

NBER WORKING PAPER SERIES

HOW DOES RISK SELECTION RESPOND TO RISK ADJUSTMENT? EVIDENCE
FROM THE MEDICARE ADVANTAGE PROGRAM

Jason Brown
Mark Duggan
Ilyana Kuziemko
William Woolston

Working Paper 16977
<http://www.nber.org/papers/w16977>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2011

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, 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, Princeton, RAND, Rice, Stanford, Wisconsin, Yale, and the NBER Health Care meetings for helpful comments and feedback. Views expressed here are solely those of the authors and not of the institutions with which they are affiliated, nor of the National Bureau of Economic Research, and all errors are our own.

© 2011 by Jason Brown, Mark Duggan, Ilyana Kuziemko, and William Woolston. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How does Risk Selection Respond to Risk Adjustment? Evidence from the Medicare Advantage Program

Jason Brown, Mark Duggan, Ilyana Kuziemko, and William Woolston

NBER Working Paper No. 16977

April 2011

JEL No. H51,I11,I18

ABSTRACT

Governments often contract with private firms to provide public services such as health care and education. To decrease firms' incentives to selectively enroll low-cost individuals, governments frequently "risk-adjust" payments to firms based on enrollees' characteristics. We model how risk adjustment affects selection and differential payments---the government's payments to a firm for covering an individual minus the counterfactual cost had the government directly covered her. We show that firms reduce selection along dimensions included in the risk-adjustment formula, while increasing selection along excluded dimensions. These responses can actually increase differential payments relative to pre-risk-adjustment levels and thus risk adjustment can raise the total cost to the government of providing the public service. We confirm both selection predictions using individual-level data from Medicare, which in 2004 began risk-adjusting payments to private Medicare Advantage plans. We find that differential payments actually rise after risk adjustment and estimate that they totaled \$30 billion in 2006, or nearly eight percent of total Medicare spending.

Jason Brown
U.S. Department of the Treasury
1500 Pennsylvania Ave, NW
Washington, D.C. 20220
Jason.Brown@do.treas.gov

Mark Duggan
University of Maryland
Department of Economics
3115L Tydings Hall
College Park, MD 20742
and NBER
duggan@econ.bsos.umd.edu

Ilyana Kuziemko
Princeton University
808 Columbus Ave., Apt 21D
New York, NY 10025
and NBER
kuziemko@princeton.edu

William Woolston
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
william.woolston@stanford.edu

1 Introduction

Instead of directly administering public services such as education and health care, governments have in recent years increasingly contracted with private firms or other organizations to provide these services. In the United States, charter school enrollment has tripled since 2000, and today the Medicare and Medicaid programs contract with HMOs and other private plans to cover nearly half of the programs' 90 million beneficiaries.

The main argument in favor of “contracting out” is that privatization can enhance the quality or efficiency of public services. However, a large body of research highlights the challenges often associated with writing a complete contract that aligns the incentives of both the government and the private firm (Hart *et al.*, 1997, Laffont and Tirole, 1993). For example, giving firms a fixed payment per enrollee creates strong cost-control incentives but can also encourage a firm to risk-select, or “cream-skim,” low-cost individuals. This concern is especially acute in the health care sector, where costs are highly skewed and thus create substantial scope for advantageous selection (Newhouse, 1996).

To address this problem, governments often turn to risk-adjustment models, which link the payment a firm receives from the government for covering an individual to her expected costs. Ideally, these models would base payments on factors exogenous to the firm—so that it is not reimbursed at the margin and thus has strong cost-control incentives—but would still explain a substantial share of individual cost variation. Such a model could remove firms' incentives to selectively enroll profitable customers and instead encourage them to compete on price and quality.¹

This paper examines how firms' incentives to risk-select change with risk adjustment and how this response affects the total cost to the government of providing the public service. We formulate a model that shows that if firms' risk selection does not change, introducing risk adjustment will typically reduce *differential payments*—the payment given to the firm to cover an individual minus the counterfactual cost had the government directly covered her. However, firms' risk-selection incentives do change with risk adjustment, and their response—decreasing screening efforts along dimensions included in the formula while increasing efforts along excluded dimensions—can more than offset the reduction in differential payments noted above. That is, risk adjustment can actually increase differential payments and thus

¹There is a large 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 how firms react to them is an important empirical question.

the total cost to the government.

We apply the insights of the model to the Medicare program, which since the 1980s has offered beneficiaries a choice between a traditional government-administered fee-for-service (FFS) program and private Medicare Advantage (MA) plans. MA plans, which now cover more than one-fourth of Medicare’s 47 million beneficiaries, must be offered at the same price to all Medicare beneficiaries in the plan’s geographic area of operation. However, past work has shown that despite these open-enrollment requirements, MA plans have nonetheless disproportionately enrolled low-cost individuals.² In response to this evidence, in 2004 the Medicare program began to base capitation payments to MA plans on an individual’s *risk score*, calculated using a formula that distills the medical conditions documented by health care providers into more than seventy disease categories, as well as adjusting for age, gender, and other demographic variables.

To the best of our knowledge, we present the first empirical analysis of the effect of Medicare’s risk adjustment on selection and differential payments. Using individual and claim-level administrative data for a sample of more than 55,000 Medicare recipients from 1994 to 2006, we find that after risk adjustment, risk scores among beneficiaries switching into MA plans rose relative to those remaining in FFS, consistent with firms reducing their efforts to select along characteristics *included* in the risk formula. Moreover, *conditional on the risk score*, the actual costs of those switching to MA fell after risk adjustment, consistent with MA enrollees being more intensely selected along dimensions *excluded* from the formula. Put another way, MA enrollees shifted from being low cost to being low cost conditional on their risk score. Indeed, after risk adjustment, those switching into MA plans are over \$1,200 “cheaper” than the risk-adjustment formula predicts them to be.

Turning from selection to differential payments, we begin by testing the prediction from our model that risk adjustment decreases differential payments if selection patterns do not change in response to the formula. Indeed, we show that applying the risk-adjustment formula to those switching into MA plans before 2004 would have decreased differential payments by over \$800 dollars per enrollee. However, differential payments actually rise by roughly \$1500 after 2003, even after removing additional payments that MA plans were given to ease the transition to risk adjustment.

In our model, plans incur screening costs to improve selection along different dimensions of health, but we do not specify how they do so. One possible selection strategy would be to differentially retain profitable enrollees by improving the quality of their benefits in ways the government does not observe or regulate. Throughout our sample period we find that,

²See, for example, Langwell and Hadley (1989), Physician Payment Review Commission (1997), Mello *et al.* (2003) and Batata (2004).

relative to FFS enrollees, MA enrollees' satisfaction with their health care is more positively correlated with their self-reported health. We interpret these results as consistent with MA plans dedicating resources to retain their healthiest (and presumably most profitable) enrollees.

Given the evidence we present suggesting that risk adjustment “didn’t work,” a natural question is if it can be improved. We close the paper by discussing the challenges inherent in improving the current risk-adjustment formula. In particular, we show that future recalibrations will likely lead to upwardly biased payments to MA plans. We also examine drawbacks to simply adding more detail to the formula.

Our results contribute to the literature on the unintended consequences of making contracts more specific. 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 and Baker, 1992) 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 information to determine prices can sometimes aggravate problems associated with asymmetric information, as in Einav and Finkelstein (2011).

Although we focus on the Medicare program, our results speak to the challenges facing risk adjustment more generally, an important public policy concern given that the use of risk adjustment is about to substantially increase. The 2010 Patient Protection and Affordable Care Act (ACA) requires risk adjustment in the individual and small group health insurance markets starting in 2014. In fact, the legislative language of the ACA suggests that risk adjustment in these markets may use similar “criteria and methods” to those used in Medicare Advantage.³ Beyond health insurance, recent work (e.g., Epple and Romano, 2008) has even suggested conditioning school voucher payments on students' ability, the educational analogue to conditioning capitation payments on disease scores.

Our results also highlight potential distributional concerns from privatizing social insurance programs. The selection and differential payment patterns suggest not only that Medicare may have overpaid private plans for the coverage of MA enrollees relative to what they would have cost in the FFS program, but also that resources may have shifted from relatively high-cost enrollees (who are more likely to remain in FFS) to relatively low-cost

³*Section 1343 — Risk Adjustment* of the ACA includes the following paragraph (b): “CRITERIA AND METHODS. The Secretary [of Health and Human Services], in consultation with States, shall establish criteria and methods to be used in carrying out the risk adjustment activities under this section. The Secretary may utilize criteria and methods similar to the criteria and methods utilized under part C [Medicare Advantage] or D of title XVIII of the Social Security Act.” Note that a 1965 amendment to the 1935 Social Security Act established Medicare, and thus Medicare reforms such as Medicare Advantage are actually amendments to the Social Security Act.

enrollees (who are more likely to switch to MA). Moreover, consistent with our satisfaction results, the Kaiser Family Foundation (2010) documents that MA contracts provide many up-front benefits that healthy individuals enjoy (e.g., 57 percent offer free or discounted gym memberships), but worse cost-sharing than does FFS for serious medical conditions. As such, compared to the FFS population, benefits *among MA enrollees* are more skewed toward the healthy. Both of these distributional effects will tend to diminish Medicare's ability to smooth the utility consequences of variation in health status. A large literature examines when governments should and should not contract out a given service (see Megginson and Netter, 2001 and Vickers and Yarrow, 1991 for reviews), and our results suggest that the extent to which the service serves as social insurance should influence that decision.

Finally, because of its sheer size, Medicare Advantage is of interest in its own right. Controlling Medicare costs is viewed as the most important factor in improving the long-term fiscal position of the U.S. government. We estimate that differential payments to MA plans in 2006 totaled about \$30 billion, or nearly eight percent of total Medicare spending that year. The MA program also appears to be at a crossroads: while there has been long-standing concern among some government agencies regarding the fiscal impact of the Medicare Advantage program and indeed the ACA cuts payments to some plans, the House of Representatives passed a budget measure on April 15, 2011 that would end FFS Medicare and have private plans cover all beneficiaries.⁴

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 to adjust capitation payments to MA plans. Section 3 presents the model. Section 4 describes the data. Sections 5, 6 and 7 present the empirical results on selection, differential payments, and satisfaction, respectively. Section 8 discusses the challenges to improving risk adjustment and Section 9 offers concluding remarks.

2 Background on Medicare Advantage capitation payments and risk adjustment

Since its inception in the mid-1960s, Medicare has provided coverage for enrollees' care in hospitals and other inpatient facilities through Part A of the program, while physician

⁴The proposed ACA cuts to MA payments echoes the recommendations of Medicare Payment Advisory Commission (2009), as well as the Commission's reports from earlier years. On the other side of the debate, the House bill, authored by Representative Paul Ryan, ends FFS Medicare for new enrollees starting in 2022 and replaces it with vouchers to purchase private health insurance as in the current MA program (see http://www.cbo.gov/ftpdocs/121xx/doc12128/04-05-Ryan_Letter.pdf for the Congressional Budget Officer's summary and preliminary analysis of the proposal). President Obama made opposition to this proposal one of the main themes of his April 13, 2011 address on the federal budget.

services and outpatient care are covered through Medicare Part B. Beginning in the 1980s, enrollees have had the choice between this traditional fee-for-service system and private MA plans. Plans must provide benefits generally comparable to Part A and B Medicare, but otherwise are free to coordinate patient care as they deem most effective and thus can have varying benefits, cost-sharing arrangements, and provider networks. Instead of reimbursing providers for MA enrollees' Part A and B services, the Medicare program pays MA plans a fixed capitation payment to cover these costs, and plans are thus the residual claimants if actual costs are above or below the level of the capitation payment. Since 2006, Medicare Part D has provided enrollees coverage for prescription drugs, though all of our analysis will focus on Part A and B, as these are the services MA plans are required to provide.⁵

The capitation payment to an MA plan for covering an individual is thus based on the estimated Part A and B payments had FFS Medicare covered her directly. The Centers for Medicare and Medicaid Services (CMS), the government agency that administers Medicare, has generally focused on adjusting payments along two dimensions of expected cost: individual attributes and geography.⁶ Individual attributes are used to generate an individual-level *risk score*. This risk score is then multiplied by a county-level *benchmark*, and capitation payments for individual i in year t in county c are given by: $Capitation\ payment_{itc} = Risk\ score_{it} \times Benchmark_{ct}$.⁷ Below we describe how the methodology for calculating risk scores has changed, as well as how county-level benchmarks have evolved.

2.1 Risk adjustment before 2004

Throughout the 1980s and 1990s, county benchmarks were generally set to 95 percent of county FFS costs (generally calculated based on a moving average over the past eight years to limit large fluctuations for small counties), as it was believed that MA plans should be

⁵MA 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. We discuss in the Conclusion how future work might examine the effect of Medicare Part D on the demand for MA plans.

⁶Since the 1980s, the names of both the MA program and the agency that administers Medicare have changed. For simplicity, we will always refer to private Medicare plans as "MA plans" and to the agency that governs Medicare as "CMS."

