultural loads were more cost effective and/or farmers were provided with greater flexibility to reduce their loads. Another large group of statements indicated that there is a strong preference for maintaining voluntary measures supplemented by incentives as opposed to mandatory regulations for agriculture.

In some blocks, these More Difficult annotations were used to explain why states focused more on load reductions in other sectors in their WIPs. In contrast, about 24% of statements about Agriculture indicated that more of the state's load should be placed on agriculture because load reduction was More Cost Effective/Flexible in the sector. These responses tended to co-occur with statements about limited capacity and strong regulatory requirements for stormwater and wastewater sectors (see Section 6.4.2 Stormwater and Wastewater Sectors in WIP Design/Allocation).

In 6 blocks of text, respondents indicated that farmers had a strong preference for voluntary over mandatory approaches to load reduction in the sector, and that this could be used to motivate them to accept more voluntary practices now in order to avoid the threat of mandatory practices of load reduction goals were not met. Interestingly, these blocks were found in descriptions of the WIP Design/Allocation process for 4 states, including both Bay Coastal states and two Non-Bay States. These statements were not used to support any contentions about the allocation of loading by sector, but this political factor was considered important in the selection of BMPs for Agriculture. The last group of annotations for this sector was largely comprised of statements about the proportion of agriculture in the state's nutrient load and remaining capacity for loading reduction in the sector. For the most part, these blocks supported increasing the load on agriculture, as it was generally depicted as having high remaining capacity for load reduction compared to other sectors.

6.4.2 Stormwater and Wastewater Sectors in WIP Design/Allocation

Stormwater and wastewater were both discussed primarily as urban problems, although some rural concerns about sewage were mentioned in response to the question about BMPs selection (see Section 6.5 BMP Selection Aggregated). Both sectors are regulated directly under the Clean Water Act as well as through the CBP TMDL. Stormwater is regulated through the MS4 permitting system, while wastewater is regulated through the EPA's Effluent Guidelines. There were fewer categories of statements for

Stormwater than for Agriculture , and all 6 could be distilled into two major groups. As shown in Figure 44, the majority of these statements indicated that stormwater was too expensive or difficult and so was not emphasized in the state's WIP. Within this group, most statements referred specifically stormwater retrofits, or upgrading existing development to reduce stormwater run-off, as opposed to ensuring high standards for new developments, which are already regulated. A few blocks in this category also mentioned that the capacity for load reduction was low in the stormwater sector and so the state's WIP focused on other sectors where capacity was higher.

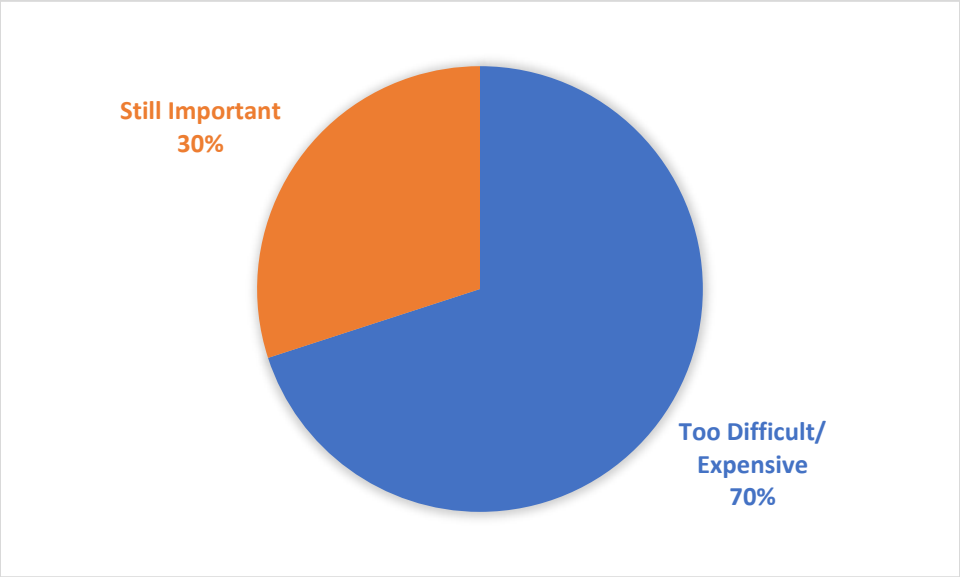


Figure 44: Statements about Stormwater as a Sector in the WIP Design/Allocation Process (% Mentions). The stormwater sector was overwhelmingly viewed as too difficult or expensive to include as a major source of load reduction in WIPs, though about a third of statements indicated that it was still important to include stormwater-related BMPs in WIPs because reductions in the sector were needed to meet loading goals.

About one third of the blocks discussing stormwater in the context of WIP Design/Allocation argued that it was still important and should be emphasized more in WIPs. Co-benefits, including improvements in human health and reduction of visible pollution in waterways were mentioned most often as rationale to continue to improve on stormwater in spite of the difficulties described above. A few blocks also mentioned that reducing loadings in other sectors would not be enough to meet the goals for 2025 and so stormwater remained important even if it was a relatively small proportion of the state's load.

Although there was some variation, statements about wastewater in the WIP Design process all pointed in a similar direction, though they were made differently. Again, 6 specific categories were aggregated into two broad groups for this report (see Figure 45). Just over half of these blocks indicated that reducing nutrient loads through improvements in wastewater treatment was easier than reducing loadings in other sectors, at least in Phases I and II of the WIPs. Most, however, agree with the other group of statements which indicated that there is little capacity to reduce loads further via the wastewater sector. This is due to both diminishing returns to treatment plants and reduced capacity for improvement as many of the most important plants have already been upgraded in the region. Both types of statements focused almost exclusively on large-scale wastewater treatment plants in urban areas, though at least one agricultural wastewater treatment plant was mentioned.

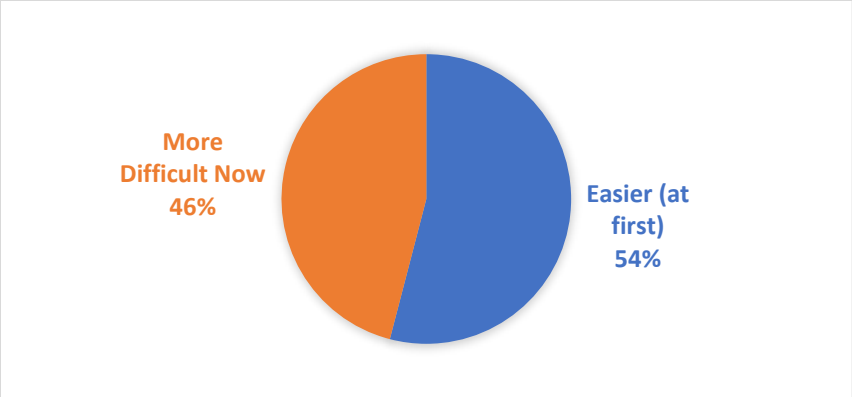


Figure 45: Statements about Wastewater as a Sector in the WIP Design/Allocation Process. Wastewater was viewed as a success in early phases, when large reductions in load were attained through investment in major wastewater treatment plants but most statements recognized that there was limited capacity or high marginal costs to further reduce loading in the sector.

6.4.3 Other Sectors in WIP Design/Allocation

Most other sectors were not described as extensively in relation to the WIP Design/Allocation process. Atmospheric deposition was mentioned briefly in a few blocks as an area where nutrient reduction had occurred in the early phases of the TMDL but for which there was little remaining capacity. Interestingly, these reductions were a side effect of tighter air pollution controls under the Clean Air Act, rather than the CBP TMDL. Descriptions of the "urban" sector generally conformed to the above analysis, suggesting that urban was easier than agriculture, likely because it involves mandatory stormwater and wastewater permitting systems. Similarly, point source pollution—largely wastewater—was viewed as easier because it was regulated whereas non-point source—mainly agriculture—was considered more difficult because BMP adoption was voluntary with incentives rather than mandatory. This supports the idea that there is a difference in terminology used by respondents but that this does not affect their overall evaluation of the role of different sectors in the WIP Design/Allocation process.

6.5 BMP Selection Aggregated

When asked to describe how BMPs are selected for WIPs, responses are similar to factors described in Section 6.1 WIP Design Aggregated, but variables associated with learning from the past drop out and are replaced by descriptions of BMPs as Popular with stakeholders or Preferred by decision makers. In fact, Popular/Preferred is the most common reason to include BMPs in WIPs (or exclude if NOT Popular/Preferred), with regulators, farmers, and the general public mentioned most often as groups whose BMP preferences mattered most. This may again relate to the fact that agricultural BMPs are voluntary, as are a number of other BMPs like private green infrastructure. Effectiveness in reducing nutrient loading, Feasible/Cheap/Easy to Implement, and Co-benefits/Co-costs were also mentioned frequently. Co-benefits were mentioned more often by respondents describing BMP selection for New York and West Virginia, and Feasibility was a more common rationale for BMP selection in West Virginia compared to all other states. Funding and CAST-related variables come up as important factors more when respondents describe BMP selection specifically than they did in general descriptions of the WIP Design Process.

After describing the WIP Design/Allocation processes generally, we asked respondents to specifically explain why different BMPs might be selected for their state's WIP. This question produced a wealth of information about many different types of BMPs. Here we review the various factors that respondents

identified as generally important when selecting BMPs for the WIPs. For instance, if a respondent said that cover crops were likely to be included in the WIP because they were popular with farmers, that would show up here as a mention of Popular/Preferred = Reason to Include. Note that the actual decision process involves weighing multiple factors against each other, so for any given BMP factors that provide Reasons to Include may be balanced against factors that provide Reasons to Exclude. We will examine the relative influence of factors in Section 6.6 BMP Selection by Type of BMP. For now, we simply review the importance of each factor in itself through the total number of mentions.

The factors affecting BMP selection are shown in Figure 46. Popularity with a particular stakeholder group or simply being preferred by regulators was the most common reason for including a BMP in a WIP. Exclusion of BMPs that were not Popular/Preferred was also a common response, but not more so than other reasons to omit BMPs. Effectiveness in reducing loading was the second most common reason to include a BMP in a WIP, and lack of Effectiveness was a noted reason to exclude BMPs. Respondents varied in whether Effectiveness referred to load reductions "on the ground" or "in CAST". This will be discussed in greater detail below. Some blocks specified Cost Effectiveness or the "biggest bang for the buck" as a reason to include a BMP in a WIP, with about as many saying that lack of Cost Effectiveness was a reason to exclude. Similarly, Feasibility/(Low) Costs/Ease of Implementation was described as an important factor whether positive (reason to include) or negative (reason to exclude). There is some overlap between these last three factors, as Cost Effectiveness encompasses both Effectiveness and how cheap or easy the BMP is to implement. The other major factor was Co-benefits (reason to include) or Co-costs (reason to exclude) associated with the BMP. Co-benefits were described in Section 4.3 Co-benefits. **Co-costs** are negative unintended effects of BMPs on some stakeholder group. This will be covered more in the next section.

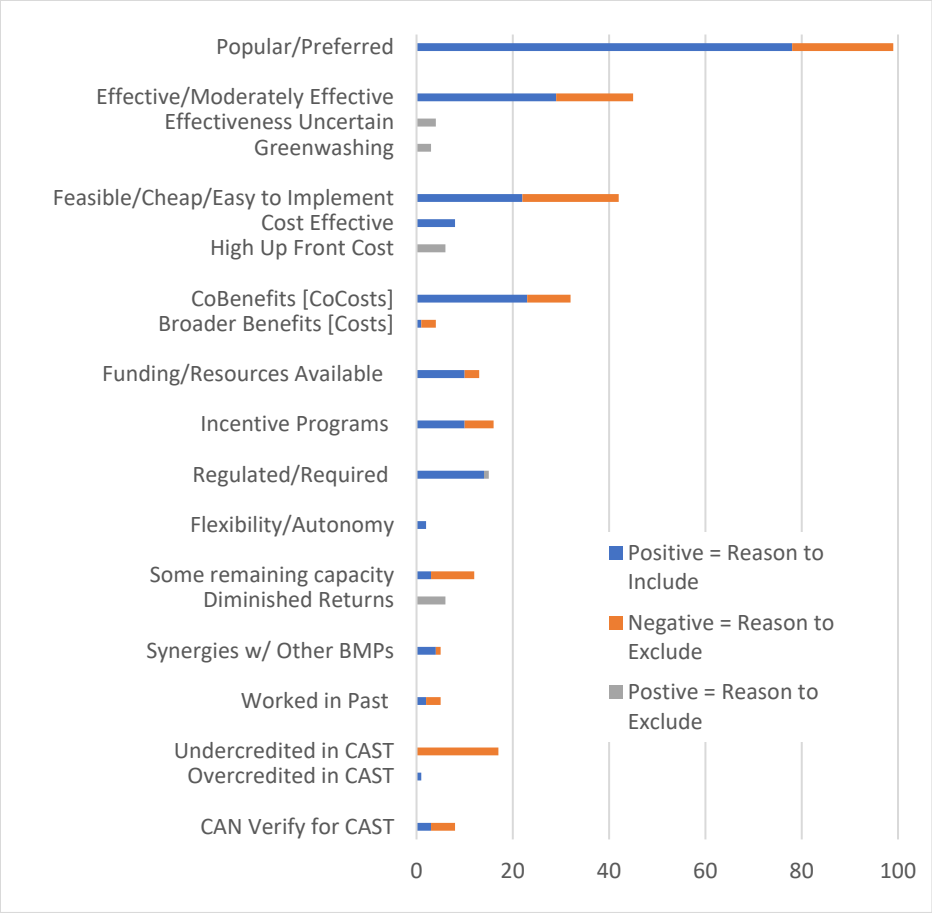


Figure 46: Factors considered when selecting BMPs (# Mentions). Popular/Preferred was by far the most common reason for selecting BMPs, Effectiveness in reducing nutrient loading, Feasible/Cheap/Easy to Implement, and Co-benefits/Co-costs were also mentioned frequently. Most of these results are similar to factors described in Section 6.1 WIP Design Aggregated but learning from the past drops out and is replaced by Popular/Preferred. Funding and CAST-related variables also come up as important factors more when respondents describe BMP selection specifically.

Other factors were mentioned much less often and are generally more specific versions of the major factors described above. The availability of funding or other resources for implementation, incentive programs (especially for voluntary BMPs), flexibility and autonomy for stakeholders are all factors that make a BMP more (or less) feasible and synergies with other BMPs is similar to the idea of co-benefits. There is an interesting variation in the perception of voluntary vs. regulatory mechanisms. As noted above, stormwater and wastewater BMPs are usually required for regulation. This was typically stated as a reason to include related BMPs in the WIP without any statements about effectiveness or preferences. For the agricultural sector, however, there was a clear preference for voluntary programs with incentives, and in one block antipathy toward mandatory requirements for BMPs. Discussions of diminished returns and remaining capacity (e.g., land available, areas where new treatment plants could be added) also paralleled the sectoral discussion, though there were concerns about limits on some agricultural and multi-sector BMPs as well (see next section). It is interesting that whether or not a BMP was observed to "work in the past" was mentioned in so few blocks, given how central past experience was in the general descriptions of the WIP Design/Allocation Process. It may be that respondents conflate Popular/Preferred with What Works/Doesn't.

The last set of factors in Figure 46 deals with accounting for BMPs in CAST. Similar to Effectiveness, the ability to Verify BMPs so that they could be credited against the state's loading goals was mentioned as a reason to include, while difficulties with Verification were described as reasons to exclude a BMP from the state's WIP.

The only large difference among states as they described factors affecting selection of specific BMPs is greater emphasis on "Popular/Preferred" as a reason for selection for Delaware, Maryland, New York, and Virginia, which mention this factor in 2-3 times as many blocks as Pennsylvania, Virginia, or Washington, DC. In addition, co-benefits were mentioned as a reason to include BMPs in the state's WIP much more often for New York (m = 12) and West Virginia (m = 12). Co-benefits were associated with BMPs in less than 5 blocks for all other jurisdictions. The only other large difference is that blocks of text describing the West Virginia design process emphasize feasibility and ease of implementation in both the positive and negative incarnations much more frequently than for other states (2-6 times more mentions). Otherwise, we received fewer responses in this category generally for Maryland and Washington, DC. The former may be another vestige of the sampling bias described in our methods section. The latter is likely because urban Washington, DC has many fewer BMPs to choose from than states with agricultural sectors.

Breaking down the Popular/Preferred responses further, Figure 47 shows which stakeholder groups were mentioned in conjunction with these code groups. Being "preferred" by regulators received the largest number of mentions, though there were very few BMPs that were omitted because they were not preferred by regulators. Popularity among environmentalists and related NGOs was mentioned for a few BMPs, and both the positive and negative incarnations were mentioned for Urban/Suburban populations. Overall, however, popularity with farmers and the general public were mentioned much more often than with other stakeholder groups, as both as a reason to include if a BMP was popular with these groups and a reason to exclude if it was not popular with these groups. Results for farmers were more evenly split between the positive and negative incarnations of this variable, and their importance generally reflects the voluntary nature of BMP implementation for the Agricultural sector. That is, decision makers realize that they cannot force farmers to implement BMPs, they can only provide incentives, which makes them more likely to take farmer preferences into account. That said, the high number of blocks for farmers may also be skewed by the fact that so many BMPs are designed for the agricultural sector and therefore they received more attention from respondents.

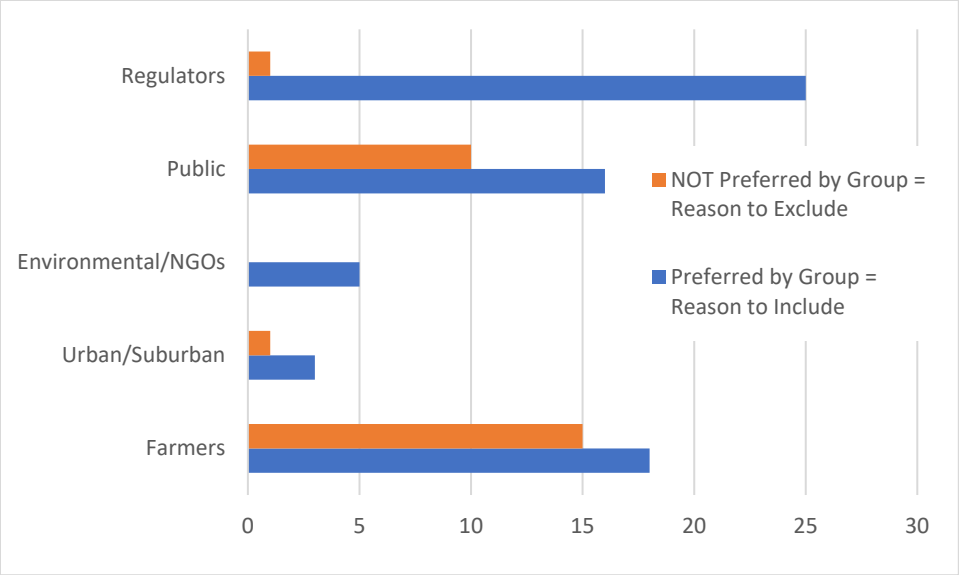


Figure 47: Popular/Preferred Categories by Stakeholder Group

6.6 BMP Selection by Type of BMP

Wastewater has the fewest options and was mentioned less than most other sectors, even though it was mentioned frequently in Section 6.4 Sectoral Aspects of WIP Design/Allocation. MS4 Permits were mentioned frequently for stormwater but other types of BMPs in the category received equal attention. Cover Crops, AG Waste/CAFO, and Nutrient Management were mentioned much more often than other agricultural BMPs. Buffers were mentioned most often in the multi-sector BMP category.

Within the WIP Design/Allocation section, respondents described factors affecting the selection of 22 main types of BMPs. Of these, 2 applied to the wastewater sector, 6 applied to the stormwater sector, 8 to the agriculture Sector, and 7 could apply across multiple sectors. In this section, we first review the types of BMPs by number of mentions and then go into more detail on each type of BMP by sector.

Total number of mentions per type of BMP are shown in Figure 48. Wastewater has the fewest BMP options, which apply to either the design of treatment plants (m = 9) or the design of septic and sewer systems (m =8). Since this sector is already heavily regulated, choices are often limited. Stormwater is also regulated at the national level, but there are a number of options for this sector, including public and private green infrastructure, stormwater management, and retrofits or removing impervious surface. Technically, the MS4 permitting system is not a BMP. It establishes limits on the amount and quality of stormwater discharge for cities, towns, and other municipalities. However, respondents frequently talked about MS4 permits in conjunction with WIP Design because they set a lower limit on the amount of nutrient reduction that can be accounted for in the stormwater sector. Therefore, we discuss them here even though they do not exactly fit. Most stormwater BMPs were mentioned in about the same number of blocks, so it does not appear that some receive more attention than others. A few other stormwater-related BMPs were also mentioned, including manufactured treatment devices (MTDs) and water diversion.

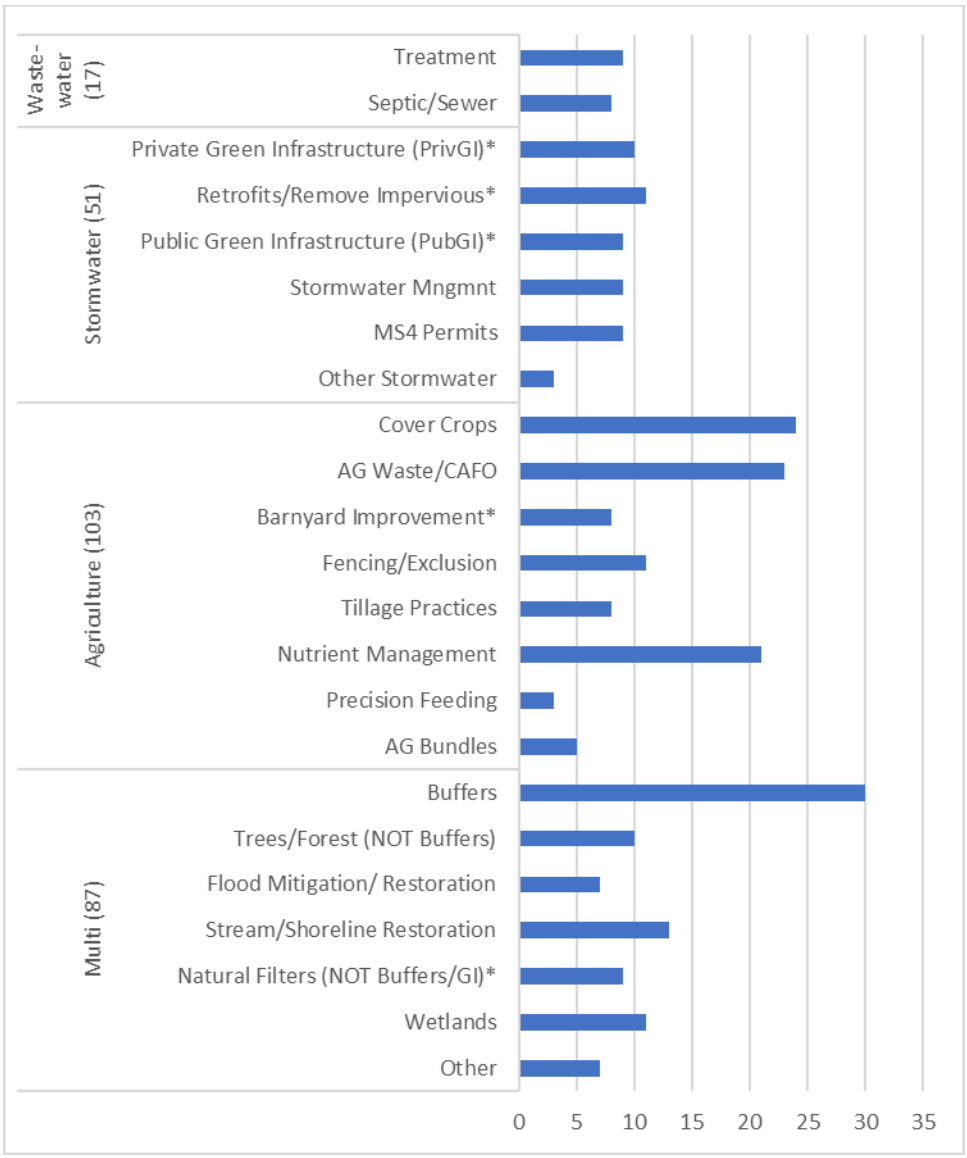


*Figure 48: Types of BMPs Described (# Mentions; * Indicates heterogeneous grouping). Wastewater has the fewest options and was mentioned less than most other sectors, even though it was mentioned frequently in Section 6.4 Sectoral Aspects of WIP Design/Allocation. MS4 Permits were mentioned frequently for stormwater but other types of BMPs in the category received equal attention. Cover Crops, AG Waste/CAFO, and Nutrient Management were mentioned much more often than other agricultural BMPs. Buffers were mentioned most often in the multi-sector BMP category.*

Some BMPs definitely receive more attention than others in the agricultural sector. Cover crops, agricultural waste and CAFO-related BMPs, and nutrient management plans were mentioned in more blocks than almost any other set of BMPs. Buffers (forest and grass) were mentioned in more blocks than these three categories, but while most buffers are used in the agricultural context, some are not, which places this type of BMP in the multiple sectors category. Other agricultural BMPs received less attention, including barnyard improvement, fencing/exclusion of livestock from streams and stream banks, tillage practices designed to minimize runoff and precision feeding to reduce animal waste. Interestingly, bundles of agricultural BMPs were mentioned in a few blocks. This is the idea that considering BMPs in combination could improve both nutrient reduction and farmer profits over considering them separately.

Other BMPs that could be used in multiple sectors were mentioned less often that buffers but about the same as wastewater or stormwater BMPs. This included forestry related practices other than buffers, such as forest conservation and urban tree canopy, flood mitigation and restoration efforts, stream and shoreline restoration, preservation or renewal of wetlands, and use of natural filters other than buffers or green infrastructure. With fewer than 5 mentions, street sweeping, structural/permanent facilities, and cultural/non-technical changes were aggregated into the Multi/Other category. The rest of this section looks at how each of these types of BMPs was described in the context of WIP Design/Allocation.

6.6.1 Wastewater BMPs in WIP Design/Allocation

Only two types of BMPs were identified for wastewater: Treatment Plants and Septic/Sewer Improvements. Descriptions of the former paralleled discussions of Wastewater as a sector in Section 6.4 Sectoral Aspects of WIP Design/Allocation, particularly indicating that there is limited capacity or increasing marginal costs to reduce loading further via sewage treatment. Comments about Septic/Sewer Improvements were more like discussions of the Stormwater sector, in that these BMPs were viewed as difficult and expensive although some blocks of text mentioned co-benefits like improved human health.

As noted above, descriptions of wastewater BMPs fall into two broad categories: Sewage Treatment and Septic/Sewer Improvements. It is difficult to compare these types of BMPs directly because there are few blocks of text describing them and, of those, there is little overlap in the variables coded. It is interesting that no blocks of text describe either type of BMP as “popular/preferred” or “NOT popular/preferred”.

Codes for Sewage Treatment mirror findings for the wastewater sector as a whole. Treatment is described as an easy way to achieve large reductions in nutrient loading in 5 blocks of text, though a few blocks mention that high up-front costs and the regressive nature of sewage taxes could be considered drawbacks. Even so, 5 blocks of text also mention diminishing returns/increasing marginal costs to improved sewage treatment and another 5 mention that capacity for load reduction using sewage treatment is now very low. In other words, it would be expensive to improve loading reductions in the high-efficiency plants (diminishing returns/increasing marginal costs), but also most of the sewage system that can be upgraded to high-efficiency plants already have been (low capacity). Only one block of text mentions that there may be additional capacity for further reduction in loads via improvements in wastewater treatment technologies.

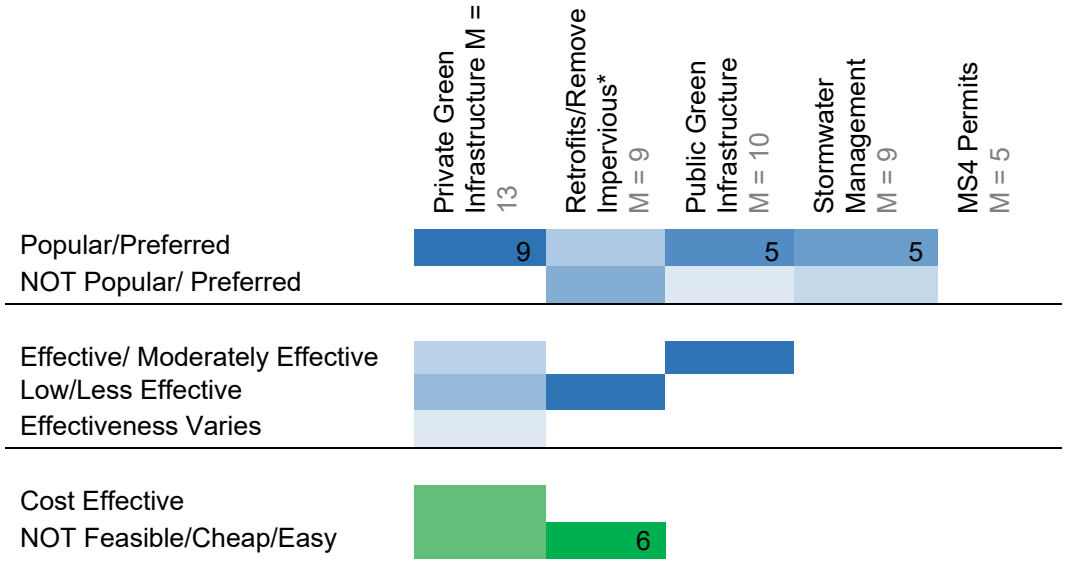
In contrast, comments about septic and sewer improvements diverge from the general statements on wastewater and are more similar to concerns about stormwater retrofits. Although these improvements are mentioned as being moderately effective for load reduction in one block of text, the majority (m = 9) describe them as expensive, with high up-front costs and potential to be regressive if costs are spread out evenly across the population. Another block mentions that sewer and septic improvements are also under credited in CAST. All in all, these BMPs appear less likely to be included in WIPs, although one block of text did mention their co-benefits in terms of human health and cleaner drinking water. There may be a divergence in views here because treatment tends to be an urban option, while septic retrofits and sewer improvements are more rural or suburban options. Large cities have both the incentives and the resources to invest the large-scale, highly effective treatment plants and to pay the professional workforce required to maintain both treatment plants and sewage systems. Rural and poorer suburban areas lack said resources and are also not able to take advantages of the economies of scale available to cities.

Breaking down these BMPs by state cuts the data to fine for either confidentiality or reliability, but it is interesting to note that wastewater treatment was not mentioned at all for WIP Design in Delaware or Pennsylvania, and septic/sewer improvements were not mentioned for New York, Virginia, or Washington, DC. Both types of BMPs had more blocks of text for West Virginia than any other state. Septic/sewer improvement was also mentioned more often for Delaware. It will be interesting to compare these outcomes to WIPs from these states to see if they are discussed more extensively there as well.

6.6.2 Stormwater BMPs in WIP Design/Allocation

Private Green Infrastructure like rain gardens and pollinator plots were described as popular/preferred more often than other BMPs but responses varied on the effectiveness and costs/benefits of these practices. Public Green Infrastructure and Stormwater Management were also described as Popular/Preferred more often than not. Retrofits/Remove Impervious was mixed on Popular/Preferred but responses clearly indicated that this type of BMP was expensive. MS4s were not mentioned often and, when they were it was mainly to described them as a mandatory requirement.

Responses describing the role of stormwater BMPs in WIP Design/Allocation are more varied. As shown in Figure 49, Private Green Infrastructure was most often perceived as popular/preferred among the different types of stormwater BMPs. This includes things like rain gardens, green roofs, pollinator plots, and other changes to reduce nutrient runoff from private property. While these BMPs can also be used on public property, to be included in this group blocks specifically described them as practices undertaken by homeowners or businesses. Evaluations of the effectiveness of these BMPs was divided, however, and there were also contrasting comments on the costs of private green infrastructure. Some of this may be based on variations among the BMPs in this category. The number of blocks for other factors were low for this category, but there is some indication that private green infrastructure might be less likely to be included in WIPs because it is not easy to get funding/resources for implementation and these BMPs are perceived to be under-credited in CAST. The fact that these BMPs are voluntary with incentives was also mentioned more as an aspect of WIP design than a factor affecting the selection of this type of BMP per se.



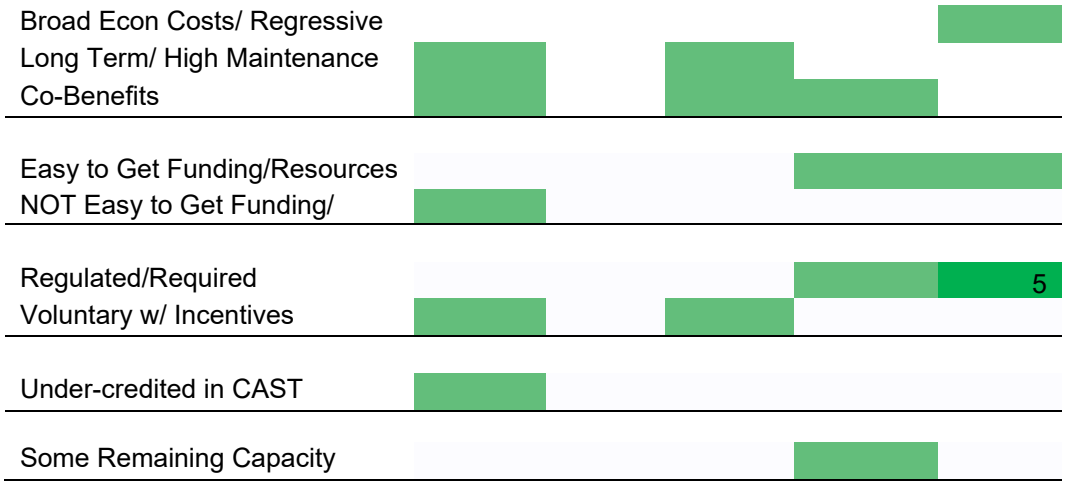


Figure 49: Heatmap of Average Mentions (blue) or Any Mentions (green) per BMP for Stormwater (# Mentions provided if > 5). Private Green Infrastructure like rain gardens and pollinator plots were described as popular/preferred more often than other BMPs but responses varied on the effectiveness and costs/benefits of these practices. Public Green Infrastructure and Stormwater Management were also described as Popular/Preferred more often than not. Retrofits/Remove Impervious was mixed on Popular/Preferred but responses clearly indicated that this type of BMP was expensive. MS4s were not mentioned often and, when they were it was mainly to described them as a mandatory requirement.

Responses were divided on whether retrofitting/removing impervious surfaces is a popular or preferred set of practices. It was uniformly viewed as both ineffective (m < 5) and expensive (m = 6), however, suggesting that these BMPs are will not be a major component of most WIPs. In contrast, public green infrastructure like bioretention swales and permeable pavements were also split on the Popular/Preferred axis but this was more heavily weighted to Popular/Preferred. Public GI was also uniformly considered to be Effective/Moderately Effective. No blocks of text described the costs these BMPs specifically, but a few mentioned that they were good long-term options if maintained and that they could provide co-benefits. Like private green infrastructure, it was also mentioned that some public green infrastructure projects were voluntary at the municipal level and that incentives for implementation would need to be provided in the WIP.

Stormwater management is an umbrella term that could include the use of other types of BMPs, including green infrastructure. It is a broad approach to reducing stormwater runoff that employs multiple technologies strategically across a given space (e.g., town, city, region). We code it separately here when the term is used specifically by respondents. Like public green infrastructure, stormwater management is split along the Popular/Preferred axis but more blocks of text fall on the positive side than the negative. Little else was said about this type of BMP, including no statements about effectiveness. There were a few mentions of positives like co-benefits and the availability of resources, as well as some remaining capacity for load reduction using this approach. Regulatory requirements for stormwater management were also mentioned, indicating that this set of BMPs might be included in WIPs because they are required in other regulatory contexts (i.e., flood prevention, public health).

Less information was provided on MS4 permits, but as mentioned above, these are not technically BMPs though they are certainly considered in the WIP design process. The most frequently mentioned aspect of MS4s was the fact that they are required by regulation (m = 8), which was generally seen as a positive in terms of both design and implementation. Some concern was expressed about the potentially regressive or growth-inhibiting effects of MS4s, but it was also noted that funding is relatively easy to

find. Of course, with so few blocks, it is possible that these statements reflect geographic variation or other differences, so, again, more research would be needed to substantiate these findings.

Examining variation by state for stormwater BMPs reinforces our concerns about lack of information for Maryland, where only MS4 permits were mentioned in the section on BMP selection for WIPs. Otherwise, West Virginia and Washington, DC had the highest number of blocks addressing stormwater generally (m = 7 for both, 2-3x greater than for other states). Most blocks for West Virginia focused on private green infrastructure or retrofitting/removing impervious surfaces, while blocks for Washington, DC were more evenly spread among the different types of BMPs, though public green infrastructure received marginally more attention. Here again, DC's focus on stormwater broadly can be explained by their urban geography. Delaware was mainly focused on retrofits/removal, with some mentions of the other types of stormwater BMPs. Responses from Virginia were more heavily weighted toward both public and private green infrastructure, with some attention to stormwater management. Pennsylvania had the second fewest blocks in this category after Maryland, with statements spread fairly evenly across all types of BMPs except for the green infrastructure BMPs, which were not discussed hardly at all.

Methods Note: Blocks of text describing whether a set of practices is popular/preferred or evaluating effectiveness are more numerous than blocks for other variables, so we present them as averages in the heatmap but provide only information on presence/absence of blocks with for other descriptors. We omit information on "Other Stormwater BMPs" entirely because the number of blocks for this category is so small and the variety of BMPs covered it is so large. Except in a few cases where mentions are larger than 5, findings on these factors should be interpreted as potentially important but more research would be needed to determine if the perception is widely held. Additional data collection is also indicated for types of BMPs with contradictory patterns (i.e., some blocks indicate that they are popular/preferred others indicate that they are NOT popular/preferred). These differences may be due to a number of factors, including the state whose WIP Design process is being described, heterogeneity of the practices within the set, and the stakeholder group considered (e.g., a practice may be popular with some stakeholders but not others).

6.6.3 Agricultural BMPs

Among Agricultural BMPs, Cover Crops were most often described as Popular/Preferred. Assessment of their Effectiveness in reducing loading was mixed, but there was agreement on other positive attributes such as Feasibility, Co-benefits, Funding Availability, and Worked in Past. Most other Agricultural BMPs mentioned were also described much more often as Popular/Preferred than NOT Popular/Preferred. Precision Feeding was the one exception, but there was little other information provided on this BMP. Waste storage was only described as Effective/Moderately Effective and Barnyard Improvements and Removal/Disposal were only described as Low/Less Effective. Effectiveness was mixed for remaining BMPs. Credits in CAST and/or ability to Verify for CAST were mentioned as important for most of these types of BMPs.

As noted above, agricultural BMPs were discussed more and in greater detail than most other BMPs. Figure 50 reviews the number of mentions for agricultural BMPs from Figure 48 with more detailed information on the AG Waste/CAFO category in particular. Once we differentiate between these different types of waste management practices, mentions of these categories are more commensurate with other types of BMPs, though the broad category of Agricultural Waste Management still gets a

fairly high number of mentions. As noted above, “Cover Crops” was the most frequently mentioned type of agricultural BMP but respondents provided little information about different types of cover crops, planting methods, etc. “Nutrient Management” also received a lot of attention, but different approaches were not described.

