Project Report

Evaluation of Phase II of the Medicare Advantage Value-Based Insurance Design Model Test

Appendixes

Dmitry Khodyakov, Christine Eibner, Erin Audrey Taylor, Rebecca Anhang Price, Christine Buttorff, Matthew Cefalu, Brian G. Vegetabile, Julia Bandini, Monique Martineau, Catherine C. Cohen, Michael Dworsky, Marika Booth, Alice Y. Kim, Julie Lai, Shiyuan Zhang, Afshin Rastegar, Stephanie Dellva, Nabeel Qureshi, Priya Gandhi, Courtney Armstrong, Daniel Schwam, Natalie Ernecoff, Anagha Alka Tolpadi

RAND Health Care

PR-A1881-1
October 2022

Prepared for the Centers for Medicare & Medicaid Services
Center for Medicare & Medicaid Innovation
Under Research, Measurement, Assessment, Design, and Analysis Contract Number 75FCMC19D0093, Order Number 75FCMC20F0001

About This Appendix

This Appendix supplements the Evaluation of Phase II of the Medicare Advantage (MA) Value-Based Insurance Design (VBID) Model Test, initiated by the Center for Medicare & Medicaid Innovation (CMMI), for the years 2020 and 2021. The evaluation was funded by CMMI under Research, Measurement, Assessment, Design, and Analysis Contract Number 75FCMC19D0093, Order Number 75FCMC20F0001, for which Julia Driessen is the contracting officer's representative. It was carried out within the Payment, Cost, and Coverage Program in RAND Health Care.

RAND Health Care, a division of the RAND Corporation, promotes healthier societies by improving health care systems in the United States and other countries. We do this by providing health care decisionmakers, practitioners, and consumers with actionable, rigorous, objective evidence to support their most complex decisions. For more information, see www.rand.org/health-care, or contact

RAND Health Care Communications
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138
(310) 393-0411, ext. 7775
RAND_Health-Care@rand.org

Contents

Figures and Tables

Figures

Tables

Appendix A. Research Questions

This Appendix lists research questions, specifies outcome types, and identifies analytic approaches used to answer research questions in each report chapter.

Table A.1. Research Questions, Outcomes, and Approaches Grouped by Chapter

Chapter	Research Questions	Task Order Question Wording	Outcomes	Approaches
2	• How are VBID plans and participating POs different from nonparticipants? • How do their plan characteristics differ?	How are the participating plans different from MA plans in general in terms of benefit packages, premiums, and other characteristics? (TORP 5)	Participation	Descriptive, Thematic
		Are participating Medicare Advantage Organizations (MAOs) different from non-participating and those that leave? (TORP 8.2)		
	• Why did POs choose to participate—or not?	What are the primary reasons that plans chose to participate in VBID? (TORP 8.1)	Participation	Thematic
		What are the primary reasons plans chose not to participate in VBID? (TORP 8.1a)	Participation	Thematic
	• What processes and staff were involved in these decisions?	What was the decisionmaking process in the plan and what areas and levels of plan management were involved? (TORP 8.1b)	Participation	Thematic
3	• What interventions did POs implement, and what groups of beneficiaries did they target?	N/A	Participation	Descriptive, Thematic
	• Why did POs choose these interventions?	N/A	Participation	Descriptive, Thematic
4	• What are POs' and vendors' implementation experiences with the BDI component?	What are the implementation experiences plans face in operationalizing their interventions, communicating and engaging enrollees initially and on an ongoing basis, tracking uptake, modifying operating processes and IT systems, etc.? (TORP 9.1)	Implementation	Thematic
		What are [vendors'] experiences with the VBID Model? (TORP 10.3)	Experiences	Thematic
	• Do these experiences vary by intervention?	How did these implementation experiences vary by intervention and component? (TORP 9.1a)	Implementation	Thematic
		Do they [vendors' experiences] vary by intervention? (TORP 10.3a)	Experiences	Thematic
	• How did the coronavirus disease 2019 (COVID-19) pandemic affect the BDI component implementation?	N/A	Participation	Descriptive

Chapter	Research Questions	Task Order Question Wording	Outcomes	Approaches
5	• What proportion of plan enrollees are eligible for VBID and receive benefits? How does this change over time?	Do participating plans enroll more or fewer enrollees over the course of the model test, and why? (TORP 1.1)	Enrollment & Eligibility	Descriptive, Thematic, DD
	• Are targeted beneficiaries electing to participate in VBID?	What is the proportion of enrollees eligible for VBID under this model test, and what is the proportion of those eligible beneficiaries actually receiving VBID benefits? (TORP 4)	Enrollment & Eligibility	Descriptive
		What is the VBID uptake among targeted beneficiaries? (TORP 10.2)	Experiences	Descriptive, Thematic, DD
	• How does the model test affect plan bids for Parts C and D? What variables factor into bid changes?	What is the model's effect on plans' bids, for Parts C and D? (TORP 3.2)	Cost	DD
	• How does the model affect premiums and supplemental benefits?	What is the model's impact (if any) on targeted enrollees' and non-targeted enrollees' premiums, and the availability of supplemental benefits for non-targeted enrollees in participating plans? (TORP 3.3)	Cost	DD
		What factors or variables are driving any increases or decreases in plans' bids? (TORP 3.4)	Cost	DD
6	• What palliative care, transitional concurrent care (TCC), and hospice supplemental benefits do participating POs offer as part of the model test?	N/A	Participation	Descriptive, Thematic
	• Why did hospices join VBID PO's networks?	N/A	Participation	Descriptive, Thematic
	• How are networks of hospices being built, and what do they look like?	How do plans identify in-network hospices, and what do networks look like? (TORP 14.5)	Participation	Descriptive, Thematic
	• How are payment arrangements being handled?	What payment arrangements are used? (TORP 14.5a)	Participation	Thematic
7	• What did POs need to do to implement the Hospice component into their plans?	What implementation and operational adaptations are needed by plans to participate in the model for hospice? (TORP 9.2)	Implementation	Thematic
	• Do in-network hospices need to operate differently under VBID?	How do in-network hospices deliver care differently, relative their standard care delivery? (TORP 14.6)	Implementation	Thematic
	• How do in- and out-of-network (OON) hospices perceive the Hospice component of the model test?	What are hospices' perspectives on the hospice benefit component, both among those deemed in-network by at least one participating plan, and those who are out of network? What are their perceptions as to how it affected end-of-life care? What changes would they like	Implementation	Thematic

Chapter	Research Questions	Task Order Question Wording	Outcomes	Approaches
		to see in the model, and/or are there any changes or adjustments they intend to make in terms of how they engage with the model? What do hospices perceive as the benefits and drawbacks of engaging with plans as part of the Model, both anticipated and unanticipated? (TORP 14.7)		
		How are hospices affected by the hospice benefit component of the model, both among those with in-network status and those without? How does the model affect their census level and case mix? What are the significant implementation and operational adaptations needed to accommodate the hospice benefit component of the model? What do hospices perceive as the benefits and drawbacks of engaging with plans as part of the hospice benefit component of the model, both anticipated and unanticipated? (TORP 14.8)	Implementation	Descriptive, Thematic
	• How did the COVID-19 pandemic affect the Hospice component implementation?	N/A	Implementation	Thematic
8	• How does enrollment in participating plans change over time? Why?	Do participating plans enroll more or fewer enrollees over the course of the model test, and why? (TORP 1.1)	Enrollment & Eligibility	Descriptive, Thematic, DD
	• How does the model test affect plan bids for Parts C and D? What variables factor into bid changes?	What is the model's effect on plans' bids, for Parts C and D? (TORP 3.2)	Cost	DD
	• How does it affect premiums and supplemental benefits?	What is the model's impact (if any) on targeted enrollees' and non-targeted enrollees' premiums, and the availability of supplemental benefits for non-targeted enrollees in participating plans? (TORP 3.3)	Cost	DD
		What factors or variables are driving any increases or decreases in plans' bids? (TORP 3.4)	Cost	DD
9	• How did POs implement the WHP requirement?	What are the implementation experiences plans face in operationalizing their interventions, communicating and engaging enrollees initially and on an ongoing basis, tracking uptake, modifying operating processes and IT systems, etc.? (TORP 9.1)	Implementation	Thematic
	• How did the COVID-19 pandemic affect WHP implementation?	N/A	Implementation	Thematic

NOTE: N/A signifies that a research question was not explicitly included in the task order request for proposal (TORP).

Appendix B. Primary Data Collection and Analysis

This appendix describes the primary data collection and analysis techniques used in this report. All research activities have been reviewed by the RAND Human Subjects Protection Committee, which determined this project to be exempt from additional review.

In 2021–2022, we conducted a series of semistructured interviews with model-participating and nonparticipating parent organizations (POs), vendors, and in-network and OON hospices to help explain their decisions about model participation, explore Value-Based Insurance Design (VBID) implementation experiences, describe how and why VBID implementation was associated key model outcomes, and comprehensively address research questions that focus on the nature, context, implementation, and expected outcomes of various VBID model components.

All interviews were conducted using a similar approach. For each of the five groups noted above, we reached out to contacts at each organization via email and provided them with a brief description of the interview, its purpose, and logistical details and a signed endorsement letter from the Centers for Medicare & Medicaid Services (CMS). We conducted follow-up outreach activities by email and phone with up to three attempts to reach those who had not responded to our invitations.

We used a small group approach to the interviews. We allowed contacts at each organization to invite colleagues most knowledgeable about VBID to participate in the interviews. During the scheduling phase, we sent the consent form via email. We obtained verbal consent and answered any questions prior to beginning the interview.

Each interview was conducted virtually by a team that included up to three researchers and one research assistant who took notes. All but one interview were audio recorded and professionally transcribed. Close-to-verbatim notes were taken during one interview in which PO representatives declined to have their interview recorded.

We provide additional descriptions of our sampling and data collection processes for each of the groups below.

Sampling and Data Collection

Nonparticipating Parent Organizations

In spring 2021, we conducted a series of 45- to 60-minute interviews with nonparticipating POs (NPPOs) to understand their reasons for not joining the model test and their decisionmaking processes. Interviews with NPPOs were designed to help us answer the following research questions:

- Why did POs choose not to participate in VBID?
- What processes and staff were involved in these decisions?
- How are VBID plans and participating POs different from nonparticipants?

We recruited a diverse sample of eligible POs that had decided not to join the VBID model test. Enrollment in Medicare Advantage (MA) plans is heavily concentrated in a few national POs. However, there are a variety of smaller regional and state-based plans delivering MA benefits. Given this tiered structure to the market, our sample purposefully included large national POs as well as regional POs from different parts of the country.

We first applied the model test eligibility criteria to determine the list of eligible but nonparticipating POs (see Appendix D) and excluded all POs that had previously participated in the VBID model test. We classified POs into national, regional, or state on the basis of their service area information, using a threshold of the number of states in which the PO has contracts: state-based POs (up to two states), regional POs (three to eight states), national POs (nine or more states). We classified all states into four regions according to the U.S. Census Bureau's classifications: West, Midwest, Northeast, and South. For POs with operations in multiple states, we grouped them into the region where they have the majority of operations. We then randomly chose nonparticipants from each of these categories of POs until we assembled a diverse sample of ten POs (two national, two regional, and six state based from different regions) that agreed to share their perspectives. We considered a target sample size of ten interviews to be adequate for reaching saturation: "the point in data collection when no additional issues or insights are identified and data begin to repeat so that further data collection is redundant, signifying that an adequate sample size is reached" (Hennink and Kaiser, 2022, p. 2). Previous research shows that a sample size as small as six could be enough to identify main themes (Guest, Bunce, and Johnson, 2006) and that saturation could be achieved after nine interviews (Hennink, Kaiser, and Marconi, 2016; Hennink and Kaiser, 2022). We used email to contact all four national POs and 58 regional and state POs to assemble our sample.

Between March and April 2021, we conducted interviews with 33 representatives from nine POs that did not participate in the VBID model test. Representatives of one PO canceled a scheduled interview. Three additional POs provided feedback via email as to why they chose not to participate in VBID. Of the 12 POs that we interviewed or received written feedback from, two were national, two were regional, and eight were state based (three from the West, two from the Midwest, two from the South, and one from the Northeast).

Interviews with NPPOs followed a semistructured format covering such topics as

- decisionmaking related to model test participation
- awareness of the VBID model
- reasons for not participating in the VBID model test
- thoughts on various model components and their possible impact on plans, providers, and beneficiaries
- suggestions on how to make the VBID model more attractive to MA plans.

Participating Parent Organizations

We administered pre-interview questionnaires and conducted interviews with POs that participated in the model test in 2021 to understand reasons for joining the model test, the process and reasons behind design and implementation decisions, experiences with specific model components, barriers encountered and strategies used to overcome them, and impact of interventions on beneficiary health outcomes and costs. The questionnaires were developed after the review of POs' model test application materials and informed by the results of PO interviews we conducted during Phase I of the VBID model test.

Interviews with participating POs were designed to help us answer the following research questions:

Model Participation

- Why did POs choose to participate—or not? What processes and staff were involved in these decisions?
- How are VBID plans and participating POs different from nonparticipants? How do their plan characteristics differ?

Implementation

- What interventions did POs implement, and what groups of beneficiaries did they target?
- Why did POs choose these interventions?

Benefit Design Innovations Component

- What are POs' and vendors' implementation experiences with the Benefit Design Innovations (BDI) component?
- Do these experiences vary by intervention?
- How did the coronavirus disease 2019 (COVID-19) affect implementation of the BDI component?
- What proportion of plan enrollees are eligible for VBID and receive benefits? How does this change over time?
- Are targeted beneficiaries electing to participate in VBID?

Hospice

- What palliative care, TCC, and hospice supplemental benefits do participating POs offer as part of the model test?
- Why did hospices join VBID POs' networks?
- How are networks of hospices being built, and what do they look like?
- How are payment arrangements being handled?
- What did POs need to do to implement the Hospice component into their plans?

- Do in-network hospices need to operate differently under VBID?
- How do in-network and OON hospices perceive the Hospice component of the model test?
- How does enrollment in participating plans change over time? Why?
- How did COVID-19 affect the Hospice component implementation?

Wellness and Health Care Planning Requirement

- How did POs implement the wellness and health care planning (WHP) requirement?
- How did COVID-19 affect WHP implementation?

We invited all 19 POs that participated in the VBID model test in 2021 to participate in a two-hour semistructured interview. Between June and August 2021, we conducted interviews with 80 representatives from 18 of the 19 POs. One PO did not respond to our interview invitation. POs could choose to complete one two-hour interview or schedule two separate hour-long interviews.

During the interviews, we discussed POs' responses to the pre-interview questionnaires and asked additional open-ended questions covering such topics as

- reasons for implementing or not implementing different VBID model components and subcomponents
- implementation experiences, successes, and challenges
- WHP activities
- impact of COVID-19 on the model implementation
- intervention uptake among beneficiaries.

By design, this round of PO interviews was not focused on POs' perspectives on the model test outcomes or the extent to which the observed outcomes met their expectations and their thoughts about the future impact of the model; these perspectives were captured in POs' responses to the pre-interview questions and did surface in a small number of interviews that had time for such discussion. We plan to structure future interviews so that there is sufficient time to focus on POs' perceptions of the achieved and expected outcomes.

Vendors

During participating PO interviews, we learned that vendors played an important role in helping POs implement their VBID interventions. As a result, we refined our evaluation design to include interviews with vendors that provided either health-related or non-health-related services to beneficiaries as part of the VBID model test. Interviews with vendors were designed to answer the following research questions:

- What are providers' (vendors') experiences with the VBID model? Do they vary by intervention?
- What strategies are successful in implementing WHP?

We asked 10 POs to provide us with the names and contact information of 19 third-party vendors that were mentioned during the interviews. We received contact information for 16 vendors from nine POs and invited all of them to participate in an interview.

In fall 2021, we interviewed 17 representatives from ten vendors offering services to seven POs. This sample size enabled us to reach thematic saturation, typically reached between the sixth and twelfth interview (Guest, Bunce, and Johnson, 2006). Among the ten organizations interviewed, five were vendors of primarily health-related services and benefits, such as a provider of fall risk assessments, and five were vendors of non–primarily health-related services, such as a nonemergency medical transportation broker.

Vendor interviews lasted 45–60 minutes and followed a semistructured format covering such topics as

- description of services provided as part of the MA VBID model test
- VBID implementation experiences
- challenges and facilitators associated with the delivery of VBID services
- perceived impact of the interventions.

In- and Out-of-Network Hospices

We conducted interviews with the representatives of hospices that were and were not part of the hospice networks of POs that implemented the Hospice component of the model test. We wanted to understand their perspectives on MA hospice integration and the likelihood of their future participation in VBID. Interviews with in-network and OON hospices were designed to help us answer the following research questions:

In-Network Hospices Only

- What are the significant implementation and operational adaptations needed by plans to participate in the model for hospice?
- How does this model affect the way hospice care is introduced to MA enrollees deemed potentially eligible, including the timing of the initial discussion of hospice and the approach to introducing the topic and options?
- In what ways do hospices designated as in-network by plans participating in the Hospice benefit component deliver hospice care differently to beneficiaries in the model, relative to the hospice program's standard care delivery approach?
- What payment and delivery innovations emerged as a result of the Hospice benefit component of the model, in terms of how participating plans addressed serious illness and end-of-life care, and as seen in the arrangements between plans and hospices?

Out-of-Network Hospices Only

- Are there any unanticipated effects of the model?

Both In- and Out-of-Network Hospices

- What are hospices' perspectives on the Hospice benefit component, both those deemed in-network by at least one participating plan and those that are OON? What are their perceptions as to how it affected end-of-life care? What changes would they like to see in the model, and/or are there any changes or adjustments they intend to make in terms of how they engage with the model? What do hospices perceive as the benefits and drawbacks of engaging with plans as part of the model, both anticipated and unanticipated?
- How are hospices affected by the Hospice benefit component of the model, both those with in-network status and those without? How does the model affect their census level and case mix? What are the significant implementation and operational adaptations needed to accommodate the Hospice benefit component of the model? What do hospices perceive as the benefits and drawbacks of engaging with plans as part of the Hospice benefit component of the model, both anticipated and unanticipated?

We assembled a diverse sample of in-network and OON hospices to achieve thematic saturation. We used hospice network lists that POs submitted to CMS to identify in-network hospices. OON hospices were identified as those providing care in POs' service areas and having served at least one VBID beneficiary through July 2021. We excluded hospices that provided care to fewer than 50 beneficiaries in the prior year. We sorted both in-network and OON hospices for each PO in descending order by the number of VBID beneficiaries served and sequentially reached out to in-network and OON hospices for each PO. We prioritized large hospice chains and hospices that provided care to beneficiaries from more than one PO participating in the VBID Hospice component, either as in-network or OON.

To assemble a hospice sample of ten in-network and ten OON hospices, we contacted five national hospice chains, 42 OON hospices, and 28 in-network hospices. We were able to schedule 45- to 60-minute interviews with representatives of 23 hospices. Of these 23 hospices, 13 were in-network, six were OON, and four were chains. Three of the four chains had both in-network and OON hospices, and one chain had only OON hospices in the model test. We asked all in-network hospices, including all chains, to complete a short online questionnaire in advance of the scheduled interview to help guide the discussion of implementation challenges, facilitators, and expected outcomes.

Between November 2021 and January 2022 we conducted interviews with 42 representatives of 13 in-network hospices, 13 representatives of four national hospice chains, and 11 representatives of six OON hospices. Using the information received during the interviews, we classified all but one hospice chain as OON providers, meaning that our final sample included 14 in-network hospices and nine OON hospices. Interviewed hospices provided care to beneficiaries from all nine POs participating in the Hospice component of the model test and were located in different parts of the country. Our hospice sample was diverse with regard to ownership status (12 for-profit and nine nonprofit hospices, as well as four listed as having "other" profit status)

and size (one hospice with 50 to 109 Medicare beneficiaries per year, eight with 230 to 510 per year, and 13 hospices with 519 or more per year).

Interviews with in-network and OON hospices followed a semistructured format covering such topics as

- reasons for joining or not joining the hospice network of one or more POs participating in VBID hospice
- the process of negotiating new contracts and working with POs
- implementation experiences, successes, and challenges
- experiences working with the POs as an OON hospice
- changes in care delivery as a result of VBID hospice
- thoughts about model achieved and expected future outcomes, including any unintended outcomes.

Data Analysis

We used descriptive statistics to analyze responses from the pre-interview questionnaire. Given the types of questions included, we used medians as a measure of central tendency and provided frequency distributions to show a range of answers. We used these survey findings to guide our qualitative analysis by identifying barriers, facilitators, and expected outcomes that were most frequently mentioned across organizations.

We used a thematic analysis of interview data to provide more in-depth responses to the research questions. We coded the interview transcripts using Dedoose, a qualitative software program, to facilitate systematic team-based coding. All data analysis processes described below apply to all qualitative data collection activities. Our qualitative analysis process included several steps to ensure the trustworthiness, rigor, and validity of the qualitative data collected (Shenton, 2004), such as discussions of emerging themes throughout the data collection and data analysis process, establishment of an interrater reliability across members of the qualitative coding team, and discussions around the comparison of emerging themes from the prior VBID model test evaluation.

For each group of interviews, we developed an initial codebook based on the interview protocol. We structured the codebook in alignment with important themes from the previous two VBID model test evaluation reports (Eibner et al., 2018; Eibner et al., 2020), including reasons for participating in VBID and implementation challenges encountered, while also providing opportunities for themes to emerge from the data collected in 2021. We created an "other" code for text that did not fall under a given code and subsequently categorized excerpts from the "other" code into newly created codes. Team members discussed the initial codebook, and one team member used this codebook to code the first five transcripts to identify any emerging themes and explore the need for creating new codes. Because of the overlap in the questions we asked across interview types (e.g., participating and nonparticipating POs, vendors, and hospice

interviews), we developed a similar codebook for each type of interview, using similar codes wherever possible. Because of these similarities across codebooks, we calculated a high kappa score of 0.78 using the Dedoose feature for interrater reliability for the PO interviews (McHugh, 2012), which yielded the largest volume of qualitative data. Following the establishment of interrater reliability among the coding team, three team members coded the remainder of the interview transcripts, meeting regularly to review questions and discrepancies and to discuss emerging themes.

We used a joint coding and analytic approach, in which each coder independently coded each interview transcript once coding reliability was established. Coding team members met weekly to discuss any coding questions and reach agreement on how to handle the creation of new codes. This joint coding approach is an acceptable method that we have used in previous studies (Khodyakov et al., 2014; Concannon et al., 2015). After all interviews were coded, we analyzed the interview data thematically to answer relevant research questions. One researcher was assigned to answer a given research question by synthesizing findings across the interviews coded by different coding team members, which helped us reconcile any remaining coding discrepancies between different data coders.

We used thematic analysis techniques (Guest, MacQueen, and Namey, 2012) to compare themes and explore patterns and variation in POs', vendors', and hospices' perspectives on and experiences with the model test. We also compared emerging themes from this evaluation with our previous VBID model test evaluation findings and identified new themes specific to data collected in 2021 as part of this model test. For example, we tracked experiences with the model test longitudinally by examining whether responses to our questions regarding model test participation decisionmaking as well as implementation challenges and facilitators changed over time. Our expectations about the Hospice component implementation and outcomes were primarily based on the limited literature rather than previous evaluation findings because it was a new model test component in 2021. However, in analyzing the hospice responses, we compared them with those provided by POs that implemented the Hospice component. To illustrate, POs and hospices described different primary reasons for participating in VBID Hospice. Many POs mentioned the alignment of the Hospice component with their organizational missions, while many hospices highlighted the importance of being included in PO hospice networks to ensure ongoing financial viability.

In line with the mixed-methods nature of this evaluation, we integrated quantitative and qualitative analytic techniques to ensure the rigor of our findings. When possible, we "quantified" qualitative data to identify the most and least frequently mentioned themes, which may be treated as a marker of relative importance (Buetow, 2010), and used the most salient themes to guide our presentation of findings. We also used the responses from the pre-interview questionnaire to identify the most commonly cited implementation barriers, which in turn guided our analysis of implementation challenges that were raised during the interviews. In addition to triangulating our findings by looking, for example, at the differences between PO and hospice

perspectives on the model test, we looked for concordance between the results of primary and secondary data analysis. To illustrate, in the report chapters that focused on the model test outcomes, we integrated quantitative and qualitative results by using secondary data analysis to provide impact estimates and by relying on primary data to explain *how* and *why* VBID might have affected key outcomes.

Appendix C. Statistical Approach

Our statistical approach uses methods that focus on isolating the short-term impact of VBID in the presence of different patterns of participation, varying configurations of offerings of services within participating plans represented in the BDI component, and participation in the Hospice component of the evaluation. As described in the main text, BDI encompasses a range of intervention types, such as Rewards and Incentives (RI), reduced cost sharing for high-value services, and rebates. While conducting separate analyses at the BDI subcomponent level was out of scope for the current report, we expect to address this issue in a future report. The overall approach is a weighted difference-in-differences (DD) design to identify the effects of VBID participation. Our DD approach aligns most closely with those described in Callaway and Sant'Anna (2021), with modifications to allow for the inclusion of balancing weights and to accommodate the fact that some plans leave the VBID model test before its conclusion.

There are a few challenges to successfully estimating the average impact of VBID participation given the observational nature of the VBID model test. First, plans' fidelity of implementation and beneficiaries' uptake of the proposed intervention may be varied. For this reason, all analyses, unless otherwise noted, will be based on the intention-to-treat principle—that is, plans will be analyzed based on their proposed interventions regardless of fidelity or uptake. This allows us to estimate the effectiveness of VBID participation under real-world implementation of the interventions. We do not estimate the efficacy of the interventions, which would measure the effect of VBID participation under ideal circumstances of perfect fidelity and uptake. Second, as alluded to above, plans that chose to participate in the VBID model may not be representative of all eligible plans on *observable* characteristics (i.e., there may be selection on observables). Failure to account for observable differences between the plans may lead to bias in our analyses, which is often referred to as confounding bias. Next, plans were allowed to join and leave the VBID model test on a year-to-year basis, which we refer to here as "staggered adoption" and "discontinuation," respectively. Finally, plans that chose to participate in the VBID model may differ in *unobservable* ways from those that did not (i.e., selection on unobservables). To address these analytic concerns, this evaluation combines entropy balancing on observables with the DD framework established in Callaway and Sant'Anna (2021), which allows DD designs with differing patterns of participation.

Specifically, our DD framework relies on the estimation of a separate DD model for groups of participating plans, where the groups of plans are defined by their patterns of participation in the VBID model test. The estimates from these separate DD models are combined to provide aggregate summaries of the effect of VBID participation. Each of the models requires the identification of a comparison group, and we use entropy balancing weights such that the weighted comparison group is as similar to the group of participating plans on observable

characteristics as possible. This analytic strategy can be summarized into four distinct stages, which are described in detail in subsequent sections:

1. Definition of groups of participating plans and the effects of interest
2. Identification of nonparticipating plans that are eligible for VBID
3. Construction of outcome-specific comparison groups using entropy balancing for each of the groups in #1 using the comparisons identified in #2
4. Estimation and summarization of DD models using the comparison groups derived in #3.

Several POs participated in the Phase I (2017–2019) iteration of the model test and thus had matched comparators in RAND's prior evaluation. We disregarded these previous matches for the current evaluation as this evaluation uses a different analytic strategy from the previous evaluation. For the purposes of this evaluation, participation in the previous version of VBID was considered pre-participation activity. We assessed the effect of the current version of VBID in isolation; thus, we estimated the effect of any additional changes to their interventions. Further discussion of this is provided at the end of this appendix.

Defining Groups of Participating Plans

We limited our analyses to Medicare Advantage–Prescription Drug (MAPD) plans because very few MA-only plans participated, and we expected substantial differences in the design and structure of MAPD and MA-only plans owing to Part D coverages. Analyses including MA-only plans were conducted as sensitivity analyses.

Here, we describe groups of plans participating in BDI and corresponding effects of BDI. Suppose plans are observed for time points $t = 2017, \ldots, 2021$ and let a_t be a binary indicator that is defined to be one if a plan participates in VBID in year t and zero otherwise. We assume that 2020 is the first year that plans are eligible to participate in the VBID model test, and we consider participation in Phase I of VBID (2017–2019) as pre-participation activity. Define $y_{it}(\boldsymbol{a})$ for $2017 \leq t \leq 2021$ as the outcome that would have been observed for unit *i* at time *t* if the unit follows BDI participation pattern $\boldsymbol{a} = (a_{2020}, a_{2021})$. There are four possible BDI participation patterns $\boldsymbol{a}$ as of 2021, which we describe in Table C.1.

Table C.1. Four Possible BDI Participation Patterns as of 2021

Participation Pattern	Participated in 2020 (a_{2020})	Participated in 2021 (a_{2021})	Analytic Sample Size[a]
Comparison	0	0	—
Participated in both 2020 and 2021	1	1	89
Participated only in 2020	1	0	39
Participated only in 2021	0	1	257

[a] The analytic sample sizes represent the number of plans after accounting for plan consolidations and data cleaning.

Plans that participated only in the Hospice component are considered comparison plans for the evaluation of BDI since plans could opt to participate in BDI only, the Hospice component only, or both. The group of plans that discontinued participation in the BDI component of the VBID model test (plans that participated only in 2020) represents a potential analytic concern as it is a departure from the current DD literature in staggered adoption. Existing methods for DD with multiple adoption points either explicitly or implicitly assume that once participating, always participating, and do not allow for discontinuation of an intervention. To allow for this possibility, we extend the methodology of Callaway and Sant'Anna (2021) to allow for discontinuations. Note that effects were not estimated for this group after discontinuation.

Define average treatment effects for each BDI participation history $\boldsymbol{A} = \boldsymbol{a}$ against the comparison group of nonparticipants as

$$ATT(\boldsymbol{a}, t) = E[y_{it}(\boldsymbol{a}) - y_{it}(\mathbf{0}) \mid \boldsymbol{A} = \boldsymbol{a}].$$

These $ATT(\boldsymbol{a}, t)$ can be estimated using comparisons between a group of participating plans defined by the BDI participation pattern $\boldsymbol{A} = \boldsymbol{a}$ and a group with $\boldsymbol{A} = \mathbf{0}$. These estimates can then be used as the building blocks for different overall effects that summarize the group-time effects, that is,

$$\sum_{\boldsymbol{a} \in \mathcal{A}} \sum_{t=2020}^{2021} w(\boldsymbol{a}, t) * ATT(\boldsymbol{a}, t),$$

where $\mathcal{A}$ represents all possible participation histories and $w(\boldsymbol{a}, t)$ represents a set of weights. The choice of $w(\boldsymbol{a}, t)$ can be varied to answer different research questions. While this framework allows for arbitrary choices of w, we consider two definitions that facilitate interpretation for this evaluation. We consider the following two types of aggregated effects:

1. The effect of participating *e* time periods after initial adoption among those that participated for at least *e* time periods. These effects allow us to estimate how the effect of BDI participation changes over time, as plans gain experience with the model test and the set of participating plans change.
    a. The weights, $w(\boldsymbol{a}, t)$, are defined as the proportion of plans in each group among those plans that participated for at least *e* time periods.
2. The effect in year *t* among plans participating in year *t.* These effects allow us to estimate the average effect of BDI participation in each calendar year, across all participating plans in that year.
    b. The weights, $w(\boldsymbol{a}, t)$, are defined as the proportion of plans in each group among plans that participated in year *t.*

The weights used in this report are based on the number of MAPD plans that were included in our final models. There were 140 plans participating in BDI in 2020; two of these plans were excluded because they were MA-only plans, and 10 others were excluded because they did not exist before 2020 (i.e., these plans have no pre-participation data). Of the remaining 128 MAPD plans, 89 participated in BDI in both 2020 and 2021, and 39 participated only in 2020. There were 377 plans that participated in BDI in 2021; three of these plans were excluded because they were MA-only plans, and 28 other plans were excluded because they did not exist before participation. Of the remaining 346 MAPD plans, 89 participated in BDI in both 2020 and 2021, and 257 participated only in 2021. Tables C.2 and C.3 provide the weights used in this report based on the total analytic sample of 385 MAPD plans.

