NBER WORKING PAPER SERIES

SIMPLIFYING AND IMPROVING THE PERFORMANCE OF RISK ADJUSTMENT SYSTEMS

Thomas G. McGuire
Anna L. Zink
Sherri Rose

Working Paper 26736
http://www.nber.org/papers/w26736

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2020

This research was supported by a grant from the Laura and John Arnold Foundation and NIH Director's New Innovator Award DP2-MD012722. The views presented here are those of the authors and not necessarily those of the Laura and John Arnold Foundation, its directors, officers, or staff. We thank Konstantin Beck, Michael Chernew, Randy Ellis, Lukas Kauer, Tim Layton, Joseph Newhouse, Eran Politzer, Sonja Schillo, Richard van Kleef and participants at the Risk Adjustment Network Meeting in Portland, Maine, September 24-26, 2019 for helpful comments. We are grateful to Tram Nham for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Simplifying and Improving the Performance of Risk Adjustment Systems
Thomas G. McGuire, Anna L. Zink, and Sherri Rose
NBER Working Paper No. 26736
February 2020
JEL No. I11,I13,I18

ABSTRACT

Risk-adjustment systems used to pay health plans in individual health insurance markets have evolved towards better "fit" of payments to plan spending, at the individual and group levels, generally achieved by adding variables used for risk adjustment. Adding variables demands further plan and provider-supplied data. Some data called for in the more complex systems may be easily manipulated by providers, leading to unintended "upcoding" or to unnecessary service utilization. While these drawbacks are recognized, they are hard to quantify and are difficult to balance against the concrete, measurable improvements in fit that may be attained by adding variables to the formula. This paper takes a different approach to improving the performance of health plan payment systems. Using the HHS-HHC V0519 model of plan payment in the Marketplaces as a starting point, we constrain fit at the individual and group level to be as good or better than the current payment model while reducing the number of variables called for in the model. Opportunities for simplification are created by the introduction of three elements in design of plan payment: reinsurance (based on high spending or plan losses), constrained regressions, and powerful machine learning methods for variable selection. We first drop all variables relying on drug claims. Further major reductions in the number of diagnostic-based risk adjustors are possible using machine learning integrated with our constrained regressions. The fit performance of our simpler alternatives is as good or better than the current HHS-HHC V0519 formula.

Thomas G. McGuire
Harvard Medical School
Department of Health Care Policy
180 Longwood Avenue
Boston, MA 02115
and NBER
mcguire@hcp.med.harvard.edu

Sherri Rose
Harvard Medical School
Department of Health Care Policy
180 Longwood Avenue
Boston, MA 02115
rose@hcp.med.harvard.edu

Anna L. Zink
Harvard Medical School
Department of Health Care Policy
180 Longwood Avenue
Boston, MA 02115
azink@g.harvard.edu

1. Introduction

Germany, The Netherlands, and Switzerland, among other countries, purvey health insurance to their residents through individual health insurance markets. In the U.S., Medicare Advantage and the Marketplaces created by the Affordable Care Act (ACA) of 2010 do the same. In all of these countries and sectors, plans charge premiums, and a regulator combines them with public subsidies and disperses funds to plans using a "risk adjustment" formula. Switzerland and the Marketplaces apply forms of risk sharing for high-cost cases to the payment formula, and Germany and The Netherlands are scheduled to add risk sharing to their plan payment formula. Broadly, these mechanisms seek to match plan payments to plan costs to improve the functioning of health insurance markets by minimizing incentives related to adverse selection.

Reflecting their salience to health policy, the past 30 years have seen a great deal of research on risk adjustment and health plan payment, including theoretical research linking risk adjustment to economic objectives of plan payment; empirical research on classification systems and estimation methods; empirical evaluation of actual systems; and simulation methods for evaluating the functioning of these payment models (McGuire and Van Kleef 2018). A primary and concrete outcome of the empirical research has been improved "fit" of payments to plan costs at the individual and group level. Fit at the individual level is predominantly measured by the R^2 of the prediction model, and fit at the group level by comparing average plan costs to average payments for groups of interest, such as those with a certain illness.[1] Fit of the payment models has improved over the years in all countries and sectors mentioned above. For example, Medicare's risk-adjustment system for paying private health plans evolved from 1) using only demographic data, to

[1] In the U.S. the group measure is frequently the ratio of average payments to average costs, referred to as the "predictive ratio." In Europe the group measure is typically the difference between average payments and average costs, referred to as "over or undercompensation."

2) adding diagnostic data from hospital claims, to 3) adding diagnostic data from office-based encounters. The R^2 of these models improved from about 1% in the early years to 11% in the current payment model (Pope et al. 2011).

Fit has another attraction as a target for improvement: it is measurable, with cardinal metrics. The magnitude of reduction in undercompensation for persons with cancer falling from $2,000 per year to $1,000 per year, for example, can be readily understood and appreciated. At the same time, researchers and policymakers recognize that gains in fit have come at a cost in terms of data demands, administrative complexity, creation of opportunities for gaming the system by "upcoding" clinical data used for payment, and inducing provision of extra services to achieve higher risk scores. While these side effects are acknowledged, there is no consensus on how to measure them.[2] In the tradeoff between improvements in fit and reduction in incentives for cost control, for example, a regulator must weigh measurable improvements in fit against a speculative degradation in incentives for upcoding and excess costs.