⁷County-level adjustment changed slightly in 2006, when MA plans were required to submit bids to CMS to cover the Part A and B costs of the average-cost individual in a given county. These were not, however, competitive bids in the traditional sense: instead of comparing MA plans' bids to each other and basing payments on the lowest bid made, CMS based payments on the difference between a plan's bid and the statutory county benchmark and plans that bid below the benchmark would receive 75 percent of the difference to spend on additional benefits for enrollees. Plans that bid at or above the benchmark would receive the statutory benchmark. CMS estimates this reform lowered MA payments by 3.4 percent in 2006, and, as described in the data section, we adjust our empirical estimates accordingly (Medicare Payment Advisory Commission, 2007).

able to deliver services more efficiently than FFS (Medicare Payment Advisory Commission, 2009). However, the 1997 Balanced Budget Act began to weaken the strict link between benchmarks and FFS costs, and beginning in the late 1990s benchmarks were raised beyond FFS costs in areas of low MA penetration in an effort to expand access to MA plans. By the end of 2003, benchmarks were roughly 103 percent of average FFS spending (Medicare Payment Advisory Commission, 2004).

During the 1980s and 1990s, CMS used a “demographic model” to generate individual-level risk scores, so-called because it included only demographic variables (gender, age, and disability and Medicaid status) as opposed to disease or health conditions. Then as now, CMS does not require MA plans to report cost or claims data (doing so might be seen as being in tension with giving MA plans’ freedom to coordinate care as they deem most effective), so it used the FFS population to estimate how each of these demographic factors contributed to average costs. Essentially, CMS regressed total Part A and B payments made on behalf of an FFS enrollee on the above demographic variables (including many interactions). CMS found that the model explained only one percent of the variation in payments among the FFS population (Pope *et al.*, 2004). Given the difficulties in out-of-sample prediction, it is unlikely that the model explained any more of the variation among the MA population, which, of course, is the population whose payments are actually determined by the formula.

Because of the low predictive power of the demographic model, there was substantial scope for MA plans to target enrollees who were healthier—and thus cheaper—than the demographic model would predict. Indeed, previous research has shown that during this period MA plans were able to attract patients who were far less costly than the FFS population in general or than predicted by the demographic model. Estimates suggest that individuals switching from traditional FFS to MA had medical costs between 20 and 37 percent lower than individuals who remained in FFS.⁸

Federal policymakers reacted to this evidence by enhancing the risk-adjustment procedure. In 2000, CMS introduced the principal inpatient diagnostic cost group (PIP-DCG) model. Due to the lack of MA cost data, the model used inpatient diagnoses (thus excluding outpatient and physician diagnoses) documented on FFS claims data to predict FFS costs the following year. As MA plans do not submit claims data, applying this model to the MA population required that MA plans submit “encounter data,” which document an enrollee’s diagnoses. CMS found that the PIP-DCG model could explain 6.2 percent of the variation in FFS costs.

⁸This 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).

Between 2000 and 2003, risk scores were calculated as a 90/10 blend of the demographic and PIP-DCG models: $Risk\ score = 0.9 * Demographic\ score + 0.1 * PIP-DCG\ score$. Thus, the introduction of the PIP-DCG model raised the portion of MA cost variation explained by risk scores from one to $0.9 * 1 + 0.1 * 6.2 = 1.5$ percent.

2.2 Risk adjustment after 2003

The difference between county benchmarks and FFS costs continued to grow after 2003. By 2009, benchmarks reached 118 percent of county FFS costs (Medicare Payment Advisory Commission, 2009).

In 2004, CMS introduced a more comprehensive risk-adjustment regime that is based on the hierarchical condition categories (HCC) model. Like the PIP-DCG model, the HCC model uses claims data from the FFS population to calibrate a model that predicts FFS costs in the following year, though the HCC model accounts for not just inpatient claims, but physician and outpatient claims as well. The model distills the roughly 15,000 possible ICD-9 codes providers list on claims into seventy disease categories.⁹ Initially, the HCC model was blended with the demographic model, with the HCC model accounting for 30, 50 and 75 percent of the total risk score in 2004, 2005, and 2006, respectively. From 2007 onward, risk scores have been based entirely on the HCC model.

CMS found that when FFS data are used to calculate HCC scores, the HCC score explains 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. Moreover, 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 the formula systematically under-predicts costs for those with above-average costs, and over-predicts costs for those with below-average costs.

Evaluation and application of the PIP-DCG and HCC models is complicated by the lack

⁹ICD-9 is an abbreviation for International Statistical Classification of Diseases and Related Health Problems Version IX. 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>. The full HCC model used 184 condition categories, though CMS elected to use the 70 with the most predictive power. The three most common conditions in our sample are “diabetes without complications,” “chronic obstructive pulmonary disease,” and “congestive heart failure.” County benchmarks are published annually in the Medicare Advantage “ratebooks” and ratebooks from 1990 to 2011 are all available at: <http://www.cms.gov/MedicareAdvtgSpecRateStats/RSD/list.asp>.

of cost or claims data from MA plans. Whether the model performs as well on the FFS population as it does on the MA population depends on at least two key assumptions: first, that the coding practices MA plans use in generating encounter data to record the conditions used in the formula are the same as those used by FFS providers on claims data; second, that differences in the MA and FFS populations can be fully accounted for by these conditions.

CMS has done extensive research related to the first assumption. They have found that MA plans exhibit greater “coding intensity” in documenting disease conditions, so that an MA enrollee’s risk score grows substantially faster than an FFS enrollee’s risk score (Center for Medicare and Medicaid Services, 2008).¹⁰ As risk scores are based on disease conditions from the previous year, the risk score of an enrollee the year he switches from FFS to MA is based on FFS provider claims data and is thus free of intensive coding. However, in subsequent years, risk scores are based on MA plans’ coding practices, which are far more aggressive than those assumed when the HCC model was calibrated.

The HCC model will also offer an upwardly biased estimate of the counterfactual FFS costs of MA enrollees if these enrollees are positively selected along dimensions not included in the model. In fact, the introduction of risk adjustment will incentivize firms to selectively target individuals who they expect to have low costs *conditional on their risk score*. The next section presents a formal model of how firms’ screening incentives change upon the introduction of risk adjustment.

3 Model

In this section, we model private insurers’ incentives to screen customers based on underlying cost and how these incentives change under risk adjustment. To focus on predicting how firms will react to the introduction of risk adjustment, we take as given the basic contours of the risk-adjustment formula used by CMS and described in the previous section, as opposed to exploring the optimal formula, as in Glazer and McGuire (2000) and others, though we discuss potential improvements to the current formula in Section 8.

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

¹⁰Note 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. The CMS study focuses on the difference in growth rates between those who stay in FFS for at least two years in a row and those who stay in MA for at least two years in a row, in order to eliminate the effect of compositional changes. One reason that MA risk scores might grow faster is that the health of MA “stayers” is deteriorating faster than that of their FFS counterparts. However, CMS explicitly dismisses this possibility (Center for Medicare and Medicaid Services, 2009).

at least some scope to selectively enroll individuals, whether through targeted advertising, designing benefit and cost-sharing parameters that appeal to some groups but not others, or more readily facilitating the enrollment of certain individuals. As such, we do not model the consumer side of the enrollment decision and instead focus on firms' decision to incur the costs associated with these screening activities in return for enrolling a selected subsample from the Medicare population. In our model, firms 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.

3.1 Basic framework and assumptions

3.1.1 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.¹¹ Both v and b are in units of absolute dollars.¹² 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 medical costs. Note that v is the residual with respect to the *government's* model; it need not be the residual with respect to the prediction of the individual herself or the private plan. We discuss the conditional variance of v in greater detail when we lay out our assumptions regarding screening costs.

In our baseline model, 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 firms 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.¹³

¹¹We return to the question of dynamics in Section 8 when we discuss recalibrating the risk-adjustment model over time.

¹²Note 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.

¹³Whether 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

3.1.2 Capitation payments and risk adjustment

Without risk adjustment, firms 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 $p''(\cdot) = 0$.¹⁴

Given that $\mathbb{E}(v|b) = 0$, total conditional expected costs m are given by $\mathbb{E}(b + v|b) = b$ and thus the government would like to ideally set $p(b) = b$. However, based on the evidence presented in Section 2, the government is not able to perfectly estimate b . In particular, individuals with higher (lower) risk scores are under- (over-) compensated. As such, we assume that while the government's estimate of costs grows with b it grows less than one-for-one, so $p'(b) \in (0, 1)$. Note that the main results on how selection responds to risk adjustment and the resulting effect on differential payments require only that $p' > 0$; assuming $p' < 1$ generates additional empirical predictions.¹⁵

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.¹⁶

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

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.

¹⁴In particular, our proofs of Proposition 1 (that risk adjustment causes selection to fall along the b margin and rise along the v margin), Proposition 2 (that there is positive selection along the b margin both before and after risk adjustment) and Proposition 4 (that the effect of risk adjustment on differential payments is ambiguous) do not depend on the linearity of $p(\cdot)$.

¹⁵Specifically, Proposition 1 and Proposition 4 do not depend on $p' < 1$. We provide a simple microfoundation for why p' will be typically less than one in the Appendix. In particular, we assume that the government observes a mismeasured version of b , which results in the standard errors-in-variables attenuation bias.

¹⁶As we discuss in Section 2, firms 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.

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

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

3.1.3 Screening costs

Though we discuss profit-maximization in greater detail shortly, firm profits are obviously a function of an individual’s cost $m_i = b_i + v_i$, and thus firms will have preferences over the b and v values of their enrollees. However, MA plans are required to accept any patient in their geographic coverage area who chooses to enroll, and thus selectively enrolling certain individuals in an open-enrollment regime will entail screening costs. As noted earlier, costs may include targeted advertising and recruiting, designing contracts that specifically appeal to one profitable group but whose terms make it unattractive to another *ex ante* profitable group, or even the risk of government sanctions for violating open-enrollment regulations.

We assume that the per capita screening cost c a firm 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 $\bar{v} = 0$). Enrolling individuals 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 v is measured in absolute dollars, it should be easier to find, say, a cancer patient with costs \$100 below expectation than someone without a single documented disease condition with costs \$100 below expectation.

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

$$\mathbb{E}(\pi) = (1 - \Omega)\bar{p} + \Omega p(b) - b - v - c(b, v). \quad (1)$$

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

3.2 Results

We begin with our main selection result, which characterizes how firms will react to a change in risk adjustment. The proof of this and all other results are in the Appendix.

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

- (i) *Firms decrease screening along the b margin and thus the average value of b among their enrollees rises (“extensive-margin” selection decreases).*
- (ii) *Firms increase screening along the v margin and thus the average value of v among their enrollees falls (“intensive-margin” selection increases).*

As risk adjustment makes capitation payments a positive function of b , firms 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 risk score, which is the government’s approximation 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.¹⁷ We show as a corollary to Proposition 1 that the effect of increasing risk adjustment on selection with respect to total costs $m = b + v$ is ambiguous—that is, it is possible that the increased intensive-margin screening along the v -margin can offset the decreased extensive-margin screening along the b -margin.

The next result demonstrates that increasing risk adjustment does not lead firms to indiscriminately enroll individuals with high b values.

Proposition 2. *For any $\Omega \in [0, 1]$ and any $p(\cdot)$ such that $p' < 1$, firms will enroll only individuals with $b < \bar{b}$.*

This result is the first to depend on $p' < 1$. As we show in the Appendix, if $p' < 1$, then with or without risk adjustment, individuals with $b > \bar{b}$ have capitation payments below their expected costs, and firms have the incentive to engage in screening to avoid such *ex ante* unprofitable enrollees.¹⁸

¹⁷The 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.

¹⁸Similar to the discussion in footnote 17, this prediction technically applies to b , whereas we can only observe $p(b)$. In a corollary to the proof, however, we show that $b < \bar{b} \Leftrightarrow p(b) < \mathbb{E}(p(b))$.

We show as a corollary to Proposition 2 that firms positively select with respect to overall costs $m_i = b_i + v_i$ both before and after an increase in risk adjustment. Of course, as shown in Proposition 1, the relative intensity of b versus v screening depends on the degree of risk adjustment, but overall positive selection holds for any $\Omega \in [0, 1]$.

We now turn to examining how differential payments change with risk adjustment. We begin by showing how increasing risk adjustment affects a firm's differential payments if selection is held fixed—that is, the firm's b and v (the average b_i and v_i among the firm's enrollees) do not change in response to a change in risk adjustment.

Proposition 3. *For $\Omega_0 < \Omega_1$, moving from Ω_0 to Ω_1 will always decrease differential payments if b and v are held fixed at their equilibrium values under Ω_0 .*

This proposition shows that risk adjustment reduces differential payments if selection remains unchanged. It yields an additional testable prediction—namely, that applying the risk-adjustment model to the MA population before the introduction of risk adjustment would have reduced differential payments.

While the results and corollaries so far allow us to derive empirical tests of the model, they do not necessarily have direct public-policy relevance. After all, the key test of risk adjustment from the government's perspective is not how it affects selection patterns (Propositions 1 and 2), or how it performs in the artificial scenario when firms cannot react (Proposition 3). Rather, it is whether risk adjustment reduces differential payments when firms can optimally change their screening patterns, a question our final result examines.