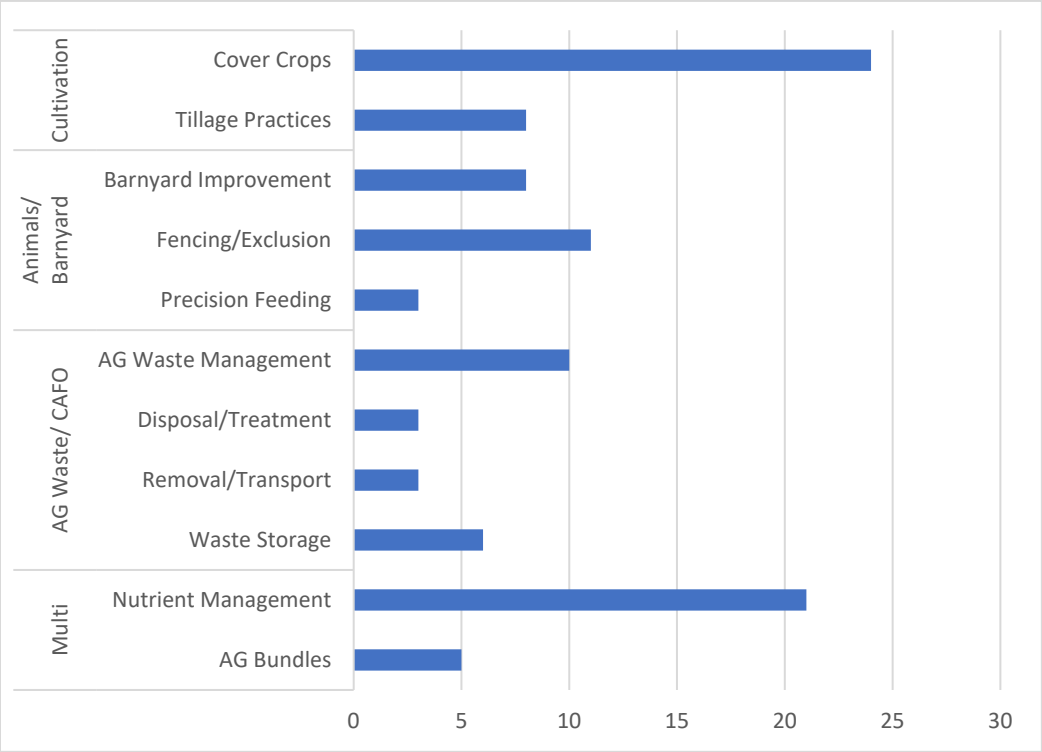


Figure 50: Further Breakdown of Agricultural BMPs (# Mention). More detailed breakdown of Agricultural BMPs shows four main types of AG Waste/CAFO BMPs, leaving Cover Crops and Nutrient Management as the most frequently mentioned Agricultural BMPs. Although there are multiple types of Cover Crop BMPs and several approaches to Nutrient Management respondents did not differentiate among them, so we cannot disaggregate these variables any further.

Like our analyses for stormwater and wastewater BMPs, breaking down factors influencing BMP selection by sub-category cuts the data too fine for definitive results for most sub-categories of BMPs. However, as shown in Figure 51, there are some clear if unsurprising results for Cover Crops. It is plainly considered to be the most popular of all of the BMPs mentioned, whether agricultural or not, with 36 blocks of text indicating that this approach is popular/preferred. Reasons for this popularity include perceived co-benefits (m = 7), whether in terms of local water quality or increased profits to farmers, ease of access to funding (m = 5), and the voluntary nature of the practice (m = 8). Responses were mixed in their evaluation of the effectiveness of this BMP, however, with 9 blocks indicating that it is Effective/Moderately Effective and 7 indicating that it is Low/Less Effective. Although the details of implementation matter for effectiveness, these differences are not attributed to implementation per se and seem to reflect assessments of cover crops generally. Feasibility, past success, and the ability to verify the BMP (and therefor count its effects on the state’s loading goals in CAST) are other positive aspects of cover crops that were mentioned. A few blocks mentioned drawbacks, such as cover crops being under-credited in CAST and limited room for expanding capacity for this BMP. Again, the number of blocks of text for these responses is quite low, indicating the need for future research to verify the findings.


Cultivation
Animals/Barnyard
AG Waste/CAFO
Multi
Cover Crops
Tillage Practices
Barnyard Improvement
Fencing/ Exclusion
Precision Feeding
AG Waste Management
Disposal/ Treatment
Removal/ Transport
Waste Storage
Nutrient Management
AG Bundles
Popular/ Preferred
NOT Popular/ Preferred
36
5
7
Effective/Moderately Effective
Mixed/Variable Effectiveness
Low/Less Effective
9
8
7
Feasible/Cheap/Easy to Implement
NOT Feasible/Cheap/Easy
Cost Effective
High Up-Front Costs
Flexibility/ Autonomy
Structural/ Long Term/ High Maintenance
Diminished Returns
Co-benefits
Risks
7
5
Funding Available
5
Voluntary w/ Incentives = Good
Regulated/No Incentives = Low uptake
8
5
Remaining Capacity HIGH
SOME Remaining Capacity
Remaining Capacity LOW

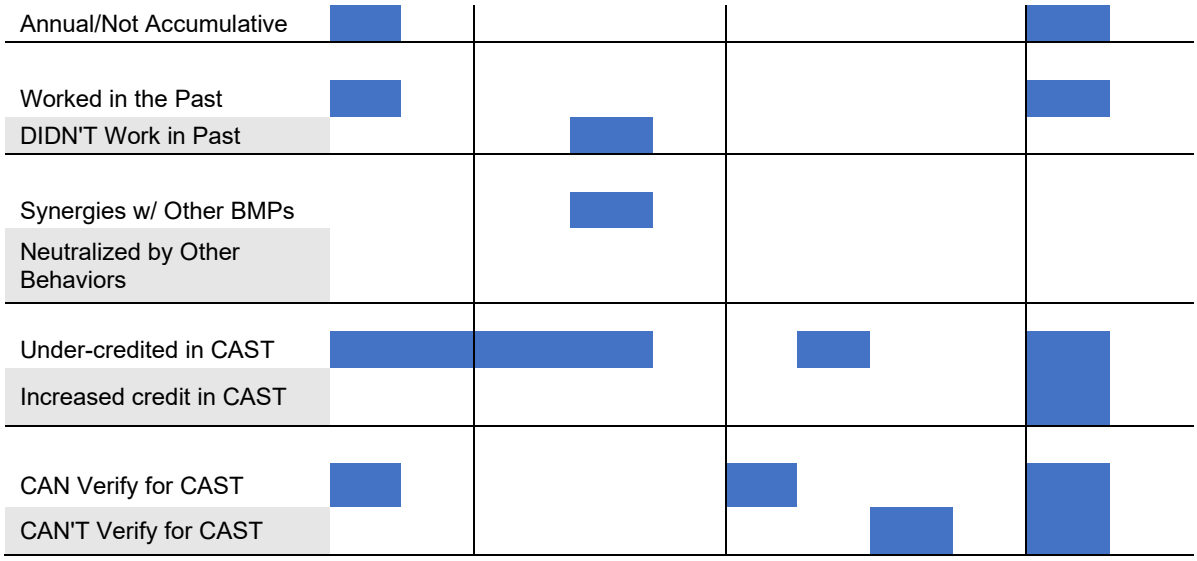


Figure 51: Heatmap of Average Mentions (Green) and Any Mentions (Blue) per Category for Agricultural BMPs (# Mentions provided if > 5). Cover crops were most often described as Popular/Preferred. Assessment of their Effectiveness in reducing loading was mixed but there was agreement on other positive attributes such as Feasibility, Co-benefits, Funding Availability, and Worked in Past. Most other Agricultural BMPs mentioned were also described much more often as Popular/Preferred than NOT Popular/Preferred. Precision Feeding was the one exception, but there was little other information provided on this BMP. Waste storage was only described as Effective/Moderately Effective and Barnyard Improvements and Removal/Disposal were only described as Low/Less Effective. Effectiveness was mixed for remaining BMPs. Credits in CAST and/or ability to Verify for CAST were mentioned as important for most of these types of BMPs.

The only other agricultural BMP with enough blocks of text to evaluate confidently is Nutrient Management. Seven blocks described the practice as popular/preferred, with no blocks describing it as NOT popular/preferred. Respondents also evaluated the practice as Effective/Moderately Effective in 8 blocks of text with only on Low/Less Effective evaluation. Reasons for this popularity are similar to those for cover crops (Co-benefits, m = 5; Voluntary w/ Incentives, m = 5) but the number of blocks is much lower and there are some areas of contention, so additional research is indicated.

Among the remaining BMPs, stream-bank fencing/excluding livestock from streams was also relatively popular (m = 5) though seen to have mixed effectiveness. Available funding appears to be the main factor in its favor, with some drawbacks mentioned. Tillage practices (no till, conservation tillage), barnyard improvement, and AG Bundles scored moderately high on popular/preferred though assessments of effectiveness were either mixed or non-existent. Barnyard improvement was the only type of BMP seen to provide long-term benefits, which was a positive in terms of perceived effect but negative due to high upfront costs, high maintenance costs, and the difficulty of verification in CAST. Other agricultural BMPs were not described enough to make even preliminary determinations but anecdotally, further research into the storage, treatment, and transport of poultry litter and other animal waste may be useful.

If we look at patterns across different states, there is little that can be said about separate factors, but we can at least describe the level of attention paid to each type of BMP for each state (see Figure 52). As explained above, Washington, DC is an urban area, so there were no mentions of agricultural BMPs for this jurisdiction's WIP Design/Allocation process. Cover crops was the only type of BMP mentioned for all other states. Tillage practices, nutrient management, and AG Waste/CAFO were mentioned 5 of the 7 jurisdictions, plus a few blocks describing WIP design across all states. Nutrient management came up

more often for Delaware (m = 5) and New York (m = 6) and AG Waste/CAFO was mentioned more frequently for Pennsylvania (m = 5) and West Virginia (m = 10) . In West Virginia, AG Waste/CAFO was specifically tied to concerns about poultry litter in some blocks of text. Other agricultural BMPs were only mentioned for 2-4 of the 7 jurisdictions. Livestock-related BMPs other than CAFO in particular were only mentioned in states with commercial-scale poultry, cattle, or dairy production.

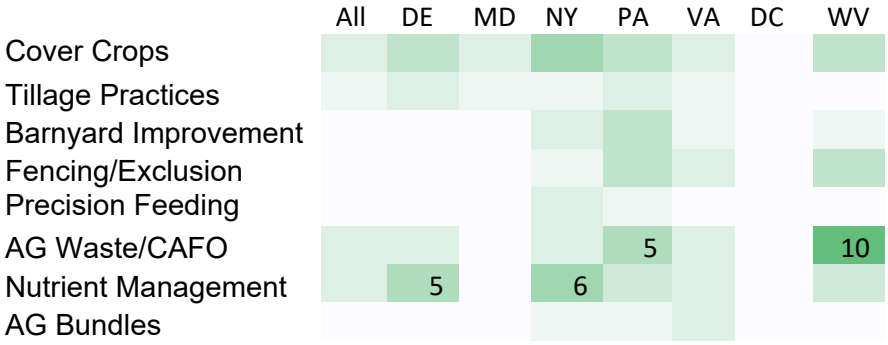


Figure 52: Heatmap of Mentions of Type of BMP by State WIP Design (# Mentions provided if > 5). Cover crops was the only type of BMP mentioned for all states other than Washington, DC. Tillage Practices were also mentioned for most states but not often. Mentions of AG Waste/CAFO and Nutrient Management were concentrated in states with commercial-scale poultry, cattle or dairy production.

6.6.4 Multi-Sector BMPs

Most types of Multi-Sector BMPs were described as Popular/Preferred more than NOT popular Preferred. Forest Buffers were exceptional in that most statements described them as NOT Popular Preferred in spite of their perceived Effectiveness and Co-benefits. Co-costs like loss of agricultural land, Long Term/High Maintenance costs and past failures may explain why this type of BMP is less prevalent in WIPs. Forest Buffers were also one of the few types of BMPs deemed OVER-credited in CAST. Grass Buffers were described consistently as Low/Less Effective but also Feasible/Cheap/Easy, which may explain why they were uniformly considered Popular/Preferred. Assessments of Popular/Preferred were mixed for Wetlands, Riparian Buffers, and Trees/Forest (NOT Buffers) BMPs, again in spite of high perceived Effectiveness and Co-Benefits. Reasons for NOT Popular/Preferred assessments are similar to those for Forest Buffers. Other Popular/Preferred BMPs like Stream/Shoreline Restoration also were associated with Long term/High Maintenance and Co-Costs, however, so more research is needed to better understand why BMPs are selected.

Several types of BMPs are used in more than one sector. Buffers can be used stormwater management and to reduce nutrient runoff from yards or farms. That said, the majority of the comments about buffers referred to their use in agriculture. As shown in Figure 53, respondents discussed three “types” of buffers. Some distinguished between forest buffers (m = 18) and grass buffers (m = 8), while others used the more generic term riparian buffer (m = 13), which probably refers to forest buffers, but we do not have sufficient information to be sure in any given block of text. Forestry projects other than buffers are varied, including conserving existing forests, creating urban tree canopies, or transforming agricultural land to forests. Depending on their location, the first two may help to reduce urban stormwater or reduce runoff/increase nutrient uptake from agriculture. Unfortunately, we had to aggregate these disparate forestry projects due to low numbers of blocks per practice, which makes interpreting these data more difficult. Natural filters other than buffers or green infrastructure is also a mixed back of BMPs, including dry ponds, biochar, oyster seeding, drainage ditches, and conservation

landscaping. Most of these were mentioned only a few times and so were aggregated into a single category.

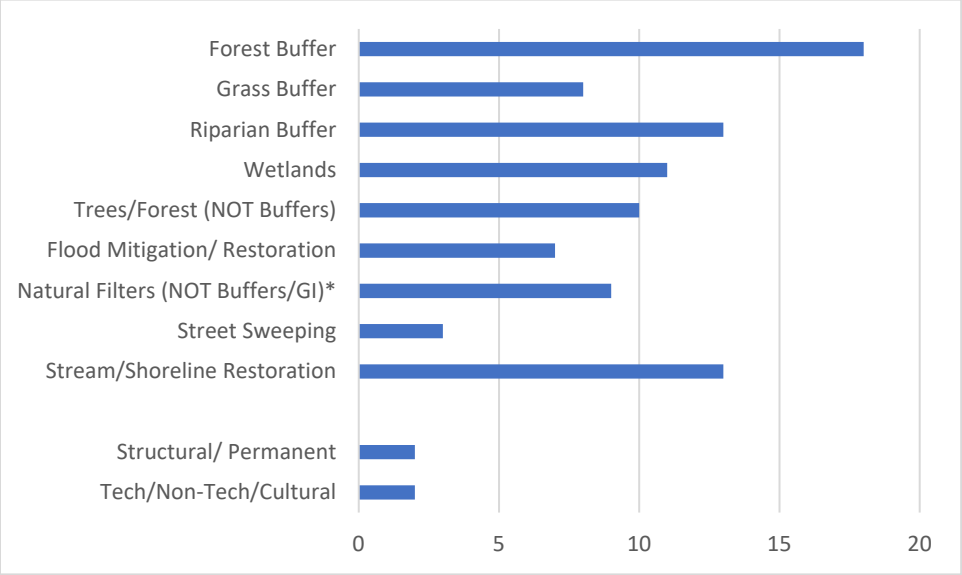


Figure 53: Further Breakdown of Multi-Sector BMPs (# Mentions). When broken down into different sub-types, Forest Buffers are still mentioned more than any other Multi-Sector BMP and generic Riparian Buffers are also mentioned quite often. Grass Buffers have limited geographic application. They were mentioned less often than Wetlands, Trees/Forests (NOT Buffers), Natural Filters, and Stream/Shoreline Restoration. Street sweeping is also a primarily urban BMP that was only mentioned a few times. General references to Structural/Permanent and Tech/Non-Tech/Cultural BMPs occurred a few times as well.

On the other hand, practices such as stream or shoreline restoration, flood mitigation and floodplain restoration, and wetlands restoration were generally discussed as unitary categories of BMPs, so interpretation of related variables should be easier. Street sweeping was also mentioned only a few times. It is reported here, but we will not perform more detailed analysis because the number of blocks of text are two small. Similarly, a few blocks of text discussed categories of BMPs by their characteristics, including broadly describing WIP Design factors for BMPs that were non-technical or cultural and BMPs that were structural or permanent. With so few mentions, we only point out that at least some respondents characterized BMPs in this way and believed that these characteristics were important for WIP Design.

Most of the factors described as affecting the selection of multi-sector BMPs are the same as those mentioned in previous sections. The tendency toward positive bias is also visible, although not for forest buffers or wetlands, both of which have more blocks in the “NOT popular/preferred” category (see Figure 54). Evaluations of effectiveness also tend toward the positive, including for forest buffers and wetlands, but not for grass buffers, which were uniformly recognized as being less effective than forest buffers or for natural filters. Again, the latter category is heterogeneous, so some of the BMPs in this category may be considered more effective than others. Grass buffers were mentioned as being easier than forest buffers and both flood mitigation/floodplain restoration and natural filters were uniformly described as feasible/cheap/easy while trees/forest (NOT buffers) were uniformly described as difficult or expensive. Other multi-sector BMPs were split on this variable, which may reflect differences across states or sectors.


Forest Buffer
Grass Buffer
Riparian Buffer
Wetlands
Trees/ Forest (NOT Buffers)
Flood Mitigation/ Restoration
Natural Filters
Stream/ Shoreline Restoration
Popular/Preferred
NOT Popular/Preferred
5
6
10
6
6
5
7
Effective/Moderately Effective
Mixed/Variable Effectiveness
Low/Less Effective
8
5
5
Feasible/Cheap/Easy
NOT Feasible/Cheap/Easy
Cost Effective
Long Term/High Maintenance
Co-benefits
Co-Costs
7
6
Easy to Get Funding/Resources
NOT Easy to Get Funding/Resources
Voluntary w/o Incentives = low uptake
Remaining Capacity Low
Didn't Work = Do Less
Synergies w/ Other BMPs
UNDER-credited in CAST

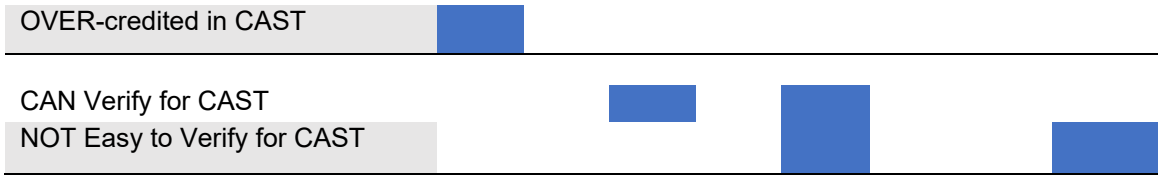


Figure 54: Heatmap of Average Mentions (Green) and Any Mentions (Blue) of Factors affecting the Selection of Multi-Sector BMPs (# Mentions provided if > 5). Most types of Multi-Sector BMPs were described as Popular/Preferred more than NOT Popular/Preferred. Forest Buffers were exceptional in that most statements described them as NOT Popular/Preferred in spite of their perceived Effectiveness and Co-benefits. Co-costs like loss of agricultural land, Long Term/High Maintenance costs and past failures may explain why this type of BMP is less prevalent in WIPs. Forest Buffers were also one of the few types of BMPs deemed OVER-credited in CAST. Assessments of Popular/Preferred were mixed for Wetlands and Trees/Forest (NOT Buffers) BMPs, again in spite of high perceived Effectiveness and Co-Benefits. Reasons for NOT Popular/Preferred assessments are similar to those for Forest Buffers. Other Popular/Preferred BMPs like Stream/Shoreline Restoration also were associated with Long term/High Maintenance and Co-Costs, however, so more research is needed to better understand why BMPs are selected.

Blocks of text were again too few to rely on results for other factors affecting selection of these BMPs in WIP Design. One interesting point, however, is that this is the only group for which "co-costs" were mentioned along with co-benefits. In the agriculture BMPs, there was a mention of potential increased risk with cover crops, but other negative unintended consequences for BMPs were not described for sector-specific BMPs. Here, negative effects were mentioned for forest and riparian buffers (e.g., shade on crops, increased habitat for pests), wetlands (e.g., bugs, other pests, smells), trees/forests (NOT buffers; e.g., increased crime from urban tree canopy, loss of agricultural heritage from conversion of farmland), and stream/shoreline restoration (e.g., loss of access to water sources, lower esthetic values). Nevertheless, co-benefits were also emphasized for some of these BMPs, including riparian buffers (m = 7; e.g., improved water quality and recreation) and flood mitigation/floodplain restoration (m = 6; e.g., reduced damage from flooding).

Unlike barnyard improvement, the long-term nature of many of these types of BMPs was also viewed as a negative insofar as there was an expectation that they would not be maintained and so resources would be wasted. In a few blocks of text, examples of failed buffers or stream restoration efforts were mentioned as rationale to minimize reliance on these BMPs in a state's WIP. The question of whether or not farmers would pull out trees from forest buffer if payments were stopped arose; there was no agreement, with some blocks indicating farmers would not uproot the trees and others indicating that they would. Capacity for these BMPs does seem to be relatively high, with only a few mentions that area for forest buffers might be limited. Funding appears available for buffers, wetlands, and stream/shoreline restoration at least in some states, though not in others. As with the other types of BMPs, concerns were raised about credits in CAST, though forest buffer was the only category where at least one block of text indicated that it was OVER-credited in the model. Verification for CAST appears as an issue again as well, though results are too sparse to interpret reliably.

6.7 Positivity Bias

Positive factors or reasons to include BMPs were mentioned more often than their negative counterpart or reasons to exclude BMPs for almost all factors described by respondents. Exceptions include Remaining Capacity LOW, which was mentioned as a reason to exclude or minimize BMPs much more often than SOME Remaining Capacity was mentioned as a reason to include them. UNDER-credited in CAST and CAN'T Verify for CAST were also mentioned more frequently than their positive counterparts, though not always as reasons to exclude BMPs from the WIP. This suggests that respondents may have some positivity bias when thinking about BMP selection, except as it relates to remaining capacity or use of CAST. Positivity bias was most clear for Popular/Preferred codes, where all jurisdictions other than PA were at least slightly more likely to mention Popular/Preferred BMPs than NOT Popular/Preferred BMPs. Most jurisdictions were also more likely to describe Effective/Moderately Effective BMPs instead of Low/Less Effective BMPs, though Maryland was split evenly between the two categories and the Multi/Independent group was more likely to describe Low/Less Effective BMPs. Positivity bias was not observed for Feasible/Cheap/Easy except possibly for Maryland. In fact, Pennsylvania and Virginia were somewhat more likely to mention NOT Feasible/Cheap/Easy BMPs and most other jurisdictions were split fairly evenly. Federal respondents showed the strongest Positive Bias for Popular/Preferred and Effectiveness indicators but did not mention Feasible/Cheap/Easy to Implement.

Consideration of these factors in aggregate across BMPs is perhaps more interesting than looking at them separately. The first thing that stands out is that respondents generally evaluated the BMPs that they discussed as Popular/Preferred (m = 64) with a much greater frequency than those identified as NOT Popular/Preferred (m = 5). Similarly, statements that a particular type of BMP was Effective/Moderately Effective (m = 31) were more common that statements that it was either Mixed/Variable Effectiveness (m = 6) or Low/Less Effective (m = 12). In fact, as shown in Figure 55, positive factors—those that increased the likelihood of selection—were mentioned more frequently than negative factors—those that decreased the likelihood of selection—almost all descriptors. This suggests that there may be some positive thinking bias in responses, as respondents may have focused more on the positive aspects of BMPs or chose to discuss BMPs that they felt were more likely to be included in the state's WIP. The one exception to this observation was the variable Under-credited in CAST, which was mentioned in 9 blocks of text and so was a much higher frequency than any other comments regarding considerations about the model in this set of BMPs. This may reflect general dissatisfaction with the way that CAST accounts for BMPs, which would fit with the finding that accounting for N,P was considered to be an important challenge.

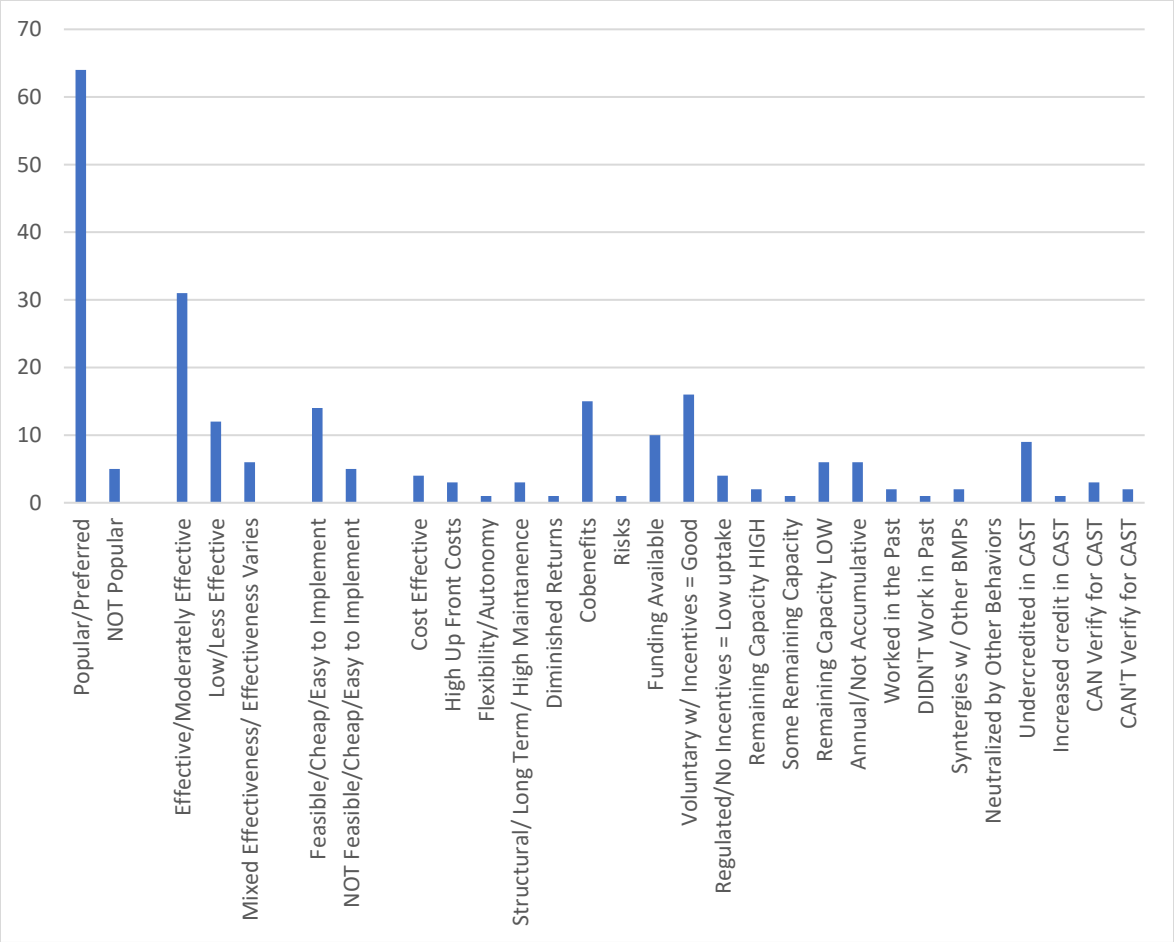


Figure 55: Factors affecting inclusion of Any Type of BMP in WIPs (# Mentions). For the most part, positive factors or reasons to include BMPs were mentioned more often than negative factors or reasons to exclude BMPs. Exceptions include Remaining Capacity LOW, which mentioned as a reason to exclude BMPs much more often than SOME Remaining Capacity was mentioned as a reason to include them. UNDER-credited in CAST and CAN'T Verify for CAST were also mentioned more frequently than their positive counterparts, though not always as reasons to exclude BMPs from the WIP.

Looking at the factors with high enough blocks of text and obvious positive/negative dichotomies, this positive bias is less consistent across states and characteristics. Tendency to talk more about positively evaluated BMPs was most consistent for the Popular/Preferred vs. NOT Popular/Preferred dichotomy, which was fairly consistent across respondent jurisdictions (see Figure 56) and general roles (see Figure 57). Only respondents from Pennsylvania, the Multi/Independent category, and the Communication general role mentioned about as many BMPs that they identified as Popular/Preferred as opposed to NOT Popular/Preferred. None of the groups mentioned NOT Popular/Preferred more than Popular/Preferred BMPs. This holds for Effective/Moderately Effective vs. Low/Less Effective when broken down by jurisdiction, though Washington, DC, West Virginia, and the Federal category were the only groups to show a strong tendency to focus on more effective BMPs. Interestingly, respondents in the Scientist General Role also mentioned only Effective/Moderately Effective BMPs, but other General Role categories were more balanced. Respondents in the Communication General Role mentioned BMPs that they classified as Low/Less Effective slightly more often than BMPs that they classified as Effective/Moderately Effective.

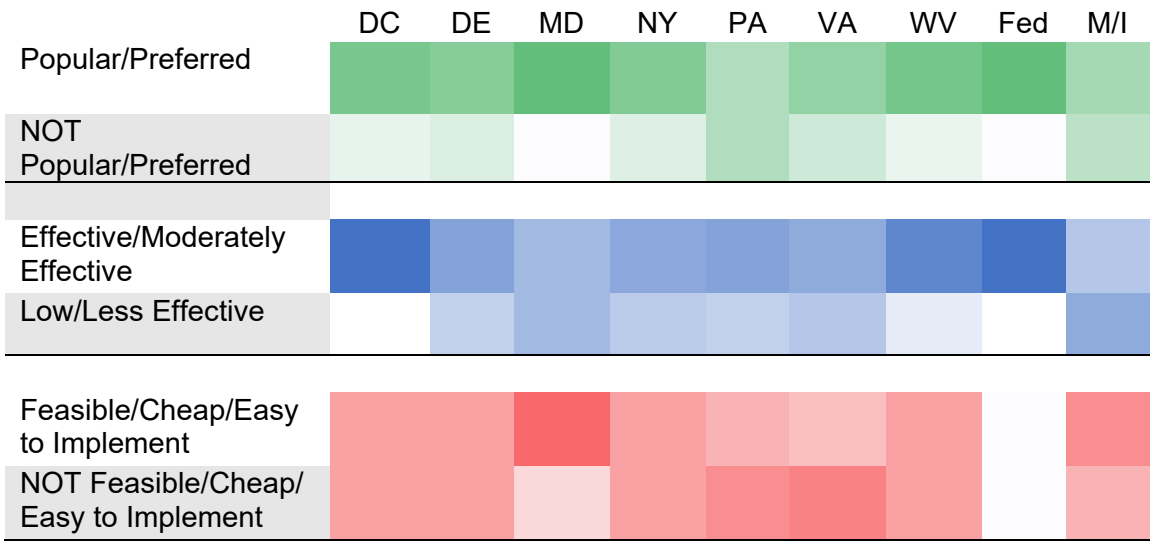


Figure 56: Heatmap of Average description of BMPs on positive vs. negative aspects of given characteristic by Respondent Jurisdiction. Positivity bias was most clear for Popular/Preferred codes, where all jurisdictions other than PA were at least slightly more likely to mention Popular/Preferred BMPs than NOT Popular/Preferred BMPs. Most jurisdictions were also more likely to describe Effective/Moderately Effective BMPs instead of Low/Less Effective BMPs, though Maryland was split evenly between the two categories and the Multi/Independent group was more likely to describe Low/Less Effective BMPs. Positivity bias was not observed for Feasible/Cheap/Easy except possibly for Maryland. In fact, Pennsylvania and Virginia were somewhat more likely to mention NOT Feasible/Cheap/Easy BMPs and most other jurisdictions were split fairly evenly. Federal respondents showed the strongest Positive Bias for Popular/Preferred and Effectiveness indicators but did not mention Feasible/Cheap/Easy to Implement.

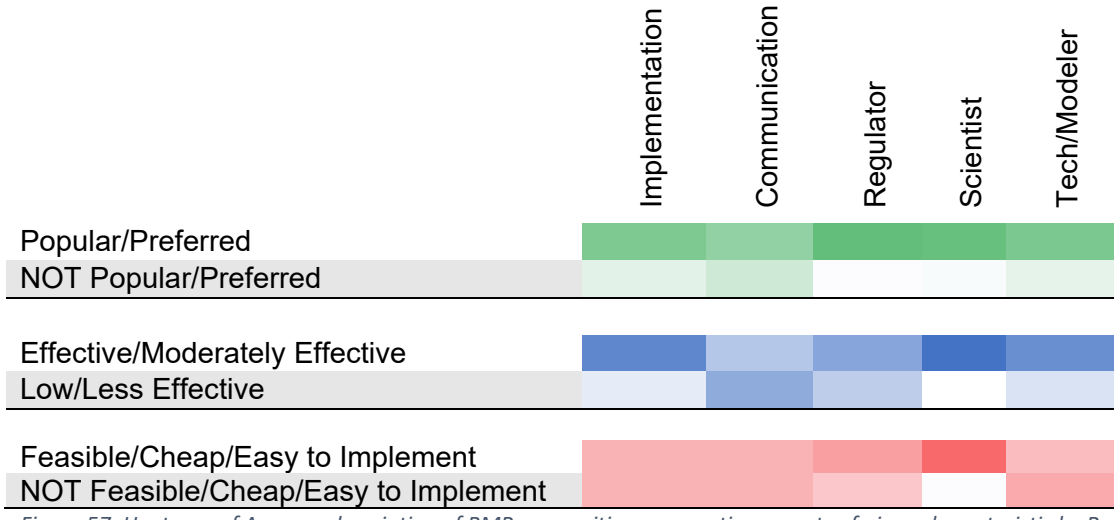


Figure 57: Heatmap of Average description of BMPs on positive vs. negative aspects of given characteristic by Respondent General Role. Positivity bias was clearest for strongest for the Popular/Preferred category of BMP attributes. Only the Scientists General Role consistently mentioned only positive variants of Popular/Preferred, Effectiveness, and Feasibility but the number of observations was low for this group. Respondents in the Communication General Role were slightly more likely to mention Popular/Preferred BMPs over NOT Popular/Preferred but were more likely to mention Low/Less Effective BMPs over Effective/Moderately Effective BMPs and were evenly split on Feasibility. Respondents in Implementation and Tech/Modeler General Roles showed at least some tendency toward positivity for Popular/Preferred and Effectiveness categories but not for Feasibility.

Positivity bias was least visible in the last set of codes shown in the figures above, with most respondent groupings describing BMPs they deemed "Feasible/Cheap/Easy to Implement" about as often as they

described BMPs that they deemed "NOT Feasible/Cheap/Easy to Implement (Expensive)". The only groupings that retained a strong positive focus on this characteristic were respondents in the Maryland Jurisdictional group and those in the Scientist General Role. Respondents from the Pennsylvania and Virginia Jurisdictions were slightly more likely to talk about BMP that they deemed NOT Feasible/Cheap/Easy to Implement, and Scientists again showed a strong focus on BMPs with positive characteristics, but since this group discussed specifics of BMP selection much less often than other groups, the number of mentions is low across all characteristics (m = 10 vs. 29-58), so the finding is not robust.

7 Implementation

While one might expect that WIP Design would guide Implementation, respondents instead tended to describe how insights into Implementation could help improve WIP Design. Although these results are interesting, we believe that they are primarily guidance for future research because the number of observations is so low relative to the scope and variation in Implementation across the watershed.

This section describes how WIPs are implemented, with particular focus on how states work to install and maintain BMPs on the ground. Further reinforcing our finding that the process is complex, there was considerable conflation between WIP implementation, WIP Effectiveness, and WIP Design, so we had to be careful when coding to differentiate among these processes. For the most part, the relationship depicted between Design and Implementation is largely unidirectional but not in the way one might expect; rather than WIP Design guiding implementation, respondents focused more on how an understanding of implementation could be used to improve WIP Design. This reinforces findings in the section on WIP Design/Allocation which indicated that implementation-related factors like popularity, feasibility, and cost effectiveness were important to consider when selecting BMPs for the WIP. That said, these factors were not always identified as most important on the implementation side. Although building on the past/learning what does and does not work was important in the WIP Design/Allocation responses, we see much less of this in the Implementation responses. Interestingly, inclusion in the WIP itself was not mentioned as a factor in the implementation process, but this may be because respondents took this connection for granted.

The rest of this section loosely parallels the sections above on WIP Design, describing factors that were thought to affect implementation first, then looking at sectoral aspects of implementation, and finally delving into the implementation of specific types of BMPs. However, because the number of blocks is so small, we rarely break results down by jurisdiction or respondent characteristics.

Methods Note: Some respondents described evidence of effectiveness when asked to describe the process of WIP Implementation. Usually, these were respondents with expertise at the TMDL/Allocation level who had little first-hand knowledge of the implementation process and were focused more on the TMDL and related environmental goals. Because of their main focus, these statements were not coded as IPROC (Implementation Process) but were instead coded as SProc (System Process; see Section 9 System Process) and/or Effectiveness (see Section 3 Overall Effectiveness). On the other hand, when asked about WIP Design, a few respondents started talking instead about WIP Implementation. This was more frequent among respondents with expertise in implementation but occurred with other areas of expertise as well. As explained in Section A.2 Data Collection and Analysis, we consolidated text about each stage in the TMDL process, in this case moving statements about WIP Implementation from the

Design section into the Implementation section, so that they could be analyzed separately. Readers should still keep in mind that these boundaries are not as clear in the actual responses.

One additional caveat to keep in mind for this section is that both the sampling process (fewer respondents with implementation expertise) and limited time for interviews resulted in fewer blocks of text for Implementation compared to Design/Allocation. It also seemed like people found it harder to describe implementation generally because there are so many disparate components of the process. For instance, WIP Design/Allocation is generally led by a single state agency which may take into account input from other agencies, but there is a clear bureaucratic scaffolding for the process. In contrast, implementation is led by the core state agency but the actual work on the ground may be facilitated by other agencies at the state or federal levels as well as various non-governmental organizations. For all these reasons, we do not report many state-level findings and even consider most of our aggregate results to be a guide for future research more than a definitive description of perceptions of the implementation process.

7.1 Implementation Aggregated

Human resources, particularly Boots on the Ground, and Monitoring/Enforcement were mentioned most often as factors that could improve implementation. Interestingly, the lack of each was cited much more often as a reason for lack of implementation than the presence of either was cited as a reason for increased implementation. Emphasis on monitoring and enforcement is an interesting contrast to a number of statements which indicate that voluntary measures are more likely to be implemented. Most of the other variables in this section parallel factors mentioned as important for WIP Design but learning from the past is much less prevalent. Exogenous legal and political factors were also mentioned much more often here than in descriptions of WIP Design. In addition to the CBP/EPA, NGOs and Local Governments were mentioned most often as actors with influence over implementation. Respondents also expressed concerns about the verification of BMPs and the way that they are accounted for using CAST, with results supporting earlier findings that a majority of responses indicate BMPs are UNDER-verified, meaning that respondents believe that actual loading has been reduced more than shown by loading estimates in CAST.