This paper takes an unconventional approach to improving the performance of health plan payment systems. Rather than treating fit as an objective, we treat the level of fit in the existing payment system as a constraint. The ready measurability of fit is critical to this approach, allowing us to formulate explicit and exact constraints on fit. By treating fit as a constraint, we are free to consider payment formula alternatives that can be compared on the basis of demands on data, complexity, and adverse incentives to upcode or provide unnecessary services, even if these adverse side effects cannot be precisely measured. In other words, this paper studies the potential of

[2] Researchers are able to estimate the magnitude of upcoding, but not, so far, the contribution of particular variables or variable sets. For estimates of the magnitude of up coding in Medicare Advantage see Jacobs and Kronick (2018) who compare changes in reported risk score to risk estimated from drug data (not used for payment in Medicare Advantage, and Geruso and Layton (2020) who compare counties with different shares of Medicare Advantage enrollment and how this affects the county-average risk score.

payment models for reducing demands on data, risks of upcoding, and risks of encouraging unnecessary service use, while matching or exceeding existing performance in terms of fit at the individual and group level. Using the complicated HHS-HHC V0519 model of plan payment in the Marketplaces as our application, we estimate individual and group-level fit with the current formula, and then impose these as constraints in our consideration of simpler plan payment formulas.

Based on recent research, we believe our plan has promise for two main reasons. First, improving fit by addition of risk adjustor variables has hit the "flat of the curve" in many countries and sectors.[3] A corollary to the finding that adding variables adds little to fit is that removing variables does little to reduce fit. Thus, attaining fit by methods other than variable addition will allow scope for dropping sets of risk adjustor variables. Modern methods for variable selection are available to help guide this task. Second, directing only a very small share of plan payments to individuals imposing the largest losses on plans improves fit substantially. While such reinsurance has an incentive effect, inclusion of gameable variables in a payment formula has incentive effects too. Pairing reinsurance with dropping sets of variables may therefore improve incentives while improving or maintaining fit at current levels.

Section 2 sets up our analyses with a review of the health plan payment system used in the Marketplaces. We explain the data, methods, and estimation results for what we refer to as the Baseline Formula that is in current use. Section 3 describes our measures of fit at the individual and group level. For group fit, we begin with constraints on fit for individuals with one of four chronic illnesses defined by Clinical Classification Software (CCS) categories: cancer, diabetes, heart disease, and mental illness. Section 4 explains the three complementary components of our approach to improvement of risk-adjusted payments. The first is to consider alternative estimation approaches

[3] We also demonstrate the flat-of-the-curve property empirically in the Marketplaces below by assessing fit with alternative numbers of risk adjustor variables chosen by variable selection methods.

tailored to ensuring group-level fit: regressions subject to group-level constraints or penalties.[4] Paired with least squares estimation methods, constraints lead to coefficient estimates that satisfy constraints at the least cost in terms of fit at the individual level. A second key feature we leverage in conjunction with these estimators is reinsurance. Even the modest reinsurance applied here, amounting to one percent of total payments, substantially improves fit at the individual level (McGuire, Van Kleef and Schillo 2020), opening opportunities to improve the payment model in other dimensions. Third, we implement variable reduction strategies – both those informed by practical factors such as availability or cost as well as data-adaptive procedures from the machine learning literature.

Section 5 presents our empirical results, with some of the most notable summarized here. We show that the set of risk adjustors drawn from drug claims introduced in the 2018 Marketplace payment formula can be dropped entirely if fit is aided by devoting 1% of funds to reinsurance. Indeed, fit at the individual level is greatly improved by the substitution of reinsurance for drug variables. Furthermore, by introducing constraints on the regression coefficients so as to completely eliminate over/undercompensation for the four chronic illness groups, under/overcompensation is also reduced for nearly all 17 major chronic illness groups defined by the CCS. We also show that individual and four chronic illness group fit can be maintained or even improved with fewer than half of the current 111 risk-adjustor variables drawn from the current classification system.

Our results have two main implications. First, major improvements in fit at the individual and group level are attainable with a small share of funds devoted to reinsurance and a small set of coefficient constraints. Second, application of these methods alters the tradeoff inherent in decisions about adding or dropping variables to create wider scope to simplify payment models and

[4] For example, constraints have been implemented for risk adjustment in Van Kleef et al. (2017), Bergquist et al. (2019), and Zink and Rose (2019), where the latter also proposed penalized regressions for risk adjustment.

drop risk adjustor variables with potentially adverse incentive effects. Improvements in the hard-to-measure but important aspects of health plan payment system performance – administrative complexity, demands on providers and plans to produce data, incentives to upcode and provide unnecessary services – can be had without sacrifice of the easy-to-measure objectives of fit.

