

How does risk selection respond to risk adjustment? New evidence from the Medicare Advantage Program*

Jason Brown
U.S. Treasury Department

Mark Duggan
Univ. of Pennsylvania and NBER

Ilyana Kuziemko
Columbia and NBER

William Woolston
Stanford University

October 18, 2012

Abstract

To combat adverse selection, governments increasingly base payments to health plans and providers on enrollees' scores from risk-adjustment formulae. In response to evidence of plan overpayments due to selection, in 2004 Medicare began to risk-adjust capitation payments to private Medicare Advantage plans. But because the variance of medical costs increases with the predicted mean, incentivizing enrollment of individuals with higher risk scores can increase the scope for enrolling "over-priced" individuals with costs significantly below the formula's prediction. Indeed, after risk adjustment, MA plans enrolled individuals with higher scores but significantly lower costs conditional on their score, and overpayments actually increased.

*We thank Richard Boylan, Doug Bernheim, Sean Creighton, David Cutler, Liran Einav, Randy Ellis, Zeke Emanuel, Gopi Shah Goda, Jonathan Gruber, Caroline Hoxby, Seema Jayachandran, Larry Katz, Robert Kocher, Jonathan Kolstad, Amanda Kowalski, Alan Krueger, Elena Nikolova, Christina Romer, Shanna Rose, Karl Scholz, Jonathan Skinner, and Luke Stein and seminar participants at Brown, Columbia, Cornell, Harvard, Houston, Northwestern, MIT, Princeton, RAND, Rice, Stanford, USC, U.S. Treasury Department, Wharton, Wisconsin, Yale, and the NBER Health Care meetings for helpful comments and feedback. We thank Nadia Tareen and Boris Vabson for their outstanding research assistance. Contact information for the authors: Jason.Brown@do.treas.gov, mduggan@wharton.upenn.edu, ik2216@columbia.edu, william.woolston@stanford.edu. Views expressed here are solely those of the authors and not of the institutions with which they are affiliated, and all errors are our own

1 Introduction

Recent health care reforms have attempted to move away from the fee-for-service (FFS) payment model—which economists have long argued incentivizes over-provision of services—by paying providers or insurers fixed capitation payments rather than reimbursing them for each service. The success of such reforms hinges on correctly aligning capitation payments with a patient’s expected cost and setting incentives for providers to deliver the appropriate care. If capitation payments are not aligned with actual costs, plans and providers will have incentives to cream-skim the most under-priced cases instead of competing on quality or cost. To more accurately equate payments with expected costs, governments and other insurance sponsors have increasingly turned to “risk adjustment”—setting payments to insurers or providers to take account of an individual’s past and current health conditions.

Both the Affordable Care Act of 2010 (ACA) and alternative proposals advocated by its opponents rely heavily on risk adjustment. Approximately 25 million people are projected to join the “insurance exchanges” established by the ACA, in which private insurers will receive capitation payments adjusted for enrollees’ health status. The law also promotes cost-control experiments such as “bundled payments,” under which providers receive a capitated, risk-adjusted payment for a certain condition such as diabetes or a hip fracture, rather than payment for each procedure.¹ Similarly, the budget passed by the House of Representatives in 2011 that repealed the ACA also called for turning Medicare into a premium-support system, which would use risk adjustment to combat adverse selection among the private plans competing for beneficiaries.² Many European countries have also used risk adjustment as they increase the roll of private actors in their universal insurance systems (Saltman and Figueras, 1998).

Despite the prominence of risk adjustment in recent health reform proposals, there has been limited empirical work assessing its effect on risk-selection or on the total cost of insuring a given population.³ We provide an assessment of the largest risk adjustment effort

¹In the U.S., many of the reforms which rely on risk adjustment are promoted as reducing health care costs. However, certainly not all cost-control proposals rely on risk adjustment. See Chandra *et al.* (2011), Blumenthal and Glaser (2007), and Baicker and Chandra (2005) on, respectively, comparative-effectiveness research, health information technology, and malpractice reform.

²This bill failed in the Senate.

³There is a large, mostly theoretical or statistical, literature on risk adjustment, and Van de Ven and Ellis (2000) and Ellis (2008) serve as excellent reviews. Recently, work has focused on “optimal” risk adjustment, following Glazer and McGuire (2000) who argue that mere predictive models (such as the one used by Medicare, on which we focus the empirical work) are fundamentally misguided because formula coefficients need to be chosen for their incentive, not predictive, properties. However, as noted by Ellis (2008), predictive models are by far the most common risk adjustment models in use today, and thus determining their effect on selection and costs is a central policy question. On the empirical side, Bundorf *et al.* (2008) provides estimates on the welfare gains to risk adjustment of health insurance *premiums*.

to date in the U.S. health care sector—Medicare’s risk adjustment of capitation payments to private Medicare Advantage (MA) plans, which the ACA suggests as the model for risk adjustment in the state-run insurance exchanges—on selection into MA plans and on the government’s total cost of financing Medicare benefits.⁴ Since the 1980s, Medicare enrollees have been able to enroll in either the traditional fee for service (FFS) program or in an MA plan, which can provide additional services but must cover the basic benefits guaranteed by traditional Medicare. For an individual in an MA plan, the government pays the plan a capitation payment meant to cover the cost of providing her Medicare benefits. Today, over one-fourth of Medicare’s 49 million enrollees receive their care through a private MA plan, and Appendix Figure 1 shows how this share has evolved during the period examined in this paper (1994 - 2006).

Before 2004, an MA enrollee’s capitation payment was, essentially, based on the average cost of FFS enrollees with the same demographic characteristics and was not adjusted for health conditions. Even though regulations required MA plans to offer the same plan at the same price to all Medicare beneficiaries in its geographical area of operation, researchers found that less costly individuals were systematically more likely to enroll in an MA plan.⁵ In response to this evidence of “differential payments” to MA plans—payments in excess of the actual cost of providing a given individual her Medicare benefits—in 2004 Medicare began to base capitation payments on an individual’s “risk score,” generated by a risk-adjustment formula that accounts for over seventy disease conditions.

We show that differential payments are actually higher for enrollees who join MA after risk adjustment than they were for enrollees joining MA before the reform. As such, risk adjustment appears to have increased overpayments and thus the government’s total cost of financing the care of Medicare enrollees. We offer a simple model of risk-selection that can account for this unexpected finding and generates several predictions for how the characteristics of enrollees joining MA plans will change in response to risk adjustment, which we test using data from the Medicare Current Beneficiary Survey (MCBS). The MCBS contains both administrative and survey data and allows us to reconstruct risk scores for the vast majority of respondents (a process we describe in Section 4) and to conduct detailed comparisons of how risk adjustment changes the characteristics of those who switch from FFS to MA.

Before risk adjustment, MA plans had an incentive to enroll individuals who were low cost on all dimensions, and thus would generally avoid individuals with the conditions that

⁴See *Section 1343 — Risk Adjustment* of the ACA legislation, which suggests that “criteria and methods” similar to the HCC model be used in the exchanges.

⁵See, e.g., Langwell and Hadley (1989), Physician Payment Review Commission (1997), Mello *et al.* (2003) and Batata (2004).

would later be included in the risk formula. After risk adjustment, MA plans have less incentive to avoid individuals with the conditions included in the formula. We show that, as predicted by the model, relative to individuals who remain in FFS, MA enrollees' risk scores increase after risk adjustment, consistent with plans no longer avoiding individuals with the conditions included in the formula.

However, the model emphasizes how selection can take place on different margins. While risk adjustment indeed decreases plans' scope for advantageous selection along the dimensions *included* in the formula, it increases the incentive to find individuals who are positively selected along dimensions *excluded* from the formula and are thus "cheap for their risk score." Indeed, as the model predicts, actual costs *conditional on the risk score* of those joining MA fall substantially after 2003, relative to those remaining in FFS.

Finally, the model demonstrates that the former effect (the decrease in advantageous selection along dimensions included in the model) can be more than offset by the latter effect (the increased selection conditional on the risk score). The key insight is that because the variance of medical costs increases with the expected mean, it is easier to find individuals with high risk scores who have, say, costs \$2,000 below their capitation payment than it is to find individuals who do not have a single documented disease condition who are \$2,000 cheaper than predicted. To take but one example from our data, pre-risk-adjustment, Hispanics were roughly \$1,200 cheaper on average than their (non-risk-adjusted) capitation payments; after risk adjustment, Hispanics with a history of congestive heart failure (one of the most common conditions included in the risk formula) are \$3,500 cheaper than their (risk-adjusted) capitation payments. Intuitively, before risk adjustment MA plans "fished" in a pond of relatively healthy enrollees with little cost variance. Risk adjustment allows them to fish in a pond of enrollees who have high costs on average but also highly variable costs. Due to this increase in variance, the ability of firms to enroll individuals with costs substantially below the formula's prediction—whether through targeted advertising or designing benefits packages that differentially appeal to certain people based on demographics or disease history—can actually increase after risk adjustment, and with it the total cost of the Medicare program.

This counterintuitive consequence of risk adjustment has, to the best of our knowledge, not been noted by other researchers, but is related to the literature on the unintended consequences of increasing the specificity of incomplete contracts. By selecting individuals with low costs conditional on their risk scores, MA firms' behavior is analogous to the worker who focuses on the contractable task to the detriment of other tasks (as in Holmstrom and Milgrom, 1991) or the instructor who "teaches to the test" at the expense of other educational goals (as in Lazear, 2006). More generally, our results suggest that using additional infor-

mation to determine prices can sometimes aggravate problems associated with asymmetric information, as in Einav and Finkelstein (2011).

We estimate that differential payments to MA plans in 2006 totaled at least \$15 billion, or over four percent of total Medicare spending that year, significantly greater than official government estimates, which assume that risk adjustment works perfectly.⁶ Given the importance of Medicare to the federal budget, the effect of risk adjustment on government expenditure is of interest in its own right. It is possible, however, that despite increasing government overpayments, risk adjustment was welfare-enhancing because those overpayments generated greater producer or consumer surplus than before risk adjustment. While our evidence on these questions is less definitive than that on government expenditure, after analyzing a variety of data sources—on consumer satisfaction and health care utilization, vital statistics data on mortality rates, and evidence from the historical record—we find no evidence that producers or consumers in the MA market enjoy increases in surplus that offset the increase in government costs or that, more generally, the Medicare program as a whole functions more efficiently or provides greater insurance value after risk adjustment.

The remainder of the paper is organized as follows. Section 2 provides background information on the MA program and the risk-adjustment formula Medicare currently uses. Section 3 presents the intuition and results from the model. Section 4 describes the data. Sections 5 and 6 present the empirical results on selection and differential payments, respectively. Section 7 explores potential mechanisms by which MA plans might be able to differentially select certain enrollees. Section 8 explores the welfare consequences of risk adjustment and discusses ways to improve it, and Section 9 concludes.

2 Background on Medicare Advantage capitation payments and risk adjustment

Since the 1980s, Medicare enrollees have had the choice between the traditional fee-for-service (FFS) system and private MA plans. Plans must accept all applicants, charge all enrollees the same premium, and provide benefits generally comparable to traditional Medicare, but otherwise are free to coordinate patient care and thus can have varying benefits, cost-sharing arrangements, and provider networks. The Medicare program pays MA plans a fixed capitation payment to cover these costs, and plans are, essentially, the residual claimants if actual

⁶Unless otherwise stated, all dollars amounts reported in the paper are adjusted to 2007 dollars using the CPI-U. Note that MedPAC, which annually publishes estimates of “overpayments” to MA plans that receive considerable media attention, explicitly states that their estimates assume no health differences between MA and FFS enrollees after accounting for risk adjustment. Though, as we note later, MedPACs 2012 report suggests they are interested in relaxing this assumption.

costs are above or below the level of the capitation payment.⁷ Since 2006, Medicare Part D has provided enrollees coverage for prescription drugs, though all of our analysis will focus on Part A (hospital and inpatient) and B (physician and outpatient), as these are the services MA plans are required to provide.⁸

The capitation payment to an MA plan for covering an individual is based on the estimated Part A and B payments had FFS Medicare covered her directly. During the 1980s and 1990s, the Center for Medicare and Medicaid Services (CMS)—the agency that administers Medicare—used a “demographic model” to perform this estimation, so-called because it included only demographic variables (gender, age, and disability, Medicaid and institutional status) as opposed to disease or health conditions. Then as now, CMS did not require MA plans to report cost or claims data—doing so might be seen as undermining plans’ freedom to coordinate care as they deem most effective—so it used FFS data to regress total Part A and B spending the following year on these demographic factors (and many of their interactions). Again due to lack of MA cost data, the predictive power of the model can only be evaluated on the FFS population, even though MA enrollees are actually the group being risk adjusted—CMS found that only one percent of FFS costs were explained by the demographic model (Pope *et al.*, 2004).

In response to research showing that MA plans systematically enrolled beneficiaries who were cheaper than the demographic model would predict, CMS attempted to enhance the risk-adjustment procedure.⁹ In 2000, CMS experimented with making ten percent of capitation payments dependent on inpatient claims data, raising the R^2 of the formula from one to 1.5 percent.¹⁰ More significantly, in 2004—which for convenience we term the “start” of risk adjustment, even though the earlier models were also forms of risk adjustment—CMS introduced the hierarchical condition categories (HCC) model, still in use today. Again due to lack of MA cost or claims data, the HCC model, like the demographic model, uses data from the FFS population to calibrate a model that predicts FFS costs in the following year, but, importantly, the HCC model accounts for not just demographic data but also the dis-

⁷An important exception is the cost of hospice care, which FFS Medicare covers even for MA enrollees.

⁸MA plans that provide prescription drug coverage receive a separate capitation payment in return. All of our analysis on the fiscal impact of MA plans considers only the payments made to plans for covering Part A and B services.

⁹Estimates suggest that individuals switching from traditional FFS to MA had medical costs between 20 and 37 percent lower than observably similar individuals who remained in FFS. This range is taken from the estimates in Langwell and Hadley (1989), Physician Payment Review Commission (1997), Mello *et al.* (2003) and Batata (2004). Related research has found evidence of favorable selection into private Medigap plans during this period as well, which provide supplemental coverage to enrollees’ in traditional Medicare (Fang *et al.*, 2008).

¹⁰This Principal Inpatient Diagnostic Cost Group (or PIP-DCG) model itself had an R^2 of 6.2, but as it only accounted for ten percent of the capitation payment, the overall R^2 was $.9 * 1 + .1 * 6.2 = 1.52$.

ease conditions noted on FFS providers' claims. The model distills the roughly 15,000 ICD-9 codes that providers list on claims into seventy disease-category indicator variables, the most common of which are described in Appendix Table 1. Initially, the HCC model was blended with the demographic model, with the HCC model accounting for 30, 50, 75 and 100 percent of the total risk score in, respectively, 2004, 2005, 2006, and 2007 or later.

CMS found that within the FFS population, the HCC score explained eleven percent of FFS costs the following year (Pope *et al.*, 2004). Newhouse *et al.* (1997) and Van de ven and Ellis (2000) survey the literature and conclude that the lower bound on the percent of cost variation plans are able to predict is between 20 and 25 percent, suggesting there is still potential room for risk selection even if the model were to perform as well on the MA population as it does on the FFS population. Similarly, both prospective reports commissioned by CMS in 2000 and 2004 (Pope *et al.*, 2000 and Pope *et al.*, 2004) and more recent work using data from 2004 to 2006 (Frogner *et al.*, 2011) have found that—again, looking only at the FFS population—the formula systematically under-predicts costs for those with the most serious health conditions, a fact we return to in Section 5.

Of course, what matters for overpayments is how well the formula performs on the MA population, not the FFS population, and it will likely perform worse on this group for at least three reasons. First, out-of-sample prediction is more difficult than in-sample prediction. Second, CMS has found that MA plans exhibit greater “coding intensity” in documenting disease conditions than do FFS providers; because the model is calibrated to the coding practice of the latter group, “coding intensity” increases MA overpayments. For example, what an FFS provider might code as “diabetes” an MA plan would call “diabetes with complications,” thus increasing the enrollee’s risk score and capitation payment.¹¹ For empirical reasons we discuss later, most of our results do not include the effects of intensive-coding, though it is a central concern for any model, such as the HCC formula, that requires plans to submit information over which they have some discretion, as opposed to the largely pre-determined factors that comprise the demographic model.

Third, and central to our paper, the performance of the HCC model in predicting costs for the FFS population is not necessarily reflective of its performance for the MA population because those joining MA could be exactly those for whom the formula overpredicts costs. As the model in the next section demonstrates, the introduction of risk adjustment

¹¹See www.cms.gov/MedicareAdvtgSpecRateStats/Downloads/Advance2010.pdf for CMS’s analysis on intensive coding. Note that this analysis does not mean that FFS providers are immune to the incentive to “up-code” diagnoses in order to increase reimbursements, a practice documented by Silverman and Skinner (2004) and Dafny (2005), but merely that they do not do so as intensely as MA plans. Interestingly, the General Accounting Office argues in a January 2012 report that CMS’s estimates of MA plans’ intensive coding are too small (<http://www.gao.gov/assets/590/587637.pdf>.)

will incentivize plans to selectively target individuals whom they expect to have low costs conditional on their risk score.

3 Theoretical framework

In this section, we demonstrate that by incentivizing plans to select individuals who have low costs conditional on their risk scores, risk adjustment can actually increase the total cost to the government of providing a given public service. We will illustrate the key implications of the model with a “toy” example and relegate all formal propositions and proofs to the Appendix.

3.1 Cost assumptions

In a system such as Medicare, where the government directly provides insurance for a guaranteed set of benefits but also finances private plans to cover those same benefits, the total cost to the government is the sum of the direct cost of providing the benefits to individuals who choose to remain in FFS and capitation payments to private firms for individuals who instead choose to receive these benefits via an MA plan. To determine whether a capitation-payment policy increases or decreases the government’s cost of providing the guaranteed set of benefits, it is sufficient to determine whether it increases or decreases *differential payments*—the payments the government gives private plans for providing MA enrollees their Medicare benefits minus the cost had the government directly covered them.¹²

We call an individual’s *actual cost* the cost to the government had she been covered by FFS, and though it is not necessary, assume here that the plan’s cost of providing the basic FFS benefits package is the same as the government’s.¹³ We decompose this actual cost into two components: a risk score and a residual.

¹²This framework assumes that the cost to the government of directly covering a beneficiary is independent of the composition of the population enrolled in MA. But higher MA penetration might incentivize cost-control practices among providers, and these practices could spill over to how they treat their FFS beneficiaries. In this case, increased MA penetration could decrease total Medicare expenditures, even in the presence of positive differential payments as defined above. Using data from 1994 - 2001, Chernew *et al.* (2008) finds support for this hypothesis. However, using more recent data from 1999 - 2004, Nicholas (2009) finds no evidence of these cost-reducing spillovers.

¹³Whether, holding selection constant, private plans or the government can more efficiently provide a given individual with the basic benefits package certainly affects the total (public plus private) cost of providing Medicare benefits. However, it does not affect the propositions of the model: how selection reacts to a change in risk adjustment and the effect of risk adjustment on differential payments. Whether the HMO model is actually more efficient than the traditional government fee-for-service model even absent selection effects is an open question. Duggan (2004) finds that when some California counties mandated their Medicaid recipients to switch from the traditional FFS system to an HMO, costs increased by 17 percent relative to counties that retained FFS. As, within a county, individuals did not select between FFS or an HMO, selection issues are unlikely to be driving the result.

