HHS Public Access
Author manuscript
Am Econ J Appl Econ. Author manuscript; available in PMC 2020 April 01.

Published in final edited form as:
Am Econ J Appl Econ. 2019 April ; 11(2): 302–332. doi:10.1257/app.20170295.

Health Care Spending and Utilization in Public and Private Medicare

Vilsa Curto, **Liran Einav**, **Amy Finkelstein**, **Jonathan Levin**, and **Jay Bhattacharya**

Curto: Department of Health Policy and Management, T.H. Chan School of Public Health, Harvard University, 677 Huntington Avenue, Boston, MA 02115 (vcurto@hsph.harvard.edu); Einav: Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305-6072 (leinav@stanford.edu) and NBER; Finkelstein: Department of Economics, MIT, 50 Memorial Drive, Cambrdige, MA 02142-1347 (afink@mit.edu) and NBER; Levin: Graduate School of Business, Stanford University, 655 Knight Way, Stanford, CA 94305-7298 (jdlevin@stanford.edu) and NBER; Bhattacharya: School of Medicine, Stanford University, 616 Serra Street, Stanford, CA 94305-6019 (jay@stanford.edu). We are grateful to Diego Jimenez, Andelyn Russell, Daniel Salmon, and Martina Uccioli for excellent research assistance. We thank the Editor, three anonymous referees, and numerous seminar participants for helpful comments. We gratefully acknowledge support from the NSF (SES-1527942, Bhattacharya, Einav, and Levin), the NIA (R01 AG032449, Einav and Finkelstein; R37 AG036791, Bhattacharya), and the Sloan Foundation (Bhattacharya, Einav, Finkelstein, and Levin). The authors acknowledge the assistance of the Health Care Cost Institute (HCCI) and its data contributors, Aetna, Humana, and UnitedHealthcare, in providing the claims data analyzed in this study.

Abstract

We compare healthcare spending in public and private Medicare using newly available claims data from Medicare Advantage (MA) insurers. MA insurer revenues are 30 percent higher than their healthcare spending. Adjusting for enrollee mix, healthcare spending per enrollee in MA is 9 to 30 percent lower than in traditional Medicare (TM), depending on the way we define "comparable" enrollees. Spending differences primarily reflect differences in healthcare utilization, with similar reductions for "high value" and "low value" care, rather than healthcare prices. We present evidence consistent with MA plans encouraging substitution to less expensive care and engaging in utilization management. (*JEL* H11, H42, H51, I11, I13)

A long-standing question in economics concerns the appropriate roles of the public sector and private sector in providing services that society has decided are essential. This question comes up in many contexts, including education, utilities, transportation, and pensions. It is especially relevant in healthcare, where the United States is unusual among developed countries in its distinctive mix of public and private health insurance. Comparisons of public and private health insurance systems are difficult, however, since they typically do not operate at a similar scale, for the same population, in the same markets, or with the same healthcare providers.

The U.S. Medicare program in recent years has been an exception because of the “side by side” operation of public and private insurance programs. While traditional Medicare (TM) offers publicly administered insurance, a significant fraction of the over-65 Medicare population has opted out of TM in the last decade and enrolled in private insurance plans through Medicare Advantage (MA). In MA, private insurers receive capitated payments from the government for providing Medicare beneficiaries with health insurance that roughly mimics commercial health insurance for the under-65 population. Today almost a third of Medicare beneficiaries are enrolled in MA.

Empirical comparisons of MA and TM face two primary challenges. First, differences in health-care utilization between patients in MA and TM may partly or entirely reflect differences in the patient mix, rather than a “treatment effect” of MA per se. Second, historically data availability has been asymmetric: administrative claim-level data from TM are widely available to researchers, but detailed claim-level data from MA insurers has been more elusive. The primary contribution of this paper lies in our analysis of new, claim-level data from MA insurers. Specifically, we take advantage of newly available claims data from MA plans in 2010 provided by the Health Care Cost Institute (HCCI). The data consist of claims paid by three MA insurers (Aetna, Humana, and UnitedHealthcare) that cover almost 40 percent of MA enrollees. The key advantage of these data is that they contain claim-level data in MA – i.e. healthcare utilization and payments to providers – that is analogous to the existing and commonly used claims data for TM.

A simple tabulation of the MA and TM claims points to a large difference in public and private healthcare spending levels. We calculate that MA spending per enrollee-month in 2010 totaled $642, of which $590 was paid by MA insurers and the rest by enrollees out of pocket. In contrast, average spending per enrollee-month in TM was $911, of which $771 was paid directly by the Medicare program to providers. Capitated payments to the MA plans roughly track the latter amount; the MA plans in the HCCI data received on average $767 per enrollee-month. In other words, the revenue of the MA plans we observe is 30 percent higher than the payments they make for their enrollees’ healthcare. If this applied to the entire MA population in 2010 (including those outside our sample), it would imply $21 billion in annual (2010) revenue for MA insurers in excess of their spending on healthcare claims.

The bulk of our analysis compares healthcare spending and utilization for enrollees in MA and TM. To proxy for what an MA enrollee’s healthcare experience would have been like if she were (counterfactually) in MA, we construct a “comparable” group of TM enrollees. We present results from two main approaches. First, we adjust for key observables – comparing outcomes for MA and TM enrollees in the same county and with the same risk score. Medicare risk scores are based on a predictive model of healthcare spending that accounts for demographics and detailed information on prior health conditions. The county and risk score adjustment also captures the spirit in which Medicare sets reimbursement rates for MA insurers; these are the two dimensions that enter the formula by which capitation rates are computed. Second, we include an additional adjustment for unobserved health not captured by the risk score (Brown et al. 2014), which is based on mortality differences between MA and TM enrollees in the same county and with the same risk score. Without either adjustment, MA spending per enrollee-month is 30 percent lower than TM spending per enrollee-month. Holding county and risk score fixed, the spending difference becomes 25 percent, and adjusting for mortality differences further reduces it to 9 percent. None of these approaches is a panacea for concerns about selection; however, taken together they suggest a non-trivial “treatment effect” of MA on spending, albeit with some uncertainty as to the magnitude.

A key advantage of our detailed claim-level data is that they allow us to explore differences in patterns of spending and healthcare use for specific populations and for different types of care. These comparisons are qualitatively similar across our alternative adjustments. They indicate that spending differences are much greater in urban counties (where about three-quarters of MA beneficiaries enroll) than in rural counties and that lower spending in MA is present across the distribution of spending and for different types of care. Differences are smaller for inpatient care, and are particularly pronounced for care in skilled nursing facilities (SNFs).

Lower healthcare spending in MA than in TM primarily reflects lower utilization of services rather than lower payments for the same services. MA insurers’ average payment to hospitals (per admission and per day) is within one to two percent of the analogous payment in TM. Comparing payments made to the same hospital for the same diagnosis (DRG), we find that MA payments are about one percent higher than TM payments. Lower utilization in MA appears both for services where there are concerns about over-use, such as diagnostic testing and imaging, as well as for services where there are concerns about under-use, such as preventive care.

We present suggestive evidence for some potential mechanisms by which MA insurers may reduce utilization relative to TM. We find several patterns consistent with restrictions on use of the most expensive types of care and possible substitution to less expensive alternatives. For example, we find higher spending per emergency department visit in MA than in TM, which is consistent with utilization constraints in MA, so that the marginal patient admitted for care is in worse health. We also find that MA patients, relative to TM patients, are much less likely to be discharged from the hospital to post-acute care and much more likely to be discharged home. In addition, lower rates of physician visits in MA primarily reflect lower visits to specialists, with little or no difference in rates of primary care visits. Finally, inpatient surgery rates are similar in MA and TM while outpatient surgery rates are much higher in MA, which is suggestive of MA insurers substituting from inpatient to outpatient surgery. Such evidence on potential mechanisms reinforces our interpretation that differences in average spending in MA and TM by “similar” enrollees likely reflect, at least partially, an MA treatment effect. One would need a more subtle selection story, which moves beyond selection into MA on predicted spending, to explain these patterns.

Finally, we briefly examine geographic variation in MA and TM. Geographic variation in TM spending has received a great deal of attention, often interpreted as a sign of regional differences in the efficiency of healthcare delivery within TM (e.g. Gawande 2009; Skinner 2011). However, we find roughly similar levels of heterogeneity across regions in MA and

TM. Geographic variation in healthcare spending is around 20 percent higher in MA, while geographic variation in hospital prices is about 20 percent lower in MA than in TM.

After we relate our findings to the existing literature in the next sub-section, the rest of the paper proceeds as follows. Section I provides some institutional background on our setting. Section II describes our data, baseline sample, and summary statistics. Section III describes our approaches for constructing a "comparable" set of TM enrollees to compare spending in TM and MA. Section IV compares healthcare spending in MA and TM, overall and for various categories of people and spending. Section V examines differences between MA and TM enrollees in healthcare utilization and in healthcare prices, and examines some potential mechanisms for utilization reductions. The last section concludes.

Our findings relate to several literatures. The most directly related are prior comparisons of healthcare spending in MA and TM. As noted earlier, our key advance is access to detailed claims data for a large share of the MA market. Absent such data, prior studies have used a variety of approaches to infer healthcare utilization and spending differences between MA and TM. These include comparing MA plans' (mandatory) self reports of enrollee utilization to utilization measures in TM claims data (Landon et al. 2012), analyzing beneficiaries' self reports of care received in TM and in MA (Ayanian et al. 2013), analyzing hospital discharge data from New York counties experiencing MA exit (Duggan, Gruber, and Vabson 2018), and inferring cost differences from estimates of demand for MA plans and a supply-side model of the market (Curto et al. 2014). These papers have tended to find lower healthcare utilization in MA – with estimates ranging from 10 percent to 60 percent.

Our finding of similar pricing in MA and TM echoes a recent finding by Baker et al. (2016) and contrasts with the conventional wisdom that MA prices will be higher than TM prices due to the greater bargaining power enjoyed by the larger public sector (e.g. Philipson et al. 2010). It also differs from prior findings that TM prices are substantially lower than prices in the private, under-65 market both on the inpatient side (Cooper et al. 2015) and the outpatient side (Clemens and Gottlieb 2017). It seems plausible that the lower prices that private insurers pay for over-65 enrollees relative to under-65 enrollees is the consequence of regulation that is specific to the over-65 population, and requires hospitals to accept TM rates if an alternative payment rate was not negotiated (Berenson et al. 2015).

Our findings of similar geographic variation in spending and pricing in MA and TM also contrast with recent findings that geographic variation in spending in commercial (i.e. under-65) insurance is similar to TM, but stems from much larger pricing variation and lower quantity variation in commercial insurance relative to TM (Philipson et al. 2010; Institute of Medicine 2013; Cooper et al. 2015). This contrast between TM and commercial insurance has been interpreted as reflecting the lower powered incentives in the public sector relative to the private sector in constraining utilization, and monopsony power in the public sector to constrain prices relative to what the private sector can achieve (Philipson et al. 2010). Of course, there are other reasons why patterns of healthcare provision for those under 65 may differ from the patterns for the over 65. We consider this same set of facts in the context of Medicare Advantage, which arguably provides a cleaner comparison group to

TM for understanding variation under private and public regimes since MA and TM are provided to the same broad population.

Our finding that MA appears to reduce both “high value” and “low value” care in similar magnitude contributes to what we believe is an emerging, cautionary tale on the bluntness of policy instruments in the healthcare sector. Our evidence here speaks to the blunt nature of supply-side restrictions on care. Likewise, on the demand side, recent evidence suggests that high deductible plans reduce “high value” and “low value” care in equal measure (Brot-Goldberg et al. 2017), and that even targeted increases in the price of some types of care can depress care use across the board, including free preventive care services (Cabral and Cullen 2011).

Most broadly, our work is part of the large literature on the relative consequences of public and private ownership. This literature has spanned a range of disparate industries, including education, pensions, electricity, and transportation. In the specific context of healthcare, recent empirical work has emphasized that the private sector may be more efficient than the public sector at setting reimbursement prices for providers (Clemens, Gottlieb, and Molnar 2017) and at setting cost-sharing to combat moral hazard (Einav, Finkelstein, and Polyakova 2018).

I. Setting and background

The Medicare Advantage (MA) program allows Medicare beneficiaries to opt out of traditional fee-for-service Medicare coverage and enroll in private insurance plans. The program was established in the early 1980s with two goals: to expand the choices available to beneficiaries and to capture cost savings from managed care. In return for covering enrolled beneficiaries‘ healthcare expenses, private MA plans receive a risk-adjusted, capitated monthly payment from the Centers for Medicare and Medicaid Services (CMS), which is the federal agency that manages the Medicare program.