2. Baseline Marketplace Payment Formula

Creation of the Marketplaces (originally, "Exchanges") under the ACA called for construction of a health plan payment model to offset incentives for risk selection within the individual and small group insurance markets. ACA plan payment is based on a risk adjustment model to determine risk scores of enrollees and a transfer formula to redistribute premium revenues within the market in response to the risk score and other factors (Layton, Montz and Shepard 2018).[5] The payment model in the Marketplaces is arguably the most complicated in use anywhere. Our focus in this paper is the risk adjustment formula, leaving in place the other factors, such as adjustments for geographic costs and premium levels, playing into transfers among plans.

The Current Marketplace Formula: HHS-HCC V0519

The Marketplaces use the HHS-HCC formula for risk adjustment, developed by the federal Department of Health and Human Services (HHS), to predict enrollees' medical spending in the current year given enrollee demographics and a set of Hierarchical Condition Categories (HCCs) defined from diagnosis codes on medical claims. HCCs are sometimes grouped together (e.g., variable "G01" takes the value one if any of three diabetes-related HCCs are present). In total, 94

[5] The goal of the risk transfer formula is to remove anticipated premium costs attributed to risk selection while maintaining premium differences due to acceptable plan differences such as plan generosity and geographic costs. In addition to risk scores, the transfer formula considers metal level actuarial value, allowable rating factors, an induced demand factor, and a geographic cost factor. https://www.cms.gov/mmrr/Downloads/MMRR2014_004_03_a04.pdf

HCC-based variables – 75 HCCs plus 17 groups of HCCs plus 2 severity interactions explained below – are included in V0519 drawn from the underlying 128 HCCs contributing to the variables. The HHS-HCC Marketplace risk-adjustment model is "concurrent" in that it uses diagnoses from the current year to determine classification and payment, in contrast with using diagnoses from the previous year, as is done in the prospective risk adjustment systems used in Medicare Advantage and individual health insurance markets in Europe and elsewhere.[6] ICD-10 diagnosis codes are drawn from inpatient and outpatient claims involving a face-to-face meeting with a provider.[7] For example, diagnoses generated in a physician office visit are accepted for purposes of risk adjustment whereas diagnoses from a lab test are not.

The HHS-HCC V0519 formula was estimated on data from MarketScan for years 2014 and 2015; CMS then "blended" estimates from these two years to obtain the final formula. In addition, CMS, for the first time, had data available from plans (for 2016). These new data were considered, but according to CMS, taking them into account had no substantial effect on the coefficients estimated from MarketScan.[8]

The HHS-HCC formula has undergone several iterations since its inception in 2014, with HHS-HCC V0519 (2019), a slight modification of V0518 (2018), introduced for 2019.[9] In a major

[6] The rationale for the use of concurrent diagnostic information for determining a risk score in the Marketplaces was the high rate of anticipated turnover in the Marketplaces.

[7] ICD-9 diagnosis codes were used in the HHS-HCC until 2016.

[8] Regarding the final 2019 model coefficient estimates, CMS states, "The overall change in coefficients [in relation to the interim values reported prior to incorporating plan information] due to the update to the 2016 enrollee-level EDGE recalibration dataset is very small, and is further mitigated by blending of two other sets of solved coefficients from the 2014 and 2015 Truven MarketScan data." "EDGE" data is reported to CMS by the plans. See: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/2019-Updtd-Final-HHS-RA-Model-Coefficients.pdf.

[9] https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Updated-CY2018-DIY-instructions.pdf . V0518 split HCC037 (chronic hepatitis) into a 37a and 37b to distinguish types of hepatitis and, in a major change, included of risk adjustor variables based on prescription drug data. The HCC variables included in V05 are listed below in Appendix B where we present our empirical results for selected models.

earlier change, V0518 added 12 drug categories (RXC01 – RXC12) of which ten (RXC01 – RXC10) are used directly in the risk adjustment formula; with the other two used for HCC and RXC interactions only. CMS adopted a so-called "hybrid approach" in which RXC variables are included directly as risk adjustors, and, interacted with the HCCs with which the drugs are associated akin to a severity indictor.[10] The V0519 drops the RXC11-12 interactions. Drug variables are generated using National Drug Codes (NDC) from pharmacy claims with prescription filled dates within the benefit year (NDC from medical claims are not accepted).[11] Definitions of the recently introduced drug variables used for payment are contained in Appendix A.

Both V0518 and V0519 make extensive use of interactions among the HCC variables. Eight HCCs are categorized as "severe illnesses." If a patient has any of these eight severe illnesses, they receive a SEVERE flag. SEVERE interacts with 16 other HCCs or groups of HCCs to create 16 interactions, nine of which belong to a high-cost category and the other seven to a medium-cost category. The patient gets an additional variable added to their risk score for having any of the high-cost interactions or medium-cost interactions. If they have both then only the high-cost variable is added.[12] Finally, beginning with V0418 (2017), CMS introduced a variable measuring "months of

[10] Xu, Trish and Joyce (2019) describe the new drug risk adjustor variables and evaluate their incremental contribution to fit of the Marketplace payment model in a commercial health insurance data base from OptumInsight. They found an increase in R^2 from .38 to .40.