Proposition 4. *The effect of increasing Ω on a firm's average differential payment is ambiguous.*

On the one hand, risk adjustment reduces differential payments for enrollees with low b values—who, by Proposition 2, we know disproportionately join MA plans. On the other hand, plans react to an increase in risk adjustment by moving up the b distribution and down the v distribution, and the proof shows that these two effects can actually more than fully offset the decreased payments for low- b enrollees.

This result may be initially surprising, given that conditions were put on the risk-adjustment formula that would seem to make it effective in reducing differential payments. Recall that the capitation payment $p(b)$ is a positive function of b and thus a positive function of underlying costs when v is held constant. Furthermore, we require payment-neutrality, so the government is not merely handing extra money to private plans via risk adjustment. As shown in Proposition 3, these conditions are sufficient to reduce differential payments when firms do not change their selection practices in response. But even under these conditions, the endogenous response of firms can thwart the goal of lowering differential payments.

3.3 Discussion

Some assumptions or omissions of the model deserve further scrutiny. First, we do not model firm competition. Competition might lead firms to have to out-bid each other over profitable enrollees. One might thus model screening costs $c(b, v)$ as including this cost and thus as endogenous to the risk-adjustment regime.¹⁹

Second, we have assumed firms can only affect the *types*, but not the *number*, of people they enroll. As we show in the Appendix, endogenizing this margin does not change the results so long as the number of enrollees per firm is large relative to (i) a firm’s ability to enlarge its enrollment by changing b or v ; or (ii) the level of per capita profits. At least in relatively competitive settings—in which firms do not expand beyond some efficient size as instead excess demand is met by new firm entry and in which profits are minimal—these assumptions appear plausible. Moreover, empirically, the MA share of the Medicare population did not immediately change upon introduction of risk adjustment, supporting our baseline assumption.

In addition, this assumption allows us to focus on average cost and differential payments per beneficiary, which we can observe or construct using our individual- and county-level data on Medicare recipients, as opposed to firms’ total profits, which we cannot. We now turn to describing our main data sources.

4 Data

Our empirical work relies heavily on individual-level data from the Medicare Current Beneficiary Survey (MCBS) Cost and Use series from 1994 to 2006 (the most current year available). 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 each year.

Each year the MCBS follows a subsample of respondents for up to four years, thus creating a mix of cross-sectional and panel data. During our sample period, the data comprise over 55,000 unique individuals and over 150,000 person-year observations. From this sample, we

¹⁹How market power affects risk selection in the presence of adverse selection has only recently received much attention. As Olivella and Vera-Hernandez (2007) write, “most of the literature on adverse selection considers extreme cases: either perfect competition or monopoly.” Recent attempts to model the effects of market power generally focus on settings with two firms and two types of consumers (e.g., Olivella and Vera-Hernandez 2007, Biglaiser and Ma 2003 and Jack 2006). Empirically, Mehrotra *et al.* find that MA plan competition, proxied by MA market share, is positively correlated with plans’ use of advertising that targets healthy individuals.

make only minimal sampling 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 only become eligible for Medicare partway through the year, as their survey data are often incomplete.

Table 1 reports summary statistics, separately for MA and FFS enrollees. For the purposes of this table, we follow CMS and classify an individual as an MA enrollee in a given year if he spends at least half of his Medicare-eligible months in an MA plan, though in most of the empirical work we make use of the fact that we know the number of months an individual is enrolled in MA each year.²⁰ FFS enrollees are more likely to be on Medicaid or disabled, and, conditional on not being disabled, are roughly a year older. While there are no significant differences with respect to gender, MA plans attract a disproportionate share of Hispanics and, to a lesser extent, blacks, likely reflecting the fact that MA enrollment is concentrated in urban and suburban areas and still quite limited in rural and other non-metro areas. The share of MA respondents with annual income above \$20,000 (roughly the median of the sample) is about three percentage points higher than that of FFS respondents.²¹

The last variable in Table 1 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 out-of-pocket costs briefly in Section 7.²²

An important detail to discuss is that an individual's reported capitation payment in the MCBS is the *average capitation payment to an individual's MA plan*, not the capitation payment for the individual herself. As MCBS respondents are sampled so as to be representa-

²⁰See Center for Medicare and Medicaid Services (2009). While enrollees can switch mid-year, well over 90 percent of individuals we classify as being enrolled in MA in a given year spend all twelve months in MA.

²¹We do not report the average income because the survey topcodes above \$50,000.

²²One possibility we do not consider is whether changes to an individual's total Medicare expenditure affect his use of other government programs (federal or otherwise), and thus our results cannot speak to the effect on global government budgeting. For example, for individuals "dually" eligible for both Medicare and Medicaid, a decrease of \$1 in the generosity of Medicare benefits would likely save the government less than \$1 because such individuals would then rely more on the cost-sharing benefits provided by Medicaid.

tive of the Medicare population, the average *Total Medicare expenditure* for MA enrollees in Table 1 should equal in expectation *Total Medicare expenditure* for the average MA enrollee. As the Table reports, Medicare pays on average \$800 more to cover the average MA enrollee than the average FFS enrollee.²³

While the reporting of plan averages instead of individual payments still allows us to look at how average *Total Medicare expenditure* differs between MA and FFS, our empirical objective is to estimate how an individual’s Medicare expenditure changes depending on whether she is in MA or FFS. Comparing, say, *Total Medicare expenditure* the last year an individual is in FFS to the first year she is in MA would be a natural way to begin investigating how MA status affects costs to the Medicare program; unfortunately, given how MCBS actually records capitation payments, we would actually be comparing an individual’s last year in FFS with the average costs of all individuals in her MA plan the following year.

However, we are greatly aided by the fact that HCC scores are based on diagnoses documented on medical claims in the *prior* year, and that each year an individual is in FFS, the MCBS collects all her claims data. As such, we can calculate her HCC score, and thus her capitation payment, her first year in MA.²⁴ In analyses where we do these imputations, we generally limit ourselves to those individuals who were in FFS all twelve months of a baseline year, so that we can be sure we have their complete claims history that year. Claims data are not available once an individual is in MA, so we cannot do this imputation in subsequent years. As we discuss later, given the evidence described in Section 2 on “intensive coding” by MA plans, limiting our analysis to the change in Medicare expenditure the first year an individual switches to MA likely understates any increase in differential payments.

Because we can only calculate capitation payments the first year an individual switches to MA, much of the empirical strategy focuses on transitions from FFS to MA. Appendix Table 1 shows the number of observations who are in FFS in year t and MA 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.

²³Unless otherwise stated, all dollars amounts reported in the paper are adjusted to 2007 dollars using the CPI-U.

²⁴To calculate the capitation payment, we follow the HCC risk adjustment formula published on the CMS website to calculate the risk score and multiply it by the benchmark for the individual’s county, also published on the website (see footnote 9 for the actual links). CMS provided us the actual risk scores of a sample of MCBS respondents so we could check how accurately we imputed the risk scores: the correlation between actual and imputed risk scores is above 0.96. We then multiply the risk-score by the individual’s county-level benchmark (also see footnote 9 for links). We adjust this value downward by 3.4 percent in 2006 to account for a reform in which plans were allowed to submit “bids” to CMS. Footnote 7 provided further detail on this reform.

Given that we often rely on those switching from FFS to MA for identification, we use a long pre-period in order to examine a substantial number of transitions. As discussed in Section 2, while there were some changes in MA policy between 1994 and 2003, namely the 1997 Balanced Budget Act, the major reforms were to the setting of county benchmarks, not the calculation of individual-level risk scores. However, as a robustness check, we show the magnitudes of our results are generally unchanged when we limit the sample to years after 1997.

5 Changes in selection after risk adjustment

In this section we test the predictions from the model in Section 3 regarding selection into MA plans after risk adjustment. The model offers several testable predictions. First, after risk adjustment, extensive-margin selection should fall, thus leading to an increase in risk scores, at least for values of risk scores that are profitable in expectation. Second, intensive-margin selection should increase; conditional on risk scores, total Medical expenditure for those switching to MA relative to those staying in FFS should fall after risk adjustment. Thus, third, MA recipients are positively selected with respect to over cost both before and after risk adjustment.

5.1 Quantifying the selection incentives created by the HCC model

As discussed earlier in Section 4, we focus our analysis on those in FFS for all twelve months of the baseline year, which includes many who will switch to MA the following year. For each individual, we calculate two counterfactual capitation payments were they to indeed switch to MA the following year: the first based on original demographic formula and the second based on the HCC formula.

The first column of Table 2 compares these two capitation payments. The table includes only those in the pre-risk-adjustment period, before any selection endogenous to the HCC risk scores would be incentivized. Col. (1) presents the average difference between the HCC-based capitation payment and the demographic-based capitation payment, grouped by quintile of HCC score. Of course, 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 \$3,267 less than the demographic-based capitation payment for the lowest quintile of risk score, but it would pay \$7,419 more for the highest quintile of risk score.

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) shows that doing so would not always be

profitable. For example, individuals in the highest quintile of risk scores represent on average an expected \$4,261 loss to an MA plan, consistent with CMS’s own analysis showing that HCC capitation payments for the highest-cost quintile in FFS would cover only 86 percent of these individuals’ FFS costs (Pope *et al.*, 2004). More specifically, in our MCBS sample, those above a risk score of 2.0 (the 85th percentile of the distribution) would appear to be ex ante unprofitable. Thus, increasing risk scores indiscriminately would lead plans to enroll many beneficiaries who would be very unprofitable in expectation.

5.2 Evidence from individual-level data

5.2.1 Empirical strategy

To test whether firms react to the extensive-margin incentives depicted in Table 2, we estimate the following specification on the sample of individuals who are in FFS all twelve months of a given year t :

$$Risk\ score_{it} = \beta MA_{i,t+1} \times After\ 2002_t + \gamma MA_{i,t+1} + \delta_t + \epsilon_{it}, \quad (2)$$

where i indexes the individual, t the year, $Risk\ score_{it}$ is the individual’s risk score using year t claims data to predict Medicare expenditure in year $t + 1$, $MA_{i,t+1}$ is the share of her Medicare-eligible months that the individual spends in MA in year $t + 1$, $After\ 2002$ indicates the baseline year is after 2002, and δ_t is a vector of year fixed effects. Note that while risk adjustment begins in 2004, those in FFS in the baseline year of 2003 would be switching into MA the first year under the HCC model, and are thus part of the “post-period.”

Based on the results from the model, plans will only want to increase risk scores in regions where it is profitable. According to CMS’ own analysis of the risk formula (Pope *et al.*, 2004) and Table 2, individuals are still profitable if they fall in the first four quintiles of the risk-score distribution. As such, we expect that firms will attempt to increase risk scores more in the middle part of the distribution than in the right-tail of the distribution, a prediction we test using quantile regression.

To test the intensive-margin prediction, we estimate the following equation:

$$Expenditure_{it} = \beta MA_{i,t+1} \times After\ 2002_t + \gamma MA_{i,t+1} + \lambda Risk\ score_{it} + \delta_t + \epsilon_{it}, \quad (3)$$

where $Expenditure_{it}$ is the total FFS expenditure for individual i in year t and all other notation follows that in equation (2). We predict a negative coefficient on the interaction term—conditional on the risk score, after risk adjustment those enrolling in MA should have

lower Medicare expenditure relative to those remaining in FFS.²⁵

5.2.2 Results

Figures 1 and 2 show how the distributions of risk scores change after risk adjustment, for, respectively, those remaining in FFS and those switching to MA, where we define someone as switching into MA if they spend at least half of their Medicare-eligible months the following year on MA. While there is a slight increase in risk scores after risk adjustment among the FFS population, the increase among those switching to MA is substantially greater. As predicted, this upward movement is concentrated in the middle of the distribution: the median risk score among those switching to MA increases by nearly forty percent after risk adjustment. However, risk adjustment does not increase the right tail of this distribution and thus the increase in the mean is far more muted. The lack of movement in the right tail is not surprising, given that individuals in that part of the distribution remain unprofitable even after risk adjustment,

The first five columns of Table 3 report results from estimating variants of equation (2). The first column presents the results from a basic OLS regression. The coefficients suggest 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.

As noted earlier, we expect the effect on the mean to be muted, because risk adjustment should not have increased the incentive to enroll individuals in the right-tail of the risk-score distribution. Indeed, in col. (2), merely dropping observations with risk-scores above the 95th percentile (calculated separately each year) increases the magnitude and statistical precision of the estimate. Estimating a median regression (col. 3) on the entire sample increases the coefficient by nearly a third, so that sixty percent of the pre-period gap in risk scores between FFS and MA is made up by the increase in MA risk scores after risk adjustment.