There has historically been a tension between the two goals of expanding access to MA and limiting costs (McGuire, Newhouse, and Sinaiko 2011). Insurers have tended to participate more in periods with higher payments, and to offer more plans in areas with higher payments. MA plans also enroll relatively healthier beneficiaries, complicating the problem of setting appropriate capitation rates. Reforms over the last decade have aimed to address these problems by introducing a risk scoring system to adjust plan payments based on enrollee health, and a competitive bidding system that replaced the fixed reimbursement rates used earlier. These changes, combined with an increase in capitation rates set by CMS, have coincided with the expansion of plan offerings and enrollment seen in Figure 1. Enrollment in MA tends to be especially high in urban areas; in 2010, MA penetration was 33 percent in urban counties and 18 percent in rural counties.

To participate in MA, insurers must contract with a set of healthcare providers and offer at least the same insurance benefits as traditional Medicare (TM), which covers inpatient (“Part A”) and outpatient (“Part B”) healthcare services. MA plans typically provide additional benefits as well, in the form of more generous cost sharing or supplemental coverage of dental, vision, or drug benefits. Medicare beneficiaries observe the MA plan offerings in their county of residence and can choose to enroll in any of the available MA plans during an annual “open enrollment” period every fall. The tradeoff they face in choosing between MA and TM is that MA plans typically restrict access to healthcare providers, but provide additional benefits as described above. In our data (before applying the sample restrictions described below), 73 percent of MA enrollees were in HMO or PPO plans with limited provider networks.

Every year, plans enter into a bidding process, which dictates the benefits and premium associated with each plan that is offered to beneficiaries While the precise rules by which plan bids translate to plan premiums and benefits are somewhat complicated, we summarize the key features here (see Curto et al. (2014) for a more detailed description). Each plan submits a bid b, which should be interpreted as the monthly compensation required by the plan to provide “standard” monthly coverage in the local area in which the plan is offered to an “average” Medicare beneficiary. By “standard” coverage we refer to the standard Part A and Part B financial coverage offered by TM; MA plans typically offer more comprehensive coverage, but they obtain a separate compensation for it (known as the “rebate”) on top of their bid b. As will be clearer later, by “average” beneficiary we refer to a beneficiary with an average health risk.

This bid b is then assessed against its local benchmark B, which is set administratively by CMS. In principle the benchmark B is supposed to approximate the counterfactual cost to CMS from covering an “average” beneficiary in that county through TM. In practice, the variation in benchmarks across locations departs somewhat from this principle, presumably reflecting various political economy considerations. On average in our observation period (2010), benchmark rates are higher than corresponding TM costs, and more so in some areas than in others; subsequent to our time period of analysis, the Affordable Care Act has reduced the level of these MA benchmark rates. Overall in our data (again, before applying the sample restrictions described below), the average benchmark across counties (weighted by the number of Medicare beneficiaries) is \$836 per enrollee-month, compared to an average TM cost of \$798, and this difference is lower in urban counties (benchmark of \$866 and average TM costs of \$842) than in rural counties (\$770 vs. \$716). However, in our observation period, the vast majority of plan bids are lower than the corresponding benchmarks, making MA plans financially more generous than TM, where enrollees can face large out-of-pocket costs.[1]

Capitation payment to insurers for enrolling a given enrollee in a given MA plan depends not only on the plan’s bid b but also on the enrollee’s risk score r_i, which is proportional to her predicted healthcare costs in TM over the next year. Adjusting reimbursement for risk score is a key component of CMS’ attempt to limit selection into MA by adjusting plan compensation for predictable heterogeneity in healthcare cost across beneficiaries. CMS assigns a risk score to each Medicare beneficiary based on demographic information and detailed claim-based information on chronic health conditions measured over the previous

---

[1] If $b > B$ the difference is charged as a premium to the consumer. If $b < B$, which is almost always the case empirically, 75 percent of the difference is given to the consumer through the rebate, and 25 percent is retained by CMS.

12 months. The average beneficiary's risk score is normalized to 1, so that plans obtain compensation of r_ib for covering beneficiary i. For purposes of setting MA plan payments, CMS deflates estimated risk scores for MA enrollees (by3.41 percent in 2010, which is our sample year) to reflect CMS' estimate of the "upcoding" of risk scores for MA beneficiaries (CMS 2010; Geruso and Layton 2015).

Thus, broadly speaking, plan compensation is designed to reimburse an MA insurer for the costs an enrollee would incur – based on her county and risk score – had she remained in TM. This motivates our baseline approach (described below) of comparing enrollees who are in the same county with the same risk score when comparing utilization and healthcare spending in MA and TM.

II. Data and sample construction

A. Data sources

This paper uses data from two main sources: the Health Care Cost Institute (HCCI) and the Center for Medicare and Medicaid Services (CMS). All the data pertain to spending and enrollment in 2010. Appendix A provides more details on the data and sample definition; Appendix B provides more details on the definition and construction of the specific healthcare spending and utilization variables we analyze.

The HCCI data are the key, novel data in this paper. HCCI is provided with claim-level data from three large MA insurers – Aetna, Humana, and UnitedHealthcare. HCCI pools these data (masking the individual insurers) and makes these data available for research. In 2010, these three insurers (hereafter referred to as the "HCCI insurers") covered almost 40 percent of MA enrollees: UnitedHealthcare was the largest (national market share of 18 percent), Humana was second (15 percent), and Aetna fifth (4 percent) (Kaiser Family Foundation 2010). The claim-level data reflect claims that these three insurers paid out to healthcare providers. The HCCI data also contain monthly enrollment indicators and some limited enrollee demographics (age bins, gender, and zip code).

The CMS data serve multiple roles. One role is to provide parallel claim-level data for Medicare beneficiaries enrolled in Traditional Medicare (TM). Because TM offer s fee-for service coverage, we essentially observe every healthcare claim made by TM enrollees during 2010. The TM claims data allow us to form a "benchmark" comparison of healthcare spending and utilization against which we can compare the measures obtained from HCCI.

The CMS data have a second, equally important role: providing enrollment, demographic, health and mortality data for all enrollees (TM and MA). For the universe of Medicare enrollees we can observe monthly enrollment information in TM (Parts A and/or B) or MA, risk score, demographics (zip code, age, and gender), dual eligibility status (in Medicaid and Medicare), detailed health conditions from the prior year, and mortality. The detailed CMS data on MA enrollees allow us to validate the completeness of our baseline sample in HCCI, and to adjust our comparison to TM spending for the differential demographics, health conditions, and mortality among MA enrollees compared to TM enrollees.

Finally, the CMS data contain detailed information on payments to MA insurers by CMS. This allows us to construct payments to MA plans per enrollee-month, as well as payment components.

B. Baseline sample

The HCCI data include most, but not all, MA enrollees covered by the three HCCI insurers. Based on the qualitative information that HCCI obtained from the three participating insurers, it appears that inclusion in the HCCI data was made on a plan-by-plan basis, with “highly capitated plans” left out. That is, insurance plans that pay providers on a capitated basis are omitted from the HCCI data. The HCCI data also indicate that they exclude special needs plans (SNPs), which are MA plans for individuals with specific diseases (such as end-stage liver disease, chronic heart failure, or HIV-AIDS) or certain characteristics (such as residence in a nursing home).

Ideally, we would have plan identifiers in the HCCI data, which would allow us to match this information to the plan identifiers in the CMS data, and thus know which MA plans are excluded. This would allow us to adjust for the demographics and health conditions of MA enrollees specifically enrolled in HCCI plans. However, with the exception of SNPs that are not in the HCCI data and can be identified in the CMS enrollment data, plan and insurer identifiers are omitted from the HCCI data. Instead, we rely on the fact that the MA market is localized and the use of provider capitation is most common in particular regions such as California, and construct our baseline sample by focusing on states where the HCCI data coverage appears to be approximately complete.

We judge the completeness of the HCCI data by comparing enrollment statistics for the HCCI insurers in the HCCI and CMS data. In the CMS data, we know for each MA enrollee whether he or she was enrolled in an MA plan offered by one of the HCCI insurers. This allows us to generate a pseudo HCCI enrollment data set in the CMS data, which covers all enrollees who “should” have been in the HCCI data if no plans were omitted. We then compare enrollee-month counts in this pseudo HCCI enrollment data and cross validate the actual HCCI data against it. Specifically, we compare enrollee-month counts at the state level across the two data sets, restricting the analysis to individuals who are 65 and over; we do not require individuals to be enrolled for a full year.

We define our baseline sample to be the set of 36 states where we have a close to complete sample of HCCI insurers’ enrollees, which we define to mean that the count of enrollee-months in HCCI in the state is within 10 percent of the count for the HCCI insurers in the pseudo HCCI enrollment data. In practice, in these 36 complete data states, total HCCI enrollment is within one percent of total enrollment in the pseudo HCCI enrollment data, leaving us reasonably sanguine that we have captured the entire set of MA enrollees for these three insurers. Appendix Table A1 provides more details on state-by-state enrollee-month counts in the HCCI insurers as measured in the HCCI and CMS data.

The 36 states in our baseline sample represent about 60 percent of enrollees for the HCCI insurers. As shown in Appendix Figure A1, the excluded states are disproportionately concentrated in the Western United States. Appendix Table A1 shows the MA share of total

Medicare enrollees and the HCCI insurer share of MA enrollees by state, including both the 36 complete data states and the 15 omitted states.

Table 1 shows how our baseline sample is constructed, and Panel A presents basic demographic statistics from both the CMS and HCCI data. Throughout the paper, risk scores for TM enrollees are unadjusted, while risk scores for MA enrollees are adjusted to reflect the 3.41 percent deflation CMS applies in determining MA payments, as described above and in CMS (2010, page 19).

Columns (1) through (3) present CMS data across all plans and states, while columns (4) through (6) present CMS data for our baseline sample, which is comprised of the 36 states above and omits enrollees in SNPs. In each case, we present statistics for all TM enrollees, for all MA enrollees, and then for enrollees covered by the three HCCI insurers. Columns (7) and (8) present statistics for the HCCI data, for the entire sample in column (7) and for our baseline sample in column (8).

We use Table 1 to make several observations. First, comparing columns (1)-(3) to columns(4)-(6), the 36 states that constitute the baseline sample do not seem to be very different from the overall sample, making us feel reasonably comfortable that the findings we report throughout the paper are likely to be relevant for states not covered by our baseline sample. Second, comparing column (2) to (3) or column (5) to (6), it appears that the three HCCI insurers attract enrollees that seem reasonably similar to the overall MA enrollees, suggesting that our subsequent findings may apply to the broader MA population. Third, as has been documented elsewhere, MA enrollees are slightly younger and significantly healthier than TM enrollees: their risk scores (which are proportional to their predicted healthcare spending) are about 5–10 percent lower, and their annual mortality rates are almost a third lower. This suggests that a straight comparison of TM and MA healthcare spending would be misleading, motivating the various corrections for selection we describe in the next section.

Finally, it is reassuring that, for our baseline sample, the enrollment counts and demographics (that we can measure in both data sets) are remarkably similar when measured in the pseudo HCCI enrollment data set we construct in the CMS data (column (6)) and the actual HCCI data (column (8)). This is what we would expect given our construction of a baseline sample for which the HCCI data should include all relevant MA enrollees.[2]

C. Summary statistics

Panels B and C of Table 1 report summary statistics on total healthcare spending and CMS payments per enrollee.

Spending and payments in MA.—Our first result is the size of CMS payments to MA insurers in excess of MA insurers' healthcare spending on enrollees. We define total healthcare spending as the sum of insurer healthcare spending and any out-of-pocket spending by the beneficiary. Insurer spending is based on observed payment amounts – that is, transacted prices, not list prices. Out-of-pocket spending is the amount owed by the enrollee (due to deductibles and co-insurance).[3]

---

[2] We have about 1 percent more enrollees in our HCCI sample (column (8)) than the pseudo-HCCI sample in the CMS data (column (6)). This is to be expected, given that plan assignment is missing for about 1 percent of MA enrollees in the CMS data.

Our measure of healthcare spending is a near-exhaustive measure of all healthcare claims. Specifically it covers several categories of spending: (a) inpatient spending, which is associated with providers identified as hospitals and physicians billing for treatment provided in an inpatient hospital setting; (b) outpatient spending, which also includes home health care and durable medical equipment (e.g. wheelchair rentals); and (c) skilled nursing facility (SNF) spending.[4] Average total healthcare spending per enrollee-month in MA is $642 in our baseline sample (column (8)). Of this, $590 is paid by the insurer, and $52 is owed by the enrollee.

Payments to MA insurers for "organic" MA services (i.e. for services that would be covered by TM) are $767 per enrollee-month in our baseline sample (column (6)).[5] The comparison of insurer MA revenue of $767 per enrollee-month to the insurer payments to healthcare providers of $590 suggests that net revenues for MA insurers are $177 per enrollee-month, or about 30 percent above MA insurer healthcare spending. If this applied to the entire MA population in 2010 (including those outside our sample) it would imply $21 billion in annual (2010) revenue for MA insurers in excess of their spending on healthcare claims.