Although MA plans must accept any individual who wishes to join, we assume that a firm can expend resources to influence the characteristics of its enrollees. These *screening costs* might include the costs of devising benefits packages that appeal to one group but not another, targeted advertising and recruiting, or even the risk of sanctions due to violating open-enrollment regulations. We assume that the magnitude of these screening costs depends on an individual’s risk score and residual costs. For example, enrolling individuals with an average risk score and average (zero) residual should be cheapest—plans could just open their doors and take all comers.

Importantly, because the variance of health costs increases with the mean, we assume that it is easier to find an individual with a serious disease condition who is, say, \$1,000 less expensive than her risk score would suggest than it is to find someone without a single documented disease condition who has costs \$1,000 below her risk score’s prediction. That the variance of medical costs increases with the predicted mean is well known (see, e.g., Lumley *et al.*, 2002), and can be seen in Figure 1, which plots actual costs as a function of (binned) risk scores, using our Medicare data. The difference between the mean and the 10th percentile for risk scores between zero and 0.25 is \$2,000; the corresponding difference for risk scores between 2.75 and 3.25 is an order of magnitude larger. Thus, as the risk score increases, it becomes relatively easier to find people who have (in absolute, not percentage terms) large negative residuals.

These ideas are illustrated by the toy example. Assume there are three types of beneficiaries, each representing a third of the population. Type A individuals have no conditions and, as can be seen in Appendix Table 2, have actual costs of 5. Type B individuals were diagnosed with cancer last year but are in remission; they have actual costs of 6. Type C individuals were also diagnosed with cancer last year and are currently receiving chemotherapy; they have actual costs of 13.

To keep things simple in the toy example, with respect to costs, healthy people are literally all the same whereas people with cancer can be sick in different ways. As shown in the Appendix, such extreme uniformity among the healthy is not required, but it bears repeating that it is crucial to the model that individuals with documented disease conditions (that is, higher risk scores) have greater cost variance than those without any such conditions.

3.2 Capitation payments with and without risk adjustment

For simplicity, we illustrate the shift from no risk adjustment to risk adjustment on one disease category, though the Appendix shows results more generally. Before risk adjustment, capitation payments are set equal to average costs, or $8 \left(\frac{5+6+13}{3} \right)$ in our example. We also assume risk adjustment is “payment-neutral”—that is, if the entire Medicare population

joined MA the average payment is the same with and without risk adjustment. This condition ensures that our result that risk adjustment can increase differential payments is not driven by the government simply paying extra money to firms post-risk-adjustment.

Returning to the toy example, after risk adjustment, the government pays plans the average cost in each risk category—that is, no conditions (type A) and cancer (types B and C). Capitation payments therefore equal 5 for type A individuals and 9.5 for types B and C.

Recall that it is more expensive to enroll individuals with residual costs that are substantially above or below their risk score.¹⁴ We therefore set screening costs higher for types B and C individuals (screening costs equal 2) than for type A individuals (screening costs equal 1). Put differently, it is easier to find healthy people than “healthier-than-average sick people.”

3.3 Predictions regarding selection, differential payments, and profits

We define someone as being positively selected along the “extensive margin” if he has few of the conditions included in the risk-adjustment formula and thus a low risk score. An individual is positively selected along the “intensive margin” if he has low costs along dimensions excluded from the risk score and thus a negative residual. Someone positively selected along the extensive margin will have a low risk score and thus on average low actual costs, whereas someone positively selected along the intensive margin will have low actual costs *conditional on his risk score*. In the context of our example, type A individuals have below average risk scores and hence are positively selected along the extensive margin, whereas type B individuals have low residual costs and are therefore positively selected along the intensive margin.

Prior to risk adjustment, firm profits are 2 from enrolling A, 0 for enrolling B, and -7 for enrolling C. We assume that the firms only enroll profitable individuals, so firms enroll type A individuals and earn profits of 2. Differential payments for individuals who join MA are 3.

Risk adjustment changes the incentives for MA plans. Before risk adjustment, individuals with cancer were unprofitable for firms. However, risk adjustment makes the payment a function of the risk score and thus increases payments for individuals with cancer. Hence, *after risk adjustment, the average risk score of individuals who enroll in MA increases* and thus extensive margin selection decreases. As we show in the Appendix, this result is general and does not depend on the specific values chosen in the example. In our toy example, enrolling type A individuals is no longer profitable post-risk adjustment, while enrolling

¹⁴On the other hand, screening costs are higher for individuals with risk scores that are far from the population mean. In this case, the average risk score is 8, so that types B and C have risk scores that are closer to the population average than type A. In this toy example, we assume that this effect is small.

type B individuals is, and MA enrollees' risk scores increase from five to 9.5.

While firms' incentive to avoid individuals with the health conditions included in the formula decreases after risk adjustment, the incentive to positively select along the intensive margin increases. We therefore predict that after risk adjustment intensive margin selection increases or that *actual costs conditional on the risk score will fall among those enrolling in MA relative to those remaining in FFS*. This result is also general. In our toy example, residual costs fall from 0 to -3.5.

We now turn to differential payments. First, consider the effect of risk adjustment on differential payments had selection remained fixed. In this case, type A individuals would continue to join MA post risk adjustment, and differential payments would fall from 3 to 0. Whenever risk adjustment is "payment-neutral" in the sense defined earlier, *applying the risk-adjustment formula to the pre-risk-adjustment population of MA enrollees would have decreased the total capitation payments the government would have made on their behalf*. This result is also general.

Risk adjustment changes the selection into MA plans, and this endogenous response to the formula can completely "un-do" the fiscal advantages of risk adjustment and actually increase differential payments. In our toy example, for instance, differential payments increase from 3 to 3.5 after risk adjustment. While the toy example shows that the counter-intuitive effect of risk adjustment increasing differential payments is indeed possible, we hasten to add that unlike the other outcomes from the toy example, it is not a general result. For example, had actual costs for types B and C been 7 and 12, respectively, risk adjustment would have decreased differential payments from 3 to 2.5. Hence, *the effect of risk adjustment on differential payments is ambiguous*.

Next, we consider the effect of risk adjustment on plan profits. Because risk adjustment changes the screening costs that MA plans pay in equilibrium, differential payments and plan profits need not move together. In our toy example, profits fall from 2 to 0.5. In the appendix, we show that *as long as beneficiaries were positively selected with respect to their risk score prior to risk adjustment (i.e. they have lower than average risk scores), risk adjustment will decrease MA plan profits*. This result is also general. As we do not have data on MA-specific insurer profits, we cannot directly test this result, but we return to it when we discuss welfare in Section 8.

Finally, we note that risk adjustment yields ambiguous predictions on how the total cost of beneficiaries enrolling in MA should change. On the one hand, after risk adjustment MA enrollees have higher risk scores. On the other hand, their actual costs conditional on their score will fall. In the Appendix, we show that the second effect can dominate and that risk adjustment can cause total costs of beneficiaries to fall. We hasten to add that this result is

not general—in our toy model, for example, total costs increase from 5 to 6—but is indeed possible.

3.4 Discussion

Even in the more formal treatment in the Appendix, we do not specify how plans are able to attract “over-priced” consumers, though in Section 7 empirically explore a few possibilities. Strictly speaking, how they do so is largely irrelevant to the government’s bottom line, which is our focus for much of the paper. Of course, it is highly relevant to producer and consumer surplus, which we explore in Section 8.

Along the same lines, we are also rather silent on plan competition. We believe competition is likely second-order in determining the cost to the government, as MA capitation payments are set by the risk-adjustment formula and not by competitive bidding.¹⁵ Again, however, competition likely affects how producer and consumer surplus changes as a result of risk adjustment. We also briefly explore this topic in Section 8, though emphasize our evidence on these topics is far more suggestive than definitive.¹⁶

4 Data

Our empirical work relies chiefly on individual-level data from the Medicare Current Beneficiary Survey (MCBS) Cost and Use series from 1994 to 2006, though we utilize data from several other sources as well. The MCBS links CMS administrative data to surveys from a nationally representative sample of roughly 11,000 Medicare enrollees each year. It also provides complete claims data from hospital admissions, physician visits, and all other Medicare-covered provider contact for all FFS enrollees in the sample, totaling about half a million claim-level observations annually. The MCBS follows a subsample of respondents for up to three or four years, thus creating a mix of cross-sectional and panel data. During our sample period, the data comprise more than 55,000 unique individuals and over 150,000

¹⁵In 2006 MA plans began to submit “bids” to CMS, with plans bidding below their county benchmark receiving 25 percent of that difference to use to finance extra services for their enrollees. The MCBS does not record plans’ bids and thus we cannot perfectly calculate an individual’s capitation payment even knowing the risk score and benchmark in 2006. MedPAC estimates that this bidding reform reduced capitation payments to MA plans by 3.6 percent in 2006. To be conservative, in the empirical work, we reduce capitation payments by five percent in 2006.

¹⁶Past studies have explored consumer surplus and profits in the MA market, as well as competition, though all are from the pre-risk-adjustment era. Hall (2011) finds that between 1999 and 2002, annual consumer surplus surpassed \$12 billion. Town and Liu (2003) estimate that between 1993 and 2000, the MA program generated over \$18 billion in consumer surplus, and nearly three times that amount in plan profits. Lustig (2010) estimates a structural model, which suggests that in 2002-2003, the welfare loss from imperfect competition in the MA market was greater than that from adverse selection.

person-year observations.¹⁷

Importantly, the MCBS records whether an individual is in an MA plan or FFS each month he is in the sample. As noted in Section 2, MA plans do not submit claims or costs to CMS, so, as the MCBS is based on CMS data, it only contains claims and cost data for those in FFS. Otherwise, all demographic and survey data are recorded for both MA and FFS enrollees. Consistent with past work, we find that, relative to their FFS counterparts, MA enrollees—defined as those in an MA plan the majority of their Medicare-eligible months in a calendar year—are more likely to live in metro areas, are less likely to be on Medicaid or Social Security Disability Insurance and, conditional on not being on Disability, are younger.

Some of the main predictions from the theoretical framework in Section 3 involve enrollees’ risk scores, which are not included in the MCBS. We obtained risk scores from 2004 to 2006 for all MCBS respondents directly from CMS. However, testing the predictions also involves knowing what pre-period individuals’ risk scores *would have been* had the HCC formula been in place, which of course CMS had no reason to document and which we therefore must generate ourselves. We are greatly aided in this task by the fact that an individual’s risk score in year t is based on diagnoses documented on claims from year $t - 1$. As such, using CMS’s algorithm for converting claims data into risk scores, we can simulate the risk score for all MA enrollees the year immediately after they switch from FFS, as in that year their risk scores are based on FFS claims data that we observe from the previous year. Given that we know the actual risk scores from 2004 to 2006, we can check the success of our simulation in these years: the correlation between our simulated risk scores and CMS’s risks scores is over 0.96.

Another key variable in exploring the predictions of the framework is *Total Medicare expenditure*, the total cost to Medicare for individual i in year t , whether it is covering her directly via FFS or paying an MA plan to cover her. We calculate this variable by summing the reported capitation payment each month an individual is in MA and any Part A or B payments incurred over the year. Obviously, for those classified as being in MA, *Total Medicare expenditure* is determined entirely or mostly by capitation payments, and for those in FFS it is determined entirely or mostly by provider payments. Note that as the differential payment results from the model refer to how risk adjustment changes the total cost to the government, the *Total Medicare expenditure* variable refers only to costs incurred by the Medicare program and excludes individuals’ out-of-pocket costs, though we return to

¹⁷From this sample, we make only minimal restrictions. First, we do not include the less than 0.25 percent of enrollees whose Medicare eligibility is based entirely on having end-stage-renal disease, as different risk adjustment and MA-eligibility rules apply to them. Furthermore, we exclude the roughly two percent of person-year observations who join the survey in the middle of the year. The MCBS refers to these individuals as “ghost enrollees” and imputes some of their data.

out-of-pocket costs briefly in Section 7.

While summing capitation payments and Part A and B payments is in principle very simple, another limitation of the MCBS is that, perhaps for confidentiality reasons, capitation payments reported after 2003 do not consistently reflect individual-level variation in HCC scores. However, because we have already calculated individuals' risk scores, generating capitation payments after 2003 is straight-forward—we merely multiply this value by each enrollee's county benchmark, published each year by CMS.¹⁸

The need to calculate HCC scores in the pre-period means we generally limit our analysis to those individuals who were in FFS all twelve months of a baseline year, so that we observe their complete claims history that year. As such, much of the empirical strategy focuses on transitions from FFS to MA. Table 1 shows the number of observations who are in FFS in year t and in MA in year $t + 1$, as well as the number who are in FFS both years (these individuals often serve as a control group), and how these numbers change across our sample period. We observe over 1,500 individuals who switch from FFS to MA, and over 70,000 who remain in FFS over the course of two years.

While relying on individuals who switch from FFS to MA is a limitation, it is not as serious as one might assume. Most individuals in MA have recently been in FFS, so these “switchers” are the norm, not the exception, in the MA population. The large share of switchers in the MA population can be seen in Table 1: 13 percent ($1531 \div (10,040 + 1531)$) were in FFS the *previous year*. Moreover, we estimate that each year, over three-quarters of those joining MA are switching from FFS as opposed to joining MA as a new Medicare enrollee.¹⁹ Regardless, we address further potential biases arising from using the FFS-to-MA flow as a proxy for the MA stock in the next two sections.

5 How did selection patterns into MA change after risk adjustment?

In this section, we empirically test the model's predictions on selection patterns. As shown in the model, it is useful to decompose relative changes in total costs between MA and FFS

¹⁸See the online Data Appendix for further information. In practice, as we show in Appendix Table 3, using the uncorrected MCBS capitation payment has minimal effects on our results. But given the problems we document in the Data Appendix, we think researchers should exercise caution when using this variable.

¹⁹During our sample period, approximately 2.2 percent of FFS recipients switch into MA in the following year. With about 35 million in Medicare FFS during our sample period, roughly 770,000 ($35,000,000 * 0.022$) switch to MA per year. Each year during our study period approximately 2 million individuals become newly eligible for Medicare after turning 65 (ignore the 250,000 to 300,000 65-year olds who were already on Medicare because of Social Security Disability Insurance, who in general are less likely to join MA in any case). We see in the MCBS that about 11.7 percent of them are enrolled in MA in that first year, or roughly 234,000 ($0.117 * 2,000,000$). Thus, among the flow into MA each year, there is greater than a 3-to-1 ratio (770,000 to 234,000) of those who switch to FFS-to-MA versus those who join MA as soon as they become eligible for Medicare at 65.

enrollees after risk adjustment into a change in risk scores (“extensive-margin” selection) and a change in costs conditional on the risk score (“intensive-margin” selection). The model shows that, as plans are no longer incentivized to avoid enrollees with the conditions included in the risk formula, extensive-margin selection should fall, thus leading to an increase in risk scores among MA enrollees. The model also shows that intensive-margin selection should increase as plans now want to find enrollees who are “cheap” relative to their risk score; conditional on risk scores, total baseline Medicare expenditure for those switching to MA relative to those staying in FFS should fall after risk adjustment.

5.1 Quantifying the selection incentives created by the HCC model

While Section 2 gave a general sense for how the HCC risk-adjustment model worked, it is useful to quantify the change in incentives before exploring whether plans reacted to them. Col. (1) of Table 2 presents the average difference between the HCC-based capitation payment and the traditional demographic-based capitation payment using our MCBS data, with this difference broken down by percentiles of the HCC score.²⁰ Mechanically, capitation payments must, on average, rise under the HCC formula for those with higher risk scores, and col. (1) merely presents the magnitudes. For example, the HCC capitation payment would, on average, pay \$2,993 less than the demographic-based capitation payment for individuals with HCC scores in the lowest quartile, but it would pay more than \$29,000 more for the individuals in the top one percent.

Col. (1) would make it seem as though plans would be incentivized to increase risk scores over the entire risk score distribution, but col. (2), which reports actual costs minus the HCC capitation payment, shows that doing so would not always be profitable on average. For example, individuals with the highest one percent of risk scores represent, on average, a nearly \$6,000 loss to an MA plan, consistent with the work cited in Section 2 showing that HCC capitation payments for FFS enrollees with the most serious disease conditions would not fully cover these individuals’ FFS costs. Thus, increasing risk scores indiscriminately would lead plans to enroll some beneficiaries who would be highly unprofitable in expectation. While positive profits might still be possible if these individuals were very positively selected

²⁰We estimate these payments using pre-2004 data on the population of individuals in FFS all twelve months of the previous year, as we need full claims data to calculate risk scores for the following year. We use only the pre-period so that selection in reaction to the HCC model would not have already taken place. While this pre-period population might be selected because MA plans would have enrolled those most profitable under the demographic model, using data from 1994 alone, when MA penetration is under five percent, makes little difference. As described in the next Section, after 2003 MA plans enjoyed higher benchmarks as well as additional payments to ease the transition to risk adjustment, which we remove for the purposes of this table. As such, it reflects the change in incentives from “payment-neutral” risk adjustment as defined in Section 3.

along the intensive margin, these results suggest that plans might be reluctant to draw from the extreme right tail of the risk-score distribution.

Though not shown in the table, we also calculate that the share of individuals who are overpriced (have actual baseline costs less than their risk scores would predict) is 77 percent under both the demographic and the HCC model. Thus, risk adjustment does not actually decrease the number of individuals who are “over-priced”—though it obviously changes their likely characteristics—and indeed we find almost no change in MA market share after the HCC model is introduced.²¹ It is worth noting that extreme over-pricing is more common under the HCC model—more than twice as many individuals are over-priced by more than \$9,000 under the HCC than under the demographic model.

5.2 Empirical strategy

The first element of the decomposition relates to whether plans, in response to risk adjustment, are less likely to avoid individuals with the conditions included in the risk formula. The prediction that such “extensive-margin” selection should fall suggests that we should see a relative increase in MA enrollees’ risk scores after 2003 and thus a positive coefficient on the interaction term in the following regression:

$$Risk\ score_{it} = \beta MA_{it} \times After\ 2003_t + \gamma MA_{it} + \delta_t + \epsilon_{it}, \quad (1)$$

where i indexes the individual, t the year, $Risk\ score_{it}$ is the individual’s HCC score (which, by definition, uses year $t - 1$ claims data to predict Medicare expenditure in year t), MA_{it} the share of her Medicare-eligible months that the individual spends in MA in year t , $After\ 2003$ the post-period indicator, and δ_t a vector of year fixed effects.²² As will often be the case, we estimate this regression on the sample of individuals who are in FFS all twelve months of the baseline year $t - 1$ so that we can use their complete claims data that year to calculate year t risk scores.²³ The model predicts a positive coefficient on the interaction term.

²¹Note that this analysis does not suggest that 77 percent of individuals are potentially *profitable* before and after 2003, as MA might be less efficient than FFS in providing the basic Medicare benefits package. The minimal change in MA market share before and after 2003 is the main reason we do not focus on whether MA or FFS is more efficient from a social-cost perspective. As we discussed in Section 3, the results of the model go through whether MA or FFS is more efficient so long as risk-adjustment does not induce large changes in MA market share.