First, we consider the factors identified as affecting the implementation process across all states and BMPs. As shown in Figure 58, while some Implementation factors are similar to WIP Design and TMDL factors, others are quite different. Building on past experience is mentioned less often, with only one specific mention of working to expand existing programs. Feasibility remains important but is broken down into component parts such as providing incentives, finding funding, and getting stakeholder buy-in, which is similar to popular/preferred but much more targeted at those stakeholders who would need to implement the BMPs. In addition to available funding/resources (m = 5), human resources, particularly “boots on the ground” or people who can work with stakeholders to ensure implementation proved important both in terms of presence increasing uptake (m = 6) and absences decreasing uptake of BMPs (m = 6). This response was more prominent for Pennsylvania than for any other state. In several blocks of text, respondents indicated that they were not able to spend all of their financial resources if they did not have people who could go out into the field to help implement BMPs. Providing outreach (m = 5) was also an important task for those “boots on the ground”. It is interesting that more blocks indicated that voluntary BMPs were more likely to be implemented than not. Nevertheless, what stands out the most is the emphasis on the need for better monitoring and enforcement to increase

implementation (m = 2) or that too little monitoring and enforcement leads to low levels of implementation (m = 16).

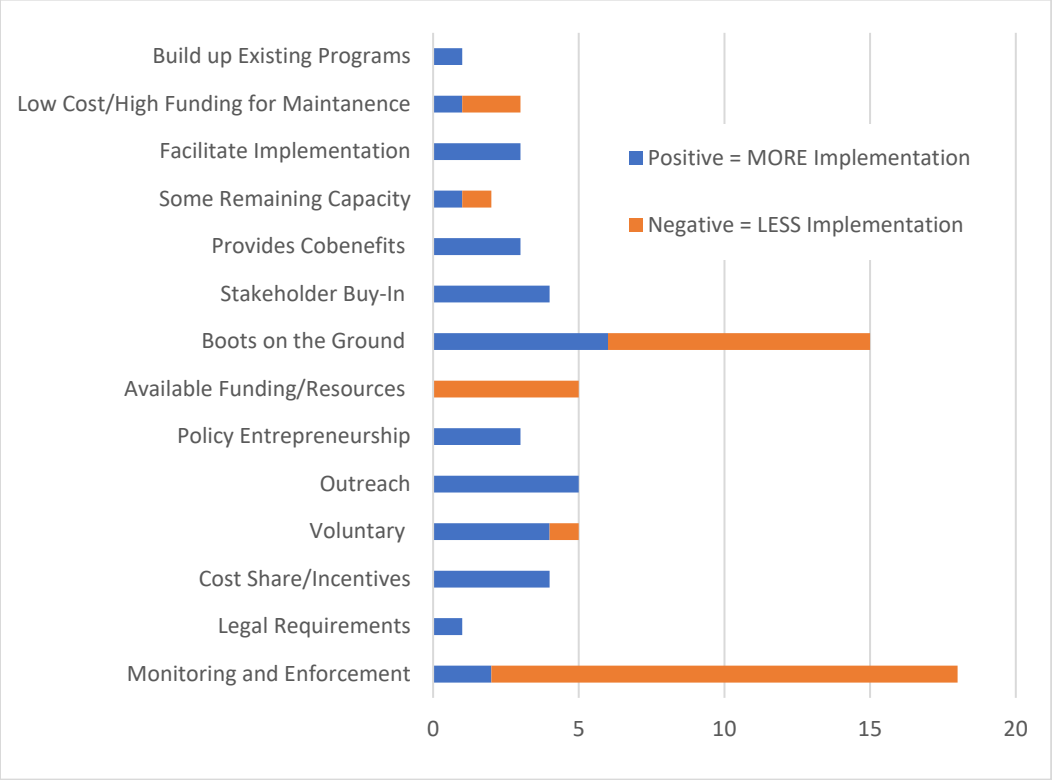


Figure 58: Major Factors Affecting Implementation (# Mentions). Human resources, particularly Boots on the Ground, and Monitoring/Enforcement were mentioned most often as factors that could improve implementation. Interestingly, the lack of each was cited much more often as a reason for lack of implementation than the presence of either was cited as a reason for increased implementation. Emphasis on monitoring and enforcement is an interesting contrast to a number of statements which indicate that voluntary measures are more likely to be implemented. Most of the other variables in this section parallel factors mentioned as important for WIP Design but learning from the past is much less prevalent. Exogenous legal and political factors were also mentioned much more often here than in descriptions of WIP Design.

There was little discussion of engagement or the hierarchy of the decision process in the section on implementation. This may be because we did not ask whether the implementation process was bottom-up or top-down, as we did for the TMDL/Allocation or WIP Design/Allocation processes. Engagement was only mentioned in a few blocks and of those, at least one indicated that it was performative (stakeholders did not really have an opportunity to influence implementation). More blocks described the implementation process as “top down” than “bottom up” as well.

Although legal requirements were not mentioned frequently as endogenous factors in WIP Implementation, exogenous legal and political factors were mentioned much more often in this section than in others. Local laws/politics were mentioned in 9 blocks of text. This group includes: 1) competition or other barriers to cooperation among local governments, 2) opposition to raising taxes or increasing regulation, 3) limited resources or attention from local leaders, 4) lack of education (among stakeholders), 5) difficulties attaining economies of scale, and 6) “hyperlocal” politics (m = 4). State laws/politics were mentioned in 9 blocks of text. Again, these were mainly barriers to implementation, such as: 1) high levels of bureaucracy/red tape, 2) partisan politics or lack of support from state legislatures, 3) laws that limit hiring or make it more difficult to retain “boots on the ground”, and 4)

high turnover in various parts of state government. National laws/politics were mentioned less often (m < 5) and referred to changes in enforcement due to shifts in presidential leadership and barriers to cooperation or information sharing among federal agencies. Five blocks of text also identified concerns about equity in the implementation process, especially in the distribution of funding at the local level.

Groups with influence over the implementation process are also quite different from those identified during the WIP Design/Allocation portion of the interviews (see Figure 59). NGOs were mentioned more often than most other groups in this context (m = 11), in large part because certain NGOs like the Chesapeake Bay Foundation, the National Fish and Wildlife Foundation, and the Upper Susquehanna Coalition are important intermediaries between stakeholders and state or federal government agencies. They provide outreach and engagement services, and some manage large government grants along with private funds for implementing BMPs. Interestingly, stakeholders and industry were mentioned less often here, and farmers were not singled out from other stakeholders as they were in the WIP Design section. Local governments were mentioned more often, however, as were other federal agencies, particularly the USDA and its various grant and extension programs. The role of the CBP/EPA was mentioned somewhat more often than for WIP Design, with responses focused primarily on either funding/support (m = 5) or enforcement (m = 7). The latter in particular included potentially reducing access to funds for states that fail to comply with their loading goals. Some mention was also made of conflict between state and federal agencies and a perception that federal agencies are out of touch with state/local level concerns about implementation.

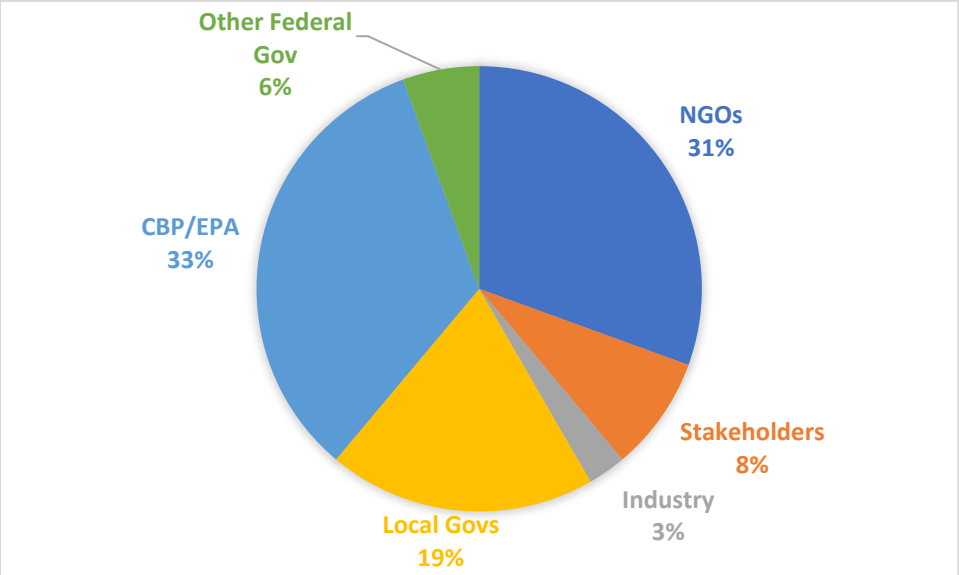


Figure 59: Groups with Influence in the WIP Implementation Process (% Mentions). The CBP/EPA was mentioned most often followed closely by NGOs when respondents described influence on Implementation. As described in Section 8.2 Funding Process, this may be because both groups play important roles in providing funds for implementation. Local Governments were also mentioned fairly often, with fewer mentions for Stakeholders, Industry, and Other Federal Government Agencies.

Although it was not discussed as often in the context of implementation, 6 blocks of text mentioned that planning at this stage was a paper process—or that much effort was put into coming up with plans for implementation but that this did not always lead to real changes on the ground. Of these, 5 indicated that the gap between planning and actual implementation reduced the legitimacy of the process while

the remainder indicated that it was still the "best we can do". Another 5 blocks of text indicated that implementation was improving, however. In a few blocks, improvement was attributed to increasing resources and better incentives to implement voluntary programs. There were also some mentions of implementation becoming more difficult, particularly as capacity limits were reached for BMPs that were considered "low hanging fruit" (combination of popular/preferred, effective, and feasible/easy). This will be discussed further in Section 9 System Process.

We were also somewhat surprised that CAST was mentioned often when respondents described the implementation process (m = 26). We did not ask about CAST specifically, as we did for the WIP design process, but it came up because CAST is also used for "accounting" or determining the amount of 2025 equivalent load reduction that states can count against their loading goals given verified implementation of BMPs. 16 blocks of text described difficulties collecting data, entering data into the National Environmental Information Exchange Network (NEIEN), or otherwise verifying BMPs so that they could be accounted for in CAST. Concerns were also expressed regarding the effects of verification on accounting, particularly whether or not BMPs on the ground were under-verified (meaning that more was being done than could be verified and therefore CAST was under-estimating the load reduction for a given state), over-verified (meaning that less was being done on the ground than was being verified, so that false verifications were leading to an over-estimation of load reduction for a given state), or correctly verified (meaning that all BMPs on the ground were verified and so estimates of changes in load reduction appropriately accounted for implementation of BMPs). The majority of blocks indicated that BMPs were under-verified, as shown in Figure 60. Note that this is not the same as accounting issues described in the section on WIP Design, which were primarily focused on the parameterization of the model (e.g., amount of load reduction ascribed to a given BMP in a given location) rather than the ability to enter all existing BMPs into the system.

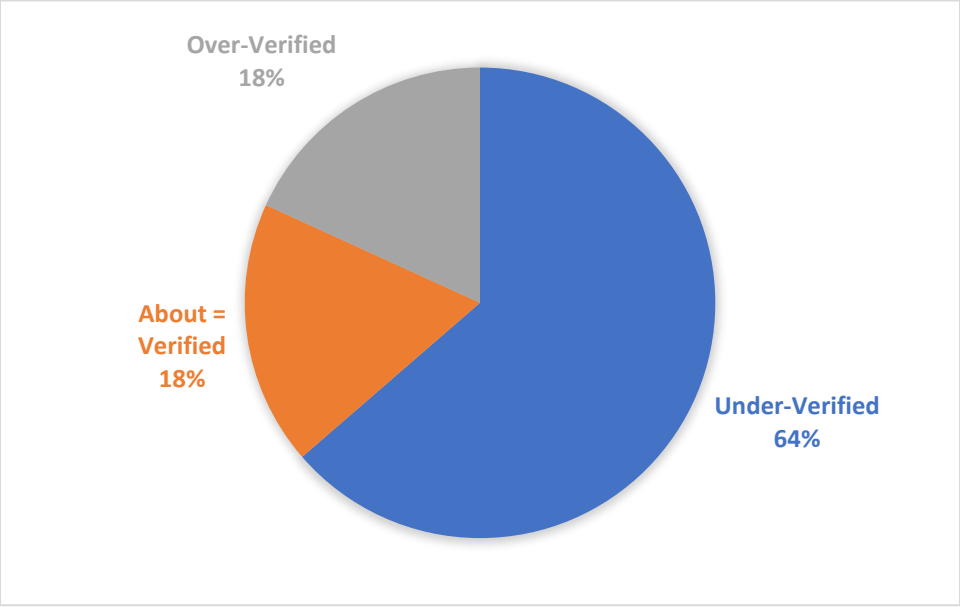


Figure 60: Perceived level of verification for implemented BMPs (% of statements on this topic). A majority of statements on this topic indicated that BMPs on the ground were not fully counted against loading goals because they were not verified and so could not be entered into CAST (UNDER-Verified), with the remaining statements split between beliefs that BMPs on the ground are close to fully verified and therefore accurately represented in CAST (About = Verified) or that there are fewer BMPs on the ground than are verified in CAST (OVER-Verified).

7.2 Sectoral aspects of WIP Implementation

Agriculture was discussed much more than any other sector when respondents described the Implementation process. Wastewater also received more attention than other sectors, followed by Stormwater and Urban/Development. Transfers of loading reductions between sectors were mentioned multiple times, particularly higher reductions in Wastewater and Stormwater to make up for low rates of reduction in Agriculture. Responses clearly indicate that Wastewater is the sector with the highest level of implementation, whereas Stormwater and Agriculture were described as having Low Implementation Rates much more often than High Implementation Rates. Agriculture was described most often as NOT Feasible/Cheap/Easy to Implement. Factors deemed most likely to increase implementation in Agriculture were more Cost-Share and Increased Outreach/Respect for Stakeholders. In part due to high levels of cost share, Equity Issues and Regressive Impacts of implementation in Agriculture were also mentioned often as negative. It is interesting that these environmental justice variables were more prominent when respondents discussed Implementation rather than WIP Design or the TMDL/Allocation Process.

Discussion of sectoral components of implementation (m = 67) focused primarily on agriculture (m = 47), which also received high levels of attention in WIP Design, though not much higher than the other major sectors (stormwater, wastewater). As shown in Figure 61, Wastewater, which received the most attention in the WIP Design section, received about half the number of mentions as Agriculture in the Implementation section (m = 23). Stormwater was mentioned even less often (m = 13) and was on par with Urban/Development (m = 13), which received much less attention in the WIP Design section. In contrast, Point (wastewater) and Non-point (stormwater, agriculture) pollution were mentioned much less often as categories in the Implementation section, and a new “sector”, or really the set of Regulated Sectors (wastewater, urban stormwater) were mentioned in one block of text. Most interesting, transfers between sectors, particularly implicit transfers of loading reduction from stormwater, wastewater, or urban sectors to make up for shortfalls in agriculture were mentioned in 6 blocks of text. This was not described at all in the section on WIP Design/Allocation and may represent the “Blame Game” that was described in the section on Challenges, although it is also consistent with finding that design is influenced by the feasibility and cost effectiveness of BMP options.

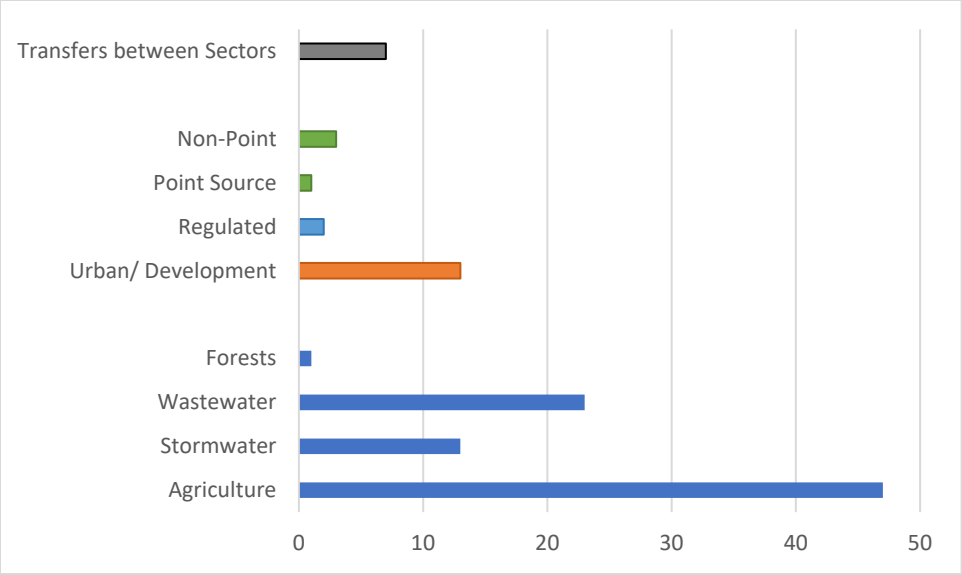


Figure 61: Blocks of Text Discussing Implementation by Sector (% Mentions for Category). Agriculture was discussed much more than any other sector when respondents described the Implementation process. Wastewater also received more attention than other sectors, followed by Stormwater and Urban/Development. Respondents mentioned the Regulated Sector as well, which we did not observe in other sections. Transfers of loading reductions between sectors also came up in statements about implementation, particularly higher reductions in Wastewater and Stormwater to make up for low rates of reduction in Agriculture.

This finding is further supported by the breakdown of specific sector-related statements about implementation (see Figure 62), which indicate low implementation for agriculture (m =7) and stormwater (m = 7) but high implementation for wastewater (m = 19). Wastewater was also the only sector seen to be highly effective in reducing nutrient loading (m = 14), with mentions of effectiveness for other sectors quite low (m < 5). The visibility of improved water quality from wastewater treatment was mentioned in 7 of the 14 statements about its high effectiveness, which contrasts with a few statements about the difficulty of collecting data or monitoring nutrient loading in other sectors. Statements about feasibility were also minimal, but mainly mirror perceptions described in the section on WIP Design/Allocation. Difficulties with implementation in the agricultural sector were mentioned in 5 blocks of text, including several discussing challenges associated with plain sect farms in Pennsylvania (e.g., Amish, Mennonite). Interestingly, political feasibility was mentioned in a few blocks, which we did not observe in the section on WIP Design/Allocation, though the importance of the "Popular/Preferred" variable may also reflect this idea of political feasibility. Co-benefits and Co-costs, as well as access to funding or other resources were also mentioned less often than we might expect in the section on implementation. This may be in part because respondents were reluctant to repeat themselves. It is also important to note that we coded statements about the funding process separately, and from that analysis we can say that the topic was still important for implementation (see Section 8.2 Funding Process).

	Agriculture	Stormwater	Wastewater	Forests	Urban/Dev
Low Implementation Rate	7	7			
High Implementation Rate			19		
Low/Less Effective					
Effective/Moderately Effective					
High/Visibly Effective			14		
Needed to Reach Goals					
Feasible/Cheap/Easy to Implement					
NOT Feasible/Cheap/Easy to Implement	5				
Politically Feasible					
NOT Politically Feasible					
Co-Benefits					
Co-Costs					
Easy to Get Funding/Resources					
NOT Easy to Get Funding/Resources					
Regressive/ Polluter DOESN'T Pay	7				
Equity Issues (Local)	11				
Voluntary/Incentives = Low Uptake					
Cost-Share = Good	8				
Increase w/ Outreach/ Respect Stakeholders	7				
Regulated = More Uptake			4		
Systemic Change/ Change Strategy	9				
Data Problems/ Need more Monitoring					
High Monitoring/Enforcement					

Figure 62: Heatmap of Mentions Describing Implementation by Sector (yellow = #M, grey = ANY M). Responses clearly indicate that Wastewater is the sector with the highest level of implementation, whereas Stormwater and Wastewater were described as having Low Implementation Rates much more often than High Implementation Rates. Given the findings in Section 6.4 Sectoral Aspects of WIP Design/Allocation, it is surprising that Agriculture was described most often as NOT

Feasible/Cheap/Easy to Implement. Factors deemed most likely to increase implementation in Agriculture were more Cost-Share and Increased Outreach/Respect Stakeholders. In part due to high levels of cost share, Equity Issues and Regressive Impacts of implementation in Agriculture were also mentioned often as negative. Observations were too small to go into detail on other sectors, but mandatory regulation was cited as a reason for higher levels of Implementation in Wastewater (Regulate = More Uptake).

There are several other interesting findings in Figure 62. For one thing, problems of equity were discussed more often than in previous sections, and respondents tended to focus on local-level concerns about the distribution of resources for implementation (m = 12) or the regressive distribution of costs (m = 7). In a few blocks, DEIJ concerns and lack of attention to the needs of under-represented groups were mentioned specifically as problems with the existing implementation process. There were also more blocks of text describing what would need to occur to improve voluntary implementation of BMPs in agriculture, including increasing or maintaining cost-share arrangements (m = 7) and improving outreach/showing respect for farmers to gain their confidence (m = 7). This is an interesting contrast to the perception that wastewater was effective because it was required by regulation (m = 4). Lastly, we see more statements that some type of systemic change is needed to improve implementation in agriculture (m = 9). This is a mixed bag of changes which can include increasing regulatory requirements for the sector, improving flexibility and innovation for agricultural BMPs, and either increasing regulations on small farms or steeply curtailing industrial-scale operations to make them more sustainable.

Breaking the sectors down by states does not yield much novel information. Agriculture was mentioned fairly often (m = 3-8) for all jurisdictions except Washington, DC (m = 0), and was mentioned much more often for Pennsylvania than for any other state (m = 8). Stormwater was mentioned in 1-2 blocks for Maryland, New York, Virginia, and West Virginia, but 4 blocks of text addressed stormwater implementation across all states combined. In fact, mentions that encompassed implementation across all states were more common for this section of the interview. Wastewater was mentioned for All states (m = 2), Maryland (m = 3), Virginia (m = 4), and West Virginia (m = 2), with 3 blocks of text describing high implementation rates for wastewater in each of three categories: All States, Maryland, and Virginia. This generally reflects the perception that Maryland and Virginia were able to meet much of their loading goals by investing in large-scale, high-quality wastewater treatment plants. Other sector-related categories were not mentioned frequently enough to break down by state.

7.3 BMP Implementation

The positivity bias observed in responses on BMP choice and WIP Design is not evident in responses about the implementation of specific BMPs. Co-benefits are still emphasized over Co-costs, but most other variables focus on negative aspects of BMPs such as NOT Popular/Preferred, Low/Less Effective, or NOT Feasible/Cheap/Easy to Implement. Respondents also described BMPs with Low Rates of Implementation much more than BMPs with High Rates of Implementation and mentioned several BMPs that had not been as easy to implement or effective as expected. In particular, Cover Crops appear much less of an obvious "win win" when respondents described their Implementation compared to the section on WIP Design/Allocation. On the other hand, Septic/Sewer Improvements were viewed more favorably in statements about implementation and some responses mentioned that there might be more capacity for wastewater treatment, especially in suburban or rural areas. One last area for future work is this idea that increasing attention to co-benefits can increase political will and willingness of stakeholders to implement certain BMPs. Though the evidence is week, the results above suggest that this is an assumption that should be tested, as it would seem that ***perceived co-benefits do not necessarily outweigh high costs/low feasibility or lack of popularity with stakeholders****. This section also highlights the* ***need for analysis of the BMP decks created in the WIP design process compared to decks of verified BMPs used in the milestone assessments****.*

Although many of the codes used to describe implementation of BMPs are similar to those used in the WIP Design/Allocation section, they do not explain why implementation is more or less likely to occur but instead provide detail on perceived implementation of BMPs. For example, while 7 blocks described at least one BMP as popular/preferred and 8 blocks described at least one BMP as NOT popular/preferred, we cannot assume that popular/preferred is necessarily linked to higher levels of implementation unless specifically stated. Thus, the bars in Figure 63 indicate high/positive, medium/varied, and low/negative levels of a given variable, rather than the influence of the variable on implementation. Interestingly, with this change in perspective, the positivity bias observed in the section on WIP Design/Allocation fades. For example, slightly more blocks of text refer to a particular type of BMP as "NOT Popular/Preferred" than "Popular/Preferred". Perceived rates of implementation are also less positively skewed, with 29 blocks indicating that implementation rates are low for specific BMPs, 6 indicating medium or varied implementation rates, and 19 indicating high rates of implementation. Discussion of effectiveness (especially in CAST) was much lower than in other sections and was weighted toward low or medium, with no highly effective BMPs identified. BMPs that were difficult or NOT feasible/cheap/easy to implement were also mentioned more frequently than those that were viewed as Feasible/cheap/easy.

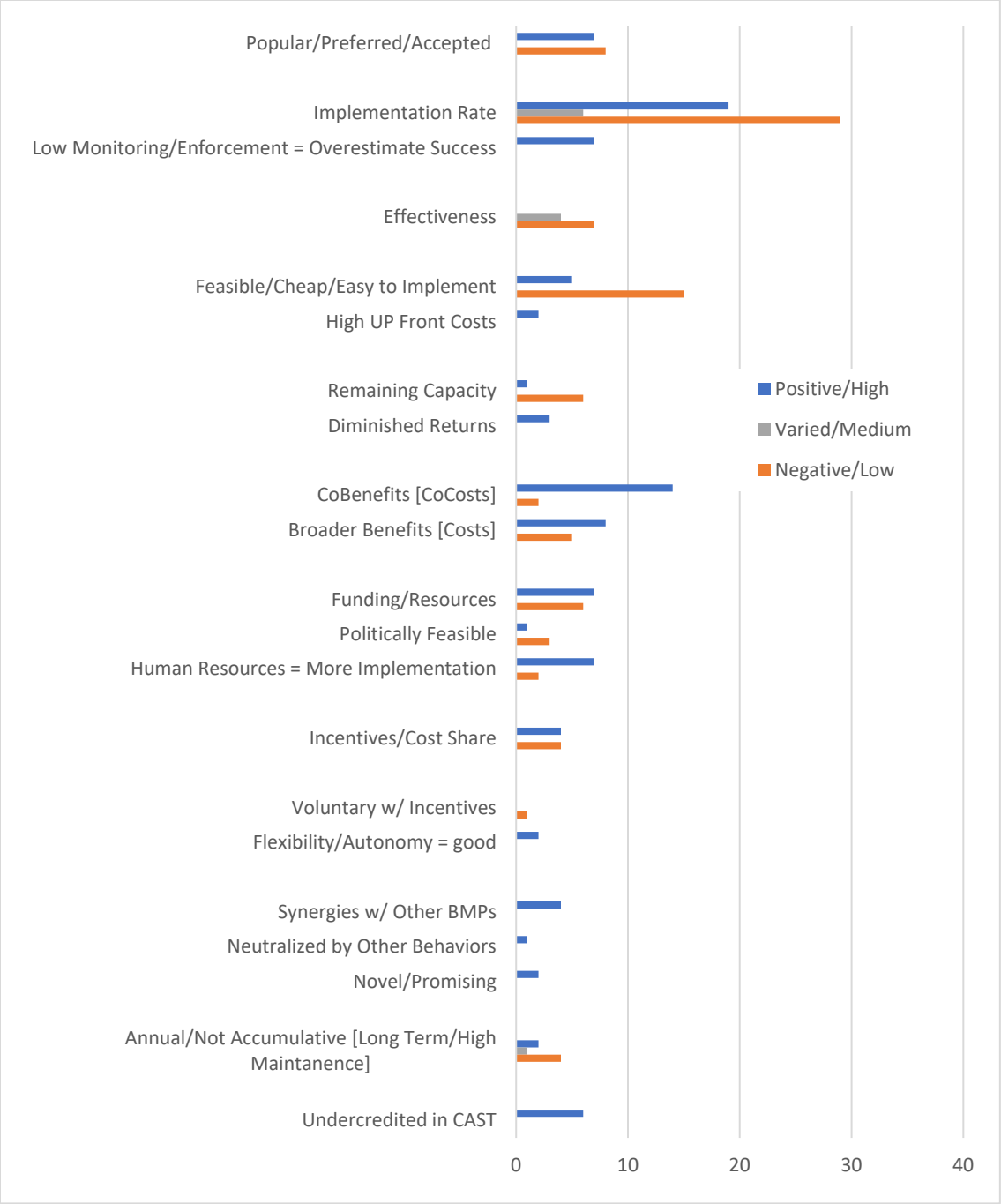


Figure 63: Factors Associated with the Implementation of Specific BMPs (# Mentions). Most comments described the rate of implementation for specific BMPs. Most other factors parallel those described for WIP Design but there is no observable positivity bias here. In fact, respondents were slightly more likely to talk about BMPs with negative evaluations of factors such as Popular/Preferred, Effectiveness, and Feasible/Cheap/Easy to Implement. Co-benefits and Broader societal Benefits remain more common than Co-costs or Broader societal Costs.

On the other hand, co-benefits were mentioned much more frequently than co-costs, as were broader benefits for society or the economy. The latter largely consisted of increased robustness in the face of climate change. Nevertheless, broader economic costs were still mentioned, particularly the regressive impacts of the way that certain BMPs are funded. Political feasibility also came up more on the negative

side for specific BMPs, including observed local pushback and missed windows of opportunity to take advantage of political will. There were also fewer statements about the political benefits of voluntary BMPs in this section, and no statements about regulatory requirements and implementation of specific BMPs. The importance of human resources—particularly boots on the ground and individuals who could conduct outreach or work with local governments—was highlighted as a factor that could increase implementation of certain types of BMPs (m = 7) or reduce implementation if absent (m = 2). Cost share was also mentioned in a similar way, with more cost share considered likely to increase implementation for some BMPs.

The last few variables in Figure 63 were only observed in a few blocks of text, but they do point to some interesting areas for future research. Synergies among BMPs were mentioned more in this section than others, though still not very much, and the potential to neutralize the effectiveness of BMPs through other behaviors also came up. A few blocks also mentioned novel or promising new BMPs, which contrast to the general antipathy to innovative BMPs found in the section on WIP Design/Allocation. Statements about annual/not accumulative BMPs getting more uptake than long/term or high maintenance BMPs reinforce those findings in the WIP Design Section. Lastly, though CAST was not mentioned very often, 6 blocks did note that a certain type of BMP was under-credited in the model.

Breaking down the BMPs by type, the first noticeable difference that jumps out is the increased attention to multi-sector BMPs and reduced attention to Agricultural BMPs when compared to the WIP Design/Allocation Section (see Figure 64). In fact, wastewater received about the same level of attention in each section, even though the number of blocks is much smaller for implementation than WIP Design/Allocation. Stormwater received much less attention, as would be expected, and MS4 permits dropped out of the discussion entirely. For agriculture, cover crops remained the dominant BMP, but there was much less discussion of others. Total number of mentions was also lower for buffers, but it remains important relative to other BMPs in this section, as were other types of forestry efforts, stream/shoreline restoration, and wetlands. Floodplain restoration and flooding mitigation did not come up in discussions of implementation, however, which is not unexpected given that this was largely described above as a BMP that should be included in future WIPs. Several of the BMPs in the "Others" categories in Figure 48 were mentioned more than once and so are shown here.

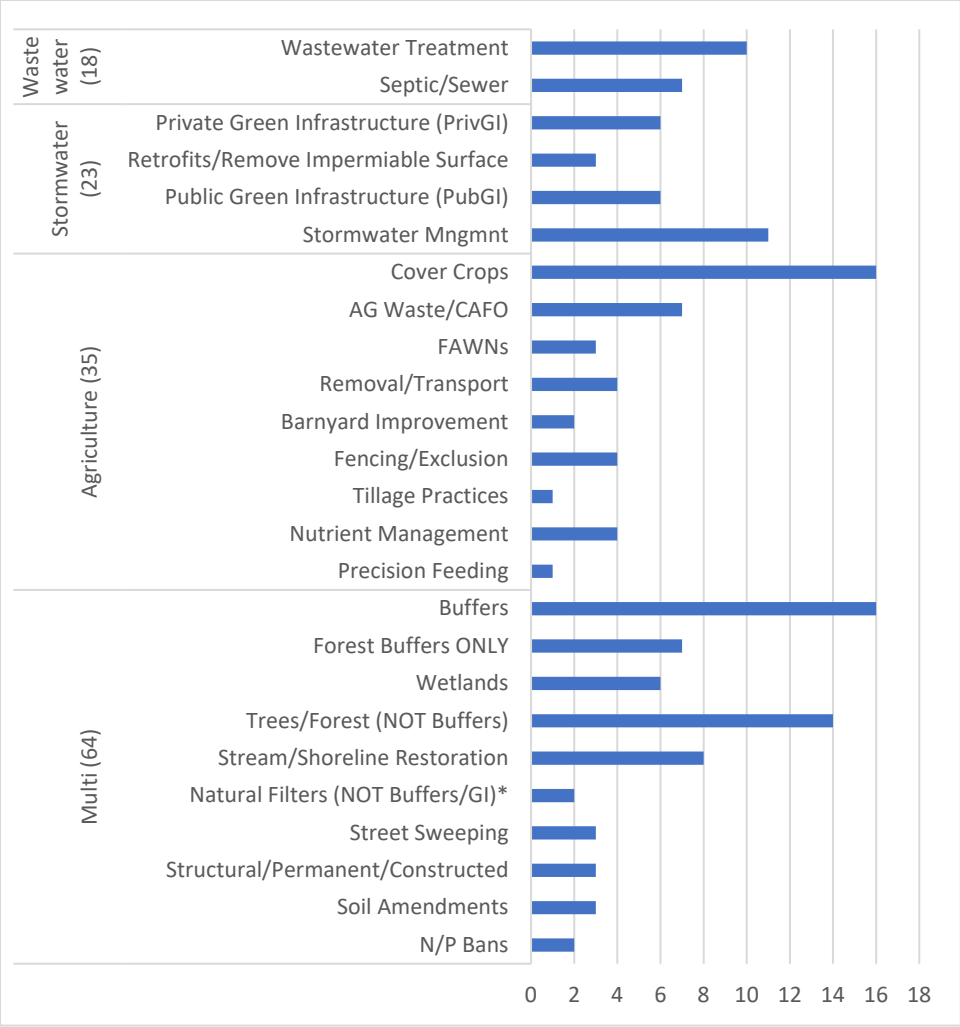


Figure 64: Types of BMPs for which Implementation was Described (#Mentions).

Finer grain cutting of the data for implementation by BMP generally leads to very low numbers of blocks per category, but some relationships are more robust than others. Figure 65 shows the breakdown of factors associated with implementation for those types of BMPs with relatively high number of blocks and, therefore, more blocks per factor than the other BMPs. Cover Crops appear much less of an obvious "win win" here compared to the section on WIP Design/Allocation. More blocks of text describe them as NOT Popular/Preferred/Accepted, perceived implementation rates are split between high and low, and levels of effectiveness are definitely on the low end. It is interesting that more blocks indicate that cover crops were less popular and/or effective than expected (overestimated success) compared to other types of BMPs. Co-benefits to cover crops were mentioned fairly often, but these include increased profits to farmers due to cost-share programs.

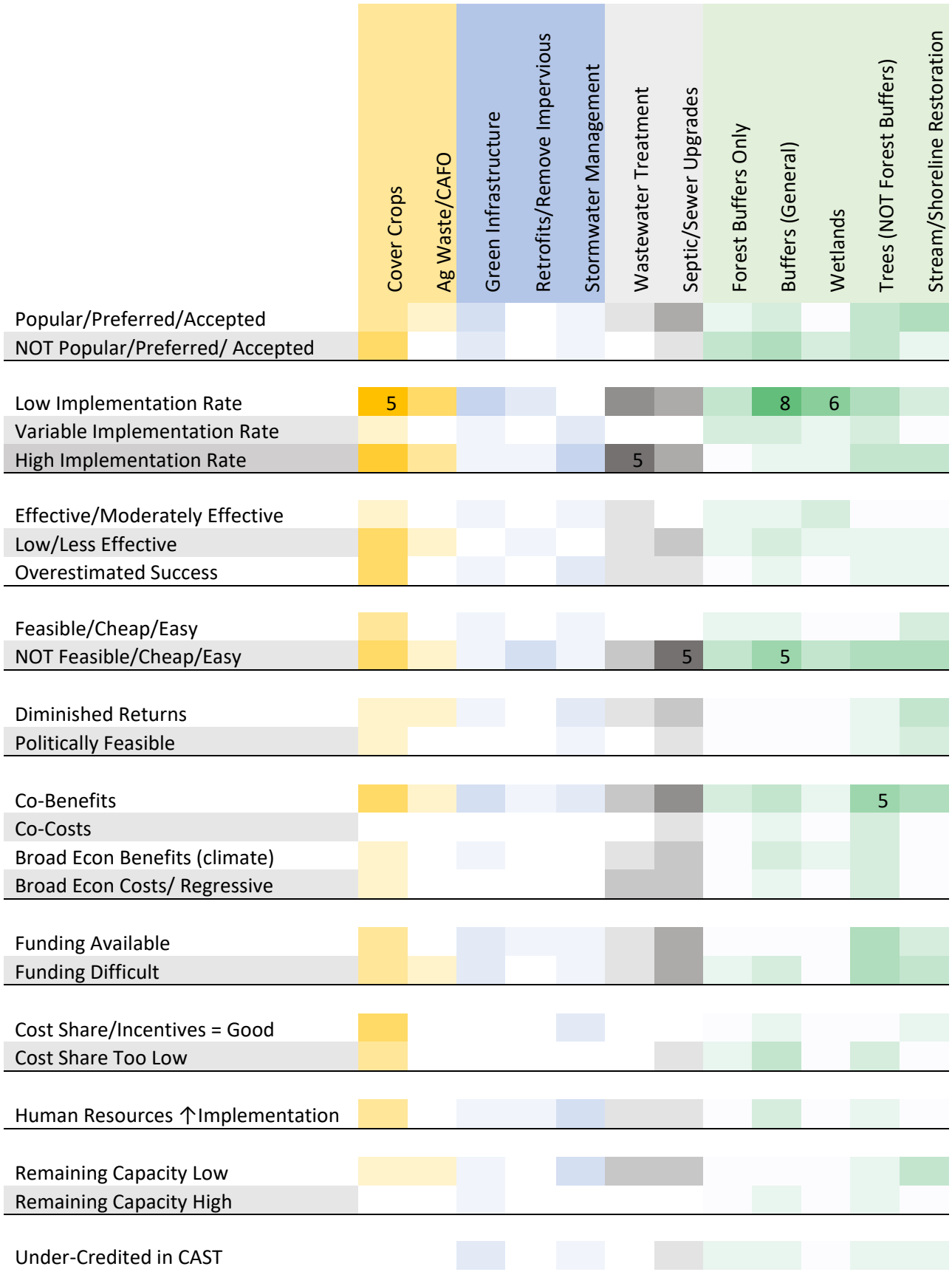

Cover Crops
Ag Waste/CAFO
Green Infrastructure
Retrofits/Remove Impervious
Stormwater Management
Wastewater Treatment
Septic/Sewer Upgrades
Forest Buffers Only
Buffers (General)
Wetlands
Trees (NOT Forest Buffers)
Stream/Shoreline Restoration
Popular/Preferred/Accepted
NOT Popular/Preferred/ Accepted
Low Implementation Rate
Variable Implementation Rate
High Implementation Rate
Effective/Moderately Effective
Low/Less Effective
Overestimated Success
Feasible/Cheap/Easy
NOT Feasible/Cheap/Easy
Diminished Returns
Politically Feasible
Co-Benefits
Co-Costs
Broad Econ Benefits (climate)
Broad Econ Costs/ Regressive
Funding Available
Funding Difficult
Cost Share/Incentives = Good
Cost Share Too Low
Human Resources ↑Implementation
Remaining Capacity Low
Remaining Capacity High
Under-Credited in CAST
5
5
8
6
5
5
5

Figure 65: Heatmap of Factors associated with the implementation of selected types of BMPs (# Mentions). More blocks of text describe Cover Crops as NOT Popular/Preferred/Accepted, perceived Implementation rates are split between High and Low, and levels of Effectiveness are primarily Low. It is interesting that more blocks indicate that cover crops were less popular and/or effective than expected (overestimated success) compared to other types of BMPs. Co-benefits to cover crops were mentioned fairly often, but these include increased profits to farmers due to cost-share programs.