Of course, MA insurers incur additional costs, such as administrative and advertising expenses, which we do not observe in our data. A rough estimate is that these additional costs are approximately 8 percent of expenditure on MA healthcare claims.[6] By comparison, the government estimates that administrative costs for Medicare (including the federal government's costs of administering Medicare Advantage) were about 1.7 percent of Medicare TM claims in 2010 (Boards of Trustees, 2011).

Spending in MA and TM: raw comparisons.—The raw summary statistics also show dramatic differences in total healthcare spending between the TM and MA populations. In our baseline sample, the average TM enrollee spends $911 per month (column (4)), while the average MA enrollee spends 30 percent less, $642 (column (8)).

---

[3] TM enrollees can purchase supplemental private insurance ("Medigap" or employer sponsored) to cover some or all of their out-of-pocket expenses. About half do so. If they do, the supplemental insurer is the primary payer of the "out-of-pocket" amount owed by the beneficiary.

[4] One (small) category of spending that is not in our measure of total spending is hospice care. This is because hospice care is billed directly to CMS even for MA enrollees, so it is observed in CMS data for both TM and MA and doesn't fully conform to the empirical exercise. In practice, we show below that the exclusion of hospice spending does not substantively affect the comparison of total spending.

[5] We define payments to MA insurers to be the sum of CMS spending on MA ($778) and additional consumer premiums for MA ($6) minus the portion of the consumer rebate that is passed on to consumers for additional services, not covered by Medicare Part A and Part B services ($17). As discussed in Section I, MA insurers typically offer more comprehensive coverage than TM, but they obtain a separate compensation ("rebate") from CMS for it. On average in our baseline sample, the consumer rebate is $51 per enrollee-month, and $34 of it is for more generous coverage of the healthcare services that would be covered by TM and that we study in the paper, while the remaining $17 of the rebate is for additional consumer benefits that are not captured by the analogous TM spending (such as premium discounts, or dental and vision coverage).

[6] This estimate is based on data from the National Association of Insurance Commissioners (NAIC) for the year 2015, which contain information at the insurer-state level on revenues and cost components for each insurer, which are part of a mandatory report by the insurer to the state's insurance commissioner (ideally we would use 2010 data, but 2015 was the first year where numbers for Medicare coverage were reported separately from other lines of insurance). Using these data, we focused on the three HCCI insurers in the 36 states that constitute our baseline sample. Our estimate of 8 percent is the overall ratio in this sample between the sum of "general and administrative costs" to the sum of "net incurred claims after reinsurance."

Figure 2 shows raw spending in MA and TM separately for each of the 36 states in our baseline sample. Spending is lower in MA in all states, but the differences range from about 3 percent lower MA spending in Alaska to over 45 percent lower MA spending in Florida and Vermont.

Geographic variation in spending within TM has attracted a great deal of attention. The "Dartmouth Atlas" findings of large differences across areas in TM spending and utilization without corresponding differences in mortality is widely viewed as indicative of the inefficiencies of the public Medicare system (Fisher et al. 2003a, 2003b; Skinner 2011; Institute of Medicine 2013). Our analysis suggests that, if anything, geographic variation in raw spending is higher in MA than in TM. The coefficient of variation across states (weighting each state by its total Medicare enrollment) is 0.136 in MA, about 20 percent higher than the 0.114 coefficient of variation we estimate in TM.[7] In Appendix Figure A2 we show that MA also exhibits the positive correlation across states between spending and mortality that has been widely documented in TM.

III. Measurement approach

A. (Standard) framework

Lower baseline spending in MA relative to TM may partly or entirely reflect differences in the beneficiaries who enroll in TM and MA. We have already seen in Table 1 that MA enrollees tend to be healthier than TM enrollees. A standard potential outcome framework is therefore useful to organize our measurement exercise. Let $MA_i = 1$ if beneficiary i is enrolled in a plan offered by one of the three HCCI insurers in MA, and $MA_i = 0$ if i is in TM. Let y_i^{TM} be the individual outcome of interest (e.g. healthcare spending per month) if she were in TM, and y_i^{MA} be the individual outcome of interest if she were in MA. We observe $y_i = y_i^{TM}$ when $MA_i = 0$, and we observe $y_i = y_i^{MA}$ when $MA_i = 1$. The individual treatment effect is $\Delta y_i = y_i^{MA} - y_i^{TM}$.

We observe (e.g., in Table 1 Panel B)

$$D = E\left[y_i^{MA} \middle| MA_i = 1\right] - E\left[y_i^{TM} \middle| MA_i = 0\right] = T + S, \quad (1)$$

where T is the average treatment effect for the MA population

$$T = E\left[y_i^{MA} - y_i^{TM} \middle| MA_i = 1\right] \quad (2)$$

and S represents the selection effect, given by

---

[7] Our analysis is at the state level rather than the Hospital Referral Region (HRR) level that is more typical in this literature. This is because many HRRs cross state boundaries and our baseline sample is limited to a subset of states.

$$S \equiv E\left[y_i^{TM} \middle| MA_i = 1\right] - E\left[y_i^{TM} \middle| MA_i = 0\right]. \quad (3)$$

A key advantage – in the context of our data – of the above representation of the selection effect is that it is only a function of y_i^{TM}; this is attractive because the set of observables is richer and more granular in the CMS data than in the HCCI data, and the above representation allows us to analyze the selection effect using CMS data alone, holding the average outcome of interest fixed in the HCCI data.

The second term in the selection equation, $E\left[y_i^{TM} \mid MA_i = 0\right]$, is directly observed in the data. The first term, $E\left[y_i^{TM} \mid MA_i = 1\right]$, is not, so would need to be estimated. Throughout the rest of the paper we report two specific strategies by which we estimate this selection term, as described below.

B. Selection on observables

Selection on “priced” observables.—In our first empirical strategy to correct for selection we reweight the TM population to match the MA population in terms of county and risk score. Within the above framework, it can be viewed as assuming that, conditional on county and risk score, MA_i is as good as random assignment. The risk score is a summary statistic based on an extremely rich set of demographic and health measures. These health measures reflect both patient health and propensity to receive healthcare – since diagnoses are only recorded if care is received (Song et al. 2010; Finkelstein, Gentzkow, and Williams 2016) – both of which may differ between TM and MA enrollees.

Specifically, consider a Medicare enrollee in county z_i with (continuous) risk score r_i, and an outcome y_i^{TM} in TM. We map r_i to a discrete risk score bin r_i', so that all Medicare beneficiaries are partitioned into a set of discrete groups, defined by their county and risk score bin $g_i = \left(z_i, r_i'\right)$. Using the sample of beneficiaries in the CMS data who are enrolled with the HCCI insurers (Table 1, column (6)), we assign each group g a weight $w_g = N_g/N$, where N_g is the number of enrollees that belong to group g and $N = \sum_g N_g$.[8] Each unweighted TM outcome

$$\bar{y}_{unweighted}^{TM} = \frac{1}{N_{TM}} \sum_{i \in TM} y_i^{TM} \quad (4)$$

is then replaced with a reweighted TM outcome

---

[8] A slight complication of this procedure arises when an MA enrollee belongs to a group for which there are no TM enrollees, which may happen in small counties and high (i.e. less common) risk scores. This applies to only 0.07 percent of enrollee-months. In such a case, we amend this procedure with an extra step, where we re-classify to such “empty” TM groups the TM enrollee in the same county whose risk score is the closest to the corresponding unmatched MA enrollee.

$$\bar{y}^{TM}_{re-weighted} = \frac{1}{\sum_{i \in TM} w_{g_i}} \sum_{i \in TM} w_{g_i} y_i, \quad (5)$$

which we compare to the corresponding MA outcome

$$\bar{y}^{MA} = \frac{1}{N_{MA}} \sum_{i \in MA} y_i^{MA}. \quad (6)$$

In addition to the transparency and simplicity of this re-weighting approach, it has the added attraction that it captures the spirit by which MA insurers are being paid by CMS. As described in Section I, CMS payments to MA insurers are based on a county-specific benchmark, and multiplied by the enrollee's risk score r_i. Our baseline approach, which reweights on precisely these two dimensions – county and risk score – can therefore be viewed as correcting for selection concerns associated with the two dimensions by which CMS varies its payments. As mentioned above, following CMS' payment policy for MA insurers during our 2010 study year, we use risk scores for MA enrollees that are deflated by 3.41 percent.

Panel A of Table 2 shows how the TM spending benchmark is affected by different ways of reweighting the TM enrollees to "look like" the MA enrollees in terms of county and risk score composition. Column (1) reproduces the raw, unweighted numbers already shown in column (4) of Table 1. Column (2) of Table 2 reweights the TM data to match the distribution of MA enrollees across counties. Average TM spending per enrollee-month increases from $911 to $942, reflecting the fact that MA enrollees are disproportionately in more expensive counties; this is primarily driven by the well-documented higher MA penetration in urban areas, in which healthcare delivery tends to be more expensive. Columns (3) and (4) add risk scores to the reweighting of the TM population, so that it matches, county by county, the risk score distribution of MA enrollees. In column (3) we match on risk score bins that are quite coarse, of width 0.5; 58 percent of MA enrollees are in the three largest bins (0.5–1, 1–1.5, and 1.5–2). In column (4) we use more granular risk score bins (of width 0.1). It is evident from column (3) (and not surprising given Table 1) that reweighting on risk scores is important, reducing the average monthly spending by 9 percent relative to reweighting on county only in column (2). However, it is quite remarkable that the much more granular matching on the risk score distribution makes little difference, with columns (3) and (4) showing essentially identical results. We will thus use the re-weighting strategy in column (4) – using county and risk bins of width 0.1 – as our first empirical strategy to correct for selection throughout the paper.

Selection on additional observables.—Although county and risk score are essentially the only variables that are currently being conditioned on for the purpose of MA payments, it seems natural to wonder about the extent to which the difference in spending between MA and TM reflects a treatment effect of MA as opposed to selection into MA by individuals who – conditional on risk score and county – have lower predicted spending due to unmeasured differences in health or preferences for healthcare. The relative importance of selection or treatment is particularly important in the context of assessing the cost implications of any expansion of the MA program to cover those currently enrolled in TM.

If we want to condition on a richer set of variables, it gets more difficult to apply the same reweighting strategy as the data become sparse and it is common to observe MA beneficiaries with a vector of characteristics for which there is no match in the TM sample. We therefore instead follow a standard approach of constructing propensity scores for enrollment in MA as a function of a rich set of observables, and then apply the reweighting strategy to the propensity score rather than to the entire vector of variables.

Specifically, given a vector of observables x_i we estimate a logit model of MA_i on x_i. That is, we assume that $p_i = \Pr(MA_i = 1) = \frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)}$ and estimate β by maximum likelihood. We estimate the logit model separately for each county, to allow the relationship between enrollment in MA and observables to differ across counties. We then use our estimate of β to generate the propensity score for individual i, denoted by $\hat{p}_i$. Appendix Figure A3 presents the distribution of the propensity score for the TM and MA populations. We then repeat the same reweighting procedure used in the first empirical strategy, but now with respect to $g_i' = (z_i, \hat{p}_i')$, where the propensity score $\hat{p}_i$ is binned into bins $\hat{p}_i'$ of width 0.01. That is, instead of assuming that (conditional on county) the risk score captures all relevant information that may affect selection, we now replace it with the propensity score of enrolling into MA.

A critical decision, obviously, regards the set of variables x_i that enter the propensity score calculation. The risk score r_i is based on a rich set of observables, including very detailed health measures as well as age, gender, and dual eligibility in Medicaid. These observables are used with a particular functional form to produce the risk score. Using the same underlying variables to generate the propensity score is a natural and less restrictive way to correct for selection, in our setting. In practice, however, the results in Panel B, column (2) show that this approach yields quite similar results to our first approach that adjusts only for selection on priced observables, which we reproduce in Panel B, column (1).

C. Using mortality to address selection on unobservable health

It is less obvious how to correct for selection on unobservables that affect the propensity to enroll in MA and may also be correlated with healthcare spending. Our main approach to address it is to leverage the fact that we can observe mortality outcomes for individuals in both TM and MA. As we saw in Table 1, mortality is lower for MA enrollees than for TM enrollees; it is also lower conditional on county and risk score (not shown). While clearly imperfect, it may provide a rough sense as to how much additional selection may affect the interpretation of the results, and this could vary for different types of healthcare utilization outcomes.

We thus continue by making the strong assumption that mortality outcomes are unaffected by enrollment in MA. Under this assumption, we can use realized mortality rates as a substitute metric to measure health risk. Such a metric would capture potential selection on health risk that is not captured by the risk score used earlier (Brown et al. 2014), and it will also be robust with respect to differential coding of health conditions, which has been shown to be more aggressive in MA relative to TM (Geruso and Layton 2015).