Table C.2. Weights for Estimating the Effect of BDI, First and Second Year of Participation

Participation Pattern (number of plans)	1st Year $w(\boldsymbol{a}, \mathbf{2020})$	1st Year $w(\boldsymbol{a}, \mathbf{2021})$	2nd Year $w(\boldsymbol{a}, \mathbf{2020})$	2nd Year $w(\boldsymbol{a}, \mathbf{2021})$
Participated in both 2020 and 2021 ($\boldsymbol{a} = (\mathbf{1}, \mathbf{1})$, N = 89)	$\frac{89}{89 + 257 + 39}$	0	0	1
Participated only in 2021 ($\boldsymbol{a} = (\mathbf{0}, \mathbf{1})$, N = 257)	0	$\frac{257}{89 + 257 + 39}$	0	0
Participated only in 2020 ($\boldsymbol{a} = (\mathbf{1}, \mathbf{0})$, N = 39)	$\frac{39}{89 + 257 + 39}$	0	0	0

Table C.3. Weights for Estimating the Effect of BDI, 2020–2021

Participation Pattern (number of plans)	2020 $w(\boldsymbol{a}, \mathbf{2020})$	2020 $w(\boldsymbol{a}, \mathbf{2021})$	2021 $w(\boldsymbol{a}, \mathbf{2020})$	2021 $w(\boldsymbol{a}, \mathbf{2021})$
Participated in both 2020 and 2021 ($\boldsymbol{a} = (\mathbf{1}, \mathbf{1})$, N = 89)	$\frac{89}{89 + 39}$	0	0	$\frac{89}{89 + 257}$
Participated only in 2021 ($\boldsymbol{a} = (\mathbf{0}, \mathbf{1})$, N = 257)	0	0	0	$\frac{257}{89 + 257}$
Participated only in 2020 ($\boldsymbol{a} = (\mathbf{1}, \mathbf{0})$, N = 39)	$\frac{39}{89 + 39}$	0	0	0

The evaluation of the Hospice component uses these same definitions of participating plans, but since the Hospice component was first introduced in 2021, there is only a single group of plans participating in the Hospice component based on the participation pattern (plans that participated only in 2021). Therefore, there is only a single relevant ATT when analyzing the Hospice component, $ATT(\boldsymbol{a} = (\mathbf{0}, \mathbf{1}), 2021)$. Plans participating only in the BDI component are

considered comparison plans for the analysis of the Hospice component. We do not analyze the WHP requirement separately from the BDI and Hospice components because WHP was mandatory and, as a result, it is difficult to disentangle its effects from other aspects of the model test. In a future report, we will address a series of research questions that CMS raised specifically pertaining to the WHP requirement.

Interpretation of Aggregate BDI Effects

The previously defined effects can be used to address different policy questions, and thus we provide some guidance for the interpretation of each effect of BDI. Note that all effects of BDI are estimated relative to a balanced comparison group that did not participate in BDI in either year. The effects are as follows:

- ***Effect of BDI in the first year of participation***: This is the average effect of BDI during the first year of implementation among plans that participated in BDI in 2020 or 2021. Thus, this effect contains all plans participating in BDI and evaluates their effect at a common relative time point.
- ***Effect of BDI in the second year of participation***: This is the average effect of BDI during the second year of participation among plans that participated in BDI for at least two years. For the current report, only plans that participated in both 2020 and 2021 are included, because this is the only group of plans that have participated in BDI for at least two years. Note that the plans that participated in both 2020 and 2021 have different characteristics and different interventions from those that participated only in 2021, such that this second-year effect may not generalize to other groups of participating plans moving forward.
- ***Effect of BDI in 2020***: This is the average effect of BDI participation in 2020 among all plans that participated in 2020, including plans that participated only in 2020 and plans that participated in 2020 and 2021. This represents the overall effectiveness of BDI in 2020 among plans that participated in BDI in 2020. Note that this effect may not generalize to future years because (a) it is an effect among the plans that participated in 2020 (an ATT) and (b) the number of plans participating in BDI has grown over time.
- ***Effect of BDI in 2021***: This is the average effect of BDI in 2021 among all plans that participated in 2021. This represents the overall effectiveness of BDI in 2021 among plans that participated in BDI in 2021. Note that this averages the effect in 2021 of plans that participated in BDI for only a single year (plans that participated only in 2021) and plans that participated in BDI for two years (plans that participated in both 2020 and 2021). In addition, this effect may not generalize to future years because of the changing landscape of participation in BDI.

Identification of Eligible Nonparticipating Plans

See Appendix D for a detailed description of the eligibility criteria used to identify eligible nonparticipating plans. As noted previously, when analyzing BDI, participants in the Hospice component are considered eligible nonparticipating plans, and vice versa.

Entropy Balancing for Outcome-Specific Comparison Groups

In this section, we describe the tools we used for finding comparable groups to each group of participating plans defined in Table C.1. Plans volunteered to participate in VBID, and those that did so differed from eligible nonparticipating plans with respect to many observable characteristics, as shown in Table D.3. We sought to construct comparison groups to minimize these differences to improve comparability between the groups.

Briefly, we use an entropy balancing approach, which achieves comparability between the VBID participating and eligible nonparticipating plans by weighting the nonparticipating plans to match the VBID group. To select the weights, we implemented an optimization approach that constrains the standardized mean difference, a measure of comparability between groups, to be small. The optimization algorithm requires a data set with no missing data, so we used a simple imputation process to fill in missing values. Note that this imputation process was used only for the derivation of weights and was not used when fitting the DD models. A detailed description of this imputation strategy is provided later. Below, we describe the entropy balancing steps in more detail, including a discussion of why we selected the entropy balancing approach as opposed to other matching approaches, and how we settled on the standardized mean differences as the metric for assessing comparability.

Matching Versus Entropy Balancing

For readers more familiar with matching methods in observational studies and less familiar with the weighting approaches (i.e., entropy balancing) that we implemented, we provide a brief discussion on how these two methods are related. Matching and weighting are popular methods for selecting comparison groups. For example, a 1:1 propensity score matching *without* replacement assigns each member of the comparison group a weight of 0 or 1, depending on whether it was identified as a match. A 1:1 propensity score matching *with* replacement assigns each member of the comparison group a weight equal to the number of times it was matched (0, 1, 2, 3, . . .). In this view, *matching is equivalent to a weighting approach that constrains the weights to be integers.* An analogous propensity score weighting approach would simply allow the weights to be nonintegers, effectively up-weighting or down-weighting comparison group members in a more continuous fashion than matching. With this view in mind, matching and weighting approaches share more similarities than differences, and weighting offers more flexibility than matching by allowing for noninteger weights and often can provide considerable computational advantages. Entropy balancing is one such weighting approach.

A primary benefit of entropy balancing over other approaches is that the analyst can have more fine-grained control over the characteristics of the weighted sample than with a matched sample. For example, in our analyses, we implement an entropy balancing procedure that uses optimization to select weights subject to constraints on the balance of the distributions between treated and control groups. As the procedure explicitly specifies the balance constraints in the optimization, it can ensure that we select a weighted sample that meets our needs. Conversely, matching is done by choosing individuals with similar characteristics and evaluating the balance of the distributions as a post hoc procedure. Checking balance in this way makes it difficult to adjust the analysis when an insufficient matching is found. A detailed description of entropy balancing is provided in the subsequent sections.

Defining Measures of Similarity

A critical part of the estimation strategy is finding balancing weights such that the weighted distribution of observable characteristics (e.g., for-profit status) in comparison plans is similar to the characteristics in the participating plans. We define the similarity between each group of participating plans in BDI and the comparison plans using the standardized difference in means (SDM) between each group's covariate distributions (note that our final approach balances variances in addition to the means). To quantify this, first define X_j as the j^{th} characteristic under consideration and consider two arbitrary groups of plans, where $g = 0$ indicates a group of comparison plans and $g = 1$ indicates a group of participating plans. Define N_g as the number of units in group $G = g$. The unweighted sample mean and variance in each group is given by

$$\bar{X}_{gj} \equiv \frac{\sum_{i:G_i=g} X_{ij}}{N_g} \quad \text{and} \quad s_{gj}^2 \equiv \frac{\sum_{i:G_i=g}\left(X_{ij}-\bar{X}_{gj}\right)^2}{N_g-1}.$$

This allows us to define the unweighted SDM for the ATTs of interest as

$$SDM_j \equiv \frac{\bar{X}_{1j}-\bar{X}_{0j}}{\sqrt{s_{1j}^2}}.$$

We can extend the above to accommodate a weighted comparison group by defining a set of weights w and letting the weighted mean be defined as

$$\bar{X}_{0j}(w) = \frac{\sum_{i:G_i=0} w_i X_{ij}}{\sum_{i:G_i=0} w_i}.$$

The weighted SDM can be defined as

$$SDM_j(w) = \frac{\bar{X}_{1j} - \bar{X}_{0j}(w_i)}{\sqrt{s_{1j}^2}}.$$

Finally, the above can accommodate arbitrary functions of random variables. Let $h(\cdot)$ be an arbitrary function, and define

$$\bar{X}_{h,gj} \equiv \frac{\sum_{i:G_i=g} h(X_{ij})}{N_g},$$

$$\sigma_{h,gj}^2 \equiv \frac{\sum_{i:G_i=g} \left(h(X_{ij}) - \bar{X}_{h,gj}\right)^2}{N_g - 1}, \text{ and}$$

$$\bar{X}_{h,0j}(w_g) = \frac{\sum_{i:G_i=0} w_{ig} h(X_{ij})}{\sum_{i:G_i=0} w_{ig}}.$$

The SDM of the random variable $h(X_j)$ can be defined as

$$SDM_{h,j}(w_g) = \frac{\bar{X}_{h,1j} - \bar{X}_{h,0j}(w_g)}{\sqrt{\sigma_{h,1j}^2}}.$$

For example, if we define the function $h(X) = (X - \bar{X})^2$ we could effectively obtain a standardized measure of how close the variance of a characteristic in the comparison group of plans is to the variance of a characteristic in the group of participating plans. In this report, we do not report balance on arbitrary functions and higher moments, but their definition will be useful for fully describing our entropy balancing algorithm.

Weight Selection—Entropy Balancing Algorithm

There are many methods for estimating weights to balance observable characteristics between two groups. These include indirect methods such as propensity scores weighting (Imbens and Rubin, 2015; Hernán and Robins, 2010; Pearl, Glymour, and Jewell, 2016) and direct methods such as stable balancing weights (Zubizarreta, 2015), covariate balancing propensity score (Imai and Ratkovic, 2014), and entropy balancing (Hainmueller, 2012). For this evaluation, we used entropy balancing, an optimization-based method of obtaining weights, as it allowed us to prespecify balance constraints on the distribution of the observable characteristics. To select weights, we modified the original entropy balancing algorithm as follows:

$$\min_w \sum_{i:G_i=0} w_i \log\left(\frac{w_i}{q_i}\right) \quad \text{subject to}$$

$$\left|SDM_{h,j}(w)\right| \leq \delta \quad \text{for each covariates } j \text{ and each function } h(\cdot),$$

$$\sum_{i:G_i=0} w_i = 1, \; w_i > 0$$

where q represents a set of "base" weights and δ represents a constraint on the maximum absolute SDM. The base weights could be a set of predefined survey weights, but, in practice, these are often set to be uniform weights (i.e., $q_i = 1/N_0$). This implementation of entropy balancing represents a departure from the original Hainmueller (2012) methodology by not requiring the constraint on the SDMs to be set to zero. Choosing δ represents a trade-off between bias and variance (Wang and Zubizarreta, 2020), and the amount of information in the weighted sample can be measured using Kish's effective sample size (Kish, 1965):

$$ESS(w) = \frac{(\sum_i w_i)^2}{\sum_i w_i^2}.$$

The amount of information can range from 1 to the original sample N. A low effective sample size implies that there may be insufficient information in the sample and that it is difficult to find comparable units between the two groups. Larger values of δ will lead to larger sample sizes, but this comes at the cost of balance between the groups. In practice, SDM values lower than $\delta = 0.1$ are often seen as sufficient for outcome estimation (Austin, 2009; Stuart, Lee, and Leacy, 2013).

Implementation of Entropy Balancing

We used entropy balancing to derive a weighted comparison group separately for each outcome and each group of BDI-participating plans defined in Table C.1. In particular, entropy balancing was used to balance a set of observable pre-participation characteristics, including pre-participation outcome trends. We note that there has been significant discussion on weighting and matching on pre-treatment trends, including the potential to admit bias into estimation; for example, see discussions in Daw and Hatfield (2018) and Zeldow and Hatfield (2019). Our decision to balance pre-participation plan characteristics and outcome trends is based on results from several studies. First, the decision to balance plan characteristics is supported by Zeldow and Hatfield (2021). An exploratory data analysis indicated our plan characteristics trends are roughly parallel over time, and, based on Zeldow and Hatfield (2021), balancing pre-participation characteristics is approximately unbiased in these situations. It is important to note that the simulations of Zeldow and Hatfield (2021) assume any deviation in the parallel trends assumption is through the observed covariate, such that balancing the covariate level is sufficient for satisfying the DD assumptions when the covariate has parallel evolution and does not deviate in the post-period (Zeldow and Hatfield, 2021). Second, we were concerned about additional unexplained differences in outcome trends (after balancing the pre-participation characteristics), so we balance pre-participation outcome trends. Several simulation studies have shown that balancing pre-period outcome trends reduces bias in DD models relative to the unadjusted DD (Daw and Hatfield, 2018; Lindner and McConnell, 2019; Arkhangelsky et al., 2021). These studies have shown that balancing outcome trends does not increase bias—simulations that show

increased bias after balancing are limited to approaches that balance outcome levels. We avoided balancing on pre-intervention outcome levels to avoid the possibility of bias in our DD models due to regression to the mean, and we proceeded with balancing pre-participation outcome trends.

We balanced the VBID and comparison groups using a range of pre-participation characteristics including beneficiary demographics, plan characteristics, characteristics of the local health care market, and pre-intervention outcome trends. Table C.5 at the end of this appendix provides the full set of characteristics that were included in our entropy balancing approach. We balanced both the first moments (i.e., $h(X) = X$) and the second moments (i.e., $h(X) = X^2$) for each characteristic and the pre-participation trends. This ensures similarity in both the mean and the variability between the participating and comparison plans. To select δ, we attempted to find a value less than 0.1 that provided an effective sample size from the comparison plans that was approximately equal to the sample size in the group of participating plans. Outcome trends were included nonparametrically as first differences over time.

Strategy for Missing Data

Two different classes of missing data existed in our data: (1) missingness in pre-participation outcome information for plans that did not yet exist and (2) rare instances of missingness in some covariate values. Table C.4 provides counts of the numbers of plans missing outcome information because the plans did not yet exist in each year for the final analytic sample.

Table C.4. Number of Plans Missing Data Because the Plans Did Not Yet Exist

Participation Pattern	2017	2018	2019	2020	Total Number of Plans
Participated in BDI in both 2020 and 2021	16	9	0	0	89
Participated in BDI only in 2020	7	1	0	0	39
Participated in BDI only in 2021	102	70	30	0	257
Hospice component	9	7	3	0	46

SOURCE: RAND analysis of VBID model test application materials and plan crosswalk data.

To handle missingness in pre-participation outcome information when constructing weights with entropy balancing, the general strategy was to ensure that the missing data were imputed to match the marginal distribution conditioned on the pattern of VBID participation. For example, to handle missingness for plans missing 2017 and 2018 data, we first fit a model $Y_{2018} = \beta_0 + \beta_{1g} I(G = g) + \beta_2 Y_{2019} + \beta_{3g} I(G = g) \times Y_{2019}$ to impute missing outcomes for 2018 and then used a similar model to impute Y_{2017}. This ensured that within each group, the imputed values were at the approximate correct level. Indicators were added to the set of balancing

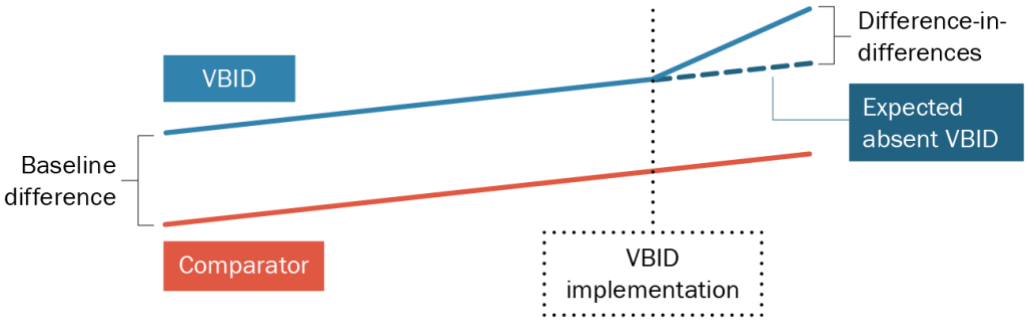
characteristics so that we balanced the proportion of plans with missingness between each group and the comparison plans.

For missing covariate information, only two characteristics had missing values: Star Ratings and standardized Medicare costs. Star Ratings were missing for 110 out of 2,281 plans in 2019, primarily driven by their contracts not existing for a sufficient period to receive a Star Rating. Standardized Medicare costs were missing for only 24 plans out of 2,281 in 2019. These low levels of missingness were imputed by sampling observations from the conditional distribution of that covariate by group, which ensures that the observations were imputed from the correct marginal distribution. The imputation of missing data was necessary for the derivation of the entropy balancing weights, but the imputations were not used directly in the DD regression models.

Difference-in-Differences

The construction of weighted comparison groups should account for selection bias associated with observed characteristics (i.e., confounding bias because of the participating and comparison plans differing with respect to observed characteristics). The DD methodology (Figure C.1) will account for any remaining unobserved differences and any secular trends in the outcomes common to both VBID-participating plans and the comparison group under a key and strong central assumption. A DD model works by assuming that the post-participation trend in the outcome for the comparison plans is a proxy for the trend in the VBID-participating plans had they not participated in VBID (the "parallel trends" assumption), and then compares the change in the pre-participation outcome with the post-participation outcome between participating and comparison plans. The parallel trends assumption is untestable, but as a proxy we will use the pre-participation outcome trends. Using the pre-participation outcome trends as a proxy for the parallel trends assumption is valid only if we assume that the similarity of the trends in the pre-period implies that the trends would be similar in the post-period absent VBID participation. The assessment of parallel trends is incorporated into the entropy-balancing algorithm; we report graphs of the pre-trends for the outcomes considered in this report in Appendixes H and I.

Figure C.1. Conceptual Illustration of Difference-in-Differences Methodology



We specified DD models to account for any time-invariant unobserved differences and any common shocks that occur during the post-intervention period. Specifically, let y_{pt} denote the outcome of plan *p* at time *t*, let $VBID_p$ indicate that plan *p* is a VBID participating plan, and let DID_{pt} denote the DD indicator for plan *p* at time *t* ($DID_{pt} = 1$ for VBID participating plans in the post-intervention period and 0 otherwise). We will use weighted DD models of the form

$$y_{pt} = \alpha_p + \eta_t + \beta_t * DID_{pt} + \delta_t X_{pt} + \epsilon_{pt},$$

where α_p is a plan-specific intercept, η_t is a time fixed effect, β_t is the effect of interest at year *t*, and ϵ_{pt} is an error term. These DD models will be fit for each of the groups of participating plans previously described such that the β_t are estimates of $ATT(\boldsymbol{a}, t)$ for those groups of plans. The choices of X_{pt} include participation in other initiatives, including Part D Senior Savings (PDSS), Special Supplemental Benefits for the Chronically Ill (SSBCI), uniformity flexibility (UF), and new primarily health-related supplemental benefits (PHRSB).

The β_t for each group of participating plans were then summarized using the previously described methodology. Variance estimates were derived using a smooth version of the bootstrap such that plans were repeatedly reweighted using a uniform Dirichlet distribution to approximate the sampling distribution. The entire estimation process, including derivation of the entropy balancing weights, was repeated for each set of bootstrap weights.

Approach for Plans That Participated in Phase I of VBID

Our analysis focuses on the effects of Phase II of the model test independently from Phase I, treating Phase I and Phase II as separate interventions. This means that for plans that participated in Phase I, we capture any marginal effects of participation in Phase II but ignore effects stemming solely from participation in Phase I. To identify the effects of Phase II over and above those of Phase I, we would ideally want a comparison group of plans that continued participation in Phase I after 2019. However, such a comparison group does not exist, because Phase I was

replaced by Phase II. Instead, we assume that if outcome trends in the pre-participation period were similar between Phase I participants and comparison plans, post-participation trends would have remained similar had Phase I continued (and Phase II had not been implemented). As with the DD approach generally, we need only make a strong assumption about outcome trends, because we account for differences in the intercept between plans using plan-level fixed effects, which absorb the average effect of Phase I for plans that participated.

Limitations

Our approach has several important limitations that should be considered when interpreting the findings. First, participation in the VBID model test is voluntary, and participants were observably different from eligible nonparticipants (see Table D.3). Although we control for a large number of relevant plan-, PO-, and community-level characteristics in our entropy-balancing algorithm, and the methods described above aim to enable causal inference in a nonrandomized research setting, the possibility remains that we could ascribe effects driven by underlying differences between participating and nonparticipating plans to the model test. This concern is heightened by the fact that the parallel trends assumption, which is critical to the validity of DD methods, is inherently untestable. For this reason, we use associational rather than causal language when discussing our findings.

Second, our analysis treats Phase II of the model test as separate and distinct from Phase I and does not give plans "credit" for effects stemming from Phase I participation. For example, if a plan participated in Phase I and experienced reductions in bids as a result, our analysis would capture this effect only if the trend in bids diverged from the Phase I pattern after Phase II was implemented. Conceptually, this approach measures the effects of participation in Phase II of the model test. However, it could underestimate the effects of participation in VBID if VBID is defined to include interventions allowed in either Phase I or Phase II. We expect that any underestimate would be small, however, because only four POs continued participation from Phase I, and one of those POs substantially revised its intervention in Phase II.

Third, beneficiary-level participation varied substantially across plans, both because some plans targeted a broader range of beneficiaries and because some plans had higher uptake conditional on beneficiaries' eligibility status. Our analysis uses an intent-to-treat (ITT) approach in which we treat all plans equally, regardless of the scope of their interventions or beneficiary-level uptake. The ITT approach is useful from CMS's perspective because it sheds light on how implementing a model of this type is likely to affect outcomes in practice, given uneven implementation and other differences across participants. However, it may not generalize to specific plans that have designed interventions with broader or narrower eligibility criteria or that have instituted policies to maximize uptake.

Similarly, the model test enabled POs to implement a wide range of benefit design changes. Other than running separate analyses for the BDI and Hospice components, we do not consider

subgroup analyses focused on specific types of BDI or Hospice implementation. Accordingly, our results must be interpreted as estimating the overall relationship between BDI or Hospice participation and key outcomes, and do not necessarily generalize to specific intervention designs, such as reducing cost sharing for high-value drugs or providing a healthy food card to people with low SES.

Table C.5. Variables Used in This Report's Analyses

Variable	Data Source	Years	Construction of Variable	Used For
County-level characteristics				
Area-level income	American Community Survey (ACS) 2019 5-year estimates	2019	Defined as median income in the past 12 months (in 2019 inflation-adjusted dollars). Merged with plan benefit package (PBP)-county file and weighted by beneficiary-months.	Entropy balancing
MA penetration	CMS	2017–2021	Derived from July County-State Penetration file from CMS. Merged with PBP-county file and weighted using beneficiary months.	Entropy balancing
Urbanicity	Rural-Urban Continuum Codes (RUCC)	2013	Values assigned using following schema: metro/urban = {1, 2, 3}; adjacent to metro/suburban = {4, 6, 8}; nonmetro/rural = {5, 7, 9}. Calculate share of counties within PBP that are urban, suburban, and rural.	Entropy balancing
Health Professional Shortage Area	Area Health Resources Files (AHRF)	2017–2021	Defined using Health Professional Shortage Area (HPSA) Primary Care code from AHRF file. Merged with PBP-county file, calculated the share of beneficiary-months in counties designated as a whole shortage area (rather than partial shortage area).	Entropy balancing, descriptives
Standardized Medicare costs per capita	AHRF	2019	Merged with PBP-county file and weighted using beneficiary-months	Entropy balancing
% population > 65	AHRF, Integrated Data Repository (IDR): dim_bene_enrlmt_snpsht_crnt & dim_geo tables	2020–2021	% age over 65 weighted by county-level enrollment in the plan	Entropy balancing
PO-level characteristics				
Blue Cross and/or Blue Shield affiliate	CMS	2017–2021	Text field search of PO- and organization-level names (see table note)	Descriptives
For-profit status	Health Plan Management System (HPMS): Contract info	2020–2021	Contract-level information on for-profit status; aggregated to the PO level, using the most common status across contracts	Descriptives

Variable	Data Source	Years	Construction of Variable	Used For
State/regional/national	CMS	2017–2021	State = 0 (≤2 states), regional = 1 (3–8 states), national = 2 (9+ states)	Descriptives
Urbanicity	RUCC; CMS	2017–2021	Counties within a PO's service area are classified according to urbanicity metric above	Descriptives
MA penetration rate	CMS	2017–2021	Average MA penetration rate for counties in PO's service area, enrollment weighted	Descriptives
Median income	ACS 2019 5-year estimates; CMS	2019	Median income across counties in PO's service area, enrollment weighted	Descriptives
PO enrollment	CMS	2017–2021	Total enrollment across eligible plans	Descriptives
Contract-level characteristics				
Star Rating (overall)	CMS Star Rating	Reporting year 2017–2021	Using reporting years (2017–2021), final overall scores were obtained for each contract	Entropy balancing
For-profit status	HPMS: Contract info	2020–2021	Contract-level information on for-profit status	Entropy balancing
Plan-level characteristics				
Enrollment	IDR, bene_fct_trans table	2017–2021	Defined as number of beneficiaries enrolled in the plan in the month of July for the respective year	Outcome
Bids—MA	CMS Office of the Actuary (OACT)	2017–2021	Defined as standardized Part A/B bid (@ 1.000).	Entropy balancing, outcome
Bids—Part D (if applicable)	OACT	2017–2021	Defined as standardized Part D bid amount	Entropy balancing, outcome
MA premiums	CMS	2017–2021	Part C premium variable from Medicare Part D landscape files	Descriptives, entropy balancing, outcome
Part D premiums (if applicable)	CMS	2017–2021	Part D total premium variable from Medicare Part D landscape files	Descriptives, entropy balancing, outcome
Cost of mandatory supplemental benefits (MSB)	OACT	2017–2021	Defined as total net per member, per month (PMPM) for additional services	Entropy balancing, outcome
Rebate dollars amount	OACT	2017–2021	Defined as the total rebate dollars per enrollee per month, from the OACT bid data file	Entropy balancing, outcome
Administrative costs (bid data)	OACT	2017–2021	Defined as the sum of MA nonbenefit expenses and total Part D nonbenefit expenses	Entropy balancing, outcome

Variable	Data Source	Years	Construction of Variable	Used For
Offers Part D	CMS PBP Benefits Data	2017–2021		Descriptives, entropy balancing
Out-of-pocket (OOP) maximum (Part C)	CMS PBP Benefits Data	2017–2021		Descriptives, entropy balancing
PDSS participant	CMS PBP Benefits Data	2021		Descriptives, control
Offers UF	CMS PBP Benefits Data	2019–2021		Descriptives, control
Offers SSBCI	CMS PBP Benefits Data	2020–2021		Descriptives, control
Offers new PHRSB	CMS PBP Benefits Data	2019–2021	Defined to indicate if a plan offers at least one new primarily health-related benefit	Descriptives, control
Type of plan	CMS	2017–2021	Plan type from publicly available CMS contract information file; preferred provider organization (PPO) = 1; otherwise= 0	Entropy balancing
Special Needs Plan (SNP)	HPMS: Plan information files	2019–2020	Identified SNPs using SNP flag in the source data	Descriptives, entropy balancing
SNP type	HPMS: Plan information files	2019–2021	Identified SNP type using SNP flag in the source data	Descriptives, entropy balancing
Urbanicity	RUCC; CMS	2017–2021	Counties within a PBP's service area are classified according to urbanicity metric above	Descriptives
Plan-level beneficiary characteristics				
Age	IDR, bene_fct_trans table	2019–2020	For beneficiaries who are continuously enrolled for 12 months in the same plan, we take the mean age of beneficiaries for each plan. Age is calculated as the beneficiary's age as of December 31 of that year.	Entropy balancing
Age	IDR, bene_fct_trans table	2017–2021	Mean age of beneficiaries in the plan during the month of July of the respective year.	Descriptives
Sex	IDR, bene_fct_trans table	2019–2020	For beneficiaries who are continuously enrolled for 12 months, we take the percentage of males for the respective year for each plan	Descriptives, entropy balancing

Variable	Data Source	Years	Construction of Variable	Used For
Race/ethnicity	Medicare Bayesian Improved Surname Geocoding dataset	2019–2020	For beneficiaries who are continuously enrolled for 12 months in the same plan, we take the mean probability of each beneficiary being a given race and roll up to the plan level	Descriptives, entropy balancing
Dual	IDR, mdcr_bene_dual_stus table, mdcr_bene_low_incm_terr, mdcr_bene_pos	2019–2020	For beneficiaries who are continuously enrolled for 12 months in the same plan, if they are flagged as dual in any of the three tables during the respective year, they are considered dual and then rolled up to the plan level	Descriptives, entropy balancing
Low-income subsidy (LIS) status	IDR, bene_fct_trans table	2017–2021	Beneficiaries with full or partial status for LIS are considered to have LIS. Beneficiaries considered are those who are enrolled in the plan during July of the respective year.	Descriptives
Average MA risk score (HCC)	IDR, mdcr_bene_risk_ptc_scre_asg	2019–2021	For beneficiaries who are continuously enrolled for 12 months in the same plan, take mean final beneficiary risk score and roll up to the plan level.	Descriptives, entropy balancing
Average Part D risk score (RxHCC)	IDR, mdcr_bene_risk_ptc_scre_asg	2019–2022	For beneficiaries who are continuously enrolled for 12 months in the same plan, take mean final beneficiary Rx risk score and roll up to the plan level.	Entropy balancing
HCC indicators for 1. Diabetes 2. Congestive heart failure (CHF) 3. Chronic obstructive pulmonary disease (COPD) 4. Cancer	IDR: mdcr_bene_risk_ptc_l_scre	2020–2021	Hierarchical Condition Categories (HCC) flags	Entropy balancing

NOTE: Many of the CMS data sets for characteristics of POs and plans are publicly available on CMS's website (CMS, 2021a). To classify POs as Blue Cross Blue Shield (BCBS) affiliates, we searched the two organization name fields for the following strings: blue; bcbs; anthem; wellpoint; highmark; hawaii medical; california physician; carefirst; independence health; health care service corporation; or usable mutual. We reviewed the categorization and removed the classification from three organizations that were not BCBS affiliates (e.g., Bluestem Communities), and added it to two organizations not picked up in this search string (e.g., Triple-S Corporation). If at least one organization was a BCBS affiliate, we assigned the PO to the BCBS group.