We take col. (3) as our preferred specification—as it accounts for the fact that we do not expect to see much movement in the right tail of the risk-score distribution but also allows us to use the entire sample—and subject it to a number of robustness checks. First, col. (4) shows that the result holds even when we employ the noisier measure of MA status used in Figure 2: whether an individual is in MA for at least half of his Medicare-eligible months the following year. Indeed, conditional on being in MA for at least half of the year and in FFS all of the previous year, the average value for *Fraction of year on MA* is 0.83, so we expect the

²⁵This specification is similar in spirit to the “unused observables” test of Finkelstein and Poterba (2006). Using the terminology of their framework, $Expenditure_{it}$ in equation (3) 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.

coefficients in this specification to be smaller.²⁶ Col. (5) shows that restricting the sample so that 1998 is the earliest year an individual would switch into an MA plan does not change the results.²⁷

The next four columns of Table 3 investigate intensive selection, and test the prediction that costs conditional on risk scores for those switching into MA relative to those staying in FFS should fall after risk adjustment. Col. (6) estimates equation (3). After risk adjustment, individuals switching into MA are over \$1,200 “cheaper” than their risk-scores predict them to be.

Col. (7) shows that while the coefficient is predictably smaller in magnitude when the MA indicator variable is used instead, it remains negative and significant. Col. (8) shows that, as with the extensive-margin results, the coefficients of interest are unchanged when the more recent sample is used.

The fact that the main effect of MA in cols. (6), (7) and (8) is close to zero suggests that differences in risk scores accounted for essentially all of the cost differences between those joining MA and those switching to FFS in the pre-period. As risk scores were designed to address exactly these differences and were obviously based on data from the pre-period, the result is not surprising. However, once risk adjustment is instituted, MA plans have an incentive to enroll individuals who have low costs conditional on their risk scores, and the coefficients on the interaction terms suggest that enrollment patterns follow exactly that pattern.

A final prediction from Section 3 regarding selection patterns is that MA enrollees should be positively selected with respect to overall cost both before and after risk adjustment. Col.

²⁶Note that the corresponding figure among all individuals in MA for at least half of given year (not just those who had been in FFS the previous year) is 0.97, which is not surprising as those just switching in may have done so after January. Throughout the sample period, the rules on enrollment and disenrollment in MA plans went through mild changes. Besides being able to change plans each year during the annual “open enrollment period” the previous November, before 2002, individuals could switch between MA plans and FFS at any point in the year, in 2002 only during the first six months, in 2003-2005 only during the first three months, and in 2006 during the first six months. Throughout the sample period, institutionalized MA enrollees could return to FFS or switch to a different MA plan regardless of the time of year. At least in the MCBS, these reforms do not correspond to any change in the number of months spent on MA conditional on spending any months at all in a given year (results available upon request).

²⁷We also experimented with several other robustness checks, which we do not report in the interest of space but for which all results are available upon request. First, we estimate quantiles other than the median. While each of these regressions yields a positive coefficient on the interaction term, the coefficient is no longer significant beyond the 79th percentile. Interestingly, this quantile is very near the point above which both CMS and our own analysis identified enrollees as representing negative expected profits to plans. Second, we estimate OLS regressions separately on the sample of enrollees above and below the “break-even” profitability point (which we estimate as a risk-score of 2.0, see notes to Table 2), finding a positive and highly significant coefficient on those below 2.0 and a positive coefficient but with a p -value > 0.9 for those above 2.0. We prefer the median regressions, however, as they do not require us to split the sample on the dependent variable.

(9) shows this prediction holds by re-estimating (3) without including the risk-score control. Their last year in FFS, those individuals joining MA before risk adjustment have costs roughly \$2,850 below those who remain in FFS. If anything, that difference widens after risk adjustment, though the negative coefficient on the interaction term does not approach statistical significance.

5.3 Evidence on post-2006 selection from county-level data

One limitation of the results in Table 3 is that our data only cover the first three years of risk adjustment. We therefore look to county-level data through 2008 (the most recent year available) to see whether selection patterns during the 2004-2006 period look similar to those during the 2007-2008 period. The county-level data do not allow us to test the specific predictions regarding extensive- and intensive-margin selection. Instead, we merely seek to show that with respect to selection on overall costs, MA enrollees appear similar before and after 2006.

We follow Batata (2004) and examine how average per capita FFS costs in a county vary as a function of the MA share of its total Medicare enrollees.²⁸ Using ideas developed by Berndt (1991) and Gruber *et al.* (1999), Batata demonstrates that, if one assumes the distribution of spending across counties differs only with respect to the mean, then the estimate for β_1 in the following specification recovers the difference in average per capita FFS costs between the population switching between MA and FFS and those remaining in FFS, with j and t indexing county and year respectively:

$$\Delta \text{Per-capita Cost}_{jt}^{FFS} = \beta_0 + \beta_1 \Delta \ln(\text{FFS share of county}_{jt}) + \epsilon_{jt}. \quad (4)$$

The positive selection with respect to overall costs we saw throughout the period—documented, for example, in col. (9) of Table 3—would suggest that when a county’s Medicare population shifts from FFS to MA its average per capita FFS costs should increase, as more low-cost enrollees leave the FFS population, driving up the average cost of those remaining and yielding a negative estimate for β_1 . However, there are several reasons why the selection effect estimated via the county-level specification should be smaller than that estimated via the individual-level specification. First, the selection effect in Table 3 is iden-

²⁸Medicare county-level per capita FFS expenditures from 2004-2008 is from http://www.cms.gov/MedicareAdvtgSpecRateStats/05_FFS_Data.asp. Medicare county enrollment and MA enrollment from 2004-2005 is from http://www.cms.gov/HealthPlanRepFileData/02_SC.asp. Total Medicare county enrollment for 2004 - 2007 is from <http://www.cms.gov/MedicareEnrpts/>. Medicare MA enrollment for 2006 and 2007 is from <http://www.cms.gov/MCRAdvPartDEnrolData/> (and then click monthly MA state-county-contract). Medicare county enrollment and MA enrollment for 2008 forward is from <http://www.cms.gov/MCRAdvPartDEnrolData/> (and then click MA state-county).

tified by individuals switching from FFS to MA, whereas β_1 in equation (4) is identified by those switching from FFS to MA as well as those switching from MA to FFS. As we show in Section 7, individuals tend to switch from MA back to FFS when they are in poor health, so these enrollees are far less positively selected and thus we expect the estimate for β_1 to be substantially smaller than the selection effect presented in the final column of Table 3. Second, the MA variable in the county-level regressions is measured via a snapshot of enrollment in July of each year, whereas the individual-level data account for how much of the year each individual spends in MA, so the county-level regressions should exhibit greater attenuation bias. Finally, the magnitudes of the county coefficient would be further depressed if, as in Chernew *et al.* (2008), high MA penetration incentivizes cost-control practices among local providers and these practices spill over to how they treat their FFS patients as well.²⁹ For these reasons, our objective is not to compare the magnitudes of the individual-level and county-level specifications, but to determine whether the magnitude of the selection effect as measured by the county data appears to change after 2006.

Table 4 reports the results from estimating equation (4). The estimated coefficient on the change in the log FFS share in the first column suggests that those switching between FFS and MA have costs roughly \$575 lower than the FFS stock. Adding county-year trends in col. (2) increases the magnitude of the selection effect slightly. Cols. (3) and (4) show that the estimated selection is essentially unchanged after 2006. Though the standard errors are too large to make precise statements about the effect size, the point-estimates themselves are very close to zero. The coefficient of -\$94 in the specification without county-year controls suggests that, if anything, MA “switchers” in 2007 and 2008 are slightly more positively selected relative to the FFS stock than they were in 2004 to 2006. Adding county-year trends changes the sign of the coefficient, but reduces its magnitude to roughly \$25. We take the results in Table 4 as presenting no evidence that selection of MA beneficiaries with respect to overall Medicare expenditure levels changed after 2006.

5.4 Discussion

The results in this section provide strong support for the predictions of the model in Section 3 regarding selection patterns. First, firms have less incentive to engage in extensive-margin selection after risk adjustment, and as a result risk scores rise for those switching into MA relative to those remaining in FFS. Moreover, this increase only occurs in regions of the risk-score distribution that support positive expected profits, providing further evidence that firms are indeed reacting to incentives. Second, conditional on the risk score, those switching to MA have lower baseline FFS spending relative to those staying in FFS after risk

²⁹Nicholas (2009), however, finds that efficiency spillovers have not persisted in more recent years.

adjustment, consistent with their being more intensely selected along dimensions excluded from the risk formula. In the MCBS data, these two effects appear to roughly cancel each other out, leaving the positive selection with respect to overall baseline costs unchanged after risk adjustment.

Recall that our model shows that if firms increase their intensive-margin selection efforts, differential payments can actually rise after risk adjustment. While risk adjustment will always decrease differential payments if selection patterns do not change, the impact on differential payments when selection patterns do change is an empirical question, which the next section explores.

6 Did risk adjustment decrease differential payments to MA plans?

In this section, we focus on how an individual’s *Total Medicare expenditure* in a given year changes as he switches from FFS to MA. Recall from Section 4 that *Total Medicare expenditure* is equal to total reimbursements to providers if he is in FFS, total capitation payments to MA plans if he is in MA, or the total of the two if he spent time in both programs. 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.

We make two adjustments to capitation payments after 2003 to isolate the effect of the introduction of risk adjustment from other changes occurring around the same time. First, as noted earlier, the growth rate of county benchmarks (the baseline value, which, multiplied by the risk score, yields capitation payments) began to rise in the later years of our sample period, in some counties considerably so. We therefore calculate capitation payments holding the growth rate of each county’s benchmark to its pre-2004 level. Second, in the years immediately following the introduction of risk adjustment, plans received a so-called “budget-neutrality” adjustment (about a ten percent increase in capitation payments) to ease the transition to risk adjustment, and we 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.³⁰ Before examining whether differential payments fell after the introduction of risk adjustment in 2004, we explore whether they would have decreased had selection into MA plans held to its pre-2004 patterns, a key prediction from Section 3.

³⁰Also note that our model assumed that the average capitation payment were the entire Medicare population to join MA would be the same with or without risk adjustment, so removing these extra payments to plans brings the empirical work in line with the theory.

6.1 Would risk adjustment have worked had selection patterns not changed?

Figure 3 overlays the distribution of the year-on-year change in *Total Medicare expenditure* for those switching from FFS to MA between 1994 and 2003 and the distribution of that change from the same population *had the HCC model been in effect*. On average during the pre-period, actual Medicare costs increase by \$2,991 for the MA joiners but by only \$1,163 for the FFS stayers. The unconditional differential increase in Medicare expenditure is therefore \$1,828. When we instead simulate the capitation payments under the HCC model for those switching to MA, differential payments shrink by over \$800. This \$800 decrease remains when we condition for the large set of control variables listed in the notes to Table 5.

In short, had selection patterns not changed, we predict that the introduction of the HCC formula would have substantially reduced differential payments. Of course, given that the formula was calibrated on this population, this is a relatively undemanding test of the risk-adjustment model. Moreover, as Section 5 demonstrates, selection patterns changed substantially after risk adjustment—first, MA enrollees’ risk scores increased once capitation payments became a function of risk scores, and, second, conditional on risk scores, pre-period FFS costs for this group fell. As we showed in Section 3, whether the change in selection patterns can completely “un-do” the risk-adjustment model is an empirical question to which we now turn.

6.2 Did risk adjustment reduce differential payments after 2003?

6.2.1 Empirical strategy

As above, we begin with a 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:

$$Expenditure_{it} = \beta MA_{it} \times After\ 2003_t + \gamma MA_{it} + \lambda X_{it} + \delta_t + f(Expenditure_{i,t-1}) + \epsilon_{it}, \quad (5)$$

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. We prefer this specification to simply regressing $\Delta Expenditure_{it}$ as the lagged expenditure controls in equation (5) can better account for the fact that medical costs typically exhibit strong regression to the mean, though we show in Appendix Table 2 that

results using $\Delta Expenditure_{it}$ look, if anything, larger in magnitude.³¹

The coefficients β and γ estimate the change in total Medicare expenditure associated with an individual switching to MA relative to his having stayed in FFS. These estimated effects are consistent only if MA is uncorrelated with ϵ . This condition implies that, conditional on our control variables, the decision to join MA is not systematically related to time-varying shocks to an individual’s expected cost to the Medicare program. If, for example, individuals join MA when their health is deteriorating and thus their expected costs are rising, β and γ will be positively biased. However, when we return to endogeneity concerns later in this section, we argue that they generally bias results against finding differential payments to MA plans.

6.2.2 Results

The first column of Table 5 shows the results from merely regressing the change in Medicare spending on the MA variable—which is allowed to have a different effect before and after risk adjustment—and year fixed effects. Total Medicare expenditure increases by roughly \$2,222 when an individual switches from FFS to MA (for the entire year) before risk adjustment, and by an additional \$1,480 after risk adjustment.

The second column regresses the level of spending on the flexible function of past spending. The MA main effect is lower than in the first column, suggesting that some of the differential payments in the pre-period may in fact have been differential regression to the mean among those switching to MA. However, the coefficient on the interaction term grows slightly in magnitude. The third column adds controls for self-reported health in the previous year, as well as a large set of demographic and other controls, listed in the table notes. 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. That the coefficient on the interaction term increases by roughly a third suggests, as we hypothesized earlier, that selection endogeneity works against finding MA differential payments, at least in the post-period.