²²Both equations (1) and (2) are parsimonious in that they do not control for demographic or other characteristics of the beneficiaries. This choice is deliberate, and it reflects the fact that MA plans are paid based on the risk scores of their beneficiaries, not their risk scores conditional on, say, age.

²³While we could potentially use the administrative risk scores provided by CMS for the post-period, throughout this section we use our simulated risk scores in both the pre- and post-periods so that any change in risk scores between the two periods cannot be driven by differences in how they are calculated. Using the actual risk scores in the post-period increases the magnitudes and statistical significance of the coefficients of interest in both the extensive- and intensive-margin analyses. As our simulated risk scores likely

The second element of the decomposition relates to the “intensive-margin” prediction that after risk adjustment, plans will enroll individuals who have low baseline costs conditional on their risk score because they are more positively selected along dimensions excluded from the formula. As such, we predict a negative coefficient on the interaction term of the following regression:

$$Expenditure_{i,t-1} = \beta MA_{it} \times After\ 2003_t + \gamma MA_{it} + \lambda Risk\ score_{it} + \delta_t + \epsilon_{it}, \quad (2)$$

where $Expenditure_{i,t-1}$ is the total FFS expenditure for individual i in year $t - 1$ and all other notation and sampling follows that in equation (1).²⁴

5.3 Results

We begin by exploring how the difference in total baseline costs change after risk adjustment among those switching to MA versus those remaining in FFS, and then decompose this effect into its extensive- and intensive-margin components. Col. (1) of Table 3 shows that before risk adjustment, those switching to MA have baseline costs \$2,847 below those who remain in FFS, consistent with past work cited earlier on the substantial positive selection into MA. Risk adjustment has no apparent effect on this selection (in fact, the point-estimate, -\$173, suggests it may have increased). Thus, while the goal of risk adjustment was to end plans’ incentives to cream-skim low-cost enrollees, we find no evidence that those joining MA after 2003 have higher costs.

The next three columns of Table 3 explore the first component of the decomposition—the change in the risk score. Col. (2) suggests that while individuals switching into MA before risk adjustment had risk scores roughly 0.31 points lower than those remaining in FFS, risk scores of those switching into MA rise after risk adjustment is introduced, making up about a third of the difference.

Based on the results from Table 2 that extreme outliers in the right-tail are still underpriced by the HCC formula, we expect the effect on the mean to be muted, as plans would still be wary of enrolling those with extreme risk scores. Indeed, in col. (3), merely dropping observations with risk scores above the 99th percentile substantially increases the magnitude of the estimate. Estimating a median regression (col. 4) on the entire sample increases the coefficient by nearly a third, so that sixty percent of the pre-period gap in risk scores between

have some measurement error, we attribute this difference to less attenuation bias when the administrative risk scores are used in the post-period.

²⁴This specification is similar in spirit to the “unused observables” test of Finkelstein and Poterba (2006). Using the terminology of their framework, $Expenditure_{i,t-1}$ in equation (2) is an “unused observable” because it is positively related to future costs to the insurer but, conditional on a beneficiary’s risk score, is not used to determine insurance premiums or capitation payments.

FFS and MA is made up by the increase in MA risk scores after risk adjustment. While we generally prefer to use a long pre-period in order to improve precision by increasing the number of individuals switching from FFS to MA, col. (5) shows that excluding observations before 1997 does not change the results. As shown in Appendix Table 3, the results are robust to using a dummy variable for MA status instead of the share of the year an enrollee is in MA.²⁵

Following Bitler *et al.* (2006), to get a clearer picture of the change across the entire distribution we estimate quantile regressions for the first through 99th quantiles, and plot the resulting coefficients in Figure 2. As predicted, the effect is especially strong for higher quantiles (due to the greater variance at higher risk scores, these individuals are very attractive after risk adjustment due to the greater scope for intensive-margin selection). But it falls close to zero right before the 99th quantile, consistent with outliers being substantially underpriced by the formula.

Given that total costs of those switching to MA relative to those remaining in FFS do not change after 2003 while their risk scores rise, then, mechanically, intensive-margin selection must have increased and the remaining columns of Table 3 merely quantify this effect. Col. (6) shows that, relative to the pre-risk-adjustment period, after 2003 individuals switching into MA versus those remaining in FFS have baseline costs over \$1,200 less than their risk scores would predict. As with the extensive-margin results, the coefficients of interest are robust to excluding years before 1997 (col. 7) and using a dummy-variable for MA status (see Appendix Table 3).²⁶

Interestingly, in its 2012 annual report to Congress, MedPAC, citing a working-paper version of our study, found nearly identical levels of intensive-margin selection levels using the universe of Medicare enrollees in 2007 and 2008.²⁷

²⁵We also tested whether using a long pre-period was biasing the results by including a trend estimated using the pre-period data. The coefficient on the variable *Fraction of year in MA* \times *year* is if anything negative, but is indistinguishable from zero ($p=0.886$). Not surprisingly, when it is included in the main regressions, the coefficient of interests increase slightly (from, for example, 0.133 in col. (2) of Table 3 to 0.155). If we break up the post-period into separate years, the coefficients, though noisy, show a monotonic increase in extensive margin selection over the post-period. The results in the table are also robust to excluding those from institutions.

²⁶We also perform the same analysis that was described in footnote 25. The coefficient on the pre-period trend is slightly negative but essentially zero ($p = 0.89$). Including it in the main regression slightly reduces the magnitude of the coefficient of interest (from -1217 in col. (6) of Table 3 to -1052), but it remains statistically significant. Like the extensive-margin results, the magnitude of the effect increases over the post-period years, though there is some slippage in 2005. The results are also robust to dropping those living in institutions.

²⁷See www.medpac.gov/documents/Jun12_EntireReport.pdf, pp. 101. MedPAC finds that “the [MA] joiners [those in FFS in 2007 and MA in 2008] had costs that were 15 percent lower than the stayers [those in FFS both years].” This effect is almost identical to our estimates of intensive-margin selection in the post-period. In the main intensive-margin specification (col. 6 of Table 3), the coefficient on the *MA* main effect is 171.5 and that on *MA* \times *After* is -1217.9, for a total effect of -1046.4 (171.5-1217.9). Average costs

5.4 Discussion and further verification

One of the central drawbacks of our identification strategy is that, because we must calculate risk scores using FFS claims data, we must sample individuals who are in FFS in a baseline year and identify our coefficients off of those who switch the following year from FFS to MA. While we have shown that such switchers are more common than those who enroll at 65, we also provide additional evidence that supports the results of this “switcher” analysis but that uses data on all those in MA or FFS in a given year (regardless of where they were the previous year).

As noted earlier, CMS provided us with actual risk scores for *all* individuals in the MCBS—whether in FFS or MA, “switcher” or “stayer”—from 2004 to 2006. Given the lack of pre-data, we cannot replicate the specification in Table 3, but we can determine whether risk scores among MA enrollees are growing faster than those in FFS during this period, as one might expect both because risk adjustment phases in during this period and because MA plans’ ability to enroll individuals with high risk scores may improve over time. Indeed, in the CMS administrative data, MA risk scores increased by 12 percent over this three-year period, while FFS risk scores increased by one percent, resulting in an MA-FFS difference in average risk scores of -0.13, -0.11, and -0.045 in 2004, 2005 and 2006, respectively. Though far noisier, when we estimate year-by-year post-period effects using our sample in Table 3, we also find a monotonically increasing pattern with much of the gain between 2005 and 2006 (the corresponding differences are -0.18, -0.17, -0.08). As such, at least along this measure, our estimates comparing those who switch to MA to those who remain in FFS closely correspond to comparisons using the *stock* of individuals in MA versus FFS.²⁸

While much of the analysis in this section requires knowing the risk score, the first result—that those switching to MA are just as positively selected on total baseline costs after risk adjustment as before—does not, and as such we can use data that does not include risk scores to verify this unconditional spending result. Table 4 takes up this task, with col. (1) replicating the original result from Table 3. We begin by using mortality as a proxy for health costs and because we have this information for everyone in the MCBS, we no longer

for FFS stayers in the post risk adjustment period are 7435.3. Thus, the MA joiners are on average 14.1 percent (-1046.4/7435.3) cheaper than the FFS stayers in 2004-2006, very close to MedPAC’s estimates. In fact, they find evidence of intensive-margin selection in 68 of the 70 HCC categories, a result that, given the relatively small sample size in the MCBS, we cannot credibly investigate ourselves. In other words, MedPAC finds that, on average, beneficiaries who enroll in MA have significantly lower costs than those on who remain on FFS even controlling for the risk score and looking within specific HCC categories. It is reassuring to know that the effects we estimate through 2006 remain in 2007 and 2008 in a much larger sample.

²⁸Readers may note that the regression estimates in Table 3 there is still a substantial difference in FFS and MA risk-scores, whereas the evidence from CMS suggests that MA risk scores essentially “caught up” to those of FFS by 2006. The faster catch-up among the MA stock than among those switching from FFS is likely due to intensive coding, which does not affect individuals who were in FFS the previous year.

need to rely on switchers. Mortality is a powerful proxy for total costs. As noted earlier, the HCC formula explains 11 percent of the variation in FFS spending; a dummy variable for dying in a calendar year explains by itself 14 percent. The assumption is that if plans are indeed enrolling higher-cost individuals, we should see an increase in relative mortality rates among MA enrollees. Col. (2) of Table 4 shows that there is no relative change in mortality rates among those in MA after 2003—while the point-estimate on the interaction term is positive, it is essentially zero with a p -value of 0.714 despite a sample of nearly 140,000 (19,000 of which are in MA).

We also move beyond the MCBS to verify these results. We gathered over 8,000,000 hospital discharge records from the twelve states between 2000 and 2006 that require hospitals to record FFS/MA status. Because these states are populous, the state-years in our data represent 42 percent of Medicare beneficiaries during this period. In col. (3), the average MA patient had \$1,500 less in charges than his FFS counterpart before 2004—consistent with positive selection in the pre-period—with only \$85 ($p=.835$) of this difference being made up in the post-period—consistent with our results that risk adjustment did not lead MA plans to enroll higher-cost individuals. Interestingly, when we do our best to replicate this regression with the MCBS switcher-analysis by using only Part A charges in 2000 to 2006, we find very similar point-estimates (col. 4).²⁹ Col. (5) shows that the hospital discharge data show very similar (null) results on relative mortality changes among MA patients after 2003.

Finally, we follow Batata (2004) and use county-level data to estimate changes in MA selection. She shows that regressing county-level changes in average FFS costs on the change in county-level MA penetration yields a measure of the difference in costs between the marginal person switching between MA and FFS and the FFS stock. While slightly different than our switcher regressions—which compare the *average* person switching from FFS to MA with the *average* person staying in FFS—one would expect these two selection measures to move in the same direction. As the final column of Table 4 shows, this difference is negative before risk adjustment, reflecting the fact that those on the margin of switching between MA and FFS have lower costs than those in FFS, and, consistent with the switcher analysis and the rest of the table, does not change after risk adjustment.

In summary, evidence from a wide variety of datasets shows that, with respect to overall

²⁹An important difference between the MCBS part A results and the hospital discharge results is that the former is aggregated across the entire year for an individual, whereas the latter is a per-episode measure. However, the share of hospital episodes accounted for by MA enrollees increased at a slightly slower (though essentially equal) rate after 2003 than did the total MA share of Medicare enrollees in these states, meaning that there is neither an increase in the per-visit cost for MA enrollees relative to FFS enrollees after 2003 nor in the average number of hospital visits per year.

average costs, those in MA are just as positively selected after risk adjustment as they were before. This null result is explained by two significant, offsetting changes—an increase in risk scores and a simultaneous decrease in costs conditional on the risk score—both of which were predicted by the model. However, our model cannot predict *a priori* whether these offsetting changes are enough to completely “un-do” risk adjustment with respect to the government’s goal of reducing differential payments.

6 Did risk adjustment decrease differential payments to MA plans?

In this section, we focus on how an individual’s annual *Total Medicare expenditure* changes as he switches from FFS to MA. Recall from Section 4 that *Total Medicare expenditure* is equal to the total cost to Medicare for an individual, whether it is covering her directly via FFS or paying an MA plan to cover her. If risk adjustment works perfectly—so that in expectation capitation payments are equal to an individual’s FFS costs—then whether an enrollee switches between FFS and MA should have no effect on his total Medicare expenditure levels.

Before exploring how differential payments changed after 2003, we begin by testing the model’s prediction that risk adjustment would have “worked”—reduced differential payments—had selection patterns not changed. To test this claim, we focus on enrollees who switched from FFS to MA in the pre-period, when selection with respect to the HCC formula would not yet have taken place. We find that capitation payments—and thus differential payments—would have decreased by \$734 dollars, on average, had the HCC formula and not the demographic formula been used on this sample. Of course, as shown in the previous section, the characteristics of those joining MA change in response to risk adjustment, and the rest of the section explores how differential payments change as a result.

6.1 Empirical strategy

To isolate the effect of the introduction of risk adjustment from other changes occurring around the same time, we make two adjustments to capitation payments after 2003. First, the growth rate of county benchmarks (the baseline value, which, multiplied by the risk score, yields capitation payments) began to rise more rapidly in the later years of our sample period. We thus calculate each county’s benchmark growth rate in the pre-period and then have the county’s benchmarks grow at this slower rate for the post-period as well. Second, in the years immediately following the introduction of risk adjustment, plans received so-called “budget-neutrality” adjustments (about a ten percent increase in the risk-adjusted portion of capitation payments) to ease the transition to risk adjustment, and we mechanically reduce

payments to remove this effect. In both cases, these adjustments increased all capitation payments by a given percent and did not depend on underlying individual conditions or characteristics. The adjustments we make obviously decrease the likelihood we would observe an increase in differential payments after risk adjustment.

We begin, as before, with the sample of beneficiaries in FFS all twelve months of a given year $t - 1$. To estimate the counterfactual Medicare expenditure for an MA joiner in year t had he remained in FFS, we examine the actual Medicare costs in year t for FFS stayers who are similar along observable dimensions. The estimating equation is:

$$Expenditure_{it} = \beta MA_{it} \times After\ 2003_t + \gamma MA_{it} + \lambda X_{it} + \delta_t + f(Expenditure_{i,t-1}) + \epsilon_{it}, \quad (3)$$

where $Expenditure_{it}$ is total Medicare expenditure for person i in year t , $f(Expenditure_{i,t-1})$ is a flexible function of lagged Medicare expenditure, and all other notation follows that in previous equations.³⁰ Note that in the intensive-margin regression we modeled an individual’s Medicare expenditure the year *before* joining MA—hypothesizing that individuals who have low baseline FFS spending conditional on their risk score would be highly attractive to MA plans after risk adjustment—whereas here we model current Medicare expenditure. While lagged Medicare expenditure is highly correlated with current Medicare expenditure and thus serves as an obvious factor on which plans would try to screen, it is the current expenditure that an MA plan must actually cover once someone has joined and thus current expenditure is what matters for estimating differential payments.

This regression model (as well as the ones in the previous section) estimate how characteristics of those switching to MA change after risk adjustment and do not take into account that the quantity of individuals in MA might change as well. As depicted in Appendix Figure 1, the MA share of the Medicare population increases in the post-period (as does the probability of switching from FFS to MA), so our ignoring the quantity margin would understate the effect of any differential payment increases as they would apply to a larger number of MA enrollees in the post-period.

6.2 Results

The first column of Table 5 shows the results from regressing the level of total Medicare spending on the MA variable, which is allowed to have a different effect before and after risk

³⁰We prefer this specification to simply regressing $\Delta Expenditure_{it}$ as the lagged expenditure controls in equation (3) can better account for the fact that medical costs typically exhibit strong regression to the mean, though results using $\Delta Expenditure_{it}$ look very similar and are reported in Appendix Table 3. The lagged Medicare expenditure controls include: lagged Part A and B expenditure and deciles of non-zero Part A payments and non-zero Part B payments as well as indicator variables for zero Part A and B payments (we found that regression to the mean differed depending on the type and level of costs). The results are not sensitive to controlling more coarsely or finely than deciles for lagged Part A and B expenditure.

adjustment, the lagged spending variables, and year fixed effects. Total Medicare expenditure increases by roughly \$905 when an individual switches from FFS to MA (for the entire year) before risk adjustment, and by an additional \$1,733 after risk adjustment.

The second column adds county fixed effects as well as demographic and other basic controls (all listed in the table notes). The coefficient on the interaction term increases to \$2,081. These controls are important if, for example, older people tend to have higher spending growth and post risk adjustment they are also more likely to join MA plans. In this case, we want to account for the fact that these older beneficiaries would have likely experienced high cost growth had they remained in FFS. Col. (3) includes measures of lagged health indicators, which has essentially no effect on the coefficient on the interaction term. Col. (4) includes health indicators from the current year. While self-reported health is not a perfect proxy for current-year health costs, this specification better accounts for potential regression to the mean in health status—if enrollees typically experience a deterioration in their health upon joining MA, then comparing current to previous year’s spending will overstate MA differential payments; however, current-year health status is endogenous to the care individuals receive in MA versus FFS and thus including it may be “over-controlling.” In practice, the two estimates are very similar.³¹

Cols. (5) and (6) subject the estimation in col. (4) to robustness checks. Winsorizing the data based on the 99th percentile in col. (5) or dropping years before 1998 in col. (6) leave the results largely unchanged. Appendix Table 3 shows the results are robust to measuring MA status with a dummy variable and using changes in *MedicareExpenditure* as the dependent variable instead of levels with lagged explanatory variables. Appendix Table 4 replicates all of Table 5, but uses our simulated HCC scores to calculate post-period capitation payments instead of the administrative HCC scores from CMS. The point estimates are almost identical.³²

³¹We explore whether individuals tend to join MA just as their health is about to deteriorate or as their spending is about to rise for other reasons, which would cause us to underestimate the costs MA plans actually face and thus to overestimate overpayments. First, if this effect were important, we should have seen a large decrease in β and γ after current health measures were added in col. (4). Second, individuals are unlikely to postpone expensive treatments until they join an MA plan because plans tend to have less generous cost-sharing for serious procedures than does FFS (see Kaiser’s report on MA benefits, <http://www.kff.org/medicare/upload/8047.pdf>). Third, we actually find no evidence of strategic timing of services for which MA is more generous than FFS, such as vision exams. Finally, we find no evidence of an “Ashenfelter dip” the year before a switch to MA—controlling for two years of lagged cost data instead of one has minimal effect on the point-estimates, though standard errors increase due to the smaller sample.

³²As with the selection results, we explore whether there are any pre-trends in the data. The coefficient on the MA trend is close to zero ($p = 0.7$). Controlling for the trend reduces the point-estimate on the col. (4) specification to \$1482, though it remains statistically significant. As with the selection results, the year-by-year point estimates in the post-period are noisy, but exhibit a positive trend (though, like the intensive-margin results, the trend is not monotonic, with slippage in 2005). The results are also robust to excluding individuals living in institutions.

Thus, our estimates imply that prior to 2004, spending was approximately \$1,000 higher for those in MA relative to what it would have been in FFS. After the move to risk adjustment, this payment difference increases approximately to \$3,000.