Appendix D. Plan Eligibility and Characteristics of Parent Organizations and Plans

In this appendix, we discuss the PO and plan eligibility criteria used to develop our sample for analyses in this report; we then describe characteristics of participating POs and plans that are summarized in Chapter 2.

Parent Organization and Plan Eligibility

We describe the criteria used to select plan benefit packages (hereafter, plans) for our analytic sample, including participating plans and nonparticipating plans that are eligible for the pool of comparison plans.

The Center for Medicare & Medicaid Innovation (CMMI) establishes the MA VBID eligibility criteria (CMS, 2020a; CMS, 2020b; CMS, 2019). The criteria, which were the same for the BDI and Hospice components of the model test, include the following:

- Plans must be an MA or MAPD plan (no stand-alone Part D plans).
- Eligible plan types were limited to Coordinated Care Plans (health maintenance organization [HMO], health maintenance organization–point of sale [HMO-POS], local preferred provider organization [PPO], or regional PPO [RPPO]) and Special Needs Plans (Chronic Condition Special Needs Plan [C-SNP], Dual-Eligible Special Needs Plan [D-SNP], or Institutional Special Needs Plan [I-SNP]).
- POs must have at least one plan with 2,000 enrollees.
    - CMMI removed this criterion for 2021.
- Plans needed to be offered in three prior open enrollment periods (OEPs).
    - CMMI relaxed this criterion for 2021 to only require that a PO have at least one plan that had been offered in three prior enrollment periods.
- Plans had to have sufficiently high performance in the application year, which included the following:
    - not being under sanction
    - the contract for the plan had to have at least a three-star overall rating
    - the plan could not be a "consistently low-performing" plan in the Medicare plan finder
    - the organization could not be an outlier in the Past Performance Review.

- Segmented plans were not allowed to enter a plan with a different intervention across segments.[1]

We made several modifications to these criteria based on conversations with CMMI where we learned that exceptions were granted, and also because the criteria shifted slightly between 2020 and 2021 (as noted above). Table D.1 shows the eligibility criteria that we applied, the data sets used to assess the criteria, and the date of the data used to make the eligibility assessment. Because VBID applications are due the year before the model test year, most of the data used to make the eligibility determination come from the year or two before participation begins. Our major changes were as follows:

- We excluded I-SNPs from the comparison group because there were no I-SNP participants in 2020 or 2021 and the beneficiaries in these plans were very different from those enrolled in VBID participating plans.
- We applied several criteria to the PO level rather than to the MA Organization or contract level when the criterion was not specific at the level to which it was applied.
- We did not use two of the performance criteria because data were not uniformly available in all years and CMMI had also granted exceptions to these criteria.

Table D.1. Criteria and Data Sources Used to Identify Plans Eligible for VBID

Criteria Category	Specific Criteria	RAND Application of Criterion	Data Set Used for Assessment	Date for 2020 Assessment	Date for 2021 Assessment
Plan type	Must be HMO, HMO-POS, Local PPO, RPPO, or any SNP (C-, D- or I-SNP)	Must be HMO, HMO-POS, Local PPO, RPPO or C/D-SNPs	Contract Info File (plan, state, county level), SNP Data	July 2019	July 2020
Enrollment[a]	At least one plan for the applicant organization has at least 2,000 enrollees; criterion dropped for 2021	Applied at the PO level for 2020 only	Enrollment File (plan, state, county level)	July 2019	N/A
Experience	At least one plan in the organization had 3+ years of experience (i.e., available in at least 3 OEPs)	PO must have at least one contract offered for 3 or more years, using 1/1/2020 or 1/1/2021 as the date to determine the three years in operation	Contract Info File (plan, state, county level)	July 2019	July 2020
Performance[b]	Plan's contract has at least a 3-Star overall rating	Applied the 3-Star rating at the contract level	Star Ratings, Summary Rating	2019 (Fall 2018 release)	2020 (Fall 2019 release)

[1] POs are allowed to divide a given plan's service area into multiple segments and to vary plan design features across these geographic units.

Criteria Category	Specific Criteria	RAND Application of Criterion	Data Set Used for Assessment	Date for 2020 Assessment	Date for 2021 Assessment
			Tab, Overall Rating		
Performance	Plan does not have a "consistently low-performing" icon on Medicare Plan Finder	Applied at the contract level	Star Ratings, Low Performing Contracts Tab	2019 (Fall 2018 release)	2020 (Fall 2019 release)
Performance	Organization offering plan is not under sanction by CMS	Not applied	Star Ratings, Summary Rating Tab, Sanction Deduction (column G)	N/A	N/A
Performance	Organization offering plan is not an outlier in CMS's Past Performance Review	Not applied	Past Performance Review Outlier Results	N/A	N/A

SOURCE: All data sources for the eligibility criteria are publicly available on CMS's website (CMS, 2021a).
NOTES: C = chronic; D = dual eligible; I = institutional.
[a] We use July enrollment and contract information files since this is the time of year when enrollment generally stabilizes.
[b] The Stars performance data for a given year are released in two files: spring of the rating year and fall of the previous year (so the 2019 Star Ratings data were released in April 2019 and November 2018). The purpose of the fall release is so that the data can be used on the Medicare Plan Finder for open enrollment, which occurs in the fall of every year prior to the plan year beginning. We use the fall release since that was available to the plans at the time of their application.

Analytic Sample

After applying the eligibility criteria, we made several additional exclusions to achieve our analytic sample used for the entropy balancing and subsequent analyses. These exclusions are as follows:

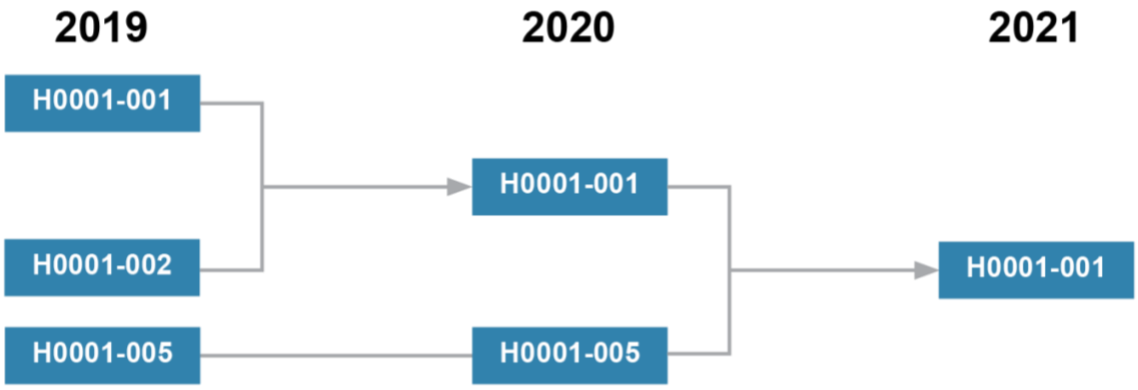
- ***End-stage renal disease (ESRD) SNPs***: We excluded ESRD C-SNPs for similar reasons that we excluded the I-SNPs—the beneficiary populations were very different and there were no participating ESRD C-SNPs in 2020 or 2021.
- ***Previous MA VBID model participants***: We excluded plans that previously participated in VBID from 2017 to 2019 and are no longer participating. These plans' decisions not to participate in 2020 align with the first year of participation for this evaluation, and we would be unable to disentangle the effect due to participation in 2020 from the effect of plans no longer participating if they were included.
- ***Part B-only***: Some MA plans offer Part B services only (no Part A or D), and we excluded these plans as there are no Part B-only VBID participants.
- ***1876 Cost plans***: There were several eligible plans that had previously been 1876 Cost plans but transitioned during the pre-period into Coordinated Care Plans. We excluded these plans because 1876 Cost plans are not eligible for the model test.

For the comparison pool only, we excluded plans that ceased operation in 2020 or 2021 and therefore have no post-VBID implementation data. We did not make a similar exclusion for participating plans. Additionally, we kept new plans with at least one year of pre-period data for analysis for both the participating and eligible nonparticipating groups. We discuss these issues with varying lengths of observation time in Appendix C.

Plan Eligibility Over Time

POs are allowed to change their contract and plan identification (ID) numbers over time for administrative, merger and acquisition, or other reasons. POs may create a new plan number and transition all beneficiaries automatically to that new plan, consolidate beneficiaries from several plans into one existing plan, or split beneficiaries from one plan into multiple plans. Since we are following plans over time, the eligible plans in each year need to be crosswalked to their counterparts in previous and subsequent years.[2] Figure D.1 shows an example of how plans can be crosswalked together. The H-numbers are the contract IDs (e.g., H0001), and the three numbers after the dash are the plan numbers (e.g., -001).

Figure D.1. Example of Plan Crosswalking



For our analyses, we crosswalked plans to their 2021 contract-plan ID since that is the latest year of data used for the quantitative analyses for this report. We aggregated variables across crosswalked plans in a given year (e.g., H0001-001 and H0001-002 in 2019) using one of the following methods:

[2] We use the CMS crosswalking files (CMS, 2021b).

- ***Participation status***: If at least one plan was a VBID model test participant, we assigned all crosswalked plans to participant status.
    - We made one exception to this rule for two nonparticipants with large enrollment in 2020 crosswalked to two participating plans in 2021 with much smaller enrollment. Because the majority of the beneficiaries in this crosswalked plan were not exposed to the intervention in 2020, we removed these plans from the analysis for 2020. Both plans contribute data to the 2021 participating plan sample.
- ***Comparison plans***: We identify comparison plans as those plans eligible in either 2020 or 2021. We restrict analyses to plans with at least one year of pre-period data (2019 for the 2020 participant analysis; 2020 for the 2021 analysis).
- ***Intervention description and participation in other initiatives flags***: These flags were aggregated the same way as the model test participation status flag, where if at least one plan had the intervention flag, the crosswalked plans were assigned the flag, within the given year.
- ***All other variables***: Other plan characteristic variables such as average risk score or OOP maximum were aggregated to the crosswalked-plan level using the enrollment-weighted mean.

In summary, the sample of 145 participating plans in 2020 was reduced to 140 plans for the descriptive tables that include the MA-only plans, and 128 plans for the analytic sample for the BDI analyses without the MA-only plans (see Appendix C for details). The sample of 418 participating plans for 2021 was reduced to 415 for descriptive tables and to 346 for regression analyses of BDI that exclude the MA-only plans (see Appendix C for details). Fifty-two hospice plans were used in the descriptive tables, and 46 were used in the regression analyses for hospice plans.

The comparison group for descriptive tables was 2,346 plans in either 2020 or 2021.

Parent Organization and Plan Characteristics

This section describes the PO and plan characteristics that are summarized in Chapter 2. Table D.2 compares characteristics of POs that participated in the VBID model test and eligible POs that did not participate in the model test.

Table D.2. Characteristics of Participating Parent Organizations

Characteristic	2020 Participating BDI	2020 Eligible Non-participating	2021 Participating BDI	2021 Participating Hospice	2021 Eligible Non-participating
Number of POs	14	103	14	9	103
% BCBS affiliate	21.4	21.4	28.6	22.2	21.4
% state	64.3**	89.3	71.4*	66.7**	89.3
% regional	7.14	6.80	7.14	11.1	6.80
% national	28.6***	3.88	21.4***	22.2**	3.88
% for-profit	50.0	37.9	50.0	44.4	38.8
Eligible plan enrollment in PO, mean (SD)	835,093*** (1,517,547)	63,279 (164,683)	854,362*** (1,657,765)	643,822*** (1,372,396)	62,068 (153,869)
MA penetration rate of counties in PO's service area, mean (SD)	52.1*** (13.9)	42.9 (10.3)	54.6*** (13.0)	54.1** (15.7)	45.9 (10.6)
Median income of counties in PO's service area, mean (SD)	29,041** (6,795)	31,960 (4,883)	28,913* (6,830)	28,306* (8,910)	31,747 (5,185)

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

NOTES: Eligible NPPOs are those eligible in either year. For 2021, four POs participated in both the BDI and the Hospice components; therefore, the numbers do not add up to 19 total participating POs. Full variable definitions can be found in Appendix C. Significant differences between the participating groups and the eligible nonparticipating groups are indicated with *** if $p < 0.01$, ** if $p < 0.05$, and * if $p < 0.1$.

Table D.3 compares characteristics of plans that participated in the VBID model test and eligible plans that did not participate in the model test.

Table D.3. Characteristics of Participating Plans

Characteristic	2020 Participating BDI	2020 Eligible Non-participating	2021 Participating BDI	2021 Participating Hospice	2021 Eligible Non-participating
Number of plans	140	2,433	376	52	2,434
% offering Part D	97.9**	91.9	99.2***	94.2	92.0
% plans that are D-SNPs	27.1***	10.0	38.2***	28.9***	10.0
% plans with $0 premium	45.0	47.5	33.8***	69.2***	48.9
Total monthly premium ($), mean (SD)	23.39* (39.75)	30.41 (46.71)	25.90* (32.85)	19.74 (37.85)	30.10 (46.50)
Average OOP maximum ($), mean (SD)	5,338** (1,466)	4,995 (1,590)	5,603*** (1,749)	4,388*** (1,599)	5,213 (1,813)
Plan % rural counties in service area, mean (SD)	5.50 (9.49)	6.62 (10.14)	7.26 (9.94)	6.56 (12.05)	6.96 (9.98)
Plan % suburban counties in service area, mean (SD)	16.3 (13.4)	17.1 (15.3)	18.1 (12.7)	12.4*** (8.8)	17.9 (15.0)
Plan % urban counties in service area, mean (SD)	78.2 (20.5)	76.3 (21.4)	74.7 (19.0)	81.1** (18.8)	75.2 (21.3)
Plan % dual beneficiaries, mean (SD)	37.6*** (39.0)	22.8 (29.7)	47.7*** (41.9)	33.5*** (42.4)	22.6 (29.8)
Plan % LIS-eligible beneficiaries, mean (SD)	37.9*** (36.2)	27.4 (29.6)	49.7*** (40.0)	12.3*** (22.9)	26.4 (29.7)
Plan average age, mean (SD)	70.1*** (4.66)	71.0 (3.95)	68.9*** (4.77)	72.7** (3.64)	71.3 (3.94)
Plan % male, mean (SD)	43.0*** (6.44)	46.2 (8.13)	42.4*** (5.98)	45.1 (7.97)	46.4 (7.56)
Plan % White, non-Hispanic, mean (SD)	63.8** (27.2)	68.2 (25.4)	63.6*** (26.0)	33.1*** (35.5)	68.6 (25.0)

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

NOTE: The participating sample size for 2020 was 145 plans, but we crosswalked and assigned plans their 2021 identifier, which consolidated the 2020 sample size to 140 plans. The full BDI sample size is 377, but one plan is missing enrollment data and other characteristics. The full eligible but not participating group sample is 2,346, but there are several plans missing enrollment and other data as well.

Significant differences between the participating groups and the eligible nonparticipating groups are indicated with *** if $p < 0.01$, ** if $p < 0.05$, and * if $p < 0.1$.

The percentage of LIS-eligible beneficiaries in the Hospice plans is particularly low because many of the participating plans are in Puerto Rico, where there is no Part D LIS.

Appendix E. Intervention Summaries

This appendix summarizes the interventions of participating POs based on the information presented in their model test applications and documentation from the implementation contractor. PO names are deidentified to protect their confidentiality and are carried over from our previous evaluation reports. Missing letters are for POs that are no longer in the model test.

In each table, the target group and year of the intervention are listed if the PO changed or added interventions between years. Then the intervention components are listed. All beneficiaries in a plan are eligible for the Cash or Monetary Rebates (called "Cash Rebates" in this report) and Hospice components, though some plans may target certain beneficiaries for pre-hospice palliative care programs (targeting criteria are noted in the intervention tables).

WHP services are described below the intervention descriptions; WHP was a requirement for this model test. Many POs interpreted the WHP component as encouraging beneficiaries to complete advance care planning (ACP) documents, detailing the types of medical treatment a beneficiary would like to receive upon becoming incapacitated. Many POs also relied on existing Medicare modalities for encouraging these discussions, which include annual wellness visits (AWVs) and health risk assessments (HRAs). AWVs, which are distinct from annual physicals, are designed to create personalized disease prevention plans and ACP documents, screen for diseases, and update medications. HRAs are often included as part of AWVs. There is no standard form, but they generally include questions such as self-rated health status, functional status items (e.g., falling, trouble walking, or difficulty urinating), depression screening questions, and items related to social needs such as transportation or housing.

PO B

PO B participated in the model test in both 2020 and 2021, offering VBID Flexibility interventions (Table E.1). PO B also participated in the previous VBID model test that ran from 2017 to 2019.

Table E.1. BDI Subcomponents Offered by PO B

Target Group	Subcomponent	Benefit[s]	Detail
Beneficiaries with COPD and/or diabetes	VBID Flexibilities	Supplemental benefits: reduced co-pays for transportation	$5 per trip for 48 trips (double the standard benefit of $10 per trip for 24 trips)
		Supplemental benefits: no co-pays for select dental services	Periodontal provided as part of routine visits and scaling/planning, and 4 lifetime periodontal procedures
		Reduced cost sharing: reduced coinsurance for diabetic testing supplies and retinal/fundus photography for diabetics	—
		High-value providers: reduced co-pays	Up to 4 high-value provider specialist visits[b]

SOURCE: RAND analysis of VBID model test application materials.
[a] Participation requirement: quarterly contact with care management team through a variety of mechanisms (phone, mail, etc.), unless beneficiary's condition is well managed.
[b] All specialists within certain specialty types that treat diabetes or COPD are eligible for the reduced co-pays (endocrinology, ophthalmology, nephrology, pulmonology, and podiatry).

WHP. These services are delivered by telephone or in person via the AWV and regular care management programs. PO B conducts outreach about these services for both beneficiaries and providers. PO B offers a $25 gift card each year for beneficiaries who engage in WHP (for 2021 and 2022). Beneficiaries must complete their AWV within the last year to be eligible to receive this incentive. There are no WHP rewards for providers.

PO C

PO C participated in the model test in both 2020 and 2021, offering BDI interventions that targeted two sets of beneficiaries (Table E.2). PO C also participated in the previous VBID model test that ran from 2017 to 2019.

Table E.2. BDI Subcomponents Offered by PO C

Target Group	Subcomponent	Benefit	Detail
Beneficiaries who qualify for LIS 1–4	VBID Flexibilities	Reduced cost sharing: outpatient mental health visits	All outpatient mental health visit co-pays are reduced 50% after beneficiary is identified as LIS
Beneficiaries with CHF and diabetes or COPD (or all 3)	VBID Flexibilities	Reduced cost sharing: Part B nebulizers	$0; do not have to participate in the RI program to receive benefit
	RI	Rewards for completing personal health review and 4 quarterly activities	Up to $200 total: $50 for personal health review; $25 for Q1, Q2, and Q3 activities; $50 for Q4 activity

SOURCE: RAND analysis of VBID model test application materials.

WHP. PO C delivers WHP services by telephone, in person, or online through the AWVs, HRAs, regular care management programs, or in-home assessments. PO C offered $20 for completing the AWV. There are no WHP rewards for providers. The wellness incentive was increased to $30 for 2021.

PO G

PO G participated in the model test in both 2020 and 2021, offering VBID Flexibility interventions (Table E.3). PO G also participated in the previous VBID model test that ran from 2017 to 2019.

Table E.3. BDI Subcomponents Offered by PO G

Target Group	Subcomponent	Benefit[a]	Detail
Beneficiaries with CHF	VBID Flexibilities	Reduced cost sharing: primary care provider (PCP) visits and cardiologist visits	$0
		Reduced cost sharing: CHF-related drugs	$0 for specified drugs[b]
		Supplemental benefits: precooked meal deliveries	$0 for up to three 14-day periods each year
		Supplemental benefits: reduced cost sharing for transportation[c]	$0 for up to 24 one-way transportation trips per year for medical appointments
	VBID Flexibilities (2021)	Supplemental benefits: medical devices	Body mass index scale and pulse oximeter
		Supplemental benefits: healthy foods allowance[c]	$25 per month for specific foods at specific retailers

SOURCE: RAND analysis of VBID model test application materials.
[a] Eligible beneficiaries are required to engage with a care management team via regular calls or visits, quarterly visits to a PCP, and an annual visit to a cardiologist to receive reduced co-pays or meal benefits. Plans with transportation benefits have participation requirements as well.
[b] Angiotensin-converting enzyme (ACE) inhibitors, angiotensin receptor blockers (ARBs), hydralazine and isosorbide dinitrate, and mineralocorticoid receptor antagonists, for all Part D benefit phases.
[c] Not available in all plans.

WHP. PO G delivers WHP services in person or by telephone through AWV, regular care management programs, or in-home assessments. PO G administrators target beneficiaries in care management for these services. PO G does not offer WHP rewards to beneficiaries or providers.

PO J

PO J participated in the model test in both 2020 and 2021, offering VBID Flexibility interventions (Table E.4). PO J also participated in the previous VBID model test in 2018 and 2019 only.

Table E.4. BDI Subcomponents Offered by PO J

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with CHF and/or coronary artery disease (CAD)	VBID Flexibilities	Reduced cost sharing: prescription drugs used to treat CHF and CAD	$0 cost sharing occurs through all Part D benefit phases except deductible • CHF medications include ACE/ARBs, beta-blockers, diuretics, and vasodilators • CAD medications include antiplatelet drugs; statins; ACE/ARBs and beta-blockers in Tiers 1–4

SOURCE: RAND analysis of VBID model test application materials.

WHP. PO J delivers WHP services by mail, by telephone, or online through a vendor that provides multiple mailings throughout the year to inform beneficiaries about filling out ACPs. PO J also has an ACP program to document and communicate wishes about end-of-life medical decisions and to help beneficiaries discuss end-of-life planning with providers and family members. Beneficiaries in a VBID-participating plan who complete the ACP program receive a onetime $25 incentive (available in each year of the model test). There are no WHP rewards for providers. There was an initial health assessment questionnaire for the WHP component in 2020, but this was removed from the WHP component for 2021. In 2020, the program was targeted to VBID members but in 2021 was rolled out to all members.

PO L

PO L participated in the model test in both 2020 and 2021, offering BDI interventions targeted to multiple subgroups of beneficiaries (Table E.5).

Table E.5. BDI Subcomponents Offered by PO L

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with dementia (2020)	VBID Flexibilities	Supplemental benefits: reduced cost sharing on in-home care	Up to 8 hours per month for $0 co-payment
Beneficiaries not adherent to diabetes, hypertension, or cholesterol medications (2020)	RI	Rewards for successful engagement (e.g., telephonic consultation for adherence counseling and to address barriers) in a medication management program	$50 for two successful engagements; annual maximum of $150 per member if member takes medications for all three conditions
Beneficiaries with LIS levels 1–4 (2021)	VBID Flexibilities	Reduced cost sharing: Part D drugs[a]	$0 cost sharing for all Part D drugs in all phases, excluding the deductible
		Supplemental benefit: healthy foods allowance[a]	$25–$55 per month, depending on the plan, at specific retailers

SOURCE: RAND analysis of VBID model test application materials.
[a] Not available in all plans.

WHP. These services are delivered by telephone, in person, or online through an internal care management program that provides an extra layer of support through phone outreach from nurses and social workers on how to initiate advance care planning. PO L representatives also encourage the discussion of ACPs with PCPs. There are no WHP rewards and incentives offered to beneficiaries or providers for 2021.

PO M

PO M participated in the model test in 2021, offering the Hospice component (Table E.6).

Table E.6. Hospice Subcomponents Offered by PO M

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Primarily through provider referral, claims algorithm, or beneficiaries electing hospice	Care management, access to social and community services, 24/7 support from care team, pain and symptom management, and spiritual/emotional support	—
TCC		Access to regular medical benefits deemed appropriate	Outpatient only
Supplemental hospice benefits		None	—

SOURCE: RAND analysis of VBID model test application materials.

WHP. PO M delivers WHP services in person or by telephone through AWVs, care management programs, and nurse outreach to talk with beneficiaries about ACP. PO M also offers provider training and an AWV toolkit for screening beneficiaries for ACP needs. There are no rewards and incentives for beneficiaries. PO M incentivizes PCPs to engage in education with patients on the importance of ACP and the annual completion of an ACP document and can earn a quality bonus for an ACP measure (up to $1,400 per year).

PO N

PO N participated in the model test in both 2020 and 2021, offering VBID Flexibility interventions to several targeted subgroups (Table E.7).

Table E.7. BDI Subcomponents Offered by PO N

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with diabetes (identified with both diagnosis code and fill of a specific diabetes medication) AND are eligible for LIS level 1 or 2	VBID Flexibilities	Reduced cost sharing: eligible medications	$0 for 90-day supply of eligible antihypertensive, statin, and antidiabetic medications
Beneficiaries with diabetes (identified with fill of specific diabetes medication) except if the beneficiary was taking metformin[a]	RI	Reward for quarterly check-in with the medication adherence program	$15 per quarterly check-in; maximum $60 per year
Beneficiaries with LIS levels 1–3 (2021)	VBID Flexibilities	Supplemental benefits: healthy foods allowance	$190 per quarter via card that can be used at selected outlets; maximum $760 per year

SOURCE: RAND analysis of VBID model test application materials.
[a] If beneficiary is taking metformin, a diagnosis code is required since metformin can be used to treat prediabetes and other conditions.

WHP. These services are delivered in person, by telephone, or online through AWVs, HRAs, care management programs, or in-home assessments. New beneficiaries are onboarded and educated regarding incentives for completing an AWV, and existing beneficiaries receive mailed communication, outreach phone calls, and reminders from staff. PO N offers $25 for completing the AWV and $15 for completing the in-home assessment. There are no WHP rewards offered to providers.

PO O

PO O participated in the model test in both 2020 and 2021, offering BDI interventions to several targeted subgroups (Table E.8).

Table E.8. BDI Subcomponents Offered by PO O

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with multiple chronic conditions (at least 6 of a specific set)[a]	VBID Flexibilities (2020)	Reduced cost sharing: outpatient cardiac/intensive cardiac rehabilitation[b]	$0
		Reduced cost sharing: outpatient pulmonary rehabilitation[b]	$0
		Reduced cost sharing: in-home visits with designated provider[c]	$0
	VBID Flexibilities (2021)	Reduced cost sharing: specialist visits[c]	$0 co-pay specialist visits for first three specialist visits in a year
	RI	Reward for engaging with designated in-home provider for the first time	$10 reward; beneficiaries already engaged with provider do not receive the incentive

SOURCE: RAND analysis of VBID model test application materials.
[a] Beneficiaries must live in specific counties.
[b] PO O decided to roll out these benefits to all beneficiaries in these plans for 2021.
[c] Participation requirements: Beneficiary must "engage" with the designated in-home provider to receive this benefit, which means allowing the provider to conduct an initial in-home visit (this was modified during the COVID-19 pandemic to include a telephonic visit).

WHP. PO O delivers WHP services in person or online through a variety of platforms including AWVs, HRAs, care management programs, or in-home assessments. There is a $10 incentive for beneficiaries to complete the ACP. There are no WHP rewards for providers.

PO P

PO P participated in the model test in both 2020 and 2021, offering the BDI component (Table E.9) and Hospice component (Table E.10).

Table E.9. BDI Subcomponents Offered by PO P

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with LIS levels 1–4	VBID Flexibilities	Supplemental benefits: healthy foods allowance	$25–$100 monthly (up from max of $50 for 2020[a] and $75 for 2021) depending on the plan, food card for purchases at national chain grocery stores; funds expire at the end of each month
Beneficiaries eligible for a medication therapy management program	RI	Reward for completion of interactive medication review (review of current medications highlights potential problematic medication use)	$25 gift card for completion of interactive medication review and $25 for medication adherence consultation
Beneficiaries with COPD using a maintenance inhaler	VBID Flexibilities	Reduced cost sharing: specific Part D Tier 3 inhalers[b]	$0 for 90-day supply or $10 for 30-day supply at retail pharmacies; one free spacer for the inhalers
	RI (2021)	Reward for completion of medication management program	$75 gift card to large, national retailer after completing 3 sessions on how to take the medications properly and basic disease management

SOURCE: RAND analysis of VBID model test application materials.
[a] In 2020, some plans offered a midyear benefit enhancement by providing additional funds to ease the impacts of COVID-19.
[b] Participation requirement: participation in COPD medication/disease management program. Cost sharing amounts are applied to the deductible and initial coverage phase.

Table E.10. Hospice Subcomponents Offered by PO P

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Provider referral or claims algorithm to identify at-risk members	Telephonic or in-person support (depending on market) from an interdisciplinary care team[a]	—
TCC		31-day ramp-down of appropriate medical services	—
Supplemental hospice benefits[b]		Non–primarily health-related services	$500 for services such as bathroom grab bars, meal preparation
		Respite care	40 hours, in 8-hour increments

SOURCE: RAND analysis of VBID model test application materials.
[a] The interdisciplinary care team coordinates care between providers, social services, and other nonclinical supports.
[b] Must select in-network providers to be eligible.

WHP. These services are delivered by telephone, in person, or online through AWVs, HRAs, care management programs, and in-home assessments. PO P also offers a digital ACP tool to all beneficiaries to create an ACP document. Beneficiaries with serious

illness (identified through a claims data algorithm) are targeted for WHP outreach. There are no WHP rewards for beneficiaries or providers.

PO Q

PO Q participated in the model test in both 2020 and 2021, offering VBID Flexibilities (Table E.11).

Table E.11. BDI Subcomponents Offered by PO Q

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with LIS levels 1–4	VBID Flexibilities[a]	Reduced cost sharing: nearly all services[b]	$0
		Supplemental benefits: over-the-counter (OTC) benefit card	$200 per quarter for Q1–Q3 and then $300 for Q4 in 2020 (boosted during COVID-19, for all MA beneficiaries); $200 in 2021

SOURCE: RAND analysis of VBID model test application materials.
[a] Participation requirement: beneficiaries must select a PCP at a specific high-value provider to receive reduced cost sharing on other services.
[b] The PCP manages the beneficiary's care, similar to a care management program.

WHP. PO Q delivers WHP services by telephone or in person through a variety of platforms, including care management programs, in-home assessments, HRAs, AWVs, and collaboration with local skilled nursing facilities. PO Q educates PCPs to include WHPs as part of the AWV. It has also received community involvement through local partnerships to provide end-of-life conversation training and support. While PO Q does not offer an incentive to beneficiaries for WHP directly, it offers a $25 gift card for beneficiaries to complete their AWV. There are no WHP rewards for providers.

PO R

PO R participated in the model test in 2021, offering both the BDI (Table E.12) and Hospice components (Table E.13).

Table E.12. BDI Subcomponents Offered by PO R

Target Group	Subcomponent	Benefit	Detail
None (all beneficiaries)	Cash Rebates	Delivered through a debit card	$50 or $160 monthly ($600 or $1,920 per year), depending on the plan; unused amounts roll over each month, and funds are available for use 180 days after disenrollment or end of benefit year

SOURCE: RAND analysis of VBID model test application materials.

Table E.13. Hospice Subcomponents Offered by PO R

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Provider referral or claims analysis. Beneficiary must have life-threatening illness with less than 12 months to live, have caregiver support at home, and be in functional decline or otherwise fail to meet hospice criteria	24/7 care team support, ACP discussions, social services and community resources, psychosocial and spiritual support, pain and symptom management, medication reconciliation, caregiver support	—
TCC		Covered medical services as needed	—
Supplemental hospice benefits[a]		In-home support	One 4-hour visit per week
		Respite care and hospice drugs	No cost sharing

SOURCE: RAND analysis of VBID model test application materials.
[a] Only for beneficiaries using in-network hospice.

WHP. These services are delivered by telephone, in person, or online via AWVs, HRAs, regular care management program interactions, in-home assessments, and education and outreach to beneficiaries and providers. There are no WHP rewards for providers or beneficiaries.