[11] When an NDC from a pharmacy claim is not available, HCPCS codes (Healthcare Common Procedure Coding System) from inpatient, outpatient, and professional medical claims with discharge dates or through dates within the benefit year can be used to create drug indicators. All our observations include drug coverage so we use only NDC codes to create drug variables.

[12] Example: someone who has both sepsis (HCC002) – a severe HCC – and metastatic cancer (HCC008) will get a flag for SEVERE_V3_x_HHS_HCC008, which is a high-cost interaction. The high-cost coefficient will then be added to their risk score. If they have sepsis and end-stage liver disease (HCC035), they will get a flag for SEVERE_V3_x_HHS_HCC035, a medium-cost interaction, and the medium-cost coefficient will be added. If they have all 3 HCCs, only the high-cost coefficient is added.

enrollment" during a contract year to contend with possible underpayment for those with partial enrollment periods.[13]

The HHS-HCC formula produces 15 sets of risk adjustment coefficients, three age-specific formulas (adult, child and infant) which include HCCs and RXCs most relevant for each age group, and five formulas specific for each coverage level in the Marketplaces (platinum, gold, silver, bronze, catastrophic). To calculate risk adjustment coefficients, relative annual plan liability is estimated using weighted OLS with age, sex, enrollment duration, and the selected set of HCCs and RXCs as predictors. Annual plan liability for each coverage tier is simulated based on annual expenditures and the relevant cost sharing schedule for that tier. To get relative plan liability, individual plan liability is divided by the average plan liability in the sample, using a predicted distribution of enrollment across coverage tiers.[14] The coefficients are adjusted post-estimation for clinical reasonableness: since the presence of a condition, or increased severity of a condition, presumes additional spending, all HCCs (or RXCs) should have a positive payment weight, and higher severity HCCs (or RXCs) should have a weight no less than lower severity HCCs (or RXCs) within the same hierarchy.[15] The final risk adjustment coefficients are multiplied by the input vector (age, sex, HCCs, RXCs, and interactions) for each enrollee to produce a risk score that feeds into the transfer formula for determining fund transfers among plans.[16] CMS reports the R^2 of V0519 estimated on MarketScan data from 2014 and 2015 to be 41%.[17]

---

[13] https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf. See pages 35-39.

[14] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4214270/pdf/mmrr2014-004-03-a03.pdf

[15] Ellis, Martins and Rose (2018).

[16] A new V06a has been proposed by CMS for payment years 2021 and beyond. See https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf.

[17] https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/2019-Updtd-Final-HHS-RA-Model-Coefficients.pdf.

In tandem with HHS-HCC formula changes, V05 restored a reinsurance function present for the first three years of the ACA by adding a high-cost adjustment to the transfer formula to cover 60% of insurer costs for claims greater than two million dollars, a feature funded by a charge based on insurer premiums.[18] Any improvement in fit conferred by reinsurance is not reflected in the CMS-reported R^2.

Data

We use the 2017 Truven MarketScan Commercial Claims and Encounters database to measure spending, HCCs and RXCs, and the 2019 HHS-HCC risk adjustment methodology to define the risk adjustment payment for each person.[19] The Truven MarketScan data, drawn from large insurers and employers, are a more recent version of the health insurance claims data used by Kautter et al. (2014) to develop the original HHS-HCC Marketplace payment system, and for subsequent revisions to the payment model. MarketScan data from 2017 were used to estimate payment weights for the Baseline Formula. Following HHS criteria, our analytic sample is composed of adults aged 21-64 who had both prescription drug and mental health coverage and who had no negative claims or claims with a capitation payment. In addition, we restrict our study population to those continuously enrolled for twelve months and who were in a non-HMO plan in the first and last month. After applying these exclusion criteria, we have 10,043,052 individuals. We then use a random sample of 5 million individuals for our analysis.

---

[18] https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/2018-RA-Model-DIY-Instructions.pdf. Layton and McGuire (2017) show this high-cost adjustment is mathematically equivalent to conventional reinsurance. In V05, risk adjustment weights are estimated without regard for the presence of this reinsurance function (which removes some costs from plan obligations). With the very high attachment point of $2m (where the reinsurance kicks in) in current Marketplace payment formula, correcting the dependent variable would make only a trivial difference in the estimated weights. With the more robust forms of reinsurance studied in this paper, it is "worth" correcting the dependent variable, plan spending prior to estimation. We do so below.