6.3 Discussion and aggregate spending calculations

Given that our identification relies on those switching to MA from FFS, a natural question is whether differential payments for this population reflect differential payments for the average MA enrollee. On the one hand, the differential payments we estimate via “switchers” understate differential payments among the entire MA stock because “intensive coding” on the part of MA plans does not affect switchers their first year in MA because risk scores are still based on FFS claims. After years of intensive coding on the part of MA plans, the difference between capitation payments and counterfactual costs in FFS should fan out further, not contract, suggesting our results would be under-estimates.³³

On the other hand, those switching into MA after 2003—off of whom the coefficients in Table 5 are identified—are presumably more highly selected with respect to the HCC formula than are those, say, still enrolled in MA in 2006 but who first left FFS in 2002, before the introduction of the HCC model. As such, average differential payments per MA enrollee after 2003 are likely a blended average of the results we find in Table 5 and the differential payment among MA enrollees who first switched in the pre-period but who remain in an MA plan in the post-period, who in fact still account for 68 percent of MA enrollees in 2006. To be conservative, we estimate a lower bound on total differential payments in 2006 by assuming that those enrollees who had switched to MA before 2004 are not selected at all with respect to the HCC formula, thus ignoring the effects of intensive coding and the fact that individuals who are “expensive” for their risk score should be the most likely to *exit* MA for FFS after 2003 (as a recent MedPAC study has in fact found).³⁴ As shown at the beginning of this section, we estimate that the HCC formula would have decreased differential payments by \$734 had selection patterns not changed, so we thus assume that differential payments decrease by this amount for each MA enrollee in 2006 who had joined before 2004.

As such, without the additional overpayments associated with the higher benchmarks and the budget neutrality payments, overpayments per MA enrollee in 2006 appear to be $.68 * -734 + .32 * 2,012 = \145 greater after risk adjustment. We would of course expect that this average effect grows over time, as the share of MA enrollees joining the program

³³See the 2010 CMS Advance Notice to MA Plans, www.cms.gov/MedicareAdvtgSpecRateStats/Downloads/Advance2010.pdf.

³⁴See www.medpac.gov/documents/Jun12_EntireReport.pdf, pp. 100-101.

after 2003 become the dominant share of the MA stock and in steady-state would approach the coefficients found in Table 5. Adding back in the additional overpayments due to higher benchmarks and budget neutrality payments (which apply to all MA enrollees, regardless of when they switched), yields a per-MA-enrollee increase in overpayments of $\$145 + \$1166 = \$1311$.³⁵ Considering overpayments were already substantial in the pre-period (our estimate is $\$954$ from Table 5 col. 4), we estimate total overpayments per MA enrollee in 2006 was $\$2265$. These results suggest a lower bound on total differential payments in 2006, when MA enrollment was 7.6 million, of $\$17.2$ billion, or 4.4 percent of *total* Medicare spending that year.³⁶

7 How does selection into MA plans take place?

In this section we explore potential mechanisms underlying the selection results. We begin by exploring the more general question of why low-cost individuals tend to be enrolled in MA plans in both the pre- and post-periods, and then focus on how plans might have been able to attract enrollees who are “cheap for their risk score” after 2003.

7.1 Why are low-cost individuals more likely to be in MA plans?

The evidence in Section 5 shows that MA plans enrolled substantially lower-cost individuals before and after risk adjustment. But how do such patterns emerge when plans must offer the same plans at the same rate to all Medicare beneficiaries in their geographical area of operation? While one can imagine several different possibilities, here we explore whether once individuals enroll in MA, the healthy ones are differentially more satisfied and less likely to return to FFS. This pattern might arise because plans actively treat healthy enrollees better than sick ones so as to differentially retain the former group, or simply because sick individuals do not like the HMO model of care.³⁷ Through reputation effects, such a result could feed back into enrollment patterns as well.

³⁵The $\$1,166$ figure is from subtracting the point-estimates on the interaction term in the final column in Table 5 from that in col. (4), as the only difference between these two models is that the final column includes budget neutrality payments and accelerated post-2003 benchmark growth.

³⁶According to CBO, Medicare spending totaled $\$392$ billion in 2006 (we adjusted to 2007 dollars, the units of the coefficients, using CPI-U).

³⁷Though we do not have the data to investigate them, other mechanisms have been suggested by past work. Targeted advertising may play a role: A report by the Kaiser Family Foundation concludes that advertisements for MA plans target healthy people (see <http://www.kff.org/medicare/upload/7805.pdf>). Health plans may delay responding to individuals they suspect are high-cost: Bauhoff (2010) finds evidence via an audit study that highly regulated German health insurance plans respond more quickly to enrollment requests from respondents residing in low-cost areas of the country, and MA plans likely have far greater flexibility than do German insurers.

The MCBS asks respondents to rate their satisfaction with their overall health care “last year” as well as specific aspects of it. As the question is asked in the fall, it is difficult to know whether individuals are answering based on their experience so far in the calendar year or the previous calendar year as well. As such, for this section we generally sample those who did not switch (either from FFS to MA or from MA to FFS) the previous year. Thus, unlike the “switcher” analysis so far in the paper, identification comes from comparing individuals who have been in MA for at least two years with individuals who have been in FFS for at least two years. Asking someone who, say, just switched from FFS to MA to rate their health care “last year” would likely shed little light on their experience so far in MA.

As we no longer rely on switchers, we do not have actual cost or claims data for those in MA.³⁸ We thus use self-reported health as an admittedly imperfect proxy and explore whether good health predicts satisfaction with one’s health care in MA more than it does in FFS:

$$Satisfaction_{it} = \beta MA_{it} \times Health_{it} + \gamma MA_{it} + \mathbf{H}_{it} + \lambda X_i + \delta_t + \epsilon_{it}, \quad (4)$$

where *Satisfaction* measures individuals’ reported satisfaction with different aspects of their health care and varies from one (very dissatisfied) to four (very satisfied), *Health* is the five-category self-reported health variable described earlier, \mathbf{H} are its corresponding fixed effects, and all other notation follows that used in previous equations. The health fixed effects account for the fact that in both MA and FFS, poor health correlates with negative feelings toward one’s health care, and thus the interaction term captures how much more or less sensitive enrollee satisfaction is to underlying health in MA versus FFS. We control for demographic characteristics in X because different groups may assess their health and health care differently.³⁹ If MA plans indeed treat their healthier enrollees better, then we would expect estimates of β to be positive.

Table 6 displays the results from estimating equation (4) via OLS.⁴⁰ We demean the *Health* variable in $MA \times Health$, so that the *MA* main effect represents the effect of MA enrollment for someone with mean self-reported health. The first row reports results when overall satisfaction serves as the dependent variable. The MA main effect is negative—suggesting that someone of average health reports lower satisfaction in MA than in FFS.

³⁸As such, we cannot test whether the specific extensive- and intensive-margin selection patterns also arise with respect to satisfaction. We would have liked, for example, to see whether MA plans treat individuals with low costs relative to their risk scores better after risk adjustment, but such detail is impossible given data limitations.

³⁹For example, Hispanics rate their health lower than non-Hispanics, even though throughout our sample period they have lower health costs.

⁴⁰Note that ordered logit gives exactly the same patterns of coefficient signs and significance levels, and results are available upon request. The sample size variation across different regressions arises from variation in the number of individuals who report not having enough experience to make a satisfaction rating as well as some questions not being asked in the earlier years.

While this result is surprising given that one might have assumed the large overpayments to MA plans would be in part passed on to enrollees and reflected in their satisfaction ratings, it is also possible that the type of person who joins MA might simply be harder to please.

We instead focus on the interaction term, which is positive and significant, indicating that good health predicts satisfaction with MA plans more than it does satisfaction with FFS. In fact, only among those who report being in “excellent” health do MA plans receive a higher rating than FFS (not shown). Moreover, relative to FFS enrollees, MA enrollees exhibit a more positive gradient of satisfaction with respect to health in all nine categories surveyed by the MCBS, and for a majority of categories this difference is statistically significant.

The last row of Table 6 investigates whether sicker MA enrollees “vote with their feet” and exit at higher rates than do sicker enrollees in FFS. Instead of satisfaction ratings, we regress whether an individual changes his coverage status—to MA if he is currently in FFS, to FFS if he is in MA—on the same set of explanatory variables. Indeed, the same pattern emerges—not only are MA enrollees less likely to retain their current coverage status in general, but this difference is especially pronounced for those in self-reported poor health.⁴¹

7.2 Possible mechanisms underlying the changes after risk adjustment

The results in Table 6 provide a potential explanation for why, in general, higher-cost enrollees wind up more often in FFS, but they do not explain how individuals with low costs relative to their risk score differentially found their way into MA plans after risk adjustment. While we cannot provide a definite answer, we believe several factors are at work.

First, plans have a wealth of data, both on their own MA enrollees and from their operations in the non-Medicare market. In fact, because plans (unlike CMS) have data on *the medical claims and costs of MA beneficiaries*, along this dimension plans have more information than the risk adjuster.

Second, CMS does not adjust for factors such as race, ethnicity and income, which are not only related to health costs but, through targeted advertising, are also relatively easy for MA plans to select on. Given that the variance in costs grows with the risk score, demographic differences that are small on average or in a relatively healthy population could be very large for groups with high risk scores or for a specific disease category. Plans certainly have the data to determine that, for example, Hispanics with heart disease are \$4,000 cheaper than

⁴¹One might assume that because those exiting MA do not appear particularly healthy relative to the FFS stock, our selection results are over-estimated and thus our differential payment results may be over-stated as well. However, differential payments are a function not only of selection but also of capitation payments, and CMS has documented that risk scores grow faster in the MA stock than in the FFS stock due to intensive coding. So, at the point when they return to FFS, MA enrollees’ capitation payments would have grown faster than their actual costs, and thus differential payments are in fact under-estimated by considering only those switching from FFS to MA. See Section 6 for further detail.

their risk score would suggest and then advertise their excellent network of cardiologists on Spanish radio stations. If demographic or other observable factors explain how costs vary from the risk score’s prediction within a disease category, then plans may have the ability to differentially enroll people who are “cheap for their risk score.”

We explore empirically whether after risk adjustment MA plans appear to target individuals with disease categories that have the most predictable variance. All else equal, plans will be able to screen within a risk cell if (a) the variance of costs within this cell is large and (b) the variance is predictable with demographic or other observable information. Due to sample-size restrictions, we focus on the ten most common disease categories. Appendix Figure 2a plots on the x -axis the variance of residual costs (actual costs minus predicted costs under the HCC model), for each of the ten most common HCC conditions.⁴² There appears to be a weak, positive relationship between residual variance and the relative increase in the probability of switching to MA, plotted on the y -axis.

As noted above, this variance is not attractive to an MA plan unless it is at least somewhat predictable. We proxy “predictable variance” with the product of (1) the variance of the residual (plotted in Appendix Figure 2a) and (2) the R^2 when we regress this residual on basic covariates (see figure notes for details). As seen in Appendix Figure 2b, the relationship between an HCC category’s predictable variance and the tendency of people in that category to switch to MA after risk adjustment is strongly positive, and statistically significant despite the small sample size.

We are more limited in testing whether specific demographic groups predicted to be cheap for their disease category differentially join MA after risk adjustment—cutting the sample of individuals switching from FFS to MA by disease category and then by race or ethnicity quickly runs into power constraints.⁴³ Fortunately in terms of data analysis, one of the most common HCC categories is “Breast, Prostate, and Colorectal Cancer” (HCC10), which, unlike the other major categories, combines several diseases; given the prevalence rates of these diseases, the large majority of women in this category will have breast cancer while the large majority of men will have prostate cancer. Moreover, past medical research (Yabroff *et al.*, 2008) has shown the annual cost of breast cancer treatment to be roughly \$1,900 cheaper than that of prostate cancer even after accounting for demographic differences between the

⁴²As usual, we can only do this on the FFS population; we further restrict the sample to FFS beneficiaries before 2003, so as to examine cost variance before any selection with respect to the HCC formula would have taken place.

⁴³For example, among Hispanics in FFS, the share with a history of congestive heart failure is the same before and after risk adjustment, whereas that share goes from zero to thirteen percent among Hispanics in MA. And even though the corresponding trends for non-Hispanic MA enrollees are slightly negative, the relative increase in Hispanics MA enrollees with a history of congestive heart failure is not statistically significant as the absolute increase is only four individuals.

two groups.⁴⁴ Therefore, we have a common disease group in which we can identify a large subgroup of individuals (women) who were underpriced before risk adjustment and are overpriced after. While all individuals in Category 10 were less likely than other Medicare beneficiaries to join MA in the pre-period, women—but not men—in this category are more likely to join after risk adjustment ($p < .02$). That the one *Demographic* \times *Disease* category we can reasonably test behaves as expected gives some assurance that plans would target other attractive subgroups within disease categories similarly.

The results in this section begin to shed light on how MA plans might actually risk-select, which our framework and even the more detailed model in the Appendix treated in a very reduced-form manner. As noted in the framework, how plans actually risk-select has little direct effect on the government’s differential payments, but could have large effects on producer and consumer surplus, topics we explore in the next section.

8 Welfare and Policy Implications

In this section we explore whether the increased overpayments after 2003 were passed on to other parties. For example, one might expect risk adjustment to increase competition for certain “cheap-for-their-risk score” Medicare recipients, and thus plans might try to woo them with higher-quality care or lower out-of-pocket costs. Alternatively, if the health insurance industry is not perfectly competitive (as, e.g., Dafny, 2010 finds), some of the overpayments could have increased plan profits. Finally, the increase in spending could have led to increased marketing or administrative expenses or the entry or expansion of less efficient insurers. Which of the above scenarios dominate is obviously central to understanding the welfare effects of risk adjustment. While the evidence we bring to bear on these questions is not as definitive as our analysis of government expenditure, it begins to shed some light on the policy’s welfare implications.⁴⁵

8.1 The effect of risk adjustment on plan profits

We do not have access to actual MA-specific profit data from insurers, a limitation that appears to be shared by all papers in the MA literature. However, our model does speak

⁴⁴In our data, the difference is \$1,200, though given the category also includes colon cancer we would expect some attenuation bias.

⁴⁵While consumer surplus, MA plans profits, and government payments likely account for the major welfare effects of risk adjustment, indirect effects on other markets are surely possible. For example, we do not examine effects on providers. While the effects are likely small given that, regardless of risk adjustment, they serve the same fixed population of Medicare enrollees, if risk adjustment has effects on aggregate MA penetration their reimbursement rates may well change as MA plans and FFS do not always reimburse providers in the same manner. We leave the exploration of any such effects to future work.

directly to plan profits—after risk adjustment, profits always fall (even if overpayments rise) because screening costs increase.⁴⁶ Note also this prediction refers to “payment-neutral” risk adjustment—which, recall, was our term for a risk-adjustment model in which, essentially, the government is not trying to systematically under- or overpay plans on average.

While we do not have actual profit data to verify the model’s prediction, evidence from plans’ reactions to risk adjustment suggest they believed payment-neutral risk adjustment would hurt their profits. The implementation of the mild precursor to the HCC model (which, recall, explained 1.5 percent of cost variation) was explicitly payment-neutral. Many MA plans formally called on CMS—which tacitly agreed the reform would hurt plans—to delay the implementation of risk adjustment or to provide extra “budget-neutrality” payments to compensate for risk adjustment.⁴⁷ Indeed, perhaps to avoid a similar backlash, as we explain in Section 2, CMS made the administrative decision to increase the risk-adjusted capitation payments by roughly ten percent (recall that we remove these extra payments in most of the Table 5 specifications). Thus, when risk adjustment is not augmented with additional payments to keep total payments constant, both CMS and the MA plans themselves expect profits to fall, a conclusion completely consistent with our model.

In our model, risk adjustment can increase government spending while simultaneously decreasing plan profits because it increases the money plans must spend attracting and retaining their beneficiaries. For example, the assumption of the model is that devising strategies that allow patients to find the cheapest diabetics (an optimal strategy after risk adjustment) is more costly than just avoiding diabetics all together (an optimal strategy before risk adjustment). This extra money could be spent in ways that increase consumer welfare, like improving the quality of medical care, or in ways that likely do not, such as engaging in targeted advertising or devising complicated screening strategies. Ultimately, the extent to which this extra government spending is passed on to consumers is an empirical question.

⁴⁶These screening costs also help explain an initially puzzling finding: that, as we showed in Section 5, after risk adjustment, in anything, individuals who joined MA had somewhat *lower* medical costs (though this effect is very close to zero and nowhere close to statistical significance). The only way to rationalize why MA plans did not enroll these individuals before risk adjustment is to note that these individuals must have been less *profitable* for MA plans prior to risk adjustment due to these additional screening costs.

⁴⁷See <http://www.cms.gov/MedicareAdvtgSpecRateStats/Downloads/Announcement2000.pdf> This same document shows that the PIP-DCG model was meant to be explicitly payment neutral: “One commenter objected to the decision to implement risk adjustment in a manner that produces savings in Medicare payments. The commenter requested that HCFA [the Health Care Financing Administration, the pre-cursor to CMS] consider implementing risk adjustment in a budget neutral manner. While budget neutrality has been mandated by law for other payment system changes, there is no provision in the law requiring that risk adjustment be budget neutral. *The purpose of implementing risk adjustment is to correct for historical payment errors caused by biased selection* [emphasis added].”

8.2 The effect of risk adjustment on consumer welfare

In this subsection, we explore whether risk adjustment improved consumer welfare along a variety of observable dimensions. To conserve space we only briefly describe the data and estimating equations and refer readers to the Appendix table notes for greater detail. It is worth emphasizing here that, through the accelerated growth of county benchmarks and “budget neutrality” payments, MA plans received increases in capitation payments in the post-period beyond those we attribute to their endogenous reaction to risk-adjustment (in Table 5, compare cols. 4, which does not include accelerated benchmark growth and budget-neutrality payments, and col. 7, which does). As such, if some of these overpayments were passed on to MA enrollees after 2003, they would almost certainly serve as an upper bound for the gains they would have enjoyed under risk adjustment absent these additional payments.

8.2.1 Comparing MA v. FFS before and after risk adjustment

Some MA plans cover a portion of their enrollees’ Medicare Part B premiums and we begin by exploring whether this benefit expands after risk adjustment. We use data on MA Part B premium reductions compiled by Mathematica and published by the American Association of Retired Persons.⁴⁸ As shown in Appendix Table 5, premium reductions increase in 2004, and overall we estimate that the average MA enrollee received a \$353 increase in annual premium reductions after risk adjustment, a small share of the increase in over-payments that we estimate in Section 6 but certainly a meaningful benefit to enrollees.

By contrast, we find no other systematic evidence that the overpayment increases after 2003 coincided with enhanced benefits for MA enrollees.⁴⁹ We begin in Appendix Table 6 by examining whether MA enrollees report higher satisfaction rates after risk adjustment, regressing each of our satisfaction measures on $MA \times After$ 2003, its lower-order terms and the controls in Table 6. The coefficient on $MA \times After$ 2003 is negative for five and positive for four, though rarely statistically significant. An important exception is that MA enrollees are significantly *less* satisfied with cost-sharing relative FFS enrollees after 2003. This result could suggest that plans, now targeting a population with a greater number of disease conditions, must engage in more stringent cost-sharing to screen out high-cost

⁴⁸We have not found government data on this feature of MA plans—MedPAC merely documents what share of plans offer premium reductions, but does not weight by the share of enrollees actually taking up these plans. See http://assets.aarp.org/rgcenter/health/2006_23_medicare.pdf as well as notes to Appendix Table 5 for more information on the data.