PO S

PO S participated in the model test in 2021, offering VBID Flexibilities (Table E.14).

Table E.14. BDI Subcomponents Offered by PO S

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with LIS levels 1–3	VBID Flexibilities	Supplemental benefits: OTC benefit card/healthy food allowance	$145 monthly via a card that can be used to purchase OTC items or a specific list of health foods at selected retailers

SOURCE: RAND analysis of VBID model test application materials.

WHP. These services are delivered in person, by telephone, through regular mailings, or online. Examples include the AWV, HRAs, care management programs, a self-guided ACP program through a vendor-provided digital platform, or conversations with a PCP and/or specialist. There are no rewards and incentives for beneficiaries. PCPs may receive up to $20 per beneficiary for conducting and documenting ACP.

PO T

PO T participated in the model test in 2021, offering the Hospice component (Table E.15).

Table E.15. Hospice Subcomponents Offered by PO T

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Provider referral	Care management, access to social and community services, 24/7 support from care team, pain and symptom management, medication reconciliation, caregiver and spiritual/emotional support	—
TCC		All regular plan medical benefits	—
Supplemental hospice benefits[a]		Respite care and hospice drugs	No cost sharing

SOURCE: RAND analysis of VBID model test application materials.
[a] Only for beneficiaries using in-network providers.

WHP. PO T delivers WHP services in person, by telephone, or online through AWVs, HRAs, care management programs, and in-home assessments. PO T does not offer WHP rewards for beneficiaries or providers.

PO U

PO U participated in the model test in both 2020 and 2021, offering the BDI component (Table E.16).

Table E.16. BDI Subcomponents Offered by PO U

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with diabetes and at least one mental health diagnosis	RI	Reward for completion of diabetes screening activities	$10 gift card for each activity completed (HbA1c, glucose testing, foot or eye exam, medical attention for nephropathy)
All beneficiaries in plan who have not yet received specific vaccines	RI	Reward for receiving vaccines	$25 gift card for each vaccine (shingles; tetanus, diphtheria, and pertussis; hepatitis A/B; meningococcal), up to maximum of $100; amount increased to $50 for 2021 for a total of $200
Beneficiaries meeting CMS eligibility criteria for medication therapy management	RI	Reward for engaging with the comprehensive or targeted medication review	$25 per quarter for engagement; increased to $50 in 2021
Beneficiaries with fall risk[a]	VBID Flexibilities	Supplemental benefits: reduced cost sharing for comprehensive fall risk evaluation	$0

SOURCE: RAND analysis of VBID model test application materials.
[a] This is determined through provider referral, care management team referral, or a claims algorithm. The claims algorithm uses a 12-month look-back for diagnoses of repeated falls or a history of falls (R26.6 or Z91.81).

WHP. PO U delivers WHP services via telephone, online, or in person through AWVs, HRAs, care management programs, and in-home assessments, and also through education for providers and beneficiaries, including monthly mailers/emails and in-person member events. Information about ACPs is available on the plan's website. There are no rewards or incentives offered for beneficiaries or providers in 2021, although PO U did offer a $25 gift card for beneficiaries in 2020 for completion of ACP.

PO V

PO V participated in the model test in 2021, offering the Hospice component (Table E.17).

Table E.17. Hospice Subcomponents Offered by PO V

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Beneficiaries diagnosed with serious illness, hospitalized with life expectancy of 12 months or less	Comprehensive care assessments, 24/7 care team support, ACP discussions, access to social services and community resources, psychosocial and spiritual support, pain and symptom management, medication reconciliation, caregiver support	—
TCC	Beneficiaries with cancer, ESRD, or end-stage liver disease (subset of hospice-eligible). They must seek TCC, and the provider must agree with and support the treatment plan and goals of care	Treatments such as chemotherapy, blood transfusions, dialysis, and paracentesis for disease states mentioned in eligibility criteria	—
Supplemental hospice benefits[a]		Respite care and hospice drugs[a]	No cost sharing

SOURCE: RAND analysis of VBID model test application materials.
[a] Drugs included in the hospice supplemental benefit are for symptom control and pain relief.

WHP. These services are delivered via telephone, in person, or online through AWV, HRAs, regular care management programs, in-home assessments, and beneficiary education through a partner provider. Providers are trained to discuss end-of-life care planning. The plan uses proactive outreach efforts in ambulatory clinics, email campaigns, and other media channels. PO V does not offer WHP incentives for beneficiaries or providers.

PO W

PO W participated in the model test in 2021, offering both the BDI (Table E.18) and Hospice components (Table E.19).

Table E.18. BDI Subcomponents Offered by PO W

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with diabetes and/or CHF and recent hospital visits (either emergency department [ED] or inpatient)	VBID Flexibilities	Supplemental benefits: new technologies—medical devices	Beneficiaries with diabetes can receive a continuous glucose monitoring device (requires provider monitoring). Beneficiaries with CHF can receive a remote patient monitoring device (requires provider monitoring)
	RI	Rewards for a variety of screenings, specialist appointments, and care management activities	Maximum $130 per year, plus $20 for ACP completion, for total possible reward of $150; screenings must be completed at PO's one-stop-shop clinics[a]
None (all beneficiaries)	Cash Rebates	Delivered through a debit card	$75 or $130 per month ($900 or $1,560 per year) depending on plan, which can be used for purchases or cash withdrawals

SOURCE: RAND analysis of VBID model test application materials.
[a] These are multidisciplinary clinics with additional nonclinical staff like social workers and nutritionists.

Table E.19. Hospice Subcomponents Offered by PO W

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Beneficiaries with a diagnosis of a life-threatening illness and a prognosis of less than 6 months to live, trouble with a variety of functional scales specified in application, and a caregiver at home; also identified through claims data algorithm	24/7 care team support, ACP discussions, social services and community resources, psychosocial and spiritual support, pain and symptom management, medication reconciliation, caregiver support	—
TCC[a]		Services will be identified as appropriate, reflective, and based on enrollees' (and/or caregivers') needs and preferences as identified and documented in the plan of care developed by the care management's interdisciplinary team	—
Supplemental hospice benefits		None	—

SOURCE: RAND analysis of VBID model test application materials.
[a] Only for beneficiaries using in-network hospice.

WHP. PO W delivers WHP services via telephone, in person, or online through AWVs, HRAs, care management programs, and in-home assessments. Additionally, the PO creates individualized care plans for all beneficiaries, which include WHP care plans that

are updated during regular provider visits. For 2021, some beneficiaries enrolled in participating plans are eligible to receive a $20 gift card for WHP activities. PO W does not offer WHP rewards for providers.

PO X

PO X participated in the model test in 2021, offering the Hospice component (Table E.20).

Table E.20. Hospice Subcomponents Offered by PO X

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Advanced illness management algorithm to predict those who may die within 12 months, or provider referral	24/7 care team support, ACP discussions, social services and community resources, psychosocial and spiritual support, pain and symptom management, medication reconciliation, caregiver support	—
TCC		Limited to specific medical services for cancer, cardiac-related conditions, dementia, respiratory-related conditions, or chronic kidney disease	Services provided for up to one month after hospice election.
Supplemental hospice benefits		Respite care and hospice drugs	No cost sharing

SOURCE: RAND analysis of VBID model test application materials.

WHP. PO X delivers WHP in person through regular care management programs and in-home assessments or by telephone through monthly outreach by care managers. The PO has a rewards program outside of VBID in which a beneficiary may earn a reward for completing an ACP.

PO Y

PO Y participated in the model test in 2021, offering both the BDI (Table E.21) and Hospice components (Table E.22).

Table E.21. BDI Subcomponents Offered by PO Y

Target Group	Subcomponent	Benefit	Detail
Beneficiaries with diabetes[a]	VBID Flexibilities	Reduced cost sharing: diabetes drugs	Tier 2 co-pays for diabetes drugs: $0 for 30-day supply; Tier 3 oral, noninsulin drugs: $25 co-pay for 30-day supply[b]
	RI	Rewards for completing diabetic screenings	$10 for completing each of three diabetic screenings (HbA1c, nephropathy, eye exam); $30 maximum per year

SOURCE: RAND analysis of VBID model test application materials.
[a] Participation requirement: beneficiaries must participate in a care management program.
[b] Plans vary in regard to which phases of the Part D benefit the reduced cost sharing applies.

Table E.22. Hospice Subcomponents Offered by PO Y

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Provider referral or claims-based algorithm	24/7 care team support, ACP discussions, social services and community resources, psychosocial and spiritual support, pain and symptom management, medication reconciliation, caregiver support	Available in three major care settings (inpatient, home, and clinic services)
TCC		Medical services for ESRD, oncology, infusion therapies, pulmonary, liver disease, rheumatology, and rehabilitation services	Limits vary by the specific service (e.g., limiting dialysis to continuing for a max of 30 days after the beneficiary elects hospice)
Supplemental hospice benefits		Respite care and hospice drugs	No cost sharing
		Safety modifications	Home and bathroom safety devices

Subcomponent	Eligibility	Benefit	Detail
		Meal support	One meal delivered per day to the member's home for a maximum of 60 meals
		Transportation	For ongoing hospice care occurring outside the member's home

SOURCE: RAND analysis of VBID model test application materials.

WHP. PO Y delivers WHP services via telephone or in person through several mechanisms, including regular care management programs, and ongoing education and outreach for beneficiaries and providers. There are no WHP rewards for beneficiaries or providers.

PO Z

PO Z participated in the model test in 2021, offering the Hospice component (Table E.23).

Table E.23. Hospice Subcomponents Offered by PO Z

Subcomponent	Eligibility	Benefit	Detail
Palliative care	Provider referral or claims algorithm	Comprehensive care assessments, access to social and community resources, 24/7 care team support, pain/symptom management, medication reconciliation, caregiver support	—
TCC		Radiation and enteral nutrition therapy, cancer curative therapies	—
Supplemental hospice benefits		Respite care	No cost sharing and an increase of 2 days for the benefit (to maximum of 7 days)
		Hospice drugs	No cost sharing
		Enteral/parenteral formula	No cost sharing

SOURCE: RAND analysis of VBID model test application materials.

WHP. These services are delivered by telephone, in person, or online through the AWV or the regular care management programs. Care managers engage beneficiaries in ACP discussions, coordinate directly with providers, or refer beneficiaries to their

PCP for ACP discussions. All beneficiaries are eligible to receive an incentive for WHP as part of an existing rewards program outside of the VBID model test. Beneficiaries can earn points for completing a PCP visit in which ACPs are discussed. There is no WHP incentive for providers.

PO AA

PO AA participated in the model test in 2020, offering the BDI component (Table E.24).

Table E.24. BDI Subcomponents Offered by PO AA

Target Group	Subcomponent	Benefit	Detail
Beneficiaries eligible for 100% LIS premium subsidy with two or more hospital admissions	RI	Reward for completion of tailored health intervention as part of a care management intervention	$100 gift card per quarter; $400 per year maximum

SOURCE: RAND analysis of VBID model test application materials.

WHP. PO AA delivers WHP services through HRAs, regular care management interactions, in-home assessments, and regular mailings. An algorithm is used to identify beneficiaries with serious illness who receive targeted WHP outreach through the care management program. Beneficiaries receive rewards of up to $165 for completing wellness activities through an existing rewards program; rewards can be redeemed from a catalog of items.

PO AB

PO AB participated in the model test in 2020, offering the BDI component (Table E.25).

Table E.25. BDI Subcomponents Offered by PO AB

Target Group	Subcomponent	Benefit	Detail
Beneficiaries eligible for LIS 1–4	VBID Flexibilities[a]	Supplemental benefits: transportation and meal support	$0 transportation (up to 48 trips/year) and $0 meals (21 meals over each 2-week occurrence, up to 84 per year)
Beneficiaries with diabetes, hypertension, or CAD who are nonadherent to at least one medication for these conditions	RI	Reward for completion of activities	$25 for social needs assessment, $5 quarterly for disease management education, $10 for completing "learn and earn" and confirming medication has been taken as prescribed; maximum of $150 annually

SOURCE: RAND analysis of VBID model test application materials.
[a] Participation requirement: beneficiaries must call PO to complete social needs assessment.

WHP. All MA enrollees receive information about advance directives in regular plan mailings. The plan requires network PCPs to document existing advance directives in the enrollee medical records. PO AB does not offer WHP rewards to beneficiaries or providers.

PO AQ

PO AQ participated in the model test in 2020, offering the BDI component (Table E.26).

Table E.26. BDI Subcomponents Offered by PO AQ

Subcomponent	Target Group	Benefit	Detail
RI	Beneficiaries with specific chronic conditions and at least $700 in total monthly drug spending	Reward for engaging in telephonic educational interventions	$10 per quarter incentive, up to four times per year; gift card is sent at the end of the year

SOURCE: RAND analysis of VBID model test application materials.

WHP. PO AQ delivers WHP services through HRAs and other ongoing modalities such as in-home assessments, regular mail, and telephone outreach. There is a $30 incentive for providers to conduct ACP discussions.

Appendix F. Perceptions of the VBID Model Test Intervention Components Among Parent Organizations That Did Not Implement Them

During the interviews with participating and nonparticipating POs, we asked PO representatives questions about their general perceptions of model test components. In this appendix, we present the perceptions of all VBID nonparticipants and those model test participants that have not implemented a particular model test component or subcomponent. In discussing pros and cons of each model test component and subcomponent, PO representatives focused primarily on the extent to which each component or subcomponent could help them address the needs of their beneficiaries, the intended outcomes of each component or subcomponent and its alignment with their organizational priorities, alternative avenues for offering the same benefit outside of VBID, and required implementation resources.

Benefit Design Innovations

VBID Flexibilities

Participating and nonparticipating PO representatives we interviewed generally agreed that additional supplemental benefits and reduced cost sharing for Part C and D benefits can help them provide more comprehensive, "whole person" care to their beneficiaries. A PO G representative stated that they try to "look at [this model] holistically and think about, based on the experience that we've had and what we're hearing from our members, what we can do to enhance the program to help them on their path to better health." Benefits of particular value to their beneficiaries, especially those with a low SES, include meal and food benefits, transportation, and $0 cost sharing for specialist visits. These benefits can help POs address their low-income members' most pressing social determinants of health and could be offered in a targeted way to them only as part of the VBID model test.

At the same time, both participating and nonparticipating PO representatives raised a number of concerns related to the implementation of the VBID Flexibilities. First, some felt that *reduced cost sharing benefits and high-value providers were not relevant to the needs of beneficiaries in their plans*—especially those in D-SNPs, because these plans already have narrower provider networks and offer $0 or low co-pays for Part C and D benefits. As an NPPO H representative explained,

> [F]or dual members, we don't willy-nilly contract with every provider in the system, because we know they are not equipped to delivering care for these dual beneficiaries. For our dual beneficiaries, we have a subset of our providers who

> we feel confident can provide the care that the duals need . . . 90 percent of our people are on $0 meds. Almost close to 100 percent of our people have $0 premiums. Almost close to 95 percent of our folks have $0 PCP co-pay/specialty co-pay/rehab co-pay. So it didn't make any sense for us.

High-value provider interventions were often considered problematic by our interviewees because of concerns related to network adequacy requirements. As an NPPO D representative explained, "There aren't a lot of endocrinologists to choose from." Indeed, PO B, which implemented a high-value provider intervention, had to adjust its design: It now considers all its in-network specialists who provide care to beneficiaries with diabetes or COPD, including endocrinologists, ophthalmologists, nephrologists, pulmonologists, and podiatrists, to be high-value providers.

Moreover, *tailoring of benefits within plans based on beneficiary chronic conditions would be resource-intensive and burdensome and could lead to beneficiary confusion*. "We would always have to be very cognizant of operational burden and things of that nature, with being a limited size plan with limited capabilities," said an NPPO G representative. NPPO B representatives were particularly concerned about stratifying beneficiaries by their health condition and offering them different benefits, because this "would be a departure from current practice." PO L and N representatives argued that administering benefits to all beneficiaries within a plan would be easier and less likely to confuse beneficiaries. Because POs want to ensure positive member experiences, representatives of NPPO B and C reported not wanting to implement VBID Flexibilities. "It's kind of almost a slippery slope of adding more targeted benefits for chronic conditions, because we realize it's going to be confusing for the members. . . . [We also need to make] sure that our benefits are very clear, that members really know what's eligible to them and what does that mean to them," said an NPPO B representative.

Rewards and Incentives

PO representatives had mixed feelings about the RI component. Some felt that ***the RI interventions could help increase beneficiary engagement in managing their own health***. Offering beneficiaries gift cards for engaging with care managers could be a stronger incentive than reducing cost sharing and could help encourage the use of preventive care (NPPO A and C). "In general, I think it's a great idea," said an NPPO C representative. "[I]t definitely drives members to get the care they need."

At the same time, representatives from two POs (G and M) and four NPPOs (B, D, F, and G) indicated that ***they offer similar benefits outside of the model test***. "I don't really understand how it differs tremendously from a rewards and incentives that we offer to our consumers through our regular plan, unless it is, again, targeted towards disease state or targeted towards low-income status," said an NPPO F representative. Moreover, NPPO E and PO G representatives stated ***they did not see the value in enhancing their rewards and incentives to the maximum model-test-allowed amount of $600 per beneficiary annually***. "We are doing a

good job through a combination of copay waiver, copay structure and benefits and care management support and don't need to add a rewards and incentives necessarily just to have one," said a PO G representative.

Cash Rebates

Our interviewees were equally divided in their perceptions of cash rebates. A key advantage of cash rebates described during the interviewees was their ***perceived simplicity for beneficiaries***: "Some kind of monetary incentive to me seems very simple, rather than trying to add on some kind of fluffy benefit of some kind. I think people would appreciate that more," said an NPPO D representative. Similar to the two POs that implemented rebates (POs R and W), an NPPO F representative felt ***that rebates could give beneficiaries more financial freedom to address their needs***: "It may be more important to them to be able to go buy a new pair of shoes or to be able to buy a small gift for their grandchild . . . to pay their utility bill, or to buy a new dress to go to church. And that, while they are not always health related, what that does for someone's mental state of mind is huge."

However, others raised a number of concerns. First, representatives from NPPOs B and H indicated that they ***preferred offering additional supplemental benefits instead of cash rebates***, which are funded by the same rebate dollars POs could receive by bidding below the benchmark, because they were a major selling point. "We feel really good about our supplemental benefits, so much so that we pride ourselves on what we say, we kind of lead almost with our supplemental benefits" (NPPO B).

Second, several interviewees worried about ***beneficiary concerns*** related to questions about why they received cash from an insurance company and possible negative tax implications (NPPO C and PO P). A PO P representative stated that

> [f]rom a premium rebate perspective, there are always concerns around the tax implications. I think that's been the biggest barrier there. What we certainly don't want to do is provide the rebates directly to members and then have them experience negative tax consequences because of something that was intended to provide them with greater financial flexibility and greater self-direction.

Although some felt that cash rebates might be easy to administer because all beneficiaries in a plan would be receiving them, others, like NPPO E representatives, were concerned about ***the administrative lift of providing very modest rebates***:

> The administrative cost of doing [rebates] . . . may not be worth the effort [for us], as opposed to adjusting benefits to capture the rebate and have it show up in terms of our final benefits. . . .My previous experience has been that the [potential rebate amounts] were somewhere $3, $5, $7, PMPM. That's not probably going to be very easy to administer nor have a really large impact to the member overall as opposed to something we could do on the benefit side.

Hospice Benefit Component

Our interviewees had somewhat mixed feelings about the Hospice component. Some felt that VBID Hospice is consistent with their organizational missions and ***has a potential to improve care coordination at the end of life and make transitions to hospice easier*** (NPPOs D and F). "If this provided the opportunity to get folks perhaps more robustly involved in palliative care earlier on, maybe with an easier transition to hospice, to me I think that would certain be a benefit" (NPPO D). Others, however, reported offering some of the Hospice component benefits, such as palliative care, outside of the model test (PO B) or cited a number of concerns and implementation challenges they expect to encounter with this model test component, especially during the COVID-19 pandemic (PO L).

First, our interviewees were concerned about their ***operational readiness to implement the Hospice component***. While some worried about having sufficient bandwidth to pursue the Hospice component (POs C and N), others worried about the complexities in identifying care that is related to the beneficiary's terminal condition (covered under the hospice benefit) and unrelated (covered under the regular medical benefit):

> I would just add that hospice is a huge opportunity for our patients, but [it is] also an area of needing ongoing attention as to paying hospice correctly and [knowing] what's hospice and not hospice in [the] Part D benefit, and so we're all journeying through that and working to keep that as clean as possible, but it also didn't quite meet that operational readiness test . . . at the time. (PO N)

The lack of clarity around operationalizing the Hospice component negatively affected POs' readiness to implement it (POs B, C, and N). A representative from PO B described the complexity of operationalizing the hospice carve-in:

> The network concerns that were there in the model . . . the questions that we're asking CMMI about how to manage the out-of-network pieces, how the data components, how the notification pieces were going to be happening, we were not too sure how that was going to be managed. So it felt like we were going to be a pass through.

A second concern for many interviewees was ***the ability to develop and manage the hospice network*** because POs have to negotiate contracts with hospices and oversee payment. Some felt that hospices might not want to participate in the model test at all (PO C). Others, like NPPO C representatives, described the need to ensure that hospices have a similar mission or share the same values prior to forming partnerships: "We would have to have such a strong partnership with that hospice organization to really make sure we're on the same mission. . . .We have [to have] the same ideas, goals, and missions in how we're going to administer this benefit."

PO S representatives worried about the financial risk their organization would undertake given the local hospice provider structure, stating that the service area is "dominated by one particular player, and the consensus from our provider team has been that we would not have good success with that."

Representatives from three NPPOs (D, H, and C) stated that they did not have a prior relationship with hospice organizations. Having a trusted relationship with a hospice provider was considered critical to implementation success by our interviewees who viewed partnering with unknown entities to be risky and prone to conflicts. According to a representative from NPPO D,

> The lynchpin here, and the fundamental assumption, is that there is a high level of trust and partnership between [POs] and providers of all specialty and all categories, which, in most cases, there is, but we all bring our own interests to the table. . . . We negotiate rates every day. And they want to get paid more, and we want to pay less. And then authorizations; the criteria that we use are based on LCDs [local coverage determinations], NCDs [national coverage determinations], and our own medical necessity criteria. And sometimes those are in conflict, and we have appeals. A huge part of our operation is working appeals from providers and members because they disagree with a decision that we've made. So just I think naturally there is this—it creates this dynamic of us versus them.

Third, some interviewees (NPPOs G, H, and C) worried about the ***administrative burden*** of collecting data and reporting data to CMS, noting that the administrative costs of operationalizing and implementing the Hospice component were large given that the intervention would affect very few beneficiaries in some of their plans. Others raised concerns about ensuring their data were provided in a CMS-approved format, which would have been overly burdensome, particularly if this reporting applied to only a small proportion of all beneficiaries. As an NPPO H representative stated,

> You have to make available I guess palliative care goals, you have to make available certain experience measures, certain process measures, on a very frequent basis. And that part would be a little bit more burdensome because we'll have to create a whole infrastructure around it. Even though we may be doing it, we are not capturing it in the [CMS-required] format. We've heard that CMS has a very specific format, and you have to conform to that format, which may mean that we may have to do something differently for a small group of patients.

Cost concerns about operationalizing the Hospice component may have been experienced differentially by smaller POs. A representative from PO O described it this way: "Like as a smaller health plan, I'll say, the cost to kind of implement some of the more advanced, for lack of a better word, offerings under VBID just wasn't feasible for us at the time. . . . I think just the hospice, in particular, I think just we weren't in a position to kind of jump into that one just the way we're positioned right now."

Wellness and Health Care Planning

Representatives from seven of nine POs we interviewed that did not participate in VBID stated that they already offer WHP or plan to do so, indicating that they do not think it would be a challenge to meet this VBID participation requirement (NPPOs H, A, D, E, F, G, and I).

Only representatives from three NPPOs worried about ***potential administrative burden*** related to operationalizing WHP within their plans (NPPO B), general resources needed to support WHP (NPPO C), or tracking WHP-related conversations or completions of end-of-life care plans (NPPO F). As a representative from NPPO F stated, "[T]he challenge is going to be tracking who accepted it and who didn't and what they did with it. That's a pretty heavy lift . . . for any plan."

Although most representatives from NPPOs described few likely WHP implementation barriers, their interpretations of what constitutes WHP differed from those of participating POs. Some NPPO representatives thought of WHP as activities that encourage the completion of ACP, living wills, and AWVs, which allow beneficiaries an opportunity to discuss their end-of-life care preferences directly with their health care provider. Others viewed WHP as care management/disease management (CM/DM) activities that emphasize healthy eating and fitness.

Like NPPOs, however, many representatives from participating POs described their organizations' strong commitment to offering WHP and ACP services prior to joining the VBID model test. Such continuity of offerings reduced the need to make major changes to implement this VBID requirement:

> We were one of the health plans that were one of the forerunners in that space [palliative care]. We had started what we called an advanced illness program years and years ago. . . . Because the Medicare Advantage team had long history of palliative care and being able to provide services in that space, [it] is comfortable having those [ACP] conversations. (PO B)

Others reported that their VBID participation has spurred them to explore new options for engaging beneficiaries and their family members in discussions around goals of care, including considering implementing new digital platforms and training programs for providers on WHP. One representative explained that their approach to WHP is

> still an ongoing process that we're looking to improve on and that we're going to have to think through as we look to add additional interventions for 2022. But we did have a really thorough conversation with the care managers to really look through our assessment and figure out where we're raising these questions and how we can tell whether or not members are taking us up on it. (PO G)

Appendix G. Analysis of Potential COVID-19 Pandemic Impacts on Evaluation Outcomes

Changes in health care utilization and other behavior brought about by the COVID-19 pandemic could pose a threat to the internal validity of our DD evaluation design if these changes happen to be correlated with VBID implementation. Such a correlation might arise if geographic areas with VBID plans had systematically different case rates from areas without VBID plans, or if changes in health care utilization due to the pandemic were different across different geographic areas. Because the start of the pandemic coincided with the first year of VBID implementation (2020), the specific concern is that the association between VBID participation and our evaluation outcomes might be confounded by changes driven by COVID.

To assess the potential for the COVID-19 pandemic to introduce bias into our evaluation design, we collected and analyzed county-level data on case rates and Medicare utilization. Case rate data were obtained from the Centers for Disease Control and Prevention ([CDC], 2022). We also constructed county-level utilization measures for the Medicare Fee-for-Service (FFS) beneficiary population that were observed during 2020, providing us with a measure of pandemic-related utilization impacts that should not be affected by VBID or other features of the MA policy environment. Under the assumption that Medicare utilization changes are driven by factors that have similar impacts on both MA and FFS beneficiaries within a county (such as beneficiary concerns about infection risk, elective procedure bans, or health care facility closures), our FFS utilization measures allow us to examine—and, ultimately, to control for—utilization changes due to the pandemic.

To provide an initial assessment of the risk that VBID interventions might be confounded by COVID-19 impacts, we compared COVID-19 case rates and FFS utilization changes between geographic areas with and without VBID participating plans present. As we discuss below, this analysis of geographic differences in COVID-19 impacts did raise some concerns that the impact of the pandemic on health care utilization may have been correlated with the service areas of plans participating in VBID. We therefore conducted sensitivity analyses of major quantitative outcomes analyzed in this report (enrollment, plan MAPD bids, total premiums) in which we added controls for the intensity of the COVID-19 pandemic to our DD regression models. We report results from these sensitivity analyses in this appendix.

Definitions and Data Sources

To assess the potential for the COVID-19 pandemic to introduce bias into our DD evaluation design, RAND analyzed county-level data on case rates and Medicare utilization.

County Definition

We classified counties with VBID-eligible plans as either counties with one or more VBID plans ("VBID counties") or counties with nonparticipating VBID-eligible plans only ("non-VBID counties") and compared COVID-19 case rates and Medicare utilization rates in each month of 2020. To measure changes in health care utilization among older adults without inadvertently capturing the impacts of VBID implementation, we focused on changes in utilization among Medicare FFS beneficiaries in each county, calculating the percentage change in utilization counts between each month of 2020 and the same month in 2019.

Table G.1 reports summary statistics on the number of counties, the average number of VBID and non-VBID enrollees, and the VBID enrollment share in VBID and non-VBID counties. We identify 2,193 counties as VBID counties (with at least one VBID plan containing 11 or more enrollees as of 2020) and 559 as non-VBID counties. Note that the non-VBID counties, as defined in this analysis, may contain small numbers of VBID enrollees who live outside their plan's service area or who are enrolled in plans with fewer than 11 enrollees in the county. In general, the non-VBID counties have lower enrollment than the VBID counties. For the VBID counties, the average VBID enrollment share (the percentage of enrollees in our analysis enrolled in VBID plans) is 37 percent.

Table G.1. County-Level Descriptive Statistics for VBID Counties and Non–VBID Counties

	VBID Counties	Non-VBID Counties
Number of counties	2,193	559
Average number of VBID plan enrollees per county	1,681	11
Average number of Non-VBID plan enrollees per county	4,641	1,335
Average county-level VBID enrollment share	37%	3%

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Enrollment figures reflect 2020 enrollment in VBID plans and nonparticipating VBID-eligible plans ("Non-VBID plans"). "Non-VBID Counties" = counties with VBID-eligible plans only. VBID plan enrollment in non-VBID counties is nonzero only if (a) county is outside that plan's contract service area or (b) county is inside contract service area but enrollment is below 11 beneficiaries. "VBID Counties" = counties in contract service area for 1+ VBID plan with 11 or more enrollees in at least 1 VBID plan. "VBID Enrollment Share" = county-level percentage of enrollees in VBID plans and nonparticipating VBID-eligible plans who are enrolled in VBID plans.

Case Rate Definition

We used data obtained from the CDC to construct county-level measures of new COVID-19 cases per capita in older adults. The CDC provided RAND with a file of deidentified data on all COVID-19 cases reported to the CDC, which included the date of diagnosis, the patient's county of residence, and the patient's age group (reported in ten-year intervals, e.g., 60–69, 70–79, 80+). To construct a case rate per capita for older adults, we calculated the total number of reported cases in adults 60 and older in each county and month (the numerator), then divided this case

count by estimates of each county's population in this age range (the denominator). County-level population denominators for adults aged 60 and over were obtained from the National Cancer Institute's SEER program.