[19] 2019 HHS Risk adjustment software downloadable from here: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors-Items/RiskModel2019

The average annual health care expenditure is about $6,700, of which 56% is in the outpatient setting, 20% is in the inpatient setting, and the remaining 24% on drug spending. Table 1 reports characteristics of the 5 million enrollee sample in 2017 and shows mean spending for each. About 52% of the population is female and the average age is 44. Only a fifth of the population is coded as having any of the HCCs included in the V0519 model, and fewer with the RXCs. Mean spending for a person with an HCC or RXC flag is much higher than for those without. Chronic conditions were identified by mappings ICD-10 diagnoses to the latest version (2019.1) of the Clinical Classification Software (CCS) created by the AHRQ (2016).[20] We find that 12% of the sample have a mental health condition, 9% heart disease, 9% diabetes, and 7% cancer. Mean spending of someone with one of the chronic conditions is two to three times higher than the overall sample mean. This spending includes costs associated with other health care needs, in addition to the chronic condition. For example, someone with heart disease spends an average of $3,780 on visits addressing their heart disease, compared to an average of $20,154 for all spending, which can include other visits and drugs.

Estimation of the Baseline Formula

In order to compare the Baseline Formula to alternatives, we estimate a fresh set of risk adjustment coefficients rather than using those supplied by CMS for the V0519 formula. (The correlation between the predicted values of our estimated Baseline Formula with those from the CMS formula is 0.96.) We modify the procedures used by CMS to suit our data. First, our dependent variable is total spending rather than the estimated spending that would be covered by the various metal level plans; second, we estimate the adult model only (ages 18-64); third, we make no post-estimation changes to regression estimates; fourth, drug indicators come from NDC codes only, not from outpatient procedure claims; and finally, we do not include a months-of-enrollment

[20] https://www.hcup-us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp

variable or weight observations by enrollment because all included individuals in the data are observed for the full 12 months. The Baseline Formula is estimated by OLS and, as are the other models studied here, uses five-fold cross validation with a sample of 5 million individuals. The cross-validated R^2 from our Baseline Formula is 36.8%. Coefficients for the Baseline Formula are contained in Appendix B.

3. Measures of Fit

Fit of risk-adjustment systems is conventionally measured by an R^2 and by a comparison of predicted to actual costs for specified groups. We follow this approach, modifying the fit measures to recognize that plan payments may include some reinsurance as well as payments stemming from the risk adjustment formula.

Individual-Level Fit

R^2 is by far the most commonly reported fit statistic; indeed, in the CMS publication describing V0519, R^2 is the *only* fit statistic reported.[21] When plan payments are simply the predicted values from a risk adjustment regression, fit at the individual level is the R^2 from the risk adjustment formula.[22] Any contribution of reinsurance or other form of risk sharing to fit is captured by a generalization of the R^2 referred to as 'Payment System Fit' (PSF).[23] PSF is an R^2-type statistic (analogous to a pseudo-R^2) measuring the degree to which plan payments for individual i, R_i, track

[21] For example, reporting on final formula for 2019, see https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/2019-Updtd-Final-HHS-RA-Model-Coefficients.pdf. The only fit statistics reported are R^2 in Table 4.

[22] An R^2 is not the only individual-level fit measure found in the risk adjustment literature, but it is the most common one. R^2 is virtually the only individual-level fit statistic reported in US publications. European researchers also sometimes, along with the R^2, report Cummings Prediction Measure (CPM), a linear version of the quadratic R^2. See Van Veen et al. (2015) for a report of the characteristics and the frequency of alternative individual-fit measures.

[23] For other applications of payment system fit see Geruso and McGuire (2016), Layton et al. (2017), McGuire, Schillo and Van Kleef (2020), Beck et al. (2019).

spending for that individual, Y_i. PSF recognizes that the payment a plan receives for an individual, R_i, can include other components, such as reinsurance, in addition to the predicted spending from a risk adjustment model:[24]

$$\text{PSF} = 1 - \frac{\sum(Y_i - R_i)^2}{\sum(Y_i - \bar{Y})^2}. \qquad (1)$$

As noted, the estimated R^2 for our Baseline Formula is 36.8%. This is the same as PSF because payments are entirely determined by risk adjustment in the Baseline Formula.

Group-Level Fit

The risk of under and over service for certain groups of enrollees is well-recognized by architects of health plan payment systems. In Europe, incentives to serve certain groups (e.g., those with multiple chronic illnesses) is typically assessed by measuring over and undercompensation (in Euros) for a group. Researchers in the U.S. concerned with the same issue form a ratio rather than a difference between predicted values and costs. The groups for which fit is reported differ across studies, countries, and insurance sector.[25]

In principle, groups should be defined at the level of action a plan can take and fit evaluated at the group so-defined (Layton et al. 2017). If a plan cannot act on a distinction, then it is

[24] In the Marketplaces, the R^2 reported from the risk-adjustment model is not the same as payment system fit. Plan premiums, geography, share of high cost cases, and other factors play a role through the transfer formula. As far as we know, the overall fit of payments to plan costs have not been officially reported. PSF in the Marketplaces has been estimated taking into account some features of the transfer formula in Layton et al. (2017).

[25] For example, in the Netherlands, individuals reporting low health status or multiple chronic illnesses have been identified as potential targets for plan underservice (Van Kleef et al. 2013; Eijkenaar et al. 2017). In the U.S., researchers have studied users of particular classes of drugs (Carey 2017a; Carey 2017b; Han and Lavetti 2017; Geruso, Layton, and Prinz 2019), users of certain hospitals (Shepard 2016), users of certain types of services (Ellis and McGuire 2007; McGuire et al. 2014).

