ACTUARIAL RESEARCH CORPORATION

6928 LITTLE RIVER TURNPIKE
SUITE E
ANNANDALE, VA 22003
703 941-7400
703 941-3951 FAX

5950 SYMPHONY WOODS ROAD
SUITE 510
COLUMBIA, MARYLAND 21044
410 740-9194
410 730-7945 FAX

To: Third Way

From: Actuarial Research Corporation

Subject: Final Scoring Memo: Bundles

Date: March 11, 2015

Policy Background

Several proposals have surfaced designed to reduce Medicare spending on hospital post-acute care. Many of these involve some form of "bundling" – grouping services into packages for which a fixed payment is made; the theory is that doing so creates incentives for providers of care to optimize the provision of that care.
In this study, we simulated the effects of one such proposal. Post-acute care (PAC) for the 180 days following a hospital discharge is bundled; payment is based on the lesser of current-law benefits or the 75th percentile of the national distribution of such benefits (adjusted for local differences in expenditure). The proposal, in which bundles are phased into the payment scheme and which includes diminishing indemnification of providers from losses, would reduce overall Medicare spending by 3.1 percent over 10 years, and would reduce Medicare Fee-for-Service (FFS) A/B benefits by 5.4 percent.

Estimation Process and Results

In this payment model, post-acute care covered by the program for Medicare FFS beneficiaries aged 65 years or older is collected into a "bundle," payment for which is capped at a set amount. The model is a modification of work done by Cutler and Ghosh (2002).[1] Following Cutler and Ghosh, Medicare services provided to a beneficiary within 180 days of the beneficiary's admission to a hospital are rolled into a single bundle for payment purposes. The bundle includes all inpatient, outpatient, Skilled Nursing Facility, home health, hospice, physician and laboratory, and durable medical equipment (DME) goods and services, regardless of where these services are provided or by whom.

Payment for a particular bundle is based upon the national distribution of current-law FFS benefits. Bundles are grouped by the Clinical Classifications Software for ICD-9-CM (CCS), which is "a tool for clustering patient diagnoses and procedures into a manageable number of clinically meaningful categories."[2] A national distribution of these bundles' current-law costs is constructed, and payment for a given bundle is capped at an adjusted value of the 75th percentile of the distribution.

The payment cap is adjusted to account for local differences in Medicare spending. Again following Cutler and Ghosh, the adjustment multiplier is the ratio of age/sex/race-adjusted Medicare FFS spending for Parts A and B at the Dartmouth Hospital Referral Region (HRR) and that for the nation as a whole.[3]

---

[1] Cutler, DM, and Ghosh, K. 2002. The potential for cost savings though bundled episode payments. N Engl J Med 366:1075-1077. March 22, 2012
[2] http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp
[3] http://www.dartmouthatlas.org/data/table.aspx?ind=225&tf=23&ch=191&loc=&loct=3&fmt=264

Bundles are phased in over 3 years. In the first year, the bundle accounting for the largest current-law spending – septicemia (except in labor) – is introduced. The next year, the next 13 bundles ranked by current-law spending are introduced. All additional bundles are added in the following year.

In order to ease the transition to the new payment methodology, a diminishing indemnification of providers is part of the proposal. In the first year that a bundle is introduced into the system, providers are paid half of the costs in excess of the adjusted cap. In the following year, that amount is reduced to a third of the excess; the following year the amount is a quarter of the excess, and in the fourth year the cap is strictly observed.
The cap for a given bundle grows over time. In this model, the cap grows at the rate of current-law Medicare spending.

Data and Methods
The data for this simulation come from the Medicare 5-percent Limited Data Set Standard Analytic Files (LDS-SAF). Although the beneficiary information in these files is encrypted or suppressed to protect the privacy of individuals enrolled in the program, this is done is a way that allows us to link across services and over time.
In addition to these claim-level data, we made use of several public information sources. We benchmarked our results to the 2014 Medicare Trustees Report (TR).[4] As noted above, we used the AHRQ CCS tool to group bundles by the diagnostic information contained in the hospital claim, and we used the Dartmouth Atlas information to group bundles into HRRs for the purpose of adjusting the spending cap.

Following the Cutler and Ghosh (C&G) approach as laid out in their article, we assigned a Clinical Classification System (CCS) number to each hospitalization that occurred in 2010 and 2011, using the primary diagnosis to make the classification. Hospitalizations identified as IRF, SNF, long-term-care, or swing-bed stays were excluded at this step. All claims – inpatient, SNF, home health, hospice, outpatient, carrier (physician and independent laboratory), and DME – were assigned a high-level (what C&G call "organ system") CCS code. For each bundle, initiated by a hospital event, we included all subsequent 2010, 2011, and 2012 claims with the same high-level CCS code and a date of service that fell within 180 days of the initiating date of admission; hospitalizations that fell within a previous bundle were not used to initiate a new bundle. Note that because we began the analysis with 2010 data, some 2010 bundles might, in fact, have belonged to a previous (2009) bundle; we were not concerned by this, as we are restricting our attention to 2011 and 2012.