Medicare Fee-for-Service Utilization Measure Definition

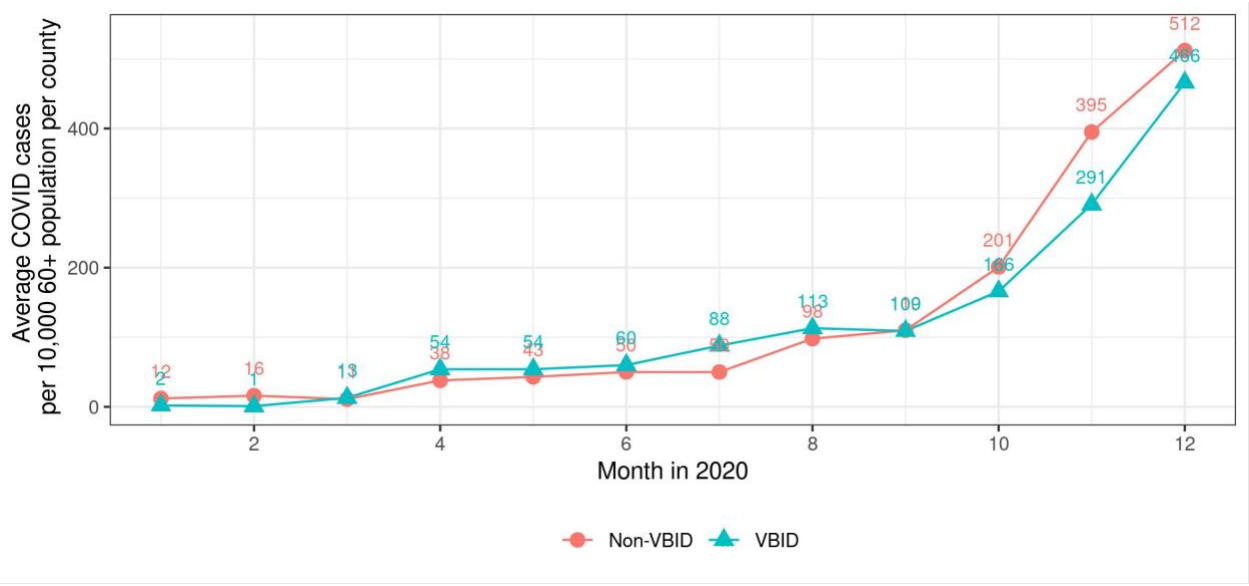
While the case rate measure introduced above can be viewed as capturing the risk of COVID-19 transmission among older adults in a county, what matters most for the VBID evaluation is the impact of the COVID-19 pandemic on health care utilization. To measure the magnitude of pandemic-related interruptions in health care utilization, we constructed a data set of year-on-year changes, at the county-month level, in the level of health care utilization among Medicare FFS beneficiaries. RAND's IDR programmers used Medicare's common working file to calculate utilization counts by county and month for each month in 2019 and 2020. We focused on inpatient stays and ED visits.

To measure utilization changes within counties attributable to the COVID-19 pandemic, we calculated the percent change in each utilization count between each month in 2019 and the same month in 2020. For example, the count of professional services visits in Montgomery County, Maryland, for June 2020 would be compared with the count of professional services visits for Montgomery County, Maryland, in June 2019. Comparing 2020 utilization with a baseline level defined by 2019 utilization provides a simple way to adjust for seasonality and population differences across counties.

Results

Figure G.1 shows the average rate of new COVID-19 cases reported to CDC in VBID and non-VBID counties for each month of 2020. COVID-19 case rates among older adults were higher in VBID counties between April and August 2020 but were higher in non-VBID counties between October and December 2020.

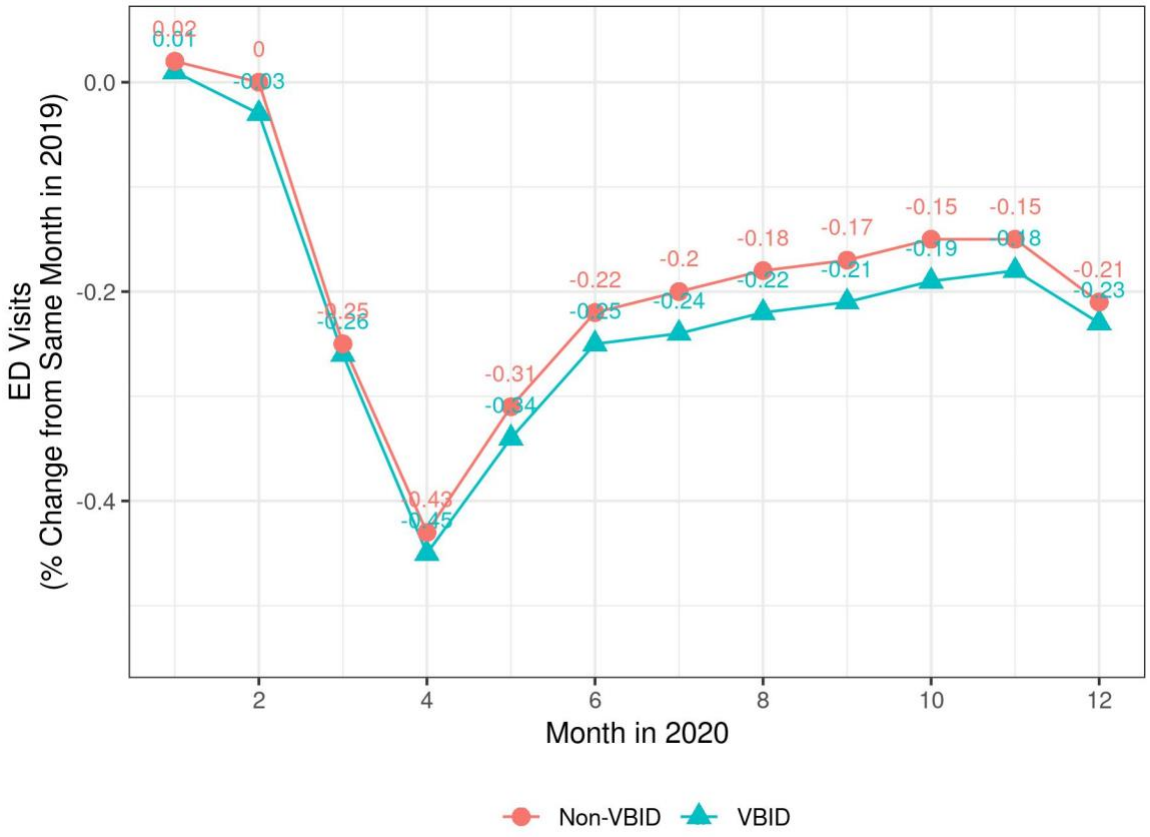
Figure G.1. Monthly New COVID-19 Case Rates per 10,000 Adults Aged 60+, by County VBID Status



SOURCE: RAND analysis of CDC COVID-19 Case Surveillance Restricted Access Detailed Data and National Cancer Institute's SEER Population Estimates.

Figure G.2 compares changes in the number of ED visits between 2019 and 2020. The two series move roughly in parallel throughout the year, but with more negative changes observed in VBID counties by 1–4 percentage points in all months.

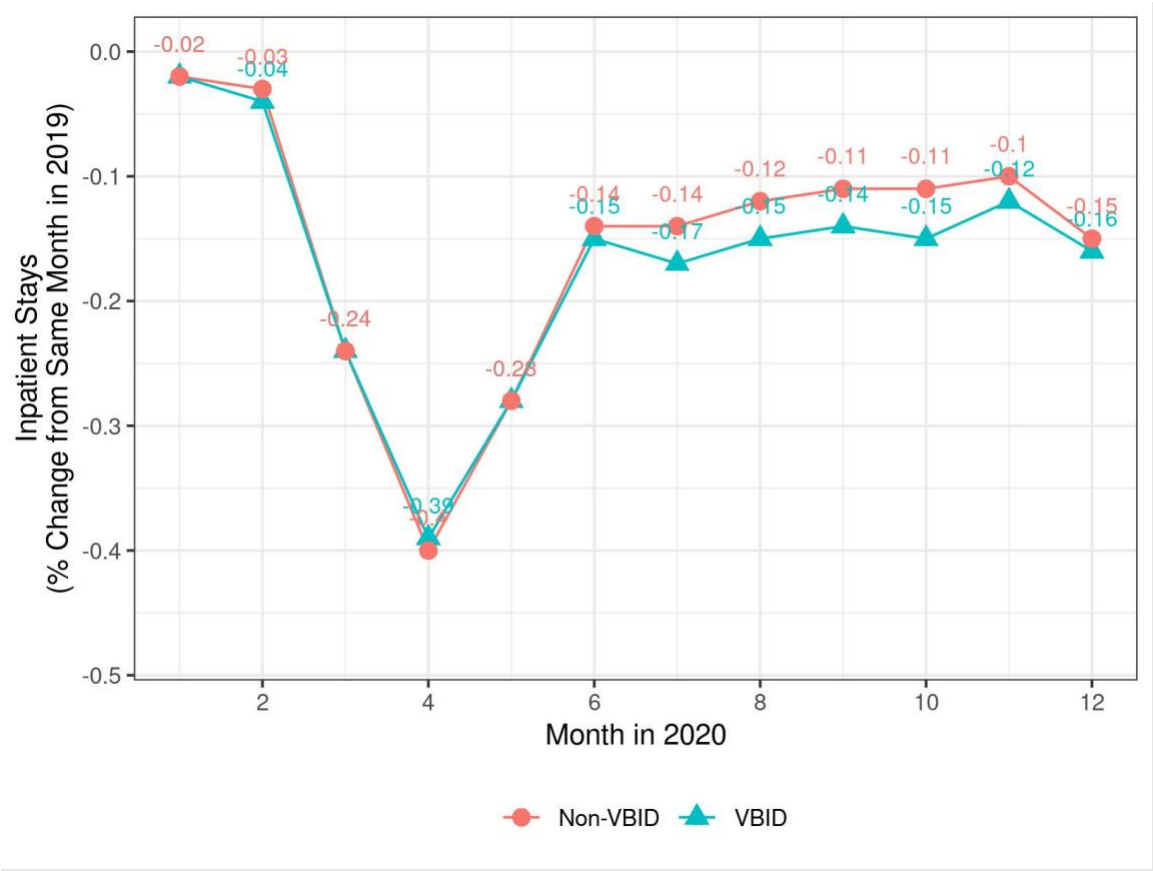
Figure G.2. Year-on-Year Change in Emergency Department Visits Between 2019 and 2020, VBID Counties vs. Non-VBID Counties



SOURCE: RAND analysis of Medicare Common Working File.
NOTE: Total number of visits for each group of counties (VBID vs. non-VBID) was calculated before calculating percent change in the group total.

Figure G.3 compares changes in the number of inpatient stays between 2019 and 2020. Here, the decline in utilization from 2019 observed in the early months of the pandemic (March, April, and May 2020) is very similar in VBID and non-VBID counties. However, inpatient stay volumes in the two groups of counties diverge in June 2020 and later months, with a larger year-on-year decline observed in VBID counties through the end of the year. The gap in the volume of stays in July and later months varies between 1 and 4 percentage points and remains around 3 to 4 percentage points between July and October.

Figure G.3. Year-on-Year Change in Inpatient Stays Between 2019 and 2020, VBID Counties vs. Non-VBID Counties



SOURCE: RAND analysis of Medicare Common Working File.
NOTE: Total number of stays for each group of counties (VBID vs. non-VBID) was calculated before calculating percent change in the group total.

Sensitivity Analyses

These findings (different patterns of case rates and larger reductions in utilization during the pandemic for counties with VBID participants) were robust to several sensitivity analyses, including the following:

- Comparing counties with high VBID enrollment with counties with low VBID enrollment (as opposed to comparing VBID and non-VBID counties)
- Defining the outcome based on claim lines as opposed to visits
- Regression-adjusting the results to control for state.

Sensitivity Analysis: Difference-in-Differences Regression Model Results With and Without COVID-19 Case Rate Controls

As noted above, differential changes in FFS utilization in VBID and non-VBID counties would appear to be a concern primarily for analyses in which utilization is the outcome measure or in which outcome measures are constructed based on utilization (e.g., quality measures or realized medical and prescription drug spending).

In this evaluation report, our quantitative outcomes are limited to plan enrollment, plan bids, beneficiary premiums, and supplemental benefit offerings. It is less clear that the COVID-19 pandemic would have major impacts on these outcomes, especially within the time frame examined (2020 and 2021 plan years) in this evaluation report. Bids for MA and Part D coverage must be submitted by June of the year prior to the plan year, which means that 2020 bids were submitted no later than June 2019 and 2021 bids were submitted no later than June 2020. Premiums for MA and Part D are determined on the basis of plan bids, benchmarks, and other inputs that are also based on information that is compiled before the beginning of the plan year.

Consequently, we do not think the COVID-19 pandemic should have any impact on plan bids or premiums in 2020. We cannot be quite so certain about bids or premiums in 2021. That said, we think the timing of the deadline for plans to submit 2021 bids (in early June 2020) makes it unlikely that POs would have had enough reliable information about the implications of the pandemic to support changes in actuarial calculations that would have been needed to justify meaningful changes in bidding strategy for 2021.

It is easier to imagine that COVID-19 might have affected enrollment, although the vast range of potential mechanisms through which the pandemic may have affected MA plan choice or election of MA versus FFS leaves us without a clear hypothesis on the direction of this impact. Although we plan to use FFS utilization measures to control for COVID-19 impacts in future analyses of utilization and related outcomes, we felt that there was not a clear justification for using ex post FFS utilization measures as control variables for modeling enrollment, bid, or premium outcomes: Projected net medical spending is based on base period data from two years before the plan year, so COVID-driven utilization changes in 2020 would most likely not be factored into the data used to develop rates.

To explore the possible impact of COVID-19 on our research design, we estimated regression models with additional control variables for a measure of the potential plan-level impact of COVID. We used the cumulative COVID-19 case rate per 10,000 adults aged 60 and over as of December 31, 2020, as a proxy for the pandemic's intensity during 2020. This variable was defined at the county level, so we used 2019 plan enrollment by county to construct a plan-level cumulative COVID-19 case rate for plans offered in multiple counties. We then interacted this plan-level COVID-19 measure with indicator variables for the 2020 and 2021 calendar years and included these interaction terms in our regression model. By adding these plan-level COVID-19 measures to our DD regressions, we effectively allowed the within-plan change in

outcomes to vary systematically with the severity of the pandemic in a plan's (enrollment-weighted) service area. Put differently, this specification controls for a dose-response-like relationship between COVID-19 cases in 2020 among older adults and changes in bid, premium, or enrollment outcomes. VBID impacts estimated in the presence of these control variables then tell us if our estimated VBID effects are due to confounding between VBID implementation and the impact of the pandemic.

Comparison of Enrollment Impacts With and Without Controls for COVID-19 Case Rate

To assess the potential for confounding between VBID and impacts of the COVID-19 pandemic, Tables G.2, G.3, and G.4 compare summary effects of BDI participation in 2020 and 2021 from models estimated with COVID-19 control variables and models estimated without COVID-19 control variables: The latter estimates are the main results reported in Chapter 5. In general, adding COVID-19 controls has little effect on our estimates. However, the negative relationship between Hospice component participation and MAPD bids, which is not statistically significant in the main models, increases and becomes statistically significant after adding COVID-19 controls.

Table G.2. Log(Enrollment) BDI Effects With and Without COVID-19 Control Variables

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Without COVID-19 controls, by intervention and plan year					
BDI participation: 2020	0.00	0.04	–0.08	0.07	0.96
BDI participation: 2021	0.06	0.03	0.00	0.12	0.06
Hospice participation: 2021	0.00	0.10	–0.21	0.20	0.96
With COVID-19 controls, by intervention and plan year					
BDI participation: 2020	–0.01	0.05	–0.11	0.08	0.82
BDI participation: 2021	0.07	0.03	0.00	0.13	0.04
Hospice participation: 2021	0.02	0.18	–0.33	0.38	0.89

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: CI = confidence interval.

Table G.3. MAPD Bid BDI Effects With and Without COVID-19 Control Variables

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Without COVID-19 controls, by intervention and plan year					
BDI participation: 2020	–5.79	3.37	–12.39	0.81	0.09
BDI participation: 2021	–5.37	2.00	–9.30	–1.44	0.01
Hospice participation: 2021	–14.79	10.02	–34.42	4.84	0.14
With COVID-19 controls, by intervention and plan year					
BDI participation: 2020	–5.45	3.77	–12.84	1.94	0.15
BDI participation: 2021	–5.68	1.82	–9.25	–2.11	0.00
Hospice participation: 2021	–21.51	9.58	–40.29	–2.72	0.02

SOURCE: RAND analysis of 2017–2021 OACT bid data.

Table G.4. Total Premium BDI Effects With and Without COVID-19 Control Variables

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Without COVID-19 controls, by intervention and plan year					
BDI participation: 2020	–0.21	0.96	–2.08	1.67	0.83
BDI participation: 2021	1.93	0.53	0.89	2.97	0.00
Hospice participation: 2021	–4.44	3.46	–11.22	2.34	0.20
With COVID-19 controls, by intervention and plan year					
BDI participation: 2020	–0.65	0.98	–2.58	1.27	0.50
BDI participation: 2021	1.69	0.55	0.61	2.76	0.00
Hospice participation: 2021	–5.35	3.24	–11.70	0.99	0.10

SOURCE: RAND analysis of 2017–2021 CMS Medicare Part D landscape files.

Tables G.2, G.3, and G.4 show that controlling for the intensity of the pandemic in plans' service areas during 2020 (as proxied by cumulative 2020 COVID-19 case rates per 10,000 adults aged 60+) does not meaningfully affect our DD regression estimates of VBID participation impacts on enrollment, bids, or premiums.

Strategy for Addressing COVID-19 Pandemic Impacts in Future Evaluation Reports

We propose to address the potential for confounding from COVID-19 impacts by including controls for county-level FFS utilization changes in DD regression analyses of VBID utilization impacts. We will include county-level control variables reflecting year-on-year changes in Medicare FFS utilization to account for differential impacts of the COVID-19 pandemic across counties. This will allow us to distinguish VBID impacts from utilization changes driven by

COVID-19 and ensure that any VBID impacts identified by our evaluation are, in fact, driven by VBID implementation.

In models with outcomes reflecting a specific care setting (e.g., inpatient or ED utilization, or some quality measures), we will include controls for changes in FFS utilization in that care setting. In other models that reflect care received across multiple settings (e.g., costs, health outcomes, or some quality measures), we will include multiple variables capturing changes in FFS utilization across multiple settings. The decision of which FFS utilization measures to include in each model will be made based on our expert and clinical judgment about which care settings are most relevant to each measure. In general, we will need to use our judgment on a case-by-case basis to choose the measurement approach that is most appropriate to specific outcomes, and these decisions will be undertaken after the specification and functional form of the baseline DD models have been finalized.

Appendix H. Enrollment Analyses

This appendix describes the methods used for the enrollment analyses presented in Chapters 5 and 8. The general approaches for the study design were described in Chapter 1, and the plan entropy balancing processes were described in Appendix C.

Balance and Parallel Trends

Plans That Participated in Both 2020 and 2021

Table H.1 provides the results of our entropy balancing algorithm for analysis of enrollment for plans that participated in both 2020 and 2021. The average absolute standardized mean difference (ASMD) before and after weighting was 0.22 and 0.03, respectively. The number of participating plans included in this analysis was 89, and the effective number of comparison plans after weighting was 106.

Table H.1. Mean, Standard Deviation, and Absolute Standardized Mean Difference Comparing Plans That Participated in Both 2020 and 2021 with Eligible Comparison Plans, Before and After Weighting

	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Plan characteristics								
MA bid	788	766	790	0.30	0.03	72.53	88.95	74.91
Part D bid	42.78	51.90	43.26	0.76	0.04	12.07	19.96	13.51
MA premiums	10.02	11.34	11.32	0.04	0.04	32.43	28.36	32.90
Part D premiums	18.31	16.41	17.47	0.09	0.04	20.94	19.57	19.90
Cost of MSB	31.15	26.20	30.11	0.19	0.04	26.18	18.15	25.07
Rebate dollars amount	105	120	107	0.32	0.04	47.51	73.65	52.62
Administrative costs	115	121	115	0.44	0.04	14.28	41.77	17.36
Part C in-network OOP maximum	5,419	5,053	5,364	0.27	0.04	1373	1,638	1,432
PPO	0.21	0.21	0.20	0.00	0.03	0.41	0.41	0.40
For profit	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
Star Rating	4.08	3.98	4.07	0.33	0.04	0.31	0.50	0.32
C-SNP	0.01	0.07	0.01	0.52	0.02	0.11	0.25	0.11
D-SNP	0.26	0.15	0.24	0.24	0.04	0.44	0.36	0.43
Moved into bonus payment	0.02	0.08	0.03	0.37	0.04	0.15	0.27	0.17
Moved out of bonus payment	0.04	0.06	0.04	0.08	0.04	0.21	0.24	0.19
Did not exist in 2017	0.18	0.29	0.20	0.29	0.04	0.38	0.45	0.40

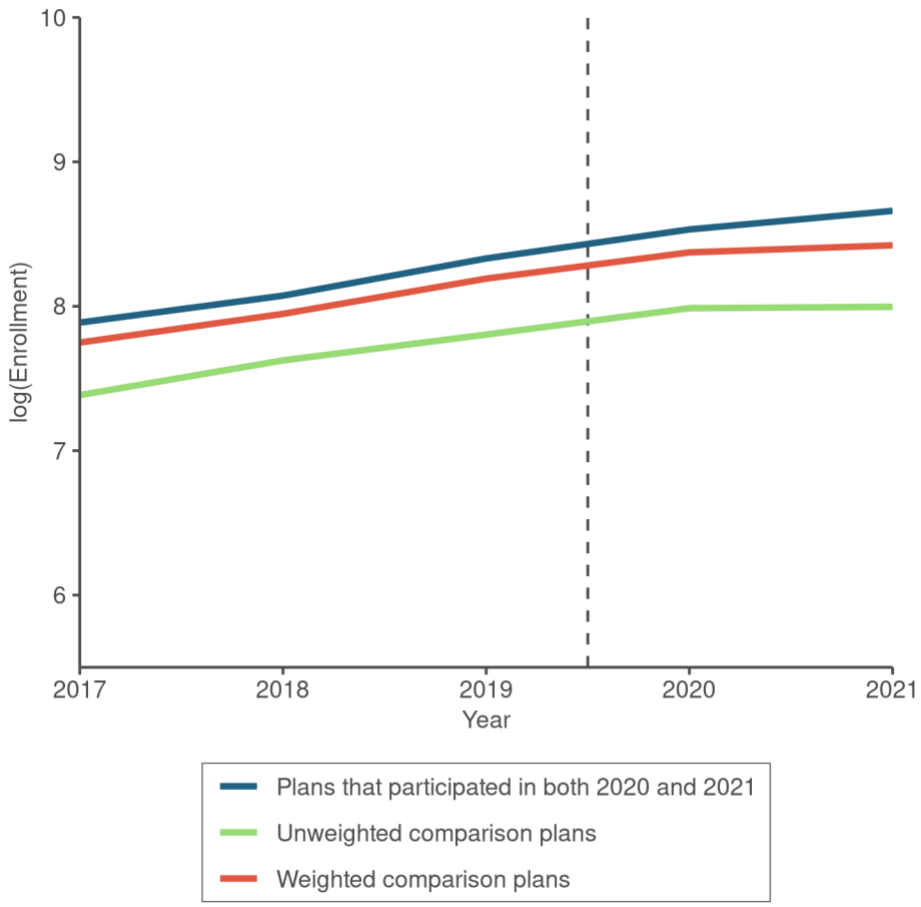
	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Did not exist in 2018	0.10	0.18	0.11	0.26	0.03	0.30	0.38	0.31
Individual characteristics at the plan level								
Average age	70.7	70.9	70.9	0.05	0.04	3.95	4.06	3.87
Percent male	0.43	0.44	0.43	0.33	0.02	0.05	0.07	0.05
Percent AI/AN	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00
Percent API	0.03	0.04	0.03	0.14	0.04	0.06	0.09	0.06
Percent Black	0.12	0.13	0.12	0.12	0.04	0.14	0.15	0.14
Percent Hispanic	0.09	0.16	0.10	0.46	0.04	0.16	0.22	0.17
Percent multirace	0.03	0.02	0.02	0.21	0.04	0.01	0.01	0.01
Percent White	0.73	0.63	0.72	0.43	0.04	0.23	0.28	0.24
Percent dually eligible for Medicaid	0.33	0.29	0.33	0.09	0.01	0.40	0.35	0.40
Average MA risk score	1.29	1.26	1.29	0.09	0.00	0.34	0.34	0.33
Average Part D risk score	1.14	1.12	1.15	0.07	0.04	0.33	0.35	0.33
Percent diabetes	0.30	0.31	0.30	0.13	0.04	0.08	0.14	0.08
Percent CHF	0.13	0.13	0.14	0.05	0.04	0.05	0.07	0.06
Percent COPD	0.19	0.17	0.18	0.20	0.04	0.08	0.10	0.08
Percent cancer	0.09	0.09	0.09	0.26	0.04	0.03	0.03	0.03
Area characteristics of the plans								
Median income	30,932	30,188	30,748	0.16	0.04	4,605	5,030	4,777
MA penetration	44.5	41.9	44.1	0.25	0.04	10.5	12.0	10.2
Percent rural	4.45	5.79	4.70	0.20	0.04	6.76	9.48	7.51
Percent suburban	14.66	15.85	15.09	0.10	0.04	11.7	14.9	12.2
Percent urban	80.9	78.4	80.2	0.15	0.04	16.9	20.7	16.9
HPSA	2.19	4.14	2.39	0.30	0.03	6.56	10.15	6.83
Standardized Medicare costs per capita	10,290	10,752	10,343	0.35	0.04	1,326	1,470	1,342
Percent over 65	0.19	0.18	0.18	0.21	0.04	0.04	0.04	0.04

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

NOTES: UWgt = unweighted; Wgt = weighted; Comp = comparison; SD = standard deviation; AI/AN = American Indian and Alaska Native; API = Asian Pacific Islander. The entropy balancing weights used in this analysis were estimated using constraints on both the first and second moments of each characteristic, ensuring balance between the participating and comparison plans on means and variances.

Figure H.1 provides the trends in the logarithm of enrollment from 2017 to 2021 for plans that participated in both 2020 and 2021 compared with eligible comparison plans before and after weighting. Pre-intervention trends are similar between the groups before and after weighting.

Figure H.1. Trends in log(Enrollment) for Plans That Participated in Both 2020 and 2021 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Plans That Participated Only in 2020

Table H.2 provides the results of our entropy balancing algorithm for analysis of enrollment for plans that participated only in 2020. The average ASMD before and after weighting was 0.37 and 0.05, respectively. The number of participating plans included in this analysis was 39, and the effective number of comparison plans after weighting was 46.

Table H.2. Mean, Standard Deviation, and Absolute Standardized Mean Difference Comparing Plans That Participated Only in 2020 with Eligible Comparison Plans, Before and After Weighting

	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Plan characteristics								
MA bid	735	782	737	0.39	0.01	119	84	118
Part D bid	42.69	52.83	43.70	0.70	0.07	14.38	18.95	14.93
MA premiums	7.70	16.57	8.52	0.55	0.05	16.17	33.29	20.69
Part D premiums	10.10	19.53	10.94	0.78	0.07	12.05	21.53	13.48
Cost of MSB	20.12	22.13	21.02	0.16	0.07	12.79	16.74	14.90
Rebate dollars amount	84.55	96.78	81.90	0.32	0.07	37.90	58.02	36.67
Administrative costs	111	116	112	0.24	0.07	23.32	38.99	28.08
Part C in-network OOP maximum	5,918	5,261	5,830	0.52	0.07	1,252	1,518	1,299
PPO	0.10	0.28	0.12	0.59	0.07	0.30	0.45	0.33
For profit	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
Star Rating	3.35	3.91	3.39	0.98	0.07	0.58	0.52	0.60
C-SNP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-SNP	0.26	0.12	0.23	0.32	0.07	0.44	0.32	0.42
Moved into bonus payment	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Moved out of bonus payment	0.18	0.07	0.15	0.28	0.07	0.38	0.26	0.36
Did not exist in 2017	0.18	0.31	0.15	0.35	0.07	0.38	0.46	0.36
Did not exist in 2018	0.03	0.20	0.04	1.08	0.07	0.16	0.40	0.19
Individual characteristics at the plan level								
Average age	69.8	71.3	69.9	0.31	0.01	4.82	3.84	4.57
Percent male	0.44	0.44	0.44	0.04	0.01	0.05	0.06	0.05
Percent AI/AN	0.00	0.00	0.00	0.61	0.07	0.00	0.00	0.00
Percent API	0.04	0.04	0.04	0.02	0.07	0.08	0.09	0.07
Percent Black	0.29	0.12	0.27	0.74	0.07	0.23	0.14	0.22
Percent Hispanic	0.19	0.12	0.20	0.22	0.04	0.31	0.18	0.30
Percent multirace	0.03	0.02	0.03	0.25	0.00	0.01	0.01	0.01
Percent White	0.45	0.69	0.46	0.93	0.04	0.26	0.26	0.28
Percent dually eligible for Medicaid	0.42	0.24	0.39	0.51	0.07	0.36	0.32	0.36
Average MA risk score	1.28	1.16	1.28	0.31	0.00	0.40	0.28	0.36
Average Part D risk score	1.21	1.03	1.19	0.57	0.07	0.30	0.26	0.29
Percent diabetes	0.33	0.27	0.33	0.81	0.06	0.07	0.08	0.07
Percent CHF	0.15	0.11	0.14	0.52	0.05	0.06	0.05	0.06
Percent COPD	0.16	0.14	0.16	0.32	0.07	0.05	0.06	0.06
Percent cancer	0.09	0.09	0.09	0.09	0.04	0.02	0.03	0.02
Area characteristics of the plans								
Median income	28,873	31,090	29,435	0.28	0.07	8,036	5,407	7,918
MA penetration	41.5	40.7	41.3	0.05	0.01	16.5	11.9	16.7

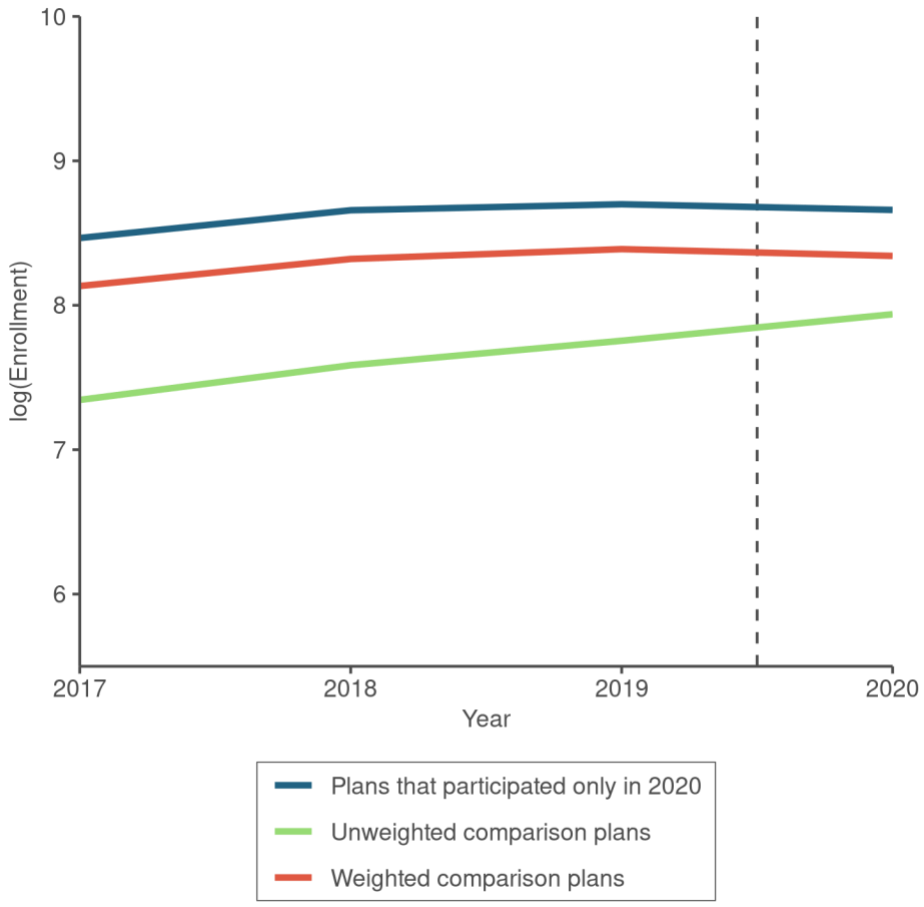
	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Percent rural	6.74	6.27	7.66	0.04	0.07	13.08	9.64	13.49
Percent suburban	16.49	17.11	15.55	0.05	0.07	13.4	15.1	13.2
Percent urban	76.8	76.6	76.8	0.01	0.00	24.2	20.6	23.5
HPSA	5.88	4.20	6.54	0.18	0.07	9.40	10.28	11.30
Standardized Medicare costs per capita	11,021	10,434	10,955	0.62	0.07	952	1,481	1,029
Percent over 65	0.18	0.18	0.17	0.00	0.07	0.03	0.04	0.03

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

NOTE: The entropy balancing weights used in this analysis were estimated using constraints on both the first and second moments of each characteristic, ensuring balance between the participating and comparison plans on means and variances.