For evaluation of an early version of the risk-adjustment model used in U.S. Marketplaces, Kautter et al. (2014, E22) computed predictive ratios for subgroups defined by predicted costs. In their evaluation of the CMS-HCC model, Pope et al. (2011) report predictive ratios for groups defined by disease, numbers of prior hospitalizations, demographic characteristics, and others.

unnecessary and possibly misleading to calculate over/undercompensation for a group based on that distinction. For example, if in deciding about which hospitals to contract with for maternity services, a plan cannot introduce distinctions in contracting decisions for complicated and uncomplicated deliveries, there is no need to check fit (or include risk adjustor variables) for those two types of deliveries separately. By contrast, a plan might be able to act on geography; for example, by assembling a higher quality network of providers in areas where the enrollees are more profitable. In this case, checking group fit by geography is merited. We recognize that the level of action plans can take is not known, and we define groups at multiple levels of aggregation in our empirical work.

To begin, we assume that plans can make distinctions based on the nature of a chronic illness, for example, by contracting for better/worse networks according to the disease specialties of physicians and hospitals.[26] Specifically, following the recent evaluation of the group fit in the Marketplaces by Layton et al. (2017), we define groups according the presence of one of four chronic illnesses: cancer, diabetes, a heart condition, and mental illness.[27] CCS groups are used to identify beneficiaries with a chronic condition because, unlike the HCCs included in the risk adjustment formula, all ICD-10 codes map to a CCS condition group. Defining a chronic condition by CCS groups identifies more people because it includes individuals with diagnostic codes not recognized in the HCC indicators used in the risk adjustment formula.[28] CCS groups can be defined by single-level diagnoses and procedure codes, or multi-level diagnoses and procedures codes.

---

[26] For study of hospital network contracting in the Massachusetts Marketplace in response to selection incentives, see Shephard (2016).

[27] One or more of these four groups have been used for purposes of assessing group fit in Bergquist et al. (2019), Layton et al. (2017), Montz et al. (2016) and McGuire et al. (2014), Rose and McGuire (2019), Zink and Rose (2019).

[28] The gap in disease coverage can be significant. Montz et al. (2016) found that HCCs in use in the Marketplace payment formula captured only 20 percent of the individuals classified with mental health disorders according the broader CCS classification.

Single-level diagnoses map ICD-9 (or 10) codes to 285 mutually exclusive and, for the most part, clinically homogeneous categories. These single-level CCS codes (and their associated ICD-9 codes) can also be mapped to 17 multi-level diagnoses categories. For example, essential hypertension and hypertension with complications are considered two separate single-level diagnosis categories, but are included in the same multi-level category: "disease of the circulatory system."[29] Our four illness groups are identified using multi-level diagnosis codes.

We expand our concern with fit to other groups. Specifically, we track fit for the complete set of 17 other multi-level diagnosis-based CCS categories to study the effect of our payment formulas on additional groups not included in the payment formula. Appendix C contains a list of multi-level CCS categories. In light of our consideration of payment formulas without drug variables, we also keep track of the over/undercompensation for individuals with health conditions treated with the drugs used in the present risk adjustment formula.

The rationale for studying groups defined by a chronic illness is two-fold: first, membership in the group is persistent and predictable, making a plan's decisions about services for that group a potential tool to affect demand for enrollment by the group; and second, a disease grouping is a plausible level of action for a plan, for example by setting level of payment or choosing provider network design more or less favorably for different illness categories.

In parallel to our generalization of fit at the individual level to PSF, we generalize the ratio measure to be the total payments, not regression predicted values, over total costs for the group g. To distinguish our measure from the conventional predicted ratio, we refer to the measure including all payments in the numerator to be a Total Payment Ratio (TPR) for group g: [30]

[29] https://www.hcup-us.ahrq.gov/toolssoftware/ccs/CCSUsersGuide.pdf
[30] McGuire et al. (2014) modify predictive ratios incorporating premium differences and risk sharing in the U.S. Marketplaces. Geruso, Layton, and Prinz (2019) modify predictive ratios and under/overcompensation measures in the same way.

$$TPR_g = \frac{\sum_g R_i}{\sum_g Y_i}, \quad (2)$$

where R_i is total revenue a plan receives for person i and Y_i is plan spending for i. A TPR less than 1.0 means the revenues for members of that group fall short of plan costs, giving the plan incentives to discriminate against members of this group. Conversely, a TPR above 1.0 indicates that a group is profitable to plan giving the plan incentives to overserve the group. TPRs near 1.0 are preferred so as to neutralize incentives based on selection at the group level.