There is not much unexpected in results for stormwater or wastewater, though wastewater treatment has not been uniformly implemented at high rates and Septic/Sewer retrofits are viewed as popular/preferred with high co-benefits but also as difficult and expensive. Results for the multi-category are similar, except that low popularity is highlighted along with high difficulty/expense, which in turn lead to low implementation in spite of recognized co-benefits. Again, not much different from the results in the section on WIP Design/Allocation except insofar as there was a greater focus on those multi-sector BMPs. This difference may show some learning over time. That is, "difficult" BMPs may have been discussed less in the WIP Design/Allocation section because decision makers have learned through experience that they are less likely to be implemented. Then again, there is still potential for positivity bias in the WIP Design section, though not here in the section on Implementation.

One last area for future work is this idea that increasing attention to co-benefits can increase political will and willingness of stakeholders to implement certain BMPs. Though the evidence is weak, the results above suggest that this is an assumption that should be tested, as it would seem that perceived co-benefits do not necessarily outweigh high costs/low feasibility or lack of popularity with stakeholders. It may be that more outreach is needed, as indicated by a number of blocks of text, but it is not clear, even in the literature on cost-benefit analysis, that valuing ecosystem services provides sufficient incentives for people to undertake major changes in their behavior or otherwise pay more to protect those co-benefits. Fuzzy-set QCA might be useful in determining what is necessary vs. what is sufficient for BMP adaption, but a much larger dataset would be needed to attain reliable results. This section also highlights the need for analysis of the BMP decks created in the WIP design process compared to decks of verified BMPs used in the milestone assessments.

Methods Note: Because we describe responses rather than respondents, it is possible that some of the differences between responses in the WIP Design/Allocation Section and the Implementation Section occur because different respondents provided more or less information in each section, depending on their area of expertise. Figure 66 shows that responses per interview were biased toward WIP Design.

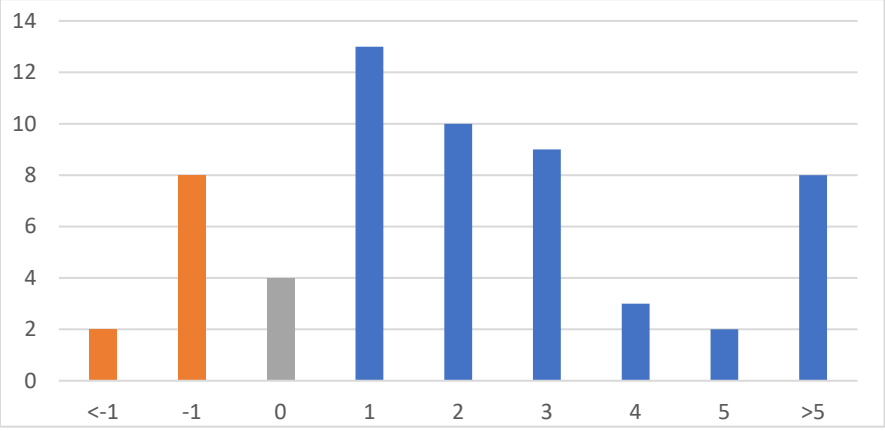


Figure 66: Difference between Number of Mentions per Respondent for WIP Design - Implementation. Negative numbers indicate the respondent made more comments about Implementation than WIP Design (m = -3). Zero indicates equal numbers of comments about WIP Design and Implementation. Positive numbers indicate that the respondent commented more*

frequently on WIP Design than Implementation (m = 11). Differences between results for Implementation and WIP Design could be biased toward the views of respondents who commented so much more frequently about WIP Design than Implementation.*

8 Other Processes

This section focuses on statements about the Modeling Process and the Funding Process, with some discussion of Climate Change, Conowingo Dam, Lawsuits, and Nutrient Trading.

A few other processes were discussed by respondents even though they were not covered in their own section of the interview instrument. As shown in Figure 67, the modeling process was most talked about (m = 106), followed by the funding process (m = 78), and the processes of negotiating amendments to the TMDL system to account for climate change (m = 41) and increased loading from the Conowingo Dam (m = 17). The role of lawsuits (m = 14) and nutrient trading (m = 9) were also discussed in some blocks of text. In this section, we will address each of these other processes in turn, focusing mainly on results that are relatively robust (m > 5) but also paying attention to areas where additional research might be useful. Since we did not ask about these processes specifically in all interviews, it is important to remember that these results will be biased toward the views of respondents who have strong opinions or higher levels of expertise on these topics.

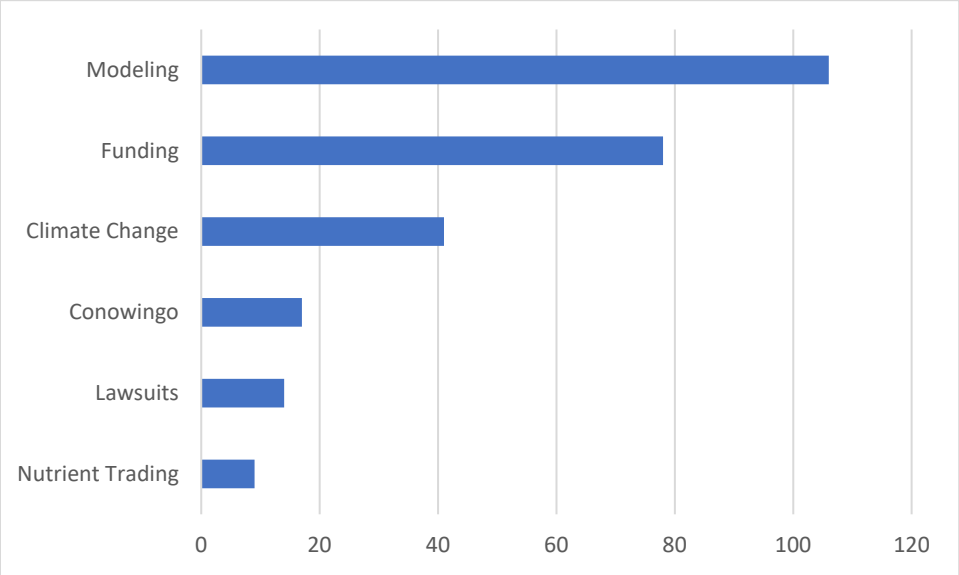


Figure 67: Annotations describing other processes (# Mentions). Of the Other Processes, Modeling was mentioned must often, followed by Funding and Climate Change. Conowingo Dam, Lawsuits, and Nutrient Trading were also discussed in some blocks of text.

8.1 Modeling Process

Respondents generally agreed that CBP Models, including CAST, were important in Chesapeake watershed governance, but the vast majority of comments evaluating the Model and/or the Modeling Process were negative. This pattern held across respondent characteristics including Jurisdiction and General Role. Dissatisfaction stemmed from the politicization and complexity of the modeling process, which parallels some of the concerns about Institutional Design Effectiveness for the CBP as a whole (see Section 3.2 Institutional Design) and related Challenges (see Section 4.1 Challenges). However, some responses also expressed frustration that relatively small changes in the model could have significant impacts on state loading goals and/or the real-world struggles of stakeholders and local governments. Respondents also identified a digital disconnect, suggesting that actors with better understanding of the models could have greater influence over governance as a whole through the modeling process. A majority of responses indicated that transparency could increase the legitimacy of the modeling process, though some argued otherwise.

Given that the role of the model was seen as central to the TMDL, WIP Design, and nutrient accounting processes, it is not surprising that the process of designing, refining, and parameterizing CAST and the Bay Models was also a focus of respondent attention. The majority of these blocks of text focused on evaluation of the model itself or the modeling process, with some also describing factors that influence model design and specification. The latter are easily summed up: 8 blocks of text indicated that the politics of allocation among states influenced the modeling process, while 5 pointed to the role of technical experts. The modeling process was also described only as Top Down/More Top Down (m = 7), with some of these blocks indicating that engagement was performative, and that communication was not satisfactory. This fits well with our findings from the framing analysis, which show that 32 blocks of text indicated a political frame for the modeling process, 12 evinced a technical frame, and only 5 described a participatory frame (see Section 2 Process Frames).

8.1.1 General Perceptions of the Model/Modeling Process

Recognizing the politicization of the modeling process is important because it is linked to dissatisfaction with the models and can point to areas where improvement has occurred or could occur in future. As shown in Figure 68, there was significant indication of general dissatisfaction with the model (m = 29) or the modeling process (m = 31), which is much larger than the few blocks that expressed satisfaction with either. If we had asked about the modeling process specifically, then it is likely that more blocks of text would indicate satisfaction or at least unconcern with the model/modeling process. Since we did not, the responses here primarily reflect the views of respondents who felt strongly about the issue. Their concerns are still important, however, and it is interesting to note that there were no statements indicating enthusiastic acceptance of the model or the modeling process (see below for more detailed breakdown of positive/negative statements by respondent characteristics). We also note that a total of 29 blocks of text indicated that respondents had trouble using CAST or previous version of the Bay Models as part of one of the three main processes described above (TMDL/Allocation, WIP Design/Allocation, Implementation).

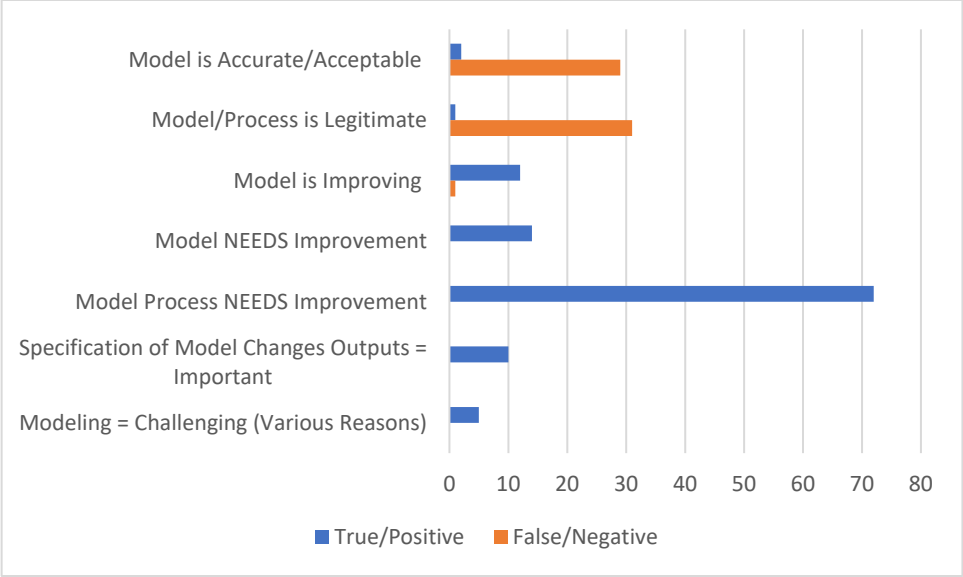


Figure 68: Qualitative Evaluation of the Bay Models (i.e., CAST) and the Modeling Process (# Mentions). Vast majority of responses that evaluate the Model indicate that it is NOT Accurate/Acceptable and that the Model or the Modeling Process is NOT Legitimate. Some blocks did indicate that the Model is Improving but many more indicated that the Model or Modeling Process NEEDS Improvement. Some comments were less value-laden, including concerns about how Model Specifications can change outputs (i.e., loading goals) in important ways and that the task of modeling such a large system is Challenging.

Before describing the factors that contributed to this dissatisfaction in greater detail, we also think it is important to note that 12 blocks of text indicated that the model/process has been improving, especially with the transition to a simpler version of the model that could be used by states and potentially by other stakeholders. That said, 14 blocks indicated that the model needs improvement and 72 suggested that the modeling process could be improved. Of these, some were general statements, but others were specific recommendations as described below. Other less value-laden statements about the model that occurred fairly often included pointing out ways that model specification could generate important (i.e., policy-relevant) changes in things like allocation of loading goals or accounting for BMPs (m = 10) and recognizing that developing the Bay Models is technically and politically challenging (m = 5).

Although some respondents were much more critical of the models/modeling process than others (13 with m* > 3, max M* = 12), 41 of the 59 respondents made some type of negative comment on this topic. In contrast, only 20 mentioned anything positive about the model or the modeling process. Interestingly, 14 of these made both positive and negative comments, and 5 were equally split between the two. Looking at the breakdown of positive and negative statements about either the model or the modeling process by the general role of our respondents, Regulators stand out as being the most critical of the model/process (m = 46; see Figure 69). Given that respondents in the "Communication" role were more critical of other processes, it is interesting to note that they were generally more positive than any other group aside from Tech/Modelers. From Figure 70, we note that respondents from Pennsylvania provide the largest number of both negative and positive blocks of text related to the model/process and that otherwise, the Federal and Multi-Independent categories also had higher numbers of positive blocks. At least some of this variation is due to just the level of attention paid to the models/modeling process by respondents from each state and the dominance of negative statements is consistent across all jurisdictional groups.

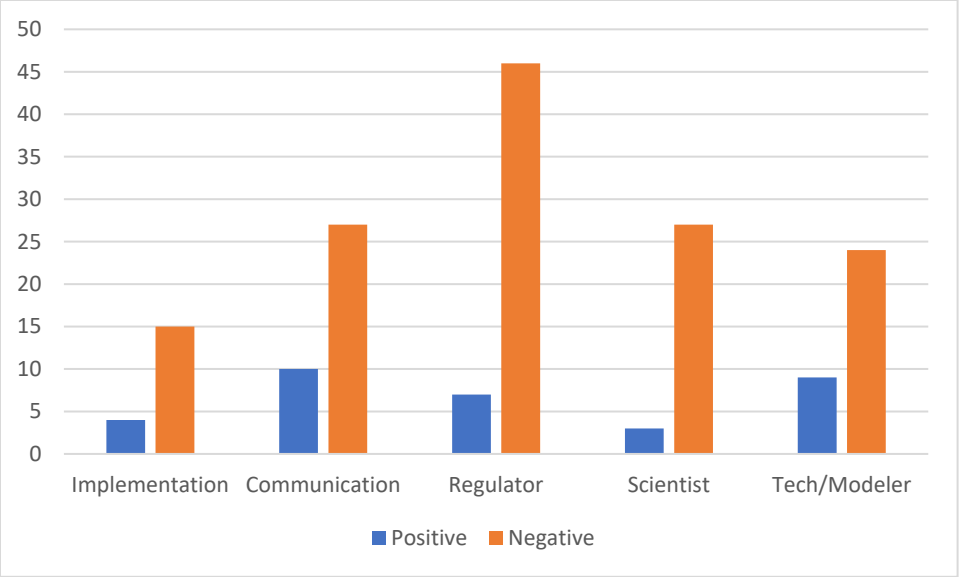


Figure 69: Positive vs. Negative statements about the Model or Modeling Process by General Role of Respondent (# Mentions). Respondents in the Regulator role made the most negative comments about the Model/Modeling Process. Respondents in the Communication and Tech/Modeler Roles made more positive comments about the Model/Modeling Process. However, the majority of comments were negative across all General Roles.

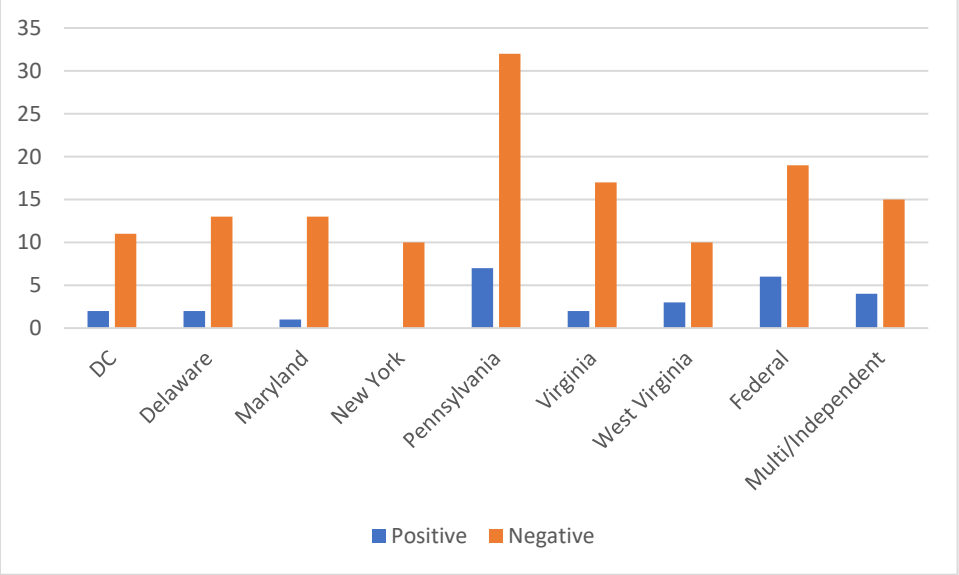


Figure 70: Positive vs. Negative Statements about the Models or the Modeling Process by Jurisdiction (# Mentions). Pennsylvania has the highest number of Negative and Positive Mentions about the Model/Modeling Process. The Federal Jurisdiction has the highest ratio of positive to negative comments. However, the majority of comments were negative across all Respondent Jurisdictions.

8.1.2 Detailed Statements about the Model/Modeling Process

Looking more closely at detailed statements about the model itself, three statements simply indicated that the model needs improvement without specifying how it could improve (see Figure 71). However, 7 blocks of text indicated that the model needs to provide estimates of co-benefits, possibly by incorporating local water quality modeling, so that users could consider benefits as well as costs when planning their WIPs. 4 other blocks of text described other desired modifications, including adding details that could increase accuracy and matching model scale to local-level planning processes. Note

that we do not include complaints about BMPs not accounted for in the model or concerns over whether or not the model is accurately accounting for BMPs in this figure, as these were already covered in previous sections, and we do not want to double-count those statements. However, from the three main processes, 33 blocks of text indicated that BMPs were under-credited in CAST or that BMPs needed to be added to CAST, while only 2 indicated that credits in cast were about right and 4 indicated that BMPs were over-credited (load reductions in CAST were higher than on the ground) or that inappropriate BMPs were included in CAST (for greenwashing).

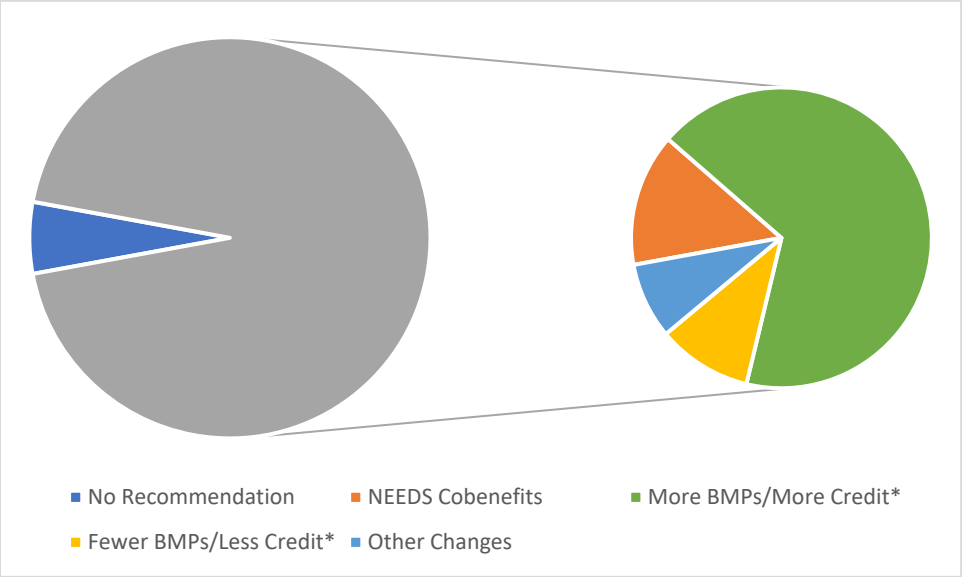


*Figure 71: Proposed "Improvements" to Models (# Mentions; * Indicates summary of information from previous sections). Most responses that criticized the Model also provided suggestions for changes that respondents deemed improving. Of these, adding more BMPs or giving existing BMPs more credit for load reductions was the most common. Many fewer responses indicated that CAST should include fewer BMPs or reduce credits for loading for some BMPs. Both of these categories are a reiteration from previous sections. Suggestions not covered elsewhere were less numerous. A number of responses did indicate that adding Co-benefits to CAST would be useful.*

Comments about the modeling process indicated respondent concerns about politicization, transparency, legitimacy, equity, and bureaucratic inertia. As noted above, 8 blocks of text indicated that modeling decisions could be used politically to affect the allocation of loading goals among states. This was associated in some blocks with credit inflation, or overvaluing the loading reduction provided by some BMPs, and was related to concerns about using the model for both planning and accounting. On the other hand, some respondents saw these political factors as beneficial. As reported in the section on the TMDL/Allocation, 9 blocks mentioned that state-level jurisdictions played a watchdog role on the model in order to ensure fairness in the process of load allocation among states. This is the other side of the politicization coin.

We also see interesting aspects of politics in the breakdown of actors with influence over the process, which heavily favors state-level jurisdictions, as shown in Figure 72. Of course, state-level influence was implicit in the main processes, but it was not mentioned explicitly. Furthermore, within the Federal grouping, 5 of 7 blocks indicated that influence was wielded through the Water Quality Goal Implementation team or related sub-committees with both state and federal representation. Although these committees often include “at large” members who do not represent government entities, only official state and federal representatives have voting or veto power. The same is true in higher-level

committees like the Management Board and the Principals' Staff Committee, which direct the GIT's activities and have final say over modeling decisions. Experts were mentioned several times here as well, as members of model-related committees and as participants on Expert Panels, which often make the most detailed decisions about things like how to assign load reduction to specific types of BMPs. Here, too, experts representing state interests are key players and the importance of expertise was highlighted in 6 blocks of text which indicated that better technical expertise could increase influence over the modeling process, whether for states or other actors. However, it appears that access is still limited; local government influence was only mentioned in one block of text, and no other stakeholders were mentioned as having influence over the modeling process at all.

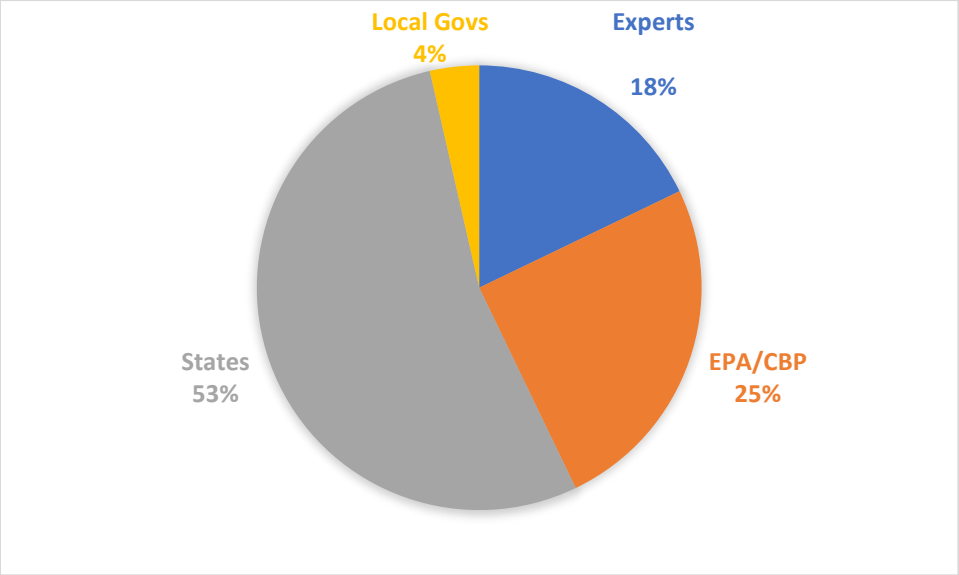


Figure 72: Actors with Influence over the Model Design Process (% Mentions). States were mentioned in more of half of the blocks of text describing actors with influence on the Modeling Process. The EPA/CBP was mentioned in a quarter of these blocks, followed closely by Experts. Local Governments were only mentioned in a few blocks. Other stakeholder groups were not mentioned at all.

Another set of annotation code groups addresses the rate of change in model development (see Figure 73). 11 blocks of text indicate that changes to the model such as adding new BMPs or shifting to a finer-scale land use model take many years to get through the bureaucratic process. On the other hand, 5 blocks of text indicated that the process was slow for changes proposed by states but faster and easier for changes proffered by the CBPO. Ironically, another 3 blocks of text indicated that the model changes much too quickly, at least when compared to implementation on the ground. This is related to some of the system-level problems with moving goal posts and political backlash that will be discussed in Section 9.3 Future Needs and Expectations. Another four blocks of text indicated that there was a process mismatch between the CBPO, state-level data collection/reporting, and other federal-level processes such as the NEIEN system. This was largely described as a hurdle that had already been overcome but which could create new difficulties if changes were made to the modeling process.

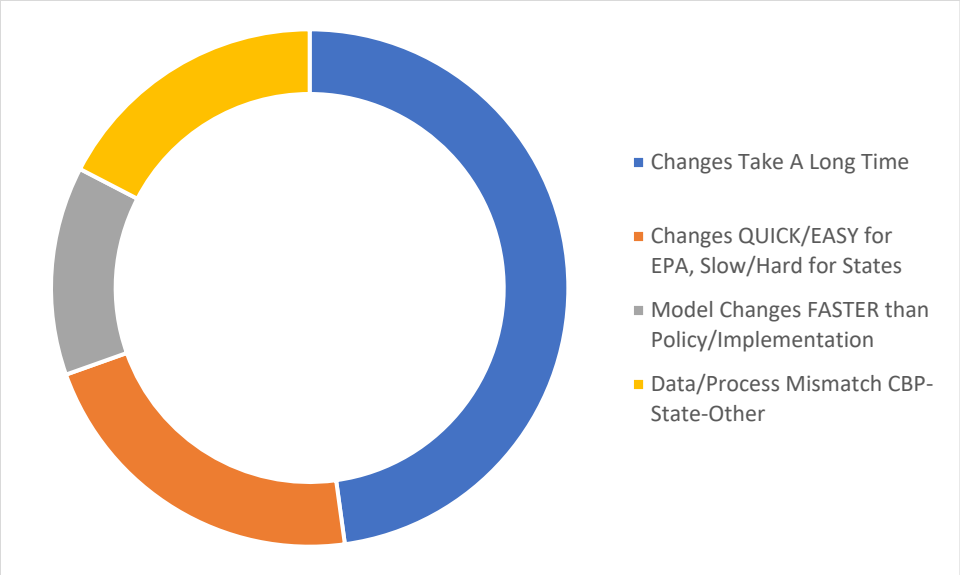


Figure 73: Statements regarding timing in the modeling process (# Mentions). Most statements indicated that changes to the model take a long time but some indicated that changes are easy for the CBP/EPA but difficult or slow when preferred by States. A small group of comments indicated that the model changes too quickly, particularly relative to on-the-ground implementation. Another concern was temporal mismatches in the data collection/reporting processes and modeling processes across different agencies.

8.1.3 Transparency and Legitimacy in the Modeling Process

The next set of variables in this section deal with transparency and legitimacy (see Figure 74). 15 blocks of text indicated that the modeling process was not transparent and 5 indicated that transparency was increasing. No blocks indicated that the process was open, clear, or inclusive. As noted above, 7 blocks indicated that it was largely top-down. Poor communication and performative engagement were also mentioned. More importantly, 28 blocks of text indicated that the modeling process was viewed as less legitimate, either because of this lack of transparency or because of concerns about modeling decisions/outcomes. Of these, 7 blocks indicated that stakeholders were less willing to accept the legitimacy of the model because outcomes did not coincide with their understanding of nutrient loading on the ground. Another 2 blocks of text indicated that stakeholders did not accept the legitimacy of the model without stipulating reasons for the statement. Only one block of text indicated that lack of transparency/unexpected outcomes did not undermine legitimacy, which coincides with statements in the section on the TMDL which suggest that more information about model decisions could reduce legitimacy if the information does not conform to stakeholder beliefs. However, 4 of the five blocks of text which indicated that legitimacy was increasing attributed the improvement to greater transparency in the modeling process.

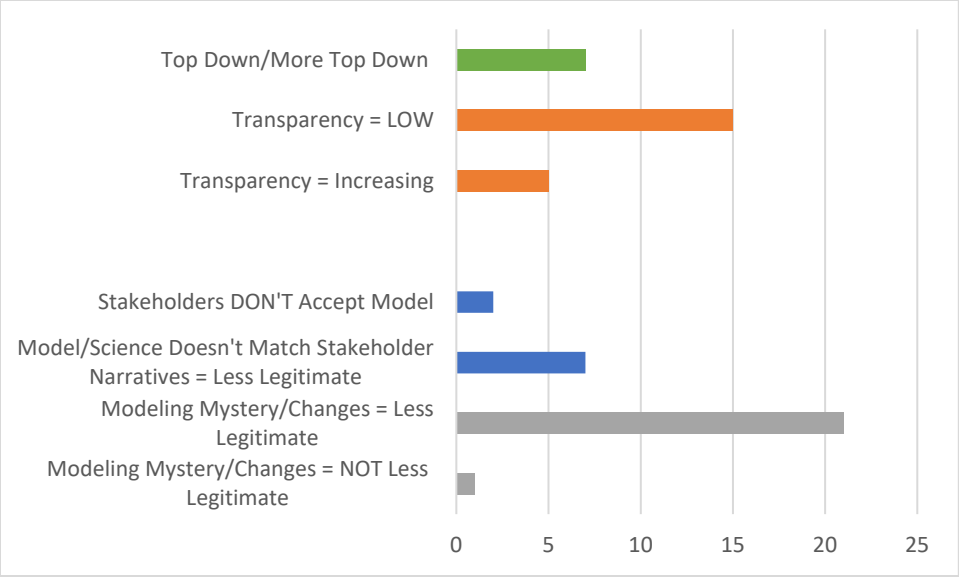


Figure 74: Statements about the Legitimacy and Transparency of the Modeling Process (# Mentions). The process was consistently described as Top Down with Low or Increasing Transparency. Most blocks describing legitimacy indicated that low transparency and/or frequent changes in the model reduced legitimacy but some did not. Some blocks also noted that stakeholders would not accept the model's legitimacy because its parameters or outputs did not match what they observed on the ground.

As a last note in this section, we also noticed considerable confusion about the purpose and use of the model in many of the interviews. Many respondents seemed to believe that the model could or should provide real-time tracking of nutrient levels and that this would be more persuasive to stakeholders in particular. We grappled with this issue ourselves and finally realized that the model is fundamentally constrained by the goals set in the CBP agreement, which means that it needs to predict changes in average maximum daily load in 2025 (the year when the TMDL is expected to be reached), rather than track changes in loading now. This is further complicated by legacy loading, which could ensure that changes in loading will not be observable for years or decades. In all, we feal this reinforces the major finding of this section, which is that **increased transparency could help to increase the legitimacy of the modeling process, but only if the CBP can also address the Political and Partnership Challenges described in Section 4.1 Challenges and otherwise improve Design Effectiveness (see Section 3.2 Institutional Design).**

8.2 Funding Process

Some of the results on the funding process parallel other sections. Co-benefits, feasibility, and effectiveness are all mentioned frequently as reasons to apply for funding, though the availability of funding or matching resources was mentioned more often. However, concerns about equity for local-level actors were mentioned much more often than for any other process. This is another source of path dependence, in that respondents frequently mentioned that communities or groups that already have resources were more likely to be successful when applying for additional resources. In part this was due to the availability of human resources, particularly time, expertise, and connections but also availability of matching funds. Several blocks of text mentioned that innovative and entrepreneurial communities could overcome regressive impacts of matching requirements through in-kind contributions. It is also interesting if not surprising that local, state, and federal influence over the funding process was commensurate with the amount of funds that each level provides (at least in terms of number of mentions).

Although there is a question about the funding process in the Implementation section of the interview instrument, due to time constraints we usually only asked this question of respondents who said that they had expertise on funding mechanisms (n = 9). However, when other respondents mentioned funding issues we coded those statements as well, so there are a mix of perceptions, but our results are still most likely to reflect the perceptions of respondents who engage with the funding process on a regular basis. Of course, the "funding process" is not monolithic and, much like implementation, is really comprised of many separate processes that may or may not be interdependent. Because funding is so complicated but also so important for implementation, we are collecting additional data on the topic from other sources. Additional survey or interview research could be beneficial as well.

8.2.1 Key Factors in the Funding Process

Factors affecting the funding process could not be easily broken down into positive/negative or high/medium/low categories. Instead, we break them down into statements about why a respondent might seek funding for a particular type of BMP or related programs, what might make an entity more likely to receive funding for a program, and what might limit funding options or make it less likely that an entity would receive funding (see Figure 75). Reasons to seek funding for a given program included the availability of matching funds (state, federal, or private; m = 13), co-benefits or other incentives (e.g., cover crop or stormwater subsidies; m = 8), and the feasibility or effectiveness of the BMP/program (e.g., cost effectiveness, stakeholder buy-in/lack of pushback, credits provided in CAST, and low levels of red tape; m = 11). Interestingly, WIPs were not mentioned as guides for seeking funding. This may be because respondents take the role of WIPs for granted but it could also be a symptom of the disconnects between WIPs as a "paper process" and on the ground implementation.

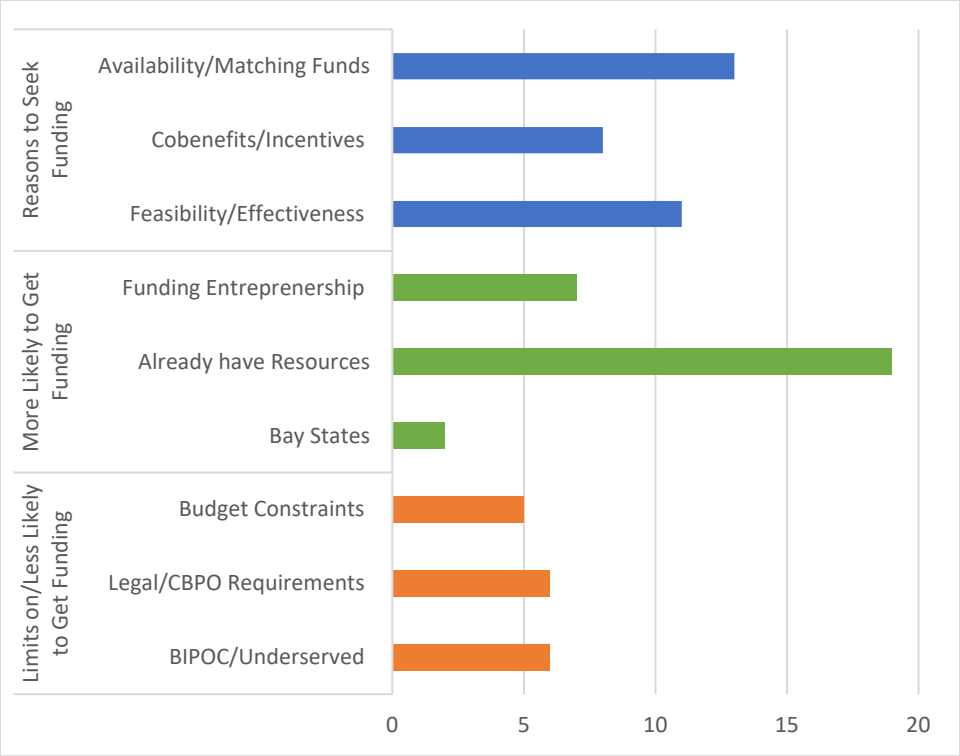


Figure 75: Factors that Affect the Funding Process (# Mentions). Reasons to seek funding for particular BMPs included the Availability of Funding/Matching Funds, Co-benefits or other Incentives for BMP adoption, and the Feasibility/Effectiveness of the BMP. Communities or organizations were more likely to receive funding if they were Entrepreneurial, Already have Resources (highest # Mentions), or were in Bay States. Limits on funding or factors that made it less likely that a group would receive funding included local, state, or federal Budget Constraints, Legal/CBPO Requirements, and membership in a BIPOC/Underserved community.

Another interesting finding here is that there was much more consideration of equity issues, particularly in terms of whether or not certain groups were more or less likely to get funding. 19 blocks of text indicated that entities (e.g., counties, conservation districts, planning district commissions, NGOs, etc.) that already have financial, organizational, and/or human resources were more likely to get funding for their projects while BIPOC/Underserved communities are still less likely to receive funding (m = 6). Of the latter, at least 3 blocks indicated that some entity was trying to increase funding to BIPOC and other under-represented groups. As in previous sections, concerns about the regressive nature of funding and particularly public subsidies for cost-share programs or taxes to pay for wastewater/stormwater upgrades were raised (m = 3), though this did not seem to impact funding per se. In four blocks of text, respondents also indicated that the choice to distribute funds equally across sub-state jurisdictions created inequities because it did not take into account the costs of reduction or the resources available to these local government agencies.

State-level equity came up again as well, with a few blocks of text mentioning that Bay States (VA, MD) are more likely to get funding than Non-Bay States. Funding entrepreneurship was seen to increase likelihood of receiving funding (m = 7) but could either be associated with entities that already have resources (especially time and expertise) to pursue funding options or those that lack resources and yet still find ways to attain funding, often through in-kind contributions to meet cost share requirements (more below).

Other limits on funding included budget constraints (m = 5), which were seen to increase competition among entities, programs, or legal constraints (m = 6). The latter include either CBPO requirements (e.g., BMP verification or grant requirements) or state and federal legislation (e.g., state hiring rules limit human resources or federal changes to cost-share provided through the Farm Bill). Of these blocks, two indicated that there are times when funds are available but cannot be spent due to said restrictions.

8.2.2 Influence and Options in the Funding Process

Legal and budgetary limits are only one source of influence over the funding process, but it is a very important one. Access to funds is important as well. These combined factors help to explain why state and federal actors are seen to have so much influence over the funding process (see Figure 76). EPA/CBP priorities were mentioned as having influence over the funding process in 30 blocks of text (including a few EPA-only programs). This was followed by other federal agencies (m = 21), including USDA programs like NRCS and EQIP, as well as the Federal Budget generally and the Farm Bill specifically. State budgets and agency priorities were mentioned as influential in 16 blocks of text. NGOs were also mentioned fairly often, mainly because some organizations like the National Wildlife Federation (NFWF) and the Upper Susquehanna Coalition (USC) can act as funding entrepreneurs, applying for funds from the government and then distributing them to various programs. Stakeholders, including farmers specifically, and local governments were least often mentioned as influential in the funding process (m < 5) which suggests that it is largely top-down, much like the implementation process more broadly.