Figure H.2 provides the trends in the logarithm of enrollment from 2017 to 2021 for plans that participated only in 2020 compared with eligible comparison plans before and after weighting. Pre-intervention trends differ slightly before weighting but are parallel after weighting.

Figure H.2. Trends in log(Enrollment) for Plans That Participated Only in 2020 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Plans That Participated Only in 2021

Table H.3 provides the results of our entropy balancing algorithm for analysis of enrollment for plans that participated only in 2021. The average ASMD before and after weighting was 0.39 and 0.06, respectively. The number of participating plans included in this analysis was 257, and the effective number of comparison plans after weighting was 261.

Table H.3. Mean, Standard Deviation, and Absolute Standardized Mean Difference Comparing Plans That Participated Only in 2021 with Eligible Comparison Plans, Before and After Weighting

	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Plan characteristics								
MA bid	843	807	836	0.44	0.08	83.83	91.12	88.12
Part D bid	35.30	46.17	36.11	1.21	0.09	8.96	17.21	10.19
MA premiums	9.90	14.69	8.10	0.20	0.08	23.70	32.21	20.51
Part D premiums	14.37	16.81	14.92	0.20	0.05	12.19	20.40	12.06
Cost of MSB	42.44	27.27	42.33	0.37	0.00	40.51	20.55	38.47
Rebate dollars amount	105	120	108	0.30	0.07	50.50	72.82	59.36
Administrative costs	162	140	165	0.61	0.07	36.77	46.96	41.54
Part C in-network OOP maximum	5,755	4,970	5,623	0.54	0.09	1,462	1,615	1,506
PPO	0.37	0.26	0.32	0.22	0.09	0.48	0.44	0.47
For profit	1.00	0.75	0.99	4.02	0.09	0.06	0.44	0.10
Star Rating	4.08	4.04	4.05	0.12	0.09	0.33	0.51	0.37
C-SNP	0.01	0.05	0.02	0.38	0.09	0.11	0.22	0.14
D-SNP	0.41	0.11	0.36	0.61	0.09	0.49	0.31	0.48
Moved into bonus payment	0.09	0.07	0.06	0.07	0.09	0.29	0.25	0.24
Moved out of bonus payment	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Did not exist in 2017	0.40	0.39	0.36	0.01	0.07	0.49	0.49	0.48
Did not exist in 2018	0.27	0.30	0.28	0.05	0.01	0.45	0.46	0.45
Individual characteristics at the plan level								
Average age	67.7	70.4	68.1	0.59	0.09	4.59	3.83	4.05
Percent male	0.42	0.45	0.43	0.45	0.09	0.06	0.07	0.06
Percent AI/AN	0.01	0.00	0.01	0.19	0.02	0.01	0.01	0.01
Percent API	0.04	0.04	0.04	0.02	0.04	0.09	0.09	0.09
Percent Black	0.17	0.12	0.18	0.28	0.09	0.17	0.14	0.17
Percent Hispanic	0.15	0.14	0.15	0.05	0.01	0.22	0.19	0.21
Percent multirace	0.03	0.02	0.03	0.53	0.04	0.01	0.01	0.01
Percent White	0.61	0.68	0.60	0.25	0.03	0.27	0.26	0.27
Percent dually eligible for Medicaid	0.50	0.23	0.49	0.63	0.01	0.43	0.31	0.42
Percent diabetes	0.33	0.30	0.34	0.46	0.01	0.08	0.14	0.10
Percent CHF	0.15	0.13	0.15	0.41	0.00	0.05	0.07	0.06

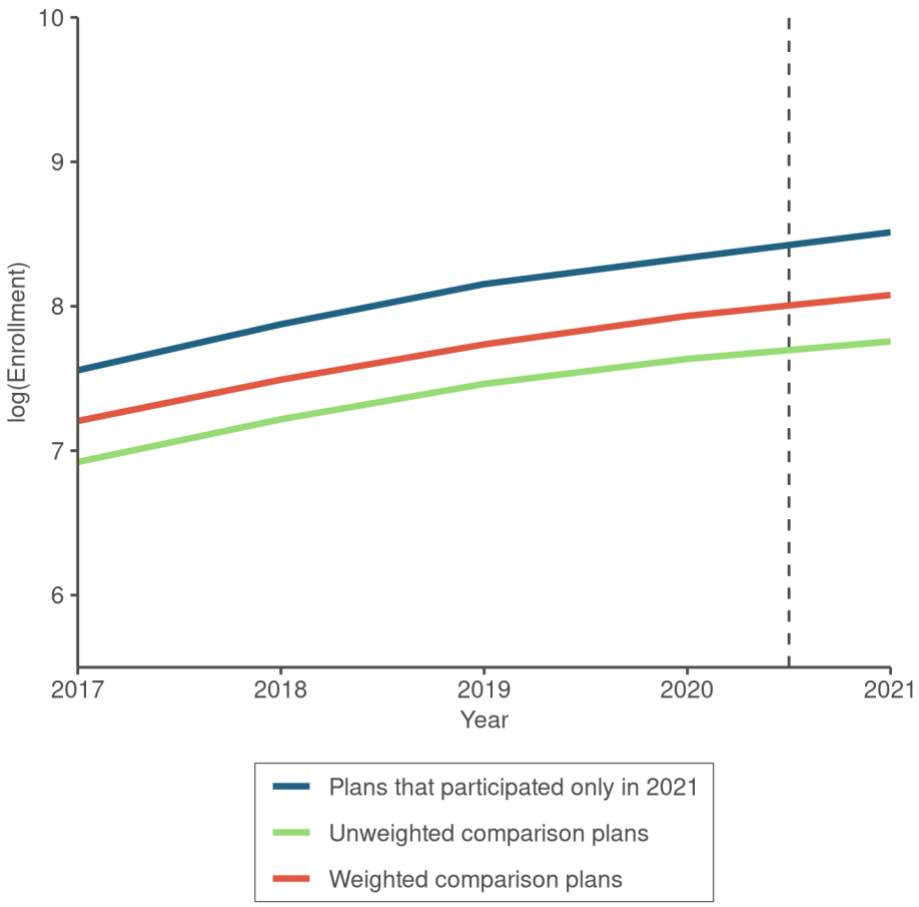
	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Percent COPD	0.21	0.15	0.20	0.68	0.09	0.09	0.08	0.08
Percent cancer	0.08	0.09	0.09	0.32	0.03	0.02	0.03	0.02
Area characteristics of the plans								
Median income	29,623	30,919	29,153	0.23	0.08	5,569	5,624	5,163
MA penetration	42.0	43.7	41.7	0.16	0.02	11.2	11.8	12.0
Percent rural	7.38	6.33	7.53	0.11	0.02	9.38	9.90	10.13
Percent suburban	17.7	16.3	18.9	0.12	0.09	12.7	14.8	13.5
Percent urban	74.9	77.4	73.6	0.13	0.07	19.2	20.7	19.4
HPSA	8.28	4.84	7.48	0.22	0.05	15.47	11.63	12.72
Standardized Medicare costs per capita	10,524	10,474	10,640	0.04	0.09	1,295	1,507	1,411
Percent over 65	0.18	0.18	0.18	0.02	0.01	0.03	0.04	0.03

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

NOTE: The entropy balancing weights used in this analysis were estimated using constraints on both the first and second moments of each characteristic, ensuring balance between the participating and comparison plans on means and variances.

Figure H.3 provides the trends in the logarithm of enrollment from 2017 to 2021 for plans that participated only in 2021 compared with eligible comparison plans before and after weighting. Pre-intervention trends are similar before and after weighting.

Figure H.3. Trends in log(Enrollment) for Plans That Participated Only in 2021 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Plans That Participated in Hospice

Table H.4 provides the results of our entropy balancing algorithm for analysis of enrollment for plans that participated in the Hospice component. The average ASMD before and after weighting was 0.37 and 0.04, respectively. The number of participating plans included in this analysis was 46, and the effective number of comparison plans after weighting was 43.

Table H.4. Mean, Standard Deviation, and Absolute Standardized Mean Difference Comparing Plans That Participated in the Hospice Component with Eligible Comparison Plans, Before and After Weighting

	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Plan characteristics								
MA bid	637	814	648	0.84	0.05	209.66	84.87	208.03
Part D bid	48.59	44.70	47.73	0.23	0.05	17.07	16.88	17.49
MA premiums	8.71	14.06	9.72	0.26	0.05	20.28	31.59	24.06
Part D premiums	15.68	16.62	15.37	0.04	0.01	26.18	19.70	25.03
Cost of MSB	31.53	29.30	32.83	0.09	0.05	25.86	24.85	26.33
Rebate dollars amount	139	118	139	0.32	0.01	65.68	69.96	64.21
Administrative costs	128	142	126	0.38	0.05	37.55	46.13	37.55
Part C in-network OOP maximum	4,393	5,087	4,314	0.44	0.05	1,590	1,606	1,532
PPO	0.17	0.27	0.15	0.25	0.05	0.38	0.44	0.36
For profit	0.70	0.78	0.72	0.18	0.05	0.46	0.42	0.45
Star Rating	4.17	4.01	4.16	0.48	0.05	0.33	0.50	0.36
C-SNP	0.04	0.05	0.04	0.02	0.03	0.20	0.21	0.19
D-SNP	0.28	0.14	0.30	0.31	0.03	0.45	0.35	0.46
Moved into bonus payment	0.02	0.07	0.03	0.33	0.05	0.15	0.26	0.17
Moved out of bonus payment	0.02	0.04	0.01	0.11	0.05	0.15	0.19	0.12
Did not exist in 2017	0.20	0.39	0.22	0.48	0.05	0.40	0.49	0.41
Did not exist in 2018	0.15	0.29	0.17	0.38	0.05	0.36	0.45	0.38
Did not exist in 2019	0.07	0.14	0.07	0.31	0.01	0.25	0.35	0.25
Individual characteristics at the plan level								
Average age	72.3	70.1	72.1	0.58	0.05	3.75	4.02	3.69
Percent male	0.44	0.44	0.44	0.06	0.05	0.07	0.07	0.08
Percent AI/AN	0.00	0.00	0.00	0.77	0.05	0.00	0.01	0.00
Percent API	0.08	0.04	0.07	0.18	0.05	0.18	0.09	0.17
Percent Black	0.06	0.13	0.06	0.50	0.04	0.13	0.15	0.13
Percent Hispanic	0.49	0.13	0.49	0.83	0.01	0.43	0.19	0.43
Percent multirace	0.01	0.02	0.01	0.52	0.05	0.02	0.01	0.01
Percent White	0.36	0.67	0.37	0.88	0.05	0.36	0.26	0.36

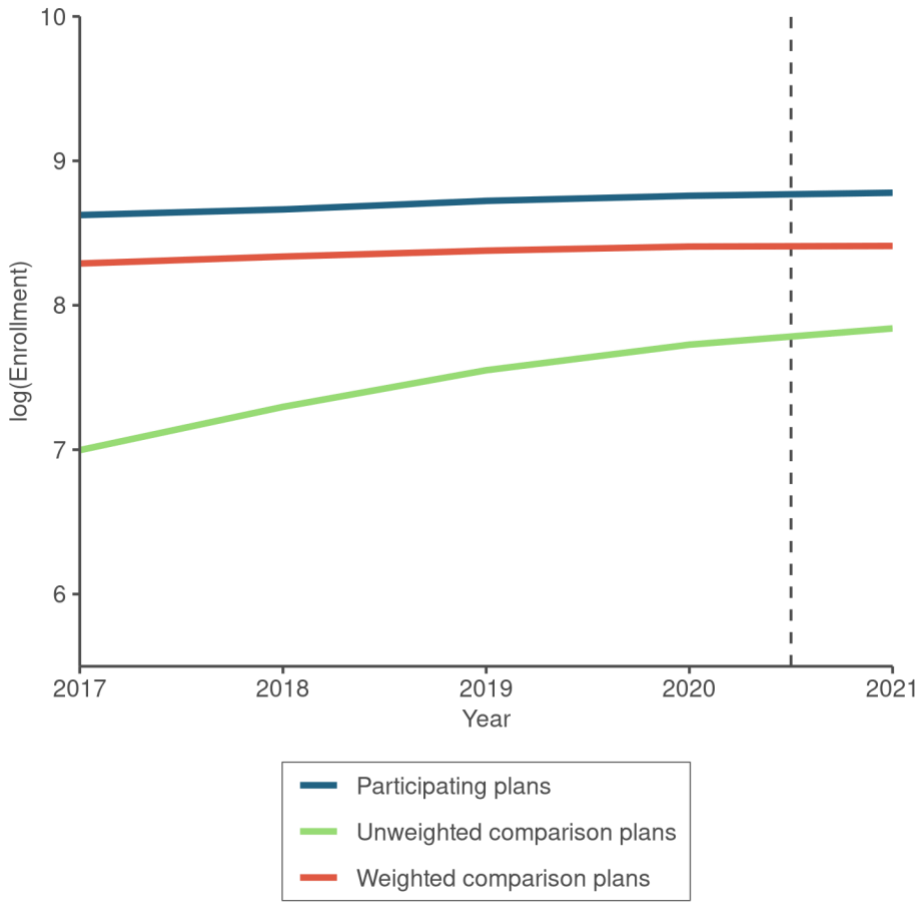
	Mean (VBID)	UWgt Mean (Comp)	Wgt Mean (Comp)	UWgt ASMD	Wgt ASMD	SD (VBID)	UWgt SD (Comp)	Wgt SD (Comp)
Percent dually eligible for Medicaid	0.31	0.26	0.33	0.13	0.05	0.40	0.34	0.41
Percent diabetes	0.35	0.30	0.36	0.44	0.04	0.12	0.13	0.13
Percent CHF	0.19	0.13	0.19	0.62	0.05	0.10	0.07	0.10
Percent COPD	0.15	0.16	0.15	0.12	0.02	0.06	0.09	0.07
Percent cancer	0.09	0.09	0.10	0.15	0.05	0.02	0.03	0.03
Area characteristics of the plans								
Median income	23,501	30,831	23,974	0.77	0.05	9,462	5,423	10,247
MA penetration	58.9	43.3	59.9	0.82	0.05	19.0	11.5	18.3
Percent rural	7.66	6.34	7.29	0.12	0.03	11.35	9.71	11.30
Percent suburban	12.4	16.6	12.7	0.47	0.03	8.75	14.64	9.38
Percent urban	79.9	77.1	80.0	0.15	0.00	18.7	20.6	18.2
HPSA	7.08	5.09	6.02	0.09	0.05	21.1	11.7	18.8
Standardized Medicare costs per capita	9,853	10,497	9,930	0.42	0.05	1,548	1,470	1,515
Percent over 65	0.19	0.18	0.19	0.30	0.05	0.03	0.04	0.03

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

NOTE: The entropy balancing weights used in this analysis were estimated using constraints on both the first and second moments of each characteristic, ensuring balance between the participating and comparison plans on means and variances.

Figure H.4 provides the trends in the logarithm of enrollment from 2017 to 2021 for plans that participated in the Hospice component compared with eligible comparison plans before and after weighting. Pre-intervention trends differ substantially before weighting but are parallel after weighting.

Figure H.4. Trends in log(Enrollment) for Plans That Participated in the Hospice Component in 2021 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Group-Specific Difference-in-Differences Regression Model Results

Tables H.5 through H.8 provide the DD model results for each group of participating plans. Enrollment was analyzed on the logarithmic scale so that model coefficients represent changes in the logarithm of enrollment.

Table H.5. Difference-in-Differences Model Results for Analysis of Plans That Participated in Both 2020 and 2021

Characteristic	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Year					
2018	0.06	0.05	–0.03	0.16	0.19
2019	0.26	0.10	0.06	0.46	0.01
2020	0.17	0.32	–0.47	0.80	0.61
2021	–0.02	0.57	–1.14	1.11	0.98
BDI participation					
2020	0.02	0.04	–0.05	0.10	0.52
2021	0.11	0.06	–0.02	0.23	0.09
Hospice participation					
2021	0.16	0.20	–0.22	0.55	0.40
UF participation					
2019	–0.31	0.17	–0.64	0.01	0.06
2020	–0.35	0.13	–0.61	–0.10	0.01
2021	–0.35	0.15	–0.64	–0.05	0.02
SSBCI participation					
2020	–0.15	0.11	–0.36	0.06	0.16
2021	–0.03	0.09	–0.20	0.13	0.68
PDSS participation					
2021	–0.35	0.08	–0.52	–0.19	0.00
New PHRSB participation					
2020	0.27	0.26	–0.23	0.77	0.30
2021	0.66	0.52	–0.36	1.69	0.21

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: Plan-level fixed effects are excluded from this table. The number of participating plans included in this analysis was 89, and the effective number of comparison plans after weighting was 106.

Table H.6. Difference-in-Differences Model Results for Analysis of Plans That Participated Only in 2020

Characteristic	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Year					
2018	–0.04	0.11	–0.24	0.17	0.72
2019	–0.28	0.23	–0.73	0.18	0.23
2020	–0.54	0.35	–1.23	0.16	0.13
BDI participation					
2020	–0.06	0.10	–0.26	0.13	0.53
UF participation					
2020	0.69	0.18	0.34	1.04	0.00
SSBCI participation					
2020	0.56	0.32	–0.07	1.19	0.08
New PHRSB participation					
2020	0.07	0.16	–0.25	0.39	0.68

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: Plan-level fixed effects are excluded from this table. The number of participating plans included in this analysis was 39, and the effective number of comparison plans after weighting was 46.

Table H.7. Difference-in-Differences Model Results for Analysis of Plans That Participated Only in 2021

Characteristic	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Year					
2018	0.09	0.04	0.01	0.16	0.02
2019	0.10	0.09	–0.07	0.28	0.25
2020	0.28	0.13	0.02	0.54	0.04
2021	0.34	0.19	–0.04	0.72	0.08
BDI participation					
2021	0.04	0.03	–0.02	0.11	0.21
Hospice participation					
2021	–0.29	0.15	–0.58	–0.01	0.05
UF participation					
2019	0.33	0.22	–0.11	0.77	0.14
2020	0.20	0.16	–0.11	0.51	0.20
2021	0.19	0.15	–0.10	0.48	0.19
SSBCI participation					
2020	0.00	0.16	–0.31	0.32	0.98
2021	0.05	0.06	–0.07	0.17	0.39
PDSS participation					
2021	–0.08	0.05	–0.19	0.03	0.15
New PHRSB participation					
2020	0.14	0.09	–0.03	0.31	0.10
2021	0.26	0.13	0.00	0.52	0.05

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: Plan-level fixed effects are excluded from this table. The number of participating plans included in this analysis was 257, and the effective number of comparison plans after weighting was 261.

Table H.8. Difference-in-Differences Model Results for Analysis of Plans That Participated in the Hospice Component

Characteristic	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Year					
2018	–0.04	0.06	–0.16	0.08	0.48
2019	–0.10	0.16	–0.40	0.21	0.53
2020	–0.01	0.21	–0.43	0.41	0.96
2021	0.00	0.31	–0.60	0.60	1.00
Hospice participation					
2021	0.00	0.10	–0.21	0.20	0.96
BDI participation					
2020	0.18	0.19	–0.19	0.55	0.35
2021	0.16	0.19	–0.22	0.54	0.42
UF participation					
2019	0.18	0.11	–0.04	0.39	0.10
2020	0.13	0.11	–0.10	0.35	0.27
2021	0.23	0.16	–0.08	0.54	0.15
SSBCI participation					
2020	–0.15	0.16	–0.46	0.17	0.36
2021	–0.22	0.13	–0.48	0.04	0.10
PDSS participation					
2021	–0.12	0.15	–0.41	0.18	0.44
New PHRSB participation					
2020	–0.04	0.10	–0.24	0.16	0.69
2021	0.02	0.21	–0.38	0.43	0.91

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: Plan-level fixed effects are excluded from this table. The number of participating plans included in this analysis was 46, and the effective number of comparison plans after weighting was 43.

Summaries of Difference-in-Differences Model Results

Table H.9 provides the aggregated effects of VBID participation, which summarize the group-specific DD model results. To understand the estimates in this table, consider the methodology described in Appendix C. The estimate for the first year of participation is a weighted average of the estimated effects from each group-specific DD estimate corresponding to the first year of participation for that group. The weights are defined in Appendix C.

Table H.9. Estimated Effect of BDI Participation on the Logarithm of Enrollment

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
By plan year					
2020	0.00	0.04	–0.08	0.07	0.96
2021	0.06	0.03	0.00	0.12	0.06
By year of participation					
1st	0.03	0.03	–0.02	0.08	0.28
2nd	0.11	0.06	–0.02	0.23	0.09

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Appendix I. Plan Bids, Premiums, and Supplemental Benefits Analyses

This appendix describes the methods used for the health care cost analyses presented in Chapters 5 and 8. The general approaches for the study design were described in Chapter 1, and the plan entropy balancing processes were described in Appendix C.

Data Sources and Variable Definitions

Outcome Measures

This section describes data sources and definitions for each of the outcome measures used to analyze health care costs. This interim evaluation focuses on three measures: plan bids, beneficiary premiums, and plans' projected PMPM costs for mandatory supplemental benefits not covered by Medicare (which we refer to as *mandatory supplemental benefit costs*). We describe each of these in turn, elaborating on high-level descriptions that were provided in Chapter 5.

MAPD Bids and Related Variables

For both MA and Part D, CMS pays plans a risk-adjusted monthly capitation amount that is determined through a competitive bidding process. Because of risk adjustment, the presence of other forms of plan payment (such as reinsurance in Part D) and other complexities of the bidding and payment system, plan payments PMPM are not identical to the bid in either MA or Part D. Even so, the plan bid in both MA and Part D can be interpreted as a measure of the plan's anticipated costs of providing coverage.

Plans submit bids for MA and Part D coverage to CMS with extensive supporting information using a structured Microsoft Excel workbook known as the Bid Pricing Tool (BPT). Bids for coverage in a given year (sometimes called the *contract year*) must be submitted in early June of the calendar year preceding the contract year: Bids for 2020 coverage were prepared and submitted by June 2019, bids for 2021 coverage were prepared and submitted by June 2020, and so on.

We focus primarily on the standardized MAPD bid, defined as the sum of standardized bids submitted by MAPD plans for MA coverage and Part D coverage. *Standardized*, in this context, means that a bid amount—which is calculated under a plan's assumptions about the risk scores of the anticipated enrollee population—is divided by the average risk score projected by the plan to yield the bid amount that the plan would submit for an enrollee population with a risk score of 1. As discussed in Chapter 5, we analyze the standardized bid amounts in MA and Part D to

facilitate comparison of bids across plans that may have different enrollee characteristics and risk scores.

OACT worked with RAND to extract bid data from BPT spreadsheets submitted for MA and Part D coverage in 2017 through 2021. Standardized MA and Part D bids are reported on the BPT spreadsheets that plans submit to CMS. Although results shown in Chapters 5 and 8 are limited to MAPD bids, this appendix breaks down VBID impacts on the MAPD bid into impacts on MA and Part D bids and on specific components of health care costs that are factored into the MA and Part D bids.

MA Bids and Bid Components

MA bids represent the projected costs of providing Medicare coverage to beneficiaries for the calendar year, including net plan spending on Medicare-covered services as well as portions of nonbenefit expenses and the gain/loss margin that are allocated to Medicare-covered services (as opposed to supplemental benefits). MA bids are submitted as monthly per-beneficiary cost estimates, which are standardized to reflect a 1.0 beneficiary risk score. Focusing on standardized bids facilitates comparison of the bids across different plans with varying risk scores.

In additional analyses reported in this appendix, we estimate how specific components of the MA bid change when plans implement VBID interventions. OACT also provided RAND with data from the BPT on projected net plan spending on Medicare-covered services, nonbenefit expenses (including administrative costs), and the plan's gain/loss margin. We also analyze the projected MA risk score. These four variables are related to the standardized MA bid as follows:

$$\text{Standardized MA bid} = \frac{\text{A/B-Covered Net PMPM} + \text{A/B Nonbenefit Expenses} + \text{A/B Gain/Loss}}{\text{Projected MA Risk Score,}}$$

where "A/B-Covered Net PMPM" denotes PMPM net medical spending on Medicare-covered services, "A/B Nonbenefit Expenses" denotes the portion of nonbenefit expenses allocated to Medicare-covered services, and A/B Gain/Loss denotes the portion of the gain/loss margin allocated to Medicare-covered services.

Part D Bids

Data on Part D bids for VBID and comparison plans were extracted by OACT from BPT spreadsheets submitted for Part D coverage in 2017 through 2021. As with MA bids, Part D bids represent the projected cost to plans of providing prescription drug coverage to beneficiaries for the calendar year, including net plan spending as well as portions of nonbenefit expenses and gain/loss margin that are allocated to coverage for Part D–covered drugs for the basic Part D

benefit. We analyze standardized Part D bids, which represent the plan's bid adjusted for an enrollee population with a risk score of 1.

In additional analyses reported in this appendix, we estimate how specific components of the standardized Part D bid change when plans implement VBID interventions. Part D bids are driven primarily by plans' projected cost of providing the *basic benefit* that is defined by statute and CMS regulations. Plans typically also offer Part D supplemental benefits that are not paid for by CMS. The Part D bid reflects the projected costs for basic coverage, along with portions of the plan's nonbenefit expenses and gain/loss margin that are allocated to basic coverage.

In addition to the standardized Part D bid, OACT provided RAND with data from the BPT on basic nonbenefit expenses (including administrative costs), the basic gain/loss margin, and the plan's projected Part D risk score. These quantities are related to the standardized bid as follows:

$$\text{Standardized Part D Bid} = \frac{\text{Part D-Covered Net PMPM} + \text{Basic Nonbenefit Expenses} + \text{Basic Gain/Loss}}{\text{Projected Part D Risk Score,}}$$

where "Part D-Covered Net PMPM" denotes PMPM net spending on prescription drugs covered under the basic benefit, "Basic Nonbenefit Expenses" denotes the portion of nonbenefit expenses allocated to the basic benefit, and "Basic Gain/Loss" denotes the portion of the gain/loss margin allocated to the basic benefit. OACT did not provide RAND with the plans' projected cost of basic Part D coverage, so we derived this cost from the bid and the other components described above.

We estimate DD regression models for each of the variables that determine the standardized bid to provide additional insight into mechanisms through which VBID may affect Part D bids.

Total (MAPD) Premiums and Related Variables

We analyzed three different beneficiary cost variables: MA premiums, Part D premiums, and total MAPD premiums. In Chapters 5 and 8, we focus primarily on total MAPD premiums, which are simply the sum of MA and Part D premiums.

We obtained monthly plan-level premium data for MA and Part D coverage for the years 2017 through 2021. We obtained premium data from the publicly available plan landscape files and from beneficiary-level IDR MA and Part D premium data for plans with missing values in the public files. We provide further background on MA and Part D premiums and describe additional variables RAND received from BPT spreadsheets, a number of which we analyze in this appendix to explore mechanisms for observed premium changes associated with VBID.

Medicare Advantage Premiums

Beneficiary premiums in MA are determined by the plan's standardized MA bid, the MA benchmark (a measure derived from average Medicare FFS spending in the plan's service area,

which is adjusted based on a plan's quality rating and other factors), and the cost of supplemental benefits not covered by Medicare Part A and Part B (including both additional services and reduced cost sharing on covered services). Premiums can be derived as a basic premium (an amount charged to cover Medicare-covered services) and a supplemental premium (an amount charged to cover mandatory and optional supplemental benefits) (CMS, 2014; Medicare Payment Advisory Commission, 2021a).

A plan that bids below its benchmark has a basic premium of zero (for MA-covered services) but may need to charge a premium to cover any supplemental benefits. For a plan that bids above its benchmark, the enrollee must pay the difference between the bid and the basic premium in addition to any supplemental premium. Most plans that bid below the benchmark receive a rebate ("the MA rebate") that is calculated as a percentage of the difference between the benchmark and the bid, with the percentage tied to the plan's quality rating and other factors (CMS, 2014; Medicare Payment Advisory Commission, 2021a). Plans are required to pass the rebate on to beneficiaries through various channels. Under typical MA rules outside of the VBID model test, these channels include spending on additional services, spending on reduced cost sharing for Medicare-covered services, buying down the Part B premium, or buying down components of the Part D premium (as we discuss below). As discussed in Chapter 1, the Cash Rebates subcomponent of the VBID model test allows plans to distribute a portion of the rebate directly to beneficiaries.

In this appendix, we estimate regression models for the MA rebate (in dollars PMPM) and three categories of MA rebate allocations with the potential to affect the MA premium: the rebate allocation for mandatory supplemental benefits (additional services), the rebate allocation for reduced cost sharing for Part A and B services, and the rebate allocation for Part B premium buydown. Reallocation of the rebate toward or away from Part D premium buydown could also affect MA premiums; estimates of VBID impacts on Part D premium buydown are discussed below in the context of the Part D premium. Because our regression models control for the benchmark (as discussed in Appendix C), changes in the MA rebate should largely track changes in the MA bid but may differ because plans do not receive a rebate if they bid above the benchmark. An increase in the MA rebate could help plans reduce MA premiums if benefit design and projected medical spending are held constant, but only if the plan chooses to use the rebate in ways that may reduce the MA premium.

Part D Premiums and Premium Components

The Part D premium reflects a *basic* premium that covers a portion of the cost of standard Part D coverage and a *supplemental* premium that covers the cost of supplemental benefits. A plan's basic Part D premium is calculated based on a comparison of the standardized plan bid with the average of standardized bids submitted by all Part D plans (Medicare Payment Advisory Commission, 2021b). Beneficiaries eligible for the Part D LIS receive premium subsidies that

may cover all or most of their plan's premium. These premium subsidies are calculated at the regional level (CMS, 2018).

In this evaluation, we use data from the CMS HPMS on the total Part D premium, the basic Part D premium, and the supplemental Part D premium charged to beneficiaries in each plan. These amounts reflect premium buydowns that plans may offer using MA rebate dollars. To explore mechanisms for Part D premium changes, we estimate regression models for four additional measures related to the total Part D premium: the basic premium before any buydown, the supplemental premium before any buydown, and the amounts allocated for basic premium buydown and supplemental premium buydown. These variables are related to the Part D premium as follows:

Part D premium = (basic premium before buydown – basic premium buydown) + (supplemental premium before buydown – supplemental premium buydown)

Examining VBID impacts on these components separately can tell us whether premium changes are driven by the costs of basic or supplemental coverage or by changes in premium buydown.

Supplemental Benefits

We descriptively characterized supplemental benefits using the publicly available PBP benefits data (CMS, 2021a). These data contain information on the benefit designs of MA plans, including premiums, deductibles, OOP maximums, and cost sharing for specific services. There are also a series of fields describing the supplemental benefits offered by each plan.

We also analyzed plans' PMPM projected costs of additional services for 2017 through 2021, which were extracted from plan BPT spreadsheets and provided to RAND by OACT. This cost variable reflects the amount of spending on mandatory supplemental benefits not covered by Medicare Parts A and B. Although mandatory supplemental benefits also encompass reductions in cost sharing on Medicare-covered services, we determined in consultation with independent actuarial experts that projected plan costs for additional services would most accurately capture the generosity of supplemental benefits.