⁴⁹Whereas it was relatively easy to remove the additional payments to MA plans after 2003 due to the budget-neutrality payments and high benchmark growth in order to estimate the effect of payment-neutral risk adjustment on overpayments, there is no obvious way to adjust downward MA enrollees’ satisfaction to simulate the effect of a payment-neutral policy. As such, these results likely overstate MA enrollees’ satisfaction post-risk-adjustment relative to the counterfactual of payment-neutral risk adjustment.

populations. It also puts the premium-reduction in Appendix Table 5 in a different light—instead of a net benefit to MA enrollees, the premium reductions might merely reflect lower actuarial value. Nor do MA enrollees in poor health report greater satisfaction after risk adjustment, as one would have expected if plans no longer have an incentive to have sick people exit. In Appendix Table 7, the coefficient on $MA \times Health \times After$ 2003 is in fact positive for seven of nine satisfaction categories, suggesting that, if anything, good health is a stronger predictor of satisfaction with MA relative to FFS after 2003. While our satisfaction measures are only a rough proxy for welfare, it is hard to conclude from these results that a large share of the overpayment increases were passed on to MA enrollees after 2003.

An interesting exception is that Hispanics in MA are more satisfied after 2003, consistent with our evidence that Hispanics are over-priced by the HCC model and would thus be highly attractive to MA plans.⁵⁰ As reported in Appendix Table 8, of the nine satisfaction categories, $Hispanic \times MA \times After$ is positive for all but one ($p=0.9$) and, despite their being a relatively small subgroup, statistically significant and positive for three and positive with p -values under 0.15 for another two.⁵¹ Thus, the evidence suggests that while MA plans may have improved service for specific groups especially attractive post-risk-adjustment, overall there is no net improvement in satisfaction among MA enrollees after 2003.⁵²

8.2.2 Did risk adjustment improve the Medicare program more broadly?

To fully assess risk adjustment we need to consider its affects on the entire Medicare program. Suppose, for example, that risk adjustment allowed for better sorting of beneficiaries between MA and FFS, and thus everyone was better off. The above analysis would obscure such an effect by not considering the benefits to FFS enrollees as well. Moreover, because the types of people switching from FFS to MA change after 2003 (they have higher risk scores, e.g.), the analysis above—comparing MA versus FFS enrollees after 2003—could be contaminated by compositional changes. We thus perform a number of analyses not subject to these critiques.

As envisioned by policy-makers, risk adjustment would make plans indifferent between enrolling sick and healthy individuals and would thus expand the health care options for those in poor health, whom MA plans would have previously avoided. As such, one would imagine

⁵⁰As noted above, Hispanics were roughly \$1,200 cheaper than their non-risk-adjusted capitation payments. After risk adjustment, Hispanics were roughly \$1,550 cheaper than their capitation payment.

⁵¹These results cannot be explained by, say, MA plans merely trying to recruit in more urban areas with a greater number of minorities. African-Americans, a group that is more expensive after risk adjustment, exhibit no such patterns: among the nine satisfaction categories, $Black \times MA \times After$ is just as often negative as positive, though never significant. Results available upon request.

⁵²Recall that there is some ambiguity in whether respondents are referring the current or previous year in evaluating their health care. Results (available upon request) in this and the next subsection are very similar if we instead use 2005 as the first post-period year.

that those in poor health—whether they enroll in FFS or MA—would be more satisfied with their health care options post-risk adjustment. We thus estimate our various satisfaction measures as a function of $Health \times After$ 2003, its lower-order terms, and the controls included in Table 6. As shown in Appendix Table 9, the coefficient on the interaction term is positive for all nine categories and significant for eight, meaning that after risk adjustment those in poor health are relatively *less* satisfied with their health care.⁵³

This result is central to the welfare effects of risk adjustment because the gradient of satisfaction with respect to health status speaks to the insurance value of Medicare. Put differently, a system in which only healthy people receive high-quality health care would not seem to allocate resources from the “good” to the “bad” state, as consumption-smoothing requires. The positive coefficients on $Health \times After$ suggest this allocation of resources between the good and bad states was potentially made worse after risk adjustment.

Next, we exploit variation in county-level MA penetration prior to risk adjustment. All else equal, in counties with greater MA penetration prior to risk adjustment, risk adjustment represents a larger intervention to the Medicare program. If risk adjustment improves the health care of Medicare enrollees generally, then average satisfaction for all beneficiaries (MA and FFS) should improve after risk adjustment in counties with higher pre-period MA penetration, relative to other counties. The results reported in Appendix Table 10 provide no evidence that risk adjustment improved overall satisfaction in MA-intensive counties.⁵⁴

Next, we turn to two additional data sources to explore whether measures of Medicare enrollees’ health care quality improved after risk adjustment relative to similar individuals outside the program. Using National Health Interview Survey data from 2000 to 2006, we follow past work and compare the “young elderly” (65-74 year-olds) to the “near elderly” (55-64 year-olds). We examine all variables in the dataset plausibly related to patient satisfaction and preventive care. Appendix Table 11 shows that none of these measures improve for the young- relative to the near-elderly, as one would have expected if risk adjustment improved

⁵³Recall that for any of the satisfaction regressions that include MA status, we have limited the sample to those who had not switched the previous year (because the satisfaction questions could refer to the current year or previous year, and asking someone who just switched to MA about her satisfaction “last year would tell us little about the performance of the MA program). In these regressions, however, we compare the overall satisfaction of individuals based on their health status regardless of their MA status, so we include all beneficiaries, including those that switched MA status. If we instead limit the sample to non-switchers, the results are similar—seven of the nine outcomes yield a positive coefficient on $Health \times After$, though, largely because the sample is smaller, only four are significant (neither of the negative coefficients are close to significant.) If we estimate the $Health \times After$ specification separately for those currently on MA versus FFS, the point-estimates for the MA sample are “more positive,” consistent with the results in Appendix Table 6.

⁵⁴Similarly, using a specification in the spirit of Batata (2004), we find no evidence that changes in MA county penetration (as opposed to pre-period levels) are more positively associated with changes in county-level satisfaction measures among all Medicare enrollees after risk adjustment than before.

care delivery generally for Medicare enrollees.⁵⁵

Combining the logic in the previous two sets of analyses, we ask whether MA-intensive counties saw mortality improvements for the young- relative to the near-elderly after 2003. To investigate this hypothesis, we obtained county-level data on mortality by age from the National Center for Health Statistics for 2000-2006.⁵⁶ As Appendix Table 12 shows, we find no evidence of differential mortality improvements among the young elderly in counties most affected by risk adjustment—the coefficients in cols. (1) and (2) vary in sign but are essentially zero. Nor do changes in MA penetration after risk adjustment predict greater mortality gains for the young-elderly relative to the near-elderly than they did before risk adjustment (col. 3 and 4). In the final two columns, we examine “long differences” to allow the effect a few years to materialize, and if anything, after 2003, increases in MA penetration are associated with a small increase in mortality for the young- relative to the near-elderly, though these results are not quite statistically significant. Though not shown, we also find no differential improvements for the young- versus near-elderly after 2003 when we aggregate across all counties.

In summary, the results from this and the previous subsection suggest a modest financial transfer to those in the MA program in the form of reduced part B premiums, but, despite the large increases in overpayments, no other detectable improvements in the care of those in MA versus FFS, nor improvements in the care of Medicare enrollees more generally after risk adjustment. A conservative calculation of the welfare costs of increased overpayments might assume that they are transfers to some party—even if we cannot identify any outside of those receiving premium reductions—and that the welfare cost is merely the dead-weight loss associated with raising the additional government revenue. Taking the \$1311 increase in overpayments per MA enrollee calculated in Section 6, the standard assumption of thirty cents of deadweight loss per dollar of tax revenue, and an MA share of the Medicare population in 2006 in our data, we estimate an annual welfare cost of \$71.66 per Medicare enrollee. If instead we assume that the entire overpayment minus the premium reduction were, say, dissipated via administrative costs or, as the model in fact suggests, socially wasteful screening costs, then the per-enrollee cost is \$246.76.⁵⁷

⁵⁵In unreported results that parallel the empirical approach in the previous subsection, we also examine whether, among NHIS Medicare enrollees, those in an MA plan report improvements in these quality measures after risk adjustment. The only significant effect is a relative decrease after 2003 in the likelihood those in MA receive flu shots, suggesting, if anything, that preventive care may have deteriorated in MA plans.

⁵⁶Previous work has examined the effect of MA more generally on mortality. For example, Gowrisankaran *et al.* (2011) examines the effect of plausibly exogenous increases in county level MA penetration rates on mortality during the 1993 to 2000 period. They find that, relative to FFS, increases in MA enrollment without drug coverage *increases* mortality, while increases in enrollment in an HMO with drug coverage has no statistically significant effect on mortality.

⁵⁷In Section 6 we calculated a per-MA-enrollee increase in overpayments of \$1311. Assuming the only cost

8.3 Improving risk adjustment

Our results suggest at least two concerns about future attempts to improve risk adjustment. The first relates to the methodology for recalibration of the risk-adjustment formula, which, essentially, just reruns the cost regressions on the HCC dummies using more recent FFS data. Our results in Table 3 on intensive-margin selection indicate that, after risk adjustment, the lowest-cost cases of each disease category join MA and are thus no longer in the FFS data used in the recalibration. As such, using more recent FFS data will create even greater positive bias for a disease category, as the FFS enrollees with the disease are more adversely selected. Put differently, if a disease condition was over-priced initially, recalibration will tend to *exacerbate* the mis-pricing.

The second note of caution is a more general reminder that we believe bears further emphasis: the R^2 of a risk-adjustment formula is not a sufficient statistic for its welfare effects (or even its effect on government payments). As noted, the model shows that the effect of an increase in R^2 on differential payments is indeed ambiguous. The risk adjustor thus needs to consider not only the predictiveness of the formula, but the incentives it creates.⁵⁸

On a more positive note, our results suggest some potential improvements. First, our framework can help predict which categories of the formula might be especially prone to screening. The results in Section 7 suggest that disease categories that have greater predictable variance will be especially problematic. The risk adjustor may wish to pay special attention to risk selection in these categories and should consider, for example, conditioning on the variables that empirically predicted an individual’s cost in Appendix Figure 2b. In particular, our results suggest that including interactions between demographic groups and certain diseases could be beneficial.

Alternatively, instead of trying to improve the predictiveness of the formula using FFS data, CMS may wish to incorporate the information embodied in MA enrollment patterns into the formula. If a certain disease group begins to “differentially disappear” from FFS and “re-appear” in MA after risk adjustment, then it is a signal that low-cost individuals in that disease group are, for whatever reason, easy to “skim,” and thus CMS could reduce the capitation payments for that category. Note that such an approach would have the exact opposite effect of the current recalibration procedure, which leads to payment increases for disease groups that experience differential migration to MA. Our model and empirical results suggests that MA enrollment patterns signal which groups are over-priced by the formula,

is the DWL of raising this revenue, then the welfare cost per Medicare enrollee is $\$1311 * .3 * 0.1822 = \71.66 . If the entire overpayment increase less the Part B premium reductions is considered waste, then the per-enrollee welfare cost is $\$71.66 + (\$1311 - \$350) * 0.1822 = \246.76 .

⁵⁸The conclusion that predictiveness is not a sufficient measure for welfare is the same reached by Glazer and McGuire (2000), but in a differential institutional setting and for different reasons.

and a recalibration that tries to maximize all informational value would have this information feed back into future calculations of the HCC disease weights.

9 Conclusion

Our analysis began with a simple framework for understanding how selection patterns would respond to an attempt to decrease differential payments to MA plans via risk adjustment. We predicted that MA plans would no longer try to avoid enrollees with the conditions included in the formula (“extensive-margin” selection falls), but they would increase efforts to enroll those with low costs conditional on the risk score (“intensive-margin” selection rises). Using individual-level data on Medicare expenditures and comparing the selection patterns for those switching to MA with those remaining in FFS, we confirmed both predictions. Indeed, despite the increases in risk scores, we find no evidence that actual costs for those joining MA increased relative to FFS after 2003. The framework also shows that because the variance of medical costs increases with the risk score, risk adjustment can potentially increase the scope for selecting individuals with costs below their capitation payment, and we indeed find that differential payments to MA plans actually increase after risk adjustment. We find little evidence that the large increases in government spending were offset by comparable gains in producer or consumer surplus.

The MA program and Medicare more generally have recently been the target of regulation and reform, and as data become available, future work might examine how MA selection and differential payments change in response. Given that MA plans often provided drug coverage even in the 1990s and early 2000s, the extension of drug benefits to the entire Medicare population in 2006 may change selection into the MA population. The Affordable Care Act more directly affects the Medicare Advantage program and aims to reduce differential payments by lowering many county benchmarks while at the same time linking capitation payments to measures of plan quality.⁵⁹

The ACA also requires risk adjustment in the state-run insurance exchanges, and future work might examine how differences in the institutional design between the exchanges and MA affect the success of risk adjustment. In particular, Medicare directly insures the FFS population and thus has access to their claims and cost data to calibrate a risk adjustment model; no such “public option” exists in the exchanges, and thus risk adjustors may lack sufficient data to estimate a model. The lack of a public option is important for another

⁵⁹However, the extent to which MA reimbursements will actually fall remains an open question. The Department of Health and Human Services recently decided to award quality bonuses to MA plans covering the vast majority of MA beneficiaries, meaning that, by 2015, an additional \$6.7 billion will flow to MA plans (see http://www.boston.com/business/healthcare/articles/2011/04/19/obama_administration_eases_pain_of_medicare_cuts/).

reason. Recall from Section 7 that both before and after risk adjustment, MA enrollees in poor health express greater dissatisfaction with their care than do their counterparts in FFS, and differentially migrate back to FFS. Plans in the insurance exchanges would seem to face similar incentives not to retain under-priced enrollees and thus might devote limited resources to their care, but these enrollees would not have a public, FFS-like option to which to return.⁶⁰ As such, while a significant portion of the cost of imperfect pricing in the MA context is borne by the government and taxpayers via higher Medicare costs, we speculate that the under-priced consumers themselves may bear more of this cost in the exchanges. Thus, while the inability of plans to send high-cost individuals to an adversely-selected public option could potentially lower government spending, its welfare effects are less clear.

We close by reflecting on what our results suggest about the effect of MA plans on the insurance value provided by the entire Medicare program. Given that throughout our sample period, individuals enrolling in MA had lower costs than those remaining in FFS, our results suggest that this imperfect pricing not only causes the government to overpay for MA enrollees, but also shifts relative Medicare expenditure from high-cost to low-cost beneficiaries, relative to a system with perfect pricing or no MA option. Our results on satisfaction also suggest MA plans, relative to FFS, may devote more resources to lower-cost than higher-cost individuals. Selection into MA would thus seem to diminish Medicare's ability to smooth the utility consequences of variation in health status and thus lower the amount of social insurance per Medicare dollar spent.

These concerns suggest that governments may wish to take special care in “contracting out” social insurance. Imperfect pricing—whereby the government overpays a private firm relative to the cost and quality of in-house production—is, of course, a potential concern every time governments contract with a private party and has received great attention in the literature (see, for example, Hart *et al.* 1997). In the case of, say, paying a supplier for office equipment, the consequences of imperfect pricing would seem limited to whatever amount the government overpaid. With social insurance programs, however, imperfect pricing can induce strategic risk-selection, potentially reintroducing the very risk against which it was intended to insure. At least in the case of Medicare, we find little evidence that risk adjustment has solved this problem.

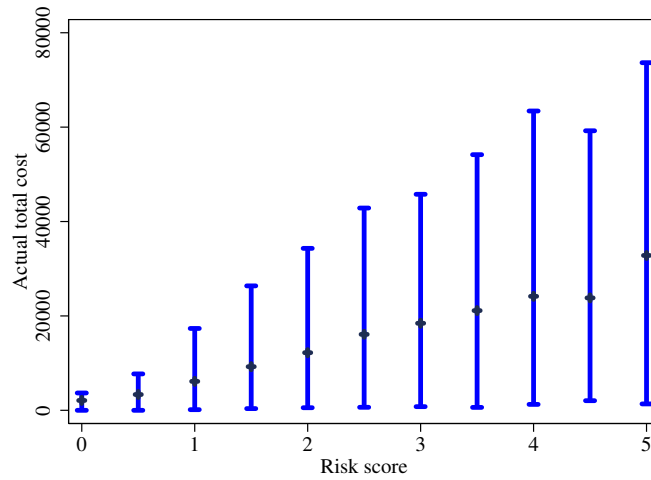
⁶⁰“Optimal risk adjustment” models, pioneered by Glazer and McGuire (2000), have focused on settings without a public option—like the exchanges—and have shown that plans’ incentives to skimp on the care of high-cost enrollees will lead any predictive risk-adjustment model calibrated on plan cost data to systematically under-estimate the cost of high-cost enrollees, further heightening the incentive to avoid such customers. For this reason they suggest that risk adjustment needs to pay high- (low-) cost enrollees more (less) than their observed costs in plan data. Note that these models generally rely on the plan data being accurately reported to the risk adjustor, whereas one might expect that firms would have the incentive to report inflated prices so as to increase capitation payments.