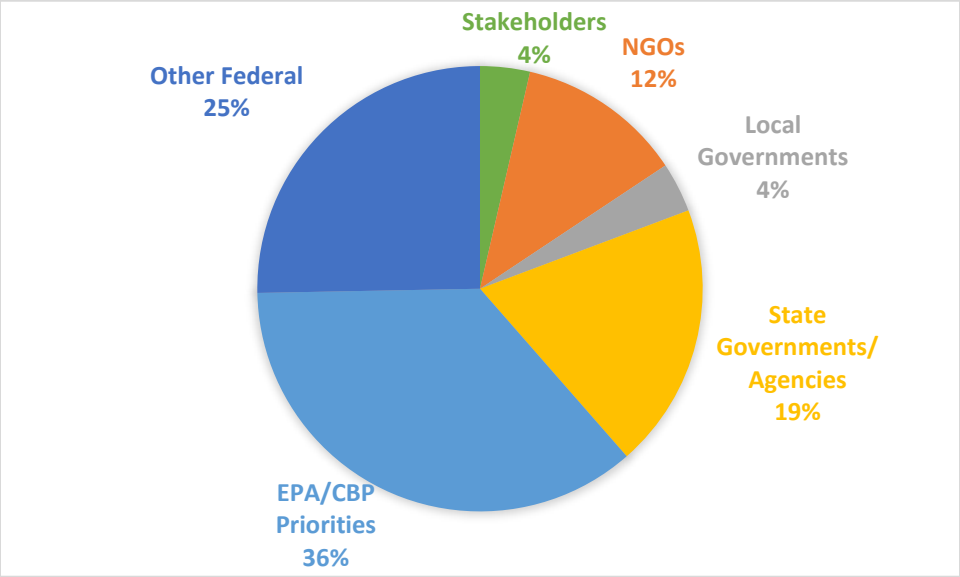


Figure 76: Entities with Influence on Funding Process (% Mention). Federal sources, whether through the EPA/CBP or other programs like USDA's EQIP were mentioned in more than 50% of blocks describing influence on funding. State Governments/Agencies and NGOs were also described as influential in more than 10% of the blocks, while Stakeholders and Local Governments were described as influential in less than 5% of the blocks.

Comparing mentions of influence with mentions of sources of funding, it is interesting that local-level funding sources were mentioned much more frequently (m = 11) than local-level influence (m = 3; see Figure 77). Otherwise, state and federal sources were mentioned in about the same proportions as in Figure 76. It will be interesting to look at actual budgetary data to determine if these perceptions are accurate.

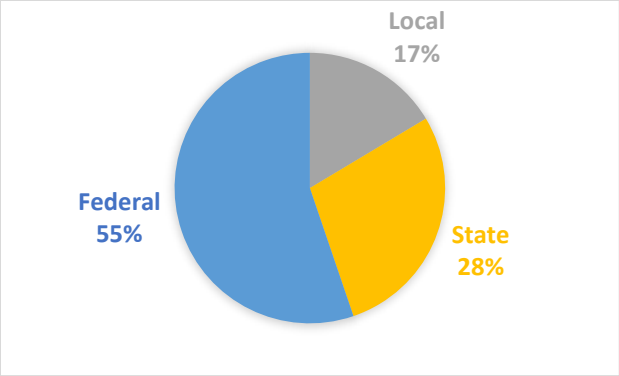


Figure 77: Sources of Funding by Level of Analysis (% Mentions). Federal sources were mentioned in over 50% of blocks, State funding sources in just over a quarter of blocks, and Local funding sources were mentioned in less than 20% of the blocks describing sources of funding.

Lastly, we had to create a whole new category of variables to capture the different types of funding or other resources described (see Figure 78). Taxes, fees, or lottery earnings were mentioned mainly in connection with funding for wastewater treatment, public green infrastructure, or other urban programs (m = 8). Grants were mentioned in 6 blocks of text, including descriptions of block grants and subsidized government loans. Cost-share was mentioned primarily for the agricultural sector. It increases funding by requiring that farmers or other stakeholders pay at least some of the costs of BMP implementation. On the one hand, cost share is thought to increase the stakeholder's stake in maintaining the BMPs, but it may also make BMP implementation more difficult, especially for long-term BMPs with high upfront costs (m = 2). Matching requirements were mentioned most often (m = 31). Similar to cost share, local governments, NGOs, or similar entities may be required to provide resources to match state or federal funds. This topic was discussed most often in the context of rural or small-town BMPs such as stream restoration, reforestation, or septic/sewer retrofits. It was generally seen as a factor that reinforced the inequalities described earlier in this section, though community ability to match funds with in-kind contributions (m = 4) was mentioned as a way that entrepreneurial communities could ameliorate the regressive impacts of matching requirements.

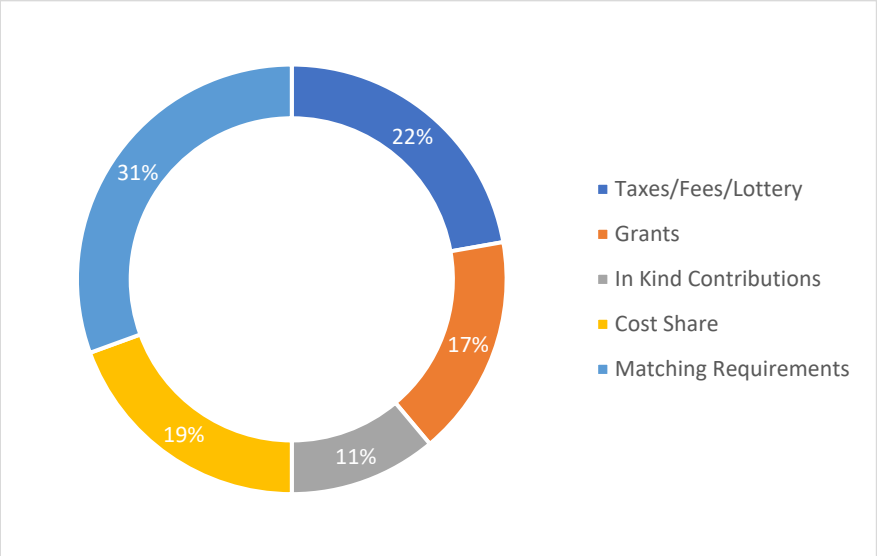


Figure 78: Types of Funding and Related Resources (% Mentions). Taxes/Fees/Lottery revenues and Grants (including subsidized loans) were the main types of government funding described. Respondents also mentioned Cost Share Programs (stakeholders

pay part of the cost) frequently, as well as Matching Requirements (local governments match funds provided), which were often associated with In Kind Contributions (matching with labor or equipment rather than financial resources).

8.3 Climate Change, Conowingo, etc.

Respondents described the process of adjusting watershed governance to accommodate Climate Change more often than any of the other processes covered here. A slight majority of responses indicated that Climate Change could be a threat to the Partnership, and a number mentioned that there may need to be changes to the WIP Design Process or CBP Goals (i.e., the TMDL) to accommodate climate change's effects on the watershed (see also Section 4.1.2 Problem Multipliers). Responses were more divided on the threat posed by the process of accommodating increasing nutrient discharges from Conowingo Dam, with specific comments focused more on intra-state equity issues and concerns about the viability of the Conowingo WIP. Mentions of Lawsuits were less frequent, with some perceiving them as windows of opportunity to improve watershed governance and others describing them impediments to cooperation. Nutrient trading was mentioned as occurring within some states but was not considered a viable option between states.

There were more blocks of text about the process of dealing with climate change within the CBP (m = 41) compared to Conowingo (m = 17). This is not surprising given that we prompted respondents to discuss climate change in the last section of the interview instrument but did not prompt them to think about Conowingo. Because of this difference, the relative numbers of mentions below should be interpreted cautiously. As noted in Section 4 Challenges and Opportunities, both climate change and the Conowingo Dam were identified as future challenges for the CBP, and Climate Change was also viewed as an opportunity in some blocks of text. Additional comments about these topics were coded as shown in Figure 79. We note that the reasons mentioned for describing climate change as an opportunity (increases in available funding and important co-benefits) are not pertinent to the Conowingo Dam issue, which may explain why there were no mentions of Conowingo as an opportunity.

Both of these topics were viewed as a challenge for the CBP because they will increase the amount of nutrients flowing to the Bay, and therefore states will have to further reduce their loading goals to reach the TMDL. In a few blocks of text, both were also viewed as politically difficult, largely for partisan reasons. Respondents varied in the impact they expected these challenges to have on the Partnership. In 6 blocks of text, respondents indicated that the increased stress created by having to reduce loading goals further would pose a threat to the partnership. However, in 5 blocks respondents indicated that this pressure would not threaten the partnership and, in a few blocks, suggested that the partnership would instead be strengthened. 3 blocks of text were more circumspect, indicating only that these topics would be difficult politically. There were more blocks of text discussing climate change in each category but, as shown in Figure 79, the distribution of blocks was similar for both challenges.

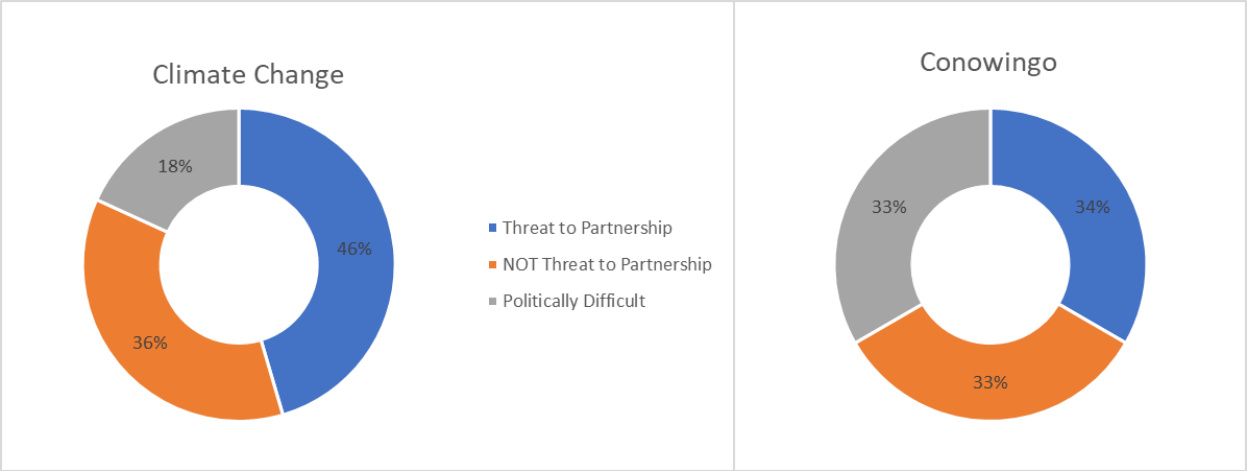


*Figure 79: Statements about Climate Change and the Conowingo Dam (# Mentions; * Indicates information already covered in previous section). Both issues were viewed as Politically Difficult. Climate change was mentioned more often as a Threat to the Partnership, but only marginally more often than NOT Threat to Partnership. Responses on Threat to Partnership are more evenly divided for Conowingo Dam.*

Otherwise, statements about each of these challenges tended to be quite different (see Figure 80). For Conowingo, most statements either expressed concerns about the separate, multi-jurisdiction WIP approach that was negotiated for the Dam in 2017[7] or indicated that the respondent's state would not be willing to reduce their own loading levels to accommodate the large release of nutrients that is expected as the Dam's capacity is exceeded. Concerns about the Conowingo WIP included perceived lack of transparency in the process and the use of the WIP to transfer resources from Bay to Non-Bay states, which is problematic from the viewpoint of Bay States. Only one block of text indicated that further changes to the WIP Design process might be needed to accommodate the Conowingo load, largely in the form of expanded BMP options to make the new WIP more palatable—and potentially more achievable.

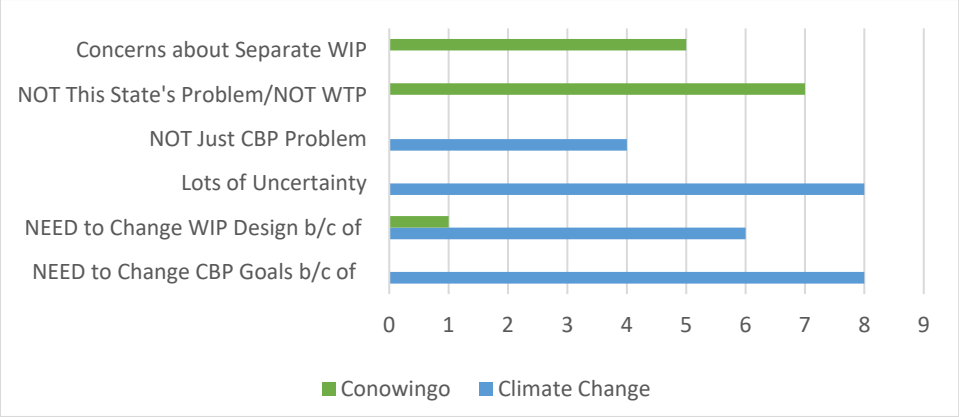


Figure 80: Statements about Conowingo Dam and Climate Change (# Mentions). Conowingo Dam is a more divisive issue, with a number of statements expressing concerns about the viability of the Conowingo WIP and more expressing the belief that the respondent's jurisdiction should not have to pay the cost of reducing nutrient loading from the Dam. Climate change was described as a source of uncertainty, a problem that was beyond the scope of the CBP, and potentially a reason to change either WIP Design or CBP goals.

[7] The multi-jurisdiction Conowingo WIP was published in 2022, after all of our interviews concluded.

In contrast, statements about climate change suggested that more sweeping changes might be needed in both the WIP Design/Allocation process (m = 6) and in overarching CBP goals (m = 8). The former was more focused on increasing BMPs like wetland restoration and green infrastructure to help with climate mitigation and adaptation as well as nutrient reductions. Building climate change into the WIPs by adding an additional cushion of load reduction as West Virginia did in their Phase III WIP was also mentioned. Changing the TMDL itself mainly referred to reducing the loading goal for the Bay in order to increase its resilience to climate change or, in some blocks, changing the TMDL based on model results that incorporate climate change projections (again, likely further reductions required). Like all aspects of the modeling process, incorporating climate change into the models is a political, as well as a technical, exercise which should be examined in greater detail, especially since we need to create our own projections using the coupled model that is the main focus of this project. Respondents also noted that there is a lot of uncertainty around climate change (m = 8), including political risks, and that it is not a problem that can be addressed solely within the CBP (m = 4).

As noted in the introduction to this section, two other sub-processes were mentioned by respondents during our interviews. Lawsuits (m = 14) and nutrient trading (m = 9) are both too small to analyze in detail. We note that respondents paid greater attention to pressuring lawsuits (m = 11) than challenging lawsuits (m = 3), which makes sense since challenges to the TMDL were settled soon after it was established in 2010 ,while lawsuits designed to pressure Pennsylvania to increase their load reduction (directly or through the EPA) are more recent. Respondents were also divided on their perception of the efficacy of pressuring lawsuits, with 4 blocks of text indicating that the suits did/will improve CBP effectiveness and 6 blocks indicating that they will not. Factors affecting lawsuits include state support, the availability of scientific evidence, and important deadlines or **political windows of opportunity** during which the lawsuit might have an effect on public opinion or politician's incentives regardless of the outcome of the case. Stakeholder buy-in was also mentioned as a factor that could prevent challenging lawsuits or at least prevent some stakeholder groups from joining in. Additional data collection and analysis are needed to determine if and how these lawsuits will be included in our models or future scenario development.

Nutrient trading was mentioned both within states (m = 9) and among states (m = 7). Within states, most mentions simply indicated whether or not a state had a nutrient trading program (m = 4) or had considered one (m = 2). One block indicated that the state program was effective while 3 blocks indicated that state programs were not effective. Trading between states has not occurred yet but was described as infeasible because transfers between states would be politically unpopular (m = 2). It was also seen to be unpopular with NGOs because of the potential for pollution havens or nutrient laundering among states/outside of the watershed (m = 6). From this, it seems that we will need to account for nutrient trading within states in our model design, but nutrient trading between states does not need to be included in our scenario design.

9 System Process

Descriptions of the governance system as a whole (or System Process) suggest that respondents viewed it differently than they did its constituent processes.

Some of the statements made in our interviews described the CBP governance process as a whole system instead of its constituent processes (m = 225). While some of these blocks of text reiterate

observations about specific processes, others place greater weight on factors mentioned in previous sections or describe attributes that are not mentioned elsewhere. These results are not definitive, but they do indicate that participants may view the entire system differently than they view its constituent parts. Of course, we cannot rule out the possibility that respondents are equating the whole system with some constituent process. Nevertheless, because emergent properties are an important element of complex systems, and our project is designed to improve understanding of the complexity in social-ecological systems, these results can point the way to important areas for further research.

Methods Note: Because we did not ask about system-level attributes aside from effectiveness, challenges, and opportunities, we cannot determine if these differences reflect emergent properties (the system being more than the sum of its parts) or the views of respondents who think about the system as a whole more than its constituent processes. The latter may be due to either extensive knowledge about the whole system or—paradoxically—lack of specific knowledge about the processes described in previous sections. Data available suggests that there is no strong bias toward system-level statements by any group of respondents (see Figure 81). In particular, there is no binomial distribution on respondent characteristics such as years of experience or depth of involvement in the CBP, which might be expected if both extensive knowledge and lack of knowledge played a role. If we break down the depth of participation variable by type of committee (i.e., advisory, administrative, core technical, and workgroup/peripheral technical), system-wide statements vary between 10-14% for all groups, except low-level participation in core technical (6.6% of blocks describe the system).

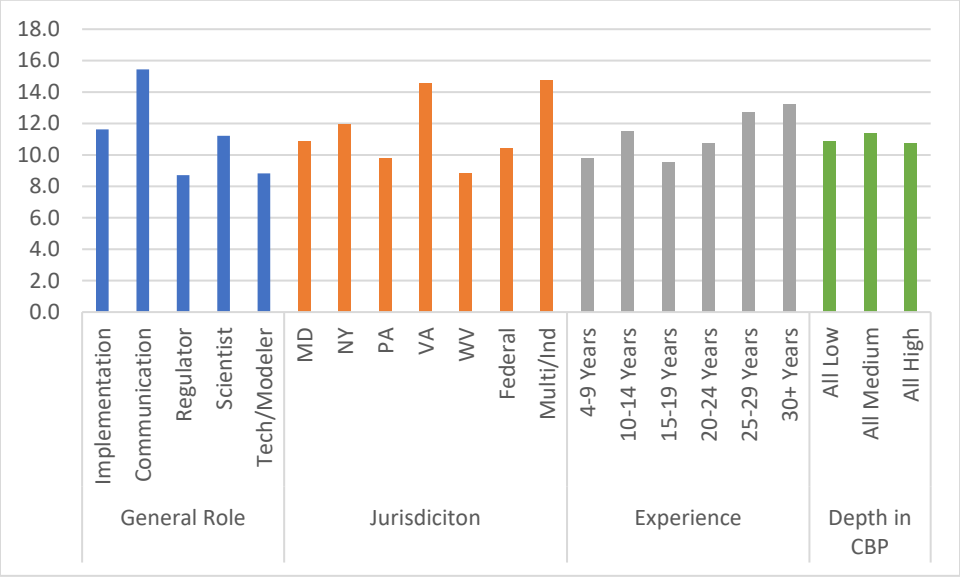


Figure 81: System Annotations as percent of Total Mentions per Respondent Group. There is no observable bias in the number of system-level annotations by respondent characteristics.

9.1 Factors and Actors

Overall, the System Process was viewed as top-down, with only performative engagement and low transparency. There were also some areas of contradiction or difference from results described in other sections. For instance, responses indicated that Co-benefits are NOT considered in System Processes, even though Co-benefits were considered important criteria for BMP selection, implementation, and funding when respondents were asked about those processes specifically. Similarly, when respondents talked about the system being Too Complex/NOT Easy to Navigate, they described this as a factor that affected the whole System Process, rather than a factor that was specific to any sub-process. The System Process was described as using Performative Engagement more often than any other Process and it was also described as providing Less Funding/Access to BIPOC on par with the Implementation/Funding Processes. Actors described as having influence in the System Process are similar to those mentioned for other Processes, with the large exception of the Media, which was only mentioned for the System Process. Local Governments, Stakeholders, Expert/Technocrats, Other States, and Other Federal actors are also not mentioned at the System Level.

Respondents only mentioned three factors affecting the System Process, so we combine these with descriptions of the System Process and compare both categories of variables across the processes described in this report (see Figure 82). Note that these are only factors and descriptors that occur for the system process; many others were described for sub-processes in previous sections. While the politics of allocation among states was seen as detrimental primarily in the TMDL Process (m = 16), it was also seen to undermine the whole system in a few blocks of text (m = 2). Although a number of blocks of text indicated that Co-benefits were an important consideration in the WIP Design/Allocation (m = 9) and Implementation/Funding (m = 6) processes, respondents indicated that Co-benefits were important but NOT considered sufficiently by the system as a whole (m = 12). 13 blocks of text indicated that the system was too complex and difficult to navigate. Although related concerns, such as bureaucratic costs and difficulties changing the models were identified for other processes, this particular factor was only mentioned for the whole system. It supports findings from our analysis of Design Effectiveness and Partnership Challenges, which also indicate that the system is complicated and difficult to navigate (see Sections 3.2 Institutional Design and 4.1.3 Partnership Challenges).

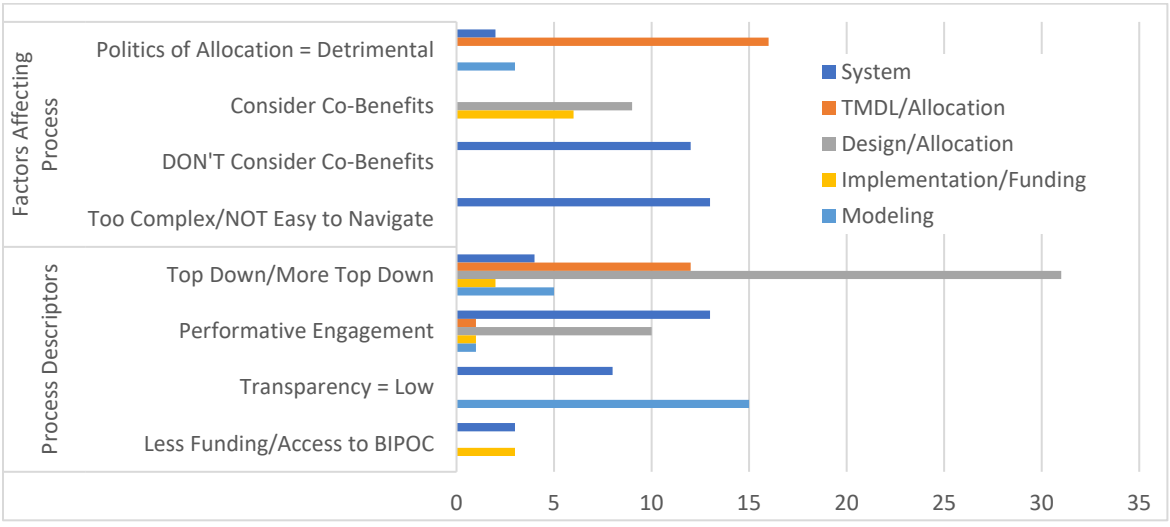


Figure 82: Comparing Factors and Descriptors across Types of Processes (# Mentions; only includes codes that apply to the System Process). While some responses indicated that the Politics of Allocation were Detrimental to System Processes, many

more described the Politics of Allocation as Detrimental to the TMDL/Allocation Process. In contrast, responses indicated that Co-benefits were considered in the WIP Design/Allocation and Implementation/Funding Processes but were NOT considered in System Processes. Similarly, when respondents talked about the system being Too Complex/NOT Easy to Navigate, they described this as a fact that affected the whole System Process, rather than a factor that was specific to any sub-process. There were no mentions that the System was a bottom-up process, but it was described as Top Down/More Top Down, along with other sub-processes. The System Process was described as using Performative Engagement more often than any other Process, it shared the Transparency = Low description only with the Modeling Process, and it was also described as providing Less Funding/Access to BIPOC on par with the Implementation/Funding Processes.

Descriptions of the whole system focused on hierarchy, transparency, and equity. The decision process was described as "Top Down/More Top Down" for the entire system in 4 blocks of text. This is less than for other processes like TMDL/Allocation and WIP Design, but, like Modeling and Funding, there are no blocks of text indicating that the system is NOT run from the top-down (where the states and federal government agencies are the "top"). Similarly, there is no mention of engagement for the entire system except for performative engagement, where advice is solicited but does not affect outcomes (m = 13). This is higher than for any of the sub-processes and only approached by the 10 mentions of performative engagement in the WIP Design Process. Given these perceptions, it is not surprising that the whole system was also described as having LOW transparency (m = 8), though this came up less than for the modeling process (m = 15). Finally, there were 5 separate mentions of equity issues for BIPOC/Under-represented groups, which were only mentioned in 3 other blocks associated with implementation/funding.

Turning to the actors who are seen to influence the CBP as a system, most groups are similar to those mentioned for other processes (see Figure 83). Farmers (m = 3) and NGOs (m = 8) are both mentioned, as are state governments/agencies (m = 9), and the CBP/EPA (m = 7). It is interesting that experts/technocrats, local governments, and other state or federal governments were not described as influential for the entire system, but it is not clear if this is because respondents really don't think these groups have influence or if these groups just did not come up during the interview. On the other hand, 11 blocks of text indicate that the media is an influential group at the system level, whereas it is not mentioned for any of the other process sections. Some of these media blocks co-occur with political will as a challenge (m = 3), political will as an opportunity (m = 3), accountability as a challenge (m =3) and co-benefits as an opportunity (m = 4).

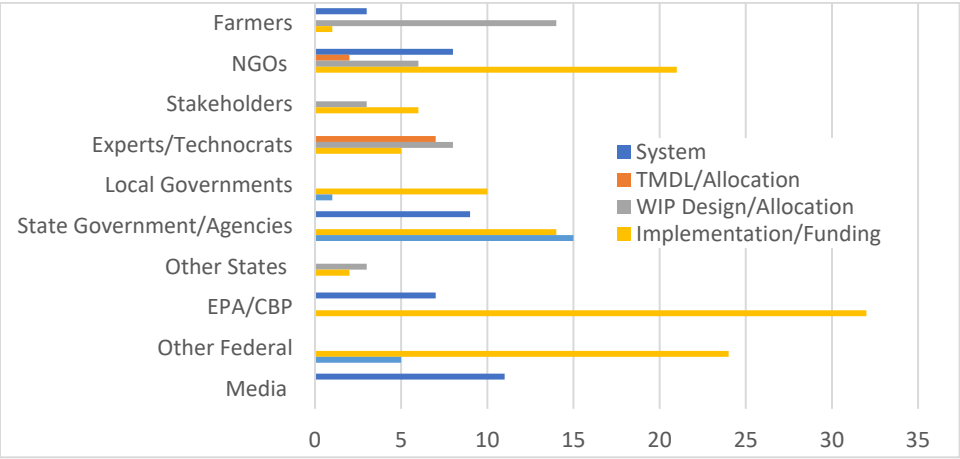


Figure 83: Comparison of Actors who Influence Different Processes (# Mentions). Actors who influence the System Process are similar to those mentioned for other Processes, with the large exception of the Media, which was only mentioned for the System

Process. Local Governments, Stakeholders, Expert/Technocrats, Other States, and Other Federal actors are also not mentioned at the System Level.

9.2. Historical Assessments

Responses describing the trajectory of the Chesapeake governance system are divided in their assessment of the trajectory of the CBP, with about 1/3 of the statements on this topic indicating that the System Process has improved since the institution of the CBP and 2/3s indicating that it has not improved or, at least, not improved enough. These specific statements echo results in Section 3 Overall Effectiveness.

Before describing respondents' views of the future of the CPB system, we also note that there were 28 blocks of text describing the historical development of the system as a whole. Much of this information was largely descriptive, but some blocks evaluated the trajectory of the system as it has evolved. Creation of the TMDL was largely attributed to states' failure to curtail loading through voluntary measures since the Chesapeake 2000 agreement. Loss of ecosystem services and some extreme events like the Pfiesteria outbreak of 1997 were listed as causes of political pressure to reduce loading, but so, too, were lawsuits against the government (state/federal) and the 2009 Executive Order from Barak Obama. The TMDL-system evolved quickly in the first few years, both in terms of new policies (the Phased system, Milestone process, E3 Scenarios, etc.) and technologies (Bay Models, scenario tools). High-level political support from state governors and facilitation by the CBC were also mentioned as important for overcoming the hurdles of creating the CBP (m = 5).

The number of blocks is small, but approximately evenly split between those who evaluate these historical trends as positive (improving effectiveness; m = 6) or negative (NOT improving effectiveness m = 10). On the positive side, 3 blocks indicated that the format of the CBP was unique and, indeed, revolutionary. These blocks tend to highlight the Milestone Assessments, Backstop Measures, and other novel approaches to accountability built into the TMDL system. A few also noted the importance of institutionalizing rules about load distribution among states (the E3 Principles, Hockey Stick, etc.) in order to minimize conflict among state-level jurisdictions. The remaining positive blocks are more circumspect, indicating that the CBP is much more effective than pre-2010 approaches to governing nutrients in the Chesapeake Bay.

On the negative side, 5 blocks indicated that the CBP did not lead to major changes in how the system worked or, at least, the changes were not as revolutionary as expected. Failure to really reduce loading from agriculture was mentioned in a few blocks and some respondents noted a decline in high-level engagement with the CBP (i.e., leadership from governors), which undercut state-level commitments to reductions. This was linked to reduced political will (esp. public pressure) and increasing partisan conflict at federal and state levels. Pushing the can down the road—specifically the fact that most states did not meet their Phase I or Phase II loading goals and/or the failure of the EPA to enforce those goals—was also mentioned in several blocks of text. One block indicated that the CBP was indeed revolutionary, but the task of load reduction had become more difficult, thereby undermining its effectiveness. Whether positive or negative, these statements provide further support for the general finding that the CBP has been partially effective and that some systemic factors may need to change to increase effectiveness.

9.3 Future Needs and Expectations

Statements about the future of the CBP System largely expressed concern about the difficulties associated with additional load reductions. Many of the specific prescriptions provided to improve governance in the future echo the Opportunities described in Section 4.2 Opportunities, including increasing Political Will, which was mentioned most often. However, there is much more emphasis on changing goals to address concerns about the future. Responses were almost equally divided between making goals more stringent to better protect the environment and making goals more realistic so that they are easier to attain. Emphasis is also higher on increasing Real or Reflexive Engagement and improving the Political Savvy of groups advocating for improved water quality in the watershed. Only 2 blocks prescribed increasing diversity/inclusion to improve System Processes, which is further evidence that, while participants might value diversity and inclusion, most do not see it as central to the CBP's mission.

Many respondent concerns about the future were already covered in Section 3 Overall Effectiveness and Section 4 Challenges and Opportunities. However, respondents also volunteered a number of perspectives on the future of the CBP as well as describing expected or desired paths forward for the CBP post-2025. First, as shown in Figure 84, concern for the future was expressed in a large number of blocks, with 32 mentioning that they expected the future to be more difficult because load reductions would have to increase/reducing loads was becoming more difficult and 3 expressing uncertainty about the future of the CBP. Of those who saw the future as being more difficult, 10 blocks expressed concerns that these problems would threaten continued cooperation through the CBP. Only 5 expected that the future would be a continuation of current trends, which is also not uniformly positive given the widespread perception that the CBP has only been partially effective to date. The positivity bias that we observed in Section 6 WIP Design and Sub-State Allocation Processes is clearly not present when respondents consider the future of the system as a whole.

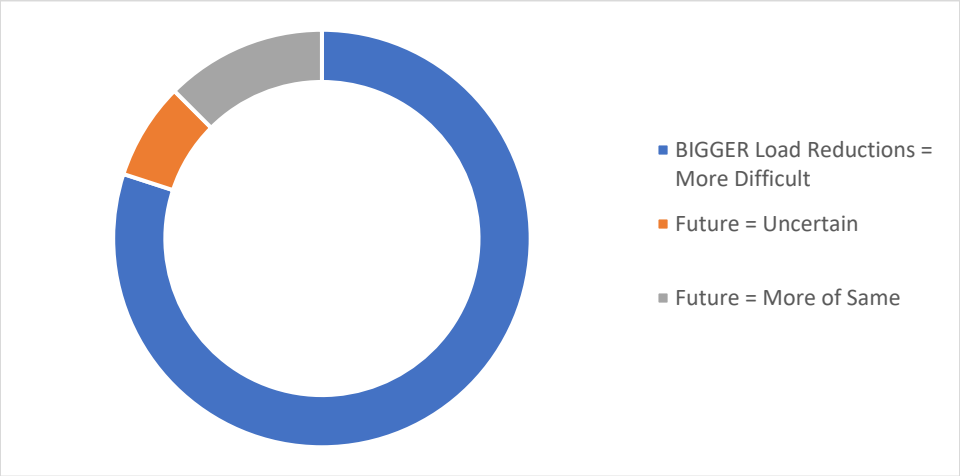


Figure 84: Future Expectations for the CBP System (# Mentions). The vast majority of statements about the future indicated that water quality governance will be more difficult because larger load reductions will be required. A smaller percentage indicated that the future was either uncertain or that it would just be "more of the same".

In addition to identifying challenges and opportunities, some blocks of text describe more specific prescriptions for improvements in the effectiveness of the CBP system. As shown in Figure 85, reducing the amount of agriculture (m = 7) or development (m = 3) in the watershed was mentioned as necessary for meeting loading goals in some blocks of text. Interestingly, 2 of the blocks advocating less

development specifically called for preservation of agricultural land. On the other hand, many more blocks of text indicated factors that need to increase to improve effectiveness. Of these, more political will (also less politicking) was mentioned most often (m = 38), which again supports findings in Section 4 Challenges and Opportunities. In fact, increasing political will co-occurred with the opportunities of Co-benefits (cc = 9) and Engagement/Outreach (cc = 7), as well as challenges Agriculture (cc = 6), Behavior Change (cc = 8), Climate Change (cc = 8), Human Resources (cc = 6), and, of course, Political Will (cc = 8). To this was added more Political Savvy (m = 6), particularly on the part of groups lobbying for more effective load reductions.

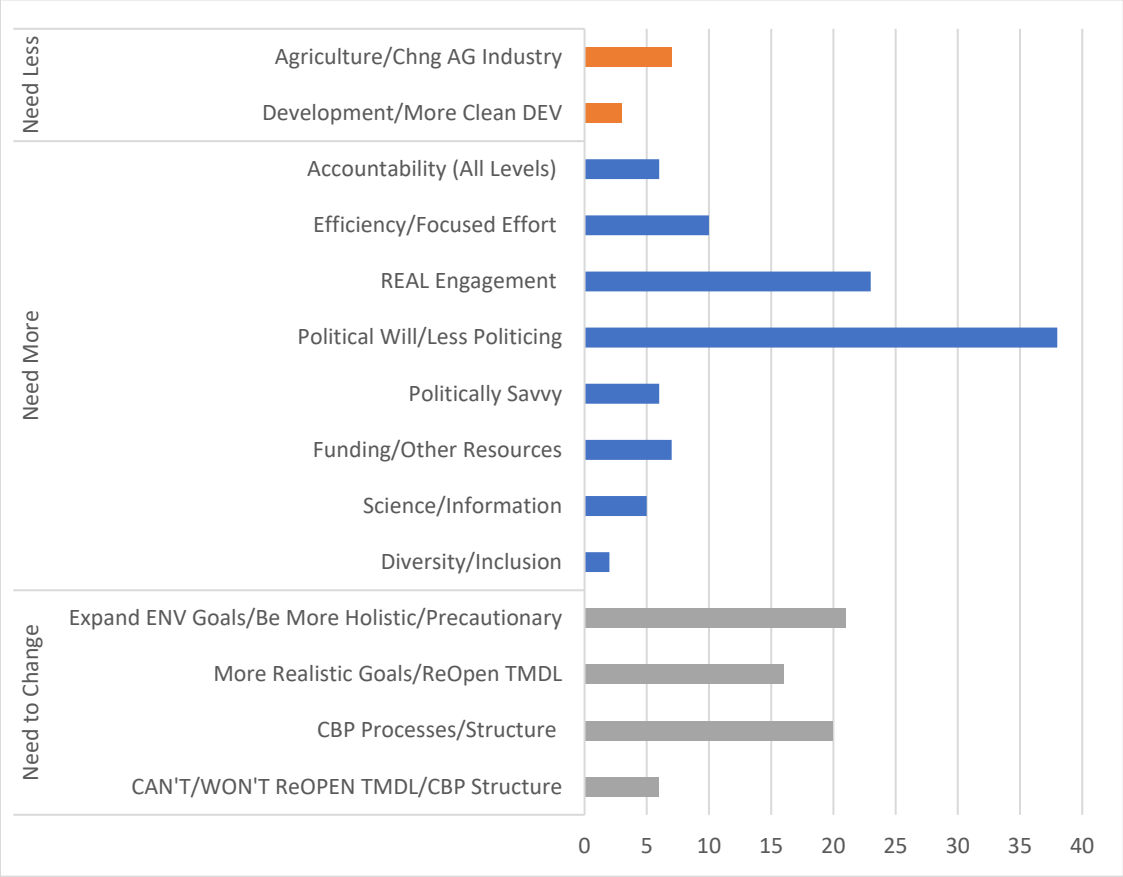


Figure 85: Prescriptions for Improving CBP Effectiveness + One Constraint (# Mentions). Many of these more specific prescriptions echo the Opportunities described in Section 4.2 Opportunities, including increasing Political Will, which was mentioned most often. However, there is much more emphasis on changing goals to address concerns about the future. Responses were almost equally divided between making goals more stringent to better protect the environment and making goals more realistic so that they are easier to attain. Emphasis is also higher on increasing Real or Reflexive Engagement and improving the Political Savvy of groups advocating for improved water quality in the watershed.

Increasing Real or Reflexive Engagement was recommended in 23 blocks of text and suggests that our numbers on performative engagement above might underestimate the issue. This was viewed as a way to increase political will but also to increase the implementation of voluntary BMPs by spreading knowledge about Co-benefits (cc = 7) and tackling the challenges of Behavior Change (cc = 6). Other overlaps occurred with the challenges of the Complexity of the CBP (cc = 6), Lack of Funding (cc = 6), and lack of Human Resources (cc = 9). Similarly, increasing Funding/Other resources (m = 7), Better Science/Information (m = 5), and Increasing Accountability (m = 7) were viewed as methods to increase implementation and thereby reduce loading. Focusing resources on areas where load reduction could be

achieved most Efficiently was also mentioned in 10 blocks of text, though we note that this would contradict concerns over equity described in sections on allocation and implementation. Only 2 blocks mentioned that Increasing Diversity/Inclusion could help improve system effectiveness, which is further evidence that, while participants might value diversity and inclusion, most do not see it as central to the CBP's mission.

The last set of "NEEDS" variables covers changes that are not strictly increasing/decreasing and also includes one constraint. The main divide in this group was between blocks of text which advocated for more Holistic and Precautionary Goals (i.e., reduce loads even further than current TMDL in order to protect ecosystems more effectively; m = 21) and more "Realistic" Goals (i.e., raising the TMDL so that states would not have to reduce their loads as much as they would under the current system; m = 16). Responses in the first category tended to focus on ecosystem function but were also linked to opportunities for Co-benefits (cc = 8). The latter was associated more with challenges, particularly Climate Change (cc = 5) and the Feasibility of the TMDL (cc = 5). 20 blocks of text also recommended some change to the CBP Process/Structure other than re-opening the TMDL. These were wide ranging but included: 1) making the system less complicated, 2) increasing flexibility, 3) increasing stability, 4) being more inclusive (of states), 5) re-evaluating the Equity Principles, and 6) changing CBP leadership. Some recommendations were contradictory, not least calls to Stop Changing Loading Targets (m = 2) and demands to shift from a Stationary TMDL to a Dynamic TMDL (m = 4).

Another 6 blocks of text indicated that major changes to the TMDL or the CBP structure were not legally or politically feasible. These statements also co-occurred more often with challenges than opportunities, but no related challenges stand out. The majority of the blocks originated from respondents in the Communication, General Role, and/or who have Expertise in Policy, Law, or Social Science. We will need to look more closely at what changes are or are not possible and why as we work on post-2025 scenarios. The responses presented here should provide a starting point for that analysis.