4. Simplifying and Improving the Payment Formula While Maintaining Fit

Our goal to simplify and improve risk adjustment while maintaining fit involves three complementary components. First, we consider alternative estimation approaches specifically tailored to group-level fit. Regressions subject to group-level constraints or penalties have been previously proposed for improving group fit in risk adjustment, and we deploy several here.[31] Another key feature we leverage in conjunction with these estimators is reinsurance.[32] Lastly, we implement variable reduction strategies – both those informed by practical factors as well as data-adaptive procedures from the machine learning literature.[33]

Constrained and Penalized Regressions

We consider two forms of constrained regression estimators. In the first, a linear constraint restricts the estimated coefficients such that, for each of the four chronic illness groups, the TPR is

[31] See van Kleef et al. (2017), Bergquist et al. (2019), and Zink and Rose (2019).

[32] As in McGuire, Schillo and Van Kleef (2020).

[33] Many machine learning methods can be implemented to reduce the number of variables considered. See, for example, James et al. (2013) for a treatment of these procedures. The techniques have been used for variable selection in health spending applications previously, including the random forests algorithm for risk adjustment in Buchner, Wasem and Schillo (2015) and Rose (2016). Lasso penalized regression is used for predicting unprofitability enrollees based on prescription use in Rose, Bergquist, and Layton (2017).

set at the value from the Baseline Formula estimated by OLS in Table 2. For example, the TPR for cancer was constrained to be equal to 0.91. Alternatively, we set the constraints such that the TPR = 1.0 for each of the four groups, ensuring that payments equal plan spending at the group level. A TPR of 1.0 improves the group fit in relation to that estimated in the Baseline Formula. In the presence of goals to fit at both the individual and group level, constrained regressions attain target group-level fit at the least cost in terms of fit at the individual level.

Constrained regressions can be equivalently formulated as penalties, and because penalties for risk adjustment have demonstrated promising performance elsewhere (Zink and Rose 2019), we also considered penalized regression with tuning parameters optimized toward group fit.[34] To our knowledge, this is the first application of penalized regression for multiple groups in risk adjustment. These alternative estimators were combined in various ways with the reinsurance and variable selection procedures described below to maintain group-fit performance.

Reinsurance

We explore two different implementations of reinsurance. In both, we set aside 1% of funds to cover 80% of plan spending above an attachment point. Conventional reinsurance payments defined on spending (e.g., covering 80% of plan spending over $500,000) can be quantified prior to risk adjustment estimation. We also consider the more targeted reinsurance based on plan losses (spending less payment, e.g., 80% of loss at the person level above $300,000). Loss-based, also referred to as "residual-based" reinsurance requires estimating residuals at the person level in a first empirical step, followed by solving for the threshold based on the funds available for reinsurance.[35]

[34] This led to λ values of 10,000, 0, 10,000, and 100,000 for the cancer, diabetes, heart disease, and mental health groups. For more information on penalized regression tailored to group fit in risk adjustment, we refer to Zink and Rose (2019).

[35] Residual-based reinsurance was proposed in Schillo et al. (2016), and explored in a Marketplace context in McGuire, Schillo and Van Kleef (2020). McGuire, Schillo and Van Kleef (2020) also simulates residual-based reinsurance in Germany and The Netherlands.

Residual-based reinsurance uses funds more efficiently in terms of individual fit than conventional reinsurance by targeting high losses as opposed to high spending per se (which in some cases is picked up by the risk adjustment formula).

Variable Selection

We estimate two major variants of the Baseline Formula. The first set drops all RXC (drug) variables, both main effects and interactions. The second set of formulas additionally drops some HCC variables using data-adaptive machine learning. There are a number of reasons dropping the drug variables should be considered. First, because of the frequency of prescription drug use, drug files are often the largest of any category of service. In our data, there are over 79 million new drug prescriptions (not refills), more than the number of outpatient visits (74m), and inpatient admissions (0.7m) combined. Dropping the drug-related variables eliminates the need to consult the enormous drug claim files for purposes of risk adjustment. Second, dropping drug variables improves incentives. CMS and others recognize that increasing payment in response to a filled prescription creates incentives to overprescribe.[36] Third, new drugs appear continuously, exposing the formula to be in need of constant refinement of variable definitions and recalibration. Fourth, drugs included in plan formularies are not the same even for plans within the same regulatory structure. And finally, using drug data for purposes of payment may decrease the elasticity of demand for a drug at the plan level, undermining the bargaining leverage plans have with respect to drug manufacturers, and increasing procurement prices.

[36] Xu, Trish and Joyce (2019) note incentive concerns with the RXC variables. For discussion of concerns from CMS perspective, see HHS-operated risk adjustment methodology meeting: Discussion paper. Retrieved from https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf. The originators of the drug-based risk adjustors have expressed concerns about incentives and proposed reforms in the way drug data are used in risk adjustment (Lamers and Vliet, 2003). One suggestion was to increase the required daily dose before a risk adjustor was recognized. The Marketplace risk adjustor is activated with the first prescription.