To implement this approach, we estimate the mortality rate (within the same calendar year of 2010) for each individual's county and risk score bin $g_i = \left(z_i, r'_i\right)$, but do so separately for MA and TM enrollees.[9] We can then construct a variable $\widehat{m}_i = m\left(z_i, r'_i, MA_i\right)$, which captures the individual's predicted mortality, and we then turn to our first empirical strategy, but replacing the individual's risk score r_i with the individual's predicted mortality $\widehat{m}_i$, and (as before) reweighting on $g''_i = \left(z_i, \widehat{m}'_i\right)$, where the predicted mortality $\widehat{m}_i$ is binned into bins $\widehat{m}'_i$.

Column (4) of Panel B of Table 2 shows results for this second approach, which additionally adjusts for unobserved health; it shows results where we follow the above exercise and use predicted mortality bins of width 0.01. In column (3) of the same panel we show an alternative, less flexible way to predict mortality, which is based on predicting mortality using risk scores, but not separately county-by-county. The results are similar to those of column (4), illustrating the point that the main results that are generated by our second empirical strategy are primarily generated by the mortality difference (conditional on risk score) between MA and TM enrollees, and not by the precise details of the procedure that adjusts for it. As can be seen, re-weighting by predicted mortality makes a significant difference. Therefore, in what follows we will report comparisons between MA and TM using two approaches: adjusting for selection on priced observables (Table 2, Panel A, column (4)), and additionally adjusting for selection on unobserved health (Table 2, Panel B, column (4)).

IV. Differences in spending in MA and TM

Overall differences.

Table 2 shows average spending differences across all our baseline sample enrollees in MA (column (5)) and comparison samples in TM. The unweighted data (column (1) of Panel A) indicate that healthcare spending in MA is $269 (30 percent) lower per enrollee-month than in TM. Adjusting for differences on priced observables (Panel A, column (4)), we estimate that healthcare spending by MA enrollees is $212 (25 percent) lower per enrollee-month than in a comparable (on county and risk score) sample of TM enrollees. Stated differently, in the spirit of CMS' capitation payment formula, if total healthcare spending of MA enrollees under TM were the same as for TM enrollees with the same risk scores in the same counties, they would cost $855 per enrollee-month, while in MA their total healthcare spending is only $642. Applying this estimate to the entire MA population in 2010 (column (2) of Table 1, which includes those outside of our baseline sample), this translates to $101.5 billion in annual (2010) healthcare spending in TM relative to $76.3 billion in healthcare spending in MA, a difference of $25.2 billion in annual healthcare spending.

---

[9] We have also estimated specifications that replace mortality rates in 2010 with mortality rates over longer horizons (up to 5 years). The overall results are quite similar and do not change much with the length over which mortality rate is measured.

The differences are still positive, but not as large, if in addition we adjust for unobserved health. Doing this (Panel B, column (4)) indicates that healthcare spending in MA is only $64 (9 percent) lower than in a comparable (on county and predicted mortality rate) sample of TM enrollees. Recall that MA insurers are paid based on risk scores, so the higher difference in spending that arises from adjusting for selection on priced observables (Panel A) is more directly associated with the profits of MA insurers from the current set of MA enrollees, while using mortality to adjust for unobserved health may be more relevant in the context of a counterfactual of moving MA enrollees to TM (or vice versa), assuming that it indeed captures much of the selection on unobserved health.

In the remaining tables we compare differences across types of consumers or care. The relative patterns are similar with either adjustment approach, although naturally the quantitative differences are smaller across the board when we additionally adjust for unobserved health.

Differences by consumer type.

Panel A of Table 3 reports the spending differences for different types of enrollees. Each row represents a different subsample of enrollees. Across the board, overall spending in MA tends to be substantially lower than the (re-weighted) TM analog; the average difference reported in Panel A of Table 2 is not driven by any specific sub-population. Yet, we see some heterogeneous effects across types of enrollees. The difference is higher in both absolute and relative terms for older beneficiaries than younger ones. The spending differences are much greater for urban counties, which is where the vast majority (77 percent) of MA beneficiaries enroll, than for rural counties. Put differently, average spending per month in MA is almost the same for rural and urban counties, but TM spending is much higher in urban counties, thus generating the differential difference. This sharp difference between urban and rural counties is also reflected in the MA revenues (i.e. in plan payments for “organic” MA services from Panel C of Table 1), which we estimate to be $205 higher than claims cost in urban counties and only $83 higher in rural ones. Figure 3 shows that states with higher TM spending have greater MA “savings” as measured by the percentage difference between MA spending and adjusted TM spending. This is consistent with the “conventional wisdom” that higher spending TM areas are less efficient or productive (e.g. Skinner 2011).

Panel B of Table 3 compares different quantiles of the MA and TM spending distributions. This allows us to assess, for example, whether the spending difference is driven by the highest spenders. We see the overall lower MA spending across all parts of the distribution, with larger percentage differences at the lower end.

Differences by spending type.

Table 4 reports spending differences across different categories of care. It shows total spending broken down into three mutually exclusive and exhaustive categories: inpatient, outpatient, and SNF. MA spending is lower in all three categories. The biggest difference is in SNF spending, where MA spending is 30–50 percent lower than TM spending for comparable enrollees. However, SNF spending accounts for only a small share (11 percent) of overall spending, so this large percentage difference does not contribute much to the overall difference in spending. The Institute of Medicine (2013) recently called attention to the fact that variation in post-acute spending is a major driver of geographic variation in TM spending. This appears to be true in MA as well, where the geographic variation in SNF spending is even larger (relative to other types of spending) than in TM.[10] We return to the SNF results when we discuss potential mechanisms for reducing healthcare use in Section V.C below.

The bottom row of Table 4 reports hospice spending in MA and TM. As noted earlier, hospice is covered by TM for both MA and TM enrollees. It is therefore not in our HCCI data on MA spending and we do not include it in our baseline "total spending" measure. It is however captured – for both MA and TM enrollees – in the CMS data. We therefore use the CMS data to measure hospice spending for both TM enrollees and enrollees in the three HCCI insurers. Because MA insurers do not bear the cost of hospice expenditures, they might have an incentive to steer patients to hospice, so that some of the lower MA spending in inpatient, outpatient, and SNF could be offset by higher spending in hospice. The bottom row of Table 4 suggests, however, that this is not the case. Hospice spending is too low to have any potential significant offset effect; moreover, it is also lower (rather than higher) for MA enrollees than for TM enrollees.

V. Differences in utilization, not in prices

In this section, we examine whether the difference in overall healthcare spending per enrollee-month between MA and TM is driven by lower healthcare utilization in MA or by the ability of MA insurers (at least the large ones, from which we have data) to negotiate lower prices, or both. One challenge in such an exercise is to conceptually separate prices from quantity or quality of care, and this challenge dictates some of the exercises we report. To preview our results, we find that quantity differences appear responsible for the entire difference; various measures of "prices" are all quite similar in MA and TM.

A. Differences in the propensity of healthcare encounters

Table 5 compares components of healthcare utilization. We examine inpatient days and admissions, days in skilled nursing facilities (SNFs), visits to the emergency department (ED), and physician visits. Inpatient and SNF utilization differences between MA and TM are similar to the analogous spending differences computed in Table 4. Conditional on an inpatient admission, length of stay is also slightly (6 percent) lower in MA. ED visits are lower in MA, reflecting lower utilization both for outpatient ED visits (ED visits that do not result in an inpatient admission) and inpatient ED visits (which do result in an inpatient admission). Physician visits in an outpatient setting are also lower in MA than in TM, with the difference approximately equally driven by the extensive margin (a lower rate of MA enrollees who see a physician at least once a month) and the intensive margin (a lower average number of physician visits by MA enrollees who visit the physician at least once).

---

[10] For example, compared to the coefficient of variation across states of 0.11 in overall (unadjusted) TM spending and 0.14 in overall MA spending (see Figure 2), we estimate a coefficient of variation in SNF (unadjusted) TM spending of 0.19, and 0.33 in SNF MA spending. By contrast, relative geographic variation in inpatient and outpatient spending in TM and MA is similar to the overall comparison (not shown).

Interestingly, additional adjustment for unobserved health essentially eliminates utilization differences for inpatient-related measures, just as it did for inpatient-related spending (Table 4). This pattern is consistent with our adjustment for unobserved health fully adjusting for health differences between TM and MA enrollees, and MA insurers having no discretion over inpatient utilization, which is fully driven by health events.

Over-used and under-used care.—In Table 6 we explore differences in potential low-value and high-value care. Panel A examines utilization of diagnostic testing and imaging services, where over use may be a concern (e.g. Brot-Goldberg et al. 2017; U.S. Government Accountability Office 2008). Panel B examines utilization of various measures of preventive care, an area where under-use may be a concern (Brot-Goldberg et al. 2017).[11] We see lower utilization in MA for both low-value and high-value care. Diagnostic tests and imaging procedures are lower in MA by similar percentages as total spending. Preventive care exhibits no obvious pattern relative to overall care; rates of most preventive care are lower in MA, although there is variation across the measures.

In Panel C we use a widely-used algorithm developed by Billings, Parikh, and Mijanovich (2000) to classify ED visits by their “appropriateness.” The algorithm uses primary diagnosis codes for the visit to distinguish between visits that represent an emergency (i.e. require care within 12 hours) and non-emergency visits (e.g. a toothache). Within emergency visits, it further distinguishes between those that require treatment in the ED (as opposed to being treatable in a primary care setting, such as a lumbar sprain). Finally, within emergency visits that require ED care, it distinguishes between those that were and were not preventable by timely ambulatory care. Appendix B provides more detail on the algorithm and its validation. The results indicate similar proportional change in each type of ED visit, irrespective of its “appropriateness.”

Overall these results suggests that MA is a relatively blunt instrument for reducing healthcare utilization, with “high value” and “low value” care showing similar proportional differences with TM. Interestingly, the bluntness of supply-side instruments such as managed care is mirrored on the demand side, where recent work suggests that high deductible health insurance plans are similarly non-discriminatory in discouraging both high-value and low-value care utilization (Brot-Goldberg et al. 2017) and Medicaid coverage for the previously-insured encourages increases in ED visits of all types, including (and perhaps particularly) non-emergency visits (Taubman et al. 2014).

B. (Lack of) Mean price differences for hospital admissions for specific diagnoses

Table 7 shows spending per encounter in MA and TM. Given the close similarity between the percentage difference in utilization measures in Table 5 and the percentage difference in the corresponding spending measure in Table 4, it is not surprising that spending per encounter is quite similar between MA and TM. Inpatient spending per admission, inpatient spending per day and SNF spending per SNF day are essentially the same in MA and TM.

---

[11] We show rates of preventive care by enrollee-month to be consistent with the analysis in the rest of the paper. Naturally, recommended care is not at a monthly level but typically at an annual (or biannual) level. The analysis looks similar if instead we examine these measures on an annual basis (not shown).

Interestingly, spending per outpatient ED visit is 9–10 percent higher in MA; this may reflect utilization management for MA patients that discourages relatively less severe cases from coming to the ED or from being admitted from the ED to the hospital. We also note that neither reweighting approach makes much difference for inpatient spending; the spending per encounter statistics are quite similar already in the raw comparison of means.

This similar spending per encounter for MA and TM enrollees suggests that prices may be similar in MA and TM. However, spending per encounter can also be affected by differences in providers seen or in reason for the visit. To hone in on differences in "prices" – or unit payment rates – we compare payments in MA and TM for admission to the *same* hospital with the *same* DRG.[12] Under TM, hospitals are paid by CMS based on a pre-set formula that is a product of a hospital-specific rate and a DRG-specific rate; it is our understanding (although no contractual data is available to verify it) that these hospitals are predominantly paid by MA insurers in a similar way. In TM, and presumably in MA as well, some accommodation for exceptions is allowed, resulting in payments that may deviate from the DRG-hospital formula rates.

We compute a parallel set of prices in MA and TM. For both, our starting unit of analysis is an admission in MA, which is characterized by a hospital and a DRG. The MA price is simply the observed (transacted) payments for the admission in the MA claims data. Construction of the TM price proceeds in two steps. First, for each MA admission, we calculate the formula price in TM, applying the PPS reimbursement formula which, as noted, is a function of the hospital and the DRG. Second, we adjust our TM formula prices to reflect average differences between TM formula and TM actual (transacted) prices since we are comparing to actual (transacted) prices in MA.[13] Appendix C provides more detail.

Figure 4 shows our estimate of the average price in TM and MA overall, and for the top 20 DRGs (by their share of MA admissions); Appendix Table A2 provides the underlying numbers. In reporting DRG-specific average prices, we weight the admissions in each DRG by the state's share of MA admissions in all DRGs, so that any differences in average prices across DRGs within MA (or within TM) reflect price differences for a common "state basket," and are not contaminated by differences in the geographic distribution of admissions by DRG across states. The national average price is computed by weighting each DRG by its (national) share of MA admissions.