Appendix A: Detailed Methods

For a large governance system like the Chesapeake watershed, it can be difficult to recruit a representative sample, especially given our interest in so many different components of the policy process. As explained below, we focused on recruiting people who were directly involved with at least one of the main processes described in this report (TMDL/Allocation, WIP Design/Allocation, or Implementation). Thus, results are mainly representative of the views of decision makers at the state and federal levels with some local and non-governmental representation. Additional data would need to be collected to assess perceptions of stakeholders, local governments, and the general public. Data on Implementation and Funding could also be expended to better describe the heterogeneity of programs that constitute these processes. Although the system is large, communities of decision makers are still tight knit, so we were very careful to maintain confidentiality during all phases of research, including sampling, data analysis, and reporting.

A.1 Sampling

We targeted the population of decision makers who were directly involved in some component of the policy process in the Chesapeake watershed. This included people who worked at or with the CBP, worked on CBP programs for state-level agencies, and those who worked directly with stakeholders on BMP implementation. We also included some respondents who did not have any official role as decision makers but who were long-time observers of the process, usually in roles as stakeholders and/or watchdogs. Beyond these general outlines, we set a goal for at least 5 interviews from federal-level agencies (e.g., EPA, CBP, etc.), independent observers, and each of the 7 state-level jurisdictions. Within these groups we worked to get a sample that would provide sufficient information on each component of the policy process: 1) setting the TMDL for the watershed, 2) allocating loading goals among the state-level jurisdictions, and 3) designing and implementing the watershed implementations plans.

Respondents were recruited using stratified snowball sampling. We collected 122 recommendations for people to interview and sent out 91 invitations to participate. 37 personal requests to participate were sent by our project colleagues (batches 1-2). The remainder were sent by Dr. Webster (batches 3-4), who would mention the name of the person who recommended the potential participant if permission was expressly provided by the recommender. If no permission was expressed, Dr. Webster simply sent an invitation requesting participation without mentioning any other names. Through this process, we were able to recruit 59 participants, with a minimum of 5 and a maximum of 9 respondents from each "jurisdiction" as shown in Figure 86. As might be expected, the response rate was much higher for personal requests from our CNH colleagues (86.5%), but even the response to invitations from Dr. Webster, who was unknown to all participants prior to the interviews, was quite high (64.8%). There is considerable variation in the response rate by jurisdiction. We do not have sufficient data to explain this variance but certainly cannot rule out artifacts or biases in the recruitment process.

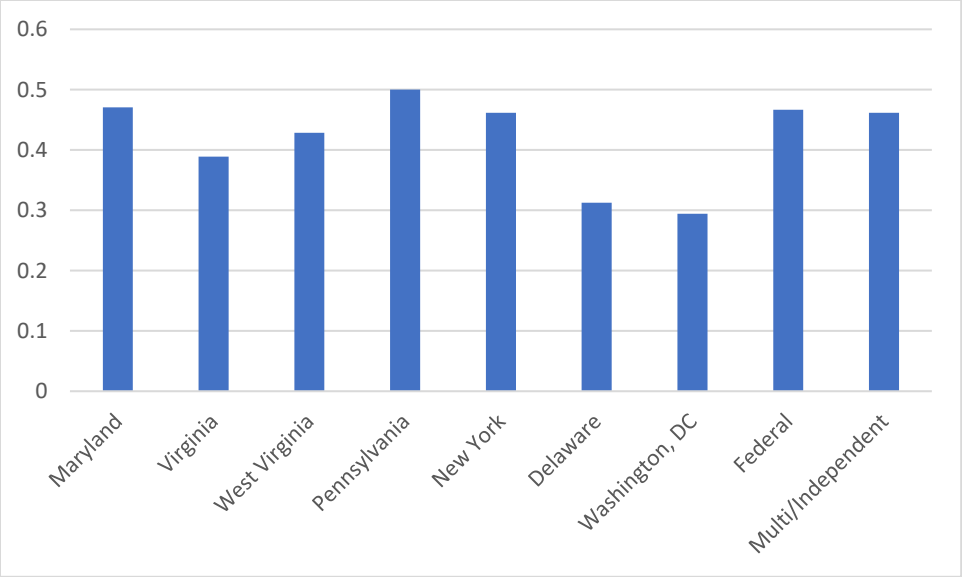
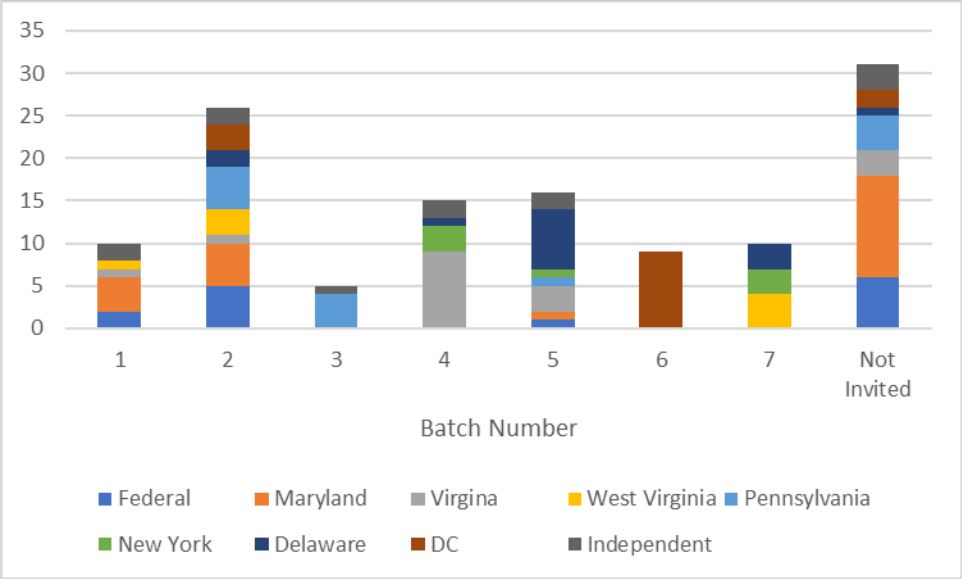


Figure 86 Percent of Invitations that were accepted by Jurisdiction (Number of Interviews in Parenthesis)

The "snowball" started with recommendations from our CNH2 colleagues with experience in the Chesapeake governance process. This produced a list of 12 recommended names, of which 10 were invited to participate in the study by our CNH colleagues on our behalf. At the end of each interview, we asked the respondent if they would be willing to recommend colleagues for our sample (see Appendix C: Annotated Instrument). This formed the basis for our second batch of invitations, which was also sent by our CNH colleagues in mid-July. After this, we chose to send out the invitations ourselves (from Dr. Webster's e-mail address), largely so that we could assess the difference in response rates but also so that we could reach out to respondents outside of our CNH colleague's close networks. Sent out in 4 batches from August to early December 2021, these invitations targeted respondents from groups that were under-represented in the first two batches, mainly based on jurisdictional affiliation and respondent expertise.

The results of this process can be seen most clearly in the distribution of interviews by jurisdiction in each batch (Figure 87) and in the changes in jurisdictional distribution from the number of individuals recommended by participants, the number we invited, and the number of interviews completed (Figure 88). In most cases, we would invite a cluster of individuals from 1-2 jurisdictions in each batch. For jurisdictions with relatively few recommended participants, all would be invited. Where we had larger numbers of recommendations, we tried to spread out our invitations by agency/expertise. Like any snowball sampling process, some recommended participants were not invited because we had already recruited enough respondents with their expertise from their jurisdiction. In a few cases, respondents were not invited because their areas of expertise appeared to be well outside the Chesapeake watershed or because they did not fit our rather broad definition of "decision makers". We also invited a larger proportion of recommended participants from jurisdictions with lower response rates to ensure our minimum of 5 respondents/jurisdiction for the sake of confidentiality (see below).



Batch	Dates Sent
1	6/25, 7/12
2	7/16
3	8/12-16
4	8/30-31
5	9/7-16
6	10/14-15
7	11/23, 12/6

Figure 87 Invitations by Batch and Jurisdiction

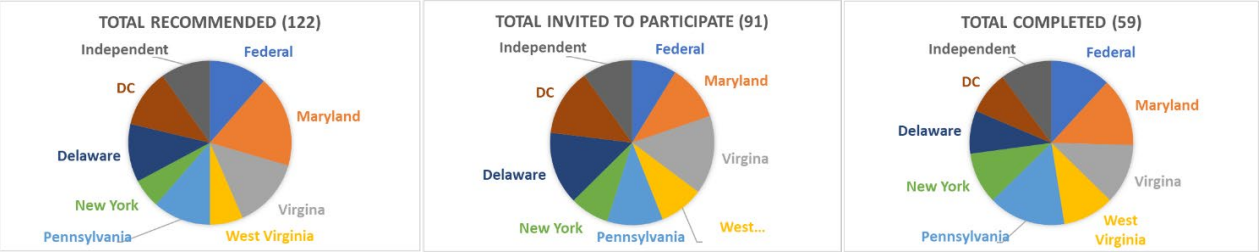


Figure 88 Share of Recommended, Invited, and Interviews Completed by Jurisdiction

While we paid particular attention to representativeness by jurisdiction and area of expertise, we did not include any over-sampling for demographic characteristics as we considered representation of the population of decision makers to be more germane to our research goals than overall diversity of representation. In other words, we accepted that our sample would likely diverge from US Census distributions on factors like education, income, race, and gender but expect that this primarily reflects attributes of the population studied rather than bias in the sampling method *per se*. Background provided by our respondents suggest that the sample is overwhelmingly college educated, with a large proportion holding advanced degrees (MA, PhD, JD). All were professionals working in environmental governance in some capacity. Just over half (n = 33) worked for government agencies at the local, state, or federal level, with the majority of these (n = 20) employed by a state-level environmental agency. The remainder of respondents (n = 26) were employed by consulting firms, universities, or other non-governmental organizations. Based on apparent rather than self-expressed gender, 34.4% of recommended participants, 39.5% of invited participants, and 42.4% of recruited participants who completed interviews were female. We did not attempt to evaluate other demographic variables, but Washington, DC stood out as being more racially and ethnically diverse than any of the other jurisdictions, based on participants recommended, invited, and recruited.

Although this research qualifies for IRB exemption, we still felt that it was important to maintain and protect confidentiality given that there could be potential negative repercussions if respondents could be identified directly or indirectly through our research process or publication of results. During the

sampling process, this entailed minimizing the sharing of identifying information on who was recruited for or participated in the study. Of course, there were several ways in which this information would be shared at least with a limited group. For instance, our CNH colleagues knew who they recommended/invited, though we did not tell them whether or not any of the people they invited chose to participate. Similarly, respondents would know who they recommended for the snowball sample but were not given access to information on whether we actually invited anyone they recommended or if anyone they recommended chose to participate. In addition, all participants were free to tell others whether or not they participated, and respondents might also gather that individuals who recommended them for participation also participated. Furthermore, given that our target population contains a number of small, close-knit groups, it is also likely that anyone who is familiar with the Chesapeake governance process would be able to make educated guesses about who participated in the study even without reading our results. Recognizing that we could not completely prevent people from guessing who was participating, we chose to increase our initial minimum of 3 respondents/jurisdiction to 5 respondents/jurisdiction, in hopes that larger groups would make it easier to obscure respondent identities when reporting results. This prolonged the recruitment period by 4 months. We also provided a 1-page disclosure statement to each potential respondent prior to the interview, as we would in any type of human subjects research (see Appendix B: Disclosure Statement).

A.2 Data Collection and Analysis

During the recruitment and interview period, data on each potential participant was collected from the internet, including title, affiliation, e-mail address, and professional biographical information. Similar information was provided by some participants when they recommended other potential participants. This information was used to track our snowball sample, identify jurisdictional affiliation/expertise, and prepare for interviews. We verified this background information at the beginning of each interview.

The interview instrument was developed in the fall of 2020 and was tested on our CNH colleagues at that time. It was designed to provide a flexible guide given the diverse areas of expertise of our respondents. Details on the instrument and questions asked in interviews can be found in Appendix B: Disclosure Statement. We tried to balance the ability to compare across participants (i.e., asking at least some of the same questions of all participants) with the need to collect more detailed data on different aspects of the policy process from participants with the relevant expertise. To this end, the instrument contained 5 sections, with the expectation that respondents might have different levels of knowledge across sections 2-4:

1. Background on the Respondent (verified/expanded information collected from the internet)
2. Setting of the TMDL/Watershed-wide loading goals and allocation of loading goals among state-level jurisdictions
3. Design of the WIP(s) and allocation of sub-state loading goals
4. WIP Implementation
5. Effectiveness and Future Challenges

Using the interview protocol as a template, Dr. Webster took detailed notes during each interview, producing one primary document per participant. Whenever possible, one of her RAs also took notes in a separate document, producing an additional .doc per participant. Interviews were not recorded in any other way. Notes were saved in dropbox folders by the month in which the interview occurred.

When all interviews were concluded, qualitative coding of the notes documents was conducted using Atlas.ti, which is a common tool for this type of analysis. Each of Dr. Webster's notes documents were given attributes based on sampling information (e.g., batch) and participant information (e.g., jurisdiction, expertise, level of experience, etc). Instruments were then coded as described in the rest of this section, starting with basic coding, then moving on to coding predetermined variables, and finally grouping annotations to create ad hoc codes. RA versions were consulted as needed to clarify anything missing in the primary document. Coded documents were then analyzed to assess the prevalence of different themes across a range of topics and further broken down by respondent characteristics (e.g. jurisdiction, general role, etc.) to assess patterns of narratives. We were particularly interested in areas of convergence and divergence of opinion within the sample.

Because of the high risk that confidentiality might be violated indirectly if individual participant's perspectives are communicated too clearly, we are careful to only report results that will not identify respondents to other participants in CBP governance. For aggregate analysis this is relatively easy, because opinions expressed are rarely identifying. However, when we break results down by jurisdiction, area of expertise, or some other respondent characteristics it is more likely that opinions could be linked to individuals. When there is some potential to reveal respondent identities by reporting results we group responses into clusters of at least 5 responses, including "other" categories which may lump together divergent statements. Although this may lead to under-representation of minority views, it was deemed necessary to protect respondent identities given that the CBP community is close knit and that there may be potential for professional reprisals if a respondent identity is revealed in association with an unpopular opinion. However, minority opinions were still coded and will be used to guide additional data collection and analysis for this project. Indeed, we will seek to validate all interview results using publicly available information such as meeting minutes, reports, datasets, news articles, and the peer reviewed literature.

A.2.1 Basic Coding

Coding in Atlas.ti can occur at the document level and quotation level. A few basic document level variables were coded first: 1) interview round or batch, 2) jurisdiction represented by the respondent, and 3) whether or not the respondent represented a jurisdiction with coastline on the Chesapeake Bay or not. Quotation-level coding essentially involves highlighting sections of interview notes and "tagging" these "quotes" with descriptive codes that are defined by the user. Then, the software can be used to identify occurrences (how often a code occurs), overall cooccurrences (how often two codes occur in proximity to each other), and cooccurrences by document-level variables. Note that occurrences and co-occurrences both indicate how often a code occurs, so without additional code-document analysis it is not possible to distinguish between a single respondent making the same comment multiple times or multiple respondents making the same comment only a few times. It is also possible to download the data at the quotation level into an .xlsx file for use in statistical software. Because we wanted to be able to assess cooccurrences at more than 2 levels, and to use tools that are not available in Atlas.ti, quotations were coded as large blocks of text, with all relevant codes directly associated with each block. This would ensure that we would retain all relevant codes with each quotation when we export the data from Atlas.ti.

For instance, if a respondent provided a list of popular BMPs in the states of Maryland and Delaware during the section of the interview on the WIP design process, we could have highlighted the entire section and coded it "PROC: WIP Design" and then highlighted the paragraphs on each state, coding

them “STATE: Maryland” and “STATE: Delaware” respectively, and then within those paragraphs highlight each of the BMPs and code them as “BMP: cover crops”, “BMP: streambank fencing,”, etc. This “hierarchical” approach would be limiting however, as Atlas.ti could only determine co-occurrences between the PROC (process) and jurisdiction (STATE) variables OR between the STATE and BMP variables, but not both. Since BMPs also come up in the section on the process of implementation (PROC: Implementation), this would be a serious limitation.

Therefore, instead of the hierarchical approach we used a “block” approach. This means that we would highlight an entire block of text relating to a specific process, time period, and state (if needed), then code all quotations within that block in the same field. For example, the coder would start by identifying a block of text such that, “PROC: WIP design, T9: Phase 6, STATE: Maryland” all apply to the quotations within the block. Then additional codes would be added to the block as appropriate, such as “DProc: Build on Previous, DBMP: Cover Crops, DBMP: Riparian Buffers, etc.”. As described in greater detail below, the prefix “D” also indicates that the quotation is associated with WIP Design, providing additional analytical capabilities. Overall, the block approach still allows us to use Atlas.ti’s built-in analytical tools but also ensures that when we export the quotation-level data all three (or more) codes will show up for the quote (each code is represented as a dummy variable in the .xlsx spreadsheet). All document-level variables also carry-over into exported quotation-level data.

Coding was completed in three rounds, each of which can be broken into five segments (one for each section of the interview instrument). The first segment of the first round of block-level coding focused on the background of the respondents (Section1). Respondent Background variables include their general role in the governance process, whether they have any regulatory experience (now or in the past), the duration of their involvement in Chesapeake water quality governance, general areas of expertise, sectoral areas of expertise, and involvement in CBP-related committees. The latter was broken down by committee and included the type and frequency of participation. Because we did not have time to collect detailed data on committee participation, these variables were coded based on the respondent’s most active period for each committee. That is, if someone chaired a committee 5 years ago but was not on that committee now, we would still score it as frequent participation at the leadership level for that committee. In instances where the general role or expertise of a respondent was not clear from the answers to questions in Section 1, answers to the “Please describe your role…” questions in subsequent sections were used to decide on background codes.

Once background coding was completed, we merged similar codes to ensure that resulting groups were larger than the minimum of 5 required for anonymous reporting. In some cases, mergers did not change the nature of the grouping (e.g., ExArea: Toxicology and ExArea: Water Quality) but in other cases, mergers were more forced and may make interpretation of results more difficult (e.g. merging ExArea: Agriculture with ExArea: Fisheries). Next, committee participation data were exported and used to create an index of participation (low, medium, high) for four areas of participation for each respondent (For want of better terms: Managerial (CBF, ExCom, PSC, MB), Advisory (STAC, LGAC, CAC), and Technical 1 (STAR, GITs), Technical 2 (Work Groups, Expert Panels)). Lastly, these background codes were used to create document groups and interview documents were classified accordingly. This allows us to break down all quotation-level codes within documents by background characteristics of participants. Detailed descriptions of each background code and document group can be found in the Summary Code Book (Appendix C: Annotated Instrument). Background codes were also used to validate our sampling procedures and identify professional biases in our sample and to investigate any

differences in perspectives across different types of general roles, areas or levels of expertise/experience, and engagement in the formal governance process.

A.2.2 Code Development

Developing codes for the Background Variables was relatively straight forward as all respondents were asked the same questions and provided approximately the same information. Some respondents were less comprehensive than others when answering background questions, however, so for some things, like expertise in a given sector, we may code background variables based on answers in other sections of the interview (particularly the "what role did you play…" questions in Sections 2-4). In addition, respondents did not always mention or could not necessarily list all of the working groups or other small committees that they served on, so we augmented the interview information with participant lists from the CPB's meeting documentation archive.

Segment 2-5 of round one covered the substantive portions of the interview protocol (Sections 2-5), which were more difficult to code because respondents provided different types of information and topics covered varied considerably based on the respondent's jurisdiction and area of expertise. The most effective approach was to create a range of code-groups that could be used to organize the different types of information in the interview notes. Each code group can be thought of as a variable that takes on different values (or codes). Using code groups also provides additional analytical options, as we can either use the code-group or its individual constituents for cooccurrence analysis.

Through an iterative process that involved reading through the interview notes multiple times, Dr. Webster developed a range of code groups and related codes. Some of these were obvious given the structure of the protocol and the nature of the information requested. As already mentioned, "Process" codes were used to indicate whether the quote applied to the TMDL/Allocation, WIP Design/Local Allocation, WIP Implementation, Modeling, Funding, Other, or System-level processes (e.g., general changes that might be needed to improve all of the other processes; PROC). Similarly, a set of codes was developed to indicate which phase of the TMDL/WIP was covered by a given quote (T#) or which phase of the model was discussed (MT#). Different types of actors were also used to create code groups, including state-level jurisdictions (STATE), federal government bodies (FED), state-level agencies (DEPT), local government bodies (LOCGOV), organizations and businesses of various types (ORGs), and specific individuals like politicians or other decision makers (IND). Members of each code-group were either identified *a priori* (e.g., STATE: Delaware, STATE: Maryland, etc.) or as they were encountered during the coding process (e.g., IND: Barak Obama, IND: Donald Trump, etc.). Other obvious codes groups were created for actions and activities. These included the different sectors governed under the TMDL (SECT), the laws, regulations, and policies that were passed during the management process (LRM), and the best management practices or other specific nutrient reduction methods, such as MS4s that are used in watershed implementation plans (BMP). A last general category that we did not expect when creating the protocol, but which was clearly important, was the different types of co-benefits that people might expect from the TMDL program and water quality governance generally (COBEN).

Developing codes that accurately depict values or relationships described by respondents required a more analytical approach. Once different value sets or frames were identified, they had to be grouped into categories that were detailed enough retain their descriptive value, yet broad enough to allow for aggregation across respondents. The simplest approach is to leverage the code group to indicate the direction of a specific value judgement. This was inspired by the question about "Challenges and

Opportunities" from Section 5 of the protocol. Most respondents actually described both challenges (CHLNG; factors that make it difficult to govern the watershed) and opportunities (OPP; factors that make it easier to govern the watershed) in other sections of the interview as well. Therefore, Dr. Webster went through and coded all factors identified in either category in detail and then used the "merge" function in Atlas.ti to consolidate similar factors into larger groups. Importantly, some factors were viewed as both challenges and opportunities, sometimes by the same respondent, so one might see "CHLNG: Political Will" attached to one quote and "OPP: Political Will" attached to another. In such cases, cooccurrence analysis should help reveal the conditions under which a given factor is considered a challenge vs. an opportunity.

More complicated value judgements or relationships were categorized using framing analysis. Three frames were coded a priori: technocratic, participatory, and political. As explained in Section 2 Process Frames, the definition of each frame varies somewhat depending on the process described by the respondent, but in general, technocratic frames describe the policy process as a technical exercise, participatory frames represent the process as engaging multiple actors at different levels including state agencies and local stakeholders, and political frames focus on the process as a means of distributing costs and benefits among different groups.

The remaining frames were developed based on annotations created during the initial round of coding. Annotations are like codes but contain synopsis of the information provided by the respondent on a particular topic. Because they are so detailed, they may only occur once or twice in the entire set of responses. Annotations were initially grouped by which part of the policy process they applied to and what type of frame they were being used to express. For instance, annotations that simply described factors that states considered when allocating loads at the sub-state level were given the prefix "DAlc" , while annotations that described the processes states used when making these substate allocations were given the prefix "DAProc". This differentiated sub-state allocation annotations from others such as those describing the WIP design process more broadly (DProc) or the more specific process of selecting BMPs to go into a WIP (DBProc). As noted above, for WIP Design and Implementation processes, blocks were identified with the relevant STATE code. However, other aspects of state-level jurisdictions were described by respondents, such as a state's role in negotiating the TMDL, the way other groups perceive the state's WIP Design Process, or the state's effectiveness in reaching their loading targets. These descripts would be tagged as annotations using the _Sta prefix (e.g., TSta MD for the role of Maryland in the TMDL process, DSta NY for unintended consequences of WIP Design, etc.). Similar groupings were identified for other sections of the interview notes, as explained in Section A.2.4 Annotations and Code Groups.

To minimize coding errors, Dr. Webster coded each section of the interview instrument separately during round 1. This provided for continuity within each section. At the end of Round 1, the Atlas.ti project contained 2,548 codes and annotations, 1,697 blocks of text, 30 code groups, and 62 document groups. Once this round was completed, Dr. Webster analyzed the annotations in each grouping and created new codes and code groups to summarize the detailed information contained in the annotations (Round 2). Dr. Webster's research assistant then went through the coded documents to check that all codes were applied as per code book documentation (Round 3). Unfortunately, she did not have time to check the thousands of annotations and annotation code groups, so those results should still be considered preliminary. The final Atlas.ti project described in this report contains 561 codes and 1777 annotations, aggregated in 43 and 970 groups respectively. Because annotations

contained multiple pieces of information, total appearance of annotations in code groups was much higher than the actual number of annotations (6516). Furthermore, 694 annotation code groups contained less than 5 annotations/group. Where possible we combined these with similar groups so that we could report as many of our findings as possible here.

A.2.3 Blocks, Mentions, and Co-occurrence Coefficients

Most of the analysis presented here is descriptive. The main variable of interest is the number of "mentions" for a particular code or code group. That is the number of times that a code or code group is attached to a block of text as identified using the coding procedures described above. In general, codes that occur or are "mentioned" more frequently can be thought of as more important, as long as the prompts provided by the interviewer through the survey instrument are kept in mind. That is, we would expect codes associated with specific questions or terms used in the survey instrument to occur more often than those that are volunteered by respondents.

Reporting code occurrences by block can possibly bias this interpretation in two other ways. First, since the code only occurs once per block, it may under-represent the importance of the concept labeled with the code to the respondent. For instance, a single "mention" could represent one sentence within a block, one paragraph within a block, or multiple paragraphs within that block. In an ideal setting we could assess the number of words associated with a code, and Atlas.ti does provide this capability. However, as explained above, we chose to prioritize analytical capability for complex co-occurrences of codes over words/code for this report. With several additional months of full-time work, we could add additional usage metrics for each code, but this is not planned at this point.

On the other hand, the block-based coding process may over-represent the importance of a code to the sample overall if one or more respondents chose to mention the related concept in separate blocks. For example, a respondent could mention the same code when they describe the TMDL process and the WIP design process or they could describe the same concept when discussing Phase II and Phase III WIPs. Thus, the number of mentions per code per respondent may be important if we are more concerned with representation across the sample as opposed to the importance of a given code to a specific individual. Where possible, we report the maximum number of mentions per code or code group (M*) and indicate where there may be important divergences between these two measures of "importance".

Co-occurrences are an important part of this analysis. In our dataset, codes or code groups co-occur when they are attached to the same block of text. We may also describe document-level co-occurrences, or the number of times two or more codes occur in the same interview document. Co-occurrences do not indicate any specific relationship between codes, except that both were seen by the respondent to be related to the same process, time period, and jurisdiction (for design/implementation), so additional information will be provided on relationships when possible. In addition, co-occurrences (cc) tend to be highly correlated with the number of mentions per code or code group, so we also use the co-occurrence co-efficients (ccf) provided by Atlas.ti:

$$ccf = cc12 / (m1 + m2 - cc12)$$

Where m1 = the number of mentions for the first code, m2 = the number of mentions for the 2nd code, and cc12 = number of co-occurrences between codes 1 and 2. In general, we consider a ccf of 0.2 or higher to be interesting, but really it depends on the codes we are comparing. Since our data are

categorical, we cannot assess the statistical significance of the co-occurrence coefficients ,and the coefficient is still skewed when there is a large difference between m1 and m2, but they do, at least help to account for the frequency of mentions.

A.2.4 Annotations and Code Groups

Annotations were initially coded using prefixes to indicate the process(es) that they related to, which allowed us to sort them into analytical categories. These matched the prefixes used for codes, but we expanded on this initial group as additional processes and sub-processes were identified (see Figure 89). If an annotation related to more than one process, then all relevant prefixes were included at the start of the annotation. The rest of the annotation was a brief summary of the content of the phrase being tagged. This provided much more information about relationships and value judgements than would be found in the formal codes described above. Once all annotations were coded, Dr. Webster went through each set of annotations by prefix and aggregated relevant annotations into code groups which captured one element of the information in the annotation. Thus, many of these annotations appear in more than one code group, since most contained more than one type of information. With a smaller or less complicated dataset, this intermediate step of annotation would not be needed, but given the amount of information analyzed here, we felt that annotations would give us a tool to ensure consistency and to develop more useful codes (via code groups). Analysis of annotation code groups is similar to that for regular codes, but we will rely more on correlation coefficients rather than co-occurrence as explained further below.

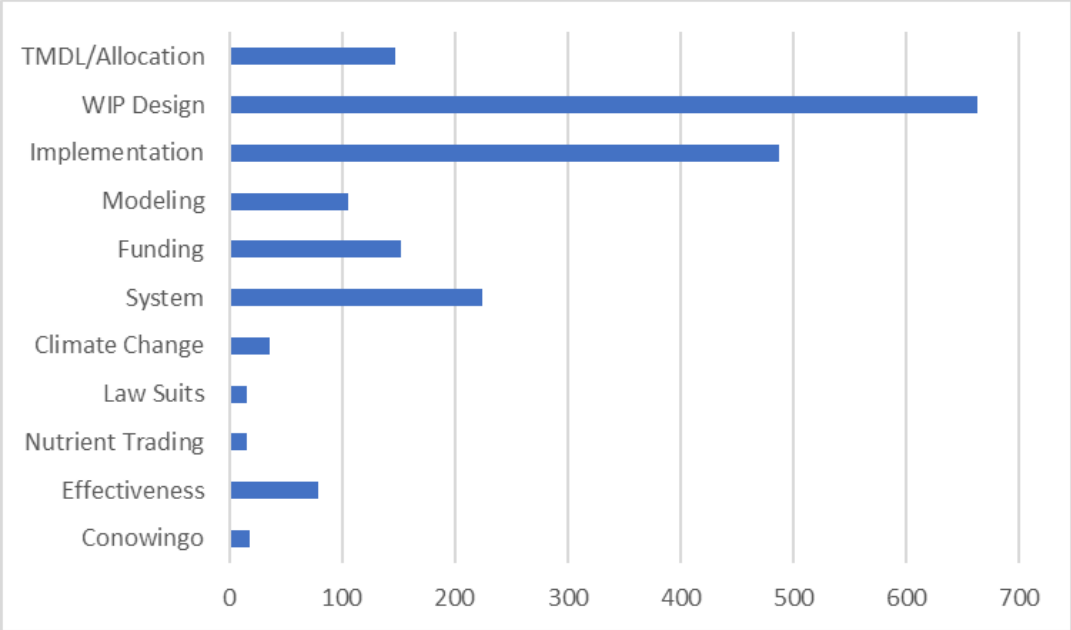


Figure 89: Annotations by Topic/Process.

To ensure consistency, we developed a more detailed prefix schema for the annotation code groups (ACGs). For the largest categories shown in Figure 4, we first created sub-categories that loosely paralleled each other (see Table 5). Then, within each sub-category we identified key variables that we would like to measure and established a system for tracking those variables across all processes. For example, for all “Proc” annotations, the addition of a 1 to the prefix indicates that this ACG refers to Endogenous Factors Influencing the Process. Different factors were then indicated with an additional

letter code, and value-based variations in the role played by said factor was indicated with another number. For example, DProc1e1: Costs/Cost Effectiveness are considered important when designing a WIP but DProc1e2: NO Costs indicates that the respondent said that costs were not considered when designing the WIP. All Prefixes were tracked in a separate excel spreadsheet so that prefixes were used consistently across process groups. For instance, FProc1e1 indicates the same thing as DProc1e1, except for the process of finding funding rather than the process of designing the WIP.

Table 5: Main Categories and Sub Categories for Annotation Coding Schema

	Main Process (_Proc)	**Allocation Process (_AProc)**	**State-Specific (_Sta)**	**Sector-Specific (_SEC)**	**BMP-Specific (_BMP)**
TMDL/ Allocation	TProc: Setting or revising watershed-wide loading target	TAProc: Distributing loading targets among state-level jurisdictions	TSta: Role of a specific state* in TProc or TAProc	NA	NA
WIP Design/ Allocation	DProc: General process used when designing the WIP for a specific state	DAProc: Process used to allocate loading targets among sub-state jurisdictions	DSta: Role of a specific state* in DProc or DAProc	DSEC: Sector-level considerations in WIP Design	DBMP: BMP-specific considerations in WIP Design
Implementation	IProc: General process used when implementing WIP	NA	ISta: Role of a specific state* in IProc	ISEC: Sector-level considerations in WIP Implementation	IBMP: BMP-specific considerations in WIP Implementation

** When that state is not the main area of expertise of the respondent*

We developed 10 variable codes for use with _Proc code groups:

1 = Endogenous Factors Influencing the Process (51 Categories w/in this Variable)
2 = How Bottom Up/Top Down is the Process (15 Categories)
3 = Monitoring/Enforcement/Accountability (NOT Verification for CAST; 7 Categories)
4 = Role of the Federal Government in Process (EPA/CBP; 12 Categories)
5 = Actors who Influence the Process (19 Categories)
6 = Sociopolitical/Exogenous Factors Influencing the Process (8 Categories)
7 = Legitimacy of the Process (39 Categories)
8 = Change in Process over Time (17 Categories)
9 = Inequalities in Process (8 Categories)
9m = Use of CAST in Process (24 Categories)

In addition to the regular _Proc variables, FProc (funding process) blocks were coded using a process-specific set of variables that described sources of funds and similar aspects of the funding process that were not mentioned for the other processes:

FProcA: Taxes/Fees Cover Costs
FProcB1: Grants
FProcB2: Subsidized Grants (Loans)
FProcC: In Kind Contributions
FProcD: Cost Share
FProcE: Matching Requirements
FProcF: Block Grants
FProcX: ANY Local Funding
FProcY: ANY State Funding
FProcZ: ANY Federal Funding

Variable codes were similar to _Proc for the _AProc code groups:

1 = Allocation Procedures (8 Categories w/in this Variable)
2 = How Bottom Up/Top Down is the Process (5 Categories)
4 = Role of the Federal Government in Process (EPA/CBP; 2 Categories)
6 = Sociopolitical/Exogenous Factors Influencing the Process (2 Categories)
9m = Use of CAST in Process (2 Categories)

Classifying code groups was somewhat different for _Sta variables. First, we created jurisdiction code groups for each process:

_Sta1 = DC
_Sta2 = DE
_Sta3 = MD
_Sta4 = NY
_Sta5 = PA
_Sta6 = VA
_Sta7 = WVA
_Sta8 = Other
_Sta9 = EPA or CBP

Then, we created variable groups using capital letters as identifiers instead of numerals such that:

Code	Description	Code	Description
A	**Proportion State in Watershed**	**G**	**Difficulty of Meeting Load Requirement**
A1	Small proportion	G1	Low Difficulty
A2	Medium Proportion	G2	Moderately Difficult
A3	Large Proportion	G3	High/More Difficult
		G4	Cannot Meet Load

B	**Gain from CBP**
B1	Little to Gain
B2	A Lot to Gain
B3	Costs/Benefits Uncertain
B4	A Lot to Lose
C	**Fragmentation (State-Level)**
C1:	Geographic Divide
C2	Political Fragmentation
C3:	Some Counties Lead
C9:	Collaborative in Some Sectors
D	**Buy-in**
D1:	State is Leader
D2	State does Minimum Required
D3:	State is Laggard/Ineffective
D4:	State Claims Leader, Still Lagging
E	**Partisanship**
E1	Republican Majority
E2	Bipartisan
E3	Democratic Majority
F	**Concern for Bay**
F1	High Concern for Bay
F2	Low Concern for Bay
F3:	Local Water Quality > Bay WQ

H	**Trust/Reputation**
H1a	State Trusted/Respected
H1b	State Not Trusted/Respected
H1c:	State Working to Gain Trust
H2a	Locals Trust State
H2b	Locals Do NOT Trust State
H3	Conflict Over BMP Details (Among States)
W	**State Resources**
W1:	Low Funding/Resources
W2:	Medium Funding/Resources
W3:	High Funding/Resources
X	**Transparency**
X1:	Transparency Increasing
X2:	Transparency Decreasing
Y	**Meeting Loading Goals**
Y1:	Likely to Meet Loading Goals
Y2:	NOT Likely to Meet Loading Goals
Z	**Proportion of Load by Sector**
Z1:	Mostly Wastewater
Z2:	Mostly Agriculture
Z3:	A Lot of Unregulated/Voluntary
Z4:	Mostly Forest

This dual-code group approach will let us examine jurisdiction-specific statements in aggregate as well as breaking them down by jurisdiction.

_SEC and _BMP variables were treated in much the same way. Annotations were first assigned to code groups based on the relevant sector or BMP grouping:

_SEC	Sectors	_BMPd	Barnyard Improvement
_SECa	Agriculture	**_BMPe**	**Buffers**
_SECb	Stormwater	_BMPe:a	Forest Buffers ONLY
_SECc	Wastewater	_BMPe:b	Grass Buffers ONLY
_SECd	Atmospheric Deposition	**_BMPf**	**Fencing/Exclusion**
_SECm	Urban	**_BMPg**	**Wetlands**

_SECy	Point Source	**_BMPh**	**Trees/Forest (NOT Buffers)**
_SECz	Non-Point	**_BMPi**	**Tillage Practices**
_BMP	**BMPs**	**_BMPj**	**Nutrient Management**
_BMPa	**Cover Crops**	**_BMPk**	**Precision Feeding**
_BMPb	**AG Waste/CAFO**	**_BMPl**	**Wastewater**
_BMPb:a	FAWMs	_BMPl:a	Treatment
_BMPb:b	Composters, Digesters, Gasification	_BMPl:b	Septic/Sewer
_BMPb:c	Removal/Transport	**_BMPm**	**Flood Mitigation/Floodplain Restoration**
_BMPb:d	Storage	**_BMPn**	**AG Bundles**
_BMPb:e	Other	**_BMPo**	**Natural Filters (NOT Buffers/GI)***
_BMPc	**Stormwater**	**_BMPp**	**Tech/Non-Tech/Cultural**
_BMPc:a	Private Green Infrastructure (PrivGI)	**_BMPq**	**Stream/Shoreline Restoration**
_BMPc:b	Retrofits/Remove Impermeable Surface	**_BMPr**	**Street Sweeping**
_BMPc:c	Public Green Infrastructure (PubGI)	**_BMPs**	**Structural/Permanent**
_BMPc:d	Stormwater Mngmnt	**_BMPt**	**Soil Amendments**
_BMPc:e:	MS4 Permits	**_BMPu**	**N/P Bans**

Then, _SEC and _BMP annotations were placed in categories based on a similar system of numbered groups to _Proc and _AProc annotations, where:

_Sector Annotations	_BMP Annotations
1 = Role of Sector in Process	1 = Factors affecting BMP selection/implementation
2 = Effectiveness/Efficiency of Sector	2 = Effectiveness/Efficiency of BMP
3 = Feasibility of Load Reduction in Sector	3 = Feasibility of BMP
4 = Political/Economic Attributes of Sector	4 = Costs and Benefits, Distributional Aspects of BMPs
5 = Funding and Equity in Sector	5 = Funding and Equity of BMPs
6 = Incentive Structure for Sector	6 = Incentives for BMPs
7 = Capacity and Synergies among Sectors	7 = Capacity and Synergies among BMPs
8 = Temporal Aspects of Sector	8 = Temporal Aspects of BMPs
9 = Monitoring and Learning	9 = Learning
9m = Use of CAST/Bay Models	9m = Use of CAST/Bay Models

Other primary groupings, such as annotations about lawsuits, climate change, and nutrient trading had a low enough number of associated blocks that complicated coding schema were not needed. We therefore describe the variables associated with these groups in the relevant results section. While the fundamental approach to analyzing annotation code groups is the same as for regular codes, Atlas.ti does not provide this functionality. Therefore, all of the analysis of code groups is undertaken using exported data from Atlas in the "Statistical" format (.csv), which treats every block of text as a row and all codes, document groups, and code groups as dummy variables for any given block. From these data we were able to create co-occurrence tables using a routine programmed in R.