Summary Statistics for Bids, Premiums, and Mandatory Supplemental Benefit Costs

Table I.1 presents the mean and standard deviation of bids, premiums, and mandatory supplemental benefit costs. Statistics are reported for 2019, the last year before the current VBID model test began. The "VBID" column reflects all plans participating in any component of VBID in either 2020 or 2021, while the "Comparison" column reflects all eligible comparison plans used as comparators for either the BDI or the Hospice component. We note that the sample size of plans varies across outcomes because of missing bid data for several plans.

Table I.1. Mean and Standard Deviation of Outcome Measures Prior to VBID Participation (2019)

Outcome Variable	VBID	Comparison
MAPD bid		
Mean	828.75	828.33
SD	87.01	91.1
N	355	1,928
MA bid		
Mean	787.93	775.25
SD	(85.76)	(87.17)
N	355	1,928
Part D bid		
Mean	40.82	53.08
SD	(10.79)	(19.21)
N	355	1,928
MAPD premium		
Mean	25.19	34.75
SD	(33.42)	(48.33)
N	355	1,928
MA premium		
Mean	9.89	15.99
SD	(24.97)	(32.99)
N	355	1,928
Part D premium		
Mean	15.31	18.76
SD	(15.46)	(21.34)
N	355	1,928
MSB net PMPM		
Mean	32.51	22.92
SD	(29.72)	(16.93)
N	355	1,928

SOURCE: RAND analysis of 2019 CMS Medicare Part D landscape files and 2019 OACT bid data.

Balance and Parallel Trends

As discussed in Appendix C, we analyzed the effects of VBID participation by defining four groups of participating plans based on the years when each plan participated in VBID and on whether the plan adopted the BDI or the Hospice component. Impacts of VBID participation for each group were estimated separately using the DD approach described in Appendix C. Briefly, plans in each group were analyzed together with a comparison group of nonparticipating plans that were also eligible for VBID. For each group of plans and each outcome, entropy balancing was used to reweight the comparison group to ensure parallel pre-intervention trends in outcomes. The entropy balancing algorithm also accounted for time-invariant plan characteristics and pre-intervention (2019) levels of selected variables other than the outcome in order to make

the comparison group as similar as possible to each group of VBID-participating plans. See Appendix C for further discussion and technical details.

In the remainder of this appendix, we first present balance tables and graphs illustrating pre-intervention trends in outcomes. We then present group-specific regression tables for each outcome. Supplementary regression results for additional outcome variables that inform our understanding of mechanisms underlying the main results are presented later in the appendix.

Balance and Parallel Trends for Analysis of MAPD Plan Bids

This section shows balance tables and graphs illustrating pre-intervention trends for the key outcomes reported in Chapters 5 and 8: MAPD bids, total (MAPD) premiums, and net PMPM plan spending on mandatory supplemental benefits not covered by Medicare.

For each outcome, we estimated separate DD models for four participant groups, resulting in a large volume of balance tables and pre-intervention trend figures. Figures showing pre-intervention trends in outcomes are provided for all models, but we omit full balance tables for most models in the interest of brevity. Instead, we summarize covariate balance for most models by reporting the average (across all covariates) ASMD between VBID and comparison plans before and after weighting. We also omit balance tables and pre-intervention trend figures for outcomes analyzed in the supplementary analyses that are reported only in this appendix (including MA bids, Part D bids, MA premiums, Part D premiums, and bid variables that feed into these outcomes).

Table I.2 shows covariate balance before and after weighting for our analysis of BDI intervention impacts on MAPD bids among plans that participated in both 2020 and 2021. Rows of the table correspond to pre-intervention (2019 plan year) covariates used in the entropy balancing algorithm. Pre-intervention changes in outcomes, which were also used in the entropy balancing algorithm, are shown graphically in Figure I.1. For each pre-intervention covariate, the table shows group means for VBID plans that participated in both 2020 and 2021 alongside both unweighted and weighted comparison group means. The ASMD between VBID and comparison groups is also reported before and after weighting, allowing a comparison of balance across variables that differ in scale. Finally, the table reports VBID and comparison group standard deviations for each variable, as we also selected entropy balancing weights to balance the variance of pre-intervention covariates between VBID and comparison groups in addition to balancing the mean. The number of participating plans included in this analysis was 89, and the effective number of comparison plans after weighting was 106.

Table I.2 shows that entropy balancing was able to dramatically improve balance on pre-intervention covariates between VBID and comparison plans: Unweighted ASMDs of as much as 0.5 standard deviation were reduced to 0.01 standard deviation or less. In addition to balancing the first moments (means) of pre-intervention covariates, weighting also leads to much better balance on the second moments (standard deviations) of pre-intervention covariates.

Table I.3 summarizes balance before and after weighting for each participant group for each of the outcomes analyzed in Chapters 5 and 8. As with our analysis of MAPD bids among plans that participated in both 2020 and 2021 (Table I.2), weighting leads to major improvements in balance between VBID and comparison groups.

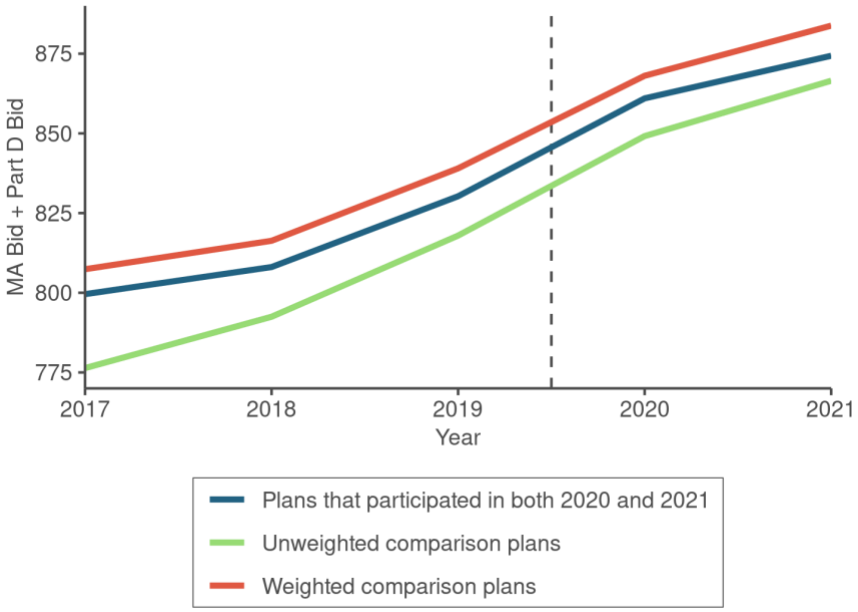
Table I.2. Mean, Standard Deviation, and Absolute Standardized Mean Differences Comparing Plans That Participated in Both 2020 and 2021 with Eligible Comparison Plans, Before and After Weighting for MAPD Bid Regression Models

	Mean (VBID)	Unweighted Mean (Comp.)	Weighted Mean (Comp.)	Unweighted ASMD	Weighted ASMD	SD (VBID)	Unweighted SD (Comp.)	Weighted SD (Comp)
Plan characteristics								
MA premiums	10.02	11.31	10.67	0.04	0.02	32.43	28.33	32.67
Part D premiums	18.31	16.41	18.62	0.09	0.01	20.94	19.56	21.13
Cost of MSB	31.15	26.29	30.63	0.19	0.02	26.18	18.27	25.63
Rebate dollars amount	105.05	120.11	106.00	0.32	0.02	47.51	73.57	47.79
Administrative costs	114.83	121.12	115.11	0.44	0.02	14.28	41.74	15.90
Part C in-network OOP maximum	5,419.08	5,056.56	5,391.61	0.26	0.02	1,373.22	1,637.08	1,403.21
PPO	0.21	0.21	0.22	0.00	0.01	0.41	0.41	0.41
For profit	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
Star Rating	4.08	3.98	4.07	0.32	0.02	0.31	0.50	0.32
C-SNP	0.01	0.07	0.01	0.52	0.02	0.11	0.25	0.11
D-SNP	0.26	0.16	0.25	0.24	0.02	0.44	0.36	0.43
Moved into bonus payment	0.02	0.08	0.03	0.37	0.02	0.15	0.27	0.16
Moved out of bonus payment	0.04	0.06	0.04	0.08	0.02	0.21	0.24	0.20
Individual characteristics at the plan level								
Average age	70.69	70.89	70.77	0.05	0.02	3.95	4.06	3.97
Percent male	0.43	0.44	0.42	0.33	0.02	0.05	0.07	0.05
Percent AI/AN	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
Percent API	0.03	0.04	0.03	0.14	0.02	0.06	0.09	0.06
Percent Black	0.12	0.13	0.12	0.12	0.02	0.14	0.15	0.14
Percent Hispanic	0.09	0.16	0.09	0.46	0.02	0.16	0.22	0.16
Percent multirace	0.03	0.02	0.02	0.20	0.02	0.01	0.01	0.01
Percent White	0.73	0.63	0.72	0.43	0.02	0.23	0.28	0.23
Percent dually eligible for Medicaid	0.33	0.29	0.34	0.09	0.02	0.40	0.35	0.40
Average MA risk score	1.29	1.26	1.29	0.09	0.02	0.34	0.34	0.33
Average Part D risk score	1.14	1.12	1.15	0.06	0.02	0.33	0.35	0.33

	Mean (VBID)	Unweighted Mean (Comp.)	Weighted Mean (Comp.)	Unweighted ASMD	Weighted ASMD	SD (VBID)	Unweighted SD (Comp.)	Weighted SD (Comp)
Percent diabetes	0.30	0.31	0.30	0.13	0.02	0.08	0.14	0.08
Percent CHF	0.13	0.13	0.14	0.05	0.02	0.05	0.07	0.06
Percent COPD	0.19	0.17	0.18	0.20	0.02	0.08	0.10	0.08
Percent cancer	0.09	0.09	0.09	0.26	0.02	0.03	0.03	0.03
Area characteristics of the plans								
Median income	30,932.28	30,183.95	30,840.17	0.16	0.02	4,604.86	5,026.48	4,692.71
MA penetration	44.54	41.88	44.33	0.25	0.02	10.51	12.02	10.68
Percent rural	4.45	5.80	4.59	0.20	0.02	6.76	9.47	7.17
Percent suburban	14.66	15.87	14.86	0.10	0.02	11.70	14.93	11.95
Percent urban	80.88	78.32	80.55	0.15	0.02	16.87	20.74	16.88
HPSA	2.19	4.15	2.32	0.30	0.02	6.56	10.14	6.79
Standardized Medicare costs per capita	10,289.67	10,754.16	10,316.20	0.35	0.02	1,325.89	1,466.55	1,320.82
Percent over 65	0.19	0.18	0.18	0.21	0.02	0.04	0.04	0.04

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: The entropy balancing weights used in this analysis were estimated using constraints on both the first and second moments of each characteristic, ensuring balance between the participating and comparison plans on means and variances.

Figure I.5. Trends in MAPD Bids for Plans That Participated in Both 2020 and 2021 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

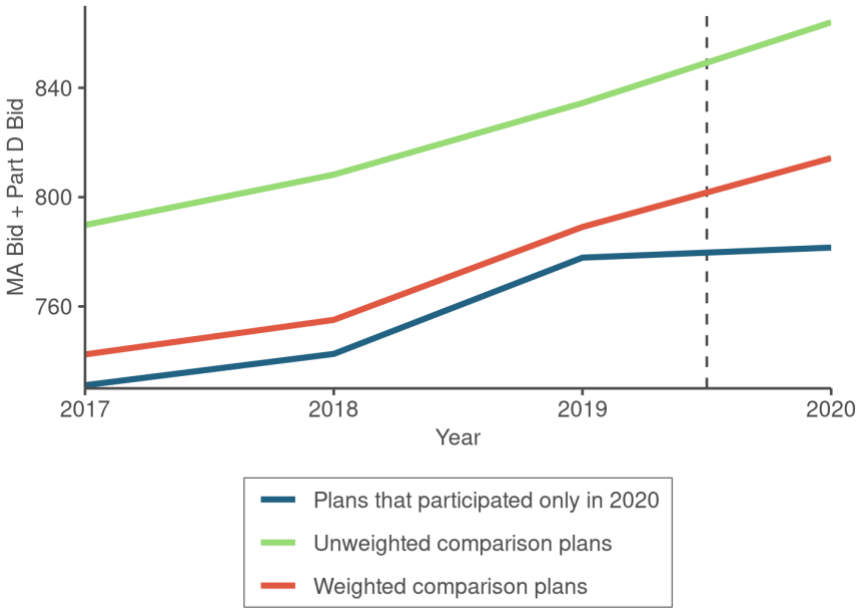
Table I.3. Balance on Pre-Intervention Plan Characteristics Before and After Weighting, by Outcome Variable and Participant Group

Outcome	Participant Group	Average ASMD Before Weighting	Average ASMD After Weighting	Number of Participating Plans	Effective Number of Comparison Plans After Weighting
MAPD bid	2020 and 2021	0.205	0.019	89	114
	2020 only	0.361	0.045	39	42
	2021 only	0.354	0.057	257	288
	Hospice	0.358	0.027	46	49
MAPD premium	2020 and 2021	0.222	0.035	89	102
	2020 only	0.373	0.054	39	60

Outcome	Participant Group	Average ASMD Before Weighting	Average ASMD After Weighting	Number of Participating Plans	Effective Number of Comparison Plans After Weighting
	2021 only	0.378	0.070	257	272
	Hospice	0.371	0.046	46	47
MSB net PMPM	2020 and 2021	0.223	0.034	89	116
	2020 only	0.373	0.053	39	47
	2021 only	0.378	0.047	257	261
	Hospice	0.375	0.074	46	49

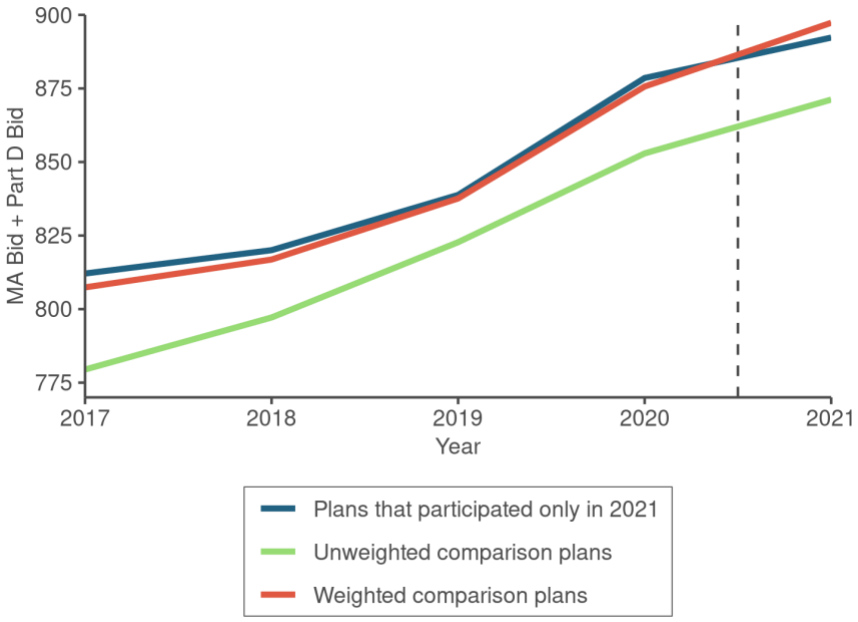
SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: Table reports average of ASMD across all variables used in weighting other than lagged outcomes. See Table C.5 for list of pre-intervention plan characteristics used in weighting. The entropy balancing weights used in this analysis were estimated using constraints on both the first and second moments of each characteristic, ensuring balance between the participating and comparison plans on means and variances. Balance was achieved on both means and variances: this table presents only differences in means in the interest of conciseness.

Figure I.2. Trends in MAPD Bids for Plans That Participated Only in 2020 and Comparison Plans, Before and After Weighting



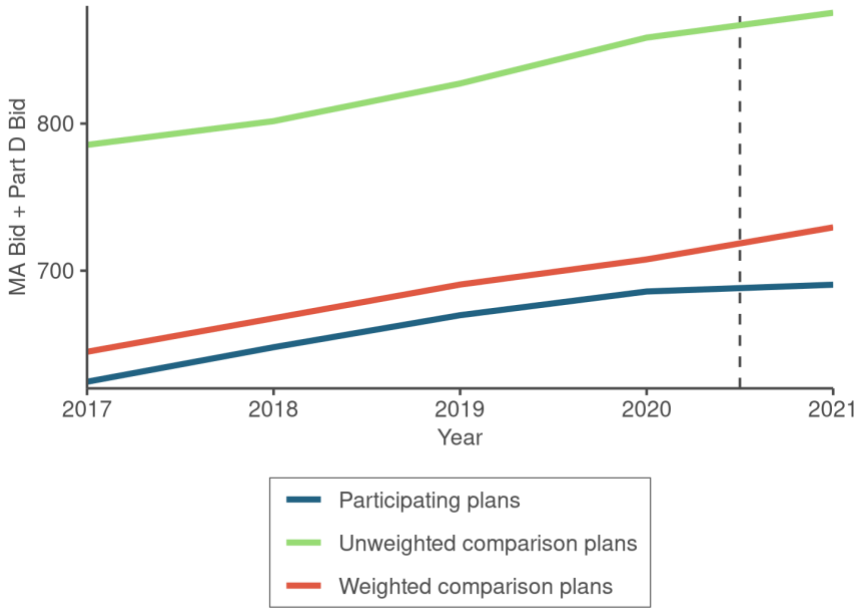
SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Figure I.3. Trends in MAPD Bids for Plans That Participated Only in 2021 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

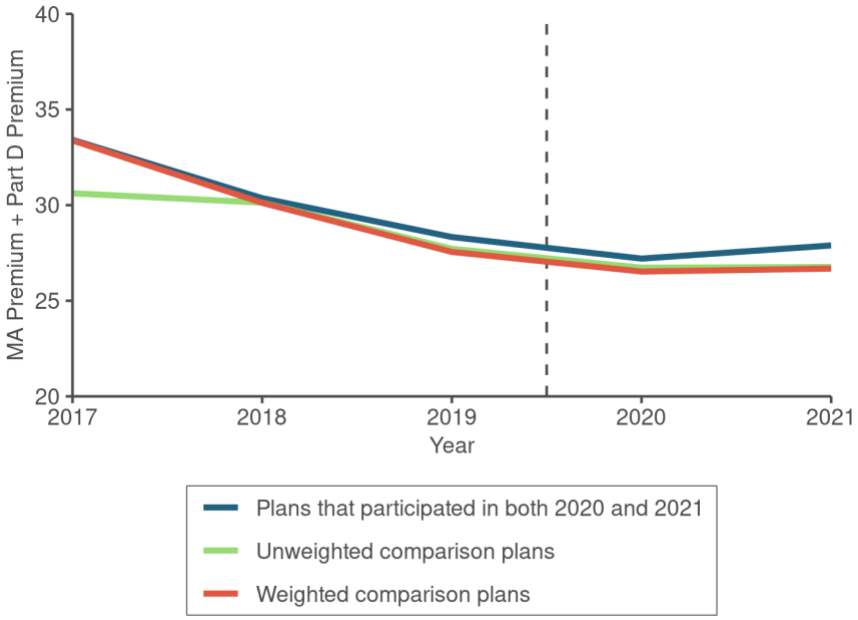
Figure I.4. Trends in MAPD Bids for Plans That Participated in Hospice and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

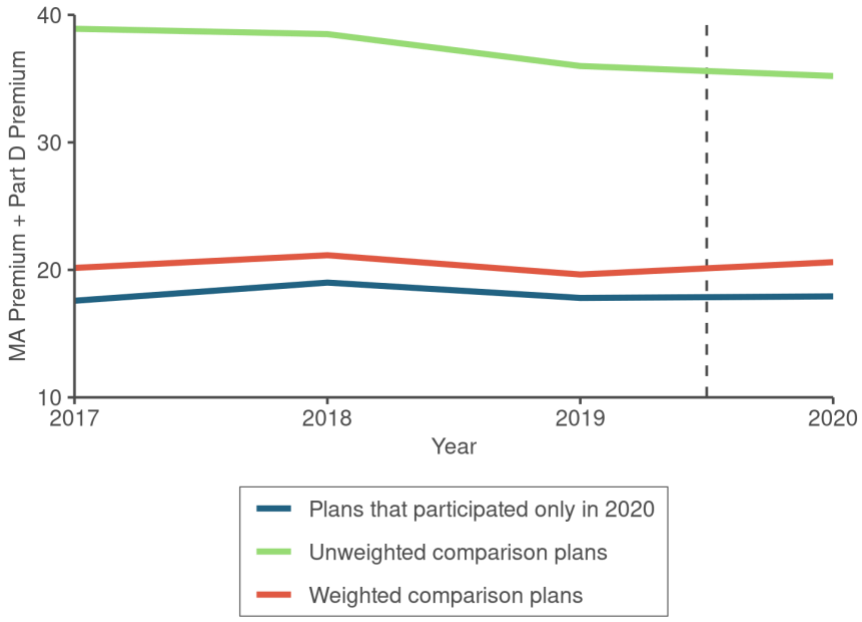
Balance and Parallel Trends for Analysis of Total (MAPD) Premiums

Figure I.5. Trends in MAPD Premiums for Plans That Participated in Both 2020 and 2021 and Comparison Plans, Before and After Weighting



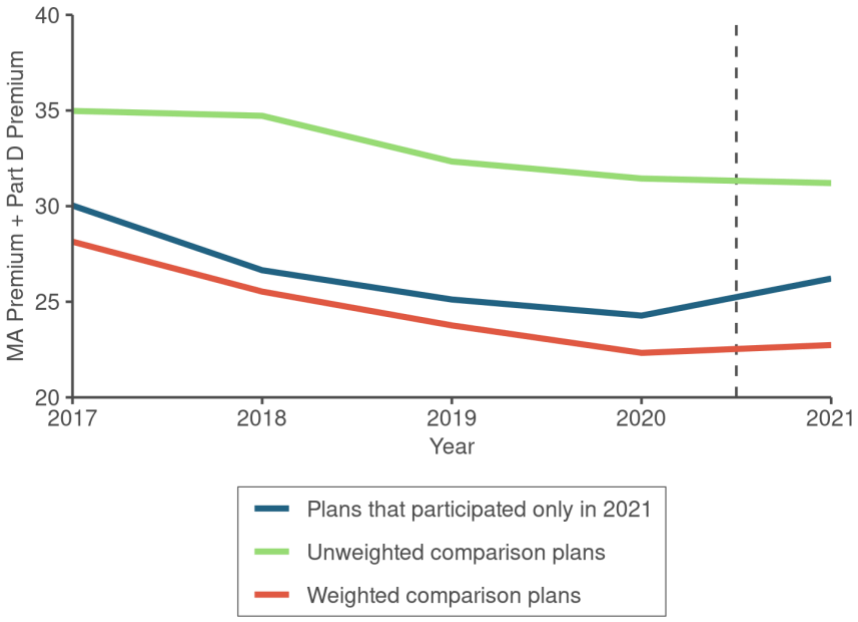
SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Figure I.6. Trends in MAPD Premiums for Plans That Participated Only in 2020 and Comparison Plans, Before and After Weighting



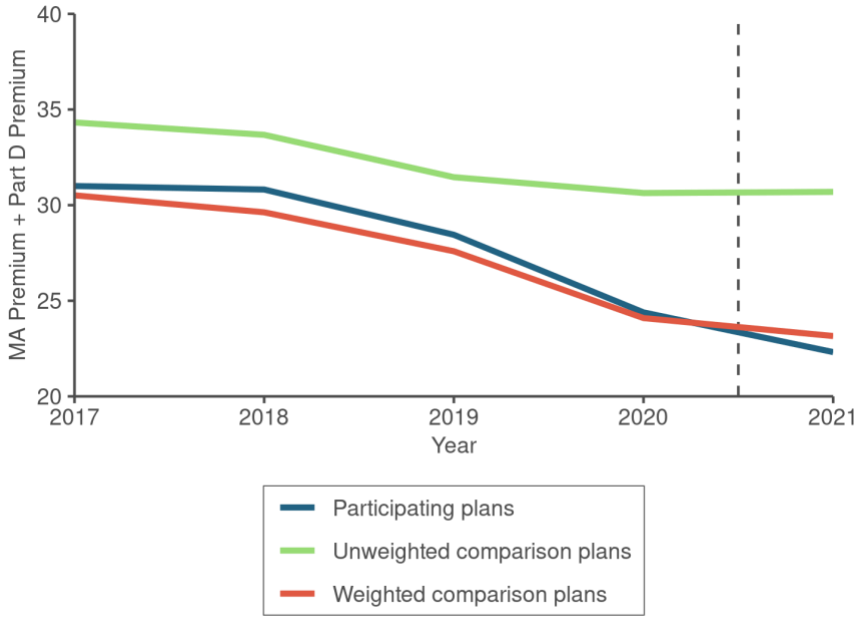
SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Figure I.7. Trends in MAPD Premiums for Plans That Participated Only in 2021 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

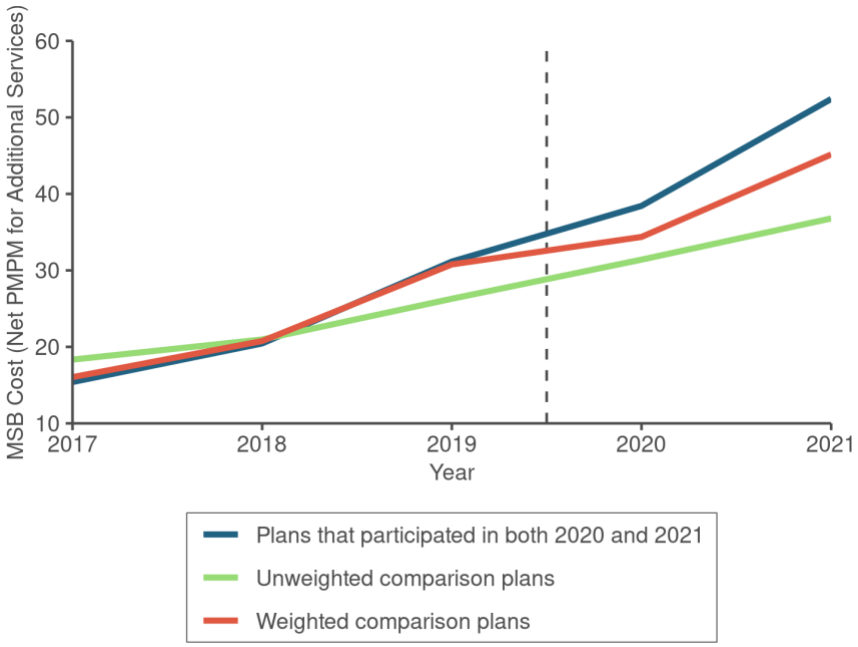
Figure I.8. Trends in MAPD Premiums for Plans That Participated in Hospice and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

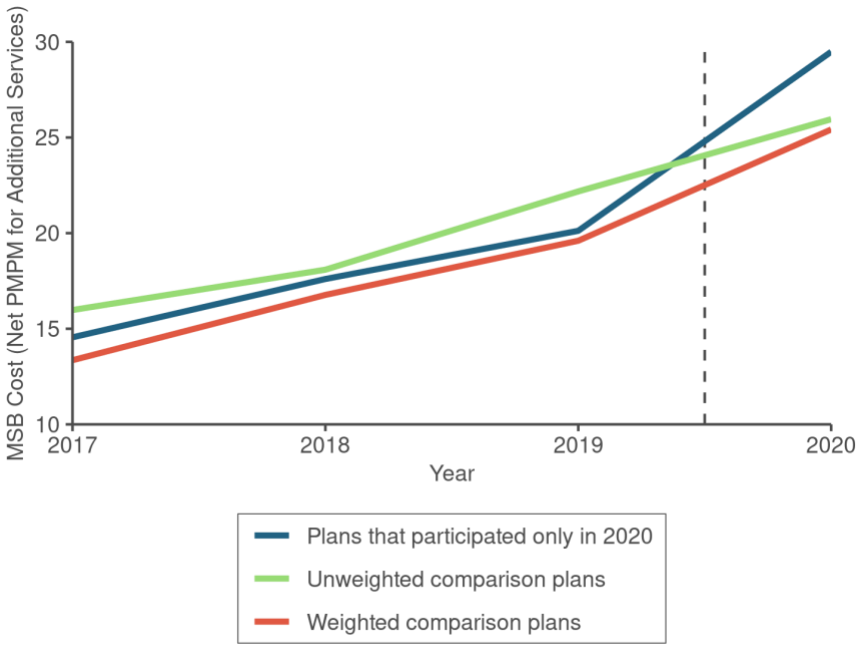
Balance and Parallel Trends for Analysis of Mandatory Supplemental Benefit Costs

Figure I.9. Trends in Mandatory Supplement Benefit Costs for Plans That Participated in Both 2020 and 2021 and Comparison Plans, Before and After Weighting



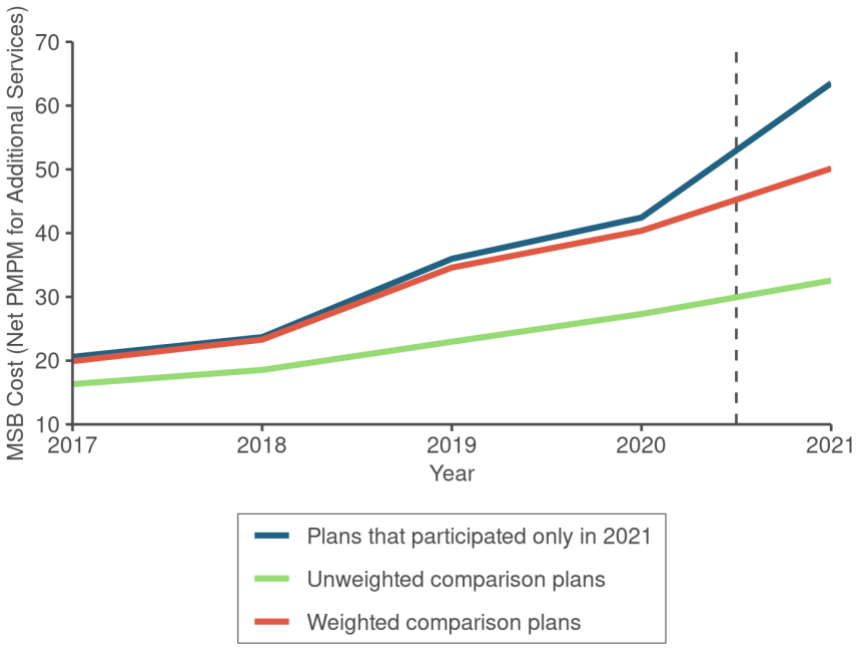
SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Figure I.10. Trends in Mandatory Supplement Benefit Costs for Plans That Participated Only in 2020 and Comparison Plans, Before and After Weighting



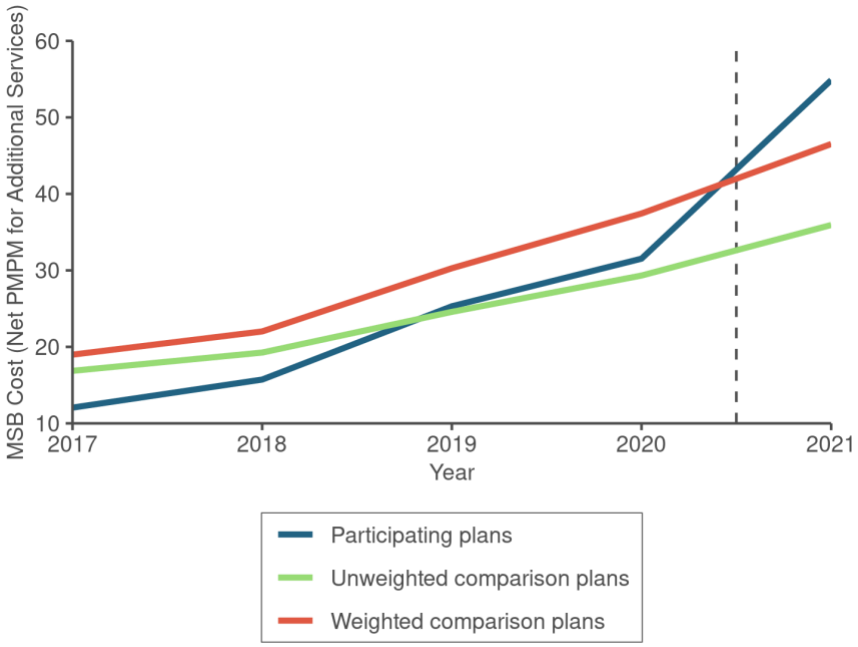
SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Figure I.11. Trends in Mandatory Supplement Benefit Costs for Plans That Participated Only in 2021 and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Figure I.12. Trends in Mandatory Supplement Benefit Costs for Plans That Participated in Hospice and Comparison Plans, Before and After Weighting



SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTES: Years to the right of the dashed vertical line are after VBID implementation. Years to the left of the dashed vertical line are pre-intervention years; changes in pre-intervention outcomes were used in the weighting algorithm. The parallel trends assumption is an inherently untestable assumption that the group trends in the post-participation period would be parallel absent VBID participation. This figure provides an assessment of the similarity of trends in the pre-participation period, which only implies similarity of the trends in the post-participation period if we assume that the similarity persists between the pre- and post-participation periods.