Inpatient prices are extremely similar in MA and TM. The national average admission price is $9,945 in TM and $10,054 in MA. The price for an average MA admission is only 1.1 percent higher in MA relative to TM. The largest difference among the top 20 DRGs is for chest pain (DRG #313), for which the average MA price is about 6 percent lower than in TM. For 10 of the top 20 DRGs, the average price in MA is within 2 percent of that in TM.

---

[12] For this pricing analysis, we focus on the approximately 4,000 hospitals in our baseline sample that are paid (by TM) under Medicare's prospective payment system (PPS). These represent about 95 percent of all inpatient admissions in MA and cover essentially all standard (non-specialty) hospitals.

[13] In principle, we could follow the exact same approach as for MA prices, and estimate transacted TM prices directly in the CMS data, where we observe TM payments for each admission, along with its hospital and DRG. In practice, however, we are constrained from doing this for two reasons: hospital identifiers are encrypted in the MA data, and our DUAs prohibit our exporting data below a minimum cell size. Fortunately, the TM hospital-specific base payment rates (which determine the TM formula payments) are available in our MA data; we are extremely grateful to Zack Cooper for providing us with this mapping. We construct actual and formula TM prices in the CMS data and use these to construct adjustment factors to reflect average differences between TM formula and actual prices by DRG or by state.

The close similarity of inpatient admission prices between MA and TM echoes similar findings by Baker et al. (2016) and is interesting given that it is frequently conjectured that because the public sector has greater bargaining power, public fee-for-service may achieve lower prices than private insurance (e.g. Philipson et al. 2010). Consistent with this conjecture, prior empirical work has shown that for the same service, TM tends to reimburse at substantially lower prices than commercial (under 65) private insurance both in the outpatient setting (Clemens and Gottlieb 2017) and the inpatient setting (Cooper et al. 2015). In contrast, we do not find that TM prices are substantially lower than MA prices.[14] One potential explanation for this discrepancy is that regulation requires hospitals to accept fee-for-service Medicare rates for Medicare beneficiaries when they are not included in the MA plan's network; as a result MA plans may have greater bargaining power - and thus obtain lower rates - than commercial plans that serve the under-65 population. Berenson et al. (2015) provide more details on this institutional environment, and report on results from a survey of hospital and MA plan executives, which are very consistent with our findings.

Geographic variation in hospital prices.—We also compare geographic variation in inpatient prices for MA and TM. We construct average state prices in MA and TM following a parallel process to what we did for measuring DRG prices; here, we weight the admissions in each state using the DRG's national share of MA admissions, so that comparisons of state-level average prices within MA (or within TM) are not contaminated by differences in the mix of DRGs across states.

Figure 5 shows the results; Appendix Table A3 shows the underlying numbers. Pricing variation across states (weighted by Medicare enrollment) is about 20 percent lower in MA than in TM. Specifically, the coefficient of variation across states is 0.067 in MA, compared to 0.082 in TM. By contrast, recent work has shown evidence of substantially higher geographic pricing variation in commercial (less than 65) private plans compared to TM (Philipson et al. 2010, Institute of Medicine 2013, Cooper et al. 2015).[15]

C. Potential channels for saving

Our results thus far strongly point to differences in utilization metrics, rather than payment rates, that are driving the overall differences in spending between TM and MA. How might MA plans reduce healthcare utilization? Some mechanisms that have been proposed include limited provider networks through which beneficiaries receive care, coordination of care programs to more efficiently deliver appropriate services and avoid excessive utilization, and financial incentives to physicians to influence the quality and quantity of services delivered (e.g. Landon et al. 2012). By contrast, in TM there are virtually no restrictions on physician clinical decisions or patient choices of care.

---

[14] Of course, our MA sample is limited to three large insurers, and their bargaining power may not be representative of smaller MA insurers; on the other hand, Cooper et al. (2015)'s analysis of commercial pricing was also limited to the same three large insurers, and there average inpatient prices were almost twice as high as in TM.

[15] Like us, this analysis focuses on pricing variation in hospitals. The recent Cooper et al. (2015) comparison of pricing variation in TM compared to commercial (i.e. private, under 65) plans also uses data from HCCI, specifically 2007–2011 data for commercial insurance. We confirmed that we replicate their finding of substantially greater variation in pricing in commercial insurance relative to TM when, as with our main analysis here, we use data only from 2010 and from the subset of 36 states in our baseline analysis. Specifically, using the MA share of admissions in each DRG to construct average prices for each state, and estimating the coefficient of variation across states weighting each state by the Medicare enrollment in that state (as in Figure 5), we estimate that pricing variation is over 50 percent larger in commercial insurance (coefficient of variation = 0.14) than in TM (coefficient of variation = 0.08).

We have already seen evidence of one “signature” of MA mechanisms to reduce care utilization: all these mechanisms should constrain patient entry into care, particularly expensive care, so that the average person using that care in MA is in worse health, and has higher cost than the average person using that care in TM. In other words, MA enrollees should have fewer encounters, but have greater spending (or utilization) per encounter. Consistent with this, we found that spending per outpatient ED visit was in fact slightly higher in MA than in TM (see Table 7).

In Table 8 we provide additional evidence consistent with restrictions on utilization. In Panel A we explore differences between TM and MA in the distribution of discharge destinations of hospitalized patients. Destinations are roughly ordered in how expensive they are (from cheaper to more expensive). Inpatients covered by MA are disproportionately discharged to less expensive destinations. In particular, discharges to SNFs (or other post-acute care) are substantially less common, while discharges home (or to home health services) are relatively more likely.

In addition to limiting use of care, MA may also constrain the type of service, encouraging use of less expensive substitutes. Panel B points to some patterns that are suggestive of such channels. First, we analyze the frequency of surgeries. We find the surgery rate to be in fact higher, not lower, in MA by about 20 to 30 percent. However, inpatient surgeries are similar and outpatient surgeries are much higher, which is suggestive of MA insurers using outpatient surgeries to substitute away from inpatient surgeries and perhaps (given the fact that overall number of surgeries is higher) from other types of expensive, non-surgical admissions as well. Second, we examine two types of physician visits: primary care and specialist visits. We already saw in Table 5 that MA enrollees are associated with fewer physician visits. The results in Table 8 show that this is driven primarily by fewer specialist visits; rates of primary care visits are similar.

VI. Conclusion

We have compared healthcare spending and utilization in public and private Medicare. This setting provides a rare opportunity for a “side by side” comparison of public and private health insurance systems operating on a similar scale, for the same population, in the same markets, and with the same providers. Novel data from the Health Care Cost Institute on the healthcare claims of MA enrollees allow us a rare look inside the “black box” of healthcare utilization and spending in MA.

We find that MA insurer revenues are 30 percent higher than their healthcare spending. Healthcare spending per enrollee-month in MA is 30 percent lower than in TM; holding enrollee county and risk score fixed, this spending difference shrinks to 25 percent, and adjusting for mortality differences further reduces it to 9 percent.

The lower spending by MA enrollees is entirely due to lower healthcare utilization. Prices appear similar in MA and TM. Where we can most directly measure this – the price of an admission for a given DRG at a given hospital – we estimate that average prices in MA are 1.1 percent higher than in TM. Reductions in utilization appear similar both for types of care where there is concern about “over use” (e.g. imaging and diagnostic tests) and where there is concern about “under use” (e.g. preventive care).

We provide suggestive evidence for some of the potential channels by which MA may reduce healthcare utilization for enrollees. We find that utilization is lower in MA but that, conditional on an encounter, spending per encounter is similar or slightly higher in MA. This suggests that MA plans restrict utilization on the margin to sicker individuals. Relatedly, individuals discharged from the hospital are much more likely to be sent home – and less likely to be sent to a post-acute care facility – if they are enrolled in MA rather than in TM. We also find evidence consistent with substitution to less expensive types of care in MA; for example, differences in specialist visits are much larger than differences in primary care visits.

Finally, in light of the widespread interest in geographic variation in healthcare spending in TM, and recent work on geographic variation in commercial (under 65) private insurance, we explore similar comparisons in MA. Although geographic variation in spending in TM is often viewed as a reflection of the inefficiencies in a public health insurance system, we find similar – in fact slightly larger – geographic variation in spending in MA compared to TM. And while recent work has emphasized the much greater geographic pricing variation in private commercial insurance than in TM, we find similar – in fact slightly smaller – geographic variation in pricing in MA compared to TM.

One natural question these findings raise is their implications for MA insurers and consumers. For insurers, our estimates from MA data indicate that their revenue exceeds their healthcare expenditures by $177 (about 30 percent) per enrollee-month. An important area for further work is to examine how this varies with competitive and other market conditions, and whether these potential cost savings may be dissipated through other forms of costs, such as the administrative costs of providing the insurance and the marketing costs of attracting enrollees.

Implications for consumers are more elusive, since the elements of their objective function are not as straightforward to define or measure. A simple revealed preference argument would suggest that consumers who choose MA are better off in MA than in TM. Other inferences are harder to make. Quality of the healthcare experience is difficult to assess; our measures of preventive care point to reductions there that are similar in magnitude to those for other forms of care. We calculated that the mean actuarial benefit to consumers (i.e. rebates that are passed on to consumers in the form of other benefits) was $51 per enrollee-month, but, of course, the rebate may be valued differently from its actuarial value, and MA plans have other attributes that will affect consumer surplus, such as limited networks. The implications of privately provided Medicare for both consumers and producers is an important area for further work.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Ayanian John Z., Landon Bruce E., Saunders Robert C., Pawlson L. Greg, and Newhouse Joseph P.. 2013 “Quality of Ambulatory Care in Medicare Advantage HMOs and Traditional Medicare.” Health Affairs 32(7): 1228–35. [PubMed: 23836738]

Baker Laurence C., Bundorf M. Kate, Devlin Aileen M., and Kessler Daniel P.. 2016 “Medicare Advantage Plans Pay Hospitals Less than Traditional Medicare Pays.” Health Affairs 35(8): 1444–51. [PubMed: 27503970]

Boards of Trustees. 2011 “2011 Annual Report of the Boards of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds.” Report by the Boards of Trustees, Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/ReportsTrustFunds/Downloads/TR2011.pdf (accessed January 9, 2018).

Berenson Robert A., Sunshine Jonathan H, Helms David, and Lawton Emily. 2015 “Why Medicare Advantage Plans Pay Hospitals Traditional Medicare Prices.” Health Affairs 34(8): 1289–95. [PubMed: 26240241]

Billings John., Parikh Nina, and Mijanovich Tod. 2000 Emergency Room Use: The New York Story. Commonwealth Fund. http://www.commonwealthfund.org/publications/issue-briefs/2000/nov/emergency-room-use–the-new-york-story

Brot-Goldberg Zarek, Chandra Amitabh, Handel Benjamin, and Kolstad Jonathan. 2017 “What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics.” Quarterly Journal of Economics 132(3): 1261–1318.

Brown Jason, Duggan Mark, Kuziemko Ilyana, and Woolston William. 2014 “How Does Risk Selection Respond to Risk adjustment? Evidence from the Medicare Advantage Program.” American Economic Review 104(10): 3335–64. [PubMed: 29533567]

Cabral Marika, and Cullen Mark. 2011 “The Effect of Insurance Coverage on Preventive Care.” http://www.marikacabral.com/Cabral_CullenPreventiveCare.pdf

Centers for Medicare and Medicaid Services. 2010 “Announcement of Calendar Year (CY) 2010 Medicare Advantage Capitation Rates and Medicare Advantage and Part D Payment Policies.” https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/downloads/announcement2010.pdf

Clemens Jeffrey, and Gottlieb Joshua. 2017 “In the Shadow of a Giant: Medicare’s Influence on Private Physician Payments.” Journal of Political Economy 125(1): 1–39. [PubMed: 28713176]

Clemens Jeffrey, Gottlieb Joshua, and Molnar Timea. 2017 “Do Health Insurers Innovate? Evidence from the Anatomy of Physician Payments.” Journal of Health Economics 55: 153–67. [PubMed: 28784289]

Cooper Zack, Craig Stuart, Gaynor Martin, and Van Reenen John. 2015 “The Price Ain’t Right? Hospital Prices and Health Spending on the Privately Insured.” NBER Working Paper 21815.

Curto Vilsa, Einav Liran, Levin Jonathan, and Bhattacharya Jay. 2014 “Can Health Insurance Competition Work? Evidence from Medicare Advantage.” NBER Working Paper 20818.

Duggan Mark, Gruber Jonathan, and Vabson Boris. 2018 “The Consequences of Health Care Privatization: Evidence from Medicare Advantage Exits.” American Economic Journal: Economic Policy 10(1): 153–86.

Einav Liran, Finkelstein Amy, and Polyakova Maria. 2018 “Private Provision of Social Insurance: Drug-specific Price Elasticities and Cost Sharing in Medicare Part D.” American Economic Journal: Economic Policy 10(3).