References

- BAICKER, K. and CHANDRA, A. (2005). The effect of malpractice liability on the delivery of health care. *Forum for Health Economics & Policy*, **8**.
- BATATA, A. (2004). The effect of HMOs on Fee-For-Service Health Care expenditures: Evidence from Medicare revisited. *Journal of Health Economics*, **23** (5), 951 – 963.
- BAUHOFF, S. (2010). Do health plans risk-select? an audit study on Germanys social health insurance. *IRB*, **17130**, 101.
- BITLER, M. P., GELBACH, J. B. and HOYNES, H. W. (2006). What mean impacts miss: Distributional effects of welfare reform experiments. *The American Economic Review*, **96** (4), 988–1012.
- BLUMENTHAL, D. and GLASER, J. (2007). Information technology comes to medicine. *New England Journal of Medicine*, **356** (24), 2527–2534.
- BUNDORF, M. K., LEVIN, J. D. and MAHONEY, N. (2008). *Pricing and Welfare in Health Plan Choice*. NBER Working Papers 14153, National Bureau of Economic Research, Inc.
- CHANDRA, A., JENA, A. B. and SKINNER, J. S. (2011). *The Pragmatists Guide to Comparative Effectiveness Research*. Working Paper 16990, National Bureau of Economic Research.
- CHERNEW, M., DECICCA, P. and TOWN, R. (2008). Managed care and medical expenditures of Medicare beneficiaries. *Journal of Health Economics*, **27** (6), 1451 – 1461.
- DAFNY, L. (2005). How do hospitals respond to price changes? *American Economic Review*, pp. 1525–1547.
- DAFNY, L. S. (2010). Are health insurance markets competitive? *The American Economic Review*, **100** (4), 1399–1431.
- DUGGAN, M. (2004). Does contracting out increase the efficiency of government programs? evidence from medicaid hmos. *Journal of Public Economics*, **88** (12), 2549–2572.
- EINAV, L. and FINKELSTEIN, A. (2011). Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic Perspectives*, **25** (1), 115–38.
- ELLIS, R. P. (2008). *Risk Adjustment in Health Care Markets: Concepts and Applications*, Wiley-VCH, pp. 177–222.
- FANG, H., KEANE, M. and SILVERMAN, D. (2008). Sources of advantageous selection: Evidence from the Medigap insurance market. *Journal of Political Economy*, **116** (2).
- FINKELSTEIN, A. and POTERBA, J. (2006). Testing for adverse selection with "unused observables".
- FROGNER, B., ANDERSON, G., COHEN, R. and ABRAMS, C. (2011). Incorporating new research into medicare risk adjustment. *Medical Care*, **49** (3), 295.
- GLAZER, J. and MCGUIRE, T. G. (2000). Optimal risk adjustment in markets with adverse selection: An application to managed care. *American Economic Review*, **90** (4), 1055–1071.
- GOWRISANKARAN, G., TOWN, R. and BARRETTE, E. (2011). Managed care, drug benefits and mortality: An analysis of the elderly. *The BE Journal of Economic Analysis & Policy*, **11** (2).
- HALL, A. (2011). Measuring the return on government spending on the Medicare managed care program. *The BE Journal of Economic Analysis & Policy*, **11** (2), 7.
- HART, O., SHLEIFER, A. and VISHNY, R. (1997). The proper scope of government: Theory

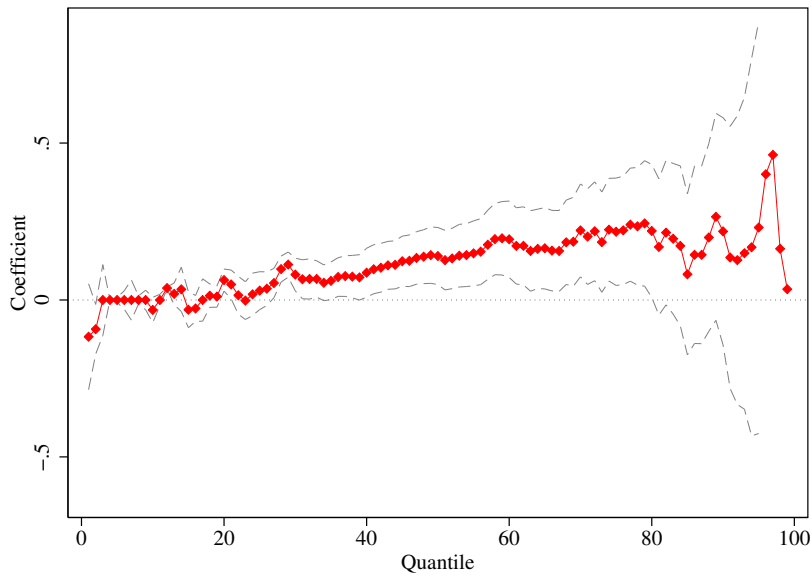
- and an application to prisons*. *Quarterly Journal of Economics*, **112** (4), 1127–1161.
- HOLMSTROM, B. and MILGROM, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization*, **7** (0), 24–52.
- LANGWELL, K. and HADLEY, J. (1989). Evaluation of the Medicare competition demonstrations. *Health Care Financing Review*, **11** (2), 65–80.
- LAZEAR, E. P. (2006). Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics*, **121** (3), 1029–1061.
- LUMLEY, T., DIEHR, P., EMERSON, S. and CHEN, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, **23** (1), 151–169.
- LUSTIG, J. (2010). Measuring welfare losses from adverse selection and imperfect competition in privatized Medicare. *Manuscript. Boston University Department of Economics*.
- MELLO, M. M., STEARNS, S. C., NORTON, E. C. and RICKETTS, T. C. (2003). Understanding biased selection in Medicare HMOs. *Health Services Research*, **38**, 961–992(32).
- NEWHOUSE, J., BUNTIN, M. and CHAPMAN, J. (1997). Risk adjustment and Medicare: Takin a closer look. *Health Affairs*, **16** (5), 26.
- NICHOLAS, L. (2009). *Does Managed Care Reduce Federal Spending? Evidence from Medicare*. Tech. Rep. 09-671, Population Studies Center Research Report.
- PHYSICIAN PAYMENT REVIEW COMMISSION (1997). Testimony of Gail R. Wilensky. Statement before the U.S. House of Representatives Subcommittee on Hearth. Committee on Ways and Means, February 24, 1997.
- POPE, G., KAUTTER, J., ELLIS, R., ASH, A., AYANIAN, J., IEZZONI, L., INGBER, M., LEVY, J. and ROBST, J. (2000). Diagnostic cost group hierarchical condition category models for Medicare risk adjustment: Final report prepared for Hearth Care Financing Administration.
- POPE, G. C., KAUTTER, J., ELLIS, R. P., ASH, A. S., AYANIAN, J. Z., IEZZONI, L. I., INGBER, M. J., LEVY, J. M. and ROBST, J. (2004). Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financing Review*, **25** (4), 119–141.
- SALTMAN, R. and FIGUERAS, J. (1998). Analyzing the evidence on European health care reforms. *Health Affairs*, **17** (2), 85–108.
- SILVERMAN, E. and SKINNER, J. (2004). Medicare upcoding and hospital ownership. *Journal of Health Economics*, **23** (2), 369–389.
- TOWN, R. and LIU, S. (2003). The welfare impact of medicare hmos. *The RAND Journal of Economics*, **34** (4), 719–736.
- VAN DE VEN, W. P. and ELLIS, R. P. (2000). Risk adjustment in competitive health plan markets. In A. J. Culyer and J. P. Newhouse (eds.), *Handbook of Health Economics*, vol. 1, *14*, 1st edn., Elsevier, pp. 755–845.
- YABROFF, K. R., LAMONT, E. B., MARIOTTO, A., WARREN, J. L., TOPOR, M., MEEKINS, A. and BROWN, M. L. (2008). Cost of care for elderly cancer patients in the Uniter States. *Journal of the National Cancer Institute*, **100**, 630–641(12).

Figure 1: Means, 10th and 90th percentiles of total costs, by risk score



Notes: All observations spent all twelve months of the previous year in FFS (so that current-year risk scores can be calculate) and no months of the current year in MA (so that all current cost data can be observed). Observations are taken only from the pre-period so that the sample is unlikely to be selected with respect to the risk score. Sample weights provided by the MCBS are used.

Figure 2: Coefficients and 90-percent confidence intervals from quantile “extensive-margin” regressions



Notes: Each point is the coefficient β from a quantile regression of the form $Risk\ score_{it} = \beta MA_i \times After_t + \gamma MA_i + \delta_t + \epsilon_{it}$. See Section 5 for further detail. Note that the confidence-intervals for quantiles greater than the 96th are suppressed so as not to compress the scale of the figure.

Table 1: Frequency distribution of transitions between FFS and MA, 1994-2006

	Baseline year t equals...				Total
	1994-1996	1997-1999	2000-2002	2003-2005	$t=1994-2005$
FFS (year t) \rightarrow FFS (year $t+1$)	19,017	18,539	18,305	17,329	73,190
FFS (year t) \rightarrow MA (year $t+1$)	566	399	102	464	1,531
MA (year t) \rightarrow FFS (year $t+1$)	102	165	457	125	849
MA (year t) \rightarrow MA (year $t+1$)	1,457	3,282	2,805	2,496	10,040
In sample both years	21,142	22,385	21,669	20,414	85,610
Left sample after baseline year	13,883	14,301	14,983	14,284	57,451
Total observations (baseline year)	35,025	36,686	36,652	34,698	143,061

Notes: An individual in a given year is classified as being on MA if she is on MA for at least half of the months for which she is Medicare eligible in that given year.

Table 2: Summarizing changes in incentives after risk adjustment

HCC score	HCC payment minus demographic payment	HCC payment minus actual Medicare expenditure
0-25 th percentile (lowest scores)	-2,993	67
25-50 th percentile	-2,406	198
50-75 th percentile	-342	549
75-99 th percentile	6,701	893
99-100 th percentile (highest scores)	29,789	-5,907
Total	491	359
Observations	54,369	54,369

Notes: All data taken from the “pre-period” before implementation of risk adjustment, among the subsample of individuals who were in the FFS system all twelve months of the previous year. Both columns use claims data from the previous year to calculate capitation payments under the HCC model for each individual. The first column follows the formula of the demographic model to calculate capitation payments for all individuals. Dollar amounts are adjusted to 2007 dollars using the CPI-U. Sample weights provided by the MCBS are used.

Table 3: Changes in selection patterns after risk adjustment

	Extensive-margin					Intensive-margin	
	(1) Cost	(2) Score	(3) Score	(4) Score	(5) Score	(6) Cost	(7) Cost
Share of year in MA	-2847.0*** [396.9]	-0.305*** [0.0355]	-0.241*** [0.0350]	-0.233*** [0.0355]	-0.238*** [0.0520]	171.5 [316.5]	191.6 [467.2]
Share of year in MA x After 2003	-172.7 [713.4]	0.106* [0.0614]	0.133** [0.0604]	0.140** [0.0552]	0.145** [0.0680]	-1217.9** [604.0]	-1280.3* [691.8]
HCC score						9903.4*** [182.5]	9691.0*** [202.2]
Estim. method	OLS	OLS	OLS	Q. reg.	Q. reg.	OLS	OLS
Outliers trimmed	No	No	Yes	No	No	No	No
1998-2006 only	No	No	No	No	Yes	No	Yes
Observations	73,054	73,054	72,274	73,054	54,646	73,054	54,646

Notes: All observations are in FFS all twelve months of the given year. Year fixed effects included in all regressions. The outcome in cols. (1), (6) and (7) is an individual's current year total Medicare expenditure. The outcome in cols. (2) through (5) is an individual's HCC score the following year, which is based on current-year claims. "Q. reg" refer to median regressions. "Outliers trimmed" excludes individuals with risk scores above the 95th percentile (where percentiles are calculated separately by year). Sample weights provided by the MCBS are used. Dollar amounts are adjusted to 2007 dollars using the CPI-U. Standard errors are clustered by the individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Changes in total baseline Medicare expenditure for MA versus FFS enrollees after risk adjustment

	(1) Costs	(2) Mortality	(3) Charges	(4) Part A	(5) Mortality	(6) FFS costs
Share of year in MA	-2862.2*** [420.6]	-0.154*** [0.0216]		-1292.1 [886.1]		
Share of year in MA x After 2003	-104.9 [757.1]	0.0154 [0.0423]		-182.7 [1019.7]		
In MA at discharge			-1497.5*** [278.4]		-0.00310*** [0.000531]	
In MA at discharge x After 2003			85.50 [409.2]		0.000757 [0.000727]	
Change in FFS share						-1220.5*** [240.3]
Change in FFS x After 2003						160.2 [274.0]
Data set used	MCBS	MCBS	Hosp.	MCBS	Hosp.	County
Years	1994-2006	1994-2006	2000-2006	2000-2006	2000-6	2000-2007
Observations	73,054	138,725	8,217,647	42,051	8,217,647	24,878

Notes: Col. (1) replicates col. (1) of Table 3. Col. (2) estimates the probability of dying in a year as a function of MA status. Col. (3) uses hospital discharge data and estimates total charges as a function of MA status. Col. (4) replicates col. (1) of Table 3 but uses only part A costs and years from 2000-2006, to be more comparable to the hospital discharge analysis. Col. (5) replicates the mortality results in col. (2) using the discharge data. Col. (6) uses county-level data and replicated the analysis in Batata (2004). She shows that in the following regression $\Delta \ln(Avg. FFS costs)_{ct} = \beta \Delta FFS share_{ct}$, where $Avg. FFS costs_{ct}$ is the average per capita costs for all FFS enrollees in county c in year t and $FFS share_{ct}$ is the share of county c 's Medicare enrollees in FFS in year t , β is an estimate for the difference between the costs of the marginal enrollee switching between MA and FFS and the average FFS enrollee. Cols. (2) and (5) report probit estimates reported as changes in probability and the other columns report OLS estimates.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Changes in differential payments after risk adjustment

	Dependent variable: Total Medicare expenditure						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of months in MA x After 2003	1733.4** [747.0]	2081.2*** [790.7]	2108.8*** [776.6]	2011.9** [811.3]	1773.1** [745.4]	1956.5** [883.1]	3177.8*** [828.1]
Share of months in MA	905.2*** [256.5]	883.5*** [289.8]	882.8*** [290.0]	954.0*** [353.8]	1359.6*** [311.6]	1016.9** [499.9]	954.5*** [353.9]
Benchmarks adjusted	Yes	Yes	Yes	Yes	Yes	Yes	No
Baseline controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Lagged health controls	No	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	No	Yes	Yes	Yes	Yes
Dept. var windsorized	No	No	No	No	Yes	No	No
Only 1998-2006	No	No	No	No	No	Yes	No
Observations	73,054	72,930	72,638	72,375	72,375	54,120	72,375

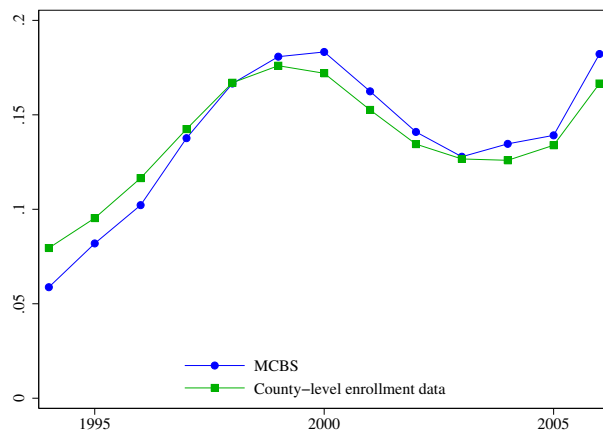
Notes: All observations are in FFS all twelve months of the previous year. Year fixed effects are included in all regressions, and county fixed effects included in col. (2) - (7). All regressions include a once-lagged dependent variable, as well as dummy variables corresponding to eleven bins of lagged Part A and B expenditure (with zero as its own bin and ten bins corresponding to ten deciles of positive Part A and B expenditure, calculated separately for each year). “Benchmarks adjusted” refers to reducing MA payments after 2003 in the following manner: the growth rate in county benchmarks is constrained to match that of the pre-period, and the “budget neutrality” adjustment meant to ease the risk-adjustment process is eliminated. Both of these adjustments make it less likely that the interaction term would have a positive coefficient, as one can see from comparing cols. (4) and (7). “MCBS cap. payments” refers to using the uncorrected capitation payments in the MCBS as a dependent variable. “Baseline controls” include the following: individual’s predicted capitation payment based on the demographic model; race and Hispanic origin; gender; age-in-year fixed effects; fixed effects for eligibility status (disabled and old-age, with and without end-stage-renal disease as a secondary condition); Medicaid status; the interaction of disability status and Medicaid status; income category fixed effects; months of Medicare eligibility; and education category fixed effects. “Lagged health controls” includes fixed effects for the five categories of lagged self reported health (excellent, very good, good, fair, poor), the lagged share of the year spent in an institution, and the lagged risk score. “Health controls” include the following: five categories of current self-reported health, the difference between current and previous-year self-reported health, an indicator variable for being alive the entire year, and the share of the year spent in an institution. The dependent variable is windsorized at the 99th percentile in col. (5). Sample weights provided by the MCBS are used. Dollar amounts are adjusted to 2007 dollars using the CPI-U. Standard errors are clustered by the individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Effect of MA enrollment and health status on enrollee satisfaction

Dependent var:		OLS coefficient estimates (clustered SEs)	
Satisfaction rating (1-4)	Obs.	In MA	MA x Health (demeaned)
Overall medical care	75,884	-0.0168** (0.00727)	0.0158** (0.00659)
Out-of-pocket costs	75,309	0.0648*** (0.00848)	0.0238*** (0.00742)
Follow-up care	69,764	0.00505 (0.00657)	0.00721 (0.00604)
Doctor's concern for your health	74,711	-0.0126* (0.00697)	0.0163*** (0.00633)
Information about your medical condition	75,539	-0.00114 (0.00658)	0.0137** (0.00592)
Access to specialists	57,187	-0.0178** (0.00721)	0.00142 (0.00653)
Questions answered over phone	48,616	-0.0167* (0.00868)	0.0198** (0.00778)
Availability of care nights and weekends	44,502	0.0115 (0.00865)	0.00955 (0.00803)
Medicare care provided in same location	69,380	0.0403*** (0.00646)	0.00973* (0.00580)
Retains coverage type next year (probit coefficients)	84,160	-0.634*** (0.0236)	0.0812*** (0.0211)

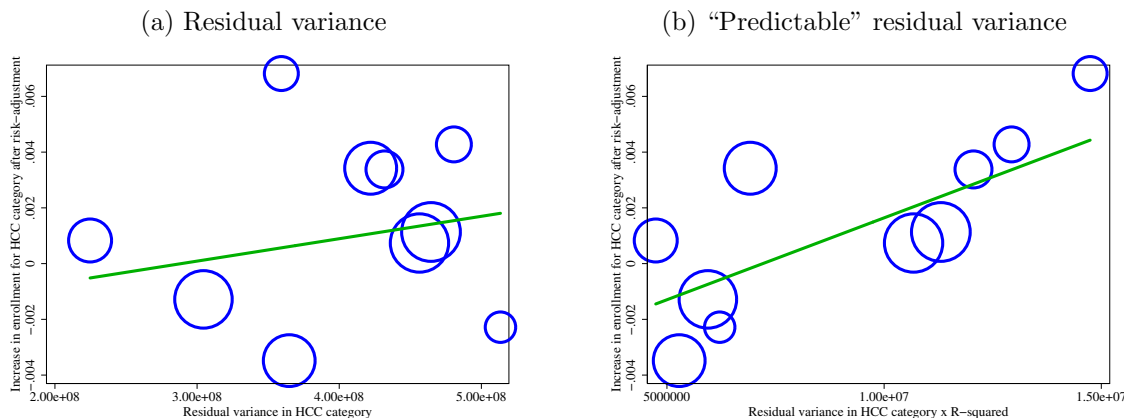
Notes: Each row represents a regression of the form: $satisfaction\ category_i = \beta_1 MA_i + \beta_2 MA_i \times Health_i + \gamma \mathbf{H}_i + \lambda \mathbf{X}_i + \epsilon_i$, where *satisfaction* takes values from one to four (“very dissatisfied,” “dissatisfied,” “satisfied,” “very satisfied”), *MA* is a dummy variable for being enrolled in Medicare Advantage at least half of all Medicare-eligible months in a give year, *Health* is a (demeaned) linear measure of the five-category self-reported health variable, \mathbf{H} is a vector of fixed effect for the five health categories (one, “poor,” up to five, “excellent”), and \mathbf{X} is a vector of basic controls: age, state-of-residence, and year fixed effects, and indicator variables for being female, disabled, or on Medicaid. As the *Health* variable is demeaned, the coefficient on the *MA* indicator variable represents the effect of being enrolled in MA for an enrollee with average health. A positive coefficient on $MA \times Health$ indicates that the relationship between satisfaction and health status for MA enrollees is greater (“more positive”) than that for FFS enrollees. Note that the sample size varies across regressions because not all questions are asked each year and there is variation in the number of individuals who respond that they do not have enough information to answer. This regression focuses only on people with the same MA status in both the baseline year and the previous year. Sample weights provided by the MCBS are used and standard errors are clustered by individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Figure 1: Share of Medicare enrollees in a Medicare Advantage plan, 1994-2006



Notes: The first series is based on our MCBS sample. The second series is from annual county-level MA penetration data published by CMS.