Appendix B: Disclosure Statement

DARTMOUTH COLLEGE

Modelling the dynamics of human and estuarine systems with regulatory feedbacks

RESEARCH PROJECT INFORMATION SHEET

This research project is being conducted by D.G. Webster from Environmental Studies Department at Dartmouth College, Hanover, NH, USA. It is a study of policy decision making regarding water quality management at the federal, state, and local levels for the Chesapeake Bay Area. Information collected will be used to create and then test a computer model to predict likely policy changes under a range of future scenarios. It is part of a larger project funded by the National Science Foundation called Modeling the Dynamics of Human and Estuarine Systems with Regulatory Feedbacks (Award #2009248). Using the Chesapeake Bay as an example, this project will combine the policy model that we are designing with biophysical models to predict how social, economic and policy changes impact water quality, and how changes in water quality influence human behavior and decision-making.

Your participation in this research is voluntary. Participation involves a 30-60 minute interview about the policy processes in the study region, with a focus on your agency/group, as well as your perceptions of the role of other actors. You may choose not to answer any or all questions and do not need to provide any reason for declining to answer. Information provided will be stored in a Dropbox folder and on an encrypted hard drive. It will only be shared among myself, my colleague Patrick Bitterman from the University of Nebraska, Lincoln, and our research assistants. Your responses will not be published or shared with other members of the project team except in aggregate form without any identifying information. We will not quote any of the information that you provide without first obtaining written permission from you. This applies whether or not the quote is attributed to you. While we appreciate the openness and transparency of the policy process in the Chesapeake watershed, some of the information collected during the interview may be political in nature and so we believe that it is important to maintain the confidentiality of our respondents.

Questions about this project may be directed to:

D.G. Webster, Associate Professor, Environmental Studies, Dartmouth College
6182 Steele Hall, Hanover, NH 03755
d.g.webster@dartmouth.edu, +1 603-646-0213

Questions about the ethical review of this project may be directed to:

Dartmouth College Committee for the Protection of Human Subjects
+1 603-646-6482 or CPHS.Tasks@Dartmouth.EDU

Appendix C: Annotated Instrument

Since we interviewed people from many different agencies and jurisdictions our primary instrument was designed to provide general guidelines only. Preliminary tests of the template were conducted in November and December of 2020. While the general structure remained the same, phrasing and focus changed over the course of the interview period. In the following, Text in BLACK was read or paraphrased for all respondents, assuming time permitted. Text in GREY indicates questions that might be omitted depending on the respondent's area of expertise or the amount of time remaining for the interview. Text in BLUE provides information on how some of the text evolved as we learned from each interview.

Section 1: Background of Respondent

Thanks for agreeing to this interview. First, I'd like to verify that you were able to read our disclosure statement (if No, provide time to read it). Any questions or concerns before we get started?

We're very early in the model development process and are mainly trying to identify sources of information—after all, if we can't measure it, we can't model it. To that end we've collected a lot of publicly available data and have a solid understanding of the technical and legal aspects of the policy process. From these interviews, we're most interested in how "behind the scenes" activities affect policy making. We can use your help to identify political, economic, or social factors that are important determinants of policy design or implementation, but which are not easily observed by outsiders.

Because respondents seem to react negatively when questions asked were not in their area of expertise, we replaced the above statement with: Since we are interviewing people who play many different roles in the CBP governance process, the questions I'll be asking are very general. Some will be right in your wheelhouse; others may be outside your area of expertise. For comparability I will be asking the same broad questions of everyone, but we'll spend most of our time on the topics that you know the most about.

There are five sections of interview questions. First, we'll ask you a bit more about yourself. Second, we'll ask about the setting of the TMDL and division of loading goals among the state-level jurisdictions. Third we will ask about the development of the WIP(s). Fourth, we'll ask about the implementation of the WIPs. Lastly, we will ask some big picture questions about the program as a whole.

A.1. Name and Title of respondent [best to fill in in advance, then confirm]:

A.2. How would you describe your role in the governance of water quality for the Chesapeake Bay Watershed?

   a. How long have you held this role?

   b. How long have you been involved in Chesapeake watershed water quality management? This included follow-up questions to get a good idea of the respondent's history in relation to the CBP.

c. What previous positions have you held in Chesapeake governance? In water quality governance elsewhere? Follow-up questions as needed to discern pre-CBP background.

A.3. Which of these committees have you served on, testified to, or observed?

- When respondents answered in the positive for any committee, we used follow up questions to get a more detailed understanding of the degree and timing of their involvement.
- Because some respondents reacted negatively to being asked about committees that were far outside of their field, we began prefacing this question by stating that we were asking all respondents to go through the same list and that it was normal for people to be heavily involved in some committees but not at all involved with others.
- Respondents with long histories at the CBP sometimes had difficulty with this section because of changes in the committee structure. When this came up, we made a note of it.

a. Chesapeake Bay Commission

b. Chesapeake Bay Program Top

   i. Executive Council

   ii. Principle staff Committee

   iii. Management Board

c. CBP Advisory

   i. Scientific and Technical Advisory Committee

   ii. Citizen's Advisory Committee

   iii. Local Government Advisory Committee

   iv. Other Advisory Committees? [prompt to name]

d. CBP Internal

   i. STAR Group (Scientific, Technical Assessment and Reporting)

   ii. Water Quality Goal Implementation Team (GIT)

   Follow-up questions used to determine involvement with any of the work groups.

   iii. Other GiTs? Workgroup participation only recorded if volunteered.

      1. Sustainable fisheries

      2. Habitat

      3. Maintain healthy wetlands

      4. Fostering Chesapeake stewardship

5. Enhance partnering, leadership, and management

iv. Other internal groups, teams, committees?

e. Other relevant experience w/ policy-relevant organizations? We found that it was easier for respondents to answer if we asked them: "Do you frequently interact with _____"?, filling in the blank with each of the groups listed below and then prompting them to list the relevant entities or provide representative examples.

i. NGOs? Because most respondents interpreted "NGOs" as "environmental organizations" we changed this to "Environmental, agricultural, or industry organizations".

ii. Federal or state agencies?

iii. Federal or state legislatures? If yes, we would follow up to determine if the role was official, including testimony to the legislature or informal such as providing information.

Section 2: TMDL and Jurisdictional Allocations

There are four more sections of questions, each of which focuses on a different part of the policy process: 1) setting the TMDL and jurisdictional allocations/planning targets, 2) designing the WIPs, 3) achieving WIP goals, and 4)future challenges. Since we're interviewing people with many different types of expertise in Chesapeake water quality governance, you may have more information on some questions than others. If a question falls outside of your area of expertise or you'd rather not answer it for any reason, feel free to decline to answer. As per the note above, we revised this statement and moved it to the beginning of the interview.

This section focuses on the creation of the TMDL and allocations among state-level jurisdictions.

B.1. What (if any) role did you play in the setting of the TMDL or allocation of loading requirements among the state-level jurisdictions in 2010? Initially we limited this question to 2010 to keep the interview time under 60 minutes. However, because we had such a wide variation in the length of involvement in the program and there have been various revisions of the TMDL/allocations since 2010, we eventually left the question more open, first asking respondents to identify WHICH of the phases/adjustments they were involved in and THEN asking how they were involved in each.

B.2. Can you briefly describe how the (Phase III) load limit for the Chesapeake watershed was set BASED ON recommendations from the CBP and CAST? As above, we ended up asking people which Phases they were involved in (usually eliciting this in Section 1) and then asked them to talk about each phase/changes between phases. We frequently omitted the reference to CAST because it either produced highly detailed and technical descriptions (from modelers) OR shut down responses (from non-modelers).

B.3. Can you describe how the (Phase III) TMDL allocations were distributed among the state-level jurisdictions? See above in re: different phases/adjustments.

Section 3: Designing WIPs

C.1. What (if any) is your role in the process of designing WIPs (e.g. selecting the BMPs) for one or more state-level jurisdictions? Prompt to describe each phase separately/consider differences among phases if applicable.

C.2. How does the state translate its TMDL allocation into local planning targets? For instance, does the state first determine sectoral allocations (wastewater, stormwater, ag, etc.)? or does it start with watershed-level allocations?

C.3. How are BMPs selected? We found it was more effective to ask:

a. What are the main criteria that were used when selecting BMPs (during the relevant phase)?

b. Are there any BMPs that have been popular, but which are not very effective at reducing loads? [Follow up for details/rationale]

c. Are there any BMPs that are very effective but have been difficult to include in the WIP? [Follow up for details/rationale]

d. Is the BMP selection top-down or bottom-up?

e. How does the state use CAST when designing the WIP/selecting BMPs? Is it primarily used by only a few people or does it get used in stakeholder engagement? How much do they focus on cost effectiveness vs. meeting their planning targets? Given that there are so many possible BMPs in CAST, what did you take as your starting point (prompt if needed: WIP from previous phase, current status, best possible combination)?

C.4. Can you describe how each of the factors listed below might facilitate or hinder effective WIP Design? This question produced interesting answers but ended up taking too much time and so was replaced by the questions in re: criteria above.

a. Technical challenges or innovations

b. Economic costs or benefits of BMPs

c. Political will generally

d. Lobbying by powerful interest groups

e. Availability of funds

f. Availability of human resources

g. Other?

Section 4: On the ground implementation of WIPs

D.1. What (if any) is your role in the process of ensuring that your jurisdiction achieves the goals set in the WIP design process described above—that is, making sure that the plans laid out in the WIP are actually put into practice on the ground? By this section, we usually had a good idea of the respondent's role and also were running short on time so would skip this question.

D.2. Can you describe how the WIP goals have been achieved, either specifically from your experience or (if no experience) in general? It proved much more effective to ask respondents to describe BMPs that were successfully implemented and those that did not work as well as expected. Follow up questions were then used to get more information on the sources of success/failure.

D.3. What (if any) is your role in the process of providing funding for WIP Implementation? Same as D.1.

D.4. Can you describe how WIP funding is distributed either specifically from your experience or (if no experience) in general? Omitted for respondents with little involvement in funding or if time was running out. Usually paraphrased based on information already provided in previous answers to save time as well. For instance, if the person mentioned that grant-writing was an important part of their role in implementation, I would ask them about the grant-writing process, whether some types of BMPs were easier to fund via grants, etc.

D.5. For the examples provided, was goal achievement limited or facilitated by any of these factors: Similar to C.4, interesting but we usually ran out of time to go through all of these. Many did still show up in other responses, however.

a. Technical challenges or innovations
b. Economic costs or benefits of BMP
c. Political will generally
d. Lobbying by powerful interest groups
e. Availability of funds
f. Availability of human resources
g. Other?

Section 5: Effectiveness and Future Challenges

In the first few interviews, this section focused more specifically on how political pressures shaped the evolution of the CBP program and what might need to change in future. We found that these questions tended to be redundant (participants usually volunteered this information in earlier sections, if they had anything to say on it) and we were also struggling to stay within the 60 minute period, so we reduced the scope to the questions shown below.

B.1. Do you think that water quality governance in the Chesapeake watershed is effective? Why or why not?

B.2. What do you think are the biggest challenges for water quality governance in the Chesapeake over the next 25 years? [prompt on Climate Change if not raised; we also started prompting on Conowingo after it came up a few times]

B.3. How might the policy process need to change to improve effectiveness and/or address those future challenges? Usually, this question was already answered in previous responses, so we dropped it.

Finally:

Thank you for your participation in this research project. We could not do this work without the important information provided by experts like yourself.

Would you be willing to be contacted again if we have any follow up questions?

Would you like to be added to a list of people who will receive updates on the progress of the project?

Would you be willing to recommend colleagues who might be willing to be interviewed for this project? We had hoped to use this question as the basis for network analysis to test the sample for potential biases but found that a more detailed set of questions would be needed to collect such information in a reliable way.

Appendix D: Respondent Backgrounds (Aggregated)

In addition to helping ensure a representative sample, information on respondent backgrounds was collected to determine how respondent characteristics affect their perceptions of the policy process. Variables described here were used throughout the report whenever observations were sufficiently high to protect confidentiality.

D.1 Employment and General Role

First, respondent's **general role** was coded based on their own description of the work they do in the CBP policy arena. As shown in Figure 90, we aggregated these roles into five categories. Boundaries between some of these categories are fuzzy and, for some individuals, placement in a category required a subjective assessment of which category was the best fit given that they played several different roles in the policy process. *Tech/Modelers* described their role as primarily working with the Bay models or related modeling technologies (e.g., GIS systems, other water quality modeling systems). *Scientists* described their role in terms of data collection and analysis. Most were natural scientists studying water quality or its effects on ecosystems, but this category also includes some social scientists. Of course, there is considerable overlap in the competencies of respondents in these two roles, as scientists often work with models or related technologies and modelers are usually scientifically competent and may even have advanced degrees in a scientific field. Some of the difference in these groups depends on employers—scientists were more likely to work for universities, tech/modelers were more likely to work for the government or private consultancies—but we primarily based this decision on how the respondent presented themselves and their work. Although we are not reporting these results by state, we did note that the Maryland group has a higher number of scientists when compared to other jurisdictions; this is at least partially due to the sampling process, which started with recommendations from colleagues at the University of Maryland and CBP.

Regulators were directly involved in government decision making. That is, they design laws or regulations and may also decide on punishment for violations. As the chart shows, they were the largest group of respondents. Most were state-level regulators, but federal and local regulators were also represented. Most of the respondents who fit in the *implementation* category also worked for the government, but their role focused more on facilitating the implementation of existing laws or regulation rather than deciding on them per se. Some members of this group also worked for NGOs or industry interest groups, but all could be thought of as "boots on the ground"; the people who work with stakeholders to get BMPs in place. Lastly, the *communication* group is a mix of individuals with different types of employers, all of whom saw communicating information as their primary role. Some focused that communication toward stakeholders and the general public (outreach); others were more focused on providing information to decision-makers. A few worked to move information in both directions. Interestingly, many other respondents also said that communication was an important part of their job, but this group only contains respondents who said that communication was their primary task.

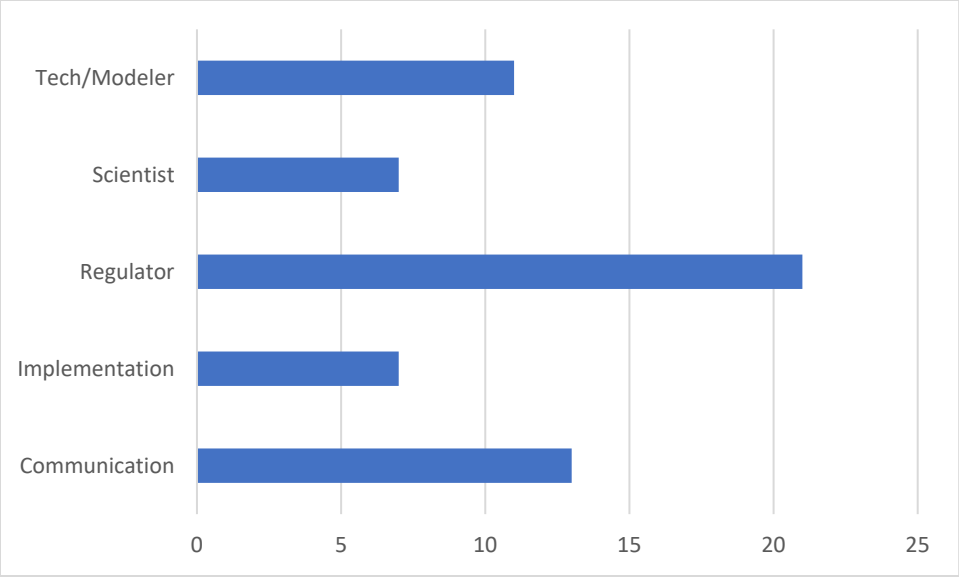


Figure 90: Number of Respondents by their General Roles

Respondents were also grouped by type of employment (see Figure 91). About 1/3 of the respondents worked for their jurisdiction's version of the department of the environment, which in all cases was also the primary state agency with responsibility for CBP-related activities (e.g., Department of Energy and Environment for Washington, DC, Department of Environmental Conservation for New York, etc.). We also recruited respondents from other state or local level agencies with connections to the CBP (e.g., departments related to agriculture, natural resources, and forestry; groups representing local planning commissions, soil and water districts, etc.). As noted in the section on Sampling, our stratified sample also includes a group of federal employees who work with the CBP or the EPA. In addition, we interviewed some people who worked for academic institutions and others who worked for private consultancies, most of which provided either technical or legal services. Lastly, to ensure a more balanced picture of the policy process, a number of our respondents were employed by non-governmental organizations, industry or municipal groups, and other private sector groups.

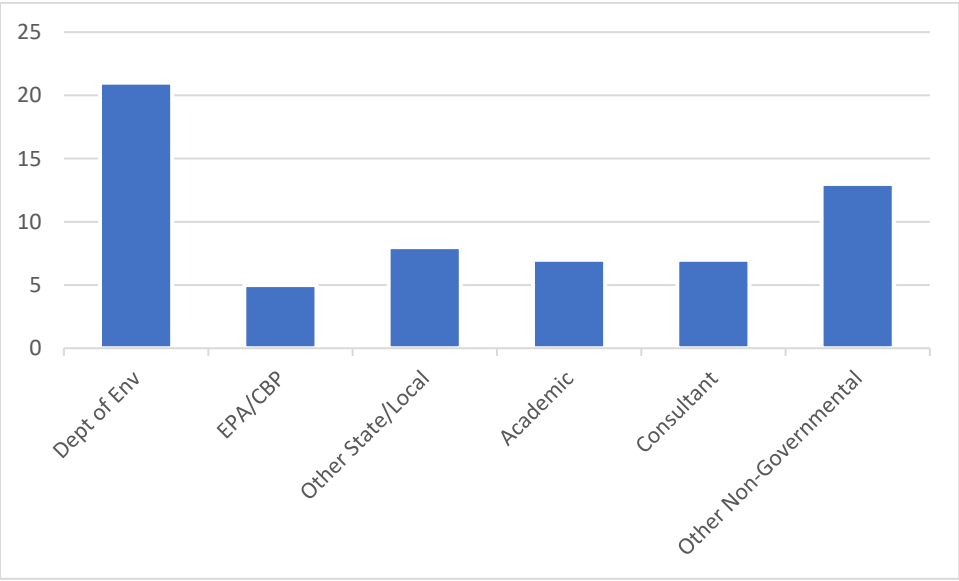


Figure 91: Number of Respondents by Employment

The above breakdown is based on current employment, but we also asked respondents about their employment history in the CBP realm. Many of the respondents had spent most of their careers working on water quality in the Chesapeake watershed. We also noted that 38 respondents mentioned working for a government regulatory agency at some point in their careers, so there was a small amount of movement between public and private sectors (5 respondents).

D.2 Experience and Expertise

For more detailed analyses, we also collected information about the respondents' length of experience working on water quality in the watershed, their various areas of expertise, and their sector(s) of expertise. As shown in Figure 92, the sample demonstrates a fairly normal distribution in length of experience, though there is a bit of a skew toward people with more than 10 years of experience. This was not entirely intentional on our part, but resulted from the snowball sampling process, as most respondents recommended others who they felt would have spent sufficient time in the region to build up the expertise needed to answer our questions. As noted above, a large majority of respondents spent their entire careers working in the watershed, though several moved from more scientific/technical work into managerial/regulatory positions, and some shifted from government to private sector employers.

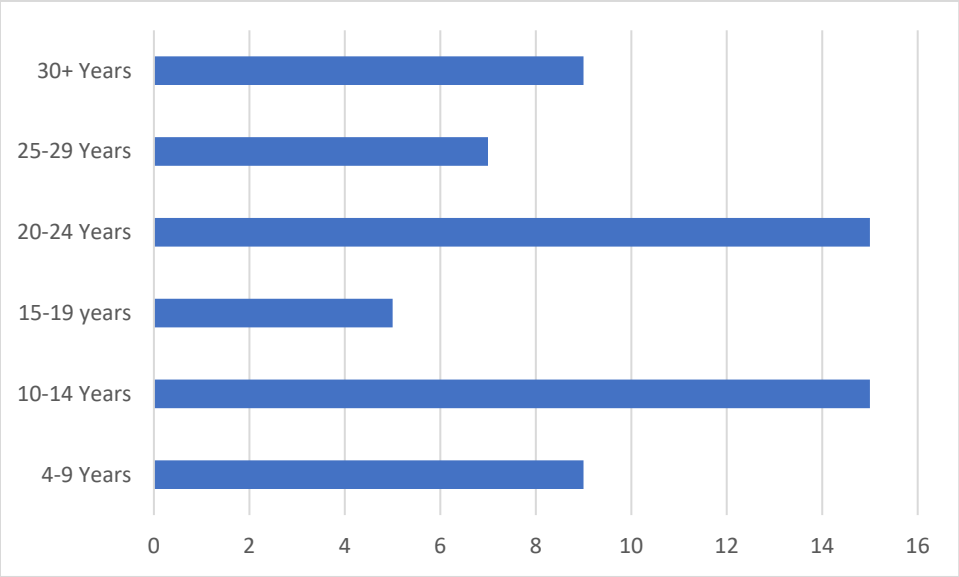


Figure 92: Number of Respondents by Length of Experience

As noted in the previous section, respondents had multiple competencies and responsibilities either within or in addition to their general role. These were coded in the (Policy) Area and Sector of Expertise variables. While each respondent was assigned only one general role, we coded all areas/sectors of expertise mentioned by the respondent. This includes direct statements that X was an area of expertise and statements indicating that the respondent had served on relevant committees, done necessary research, etc. to develop said expertise.

Figure 93 shows the distribution of expertise across policy areas. We initially divided these among the three main components of the policy process that were used to organize questions in the survey instrument (TMDL, WIP Design, and WIP Implementation), but additional areas were added to accommodate respondent answers. Thus, we were able to distinguish between expertise in setting the

total maximum daily load and expertise allocating loading targets to jurisdictions, though there was considerable overlap between these categories. Other sub-components of the policy process were also mentioned by the respondents, including monitoring and reporting for purposes of compliance, obtaining funding or other resources for implementation, and engaging in outreach or education. Modeling was also a reported expertise, whether for regulatory or scientific purposes. Similarly, a large number of respondents indicated expertise in water quality management or sciences more broadly and many evinced expertise in the areas of policy, law, social science, or economics. Other natural sciences were also represented but not in sufficient numbers to report individually.

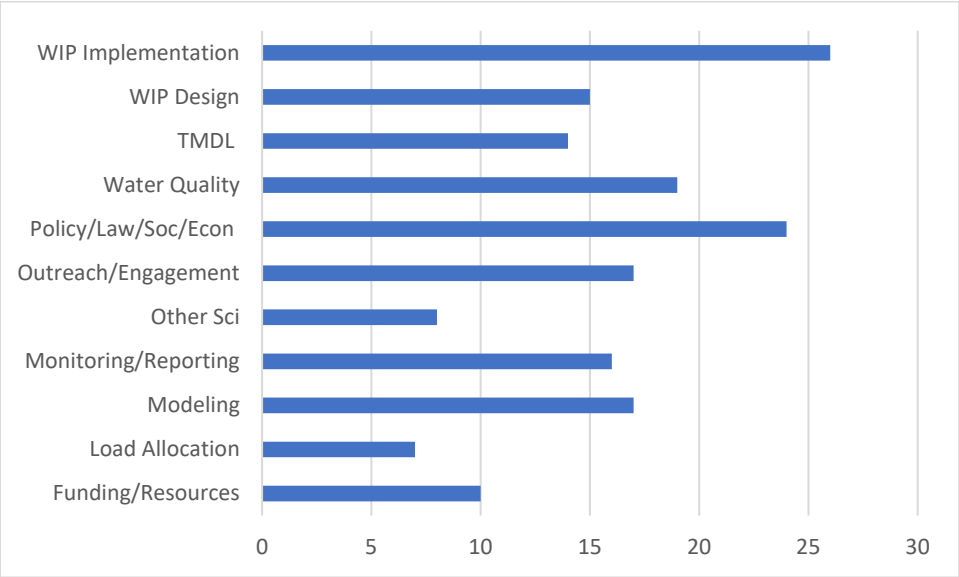


Figure 93: Number of Respondents by Area of Expertise

The wealth of expertise represented by this group of respondents can also be seen in Figure 94, which shows how diverse the areas of expertise are for each of the general roles described in the previous section. Among the "scientist" role, we see that different fields are represented from the natural and social sciences, along with some practical experience with some parts of the policy process. The "regulator" role covers a range of expertise across the policy process, but also in water quality management, modeling, and some of the sciences. Similarly, the "tech/modeler" role shows a concentration of expertise on "modeling", but expertise in the sciences and components of the policy process are also represented. The "communication" role has concentrations of expertise in outreach/engagement and WIP Implementation, as might be expected, but all members of this category also expressed expertise in the realm of policy/law/social science/economics. It is also interesting that respondents in the "implementation" role expressed expertise in finding funding/resources more than any other role. This group was also most focused in areas of expertise.

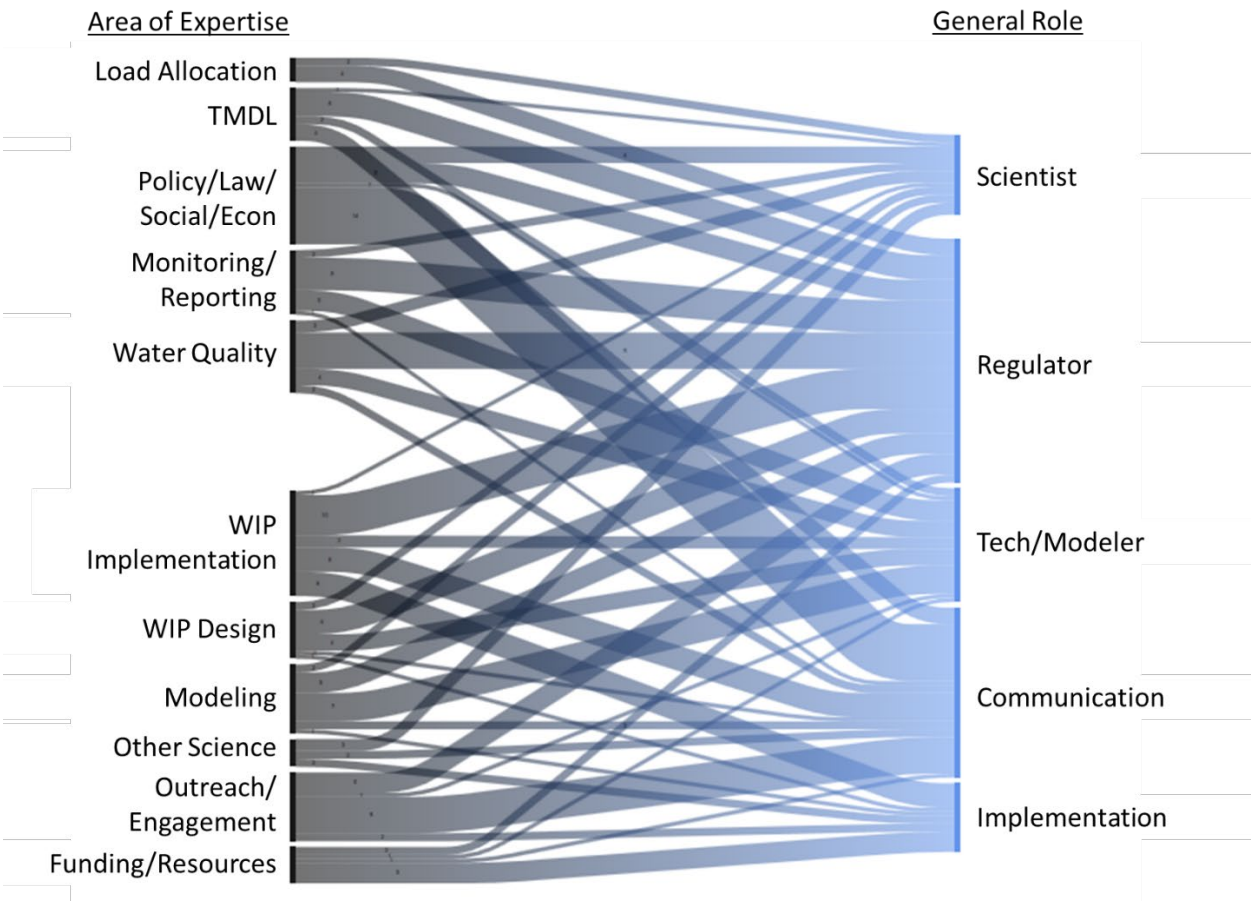


Figure 94: Sankey Diagram of Area of Expertise by General Role

For water quality governance, sectoral expertise is important as well. Figure 95 shows the distribution of sectoral expertise among our respondents. Agriculture/fisheries was the largest group, followed closely by stormwater, and wastewater. Watershed restoration was also mentioned as a sector of expertise by about 1/6 of our respondents. Forestry is the least well-represented sector of expertise. Here again, the snowball sampling method may have led to under-representation of respondents with expertise in the forestry and restoration sectors, but this may also reflect the fact that these sectors tend to be smaller components in watershed implementation plans.

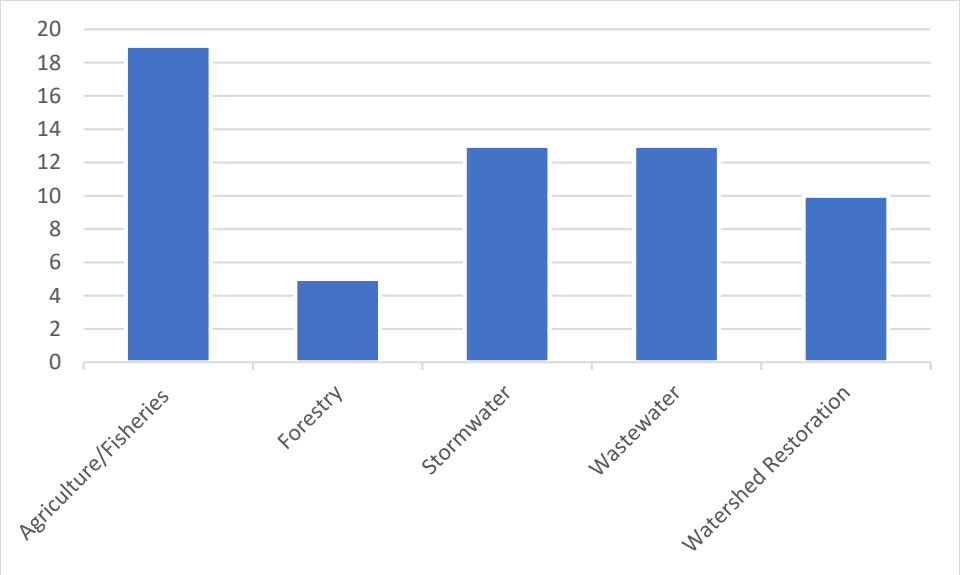


Figure 95: Breakdown of Sectoral Expertise of Respondents

D.3 Participation and Outside Engagement

The remainder of questions in the first section of the interview instrument were designed to assess how much the respondent participated in formal CBP processes and which groups they engaged with outside of the CBP. First, respondents were asked if they observed, testified to, served on, or otherwise participated in the committees listed in Figure 96. All but one of these committees is part of the CBP organizational structure. The Chesapeake Bay Commission (CBC) is a separate entity that works to coordinate legislation across its three member states (Pennsylvania, Maryland, and Virginia) and with the Federal Government. It is also a signatory to the agreement that led to the creation of the CBP and remains an important forum for developing water quality policies in the Chesapeake watershed. In addition to the committees listed in our survey instrument, respondents also volunteered information about their participation in work groups or sub-committees within the CBP and in workshops held by the Scientific and Advisory Committee (STAC). Since these responses were volunteered, it is likely that they under-represent participation in these groups; other respondents may have participated in workshops or workgroups but did not bring this up because they were not asked. In all, the breadth of participation suggests that the sample is representative of decision makers across management, advisory, and technical components of the CBP, though people engaged in detailed technical work may be under-represented. This may be either a result of the sampling procedures (especially the fact that we sought out interviewees with knowledge of the policy process), but there might also be some self-selection, as individuals with narrow technical expertise may have chosen not to accept invitations based on their own assessment of their fit w/ our criteria or unwillingness to engage with any activity that might have policy or political aspects.

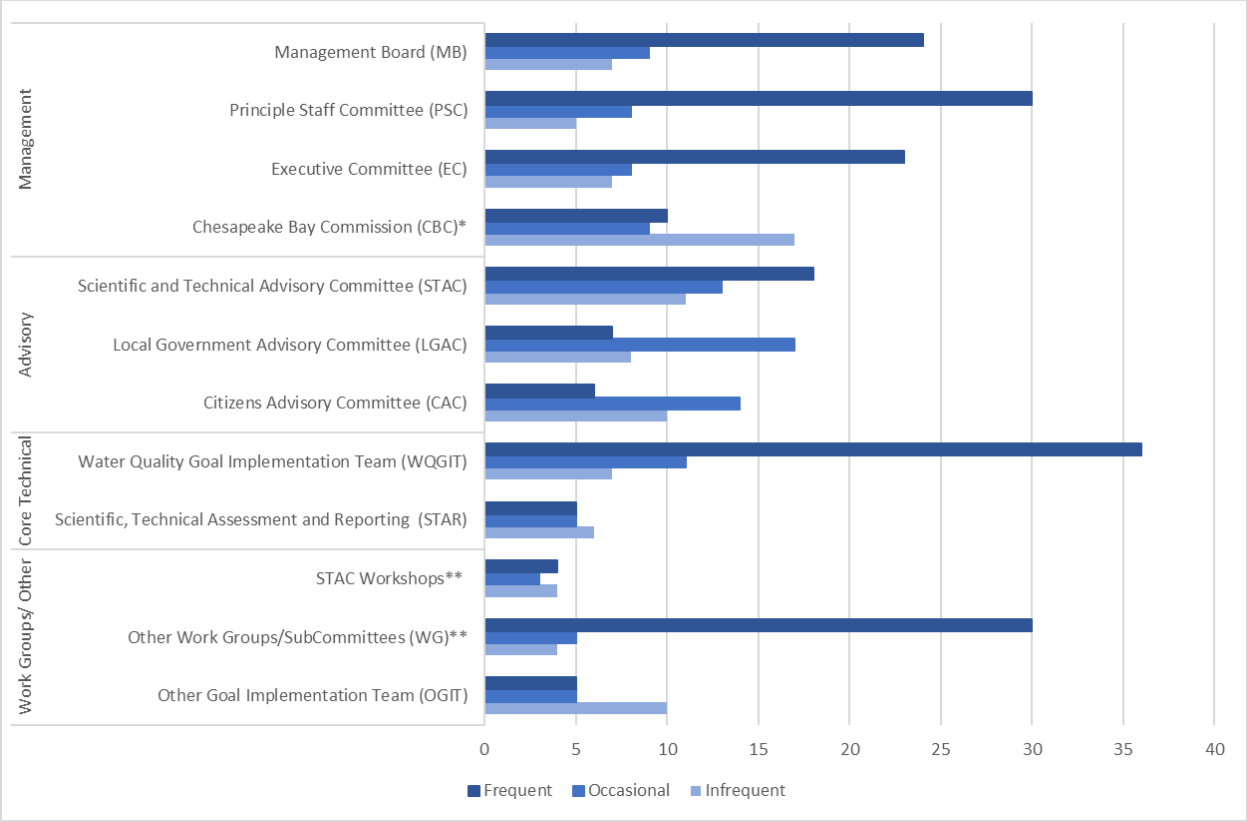


*Figure 96: Participation in CBP-Related Committees (*Not a CBP committee; **Volunteered by respondents)*

For each committee, we asked respondents to describe the type and frequency of their participation. Categories included frequent (large percentage of the meetings), occasional (more than a handful of meetings), and infrequent (only one or two meetings). This is not the best measure of frequency, but due to time limitations and the fact that the list of committees was quite long, we considered it to be sufficient for our purposes. As can be seen in Figure 96 above, frequent participation was most common in the managerial committees, STAC, the Water Quality GIT, and (mostly WQGIT) work groups, which is again consistent with the target population for our sample.

Figure 97 below shows how frequency of participation relates to the type of participation described by each respondent. While we initially only asked them if they 1) attended/observed, 2) testified/presented to, or 3) were a formal member (voting or at large) of each committee, respondents identified several other types of participation, including chairing/leading a committee (overlaps w/ formal membership), providing information or support to formal members of the committee, and helping the committee with outreach or engagement with other committees/stakeholders. Supporting and coordinating activities proved to be one of the most common types of interaction with the committees in our instrument and could either be individuals providing support or information to higher-level actors on the PSC, or Executive Council or coordinating with/supervising employees who sat on sub-committees (e.g., Water Quality GIT, workgroups, etc.). In general, we expected anyone who claimed a chair/leadership role to frequently participate in a committee, but in a few cases, respondents indicated that they chaired/lead a STAC workshop, but that they participated in STAC workshops infrequently/occasionally, which created the appearance of an anomaly.

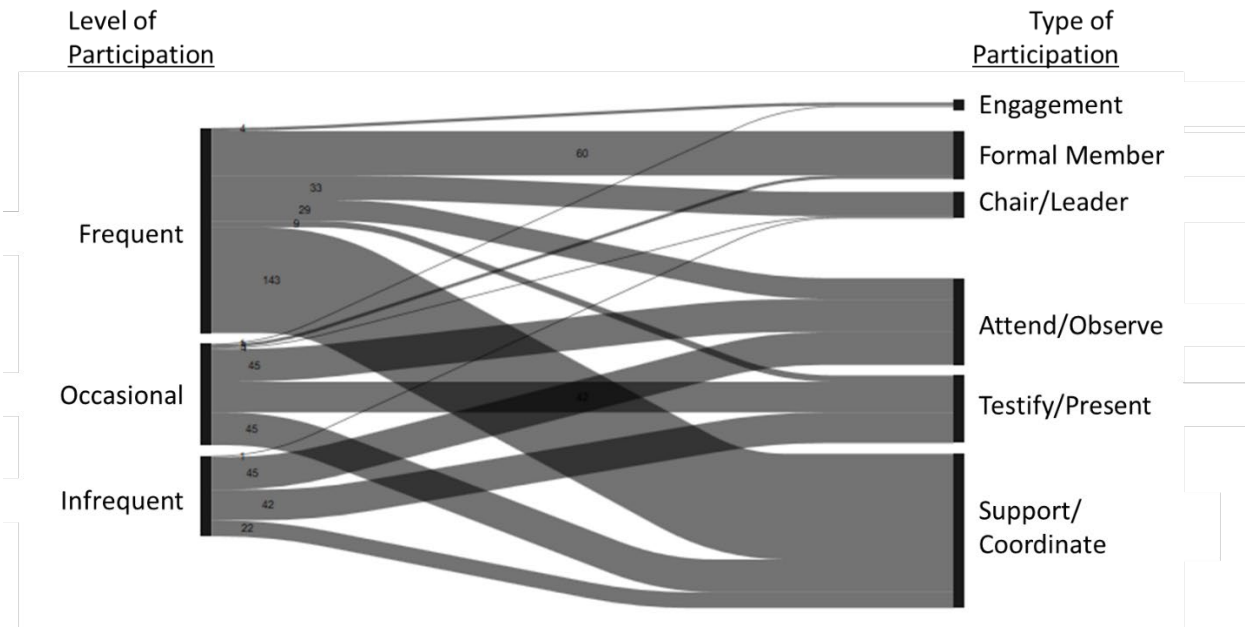


Figure 97 Sankey Diagram of Type and Frequency of Participation

Since reporting participation by committee could reveal respondent identities, we converted responses regarding the frequency and type of participation into levels of participation in four categories, as shown in Figure 98. Type of participation was weighted by frequency for each committee, and then the average was calculated for managerial, advisory, and core technical committees (see Figure 96 for committees included in each category) and normalized to a scale of 1 through 5 for comparison across all four categories. Because this process created a few small groups, we then aggregated scores such that 1-2 = low participation, 3-4 = medium participation, and 5 = high participation. Analysis of workgroup participation indicated that some respondents would have much higher levels of participation if workgroups were included in the analysis, but since we did not ask all respondents about their participation in workgroups, we could not use the same procedures as for the other committees. Instead, we used the CBP's on-line archive to search for each respondent's name and tallied the number of workgroups that they participated in at least once, based on the agendas and meeting minutes. We then normalized the simple sum of participation, which was then treated like the scores for the other categories.