Finkelstein Amy, Gentzkow Matthew, and Williams Heidi. 2016 “Sources of Geographic Variation in Health Care: Evidence from Patient Migration.” Quarterly Journal of Economics 131(4): 1681–1726. [PubMed: 28111482]

Fisher Elliott S., Wennberg David E., Stukel Therese A., Gottlieb Daniel J., Lucas FL, and Pinder Etoile L.. 2003a "The Implications of Regional Variations in Medicare Spending: the Content, Quality, and Accessibility of Care. Part 1." Annals of Internal Medicine 138(4): 273–87. [PubMed: 12585825]

Fisher Elliott S., Wennberg David E., Stukel Theresa A., Gottlieb Daniel J., Lucas FL, and Pinder. 2003b "The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care" Annals of Internal Medicine 138(4): 288–99. [PubMed: 12585826]

Gawande Atul. 2009 "The Cost Conundrum: What a Texas Town Can Teach Us about Health Care." The New Yorker, 6 1.

Geruso Michael, and Layton Timothy. 2015 "Upcoding: Evidence from Medicare on Squishy Risk Adjustment." NBER Working Paper 21222.

Institute of Medicine. 2013 "Variation in health care spending: Target decision making, not geography." Washington DC: The National Academies Press.

Kaiser Family Foundation. 2010 https://kaiserfamilyfoundation.files.wordpress.com/2013/01/8080.pdf

Landon Bruce, Zaslavsky Alan, Saunders Robert, Pawlson L. Gregory, Newhouse Joseph and Ayanian John. 2012 "Analysis of Medicare Advantage HMOs Compared with Traditional Medicare Shows Lower Use of Many Services During 2003–2009." Health Affairs 12: 2609–17.

McGuire Thomas, Newhouse Joseph P., and Sinaiko Anna D.. 2011 "An Economic History of Medicare Part C." The Milbank Quarterly 89(2): 289–332. [PubMed: 21676024]

Philipson Tomas, Goldman Dana, Seabury Seth, Lakdawalla Darius, and Lockwood Lee. 2010 "Geographic Variation in Health Care: The Role of Private Markets." Brookings Papers on Economic Activity: 325–55.

Skinner Jonathan. 2011 "Causes and Consequences of Regional Variations in Health Care" In Handbook of Health Economics. Vol. 2, eds. Pauly Mark V., McGuire Thomas G. and Barros Pedro P., 45–93. Elsevier.

Song Yunjie, Skinner Jonathan, Bynam Julie, Sutherland Jason, and Fisher Elliott. 2010 "Regional Variations in Diagnostic Practices." New England Journal of Medicine 363(1): 45–53. [PubMed: 20463332]

Newhouse Joseph, and McGuire Thomas. 2014 "How Successful in Medicare Advantage?" The Milbank Quarterly 92(2): 351–94. [PubMed: 24890251]

Taubman Sarah, Allen Heidi, Wright Bill, Baicker Katherine, and Finkelstein Amy. 2014 "Medicaid Increases Emergency-Department Use: Evidence from Oregon's Health Insurance Experiment." Science 343(6168): 263–68. [PubMed: 24385603]

U.S. Government Accountability Office. 2008 Medicare Part B Imaging Services: Rapid Spending Growth and Shift to Physician Offices Indicate Need for CMS to Consider Additional Management Practices. http://www.gao.gov/products/GAO-08-452 (accessed February 23, 2015)

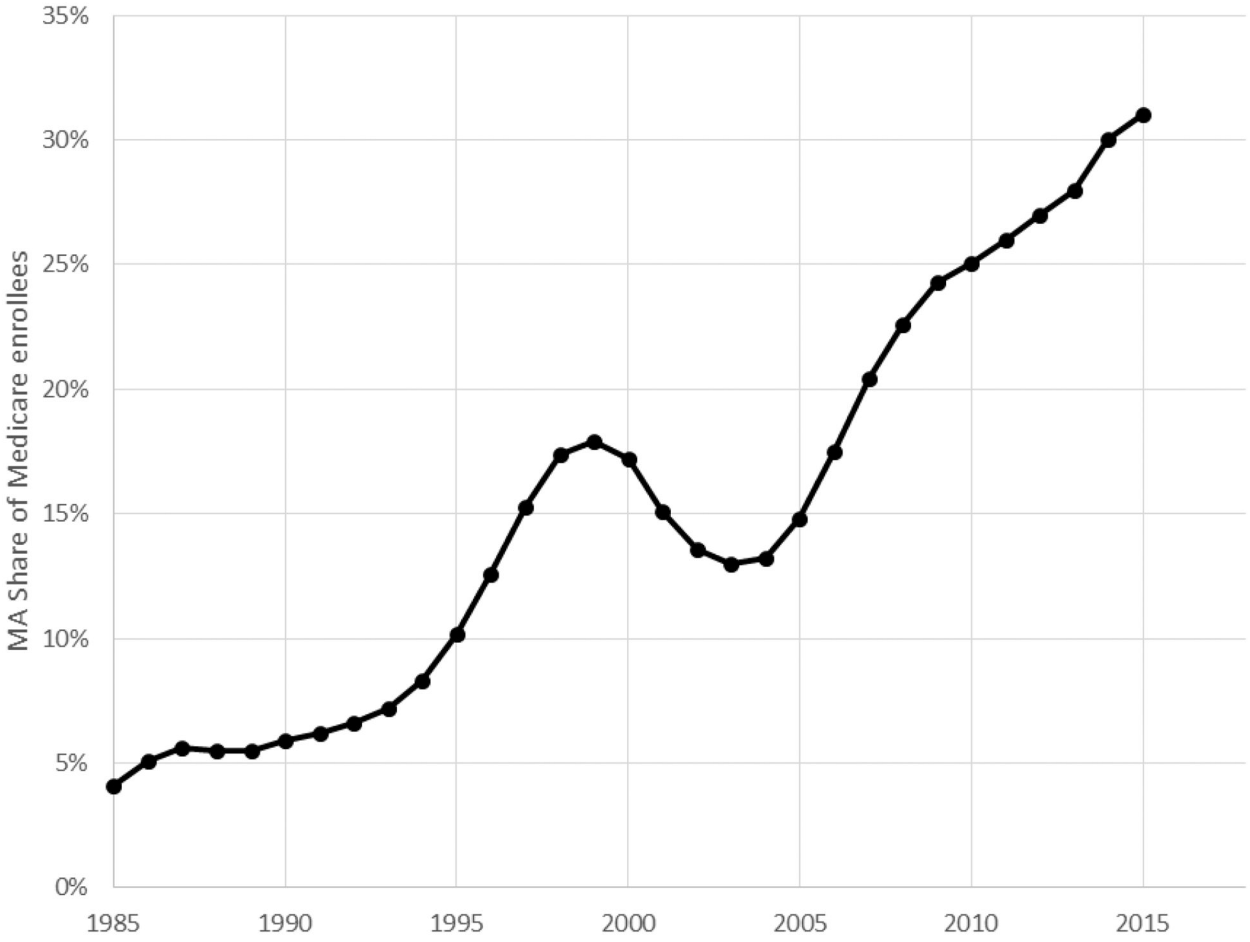

Figure 1: MA penetration over time

Figure shows the share of Medicare beneficiaries enrolled in Medicare Advantage plans, year by year. The data source is CMS' Medicare Managed Care Contract Plans Monthly Summary Reports. All data are from December of the year indicated.

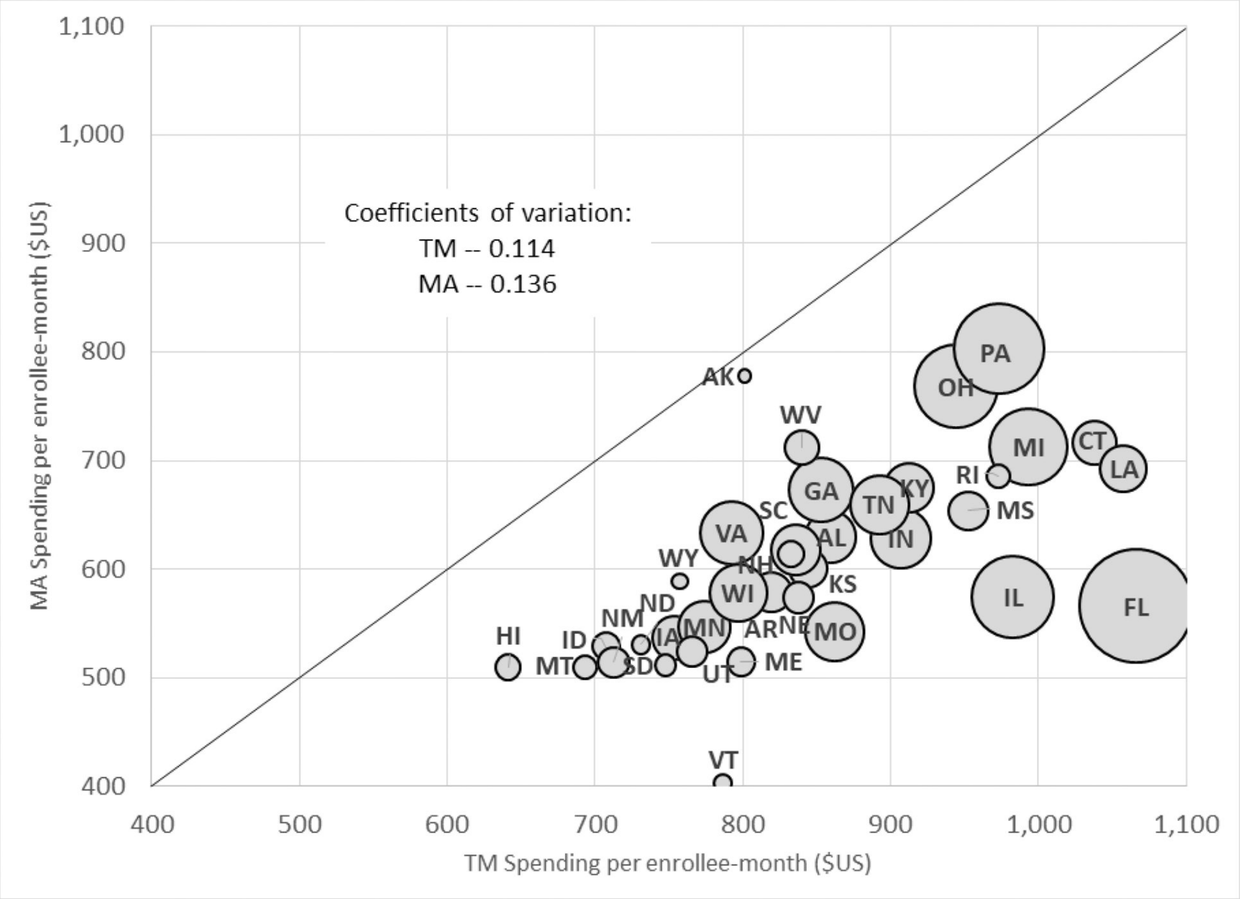

Figure 2: State-by-State Comparison of TM and MA Spending

Figure plots MA spending per enrollee-month against TM spending per enrollee-month for each of the 36 states in our baseline sample. Coefficients of variation across states in spending are computed using total Medicare enrollees in the state as a weight. The size of each bubble is proportional to the number of total Medicare enrollees in the state.