Table 1 contains a tabulation of the results of this analysis. The table shows 2011 FFS benefits attributed to a CCS episode that began in 2010 or 2011; only the 2011 spending is included. The results track C&G findings roughly. The main difference is the rank ordering by CCS, although the same CCS codes appear in the top rankings of both findings. Our inflated 5-percent sample figure, $348.9 billion, is within half a percent of the TR figure of $351.0 billion.

Again following the C&G model, we estimated the savings that would occur should bundle reimbursement be capped. We computed the distribution of national spending for each bundle, without regard to local variation in benefit levels. Then we compared the actual bundle benefit amount with the national median adjusted for benefit levels in the relevant Dartmouth Hospital Referral Region (HRR), allowing the smaller of the actual cost or the adjusted cap.

To simulate the effect of phased introduction, we rank-ordered the bundles by dollar volume and selected – only partly arbitrarily – the year in which the bundle would be introduced. Table 2 shows this phase-in.

[4] http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/ReportsTrustFunds/index.html

ACTUARIAL RESEARCH CORPORATION
6928 LITTLE RIVER TURNPIKE
SUITE E
ANNANDALE, VA 22003
703 941-7400
703 941-3951 FAX
5950 SYMPHONY WOODS ROAD
SUITE 510
COLUMBIA, MARYLAND 21044
410 740-9194
410 730-7945 FAX

Using the aggregate 2011 dollar figures, we calculated savings as a share of current-law spending. In the first year that a bundle is introduced, we recouped half of the full savings attributable to the bundle; we recouped two thirds of the full savings for bundles introduced two years earlier, three quarters of the full savings for bundles introduced 3 years earlier, and all the savings for bundles introduced 4 years earlier. In this way, 6 years after implementation of the policy the program would be achieving the ultimate percentage savings relative to current-law spending.

Results

To translate these relative savings to absolute savings, we used TR and NHE[5] projections of Medicare spending to construct a timeseries of fiscal year FFS benefits and total Medicare costs. We applied the savings percentage calculated using the static 2011 numbers to the appropriate year's FFS benefits to obtain a dollar savings and then accumulated the savings and spending over a 10-year period. This is shown in Table 3.

Medicare savings start out relatively small, partly because of the number of bundle types involved and partly because of the indemnification of providers. By the end of the 10-year period, annual savings are on the order of 8.7 percent of FFS benefits and 4.9 percent of total Medicare spending. Over the entire 10-year interval, savings on the order of 3.1 percent of total spending are achieved.

Non-Medicare estimates are shown in Table 4 and reflect the illustrative assumption that other payers in aggregate will adopt similar policies. Medicaid savings are assumed to be 0.4 times the Medicare savings as a percent of total spending, given lower payment rates. By the end of the 10-year period, Medicaid savings total $70 billion. PHI savings are assumed to be 0.6 times the Medicare value and total $190.1 billion over the 10-year period. There is no impact on out-of-pocket spending.

Savings in a model such as this are sensitive to several factors. One such factor is the location of the spending cap within the distribution of bundle costs. Many bundle groups show a pronounced distributional skew toward the upper tail, so that moving the cap to higher percentiles of the distribution produces a disproportional reduction in savings (and vice versa). Sequencing the introduction of bundles has a rather insignificant effect, as a small number of bundle groups accounts for most of program spending. Obviously, the extent of indemnification – and its phase-out – can alter the savings stream. So, too, is the manner by which caps are changed over time. In this simulation, we assumed that the cap would remain fixed relative to current-law spending (and, for that matter, that the distribution of bundle costs would remain unchanged over time); increasing the cap by less than average Medicare spending – for example, by GDP prices plus 1 percent – would effectively reduce the cap within the distribution of spending and increase the savings attributable to the policy.

Limitations

This study is subject to several limitations. First, because it is static, it does not take into account differential changes in bundle costs. For example, if physician costs rise more rapidly than hospital costs, bundles with heavier physician components will skew the distribution within a group and across all groups. Second (and as noted above) the simulation incorporates caps that move smoothly with the distribution of bundle costs over time; in the real world, such a counterfactual would be very difficult to maintain.

[5] http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsProjected.html

The policy itself retains some undeveloped aspects. For example, no mechanism is presented through which capped payments are distributed among the various providers of care, nor by which individual providers are incented to reduce the services they provide.