Col. (4) includes measures of current-year self-reported health. We prefer this specification over col. (3) as it better accounts for potential regression to the mean in health status—if

³¹The lagged Medicare expenditure controls include: lagged Medicare expenditure and quintiles 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). Note that regressing the *change* in spending is thus nested in the equation (5)—the two are equivalent if the coefficient on lagged expenditure is constrained to equal one and the coefficients on all other lagged expenditure variables are constrained to equal zero. Adding a greater number of lagged quantiles of positive Part A and B spending do not change the results, so for the sake of parsimony we use quintiles.

enrollees typically experience a deterioration in their health upon joining MA, then comparing current to previous year’s spending will over-state MA differential payments. We are aware, however, that health status will be 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.

Cols. (5), (6) and (7) subject the estimation in col. (4) to robustness checks. The coefficients predictably fall in col. (5) when the fraction of the year spent in MA is replaced by an indicator variable for having spent at least half of the year in MA, though both remain positive and highly significant. Winsorizing the data based on the 99th percentile in col. (6) or dropping years before 1998 in col. (7) leave the results largely unchanged.

6.3 Calculating the total value of differential payments

To fully measure the fiscal impact of MA enrollment, we use the actual payments to plans in col. (8) of Table 5, including the budget-neutrality payments and allowing county benchmarks to grow at their actual rate. While the coefficient on the main effect remains essentially unchanged from that in col. (4), the coefficient on the interaction term grows substantially.

We use these results to estimate total differential payments in 2006, when the MA share of the Medicare population was 16 percent and total Medicare enrollment was 43.3 million: $(840 + 3428) * 0.16 * 43.3 \text{ million} = \29.6 billion , or 7.7 percent of *total* Medicare spending on benefits in 2006.³² Of course, given the standard errors in the differential payment regressions as well as the fact that the coefficients are identified by individuals switching from FFS to MA and not the entire Medicare population, this calculation should be taken as merely our best estimate. Note, however, that whatever the level of differential payments in 2006, they have likely grown since that time, as the MA share of the Medicare population has risen substantially over the past five years.

6.4 Discussion

Given that our identification relies on those switching to MA from FFS, a natural question is whether the cost differences for this group are representative of the differences between the MA and FFS stock. For example, while readers may agree that we have indeed identified large differential payments the first year an individual is in MA, perhaps the cost-containment measures of MA plans slow cost growth thereafter, so that differential payments shrink after the initial year.

³²Medicare spending on benefits (i.e., excluding administrative costs) totaled \$385 billion in 2006 (in 2007 dollars, to match the coefficient units). See Kaiser Family Foundation (2007), which gives nominal estimates that we adjust with the CPI-U.

The MCBS does not allow us to answer this question, but, as documented earlier, CMS itself has provided evidence against this hypothesis (Center for Medicare and Medicaid Services, 2008). Specifically, they have found that the growth rate of risk scores in MA is faster than that in FFS. They attribute this phenomenon to “intensive coding”—enrollees in MA plans being diagnosed more aggressively than their FFS counterparts. Because risk scores in the first year that a beneficiary is enrolled in MA depend on conditions recorded when he is still enrolled in FFS, the effect of intensive coding can begin no sooner than an individual’s second year in MA. Given that the growth in risk scores is mechanically related to the growth in Medicare expenditure for the MA population, and that risk scores in the FFS population are, in expectation, equal to Medicare expenditure by the very construction of the risk score, this finding by CMS suggests that our estimates for new MA enrollees likely understates the average differential payments for the stock of all MA enrollees. 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.

Turning to another potential concern, while we mentioned earlier that endogeneity in equation (5) likely works against finding the results in Table 5, readers may wish for further evidence. Any bias story working in the opposite direction must argue that while those switching to MA appear relatively healthy and low-cost the year before they switch, they systematically have higher medical costs their first year in MA and thus would have also been expensive had they remained in FFS. These individuals might experience deteriorations in their health just after they switch or perhaps they put off an expensive medical procedure until joining MA.

We find such stories unlikely for several reasons. First, if those switching from FFS systematically experience a deterioration in health their first year in MA, then controlling for current-year health as we do in col. (4) should have substantially reduced the coefficients on the MA interaction and main effect. Second, individuals are unlikely to postpone expensive procedures until they join an MA plan because plans tend to have less generous cost-sharing arrangements for serious medical procedures than does FFS (Kaiser Family Foundation, 2010). In fact, we find that individuals who switch to MA were no less likely to have an eye exam their last year in FFS, even though vision coverage is generally more generous in MA.³³ Finally, consistent with there being no “Ashenfelter dip” the year before a switch to MA, when we control for the past two years of medical costs, instead of just one as in Table 5, the coefficients on the MA variables barely change, though become significant at only the ten-percent level due to the sample size falling by a half.

Considering the incentives MA plans face, these facts are not surprising. The least prof-

³³The MCBS asks about eye exams the past year but not other specific examples of medical check-ups.

itable enrollees for them in the post-period would be those who have little contact with the medical system their last year in FFS—and thus no documented HCC conditions and thus a low capitation payment—but suddenly become expensive their first year in MA. As we have shown throughout the last two sections, plans seem able to enroll the most profitable Medicare beneficiaries, though both the empirical work and model has not specified exactly how they do so. The next section seeks to shed light on a possible mechanism.

7 Enrollee satisfaction with their care as a function of health and MA status

The evidence in Section 5 shows that MA plans enrolled lower-cost individuals before and after risk adjustment. But how do they accomplish this selection, given that they must offer the same plans at the same rate to all Medicare beneficiaries in their geographical area of operation? We can imagine at least three strategies. First, they may target low-cost and thus more profitable individuals with advertisements or other recruiting efforts. Indeed, Kaiser Family Foundation (2008) concludes that advertisements for MA plans target healthy people and avoid the disabled population. Similarly, Bauhoff (2010) finds evidence via an audit study that highly regulated German health insurance firms respond more quickly to enrollment requests from respondents residing in low-cost areas of the country, and MA plans have far greater flexibility than do German firms. Second, they can design plans that disproportionately attract healthy individuals, a possibility that Glazer and McGuire (2000) and Ellis and McGuire (2007) explore for managed-care firms more generally. Finally, after individuals sign up for an MA plan, plans can treat healthy enrollees better than sick ones, so as to differentially retain the former group. Our data allow us to investigate this third hypothesis, though we think further work on these and other possible selection strategies is warranted.

7.1 Data and estimation strategy

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, we generally sample those who did not switch (either from FFS to MA or from MA to FFS) the previous year. Thus, unlike the majority of analysis so far in the paper, identification comes from cross-sectional variation—comparing individuals in MA with individuals in FFS. 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.

This sampling means we actually have little information on the medical spending of those in MA. Recall that after someone enters MA, the MCBS—and indeed the Medicare program itself—does not track their medical claims or costs, and without this information we cannot compute risk scores. As 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.

Instead, we focus on the fact that both before and after risk adjustment, MA plans enroll individuals who are healthier than average. Before risk adjustment, MA and FFS enrollees in our regression sample have mean self-reported health (from one, “poor,” to five, “excellent”) of 3.36 and 3.11, respectively. The differential shrinks slightly, to 3.34 and 3.13, after risk adjustment, but the change is not close to being statistically significant. Given the results in Section 5 that selection along pre-period spending did not change, this result is not surprising.³⁴

One way plans might achieve this selection is to make their sicker enrollees unsatisfied with their health care and thus more likely to switch out of their plan (either to FFS or to a different MA plan). We thus estimate the following equation:

$$Satisfaction_{it} = \beta MA_{it} \times Health_{it} + \gamma MA_{it} + \mathbf{H}_{it} + \lambda X_i + \delta_t + \epsilon_{it}, \quad (6)$$

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 might cause 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.

7.2 Results

Table 6 displays the results from estimating equation (6). 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.³⁵ The first row reports results when overall satisfaction serves

³⁴Of course, MA plans may have a treatment as well as a selection effect on enrollees’ health. Aizer *et al.* (2007) shows that the switch to managed care in California’s Medicaid program had deleterious effects on pre-natal care and infant health. Risk-adjustment could in fact decrease plans’ incentives to maintain enrollee health, as they now receive additional compensation if enrollees develop any of the conditions in the formula.

³⁵The 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

as the dependent variable. The MA main effect is negative—suggesting that someone of average health reports lower satisfaction in MA than in FFS. Of course, the type of person who joins MA might simply be harder to please—after all, FFS is the default and they chose to switch in the first place. As such, we do not take this coefficient to mean that MA plans in general deliver poorer services.

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, 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. In five of the nine categories (overall, out-of-pocket costs, doctor’s concern for your health, questions answered over the phone, and receiving information about your medical condition) the coefficient is significant, and a sixth (having medical care provided in the same location) has a p -value of 0.113.

We interpret these results as suggesting that MA plans focus resources on keeping their healthier enrollees relatively happier than their sicker enrollees. Any cross-sectional bias story would have to argue that MA plans differentially attract disgruntled sick people or people who become especially disgruntled when they are sick. Not only does such a story require very specific selection patterns, MA plans have no incentive to attract such individuals. We therefore believe our hypothesis, which is consistent with firms’ profit maximization, offers a more likely explanation of the results in Table 6.

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, or to FFS if he is currently 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.

The earlier results suggested that the probability of switching from FFS to MA was increasing in health, whereas this last result indicates that the probability of switching from MA to FFS is decreasing in health. As noted earlier, the county-level results are identified using both types of transitions, whereas the selection results in col. (6) of Table 3 are identified only by the first type of transition. The exit results are the basis for our earlier prediction that the county-level results would show less positive selection than the individual-level regressions based on those switching from FFS to MA.³⁶

asked in the earlier years.

³⁶One might assume that because those exiting MA do not appear particularly healthy relative to the FFS

7.3 Discussion

The results in Table 6 begin to shed light on how MA plans actually risk-select, which our model treated in a very reduced-form manner. However, given the limitations mentioned earlier, we feel this topic warrants further work. Bauhoff (2010) uses an audit design in his work on German insurance firms, and applying his methods to MA plans would seem to us an excellent first step. Moreover, while our work in this section focused on differences between MA and FFS enrollees throughout the sample period, focusing on how enrollment and recruiting strategies changed before and after the introduction of risk adjustment in 2004 would also seem like a fruitful area for future work. Given that Affordable Care Act will mandate risk adjustment in the entire small group and individual insurance markets in 2014, research on how firms actually risk-select is likely to have even greater public-policy relevance in the near future.

8 Can risk adjustment be improved?

In this section we briefly outline the challenges we see in improving risk adjustment, first in terms of future re-calibrations of the HCC model, and then with respect to adding more categories to the formula.

8.1 Recalibrating the model

Given evolution in medical treatments and research, the costs associated with diseases change over time. As such, the coefficients in any risk-adjustment model need to be recalibrated.

However, the combination of intensive selection and the lack of cost data once beneficiaries leave FFS makes recalibration especially difficult in the context of Medicare Advantage. Recall the model in Section 3, where costs are defined as $m_i = b_i + v_i$. As we showed in Proposition 1, risk adjustment will tend to increase the average value of b in the MA population, but decrease the average value of v . When the government wants to re-estimate costs conditional on b , they can only do so on the FFS population. But $E(v \mid b, FFS) > E(v \mid b)$, meaning that the government's estimate of $m_i = b_i + v_i$ will be biased upward for any value of b .

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.

In fact, recalibration will likely exacerbate mispricing. If certain categories tend to be over-priced, MA plans have a greater incentive to recruit individuals in that category. As those individuals will generally be positively selected, their leaving FFS drives up the average FFS cost in that category. Upon the next recalibration, the category is more mispriced than before, and MA plans will have ever heightened incentives to recruit individuals with this condition.

Alternatively, recalibration might benefit from using cost data from individuals in the MA population, which, as noted earlier, CMS does not currently collect. If, for example, MA beneficiaries with certain conditions are systematically less costly than their risk score suggests, the government could reduce capitation payments for these conditions. However, basing a plan's capitation payments on costs incurred recreates the very marginal-cost reimbursement the MA program was designed to avoid. The "yardstick" model of Shleifer (1985) suggests that capitation payments could be based on the submitted cost data of *other* plans, though he notes that collusion between plans could undermine the efficiency of such a system.

Note that whatever the advantages or disadvantages of relying on cost data from private plans to estimate risk models, it is this data source that risk-adjustment authorities will likely need to rely on to meet the ACA mandate that the non-elderly small-group and individual insurance markets be risk-adjusted by 2014. Because the insurance exchanges do not have a "public option," there will be no analogue to the FFS Medicare population, on which the government has extensive cost and claims data that it can use to estimate a predictive cost model.

8.2 Adding more categories to the formula

A natural reaction to the intensive-margin results in Table 3 is that the government should simply add more detail to the formula. The most obvious drawback to doing so is that it could provide firms more scope to "intensively code." If the only two conditions in a formula are heart attack and cancer, outside of actual fraud, MA plans cannot document that a patient has one of these conditions when he in fact does not. But "diabetes with complications" is far more open to interpretation; CMS reports that, relative to FFS, MA plans tend to interpret gray areas in a manner that results in higher risk scores. In addition, having extremely detailed categories would result in, essentially, paying MA plans on the margin for services performed, undercutting one of the primary rationales for the MA program.