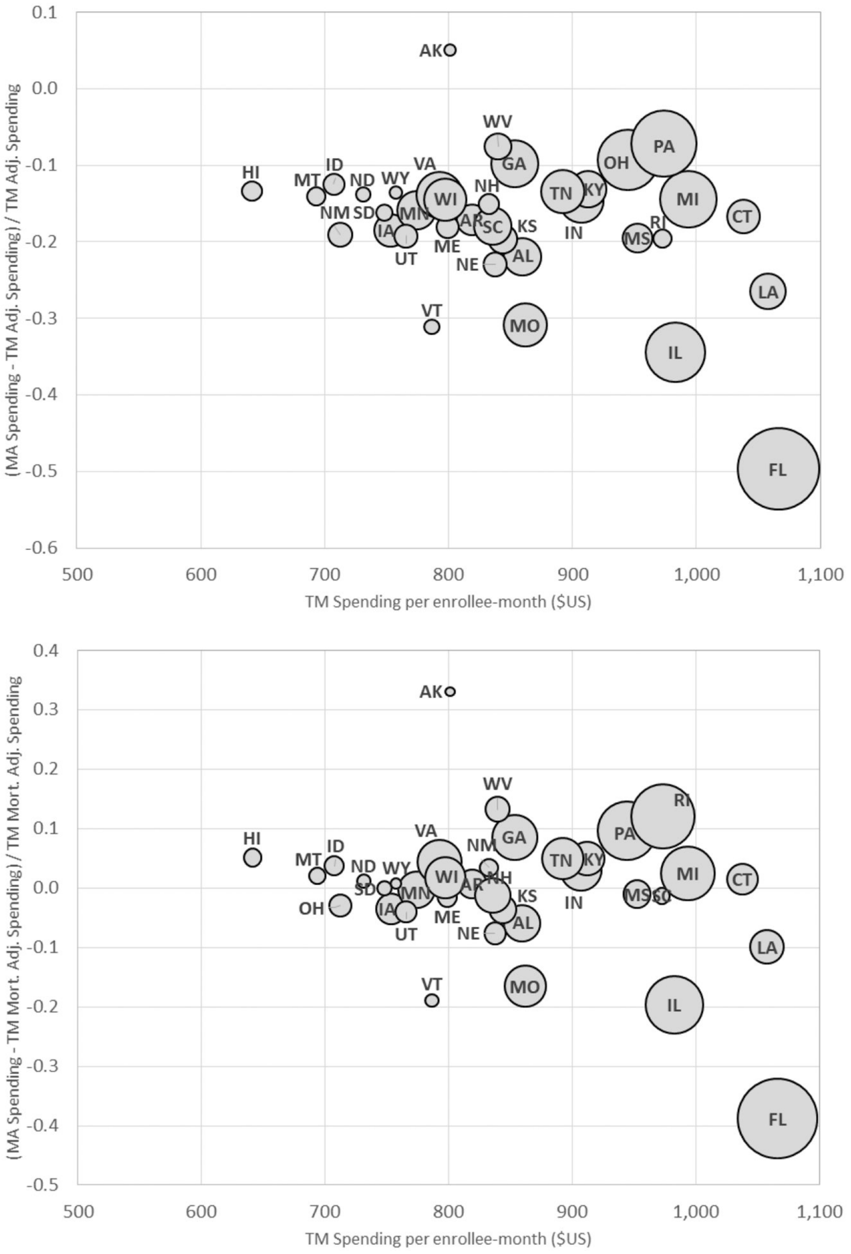

Figure 3: TM-MA Spending Differences across States

Figure plots the (percentage) difference between average MA spending and (re-weighted) TM spending per enrollee- month against average TM spending for each of the 36 states in our baseline sample. The y-axis in the top panel compares MA spending to TM spending that is re-weighted to match the MA population on county and risk score, using our preferred weighting (see Table 2, Panel A, column (4)). The bottom panel does the same but using predicted mortality to adjust for selection on unobservables (see Table 2, Panel B, column (4)), as described in Section III. The size of each bubble is proportional to the number of total Medicare enrollees in the state. The x-axis reports average (unadjusted) TM spending in the state (see Table 2, Panel A, column (1)).

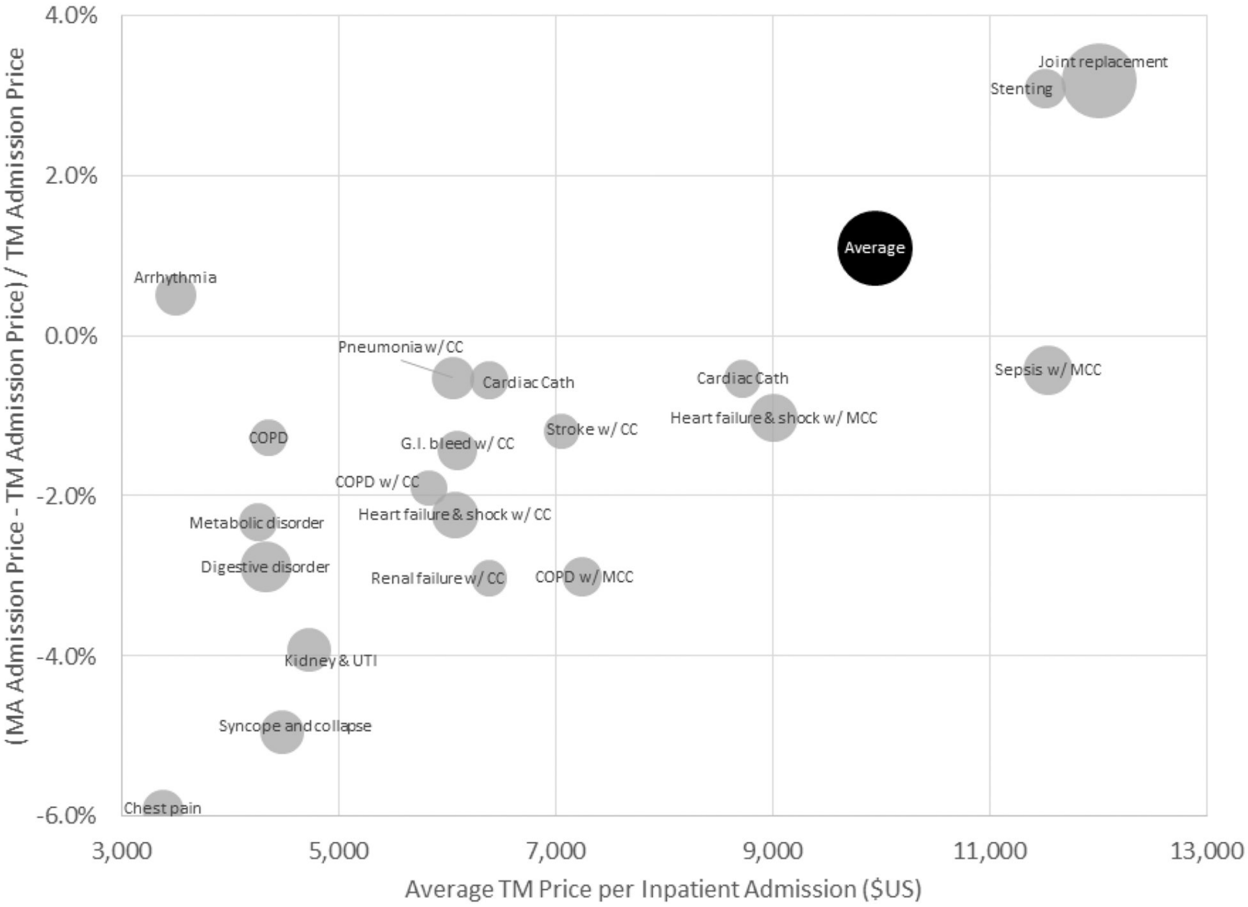

Figure 4: TM-MA price differences for inpatient admissions, across DRGs

Figure plots the (percentage) difference between average MA prices and TM prices for a hospital admission, overall and for the 20 most common DRGs in MA. Average MA or TM prices for a given DRG are computed using a common (MA) basket of state admission shares for that DRG. The national average price in MA or TM is computed by weighting each DRG (including the less common ones not shown here) by its (national) share of MA admissions. The size of each bubble (except for the overall “Average” bubble) is proportional to the number of MA admissions with that DRG.

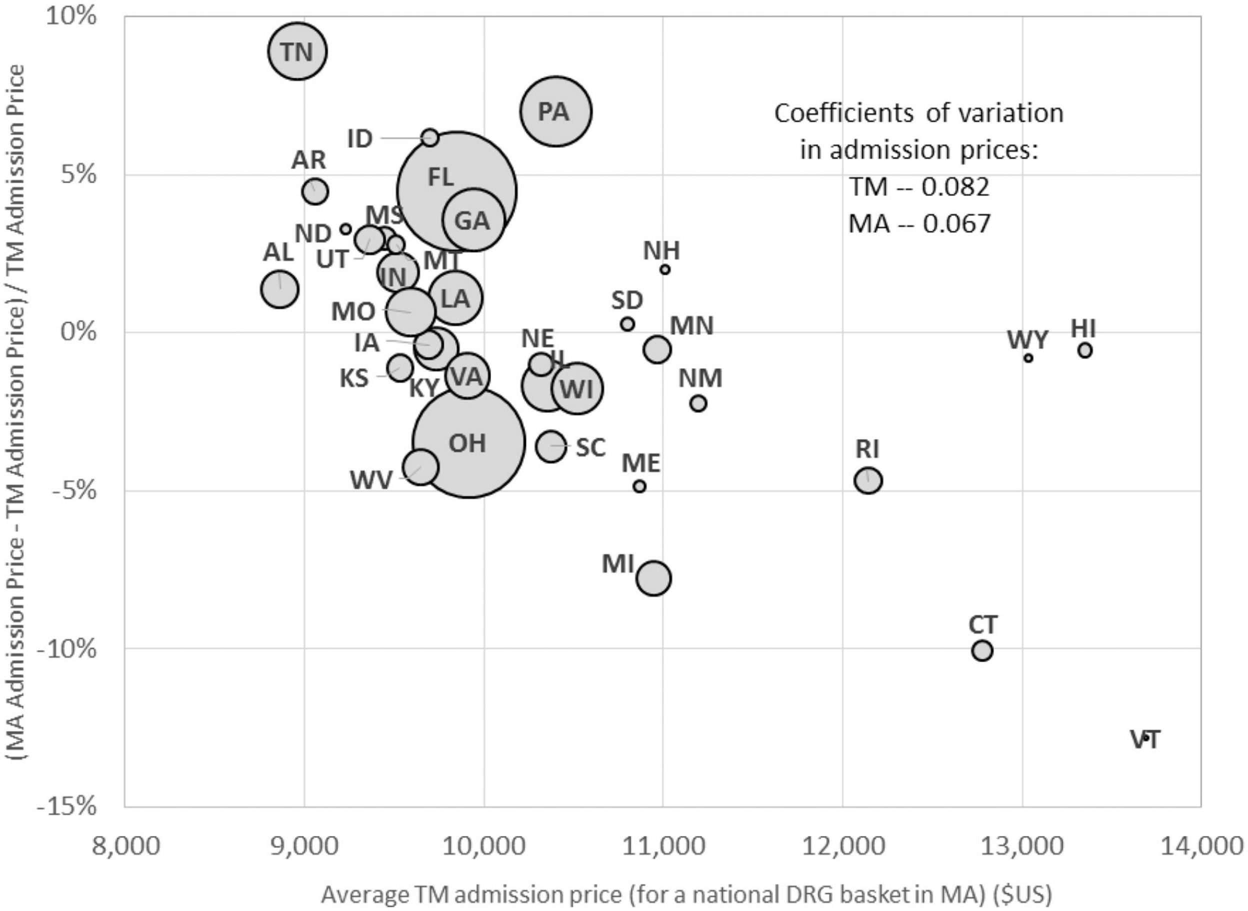

Figure 5: TM-MA price differences for inpatient admissions, across states

Figure plots the (percentage) difference between average MA prices and TM prices for a hospital admission for each state in our baseline sample (except Alaska which is omitted because it has too few inpatient admissions for us to report). Averages are computed for each state using a common (MA) "basket" of DRG admission shares. The size of each bubble is proportional to the number of MA admissions in that state. Coefficients of variation across states in prices are computed using total Medicare enrollees in the state as a weight.

Table 1:

Baseline sample

Data source / sample	All CMS[a]			Baseline CMS[b]			All HCCI[a]	Baseline HCCI[b]
	TM (1)	MA (all insurers) (2)	MA (HCCI insurers) (3)	TM (4)	MA (all insurers) (5)	MA (HCCI insurers) (6)	MA (HCCI insurers) (7)	MA (HCCI insurers) (8)
PanelA: Enrollee-level summary[c]								
No. of enrollees (000s)	26,420	10,475	3,911	15,641	5,291	2,270	2,941	2,290
Female	0.575	0.574	0.574	0.576	0.567	0.568	0.569	0.571
Age	75.4	74.6	74.5	75.4	74.3	74.1	--	--
Coarse age:[d]								
65–74	0.520	0.555	0.560	0.516	0.568	0.581	0.592	0.590
75–84	0.330	0.328	0.325	0.333	0.323	0.315	0.306	0.308
85+	0.150	0.117	0.115	0.151	0.109	0.104	0.102	0.102
Dual eligible	0.143	0.123	0.111	0.129	0.072	0.073	--	--
SNP enrollees	--	0.081	0.065	--	0.000	0.000	0.000	0.000
Riskscore	1.089	1.031	1.032	1.085	0.986	0.994	--	--
Died in 2010	0.050	0.039	0.039	0.052	0.036	0.036	--	--
Panel B: Spending per enrollee-month[e]								
No. ofenrollee-months (000s)	304,908	118,737	44,371	180,608	60,273	25,867	32,506	25,394
Total Spending ($/month)	938	--	--	911	--	--	639	642
Insurer Spending ($/month)	798	--	--	771	--	--	586	590
OOP Spending ($/month)[f]	140	--	--	140	--	--	53	52
Panel C: Payments to insurers per enrollee-month[e]								
Overall CMS expenditure ($)[g]	--	820	819	--	767	778	--	--
Actuarial value of incremental consumer benefits ($)[h]	--	63	53	--	56	51	--	--
Plan payments for organic MA services ($)[i]	--	800	806	--	751	767	--	--

Table presents summary statistics for various sample definitions. Columns (6) and (8), highlighted in gray, are comparable and are used to validate our sample construction.