Appendix Figure 2: Relationship between residual variance, “predictable residual variance,” and MA enrollment increases



Notes: Each plotted point corresponds to one of the ten most common HCC disease categories in the risk-adjustment formula and are weighted by the prevalence of the condition in the Medicare population. The y-axis in both figures depicts the increase in the probability that someone with that disease condition switches to MA from FFS after risk adjustment. That is, for each HCC category c , we estimate $MA_{it} = \beta_1 I_{i,t-1}^c \times After\ 2003 + \beta_2 I_{i,t-1}^c + \delta_t + \epsilon_{it}$, where $I_{i,t-1}^c$ is an indicator variable for having a claim for disease category c in the previous year (and thus having the current year’s risk score adjusted accordingly) and all other coefficients are as in the text. For each condition, β_1 is plotted on the y-axis. The x-axis in figure (a) plots the residual variance for each HCC category—that is, for each HCC disease category, what is the variance of $(Total\ Medicare\ Costs_i - HCC\ Capitation\ Payment)$ for individuals with HCC Condition c . Note that this variance can only be calculated for those in FFS, and we further restrict the sample to those in FFS before 2004 so as not to pick up differential selection into MA plans post-risk-adjustment. We also exclude observations above the 99th percentile of risk scores since we showed earlier that plans avoid these individuals even after risk adjustment and including them can have very large effects on the variance measures. Figure (b) is exactly the same except that the x-axis is slightly modified. For each disease category, we estimate the following equation on individuals in FFS before 2004: $Total\ Medicare\ Costs_i - HCC\ capitation\ payment = \beta \mathbf{X}_i$, where \mathbf{X} is a vector of exploratory variables including race, gender, Medicaid, disability and institutional status, as well as lagged $Total\ Medicare\ Costs_i$ and self-reported health. Thus, the R^2 from this regression is a measure of how well the errors in the HCC formula can be predicted by observables. We multiply this R^2 by the variance measures in figure (a) and plot these values on the x-axis.

Appendix Table 1: The ten most common conditions in the HCC formula

Category	Prevalence	Description	HCC weight
80	0.126	Congestive Heart Failure	0.417
108	0.124	Chronic Obstructive Pulmonary Disease	0.376
19	0.120	Diabetes without Complication	0.200
92	0.097	Specified Heart Arrhythmias	0.266
105	0.094	Vascular Disease	0.357
10	0.063	Breast, Prostate, Colorectal Cancers	0.233
83	0.046	Angina Pectoris/Old Myocardial Infarction	0.235
96	0.045	Ischemic or Unspecified Stroke	0.306
38	0.039	Rheum. Arthritis and Inflam. Connective Tissue Disease	0.322
79	0.038	Cardio-Respiratory Failure and Shock	0.692

Notes: This table is based on the FFS population, 1993-2006. The weight associated with each HCC condition is added to a person's total risk score. Given that the average benchmark is roughly \$9,345—average per capita FFS expenditure (\$8,344) multiplied by the benchmark-to-FFS markup in 2006 (1.12)—in 2006, having been diagnosed with congestive heart failure in the previous year would mean an individual's capitation payment is increased by $0.417 * \$9,345 = \$3,897$.

Appendix Table 2: “Toy” example illustrating model predictions

	No Conditions	Has Cancer	
		Benign	Malignant
<i>Model Fundamentals</i>			
True costs	5	6	13
Screening	1	2	2
Risk score	5	9.5	9.5
Residual (cost - r. score)	0	-3.5	3.5
<i>Not Risk Adjusted</i>			
Capitation Payments	8	8	8
Differential Payments	3	2	-5
Profits	2	0	-7
<i>Risk Adjusted</i>			
Capitation Payments	5	9.5	9.5
Differential Payments	0	3.5	-3.5
Profits	-1	1.5	-5.5

Notes: Boxes indicate the type of enrollee that will join MA under each regime.

Appendix Table 3: Additional robustness checks of results from Tables 3 and 5

	Table 3		Table 5		
	(1) Score	(2) Costs	(3) Tot. Mcare	(4) Δ Tot. Mcare	(5) Tot. Mcare
In MA majority of year	-0.176*** [0.0299]	229.4 [253.4]	881.2*** [293.1]		
In MA majority of year x After 2003	0.0827* [0.0477]	-1045.3** [511.3]	1497.8** [686.6]		
Share of months in MA				1331.7*** [441.5]	857.8** [356.2]
Share of months in MA x After 2003				1739.1* [961.9]	2698.5*** [748.0]
Estim. model	Qreg	OLS	OLS	OLS	OLS
MCBS cap. payments	N/A	N/A	No	No	Yes
Observations	73,054	73,054	72,375	72,375	72,375

Notes: See Tables 3 and 5 for additional detail. Cols. (1), (2) and (3) merely replicate cols. (3) and (6) of Table 3 and col. (4) of Table 5, respectively, but use a binary variable for MA status instead of the share of Medicare-eligible months an individual is in MA. Col. (4) uses the change in Total Medicare expenditure, instead of using the level and including lagged dependent variables, as in Table 5. As noted in the data Appendix, we had to re-create Total Medicare expenditure as the MCBS capitation payments after 2003 do not reflect individual-level variation in risk scores. Col. (5) merely replicated col. (7) of Table 5 to show that, in practice, the result using the uncorrected version published in the MCBS does not appreciably affect the results. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 4: Changes in differential payments after risk adjustment, using simulated instead of administrative risk scores

	Dependent variable: Total Medicare expenditure						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of months in MA x After 2003	1707.1** [750.4]	2055.3*** [797.2]	2097.0*** [782.3]	2000.0** [815.8]	1746.3** [743.1]	1942.5** [887.2]	3158.2*** [832.1]
Share of months in MA	906.6*** [256.5]	885.5*** [289.8]	885.1*** [290.0]	956.2*** [353.7]	1361.8*** [311.5]	1021.1** [499.9]	957.2*** [353.8]
Benchmarks adjusted	Yes	Yes	Yes	Yes	Yes	Yes	No
Baseline controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Lagged health controls	No	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	No	Yes	Yes	Yes	Yes
Dept. var windsorized	No	No	No	No	Yes	No	No
Only 1998-2006	No	No	No	No	No	Yes	No
Observations	73,054	72,930	72,638	72,375	72,375	54,120	72,375

Notes: This table replicates Table 5, but instead of using the risk scores provided by CMS to calculate capitation payments we use the risk scores we simulate using FFS claims data in the previous year. See notes to Table 5 for additional detail. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 5: Premium reductions received by MA recipients

Year	Pt. B monthly premium	Avg. premium for MA enrollees	Annual savings for MA enrollees	Pre- and post-period weighted averages
2001	58.50	26.91	379.08	
2002	62.10	36.80	303.60	
2003	66.33	41.81	294.25	\$329.12
2004	73.26	29.15	529.32	
2005	82.89	23.32	714.86	
2006	91.16	27.19	767.56	\$682.88

Notes: All dollars amounts are in 2007 dollars. Nominal Part B premiums are taken from the annual Medicare Trustees Report. The average premiums paid by MA enrollees is taken from analysis done by Mathematica, published by the AARP. See http://assets.aarp.org/rgcenter/health/2006_23_medicare.pdf (we take numbers from p. 8). We weight each year by MA enrollment to calculate the pre- and post-period average annual savings.

Appendix Table 6: Satisfaction measures for MA versus FFS before and after risk adjustment

	(1) gen	(2) costs	(3) follow	(4) concern	(5) info	(6) specialist	(7) phone	(8) avail	(9) sameloc
In MA majority of year x After 2003	0.0166 [0.0156]	-0.0993*** [0.0185]	-0.00189 [0.0146]	-0.0206 [0.0152]	-0.00644 [0.0142]	0.00763 [0.0153]	0.0352* [0.0188]	-0.00456 [0.0191]	-0.00434 [0.0146]
Observations	75884	75309	69764	74711	75539	57187	48616	44502	69380

Notes: Each column represents a regression of the form: $satisfaction\ category_i = \beta_1 MA_i + \beta_2 MA_i \times After_i + \lambda \mathbf{X}_i + \epsilon_i$, where *satisfaction* takes values from one to four (“very dissatisfied,” “dissatisfied,” “satisfied,” “very satisfied”), *MA* is a dummy variable for being enrolled in Medicare Advantage at least half of all Medicare-eligible months in a give year, and \mathbf{X} is a vector of basic controls: age, state-of-residence, and year fixed effects, and indicator variables for being female, disabled, or on Medicaid. The abbreviations in the column labels refer, in the same order, to the nine satisfaction categories described in Table 6. Note that the sample size varies across regressions because not all satisfaction questions are asked each year and there is variation in the number of individuals who respond that they do not have enough information to answer. This regression focuses only on people with the same MA status in both the baseline year and the previous year. Sample weights provided by the MCBS are used and standard errors are clustered by individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

50

Appendix Table 7: Satisfaction measures for MA versus FFS before and after risk adjustment, by health status

	(1) gen	(2) costs	(3) follow	(4) concern	(5) info	(6) specialist	(7) phone	(8) avail	(9) sameloc
MA x Self-reported health x After	0.0170 [0.0148]	0.00864 [0.0177]	0.0150 [0.0138]	0.0102 [0.0144]	0.0165 [0.0134]	-0.0159 [0.0146]	0.0305* [0.0175]	0.00240 [0.0183]	-0.00128 [0.0132]
Observations	75884	75309	69764	74711	75539	57187	48616	44502	69380

Notes: Each row represents a regression of the form: $satisfaction\ category_i = \beta_1 MA_i + \beta_2 MA_i \times Health_i + \beta_3 MA_i \times Post_i + \beta_4 MA_i \times Health_i \times Post_i + \gamma \mathbf{H}_i + \lambda \mathbf{X}_i + \epsilon_i$, where *satisfaction* takes values from one to four (“very dissatisfied,” “dissatisfied,” “satisfied,” “very satisfied”), *MA* is a dummy variable for being enrolled in Medicare Advantage at least half of all Medicare-eligible months in a give year, *Health* is a (demeaned) linear measure of the five-category self-reported health variable, \mathbf{H} is a vector of fixed effect for the five health categories (one, “poor,” up to five, “excellent”), and \mathbf{X} is a vector of basic controls: age, state-of-residence, and year fixed effects, and indicator variables for being female, disabled, or on Medicaid. Note that the sample size varies across regressions because not all questions are asked each year and there is variation in the number of individuals who respond that they do not have enough information to answer. This regression focuses only on people with the same MA status in both the baseline year and the previous year. Sample weights provided by the MCBS are used and standard errors are clustered by individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 8: Satisfaction measures for Hispanics and non-Hispanics before and after risk adjustment

	(1) gen	(2) costs	(3) follow	(4) concern	(5) info	(6) specialist	(7) phone	(8) avail	(9) sameloc
MA x Hispanic x After	0.0344 [0.0660]	-0.0135 [0.101]	0.0534 [0.0666]	0.160** [0.0680]	0.115* [0.0689]	0.0784 [0.0643]	0.0331 [0.0811]	0.136* [0.0790]	0.0703 [0.0677]
Observations	75884	75309	69764	74711	75539	57187	48616	44502	69380

Notes: Each column represents a regression of the form: $satisfaction\ category_i = \beta_1 MA_i + \beta_2 Hispanic_i + \beta_3 MA_i \times After_i + \beta_4 MA_i \times Hispanic_i + \beta_5 MA_i \times Hispanic_i \times After_i + \lambda \mathbf{X}_i + \epsilon_i$, where *satisfaction* takes values from one to four (“very dissatisfied,” “dissatisfied,” “satisfied,” “very satisfied”), *MA* is a dummy variable for being enrolled in Medicare Advantage at least half of all Medicare-eligible months in a give year, and \mathbf{X} is a vector of basic controls: age, state-of-residence, and year fixed effects, and indicator variables for being female, disabled, or on Medicaid. Note that the sample size varies across regressions because not all questions are asked each year and there is variation in the number of individuals who respond that they do not have enough information to answer. This regression focuses only on people with the same MA status in both the baseline year and the previous year. Sample weights provided by the MCBS are used and standard errors are clustered by individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 9: Self-reported health and satisfaction with health care before and after risk adjustment

	(1) gen	(2) costs	(3) follow	(4) concern	(5) info	(6) specialist	(7) phone	(8) avail	(9) sameloc
Health x After	0.00639 [0.00467]	0.00458 [0.00596]	0.00498 [0.00452]	0.00629 [0.00470]	0.00858* [0.00453]	0.0198*** [0.00471]	0.0150*** [0.00564]	0.00593 [0.00572]	0.0132*** [0.00459]
Observations	75884	75309	69764	74711	75539	57187	48616	44502	69380

Notes: Each column represents a regression of the form: $satisfaction\ category_i = \beta_1 Health_i + \beta_2 Health_i \times After_i + \lambda \mathbf{X}_i + \epsilon_i$, where *satisfaction* takes values from one to four (“very dissatisfied,” “dissatisfied,” “satisfied,” “very satisfied”), *Health* is a (demeaned) linear measure of the five-category self-reported health variable, and \mathbf{X} is a vector of basic controls: age, state-of-residence, and year fixed effects, and indicator variables for being female, disabled, or on Medicaid. Note that the sample size varies across regressions because not all questions are asked each year and there is variation in the number of individuals who respond that they do not have enough information to answer. Sample weights provided by the MCBS are used and standard errors are clustered by individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 10: Satisfaction measures in MA-intensive versus other counties before and after risk adjustment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	gen	costs	follow	concern	info	specialist	phone	avail	sameoloc
County's 1993 - 2003	0.0679	-0.0538	0.00558	0.0328	0.0358	-0.0391	0.0742	0.0206	0.00143
MA penetration x After 2003	[0.0624]	[0.0697]	[0.0615]	[0.0565]	[0.0527]	[0.0561]	[0.0592]	[0.0646]	[0.0572]
Observations	119662	118695	109923	117784	119071	92688	79584	69678	109206

Notes: Each column represents a regression of the form: $satisfaction\ category_i = \beta_1 MA_i + \beta_2 Hispanic_i + \beta_3 MA_i \times After_i + \beta_4 MA_i \times Hispanic_i + \beta_5 MA_i \times Hispanic_i \times After_i + \lambda \mathbf{X}_i + \epsilon_i$, where *satisfaction* takes values from one to four (“very dissatisfied,” “dissatisfied,” “satisfied,” “very satisfied”), *MA* is a dummy variable for being enrolled in Medicare Advantage at least half of all Medicare-eligible months in a give year, and \mathbf{X} is a vector of basic controls: age, state-of-residence, and year fixed effects, and indicator variables for being female, disabled, or on Medicaid. Note that the sample size varies across regressions because not all questions are asked each year and there is variation in the number of individuals who respond that they do not have enough information to answer. Sample weights provided by the MCBS are used and standard errors are clustered by individual. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 11: Quality of care measures among 55-74 year-olds in the 2000-2006 NHIS

	Unhappy with access to care regarding...				
	(1) Phone	(2) Appointment	(3) Waiting	(4) Hours	(5) Rec'd flu shot
Age 65-74 x After 2003	-0.00101 [0.00270]	0.000845 [0.00385]	0.00186 [0.00378]	-0.00198 [0.00252]	0.00888 [0.00901]
Observations	48,675	48,675	48,673	49,107	48,279

All regressions take the form $outcome_{it} = \beta Age\ 65 - 75_i \times After\ 2003_t + \alpha_i + \delta_t + \epsilon_{it}$, where α_i are a vector of age-in-year fixed effects and δ_t are a vector of year fixed effects. The first four outcomes measure self-reported dissatisfaction (a binary variable in the NHIS) with, respectively, reaching health care providers over the phone, making a timely appointment, time spent in the waiting room, and providers' hours of operation. The final variable is a binary variable for whether the respondent received a flu shot in the past twelve months. We lack of any improvement among the young elderly after risk adjustment is robust to the following specification checks. First, we excluded any of the near-elderly who are on disability, as many will qualify for Medicare. Second, we excluded those who report having no contact with health professionals in the past two years, given that otherwise it is difficult to separate the effect of not being dissatisfied with simply not seeking care. Third, we included data from 2007 to 2008 (beyond our MCBS sample period). Results from each of these checks is available from the authors.

Appendix Table 12: Mortality rates for near elderly (55-64) and young elderly (65-74) before and after risk adjustment

	Log mort. rate		Δ log mort. rate		Δ_2 log mort. rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-period MA share x Elderly x After	0.000247 [0.0164]	-0.0127 [0.0176]				
Δ MA share x Elderly x After			0.00364 [0.159]	0.0256 [0.169]		
Δ_2 MA share x Elderly x After					0.111 [0.0898]	0.107 [0.0961]
Lagged dept. var.?	No	Yes	No	No	No	No
County FE?	Yes	Yes	No	Yes	No	Yes
Observations	30,996	25,453	25,453	25,453	22,140	22,140

Notes: Data taken from county-level vital statistics data. "Pre-period MA share" is a county's average MA share between 2000 and 2003. $\Delta MA\ share_{it}$ is defined as $MA\ share_{it} - MA\ share_{i,t-1}$ and $\Delta_2 MA\ share_{it}$ is defined as $MA\ share_{it} - MA\ share_{i,t-2}$ for county i in year t , with $\Delta \ln(mortality\ rate)$ and $\Delta_2 \ln(mortality\ rate)$ defined analogously. "Elderly" is an indicator variable for being age 65-74 as opposed to 55-64. For all specifications, all lower-order terms of the triple interaction terms are also included but are not reported. We also explored results using data through 2008 (beyond the MCBS sample period) and also found no effect on elderly mortality after 2003 (results available upon request).

10 Theoretical Framework Appendix

In this Appendix, we formalize the intuition provided in Section 3. The purpose of this model is to understand how adopting risk adjustment will influence total costs to the government from offering MA plans. We therefore take as given the basic contours of the risk-adjustment formula used by CMS, as opposed to exploring the optimal formula, as in Glazer and McGuire (2000) and others.

While an MA plan must be open at the same price to all individuals in the plan’s geographic area of operation, the model assumes that, as shown in earlier work, plans have at least *some* scope to encourage individual with certain characteristics to enroll. For example, by differentially advertising in *Diabetes Forecast* (a publication of the American Diabetes Association), MA plans could increase the probability that diabetics enroll.

We emphasize that this process does not necessarily imply that the plan have access to information about the characteristics of any individual Medicare beneficiaries. Instead, plans could use information on the conditional distribution of costs in the Medicare population and employ strategies, such as targeted advertising or changing the quality of physicians in their network, to encourage beneficiaries with certain conditions to enroll. Beneficiaries, who have private information on their health type, choose to enroll in MA based on the perceived costs and benefits of the plan.⁶¹

To keep the model tractable, we do not model the consumer side of the enrollment decision and instead focus on plans’ decision to incur the costs associated with these screening activities in return for enrolling a selected subsample from the Medicare population. In our model, plans have an incentive to target individuals for whom the difference between capitation payments and expected costs is the greatest, and risk adjustment changes this set of individuals by changing how capitation payments are calculated.