A less obvious drawback is that having a more flexible HCC model would mean increasing the number of estimated coefficients for the risk-adjustment model. While the sample of individuals on which CMS can perform its estimation is large, it is not unlimited. When

choosing the complexity of a risk-adjustment model, CMS faces a tradeoff between adding more parameters to their model (and explaining more of the variance in costs) and measuring each coefficient precisely. Even if each coefficient is unbiased, having mismeasured coefficients can lead to large differential payments if MA plans are able to estimate the model more precisely than is CMS and attract patients with the over-priced conditions.³⁷

9 Conclusion

This paper began by modeling the effects of an effort to lower differential payments to private Medicare plans by risk-adjusting their capitation payments. We showed that plans respond to risk adjustment in two ways: they decrease their screening efforts along dimensions included in the model (“extensive-margin” selection falls), while increasing screening efforts along dimensions excluded from the model (“intensive-margin” selection rises). As the model demonstrates, these responses can actually thwart the government’s objective of decreasing differential payments.

We find support for the extensive- and intensive-margin predictions using individual-level data on Medicare expenditures. Moreover, we find that differential payments to MA plans are substantial and actually grew after risk adjustment. We estimate that in 2006 they totaled \$30 billion, eight percent of total Medicare expenditures. To put this number in a slightly different context, the Congressional Budget Office estimates that in 2016 the total cost of the Affordable Care Act’s insurance expansion provisions (e.g., the Medicaid expansion up to 133 percent of the poverty line and the subsidies to low-income individuals in the state insurance exchanges) will total \$114 billion.³⁸ Thus, the differential payments to MA plans are equivalent to over 26 percent of the total cost of extending insurance to roughly 30 million uninsured Americans.

As noted earlier, 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

³⁷To take an extreme example of how a very flexible risk-adjustment model can increase differential payments due to mismeasurement of the model’s coefficients suppose, for example, that CMS estimated a fully non-parametric version of the HCC model, where costs were estimated for every combination of the 70 HCC conditions. Such a model would have 2^{70} parameters, or trillions of times the number of FFS beneficiaries enrolled in a ten-year period. Here, many cells would have only one or two beneficiaries within them, and the corresponding coefficients would be measured with substantial error. Given the skewness of medical costs, some cells would be wildly over-priced, and going forward, MA plans would have strong incentives to attract individuals with these specific conditions, leading to large differential payments.

³⁸The CBO estimates the cost to be \$132 billion in 2016 dollars (see cost estimate for “insurance coverage provisions” in Table 1 (Estimate of the Effects on the Deficit of the Reconciliation Proposal Combined with H.R. 3590, as Passed by the Senate) of the CBO’s estimate of the final reconciliation legislation, which can be found at <http://www.cbo.gov/ftpdocs/113xx/doc11379/AmendReconProp.pdf>. We use the CBO’s forecasts for inflation to deflate this estimate to 2007 dollars, the units of the regression estimates.

how selection and differential payments change in response. The Medicare Prescription Drug Program (Part D) was established in 2006 and might interact with Medicare Advantage in interesting ways. Given that MA plans often provided drug coverage even before 2006, the extension of drug benefits to the entire Medicare population could 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.

Of course, MA differential payments are not pure waste and are likely passed on as profits to insurance plans and additional benefits to MA enrollees. Town and Liu (2003) estimate that between 1993 and 2000, the MA program generated over \$18 billion in consumer surplus, and Hall (2010) finds that between 1999 and 2002, annual consumer surplus surpassed \$12 billion. How total surplus is split between profits, administrative costs and consumer benefits likely depends on the level of competition in the local MA market, which we do not analyze and might indeed make for interesting future work. Town and Liu find for every \$1 of consumer surplus, the MA program generated nearly \$3 of firm profit. Lustig (2009) 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.³⁹

We close by returning to the potential distributional consequences of our results. Regardless of how the surplus described above is split, the MA program appears to expand the cost of Medicare while also transferring relative expenditures from the FFS population toward the financing of care for the MA population. As those switching into MA have, throughout the sample period, lower baseline costs and better self-reported health than do those remaining in FFS, the MA program transfers Medicare expenditure to those who likely have less need for it.

Moreover, as we show in Section 7, the gradient of satisfaction with one's health care is a more positive function of self-reported health for MA enrollees than FFS enrollees, consistent with MA plans treating their healthier (and thus more profitable) enrollees better so as to differentially retain them. Indeed, exit rates out of MA plans are differentially higher among those in poor health. Therefore, the MA program appears not only to transfer aggregate Medicare expenditures from the relatively higher-cost FFS population to the relatively lower-cost MA population, but it seems to effect a similar transfer *within the MA population*.

These results 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

³⁹Starc (2010) finds similar welfare losses due to imperfect competition in the Medigap market.

the literature (see, for example, Hart *et al.* 1997). In the case of, say, paving a road, the consequences of imperfect pricing would seem limited to whatever amount the government overpaid. With social insurance programs, however, imperfect pricing can induce private firms to cream-skin, exacerbating the utility consequences of the underlying inequality the program was initially intended to mitigate. At least in the case of Medicare, we find little evidence that risk adjustment has solved this problem.

References

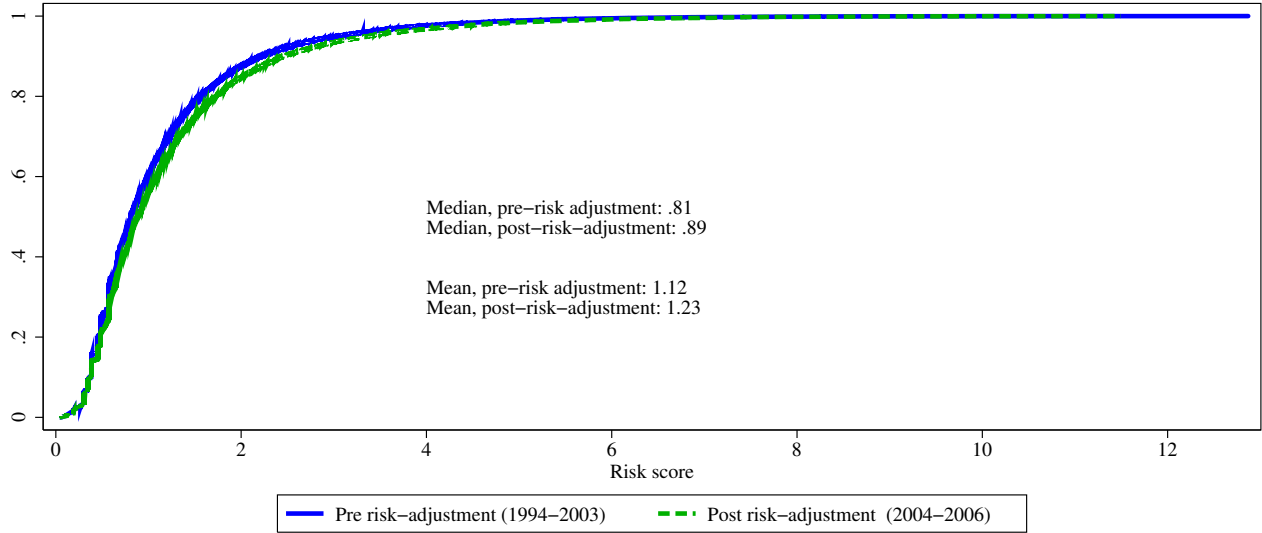
- AIZER, A., CURRIE, J. and MORETTI, E. (2007). Does managed care hurt health? evidence from medicaid mothers. *The Review of Economics and Statistics*, **89** (3), 385–399.
- BAKER, G. P. (1992). Incentive contracts and performance measurement. *Journal of Political Economy*, **100** (3), 598–614.
- BATATA, A. (2004). The effect of mhos on fee-for-service health care expenditures: Evidence from medicare revisited. *Journal of Health Economics*, **23**, 951–963.
- BAUHOFF, S. (2010). Do Health Plans Risk-select? An Audit Study on Germanys Social Health Insurance, mimeo, Harvard University.
- BERNDT, E. (1991). *The Practice of Econometrics: Classic and Contemporary*. Addison-Wesley.
- BIGLAISER, G. and MA, C.-T. A. (2003). Price and quality competition under adverse selection: Market organization and efficiency. *RAND Journal of Economics*, **34** (2), 266–86.
- CENTER FOR MEDICARE and MEDICAID SERVICES (2008). *Advance Notice of Methodological Changes for Calendar Year (CY) 2009 for Medicare Advantage (MA) Capitation Rates and Part C and Part D Payment Policies*.
- and MEDICAID SERVICES (2009). *Advance Notice of Methodological Changes for Calendar Year (CY) 2010 for Medicare Advantage (MA) Capitation Rates and Part C and Part D Payment Policies*.
- 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. S. (2005). How do hospitals respond to price changes? *American Economic Review*, **95** (5), 1525–1547.
- 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), 115–138.

- ELLIS, R. P. (2008). Risk adjustment in health care markets: Concepts and applications. In M. Lu and E. Jonsson (eds.), *Financing Health Care: New Ideas for a Changing Society*, Wiley-VCH, pp. 177–222.
- and MCGUIRE, T. G. (2007). Predictability and predictiveness in health care spending. *Journal of Health Economics*, **26** (1), 25–48.
- EPPLE, D. and ROMANO, R. (2008). Educational vouchers and cream skimming. *International Economic Review*, **49** (4), 1395–1435.
- FANG, H., KEANE, M. and SILVERMAN, D. (2008). Sources of advantageous selection: Evidence from the medigap insurance market. *Journal of Political Economy*, **116** (2), 303–350.
- FINKELSTEIN, A. and POTERBA, J. (2006). *Testing for Adverse Selection with “Unused Observables”*. NBER Working Papers 12112, National Bureau of Economic Research, Inc.
- FROGNER, B. K., ANDERSON, G., COHEN, R. A. and ABRAMS, C. (2011). Incorporating new research into medicare risk adjustment. *Medical Care*, **49** (3), 295–300.
- 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.
- GRUBER, J., LEVINE, P. and STAIGER, D. (1999). Abortion legalization and child living circumstances: Who is the marginal child? *Quarterly Journal of Economics*, **114** (1), 263–291.
- HALL, A. E. (2010). *Measuring the return on spending on the Medicare HMO program*. Finance and Economics Discussion Series 2010-31, Board of Governors of the Federal Reserve System (U.S.).
- HART, O., SHLEIFER, A. and VISHNY, R. W. (1997). The proper scope of government: Theory and an application to prisons. *The Quarterly Journal of Economics*, **112** (4), 1127–61.
- 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.
- JACK, W. (2006). Optimal risk adjustment with adverse selection and spatial competition. *Journal of Health Economics*, **25** (5), 908 – 926.
- KAISER FAMILY FOUNDATION (2007). *Medicare Spending and Financing Fact Sheet*.
- KAISER FAMILY FOUNDATION (2008). *Pitching Private Medicare Plans: An Analysis of Medicare Advantage and Prescription Drug Plan Advertising*.
- KAISER FAMILY FOUNDATION (2010). *Medicare Advantage 2010 Data Spotlight: Benefits and Cost-sharing*.

- LAFFONT, J.-J. and TIROLE, J. (1993). *A Theory of Incentives in Procurement and Regulation*. The MIT Press.
- LANGWELL, K. and HADLEY, J. P. (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. (2009). Measuring welfare losses from adverse selection and imperfect competition in privatized medicare, mimeo, Boston University.
- MEDICARE PAYMENT ADVISORY COMMISSION (2004). *Report to the Congress: Medicare Payment Policy*.
- MEDICARE PAYMENT ADVISORY COMMISSION (2007). *Report to the Congress: Promoting Greater Efficiency in Medicare*.
- MEDICARE PAYMENT ADVISORY COMMISSION (2009). *Report to the Congress: Medicare Payment Policy*.
- MEGGINSON, W. L. and NETTER, J. M. (2001). From state to market: A survey of empirical studies on privatization. *Journal of Economic Literature*, **39** (2), 321–389.
- MEHROTRA, A., GRIER, S. and DUDLEY, R. A. (). The relationship between health plan advertising and market incentives: Evidence of risk-selective behavior. *Health Affairs*, **25** (3), 759–765.
- MELLO, M., STEARNS, S., NORTON, E. and III, T. R. (2003). Understanding biased election in medicare hmos. *Health Services Review*, **38** (3), 961–992.
- NEWHOUSE, J. (1996). Reimbursing health plans and providers: Efficiency in production versus selection. *Journal of Economic Literature*, **34**, 1236–1263.
- NEWHOUSE, J. P., BUNTIN, M. B. and CHAPMAN, J. D. (1997). Risk adjustment and Medicare: taking a closer look. *Health Affairs*, **16** (5), 26–43.
- NICHOLAS, L. H. (2009). *Does Managed Care Reduce Federal Spending? Evidence from Medicare*. University of Michigan Population Studies Center 09-671.
- OLIVELLA, P. and VERA-HERNANDEZ, M. (2007). Competition among differentiated health plans under adverse selection. *Journal of Health Economics*, **26** (2), 233–250.
- PHYSICIAN PAYMENT REVIEW COMMISSION (1997). *Testimony of Gail R. Wilensky. Statement Before the US House of Representatives Subcommittee on Health. Committee on Ways and Means, February 24, 1997*.