[a] Sample include all Medicare enrollees who are 65 or older by the end of 2010.

[b] Baseline sample excludes SNP enrollees, and enrollees in the 15 states in which the number of enrollee-months in HCCI is not within 10 percent of that in CMS.

[c] At the enrollee-level, we define an individual as enrolled in TM if she is never enrolled in MA during the sample year and is enrolled in TM for at least one month of the sample year; we define her as enrolled in MA if she is enrolled in MA in any month of the year, and we assign her to an HCCI insurer if she is covered by one of them in her first month in MA. Age, dual eligibility and SNP enrollment is likewise defined based on the first month in which an enrollee is observed during the sample year.

[d] In HCCI we only have information about age in three bins: 65–74, 75–84, and 85+.

[e] We count an enrollee-month in TM if she is enrolled in TM that month and never enrolled in MA during the sample year; any enrollee-months in MA (or in HCCI insurers) are counted as such.

[f] Out of pocket (OOP) spending denotes amount owed by enrollee. For TM enrollees, OOP Spending may be partially covered by supplemental (Medigap or employer-sponsored) coverage.

[g] This includes all payments made from CMS to the MA plans, including risk-adjusted payments and rebates.

[h] This is also known as the “rebate.”

[i] The variable “Plan payments for organic MA services ($)” is equal to “Overall CMS expenditure ($)” plus additional premiums paid by the beneficiaries minus the non-cost-sharing component of the rebate.

Table 2:

Baseline reweighting

Source	CMS (TM)				HCCI (MA)	Difference	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
No. of enrollee-months (000s)	180,608	180,608	180,608	180,608	25,394		
Panel A. Reweight using risk score							
Reweight by	None	County	County & Risk score bin 0.5	County & Risk score bin 0.1	None	(5)−(4)	((5)−(4)) / (4)
Total Spending ($/month)	911	942	857	855	642	−212	−24.9%
Insurer Spending ($/month)	771	799	725	723	590	−133	−18.4%
OOP Spending ($/month)[a]	140	143	132	131	52	−79	−60.4%
Panel B. Reweight using predicted mortality							
Reweight by	County & Risk score bin 0.1	County & Prop. score bin 0.01[b]	Pred. mortality bin 0.01	County & pred. mortality bin 0.01	None	(5)−(4)	((5)−(4)) / (4)
Total Spending ($/month)	855	861	698	706	642	−64	−9.0%
Insurer Spending ($/month)	723	729	586	594	590	−4	−0.7%
OOP Spending ($/month)[a]	131	131	112	112	52	−60	−53.5%

Results based on baseline sample (see Table 1, columns 8 and 4). All statistics are at the enrollee-month level.

[a] Out of pocket (OOP) spending denotes amount owed by enrollee. For TM enrollees, OOP Spending may be partially covered by supplemental (Medigap or employer-sponsored) coverage.

[b] Propensity score is computed by running a logit regression of MA indicator on the components of the risk score formula: age, gender, Medicaid (dual) indicator, and HCC fixed effects.

Table 3:

Spending differences for different groups of enrollees

	% MA enrollees	TM, unweighted	TM, weighted[a]	TM, mort. weighted[a]	MA	Difference	
						((5)−(3)) / (3)	((5)−(4)) / (4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
No. of enrollee-months (000s)	25,394	180,608	180,608	180,608	25,394		
Total Spending	100%	911	855	706	642	−24.9%	−9.0%
Panel A. Spending ($/month) by enrollee characteristics							
Male	43%	916	857	696	673	−21.4%	−3.3%
Female	57%	907	853	713	619	−27.4%	−13.2%
65–74	56%	723	661	534	540	−18.2%	1.2%
75–84	33%	1,022	967	874	731	−24.4%	−16.4%
85+	11%	1,264	1,276	1,137	898	−29.6%	−21.0%
Urban[b]	77%	942	887	733	645	−27.3%	−12.0%
Rural[b]	23%	851	752	622	634	−15.7%	1.9%
Panel B. Realized distribution ofspending ($/month)							
Proportion w/ no spending		0.37	0.38	0.43	0.46	19.6%	7.7%
Median spending		93	84	64	38	−54.3%	−40.2%
75th pctile		332	317	262	222	−30.0%	−15.0%
90th pctile		1,314	1,233	977	849	−31.1%	−13.1%
95th pctile		3,433	3,124	2,396	2,161	−30.8%	−9.8%
97.5th pctile		8,349	7,571	5,835	5,690	−24.8%	−2.5%
99th pctile		18,510	17,332	14,672	13,614	−21.5%	−7.2%

Results based on baseline sample (See Table 1, columns 8 and 4). All statistics are at the enrollee-month level. All spending numbers are in $/month.

[a] Weighting based on our preferred weighting, as in column (4) of both panels in Table 2.

[b] Rural/urban assignment is based on whether the enrollee zip code is in an MSA.

Table 4:

Spending differences for different components of spending

	TM, unweighted (1)	TM, weighted[a] (2)	TM, mort. weighted[a] (3)	MA (4)	Difference ((4)−(2)) / (2) (5)	 ((4)−(3)) / (3) (6)
No. of enrollee-months (000s)	180,608	180,608	180,608	25,394		
Total spending[b]	911	855	706	642	−24.9%	−9.0%
Inpatient	364	333	270	269	−19.2%	−0.4%
Outpatient	452	435	371	328	−24.6%	−11.4%
Skilled Nursing Facilty (SNF)	95	86	65	45	−48.2%	−31.4%
Hospice[c]	31	32	23	24	−24.9%	1.8%

Results based on baseline sample (See Table 1, columns 8 and 4). All statistics are at the enrollee-month level. All spending numbers are in $/month.

[a] Weighting based on our preferred weighting, as in column (4) of both panels in Table 2.

[b] Total spending is the sum of inpatient, outpatient, and skilled nursing facility (SNF) spending. It doesn't include hospice.

[c] Hospice expenditures for MA enrollees are billed directly to CMS, so for MA enrollees they are in fact observed in the CMS data and not in the HCCI data.

Table 5:

Differences in healthcare utilization

	TM, unweighted	TM, weighted[a]	TM, mort. weighted[a]	MA	Difference	
					((4)−(2)) / (2)	((4)−(3)) / (3)
	(1)	(2)	(3)	(4)	(5)	(6)
Total spending ($/month)	911	855	706	642	−24.9%	−9.0%
Inpatient days	0.200	0.181	0.143	0.144	−20.6%	0.4%
Any inpatient admission	0.027	0.025	0.020	0.021	−16.0%	3.7%
Days cond'l on any	7.4	7.4	7.2	6.9	−5.5%	−3.1%
Skilled Nursing Facility (SNF) days	0.336	0.296	0.219	0.131	−55.9%	−40.3%
Days cond'l on any	47.3	46.7	45.4	20.6	−55.8%	−54.6%
Emergency Department (ED) Visits	0.049	0.045	0.037	0.038	−15.8%	1.9%
Outpatient ED visits	0.031	0.028	0.024	0.024	−14.8%	−0.2%
Inpatient ED visits	0.018	0.017	0.013	0.014	−17.5%	5.6%
Physician visits	1.22	1.21	1.10	1.01	−16.8%	−8.0%
Any physician visits	0.545	0.540	0.503	0.486	−10.0%	−3.5%
Numberofvisits cond'l on any	2.24	2.25	2.18	2.08	−7.5%	−4.6%

Results based on baseline sample (See Table 1, columns 8 and 4). All statistics are at the enrollee-month level, but all days associated with a given encounter are attributed to the original admission date, even if it extends beyond the month.

[a] Weighting based on our preferred weighting, as in column (4) of both panels in Table 2.

Table 6:

Utilization differences across different types of care

	TM, unweighted	**TM, weighted[a]**	**TM, mort. weighted[a]**	**MA**	**Difference**	
					((4)−(2)) / (2)	**((4)−(3)) / (3)**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. Testing and imaging						
Diagnostic tests	2.12	2.05	1.79	1.55	−24.4%	−13.3%
Any diagnostic test	0.35	0.34	0.31	0.293	−14.3%	−4.2%
Cond’l on any	5.97	6.00	5.84	5.29	−11.9%	−9.5%
Imaging procedures	0.66	0.64	0.57	0.52	−18.9%	−8.9%
Any imagingtest	0.18	0.17	0.16	0.154	−10.5%	−2.2%
Cond’l on any	3.75	3.71	3.62	3.37	−9.3%	−6.9%
B. Preventive care (rates per relevant population)[b]						
Flu shot	0.051	0.050	0.048	0.032	−36.7%	−35.0%
Cardiovascular screen	0.090	0.093	0.090	0.077	−16.9%	−13.7%
Colorectal cancer screen	0.009	0.010	0.010	0.008	−14.9%	−16.3%
Mammogram	0.045	0.046	0.046	0.047	2.5%	1.4%
Pap smear	0.011	0.012	0.013	0.013	7.9%	−0.5%
Hemoglobin A1ctest	0.064	0.062	0.051	0.055	−11.9%	8.3%
Blood lipids test	0.103	0.106	0.102	0.091	−14.8%	−11.4%
Eye exam	0.067	0.067	0.066	0.054	−20.4%	−18.6%
C. Appropriateness of ED Visits						
Nonemergent	0.006	0.005	0.005	0.005	−14.7%	−0.2%
Emergent						
ED care not needed (primary care treatable)	0.012	0.011	0.009	0.009	−15.8%	−0.1%
ED care needed, preventable	0.004	0.004	0.003	0.003	−18.4%	6.3%
ED care needed, not preventable	0.012	0.011	0.010	0.009	−16.6%	−1.4%
Unclassified	0.013	0.012	0.010	0.010	−19.9%	0.4%

Results based on baseline sample (See Table 1, columns 8 and 4). All statistics are at the enrollee-month level.

[a] Weighting based on our preferred weighting, as in column (4) of both panels in Table 2.

[b] Rates are per the relevant population, which is: everyone for flu shot, cardiovascular screen, and colorectal cancer screen; women for pap smear; women aged 65–74 for mammogram; and enrollees aged 65–74 with a diabetes diagnosis for hemoglobin test, blood lipids test, and eye exam.

Table 7:

Differences in spending per episode of care

	TM, unweighted	TM, weighted[a]	TM, mort. weighted[a]	MA	Difference	
					((4)−(2))/(2)	((4)−(3))/(3)
	(1)	(2)	(3)	(4)	(5)	(6)
Total spending ($/month)	911	855	706	642	−24.9%	−9.0%
Spending perSNF day	381	379	383	378	−0.2%	−1.4%
Spending per outpatient ED visit	782	768	760	837	9.0%	10.1%
Inpatient[b]:						
Spending per admission	10,134	10,151	10,206	10,093	−0.6%	−1.1%
Spending perday	1,901	1,903	1,950	1,908	0.3%	−2.1%

Results based on baseline sample (See Table 1, columns 8 and 4). All statistics are at the enrollee-month level, but all expenditures or days associated with a given encounter are attributed to the original admission date, even if it extends beyond the month.

[a] Weighting based on our preferred weighting, as in column (4) of both panels in Table 2.

[b] Inpatient spending here includes only payments to the hospital; it does not include associated physician payments as in prior tables.

Table 8:

Potential channels for cost saving

	TM, unweighted	TM, weighted[a]	TM, mort. weighted[a]	MA	Difference	
					((4)−(2)) / (2)	((4)−(3)) / (3)
	(1)	(2)	(3)	(4)	(5)	(6)
A. Hospital discharge destinations:						
Home	0.0136	0.0122	0.0104	0.0109	−10.4%	5.4%
Home health service org.	0.0053	0.0049	0.0039	0.0038	−23.3%	−4.2%
SNF	0.0067	0.0061	0.0047	0.0038	−37.6%	−17.5%
Other post-acute care	0.0014	0.0013	0.0010	0.0004	−70.5%	−63.4%
Other (incl. hospice, death)	0.0027	0.0024	0.0018	0.0018	−27.3%	−2.9%
B. Surgeries and specialists:						
Total surgeries	0.037	0.033	0.029	0.039	18.1%	33.0%
Outpatient surgeries	0.029	0.026	0.023	0.032	25.5%	41.2%
Inpatient surgeries	0.008	0.007	0.007	0.007	−7.2%	4.8%
Primary care visits	0.379	0.370	0.334	0.355	−3.8%	6.5%
Specialist visits	0.840	0.844	0.764	0.655	−22.4%	−14.3%

Results based on baseline sample (See Table 1, columns 8 and 4). All statistics are at the enrollee-month level. All spending numbers are in $/month. Panel A reports (unconditional) hospital discharge destinations.

[a] Weighting based on our preferred weighting, as in column (4) of both panels in Table 2.