10.1 Basic framework and assumptions

10.1.1 Cost of health insurance coverage

Let the cost of covering individual i in a given year be given by $m_i = b_i + v_i$, where b_i is an individual’s expected cost conditional on the variables included in the risk-adjustment formula used by the government, and v_i is the residual. As MA contracts have a year-long duration, the model is single-period, and we thus specify costs over a single year.⁶² Both v and b are in units of absolute dollars.⁶³ While $\mathbb{E}(v|b) = 0$ for all b , the conditional variance of v can vary with b , consistent with past work showing substantial heteroskedasticity in

⁶¹Note that the model does not rule out the possibility that plans use some information to actively encourage some individuals to enroll in their plan. For example, MA plans may respond more quickly to enrollment requests from respondents residing in low-cost areas, as Bauhoff (2010) finds in the German context.

⁶²We return to the question of dynamics in Section 8 when we discuss recalibrating the risk-adjustment model over time.

⁶³Note that m is the cost to the *insurer*—the cost of total medical care plus administrative costs, less the out-of-pocket costs paid by the individual—not total actual medical costs. As in Glazer and McGuire (2000), we do not model out-of-pocket costs in order to focus on selection, though we present results on individuals’ satisfaction with their out-of-pocket costs in Section 7.

medical costs. We assume that costs m are the same whether an individual is in FFS or MA. Of course, MA plans may be better or worse at controlling costs than FFS, and all of the results that follow hold when MA costs are proportional to FFS costs. However, we focus on the case where costs are identical. This assumption not only simplifies the analysis, but also allows us to more easily focus on the difference between payments to private plans for insuring person i and the counterfactual cost if the government directly covered her, which is a key parameter for evaluating the fiscal impact of private Medicare Advantage plans.⁶⁴

10.1.2 Capitation payments and risk adjustment

Without risk adjustment, plans receive a fixed payment \bar{p} for each individual they enroll. We model risk adjustment as replacing \bar{p} with a function $p(b)$, $p' > 0$, so that capitation payments become an increasing function of b . While our main results on selection and differential payments do not require that risk-adjusted payments are linear in b , this assumption corresponds to the MA setting where capitation payments are calculated by multiplying risk scores by a fixed county factor. As it allows us to generate additional empirical predictions and also simplifies the analysis, we take as a baseline assumption that $p''(\cdot) = 0$.⁶⁵

We also make risk adjustment be “payment-neutral,” that is, $\mathbb{E}(p(b)) = \bar{p}$ for the Medicare population as a whole. In other words, if the entire population joined a private plan, the government would pay the same average capitation payment with or without risk adjustment.⁶⁶

Finally, we want to allow for the degree of risk adjustment to vary, which again mirrors the actual experience of the phasing-in of risk adjustment between 2004 and 2007. We define capitation payments as $(1 - \Omega)\bar{p} + \Omega p(b_i)$, where $\Omega \in [0, 1]$ is the risk-adjusted share of the capitation payment.

As indicated in the introduction, the key objective of risk adjustment was to reduce the difference between a plan’s capitation payment for covering an individual and the cost to the government had it directly covered him via FFS. Having defined how risk adjustment affects capitation payments, we can make this concept slightly more precise.

Definition. *The “differential payment” for individual i equals*

$$\underbrace{(1 - \Omega)\bar{p} + \Omega p(b_i)}_{\text{capitation payment}} - \underbrace{(b_i + v_i)}_{\text{FFS cost}}$$

⁶⁴Whether the HMO model is actually more efficient than the fee-for-service model even absent selection effects is an open question. Duggan (2004) finds that when some California counties mandated their Medicaid recipients to switch from the traditional FFS system to an HMO, costs increased by 17 percent relative to counties that retained FFS. As, within a county, individuals did not select between FFS or an HMO, selection issues are unlikely to be driving the result.

⁶⁵In particular, our proofs of Proposition 1 (that risk adjustment causes selection to fall along the b margin and rise along the v margin) and Proposition 3 (that the effect of risk adjustment on differential payments is ambiguous) do not depend on the linearity of $p(\cdot)$.

⁶⁶As we discuss in Section 2, plans were actually given temporary payments to ease the transition into risk adjustment, but as a matter of theory, we are more interested in the steady-state results when the system returns to payment-neutral conditions. Section 6 reports our empirical results with and without these temporary payments.

10.1.3 Screening costs

Though we discuss profit-maximization in greater detail shortly, plan profits are obviously a function of an individual’s cost $m_i = b_i + v_i$, and thus plans will have preferences over the b and v values of their enrollees, even if the plan is unable to observe b and v for any potential beneficiary. However, MA plans are required to accept any patient in their geographic coverage area who chooses to enroll, and selectively encouraging certain individuals to enroll will entail screening costs. Thus, even though plans cannot directly control the characteristics of their beneficiaries, because plans can indirectly influence the population who signs up, we assume that b and v are choice variables on the part of the plan.

We assume that the per capita screening cost c a plan incurs is given by $c(b, v)$, where b and v are its enrollees’ average values of b_i and v_i . Since randomly enrolling individuals from the general population should require minimum screening costs, $c(\bar{b}, \bar{v})$ is a global minimum, where \bar{b} and \bar{v} are population averages (recall we assume $\bar{v} = 0$). Encouraging individuals to enroll who are further from the mean is costly, so $c_x < 0$ for $x < \bar{x}$ and $c_x > 0$ for $x > \bar{x}$ for $x \in \{b, v\}$. We also assume that the cost function is everywhere convex.

Finally, we assume that $c_{bv} > 0$. This assumption implies that for higher values of b , the incremental cost of reducing v falls. This assumption rules out the possibility that screening in b and v are complements. Because the variance of medical costs is typically a positive function of expected costs (see, e.g., Lumley *et al.* 2002 and Figure 1) and v is measured in absolute dollars, it should be easier to attract, say, a cancer patient with costs \$100 below what her risk score would predict than someone without a single documented disease condition with costs \$100 below what her risk score would predict.

With screening costs thus defined, we can now specify a plan’s profit function. In our baseline model, we make the simplifying assumption that plans cannot affect the number of individuals that they enroll, though we return to this assumption later in the section. Plans instead focus on maximizing the average profit per enrollee, which is a function of b and v . Thus, plans maximize the following expression:

$$\mathbb{E}(\pi) = \underbrace{(1 - \Omega)\bar{p} + \Omega p(b)}_{\text{capitation payment}} - \underbrace{\left(\frac{b + v}{\text{FFS cost}} \right)}_{\text{FFS cost}} - \underbrace{c(b, v)}_{\text{screening cost}} . \quad (5)$$

We now use this framework to prove a number of results regarding selection and differential payments.

10.2 Main Results

We begin with our main selection result, which characterizes how plans will react to a change in risk adjustment.

Proposition 1. *The following two conditions hold when the risk-adjusted share Ω of the capitation payment increases:*

- (i) *Plans decrease screening along the b margin and thus the average value of b among their enrollees rises (“extensive-margin” selection decreases).*

(ii) Plans increase screening along the v margin and thus the average value of v among their enrollees falls (“intensive-margin” selection increases).

This proposition formalizes the result from the Theoretical Framework that (1) “the risk scores of those enrolling in MA will increase relative to those remaining in FFS” (2) “actual costs conditional on the risk score will fall among those enrolling in MA relative to those remaining in FFS.”

Proof. We are required to show that $\frac{\partial b^*}{\partial \Omega} > 0$ and $\frac{\partial v^*}{\partial \Omega} < 0$, where b^* and v^* are a plan’s optimal levels of b and v . The first-order conditions from maximizing the profit expression in equation (5) with respect to b and v are given by

$$[b] : \Omega p'(b^*) - c_b(b^*, v^*) = 1 \quad (6)$$

$$[v] : -c_v(b^*, v^*) = 1 \quad (7)$$

Totally differentiating equation (6) with respect to Ω yields

$$p'(\cdot) + \Omega p''(\cdot) \frac{\partial b^*}{\partial \Omega} - c_{11}(\cdot, \cdot) \frac{\partial b^*}{\partial \Omega} - c_{12}(\cdot, \cdot) \frac{\partial v^*}{\partial \Omega} = 0 \quad (8)$$

Similarly, equation (7) yields:

$$c_{bv}(\cdot, \cdot) \frac{\partial b^*}{\partial \Omega} + c_{vv}(\cdot, \cdot) \frac{\partial v^*}{\partial \Omega} = 0$$

or

$$\frac{\partial v^*}{\partial \Omega} = -\frac{c_{bv}}{c_{vv}} \frac{\partial b^*}{\partial \Omega}. \quad (9)$$

Substituting equation (9) into (8) gives:

$$p'(\cdot) + \Omega p''(\cdot) \frac{\partial b^*}{\partial \Omega} - c_{bb}(\cdot, \cdot) \frac{\partial b^*}{\partial \Omega} - c_{bv}(\cdot, \cdot) \left(-\frac{c_{bv}}{c_{vv}} \frac{\partial b^*}{\partial \Omega}\right) = 0.$$

We can now solve for $\frac{\partial b^*}{\partial \Omega}$ and sign many of the terms:

$$\frac{\partial b^*}{\partial \Omega} = \frac{\begin{array}{c} + \text{ as cap payments increase in } b \\ \overbrace{p'} \\ + \text{ by convexity of } c(\cdot, \cdot) \end{array}}{-\Omega p''(\cdot) + \frac{\overbrace{(c_{bb}c_{vv}) - c_{bv}^2}}{c_{vv}}} \quad (10)$$

+ by convexity of $c(\cdot, \cdot)$

By assumption, $p''(\cdot) = 0$, so the entire denominator is positive. As $\frac{\partial b^*}{\partial \Omega} > 0$ and $c_{bv}, c_{vv} > 0$, equation (9) gives the result in (ii). ■

As risk adjustment makes capitation payments a positive function of b , plans will spend

less effort finding low- b enrollees and instead focus on finding low- v enrollees. We term the first result “extensive-margin” selection as it relates to the government’s risk score, which is an approximate measure of actual cost; we term the second result “intensive-margin” selection because it relates to how intensely individuals are selected conditional on the risk score.⁶⁷

Proposition 2. *For $\Omega_0 < \Omega_1$, moving from Ω_0 to Ω_1 will always decrease differential payments if (1) b and v are held fixed at their equilibrium values under Ω_0 and (2) if individuals are positively selected with respect to b under Ω_0 .*

This proposition formalizes the result from the Theoretical Framework that, “applying the risk-adjustment formula to the pre-risk-adjustment population of MA enrollees would have decreased the total capitation payments the government would have made on their behalf.”

Proof. The result is easy to show when $p(\cdot)$ is linear. Recall that $p(\cdot)$ is “payment-neutral,” so that $\mathbb{E}(p(b)) = \bar{p}$. For linear p , $\mathbb{E}(p(b)) = p(\bar{b}) = \bar{p}$, so risk adjustment does not change the payment for an individual with $b = \bar{b}$. As $p' > 0$, $p(b) < p(\bar{b}) = \bar{p}$ for all $b < \bar{b}$. So, as long as individuals are positively selected with respect to b under Ω_0 ($b < \bar{b}$), the proposition holds.

■

Proposition 3. *The effect of increasing Ω on a plan’s average differential payment is ambiguous.*

Proof. Let $\phi(\Omega)$ denote the differential payment when the risk-adjusted share of the capitation payment is set to Ω and plans are at their optimal b and v values:

$$\phi(\Omega) = \underbrace{\Omega p(b^*(\Omega)) + (1 - \Omega)\bar{p}}_{\text{capitation payment}} - \underbrace{(b^*(\Omega) + v^*(\Omega))}_{\text{actual costs}} \quad (11)$$

Differentiating with respect to Ω gives:

$$\phi'(\Omega) = \Omega p' \frac{\partial b^*}{\partial \Omega} + p(b^*) - \bar{p} - \frac{\partial b^*}{\partial \Omega} - \frac{\partial v^*}{\partial \Omega} \quad (12)$$

Rearranging and substituting $\frac{\partial v^*}{\partial \Omega} = -\frac{c_{bv}}{c_{vv}} \frac{\partial b^*}{\partial \Omega}$ from equation (9) yields

$$\phi'(\Omega) = [p(b^*) - \bar{p}] + \frac{\partial b^*}{\partial \Omega} \left(\Omega p' - 1 + \frac{c_{bv}}{c_{vv}} \right) \quad (13)$$

We showed in the proof of Proposition 2 that $p(b^*) < \bar{p}$ for any equilibrium b^* , so the first term (in brackets) is negative. However, the second term is ambiguous. While Ω and

⁶⁷The empirical work will focus on the government’s observed risk score—that is, $p(b)$ in the parlance of the model—as b itself is not observable. But as $p'(b) > 0$, Proposition (1) (i) implies that $p(b)$ will increase as well, thus giving the testable prediction that risk scores *as measured by the government* increase with an increase in risk adjustment.

p' are both by assumption less than one and $\frac{\partial b^*}{\partial \Omega} > 0$ by Proposition 1, if $\frac{c_{bv}}{c_{vv}}$ is large, the expression can indeed be positive. This condition requires c_{bv} to be sufficiently positive. ■

Endogenizing firm enrollment size

We now assume that firms maximize *total*, as opposed to per capita, profits which equal $q(b, v)\pi(b, v, \Omega)$, where π is average per capita profits as specified in equation (5) and q is the number of enrollees the firm has.

The first-order conditions with respect to b and v are now:

$$[b] : q_b(b, v)\pi(b, v, \Omega) + q(b, v) \underbrace{(\Omega p' - 1 - c_b(b, v))}_{\pi_b} \quad (14)$$

$$[v] : q_v(b, v)\pi(b, v, \Omega) + q(b, v) \underbrace{(-1 - c_v(b, v))}_{\pi_v} \quad (15)$$

Note that when the level of q is larger relative to (i) its partial derivatives or (ii) the level of per capita profits, then equations (14) and (15) reduce to the original first-order conditions of $\pi_b = \pi_v = 0$.

Overall Cost Selection

In this section, we explore how the move to risk adjustment changes selection along overall FFS costs ($b + v$). We are interested in $\frac{d(b^* + v^*)}{d\Omega}$.

Proposition 4. *If, beginning at no risk adjustment ($\Omega = 0$), increasing Ω increases a firm's average differential payment, then the effect of increasing Ω on the overall FFS costs of beneficiaries ($b + v$) is negative. That is, $\left. \frac{d(b^* + v^*)}{d\Omega} \right|_{\Omega=0} < 0$.*

Proof. From (9), we know that $\frac{dv^*}{d\Omega} = -\frac{c_{bv}}{c_{vv}} \frac{db^*}{d\Omega}$. So, we know that

$$\begin{aligned} \frac{d(b^* + v^*)}{d\Omega} &= \frac{db^*}{d\Omega} + \frac{dv^*}{d\Omega} \\ &= \frac{db^*}{d\Omega} - \frac{c_{bv}}{c_{vv}} \frac{db^*}{d\Omega} \\ &= \frac{db^*}{d\Omega} \left(1 - \frac{c_{bv}}{c_{vv}} \right). \end{aligned}$$

From Proposition 3, we know that if increasing risk adjustment causes overpayments to increase, it must be the case that $\Omega p' - 1 + \frac{c_{bv}}{c_{vv}} > 0$. This implies that

$$\Omega p' > 1 - \frac{c_{bv}}{c_{vv}}.$$

If there is no risk adjustment ($\Omega = 0$), this equation implies that $1 - \frac{c_{bv}}{c_{vv}} < 0$. Now, by Proposition 1, we know that moving to risk adjustment causes the average risk score of enrollees to rise: $\frac{db^*}{d\Omega} > 0$. Hence $\frac{d(b^*+v^*)}{d\Omega} < 0$. ■

Note that this result is true only for small changes in Ω evaluated at $\Omega = 0$. The empirical section of this paper and in the toy model, in contrast, involves moving Ω by a large amount, starting from 0. In this case, there is no guarantee that $\frac{d(b^*+v^*)}{d\Omega} < 0$. Nonetheless, this result highlights the fact that risk adjustment can cause overall cost selection to increase.

Firm Profits

In this section, we explore the effect of changing risk adjustment on firm profits.

Proposition 5. *Increasing Ω decreases a firm's average per enrollee profits so long as their enrollees are positively selected with respect to b .*

Proof. The simplest way to verify this claim is to use the envelope theorem. Write profits as a function of the amount of risk adjustment (Ω) and the characteristics of the average enrollees (b^*, v^*), which indirectly depend on Ω . The envelope theorem says that to understand how profits change with Ω , one can ignore how changing Ω influences the optimal choices of b and v . Recall from (11) that firm profits are given by

$$\pi(\Omega) = \underbrace{(1 - \Omega)\bar{p} + \Omega p(b^*(\Omega))}_{\text{capitation payment}} - \underbrace{(b^*(\Omega) + v^*(\Omega))}_{\text{FFS cost}} - \underbrace{c(b^*(\Omega), v^*(\Omega))}_{\text{screening cost}}$$

Differentiating with respect to Ω , and ignoring the dependence of b^* and v^* on Ω , we see that

$$\pi'(\Omega) = p(\cdot) - \bar{p} \tag{16}$$

As we showed in the proof of Proposition 2, this expression is negative as long as individuals are positively selected with respect to b . ■

Corollary. *So long as plans' enrollees are positively selected with respect to b , if overpayments increase after risk-adjustment, then plans' screening costs must also increase.*

Proof. From Proposition 5, we know that profits fall under these conditions. As such, the only way for overpayments to increase and for profits to fall is for screening costs to have risen. This result can also be shown analytically. ■

Data Appendix

Documents needed to calculate risk scores and capitation payments from FFS claims data

CMS provides the file mapping ICD-9 conditions to HCC categories at <http://www.cms.gov/MedicareAdvtgSpecRateStats/Downloads/RAdiagnoses.zip>. The model coefficients and algorithms can be found at <http://www.cms.gov/MedicareAdvtgSpecRateStats/Downloads/HCCsoftware07.zip>. To calculate final capitation payments, these risk scores are multiplied by “county benchmarks,” which are published annually in the Medicare Advantage “ratebooks,” and ratebooks from 1990 to 2011 are available at: <http://www.cms.gov/MedicareAdvtgSpecRateStats/RSD/list.asp>.

Capitation payments in the MCBS

Two pieces of information support our conclusion that after 2003 the MCBS capitation payment variable does not reflect variation in the risk scores. First, after risk adjustment when capitation payments were based on an individual’s risk score, there is extremely little variation in the capitation payments recorded in the MCBS for beneficiaries in the same age group, calendar year, gender, disability, Medicaid status, institutional status, plan, and county cells. For example, in 2004 and 2005, consider all individuals who are (1) enrolled in an MA plan in May of that year and (2) are in a cell (as defined above) with at least one other beneficiary in the MCBS. Of these more 1,000 individuals, more than 92 percent have capitation payments that are within \$1 of all other individuals in their cell. Second, using the actual risk scores provided to us by CMS, we show that individuals in the same cell (as defined above) who have *different* risk scores are recorded as receiving the *same* capitation payment. In 2006, the MCBS does not include plan identifiers, but the payment variable in the MCBS still does not appear to represent the actual amount of money an MA plan received. For example, there are twelve individuals who are enrolled in MA all months in 2006 and have exactly the same very low annual capitation payment (\$913.58). Yet these individuals have substantially different risk scores (one has a risk score of 1.03 while another has a risk score of 4.67) and different ages (one is 68 years old while another is 95). We speculate that the MCBS may not include capitation payments that reflect an individual’s risk score because such information would allow researchers to back out an individual’s risk score, a variable that is not included in the MCBS and that we needed to access directly from CMS itself. Nonetheless, as we show in Table 5, using the uncorrected capitation payments from the MCBS has little impact on our results.