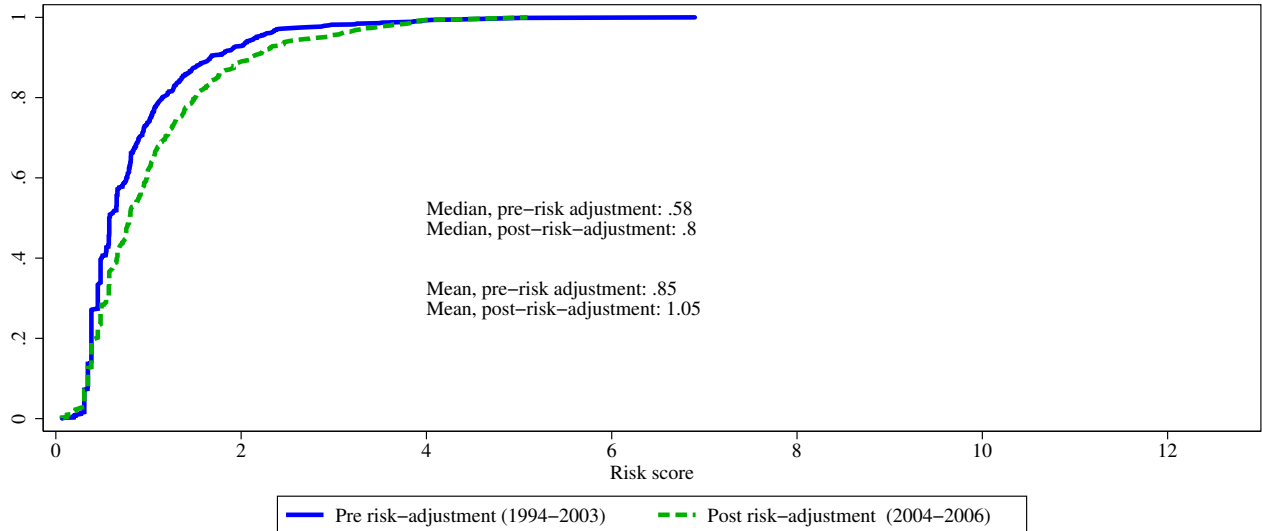
- POPE, G., KAUTTER, J., ELLIS, R., ASH, A., AYANIAN, J., IEZZONI, L., INGBER, M., LEVY, J. and ROBST, J. (2004). Risk adjustment of medicare capitation payments using the cms-hcc model. *Health Care Financing Review*, **25** (4), 119–141.
- POPE, G. C., ELLIS, R. P., ASH, A. S., AYANIAN, J. Z., BATES, D. W., BURSTIN, H., IEZZONI, L. I., MARCANTONIO, E. and WU, B. (2000). *Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment: Final Report prepared for Health Care Financing Administration*.
- SHLEIFER, A. (1985). A theory of yardstick competition. *RAND Journal of Economics*, **16** (3), 319–327.
- SILVERMAN, E. and SKINNER, J. (2004). Medicare upcoding and hospital ownership. *Journal of Health Economics*, **23** (2), 369–389.
- STARC, A. (2010). Insurer pricing and consumer welfare: Evidence from medigap, mimeo, Harvard University.
- TOWN, R. and LIU, S. (2003). The welfare impact of medicare hmos. *RAND Journal of Economics*, **34** (4), 719–36.
- VAN DE VEN, W. P. and ELLIS, R. P. (2000). Risk adjustment in competitive health plan markets. In A. Culyer and J. Newhouse (eds.), *Handbook of Health Economics*, vol. 1, Elsevier.
- VICKERS, J. and YARROW, G. (1991). Economic perspectives on privatization. *Journal of Economic Perspectives*, **5** (2), 111–32.

Figure 1: Cumulative risk score distribution for individuals remaining in FFS



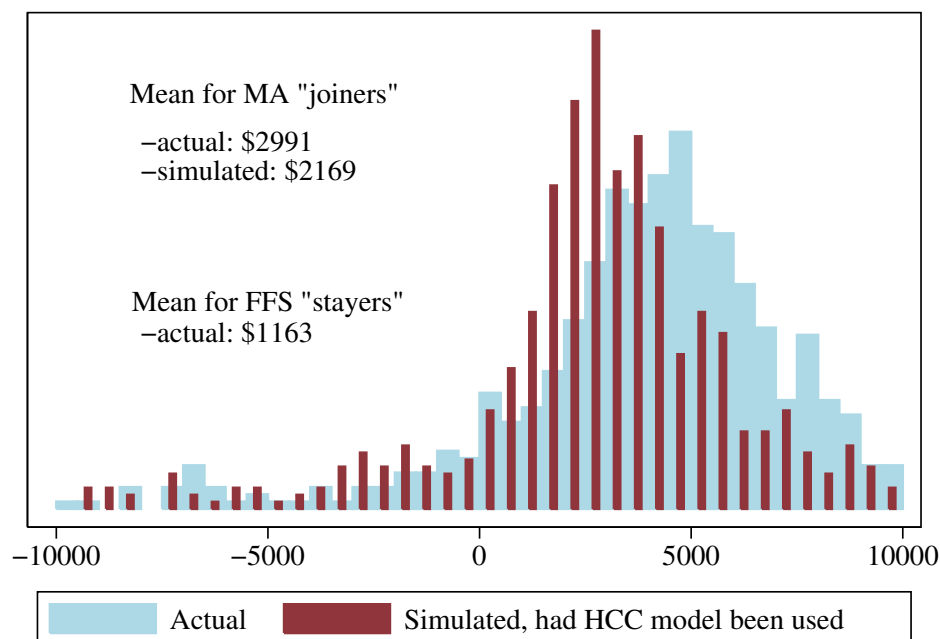
Notes: All observations spend the majority of their Medicare-eligible months in the baseline year in FFS and spent all twelve months of the previous year in FFS (so that a full year of claims data is recorded, which we use to calculate the risk score). Dollar amounts are adjusted to 2007 dollars using the CPI-U. Sample weights provided by the MCBS are used.

Figure 2: Cumulative risk score distribution for individuals switching to MA



Notes: All observations spend at least half their Medicare-eligible months in the baseline year in MA and spent all twelve months of the previous year in FFS (so that a full year of claims data is recorded, which we use to calculate the risk score). Dollar amounts are adjusted to 2007 dollars using the CPI-U. Sample weights provided by the MCBS are used.

Figure 3: Medicare expenditure increases for enrollees switching from FFS to MA, 1994 to 2003



Notes: The two histograms are based on observations that spend at least half of their Medicare-eligible months in the baseline year in MA and spent all twelve months of the previous year in FFS (so that a full year of claims data is recorded, which we use to calculate the risk score). The statistics for FFS “stayers” are based on observations that spend the majority of their Medicare-eligible months in the baseline year in FFS and also spent all twelve months of the previous year in FFS. Dollar amounts are adjusted to 2007 dollars using the CPI-U. Sample weights provided by the MCBS are used.

Table 1: Summary statistics, Medicare Current Beneficiary Survey 1994-2006

	FFS	MA	Difference: FFS - MA
On Medicaid	0.16	0.070	0.092***
Disabled	0.13	0.067	0.061***
Age	72.8	74.0	-1.12***
Age, those not disabled	76.2	75.4	0.88***
Female	0.57	0.57	-0.0012
Black	0.092	0.098	-0.0057**
Hispanic	0.018	0.034	-0.015***
Resides in a metro area	0.71	0.97	-0.25***
Self-reported health	3.12	3.35	-0.23***
Income > 20,000	0.45	0.48	-0.028***
Has a high school degree	0.75	0.74	0.0035
Total Medicare expenditure	7310.1	8044.8	-734.7***
Observations	131,339	19,320	

Notes: All observations taken from 1994-2006 MCBS. An individual is classified as in FFS if he spends the majority of his Medicare-eligible months in FFS and as in MA otherwise. Self-reported health varies from one (poor) to five (excellent). *Total Medicare expenditure* includes any capitation payments and Part A and B payments made by Medicare on behalf of an individual in a given year. Note that this table is based on all observations in the survey, instead of just those who are in the survey at least two years, a subsample used in much of the empirical analysis. Sample weights provided by the MCBS are used.

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

Table 2: Summarizing changes in incentives after risk adjustment

Quintiles of HCC score	HCC payment minus demographic payment	HCC payment minus actual Medicare expenditure
1 (lowest)	-3,267	1,391
2	-3,242	2,256
3	-2,027	2,494
4	-191	2,379
5 (highest)	7,419	-4,261
Total	-265	813
Observations	55,367	55,367

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. We only show five quantiles in the interest of space, but HCC payments minus actual Medicare expenditure becomes negative above the 85th percentile, or roughly a risk score of 2.0. 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 selection					Intensive-margin selection			
	Dependent variable: Risk score					Dependent var: Total Medicare expenditure			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fraction of next year spent in MA	-0.305*** [0.0355]	-0.180*** [0.0272]	-0.233*** [0.0325]		-0.238*** [0.0433]	171.5 [316.5]		191.6 [467.2]	-2847.0*** [396.9]
Fraction of next year in MA x After 2002	0.106* [0.0614]	0.116** [0.0509]	0.140*** [0.0507]		0.145** [0.0568]	-1217.9** [604.0]		-1280.3* [691.8]	-172.7 [713.4]
In MA majority of next year				-0.176*** [0.0272]			229.4 [253.4]		
In MA majority of next year x After 2002				0.0827* [0.0437]			-1045.3** [511.3]		
HCC score						9903.4*** [182.5]	9903.6*** [182.5]	9691.0*** [202.2]	
Mean of dept. var.	1.185	1.008	1.185	1.185	1.211	6491.9	6491.9	6630.7	6491.9
Estim. method	OLS	OLS	Q. reg.	Q. reg.	Q. reg.	OLS	OLS	OLS	OLS
Outliers trimmed	No	Yes	No	No	No	No	No	No	No
1997-2005 only	No	No	No	No	Yes	No	No	Yes	No
Observations	73,054	69,266	73,054	73,054	54,646	73,054	73,054	54,646	73,054

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) through (5) is an individual's HCC score the following year, which is based on current-year claims. The outcome in cols. (6) through (9) is an individual's current year total Medicare expenditure. "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 county per-capita FFS expenditure as a function of MA penetration

	Dept var: Δ per capita FFS expenditure			
	(1)	(2)	(3)	(4)
$\Delta \text{ Log(FFS share of county)}$	-575.9*** [146.0]	-621.1*** [162.9]	-535.9** [225.2]	-631.3*** [216.9]
$\Delta \text{ Log(FFS share of county)} \times \text{After 2006}$			-94.01 [279.6]	25.53 [344.4]
County fixed effects	No	Yes	No	Yes
Observations	15,532	15,532	15,532	15,532

Notes: The outcome variable in each regression is the change in the county's average per capita FFS expenditures. Data includes all U.S. counties (with the exception of those in Alaska and a few others with missing data on Medicare enrollment in one or more years) in each year from 2003 through 2008 (so the earlier data represent changes from 2003-2004 and the latest from 2007-2008). All specifications include year and county fixed effects, and are weighted by each county's share of the U.S. Medicare population in each year. Dollar amounts are adjusted to 2007 dollars using the CPI-U and standard errors are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Changes in differential payments after risk adjustment

	Dependent variable: Total Medicare expenditure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of months in MA	2222.0*** [397.2]	969.4*** [256.4]	646.6** [285.8]	839.0*** [311.5]		1291.1*** [278.1]	1050.9** [460.6]	838.9*** [311.5]
Share of months in MA x After 2003	1480.5* [889.1]	1642.3** [773.5]	2216.6*** [815.6]	2242.1*** [831.5]		1945.8** [758.1]	2050.4** [896.6]	3428.5*** [852.7]
In MA majority of year					759.8*** [261.6]			
In MA majority of year x After 2003					1688.7** [706.1]			
Dept var is in...	Δ s	Levels	Levels	Levels	Levels	Levels	Levels	Levels
Benchmarks adjusted	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Baseline controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	No	Yes	Yes	Yes	Yes	Yes
Dept. var windsorized	No	No	No	No	No	Yes	No	No
Only 1998-2006	No	No	No	No	No	No	Yes	No
Observations	73,054	73,054	72,638	72,375	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 all regressions after col. (2). All regressions in “levels” include a once-lagged dependent variable, as well as dummy variables corresponding to six bins of lagged Part A and B expenditure (with zero as its own bin and five bins corresponding to five quintiles 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 (8). “Baseline controls” include the following: fixed effects for the five categories of lagged self reported health (excellent, very good, good, fair, poor) indicators for gender, race and Hispanic origin; age-in-years fixed effects; income category fixed effects; marital status fixed effects; education fixed effects and fixed effects for eligibility status (disabled and old-age, with and without end-stage-renal disease as a secondary condition), Medicaid status, and the individual’s predicted capitation payment based on the demographic model. “Health controls” include the following: five categories of current self-reported health, the difference between current and previous-year self-reported health, and an indicator variable for being alive the entire year. The dependent variable is windsorized at the 99th percentile in col. (6). 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.	Enrolled in MA	MA x Health (demeaned)
Overall medical care	75,890	-0.0408*** (0.00777)	0.0144** (0.00661)
Out-of-pocket costs	75,315	0.0165* (0.00909)	0.0222*** (0.00745)
Follow-up care	69,770	-0.0191*** (0.00702)	0.00573 (0.00603)
Doctor's concern for your health	74,717	-0.0333*** (0.00748)	0.0151** (0.00634)
Information about your medical condition	75,545	-0.0245*** (0.00703)	0.0127** (0.00593)
Access to specialists	57,193	-0.0497*** (0.00772)	0.000594 (0.00646)
Questions answered over phone	48,622	-0.0279*** (0.00926)	0.0195** (0.00771)
Availability of care nights and weekends	44,507	-0.00920 (0.00916)	0.00726 (0.00796)
Medicare care provided in same location	69,386	0.0208*** (0.00695)	0.00868 (0.00574)
Retains coverage type next year (probit coefficients)	82,145	-0.545*** (0.0274)	0.0787*** (0.0213)

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. Note that this table is based on all observations in the survey, instead of just those who are in the survey at least two years, the subsample used in much of the previous analysis. 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 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.