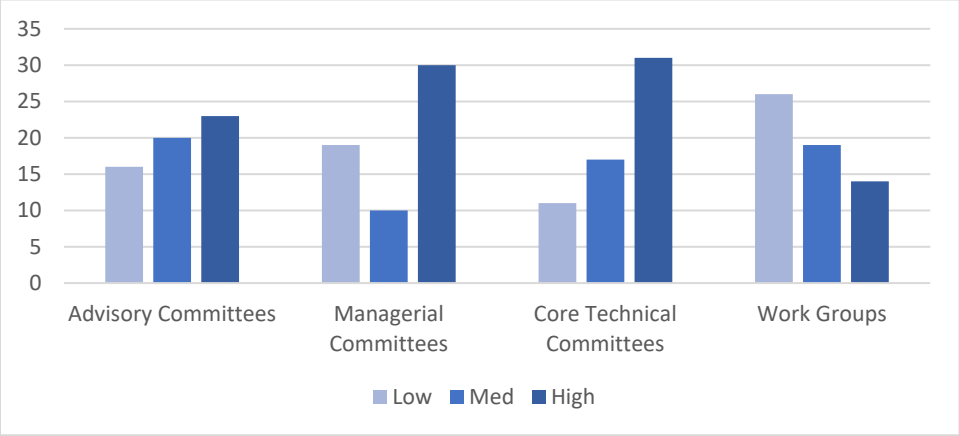


Figure 98: Level of Participation by Committee Category (Number of Respondents)

We had also hoped to create a single measure of formal participation in the Chesapeake Bay policy process, but this proved difficult as the correlation among the categories was not high (see Table 6). In other words, most respondents had high levels of participation in one or two categories, but medium or low levels in the other categories, so there were few who were either high or low participation overall. In fact, when we calculated participation scores across all committees, more than half fell into the "medium" category (see Figure 99). A histogram of the quantitative overall participation scores evinces a similar near-normal distribution, with a high concentration of respondents in the 2.5-3.1 range. In retrospect, this finding is not surprising given our sampling procedures and our primary focus on the formal policy process as governed by the CBP.

Table 6: Correlation Coefficients for Participation by Committee Type

	ADVISE	*MNGMNT*	*TEC1*	*TECH2*
ADVISE	1			
MNGMNT	0.2208	1		
TEC1	0.007424	0.299229	1	
TECH2	-0.08862	0.096157	0.355989	1

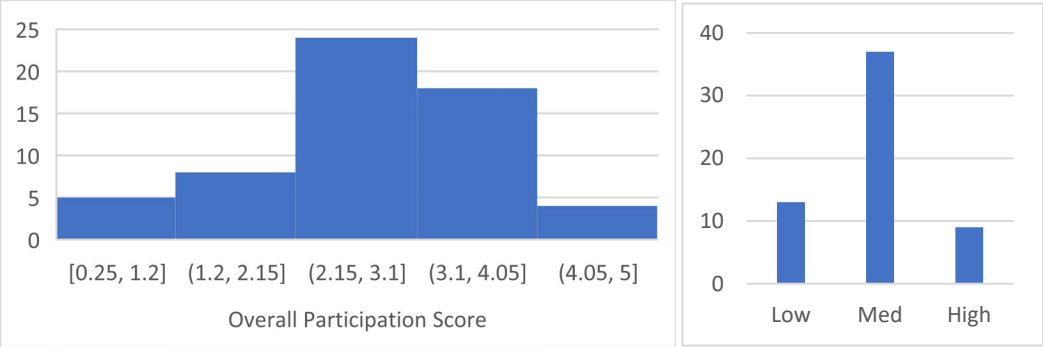


Figure 99: Overall Participation Histogram (left) and Categorical (right)

Finally, we also asked respondents to briefly describe the governmental and non-governmental entities that they interacted with as part of their job (aside from the CBP). As shown in Figure 100, most respondents reported at least some work-related interaction with state and federal bureaucratic agencies like the EPA or their states agricultural agency. Interactions with environmental NGOs was also high, but fewer respondents reported interactions with other types of private-sector organizations. In both cases, interactions with private organizations could include listening to their representatives, answering their questions, or serving with them on committees. Reports of interaction with legislative bodies were moderate, with more interactions at the state than the federal level. While some respondents had formally testified to legislative bodies or served on committees, most had less formal roles providing information to legislators. A few respondents reported similar interactions with the executive branch at the state level, notably a governor's office and related advisory committees.

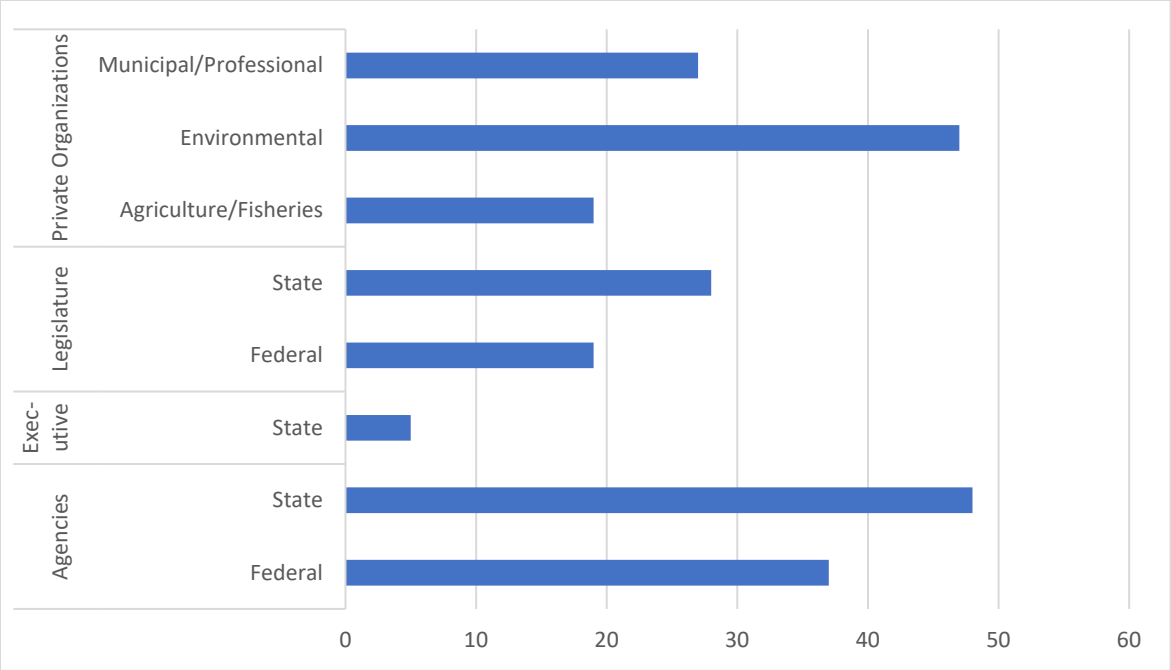


Figure 100: Areas of Interaction Outside of the CBP (Number of Respondents)

Appendix E: Processes and other General Variables

This section covers two types of codes: Block Identifiers and Basic Codes. As explained in Appendix A: Detailed Methods, when a "block of text" is mentioned in this report, it refers to all of the text in an interview that discussed a specific process, for a given time-period and (for WIP Design and Implementation) a state-level jurisdiction. Thus, there are three sets of variables that we used for Block Identifiers: Process Codes, Timeline Codes, and Jurisdiction Codes. Descriptive statistics about these Block Identifiers help us to control for variation in the number of blocks of text per process (e.g., WIP Design has many more blocks of text than other processes) and to make sure we are comparing the same time periods (e.g., Phases of the TMDL or the Model). Jurisdiction Codes were also very important for comparing WIP Design and Implementation processes across states. In contrast to Block Identifiers, Basic Codes, were less central to our analysis but they provide guidance for future research. They tabulate the number of times specific people, places, government agencies, and laws, regulations, or management measures were mentioned. We have already used this information to begin compiling a database of laws, regulations, and management measures for the watershed.

E.1 Block Identifiers

Block identifiers were used to organize our coding schema and will be useful in the analysis of other variables. They include process codes, timeline codes, and state-level jurisdiction codes. While we report the number of blocks of text per code throughout this section, this information should not be regarded as an indication of importance per se, since blocks of text were also delineated based on the questions in the survey instrument. When information was provided that did not match the question, it would be moved to the pertinent block ex post. For instance, if a respondent gave us information about WIP Implementation (Section 4) when we asked about WIP Design (Section 3), this text would be greyed out in Section 4, copied, and then pasted and highlighted under the relevant question in Section 4. Greyed

out text would not be coded. This created consistent blocks of text by question, process, timeline, and jurisdiction for Sections 2-5. None of these codes were used for respondent background (Section 1).

E.1.1 Process Codes

Process codes indicated the main component of the policy process that the respondent was describing in a given block of text. Based on our reading of the literature and of government documents, we originally focused on five main processes: 1) setting the total maximum daily load (TMDL) for the watershed, 2) allocating the TMDL loading targets among the state-level jurisdictions, 3) allocating state-level targets to jurisdictions within each state, 4) designing watershed implementation plans (WIPs), and 5) implementing those plans, which includes state and federal monitoring and enforcing the placement of best management practices and other policy measures contained in the WIPs. If these processes were only mentioned briefly in blocks of text about other processes, they were tagged with LRM codes as discussed in the section on Basic Codes.

After the first few interviews, it became clear that the majority of our respondents did not view these as separate processes, nor did they perceive the overall process as a linear movement from one stage to another (repeating with each new phase, see temporal codes below). Instead, respondents generally viewed setting the TMDL and allocating targets among state-level jurisdictions as a combined process that was highly iterative with WIP design. Similarly, WIP design and allocation of targets, goals, or responsibility among sub-state jurisdictions was also viewed as an iterative process—involving different actors than the TMDL/Allocation process. This combined WIP Design/Allocation process differed substantially across jurisdictions. There was also some overlap between implementation and WIP design, though this was less consistent. Modeling, or developing and parameterizing the various computer models used in setting the TMDL and assigning credits to state-level jurisdiction was also seen as an important, integral part of the policy process. Obtaining funding to implement BMPs was also a key sub-component of both the WIP design and WIP implementation processes. A few respondents also mentioned four other processes: climate change adjustments, the Conowingo Dam WIP, and lawsuits (against the EPA or state governments) and nutrient trading programs. Lastly, many respondents made statements about the functioning of the system as a whole, so we created this as a separate system process code as well.

Figure 101 shows the breakdown of blocks and annotations by the process to which they apply. We do not read too much into these differences because the number of blocks is at least partially determined by the number of questions answered in each section. However, there are some interesting observations. First, a large number of blocks (129) discuss the setting of the watershed-wide loading targets (TMDL) in combination with allocation of state-level loading targets compared to only 18 that directly address the setting of the TMDL alone. This is particularly interesting given that our questions separate out these two processes, but, when asked only about the setting of the TMDL, most respondents answered about the TMDL and state-level allocation together, rather than only answering about the TMDL and then discussing allocation when prompted by the next question in Section 2. WIP design and local allocation processes grab the lion's share of attention in our results with 267 blocks of text. Here again, we asked about WIP design and allocation of loading targets to sub-state-level jurisdictions separately but, in every case, respondents answered about the two together, so there is no "Design Only" category. We did distinguish between comments about WIP Design and Sub-state Allocation when coding, but respondents clearly addressed both together.

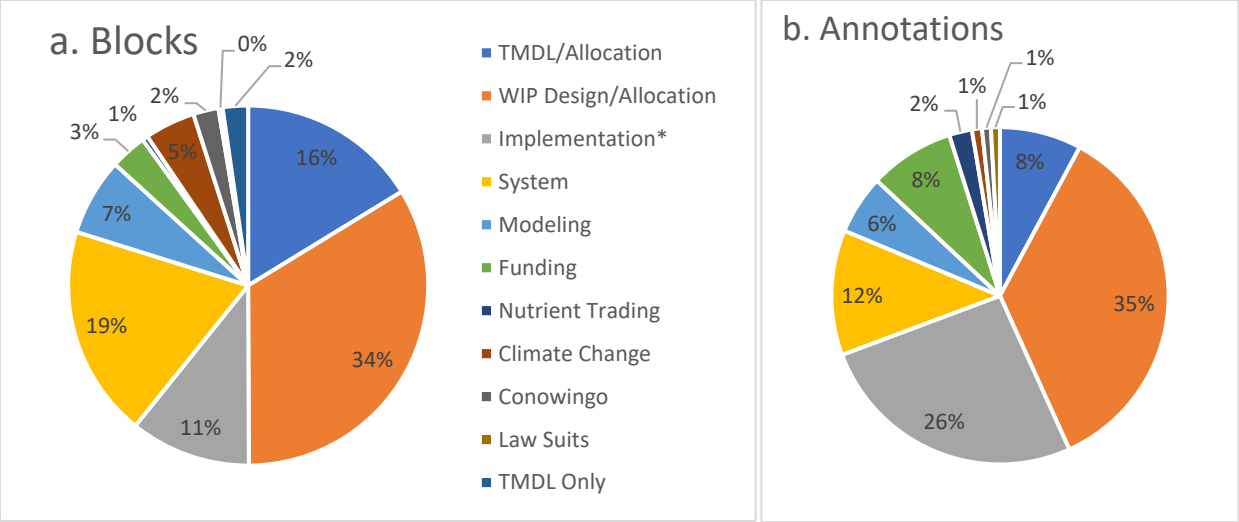


*Figure 101: Process Codes by Number of Blocks (a) and Number of Annotations (b) (*Includes Enforcement)*

The WIP implementation process only has 86 blocks of text. This is likely a reflection of both our respondent's areas of expertise (see Section D.2 Experience and Expertise) and respondent fatigue or lack of time to answer all of the questions in the interview. That is, while all respondents were asked all of the questions in Sections 1 and 2 (background and TMDL respectively), answers to the questions in Sections 1-3 often left little time for questions in Sections 4 and 5 (WIP implementation and effectiveness/challenges). Exceptions were respondents whose primary area of expertise was in implementation. They usually took much less time with Sections 2 and 3, leaving more time for questions on implementation in Section 4. Both system and modeling processes could occur anywhere in the interview, though we did specifically ask about use of CAST in Section 3 (only as it was used in the WIP Design Process, which does not count as Modeling Process) and most answers in Section 5 on effectiveness and challenges were coded as system-level processes. This helps to explain why a larger proportion of annotations discuss annotations compared to the breakdown by blocks of text. Similarly, funding as a separate sub-process usually occurred in Section 4 on implementation but would sometimes be mentioned in other sections. Climate change, Conowingo Dam, nutrient trading and lawsuits were seldom discussed specifically in blocks of text, but we find them analytically interesting and so highlighted them when they occurred. More analysis of these processes can be found in later sections of the report.

E.1.2 Timeline Codes

Due to time constraints, most of the questions in our interview protocol focused on Phase III of the CBP TMDL policy process, but a number of our respondents had experience with earlier phases, and we wanted to take advantage of this knowledge. Therefore, we did not restrict answers and instead coded blocks based on what portion of the policy timeline they referred to as well as which process was discussed. As shown in Figure 102, a small portion of the blocks referred to the history of policy making including the Chesapeake Bay 2000 Agreement and earlier events. Some people also discussed the run up to the current TMDL-based system (Pre-Phase I). Then there was Phase I of the Watershed Implementation Plans, which were approved in 2010 and remained active until Phase II WIPs were approved in 2012. State-level jurisdictions continued to implement their Phase II WIPs until the Phase III WIPs were approved in 2018. Clearly, the majority of time-specific text blocks focus on Phase III, but

some of the other phases and the time between phases were also mentioned. Additionally, separate discussions and agreements related to challenges posed by climate change and increased nutrient flows from Conowingo Dam were dealt with in 2017 and were given their own codes. Since the Phase III WIP will expire in 2025, any predictions about the future of governance were coded as Post-2025. Lastly, since we primarily focused on phases in the TMDL and WIP Design sections, there are also a number of blocks that do not have any specific time associated or, in the case of some system-process blocks, refer to multiple time periods.

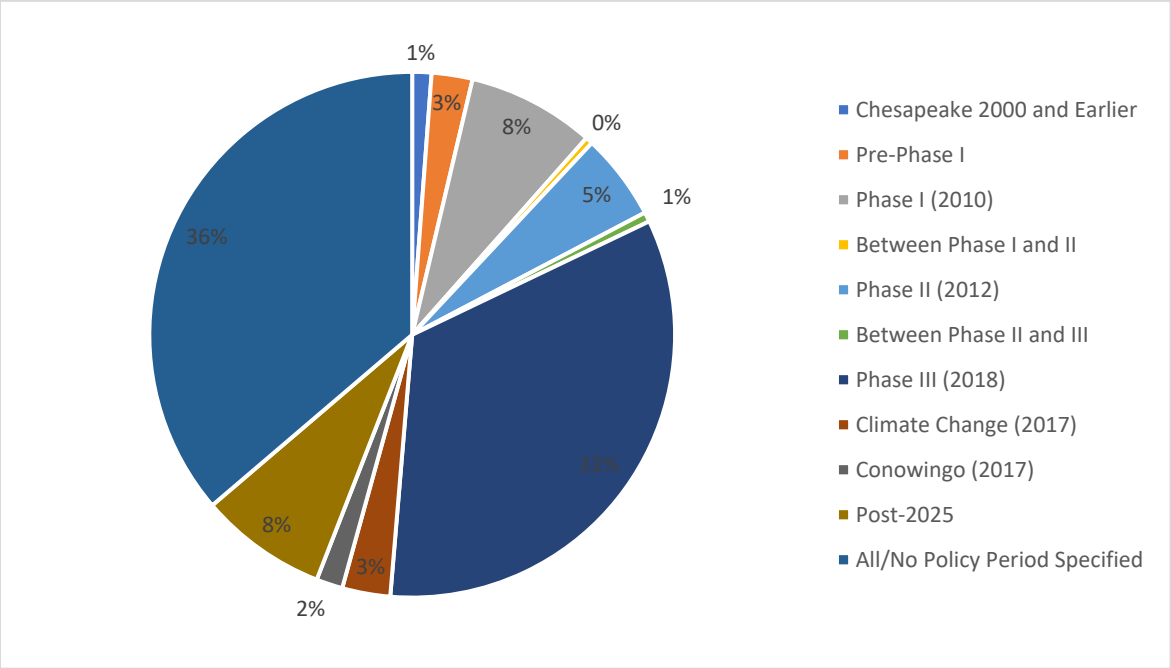


Figure 102: Breakdown of Time Periods by Text Block

In addition to the policy timeline, there is also a modeling timeline that came up in 44 blocks of text. These were volunteered by respondents and so the list is not comprehensive, and it is not always easy to match respondent intention to a specific time period (See Figure 103). In the 1980s, a suite of coupled models was developed for the Chesapeake policy process, with the first used in establishing voluntary loading goals as early as 1987. Among these "Bay Models", the Watershed Model was most important for allocating loading goals. The Watershed Model went through different "phases" as new versions were added, parameters changed, etc. To avoid confusion, we use Roman numerals to indicate a phase in the policy process and Arabic numerals for phases in the Watershed Model. The Phase 5.3 Watershed Model was used to set the TMDL for the watershed in 2010 and to develop the Phase I WIPs in the same year. The Phase II WIPs were developed using the Phase 5.3.2 Watershed Model, but states also had access to the Chesapeake Approximation and Scenario Tool (also called CAST), which was designed as a more user-friendly system that all jurisdictions could use to estimate the effect that implementing different sets (or decks) of best management practices would have on their loads. In Phase II, states used this tool to develop several initial decks, which were sent to the Chesapeake Bay Program Office (CBPO) to run through the Watershed Model v5.3.2 to get official estimates. This process would be repeated a few times to develop the final Phase II WIPs (Linker et al. 2002; Lim 2021).

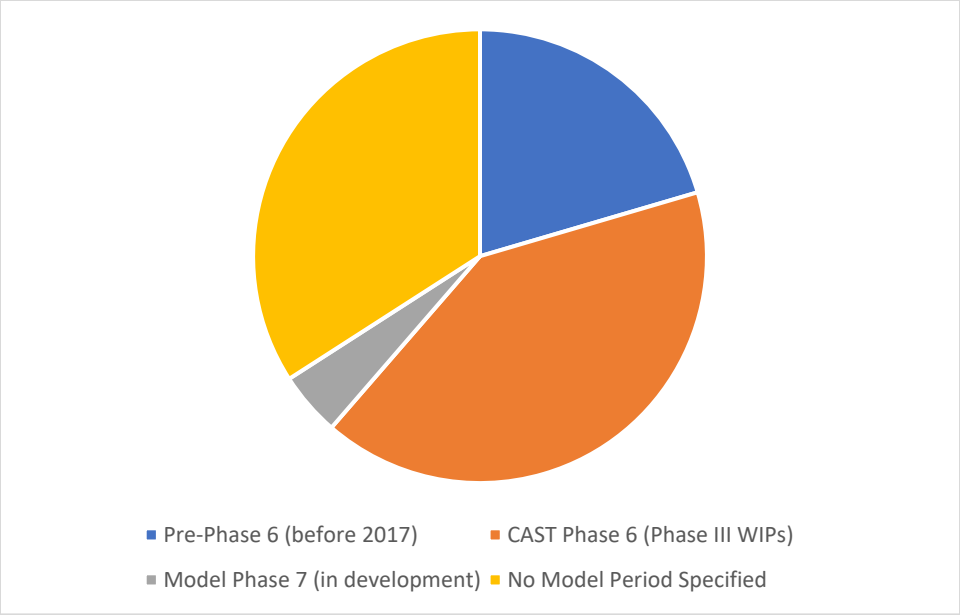


Figure 103: Breakdown of the Modeling Timeline by Number of Blocks of Text

For Phase III WIPs, the CBPO developed a new simplified watershed model that could be used directly by the partnership without a separate scenario tool. Since the interface was similar to the existing scenario tool CAST, the Phase 6 version of the Watershed Model was also called CAST, with the "A" in "CAST" changed from "Approximation" to "Assessment". Many respondents use the term "Phase 6 of CAST" instead of "Phase 6 of the Watershed Model", which is somewhat confusing, as CAST itself has not gone through 6 phases per se. Extensive testing by the CBPO showed that CAST or the Phase 6 of the Watershed Model could predict loading, as well as the more complicated models used in previous phases. The model is updated every 2 years, with the current version in use labeled CAST 2019. States can now use CAST to develop their WIPs without sending decks to the CBPO to estimate loading changes.

Since most of our questions focused on Phase III WIPs, most of the comments about the model also focused on the Phase 6 version of the Watershed Model or CAST. As noted above, respondents also referred to the "Bay Models" frequently. Where possible, we translated this into a given phase of the Watershed Model based on the time frame that they were talking about or lumped these annotations in with the "Any/All" model phases category. A few respondents also considered modeling in the future, including the Phase 7 version of CAST which is expected to be ready for use by 2027. Much like the policy timeline, there were also a number of blocks about the modeling process or use of the models that were not associated with any specific version of the models.

E.1.3 Jurisdiction Codes

In the sections on WIP Design and WIP Implementation, blocks were also designated based on the state-level jurisdiction described (see Figure 104). However, if only a brief comment was made about a jurisdiction anywhere in the interview document, it was treated as an annotation and given an _Sta prefix. These will be described later in the section on annotations. Respondents representing a particular jurisdiction focused on describing their jurisdiction's practices, but respondents from federal agencies or who fit in the independent/multi-jurisdictional group usually talked at length about design or implementation in more than one jurisdiction.

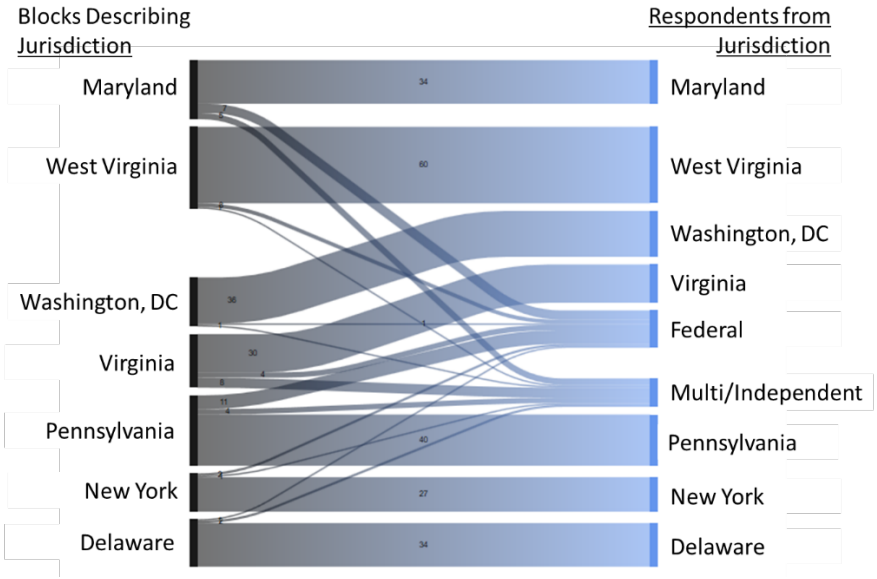


Figure 104: Sankey Diagram of Jurisdiction by Text Block and Jurisdiction of Respondent

West Virginia stands out as having the highest number of blocks and the highest number of blocks/respondent from their jurisdiction (12.8), but this could reflect a number of factors, including the length of answers per question (which usually limited the number of questions asked and the number of text blocks identified), the area of expertise of the respondents (design and implementation sections had the highest number of questions, so individuals with expertise in these areas would provide more blocks of text with jurisdictional codes), amount of time spent in the interview (most respondents went for a full 60 minutes, but some stayed on for up to 90 minutes), and various other factors that are exogenous to the study method. Nevertheless, we report these numbers here as a reminder that the amount of information about each jurisdiction may vary and that, while the total sample is quite large (59), the sample size per jurisdiction is small (5-9) and so analysis based on jurisdiction should be interpreted cautiously.

E.2 Basic Codes

In preliminary rounds of coding, we created codes for some basic variables that might be of interest in a particular item re-occurred and which will be useful when we compare our interview results with other sources of information, such as CBP meeting minutes and news reports. As shown in Table 7, 33 specific locations were mentioned 63 times in different capacities. Of these, only the James River (6) and Lancaster, PA (5) were mentioned 5 or more times. Keep in mind that "mentioned" does not mean how many people talked about the location, but how many times at least one respondent talked about it. Similarly, most individuals were only mentioned once, though Barak Obama (5) and Donald Trump (7) were mentioned more than five times. Gary Shenk (6), Rich Batiuk (6), and Olivia Devereux (5) were all mentioned more than 5 times for their work with the CBP and/or state-level jurisdictions. Among private sector organizations, the Chesapeake Bay Foundation (7), National Fish and Wildlife Foundation (11), farm lobby (7), TetraTech (8), and the Upper Susquehanna Coalition (17), were the only groups mentioned more than 5 times.

Table 7: Basic Variables (#C = Number of distinct codes, #M = total number of "mentions" or times one of the codes in this category was tagged with a block of text

Variable	#C	#M	Description

Location	33	63	Any specific geographic entity mentioned by a respondent as important in the policy process (e.g., “Lancaster, PA”, “Marietta Station”, “the Delaware River”)
Individuals	42	84	Any individuals mentioned specifically by respondents because they were perceived to be important in the policy process. Mostly these were politicians (US presidents, state governors).
Organizations/ Private Entities	32	106	Any private (non-government) entity mentioned by respondents. Includes NGOs, industry interest groups, corporations, municipal groups, universities, news sources, and think tanks.
Local Government Entities	11	58	Local government bodies mentioned by respondents. Includes counties, planning commissions/boards, conservation districts, etc.
State-Level Entities	10	67	Generic classification of state-level agencies and entities mentioned by respondents (as opposed to agencies employing respondents). Since each jurisdiction has its own specific nomenclature, these were loosely grouped (e.g., environment, agriculture, education, transportation, water). Governors’ offices were also included in this category.
Federal Entities	15	170	Any federal agency, department, or organization mentioned by respondents.
CBP Committees	13	82	Any committee, workgroup, or similar entity organized under the rubric of the CBP.
Laws, Regulations, and Management Measures	125	392	Local, state, or federal laws, regulations, or management measures related to water quality in the CBW. Includes government outreach and education programs as well as grant and funding opportunities

Switching to government entities, for local decision-making, Conservation Districts (22), County Governments (6), Planning District Commissions (10), and Soil and Water Districts (5) were mentioned at least 5 times. State-level agencies were mentioned a bit more often (10 agency codes over 67 blocks total). As might be expected given their central role in all of the policy processes covered in the interview, environmental agencies were mentioned most frequently (m = 27). Agriculture agencies were mentioned 17 times, while conservation agencies (7), water boards/agencies (5), and governor’s offices (6) were also mentioned at least 5 times. As shown in Figure 18, federal level entities were mentioned more often. Unsurprisingly, both the EPA (66) and the CBP (51) were mentioned more frequently, with fewer mentions for other agencies. It is interesting that more than 5 respondents distinguished between “the EPA” and “EPA Region 3”, and that while the USDA was mentioned 10 times, its subsidiary agency, the Natural Resources Conservation Service was mentioned 15 times. The National Environmental Information Exchange Network (NEIEN) was mentioned 6 times. States use the NEIEN system to report their verified BMPs. Within the CBP, the Principals’ Staff Committee (21) was mentioned most frequently, followed by the Water Quality Goal Implementation Team (13), with STAC (8), the Executive Committee (5), and all Expert Panels (7) also receiving enough mentions for us to list them in public reports.

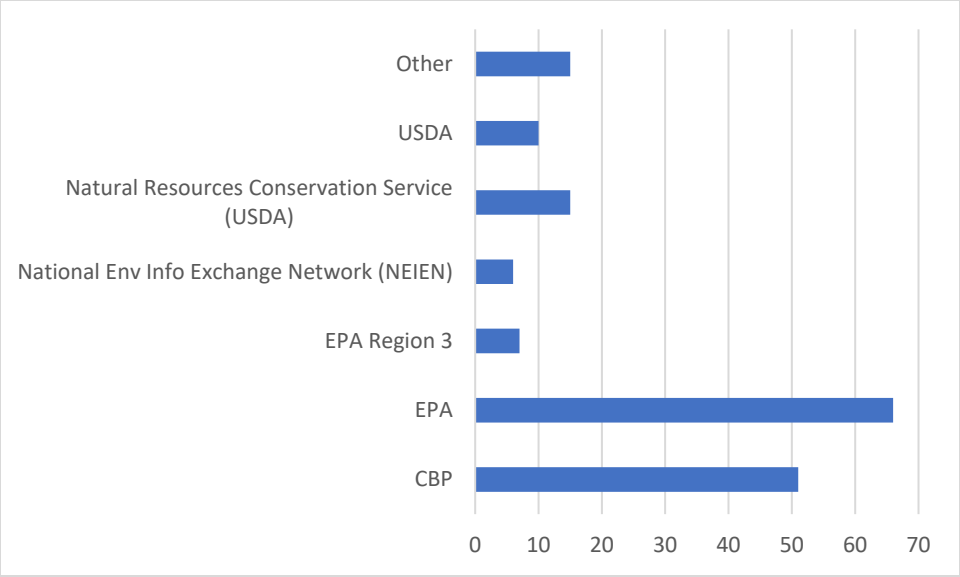


Figure 105: Federal Level Policies Mentioned by Respondents

Lastly, laws, regulations, and other policy measures (LRM) mentioned by respondents varied widely, as summarized in Figure 106. 19 state and 2 local programs for supporting BMP implementation through funding or outreach were mentioned fewer than 5 times, so they are aggregated into the Other State Funding/Outreach category (m = 34). However, several states have "Flush Taxes" (m = 5) that were implemented to raise funds for wastewater treatment and sewage programs, so that policy was mentioned more often. Respondents also mentioned 16 state-level laws and regulations (m = 21) and 5 state-level fundraising mechanisms in addition to the flush tax (m = 6). Federal-level policies that are not within the jurisdiction of the CBP were also mentioned. Of these, the Clean Water Act (m = 12) was mentioned most often, followed by the USDA's Conservation Reserve Enhancement Program (CREP; 6) and President Barak Obama's 2009 Executive Order (m = 5) which spurred the creation of the TMDL and related WIPs. Otherwise, 4 federal funding and outreach programs were mentioned 5 times total. All were programs run by the USDA. 10 other federal laws or regulations were mentioned a total of 19 times.

Although relatively few blocks addressed lawsuits as a separate process (i.e., explained how and why lawsuits were brought against the federal or state governments), lawsuits were mentioned 22 times total, indicating that respondents saw them as part of the other policy processes as well as processes on their own. The majority of the blocks in this category referred to litigation designed to pressure the EPA to take steps to improve water quality in the Bay but some referred to suits contesting the TMDL. These lawsuits blur the line between levels of analysis because participants included private groups/NGOs, state governments, and the federal government. Although NGO efforts to assist w/ regulation were not mentioned as often, 12 NGO-run funding and support programs were mentioned a total of 14 times. Although they are not officially "regulatory" these programs still can be considered part of the governance system, which includes private as well as public efforts that shape stakeholder behavior.

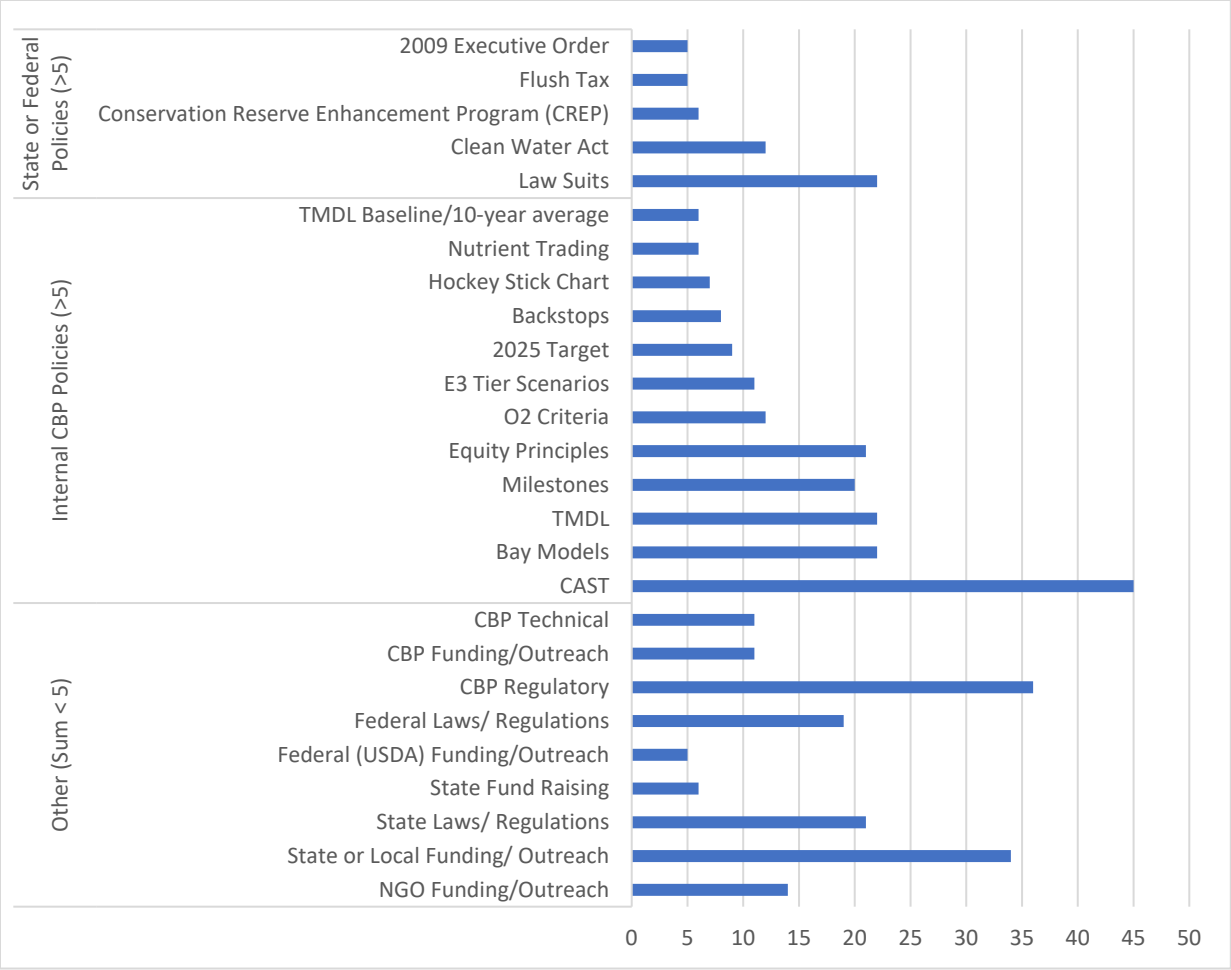


Figure 106: Laws, Regulations, and Management Measures Mentioned by Respondents (n >5)

All other laws, regulations, and management measures mentioned by respondents were established under the auspices of the CBP. Given the discussion of the modeling timeline, it might not be surprising that "CAST" (m = 45) or the more general "Bay Models" (m = 22) were mentioned in more blocks than any other LRM. Again, we did ask specifically about the use of CAST in the WIP design process, but mentions were found in other sections as well. Modeling-related policies were also mentioned frequently. These included policies that affected the setting of the watershed-wide targets to attain the 2025 Target (m = 9) for the TMDL (i.e., selection of the 10-year time period over which loads were averaged to provide a baseline (m = 6) and setting of the oxygen (O2) minimum criteria (m = 12) as a major determinant of water quality). Policies that affected allocation of loading targets among state level jurisdictions included the "Hockey Stick Chart" (m = 7) and "E3 Tier Scenarios" (m = 11) that were used to implement the "Equity Principles" (m = 21) that CBP members agreed to as their allocation criteria. 9 other technical or model-related policies were mentioned a total of 11 times during interview segments that were not designated with modeling process codes.

Other policies that fit into the "regulatory" sub-category included the TMDL itself, which was mentioned 22 times when respondents were discussing other components of the policy process. Nutrient trading was mentioned 6 times outside of the nutrient trading-specific blocks described under General Codes. The "Equity Principles" (m = 21) mentioned above could also be considered regulatory. In addition,

monitoring via the two-year Milestone Assessments was mentioned 20 times, and the EPA's option to punish jurisdictions that do not make their loading requirement using "Backstops" was mentioned 8 times. 19 other CBP regulatory decisions were mentioned a total of 36 times, making this the largest "Other LRM" category. Lastly, 7 CBP funding and outreach programs were mentioned in 11 blocks of text. It is interesting that respondents tended to focus more on funding and outreach programs at the state level while describing regulatory policies more both for the CBP and non-CBP federal entities. This may reflect the preponderance of state-level expertise in the sample as well as the diversity of programs used by states to help implement their WIPs.

Appendix F: List of Tables

Appendix G: List of Figures