Group-Specific Difference-in-Differences Regression Model Results

Below we present results of our empirical analysis for *bids*, *premiums*, and *supplemental benefits*. We begin by evaluating the assumption that trends in outcomes were parallel between VBID and comparison plans before VBID implementation. We also present descriptive figures illustrating how outcomes for VBID plans changed after VBID implementation. We then present regression tables showing the results discussed in the report. Finally, we discuss limitations of our analysis and report sensitivity analyses intended to address those limitations.

Models for all outcomes were estimated using weighted least squares with *entropy balancing* weights constructed as described in Appendix C. Statistical inference was based on a cluster bootstrap using 30 bootstrap resamples (with clustering at the plan level), allowing for arbitrary correlation of the error term within plans over time. See Appendix C for further discussion.

Difference-in-Differences Regression Model Results for MAPD Bids

Table I.4. Difference-in-Differences Model Results for Analysis of Plans That Participated in Both 2020 and 2021: MAPD Bids

Outcome and Participant Group	2020 Estimate	2020 Standard Error	2020 95% CI Lower Bound	2020 95% CI Upper Bound	2020 p-value	2021 Estimate	2021 Standard Error	2021 95% CI Lower Bound	2021 95% CI Upper Bound	2021 p-value
MAPD bids										
BDI 2020 and 2021	1.46	3.02	–4.45	7.37	0.63	–0.71	4.01	–8.58	7.15	0.86
BDI 2020 only	–22.35	8.74	–39.47	–5.23	0.01	N/A	N/A	N/A	N/A	N/A
BDI 2021 only	N/A	N/A	N/A	N/A	N/A	–6.98	2.24	–11.37	–2.60	0.00
Hospice	N/A	N/A	N/A	N/A	N/A	–14.79	10.02	–34.42	4.84	0.14
Total (MAPD premiums)										
BDI 2020 and 2021	0.38	0.95	–1.48	2.24	0.69	1.03	1.14	–1.20	3.25	0.37
BDI 2020 only	–1.54	2.30	–6.05	2.97	0.50	N/A	N/A	N/A	N/A	N/A
BDI 2021 only	N/A	N/A	N/A	N/A	N/A	2.25	0.55	1.17	3.33	0.00
Hospice	N/A	N/A	N/A	N/A	N/A	–4.44	3.46	–11.22	2.34	0.20
MSB cost										
BDI 2020 and 2021	3.63	2.33	–0.93	8.18	0.12	6.53	3.47	–0.27	13.34	0.06
BDI 2020 only	1.78	2.93	–3.97	7.52	0.54	N/A	N/A	N/A	N/A	N/A
BDI 2021 only	N/A	N/A	N/A	N/A	N/A	13.02	1.75	9.58	16.46	0.00
Hospice	N/A	N/A	N/A	N/A	N/A	1.91	5.15	–8.17	12.00	0.71

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.
NOTE: Plan-level fixed effects are excluded from the table.

Summaries of Difference-in-Differences Model Results

For plans that participated in BDI interventions, we summarized the estimated group-specific impacts of VBID participation by calculating average effects by plan year (2020 or 2021) and by year of BDI participation (first year or second year). The summary parameters reflecting BDI effects for 2020 and 2021 were presented graphically and discussed in Chapter 5. Effects by year associated with BDI participation (first year or second year) were not discussed in the report, and we note that these effects should not be interpreted as reflecting dynamic effects of BDI interventions since changes in composition of participating plans between 2020 and 2021 also contribute to differences between the first-year and second-year effects. Plans participating in both 2020 and 2021 are the only plans for which we currently have evidence on dynamic effects of BDI participation. Readers interested in these dynamics should examine Figures I.1, I.5, and I.9 above.

This section reports the underlying regression coefficients, standard errors, and 95-percent CIs for these parameters. Statistical inference was based on a cluster bootstrap using 500 bootstrap resamples (with clustering at the plan level), allowing for arbitrary correlation of the error term within plans over time. See Appendix C for further discussion.

We note that, in this interim evaluation, there was no need to summarize or aggregate estimated impacts of Hospice component participation across multiple groups of plans with different participation histories because there was only one year of model implementation for these plans.

Summaries of Difference-in-Differences Regression Model Results for MAPD Bids

Table I.5. Estimated Association Between VBID Participation and MAPD Bids

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
By plan year					
2020	–5.79	3.37	–12.39	0.81	0.09
2021	–5.37	2.00	–9.30	–1.44	0.01
By year of participation					
First	–6.59	1.88	–10.27	–2.91	0.00
Second	–0.71	4.01	–8.58	7.15	0.86

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Summaries of Difference-in-Differences Regression Model Results for Total (MAPD) Premiums

Table I.6. Estimated Association Between VBID Participation and MAPD Premiums

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
By plan year					
2020	–0.21	0.96	–2.08	1.67	0.83
2021	1.93	0.53	0.89	2.97	0.00
By year of participation					
First	1.43	0.49	0.48	2.39	0.00
Second	1.03	1.14	–1.20	3.25	0.37

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Summaries of Difference-in-Differences Regression Model Results for Mandatory Supplemental Benefit Costs

Table I.7. Estimated Association Between VBID Participation and Mandatory Supplemental Benefit Costs

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
By plan year					
2020	3.06	1.79	–0.44	6.57	0.09
2021	11.35	1.54	8.34	14.36	0.00
By year of participation					
First	9.71	1.27	7.23	12.19	0.00
Second	6.53	3.47	–0.27	13.34	0.06

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Additional Results on Supplemental Benefit Offerings

Table I.8 shows the average number of supplemental benefits offered by plans, across the different participation statuses. See Chapter 5 for discussion of these findings.

Table I.8. Average Number of Supplemental Benefits Offered, by Participation Status and Year

Participation Status	2019	2020	2021
2020 only	14.6	16.1	17.0
2021 only	18.1	19.3	19.3
Both 2020 and 2021	19.2	20.7	21.1
Hospice	16.5	17.6	19.0
Eligible nonparticipants	15.8	17.5	18.7

SOURCE: RAND analysis of publicly available PBP benefits data.
NOTE: CMS added three supplemental benefits to the list of potential supplemental benefits in 2020.

Tables I.9 through I.12 present the percentage of beneficiaries enrolled in plans in each participation category, who had access to specific supplemental benefits in that year (2019, 2020, and 2021). These tables show the supplemental benefits offered in each year, providing additional context for the overall numbers of supplemental benefits presented in the main report and shown above in Table I.8. In addition, the tables provide more detail on the percentage of beneficiaries who had access to each specific supplemental benefit in each year, since this varies substantially by benefit and (to a lesser extent) across the BDI and Hospice participation categories. Tables I.9 and I.11 present these results for a subset of supplemental benefits that had at least 15 percent of enrollees with access in one or more of the plan participation categories. Tables I.10 and I.12 present the percentage of enrollees with access to the new primarily health-related benefits.

Table I.9. Percentage of Enrollees Offered Specific Supplemental Benefits by BDI Participation Status, 2019–2021

	2020 Only			2021 Only			Both 2020 and 2021			Eligible Nonparticipants		
	2019	2020	2021	2019	2020	2021	2019	2020	2021	2019	2020	2021
Additional days	63	63	62	100	100	100	99	98	98	89	88	89
Worldwide emergency coverage	94	94	98	99	100	100	99	98	98	96	96	97
Routine care	18	20	25	24	33	42	27	28	30	13	15	17
Routine foot care	63	62	62	48	50	57	44	56	51	44	46	47
Additional telehealth benefit for Part B services	N/A	61	100	N/A	100	100	N/A	87	96	N/A	70	95
Plan-approved health-related location	40	41	45	63	63	70	54	67	64	36	38	41
Acupuncture	38	35	37	30	25	27	25	32	32	23	25	25
Limited duration meal benefit	31	30	33	85	91	92	49	79	82	36	43	55
Annual physical exam	77	76	74	56	98	99	82	99	99	75	85	88
Health education	18	20	22	6	6	6	11	12	16	32	31	34
Nutritional/dietary benefit	12	20	22	5	6	6	5	9	9	15	14	16
Additional sessions of smoking and tobacco cessation counseling	0	0	0	10	11	14	44	47	44	17	18	18
Fitness benefit	95	95	97	94	98	99	97	99	99	85	88	91
Enhanced disease management	0	0	0	2	1	1	3	5	5	7	7	7
Telemonitoring services	0	0	0	1	1	1	0	0	0	8	7	6
Remote access technologies (including web/phone-based technologies and nursing hotline)	98	87	87	99	29	39	100	37	29	92	79	80
Counseling services	0	0	0	2	1	1	18	15	11	3	3	4
In-home safety assessment	0	0	0	0	0	0	18	15	11	3	4	4

	2020 Only			2021 Only			Both 2020 and 2021			Eligible Nonparticipants		
	2019	2020	2021	2019	2020	2021	2019	2020	2021	2019	2020	2021
PERS	6	7	7	29	36	44	12	17	20	17	20	23
MNT	11	11	0	13	14	1	7	10	0	6	6	6
Alternative therapies	18	20	22	0	0	0	0	3	3	2	3	3
Therapeutic massage	N/A	20	22	N/A	0	0	N/A	1	0	N/A	6	2

SOURCE: RAND analysis of publicly available PBP benefits data.
NOTE: N/A = not applicable; the benefit was not included in the plan data for that year.

Table I.10. Percentage of Enrollees Offered New Primarily Health-Related Benefits by BDI Participation Status, 2019–2021

	2020 Only			2021 Only			Both 2020 and 2021			Eligible Nonparticipants		
	2019	2020	2021	2019	2020	2021	2019	2020	2021	2019	2020	2021
Home and safety devices	0	0	0	1	1	1	18	18	13	5	8	9
Transportation	40	41	45	63	63	70	54	67	64	36	38	41
OTC items	76	62	67	95	96	96	79	96	97	62	74	81
Adult day health services	0	0	0	0	0	0	0	0	0	0	4	4
Home-based palliative care	0	0	0	0	0	0	0	15	11	1	2	3
In-home support services	1	12	12	0	3	6	0	5	8	1	6	9
Caregiver support	0	0	0	0	17	0	0	22	15	0	2	3

SOURCE: RAND analysis of publicly available PBP benefits data.

Table I.11. Percentage of Enrollees Offered Specific Supplemental Benefits by Hospice Participation Status, 2019–2021

	Hospice			Eligible Nonparticipants		
	2019	2020	2021	2019	2020	2021
Additional inpatient days	75	75	75	89	88	89
Worldwide emergency coverage	100	100	100	96	96	97
Routine care	64	64	64	13	15	17
Routine foot care	56	56	62	44	46	47
Additional telehealth benefit for Part B services	N/A	42	59	N/A	70	95
Plan-approved health-related location	67	67	57	36	38	41
Acupuncture	58	55	56	23	25	25
Limited duration meal benefit	42	41	36	36	43	55
Annual physical exam	38	38	38	75	85	88
Health education	75	74	74	32	31	34
Nutritional/dietary benefit	15	31	30	15	14	16
Additional sessions of smoking and tobacco cessation counseling	45	44	44	17	18	18
Fitness benefit	41	40	42	85	88	91
Enhanced disease management	17	18	18	7	7	7
Telemonitoring services	5	5	5	8	7	6
Remote access technologies	98	68	74	92	79	80
Counseling services	21	21	21	3	3	4
In-home safety assessment	0	0	6	3	4	4
PERS	1	1	10	17	20	23
MNT	5	5	11	6	6	6
Alternative therapies	16	16	16	2	3	3
Therapeutic massage	N/A	0	0	N/A	6	2

SOURCE: RAND analysis of publicly available PBP benefits data.
NOTE: N/A = not applicable; the benefit was not included in the PBP benefits data for that year.

Table I.12. Percentage of Enrollees Offered New Primarily Health-Related Benefits by Hospice Participation Status, 2019–2021

	Hospice			Eligible Nonparticipants		
	2019	**2020**	**2021**	**2019**	**2020**	**2021**
Home and safety devices	0	1	7	5	8	9
Transportation	67	67	74	36	38	41
OTC Items	81	88	89	62	74	81
Adult day health services	0	0	0	0	4	4
Home-based palliative care	0	4	4	1	2	3
In-home support services	0	6	3	1	6	9
Caregiver support	0	0	0	0	2	3

SOURCE: RAND analysis of publicly available PBP benefits data.

Additional Regression Results Analyzing Mechanisms for Bid and Premium Changes

The MAPD bid is the sum of the standardized MA and Part D bids submitted by each plan. In this section, we present DD regression results for models analyzing MA and Part D bids separately, as well as results for selected additional variables that affect the MA and Part D bids. Impacts of BDI interventions are reported as summary parameters aggregated across the different intervention groups. Impacts of the Hospice component, which are estimated for one intervention group only, do not need to be aggregated.

Definitions of bid and premium components analyzed in these regression models were presented at the start of this appendix (in the subsection titled "Outcome Measures"), and results of note were discussed in Chapters 5 and 8. Results of interest are discussed in those chapters.

Difference-in-Differences Regression Model Results for Factors Explaining Changes in MAPD Bids

Factors Explaining Changes in MAPD Bids: MA Bids

Table I.13. Estimated Association Between VBID Participation and MA Bids

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
BDI effect by year					
2020	–7.30	3.78	–14.71	0.11	0.05
2021	–8.78	2.02	–12.74	–4.81	0.00
Hospice effect					
2021	–22.40	8.02	–38.12	–6.68	0.01

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Table I.14. Estimated Association Between VBID Participation (BDI) and MA Bid Components

Outcome	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Medicare-covered net PMPM					
BDI effect by year					
2020	–12.14	7.54	–26.91	2.64	0.11
2021	7.33	6.40	–5.22	19.87	0.25
Hospice effect					
2021	-26.79	16.31	-58.76	5.19	0.10
Nonbenefit expenses allocated to MA-covered services					
BDI effect by year					
2020	1.83	1.75	–1.60	5.26	0.30
2021	–5.46	1.24	–7.90	–3.03	0.00
Hospice effect					
2021	4.03	7.46	–10.60	18.65	0.59
Gain/loss alloc. to MA-covered services					
BDI effect by year					
2020	–4.96	5.75	–16.23	6.31	0.39
2021	1.71	4.32	–6.75	10.18	0.69
Hospice effect					
2021	–25.23	16.72	–57.99	7.53	0.13
Projected MA risk score					
BDI effect by year					
2020	0.00	0.01	-0.01	0.02	0.83
2021	0.02	0.01	0.01	0.04	0.00
Hospice effect					
2021	–0.05	0.04	–0.12	0.03	0.24

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Factors Explaining Changes in MAPD Bids: Part D Bids

Table I.15. Estimated Association Between VBID Participation and Part D Bids

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
BDI effect by year					
2020	1.22	1.01	–0.76	3.21	0.23
2021	4.77	0.53	3.73	5.81	0.00
Hospice effect					
2021	2.82	2.23	–1.56	7.20	0.21

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Table I.16. Estimated Association Between VBID Participation and Part D Bid Components

Outcome	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Part D standard coverage net PMPM					
BDI effect by year					
2020	–1.67	0.77	–3.18	–0.15	0.03
2021	4.75	0.55	3.68	5.82	0.00
Hospice effect					
2021	1.26	2.11	–2.88	5.41	0.55
Part D basic nonbenefit expenses					
BDI effect by year					
2020	–0.08	0.31	–0.68	0.53	0.80
2021	0.20	0.18	–0.15	0.56	0.26
Hospice effect					
2021	–2.77	1.06	–4.84	–0.70	0.01
Part D supplement nonbenefit expenses					
BDI effect by year					
2020	0.28	0.13	0.03	0.53	0.03
2021	–0.72	0.08	–0.87	–0.57	0.00
Hospice effect					
2021	–0.28	0.41	–1.08	0.53	0.50
Part D basic gain-loss					
BDI effect by year					
2020	–0.04	0.16	–0.37	0.28	0.79
2021	–0.29	0.11	–0.51	–0.08	0.01
Hospice effect					
2021	1.26	0.68	-0.08	2.59	0.06
Part D projected risk					
BDI effect by year					
2020	0.00	0.01	–0.02	0.01	0.83
2021	–0.01	0.01	–0.02	0.00	0.05
Hospice effect					
2021	–0.03	0.03	–0.08	0.02	0.21

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Difference-in-Differences Regression Model Results for Factors Explaining Changes Total (MAPD) Premiums

Factors Explaining Changes in Total (MAPD) Premiums: MA Premiums

Table I.17. Estimated Association Between VBID Participation and MA Premiums

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
BDI effect by year					
2020	0.07	0.85	–1.60	1.73	0.94
2021	–0.10	0.58	–1.23	1.03	0.87
Hospice effect					
2021	–4.61	3.43	–11.33	2.10	0.18

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Table I.18. Estimated Association Between VBID Participation and MA Premium Components

Outcome	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
MA rebate					
BDI effect by year					
2020	4.95	3.17	–1.26	11.15	0.12
2021	16.34	2.17	12.08	20.60	0.00
Hospice effect					
2021	11.54	8.38	–4.89	27.97	0.17
Rebate allocation: reduced Part A/B cost sharing					
BDI effect by year					
2020	1.10	1.11	–1.08	3.27	0.32
2021	1.50	0.84	–0.14	3.15	0.07
Hospice effect					
2021	–1.64	4.80	–11.05	7.76	0.73
Rebate allocation: MSB (additional services)					
BDI effect by year					
2020	3.60	2.22	–0.75	7.94	0.10
2021	14.56	1.95	10.74	18.39	0.00
Hospice effect					
2021	5.98	6.21	–6.19	18.15	0.34
Rebate allocation: Part B premium buydown					
BDI effect by year					
2020	0.46	0.44	–0.40	1.32	0.29
2021	–0.70	0.41	–1.51	0.11	0.09
Hospice effect					
2021	–1.23	1.68	–4.53	2.07	0.47

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Factors Explaining Changes in Total (MAPD) Premiums: Part D Premiums

Table I.19. Estimated Association Between VBID Participation and Part D Premiums

Effect	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
BDI effect by year					
2020	–0.47	0.78	–2.00	1.07	0.55
2021	1.53	0.46	0.62	2.43	0.00
Hospice effect					
2021	–2.65	3.32	–9.16	3.86	0.42

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Table I.20. Estimated Association Between VBID Participation and Part D Premium Components

Outcome	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound	p-value
Part D basic premium before buydown					
BDI effect by year					
2020	–1.04	0.69	–2.39	0.31	0.13
2021	3.52	0.43	2.67	4.37	0.00
Hospice effect					
2021	4.15	2.61	-0.96	9.26	0.11
Part D supplemental premium before buydown					
BDI effect by year					
2020	1.28	0.52	0.26	2.31	0.01
2021	–1.99	0.34	–2.66	–1.32	0.00
Hospice effect					
2021	–0.88	1.78	–4.36	2.61	0.62
Part D basic buydown					
BDI effect by year					
2020	–0.52	0.69	–1.87	0.83	0.45
2021	2.44	0.49	1.48	3.40	0.00
Hospice effect					
2021	9.78	3.44	3.03	16.53	0.00
Part D premium supplemental buydown					
BDI effect by year					
2020	0.28	0.62	–0.95	1.50	0.66
2021	–2.10	0.40	–2.88	–1.32	0.00
Hospice effect					
2021	–2.83	3.05	–8.80	3.14	0.35

SOURCE: RAND analysis of CMS and other data. The complete list of data sources and variables is in Table C.5 in Appendix C.

Abbreviations

ACP	advance care plan
ACS	American Community Survey
AHRF	Area Health Resources Files
AI/AN	American Indian and Alaska Native
API	Asian Pacific Islander
ASMD	absolute standardized mean difference
AWV	annual wellness visit
BCBS	Blue Cross Blue Shield
BDI	Benefit Design Innovations
BPT	Bid Pricing Tool
CAD	coronary artery disease
CDC	Centers for Disease Control and Prevention
CHF	congestive heart failure
CI	confidence interval
CM	care management
CMMI	Center for Medicare & Medicaid Innovation
CMS	Centers for Medicare & Medicaid Services
Comp	comparison
COPD	chronic obstructive pulmonary disease
C-SNP	Chronic Condition Special Needs Plan
DD	difference-in-differences
DM	disease management
D-SNP	Dual Eligible Special Needs Plan
ED	emergency department
ESRD	end-stage renal disease
FFS	Fee-for-Service
HCC	Hierarchical Condition Categories
HPMS	Health Plan Management System
HPSA	Health Professional Shortage Area
HRA	health risk assessment
ID	identification
IDR	Integrated Data Repository
I-SNP	Institutional Special Needs Plan
LIS	low-income subsidy
MA	Medicare Advantage

MAO	Medicare Advantage Organization
MAPD	Medicare Advantage plan with Part D coverage
MNT	medical nutrition therapy
MSB	mandatory supplemental benefits
NPPO	nonparticipating parent organization
OACT	CMS Office of the Actuary
OON	out-of-network
OOP	out-of-pocket
OTC	over the counter
PBP	plan benefit package
PCP	primary care provider
PDSS	Part D Senior Savings
PERS	personal emergency response system
PHRSB	Primarily Health-Related Supplemental Benefits
PMPM	per member, per month
PO	parent organization
PPO	preferred provider organization
RI	rewards and incentives
RUCC	Rural-Urban Continuum Codes
SD	standard deviation
SDM	standardized difference in means
SNP	Special Needs Plan
SSBCI	Special Supplemental Benefits for the Chronically Ill
TCC	transitional concurrent care
UF	Uniformity Flexibility
UWgt	unweighted
VBID	Value-Based Insurance Design
Wgt	weighted
WHP	Wellness and Health Care Planning

References

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager, "Synthetic Difference-in-Differences," *American Economic Review*, Vol. 111, No. 12, 2021, pp. 4088–4118.

Austin, P. C., "Balance Diagnostics for Comparing the Distribution of Baseline Covariates Between Treatment Groups in Propensity-Score Matched Samples," *Statistics in Medicine*, Vol. 28, No. 25, 2009, pp. 3083–3107.

Buetow, Stephen, "Thematic Analysis and Its Reconceptualization as 'Saliency Analysis,'" *Journal of Health Services Research and Policy*, Vol. 15, No. 2, 2010, pp. 123–125.

Callaway, Brantly, and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, Vol. 225, No. 2, 2021, pp. 200–230.

Centers for Disease Control and Prevention, "COVID Data Tracker," webpage, 2022. As of March 2, 2022: https://covid.cdc.gov/covid-data-tracker

Centers for Medicare & Medicaid Services, *Medicare Prescription Drug Benefit Manual: Chapter 13 - Premium and Cost-Sharing Subsidies for Low-Income Individuals*, 2018. As of March 22, 2022: https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/Downloads/Chapter-13-Premium-and-Cost-Sharing-Subsidies-for-Low-Income-Individuals-v09-14-2018.pdf

Centers for Medicare & Medicaid Services, *Value-Based Insurance Design Model: Request for Applications for CY 2020*, 2019. As of March 22, 2022: https://innovation.cms.gov/files/x/vbid-rfa2020.pdf

Centers for Medicare & Medicaid Services, *Value-Based Insurance Design Model Incorporation of the Medicare Hospice Benefit into Medicare Advantage: CY 2021 Request for Applications*, 2020a. As of March 22, 2022: https://innovation.cms.gov/files/x/vbid-hospice-rfa2021.pdf

Centers for Medicare & Medicaid Services, *Value-Based Insurance Design Model Request for Applications for CY 2021*, 2020b. As of March 4, 2022: https://innovation.cms.gov/files/x/vbid-rfa2021.pdf

Centers for Medicare & Medicaid Services, "Benefits Data," webpage, 2021a. As of March 2, 2022:
https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/Benefits-Data

Centers for Medicare & Medicaid Services, "Plan Crosswalks," webpage, 2021b. As of February 6, 2022:
https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/Plan-Crosswalks

Concannon, Thomas W., Dmitry Khodyakov, Virginia Kotzias, Gavin Fahey, Jennifer Graff, and Robert W. Dubois, *Employer, Insurer, and Industry Perspectives on Patient-Centered Comparative Effectiveness Research: Final Report*, Santa Monica, Calif.: RAND Corporation, RR-1242-PCORI, 2015. As of February 6, 2022:
https://www.rand.org/pubs/research_reports/RR1242.html

Daw, Jamie R., and Laura A. Hatfield, "Matching and Regression to the Mean in Difference-in-Differences Analysis," *Health Services Research*, Vol. 53, No. 6, December 2018, pp. 4138–4156.

Eibner, Christine, Dmitry Khodyakov, Erin Audrey Taylor, Christine Buttorff, Courtney Armstrong, Marika Booth, Kathryn E. Bouskill, Matthew Cefalu, Stephanie Dellva, Michael Dworsky, Federico Girosi, Ann C. Haas, Monique Martineau, Blen Eshete-Roesler, Alice Kim, Julie Lai, Afshin Rastegar, Daniel Schwam, Tisamarie Sherry, and Shiyuan Zhang, *Evaluation Report of the First Three Years (2017–2019) of the Medicare Advantage Value-Based Insurance Design Model Test*, Santa Monica, Calif.: RAND Corporation, EP-68316, 2020. As of March 2, 2022:
https://www.rand.org/pubs/external_publications/EP68316.html

Eibner, Christine, Dmitry Khodyakov, Erin Audrey Taylor, Christine Buttorff, Courtney Armstrong, Marika Booth, Kathryn E. Bouskill, Matthew Cefalu, Stephanie Dellva, Blen Eshete-Roesler, Alice Kim, Julie Lai, Afshin Rastegar, and Christopher Whaley, *First Annual Evaluation Report of the Medicare Advantage Value-Based Insurance Design Model Test: 2017, Year 1 of the Intervention*, Santa Monica, Calif.: RAND Corporation, EP-68317, 2018. As of March 2, 2022:
https://www.rand.org/pubs/external_publications/EP68317.html

Guest, Greg, Arwen Bunce, and Laura Johnson, "How Many Interviews Are Enough? An Experiment with Data Saturation and Variability," *Field Methods*, Vol. 18, No. 1, 2006, pp. 59–82.

Guest, Greg, Kathleen M. MacQueen, and Emily E. Namey, "Comparing Thematic Data," in Greg Guest, Kathleen M. MacQueen, and Emily E. Namey, eds., *Applied Thematic Analysis*, Thousand Oaks, Calif.: SAGE Publications, Inc., 2012, pp. 161–186.

Hainmueller, Jens, "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, Vol. 20, No. 1, 2012, pp. 25–46.

Hennink, Monique, and Bonnie N. Kaiser, "Sample Sizes for Saturation in Qualitative Research: A Systematic Review of Empirical Tests," *Social Science & Medicine*, Vol. 292, 2022, p. 114523.

Hennink, Monique M., Bonnie N. Kaiser, and Vincent C. Marconi, "Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough?" *Qualitative Health Research*, Vol. 27, No. 4, 2016, pp. 591–608.

Hernán, Miguel A., and James M. Robins, *Causal Inference*, Boca Raton, Fla.: Chapman & Hall/CRC, 2010.

Imai, Kosuke, and Marc Ratkovic, "Covariate Balancing Propensity Score," *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, Vol. 76, No. 1, 2014, pp. 243–263.

Imbens, Guido W., and Donald B. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences*, Cambridge: Cambridge University Press, 2015.

Khodyakov, Dmitry, Lori Uscher-Pines, Suchita A. Lorick, Megan C. Lindley, Victoria Shier, and Katherine Harris, "A Qualitative Analysis of the Impact of Healthcare Personnel Influenza Vaccination Requirements in California," *Vaccine*, Vol. 32, No. 25, May 23, 2014, pp. 3082–3087.

Kish, Leslie, *Survey Sampling*, New York: John Wiley & Sons, 1965.

Lindner, Stephan, and K. John McConnell, "Difference-in-Differences and Matching on Outcomes: A Tale of Two Unobservables," *Health Services and Outcomes Research Methodology*, Vol. 19, No. 2, 2019, pp. 127–144.

McHugh, Mary L., "Interrater Reliability: The Kappa Statistic," *Biochem Med (Zagreb)*, Vol. 22, No. 3, 2012, pp. 276–282.

Medicare Payment Advisory Commission, *Medicare Advantage Program Payment System*, 2021a. As of March 22, 2022: https://www.medpac.gov/wp-content/uploads/2021/11/medpac_payment_basics_21_ma_final_sec.pdf

Medicare Payment Advisory Commission, *Part D Payment System*, 2021b. As of March 22, 2022: https://www.medpac.gov/wp-content/uploads/2021/11/medpac_payment_basics_21_partd_final_sec.pdf

Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell, *Causal Inference in Statistics: A Primer*, Chichester, UK: John Wiley & Sons, 2016.

Shenton, Andrew K., "Strategies for Ensuring Trustworthiness in Qualitative Research Projects," *Education for Information*, Vol. 22, No. 2, 2004, pp. 63–75.

Stuart, E. A., B. K. Lee, and F. P. Leacy, "Prognostic Score-Based Balance Measures Can Be a Useful Diagnostic for Propensity Score Methods in Comparative Effectiveness Research," *Journal of Clinical Epidemiology*, Vol. 66, No. 8 Supp., 2013, pp. S84–S90.e81.

Wang, Yixin, and Jose R. Zubizarreta, "Minimal Dispersion Approximately Balancing Weights: Asymptotic Properties and Practical Considerations," *Biometrika*, Vol. 107, No. 1, 2020, pp. 93–105.

Zeldow, Bret, and Laura A. Hatfield, "Confounding and Regression Adjustment in Difference-in-Differences Studies," *arXiv preprint*, 2019.

Zeldow, Bret, and Laura A. Hatfield, "Confounding and Regression Adjustment in Difference-in-Differences Studies," *Health Services Research*, Vol. 56, No. 5, October 2021, pp. 932–941.

Zubizarreta, José R., "Stable Weights That Balance Covariates for Estimation with Incomplete Outcome Data," *Journal of the American Statistical Association*, Vol. 110, No. 511, 2015, pp. 910–922.