Appendix Table 2: Differential payment regressions (Table 5 results, estimated in changes)

	Dependent variable: Δ Total Medicare expenditure						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of months in MA	2222.0*** [397.2]	2292.6*** [440.3]	2508.8*** [450.7]		2989.5*** [441.2]	2404.4*** [628.3]	2507.7*** [450.7]
Share of months in MA x After 2003	1480.5* [889.1]	1921.3** [923.3]	1872.5** [941.9]		1630.3* [895.4]	1936.9* [1036.1]	3059.1*** [954.6]
In MA majority of year				2093.2*** [376.0]			
In MA majority of year x After 2003				1441.3* [803.5]			
Benchmarks adjusted	Yes	Yes	Yes	Yes	Yes	Yes	No
Baseline controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	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	72,757	72,493	72,493	72,375	54,163	72,493

Notes: This table is the analogue to Table 5, except that the outcome variable is the change in total Medicare expenditure. 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 all regressions after col. (1). Standard errors are clustered by the individual. “Post-2003 adjustment” 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. (3) and (7). “Baseline controls” include the following: fixed effects for the five categories of lagged self reported health (excellent, very good, good, fair, poor) indicators for gender, race and Hispanic origin; age-in-years fixed effects; income category fixed effects; marital status fixed effects; education fixed effects and fixed effects for eligibility status (disabled and old-age, with and without end-stage renal disease as a secondary condition), Medicaid status, and the individual’s predicted capitation payment based on the demographic model. “Health controls” include the following: five categories of current self-reported health, the difference between current and previous-year self-reported health, and an indicator variable for being alive the entire 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Proofs of Propositions

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

- (i) *Firms decrease screening along the b margin and thus the average value of b among their enrollees rises.*
- (ii) *Firms increase screening along the v margin and thus the average value of v among their enrollees falls.*

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 firm's optimal levels of b and v .

The first-order conditions from maximizing the profit expression in equation (1) with respect to b and v are given by

$$[b] : \Omega p'(b^*) - c_b(b^*, v^*) = 1 \quad (7)$$

$$[v] : -c_v(b^*, v^*) = 1 \quad (8)$$

Totally differentiating equation (7) with respect to Ω yields:

$$p'(\cdot) + \Omega p''(\cdot) \frac{\partial b^*}{\partial \Omega} - c_{11}(\cdot, \cdot) \frac{\partial b^*}{\partial \Omega} - c_{12}(\cdot, \cdot) \frac{\partial v^*}{\partial \Omega} = 0 \quad (9)$$

Similarly, equation (8) yields:

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

or

$$\frac{\partial v^*}{\partial \Omega} = -\frac{c_{bv}}{c_{vv}} \frac{\partial b^*}{\partial \Omega}. \quad (10)$$

Substituting equation (10) into (9) 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:

$$\begin{aligned} \frac{\partial b^*}{\partial \Omega} = & \frac{\overbrace{p'}^{+ \text{ as cap payments increase in } b}}{\underbrace{-\Omega p''(\cdot) + \frac{(c_{bb}c_{vv}) - c_{bv}^2}{c_{vv}}}_{+ \text{ by convexity of } c(\cdot, \cdot)}} \quad (11) \end{aligned}$$

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 (10) gives the result in (ii). ■

Corollary 1. *The effect of increasing Ω on firms' overall insurance costs m is ambiguous.*

Proof. Differentiating total costs $m = c + b$ by Ω yields

$$\frac{\partial m}{\partial \Omega} = \frac{\partial b}{\partial \Omega} + \frac{\partial v}{\partial \Omega},$$

or, via substitution from equation (10),

$$\frac{\partial m}{\partial \Omega} = \left(1 - \frac{c_{12}}{c_{22}}\right) \frac{\partial b}{\partial \Omega}.$$

While Proposition 1 guarantees $\frac{\partial b}{\partial \Omega}$ is positive, $(1 - \frac{c_{12}}{c_{22}})$ can be positive or negative. ■

Proposition 2. *For any $\Omega \in [0, 1]$ and any $p(\cdot)$ such that $p'(\cdot) < 1$, firms will enroll only individuals with $b < \bar{b}$.*

Proof. It is sufficient to prove this claim holds for $\Omega = 1$ since we know from Proposition 1 that $\frac{\partial b^*}{\partial \Omega} > 0$.

The first-order condition for b in equation (7) with $\Omega = 1$ is $p'(b^*) - c_b(b^*, v^*) = 1$. As $p' < 1$, $c_b(b^*, v^*)$ must be negative. By assumption c_b is negative if and only if $b < \bar{b}$ (recall that screening costs grow as a function of distance from the mean), so the first-order condition only holds for b below $\bar{b}$. ■

Corollary 2. *For any $\Omega \in [0, 1]$ and any $p(\cdot)$ such that $p' < 1$, firms will enroll individuals with $p(b) < \mathbb{E}(p(b))$.*

Proof. From Proposition 2, firms always choose $b < \bar{b}$. By linearity of $p(\cdot)$, $\mathbb{E}(p(b)) = p(\bar{b})$, and since $p'(\cdot) > 0$, $p(b) < p(\bar{b})$ for all $b < \bar{b}$. ■

Corollary 3. *For any $\Omega \in [0, 1]$ and any $p(\cdot)$ such that $p' < 1$, firms always positively select with respect to total costs $m = b + v$.*

Proof. From Proposition 2, we know firms always positively select with respect to b so it is sufficient to show that they always positively select with respect to v as well. The first-order condition with respect to v is $-c_v(b^*, v^*) = 1$. So, c_v must be negative, and $c_v < 0$ iff $v < \bar{v}$. ■

Proposition 3. *For $\Omega_0 < \Omega_1$, moving from Ω_0 to Ω_1 will always decrease differential payments if b and v are held fixed at their equilibrium values under Ω_0 .*

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 $b < \bar{b}$. From Proposition 2 we know that $b < \bar{b}$ in equilibrium, so $p(b) < \bar{p}$ for any equilibrium b . ■

Proposition 4. *The effect of increasing Ω on a firm'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 firms are at their optimal b and v values:

$$\phi(\Omega) = \underbrace{\Omega p(b^*(\Omega)) + (1 - \Omega)\bar{p}}_{\text{capitation payment}} - \underbrace{(b^*(\Omega) + v^*(\Omega))}_{\text{actual costs}} \quad (12)$$

Differentiating with respect to Ω gives:

$$\phi'(\Omega) = \Omega p' \frac{\partial b^*}{\partial \Omega} + p(b^*) - \bar{p} - \frac{\partial b^*}{\partial \Omega} - \frac{\partial v^*}{\partial \Omega} \quad (13)$$

Rearranging and substituting $\frac{\partial v^*}{\partial \Omega} = -\frac{c_{bv}}{c_{vv}} \frac{\partial b^*}{\partial \Omega}$ from equation (10) yields

$$\phi'(\Omega) = [p(b^*) - \bar{p}] + \frac{\partial b^*}{\partial \Omega} (\Omega p' - 1 + \frac{c_{bv}}{c_{vv}}) \quad (14)$$

We showed in the proof of Proposition 3 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 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 (1) and q is the number of enrollees the firm has.

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

$$[b] : q_b(b, v)\pi(b, v, \Omega) + q(b, v) \underbrace{(\Omega p' - 1 - c_b(b, v))}_{\pi_b} \quad (15)$$

$$[v] : q_v(b, v)\pi(b, v, \Omega) + q(b, v) \underbrace{(-1 - c_v(b, v))}_{\pi_v} \quad (16)$$

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 (15) and (16) reduce to the original first-order conditions of $\pi_b = \pi_v = 0$.

Why $p(\cdot) \in (0, 1)$

Here, we explain why we assume that although the government's estimate of costs grows with b it grows less than one-for-one, or $p'(b) \in (0, 1)$.

First, assume that the government observes b and c for each individual i and wishes to create a payment function, $p(b)$ that best predicts c_i for each individual i . Since we assumed that $c = b + v$ and $\mathbb{E}(v|b) = 0$, if the government ran a regression of $c_i = \alpha_0 + \alpha_1 b_i + \epsilon_{1i}$, the government would find $\hat{\alpha}_1 \xrightarrow{p} 1$.⁴⁰ That is, without measurement error, the government would set $p'(\cdot) = 1$.

Now, the government observes only a mismeasured version of b , $\hat{b} = b + \varepsilon$, with ε independent of b and v . This ε comes from the fact that the government risk adjusters have only finite sample size, so it cannot run a totally non-parametric regression of all of the HCC indicators on costs. By imposing the HCC model, the government ensures that its composite measure of health (the risk score; $\hat{b}$) does not perfectly equal the composite measure of health (b). Given the know value $\hat{b}$, the government wishes to construct a function $\tilde{p}(\hat{b})$ that best predicts c_i for each individual i . If the government ran a regression of $c = \beta_0 + \beta_1 \hat{b} + \epsilon_{2i}$, the government would find $\hat{\beta}_1 \xrightarrow{p} \beta_1 < 1$ due to the measurement error in $\hat{b}$.⁴¹ So, we know that $p'(\hat{b}) < 1$. In particular, $p(\hat{b}) = \beta_0 + \beta_1 \hat{b}$. For our model, we want to know $p(b)$, how the

⁴⁰Mathematically, we know that

$$\begin{aligned} \hat{\alpha}_1 \xrightarrow{p} \alpha_1 &= \frac{\text{cov}(c, b)}{\text{var}(b)} \\ &= \frac{\text{cov}(b + v, b)}{\text{var}(b)} \\ &= \frac{\text{cov}(b, b)}{\text{var}(b)} + \frac{\text{cov}(v, b)}{\text{var}(b)} \\ &= 1 + \frac{\text{cov}(v, b)}{\text{var}(b)} \\ &= 1 \end{aligned}$$

⁴¹Mathematically, we know that

$$\begin{aligned} \hat{\beta}_1 \xrightarrow{p} \beta_1 &= \frac{\text{cov}(c, \hat{b})}{\text{var}(\hat{b})} \\ &= \frac{\text{cov}(c, b + \varepsilon)}{\text{var}(b + \varepsilon)} \\ &= \frac{\text{cov}(c, b)}{\text{var}(b + \varepsilon)} + \underbrace{\frac{\text{cov}(c, \varepsilon)}{\text{var}(b + \varepsilon)}}_{=0 \text{ since } \varepsilon \text{ is independent of } b \text{ and } v} \\ &= \frac{\text{cov}(c, b)}{\underbrace{\text{var}(b) + \text{var}(\varepsilon)}} \\ &\quad \text{no covariance term since } \varepsilon \text{ is independent of } b \\ &< 1 \text{ since the numerator is the same as above, but the denom is larger} \end{aligned}$$

average cap payment for a person with a given b . This equals

$$\begin{aligned}
 p(b) &= \mathbb{E}(\tilde{p}(b + \varepsilon)) \\
 &= \int_{-\infty}^{\infty} \tilde{p}(b + \varepsilon) \underbrace{f(\varepsilon)}_{\text{density of } \varepsilon} d\varepsilon \\
 &= \int_{-\infty}^{\infty} \beta_0 + \beta_1(b + \varepsilon) \cdot f(\varepsilon) d\varepsilon \\
 &= \beta_0 + \beta_1 \int_{-\infty}^{\infty} (b + \varepsilon) \cdot f(\varepsilon) d\varepsilon \\
 &= \beta_0 + \beta_1 b + \underbrace{\int_{-\infty}^{\infty} \varepsilon \cdot f(\varepsilon) d\varepsilon}_{\text{integrates to 0}} \\
 &= \beta_0 + \beta_1 b
 \end{aligned}$$

So, $p'(b) = \beta_1 < 1$.