We also rely on machine learning methods to guide removing HCCs in the risk adjustment formula after having already dropped the drug variables. Here we use a lasso penalized regression within the cross-validation folds of the constrained or penalized regressions to select a maximum of 20, 30, 40, or 50 HCC variables while keeping all age and sex category variables.[37] Earlier research demonstrates that smaller sets of HCCs chosen via data-adaptive driven techniques can maintain strong overall individual-level fit (Rose 2016; Rose, Bergquist and Layton 2017), but the potential of variable-reduction to maintain group fit has not yet been studied. Final formulas drop the drug variables and then drop some HCC variables while using reinsurance based on spending and alternatively based of residual spending (after risk-adjustment payment) equal to about 1% of funds.

5. Results

We present results for the range of formulas considered and then provide additional information on the best-performing formula, the regression with tightened constraints on the TPR for the four chronic illness groups.

Overall Results

Results for the Baseline Formula and constrained and penalized regression methods using reinsurance and variable selection techniques are shown in Table 2. The first row of the table contains the results for the Baseline Formula. Our measure of individual fit, PSF, is the same as R^2 from the regression equation because no other factors affect payment in the Baseline Formula. All other formulas incorporate reinsurance so PSF diverges from the regression R^2. The regression R^2 for the other formulas reports the explained variance attributable to the risk adjustor variables after deducting reinsurance payments from plan spending.

[37] When ties occur such that more than the pre-specified number of variables would be selected, the algorithm increases the value of the regularization tuning parameter until ties are excluded.

First, we consider the effects on individual and group-level fit dropping the drug variables.[38] We find that all alternative methods using reinsurance maintain or even improve upon individual and group-level fit. Constrained regression with TPR's set to the baseline values and penalized regression have PSFs (and R^2) well above the Baseline Formula. In both the constrained regression with TPRs set at baseline and the penalized regression with penalties tuned to hit or exceed baseline values, PSF improves from 36.8% to 55.6% and 55.4%, respectively. At the same time, group fit is maintained in the constrained regression (by construction). For the penalized regression we even see improvement in group fit, with the mental health TPR increasing to 0.90 from the 0.81 in the Baseline Formula. The final column in Table 2 summarizes the impacts on all 17 CCS groups. The 23.5% improved value for the constrained regression with TPRs set at baseline indicates that of the 17 CCS groups, 23.5%, i.e. four, saw their TPRs move towards 1.0, the balance of the CCS moving away from 1.0.

We can ask more of the constrained regression approach by constraining the TPRs to 1.0 for each of the four chronic illness groups. Even with this stricter set of constraints, completely eliminating selection incentives at the group level for these four conditions, we maintain high levels of individual fit with conventional reinsurance. With the tightened constraints, the PSF still substantially exceeds the Baseline Formula (52.7% compared to 36.8%) and the R^2 drops only negligibly to 36.4% from 36.8%. To compare, running an OLS regression on the reduced set of variables with reinsurance gives an R^2 of 40.0% and a PSF of 55.7%. Note that not only are the TPRs improved for the four chronic illness groups subject to the explicit constraints, but imposing these constraints improves the fit for other chronic illness groups too. Specifically, the impact of the

[38] The incremental improvement in R^2 from simply adding the set of RXC variables to the HCC variables used in the Marketplace model is about 2 percentage points, a finding in accord with the report of Xu, Trish and Joyce (2019).

constraints on all the CCS categories is generally positive, with 88% (i.e., 15 of 17) improving in relation to the Baseline Formula. We revisit this promising approach in the next subsection where we will characterize in more detail the effects of the tighter constraints on other CCS groups, those with no CCS indicator, and those with illnesses related to the drug variables.

We next compare fit using residual reinsurance rather than conventional reinsurance. Using constrained regression with TPRs set to baseline values, we estimate individual fit under both reinsurance methods (with group fit maintained by the constraints). We find that residual reinsurance improves PSF from 52.7% to 53.2%. While this improvement in individual fit is small, it comes "free" by better directing the same funds set aside for conventional reinsurance. Residual reinsurance also improves individual fit without affecting the number of CCS groups positively or negatively affected in relation to baseline.

We subsequently remove HCCs from the risk adjustment formula using data-adaptive variable selection. As discussed in Section 3, we implement a lasso penalized regression within the cross-validation folds to select a maximum of 20, 30, 40, or 50 HCC variables (in addition to all age and sex category variables) and find corresponding sets of 18, 30, 38, and 49 HCCs. The list of HCCs selected for each of these variables sets can be found in Table 3. In general, we find that conditions that are expensive and prevalent (e.g., diabetes and the high-severity flag) appear in the most effective minimal set of variables. As we allow more HCCs, we see rare and expensive conditions included. For example, hemophilia (coded in less than .01% of the sample with average spending of $286,000) is added among the 12 new variables appearing between the set of 18 to the set of 30. Cystic fibrosis, (also occurring in less than .01% with half the mean spending of hemophilia) comes in when we expand from 38 to 49 HCCs. However, these variable sets should not be overinterpreted as they were selected based on their predictiveness for the outcome. For

example, selected variables could be proxies for factors that are not measured in the data. Other empirical considerations, such as collinearity, could lead to a predictive variable not being selected.

Our alternative methods maintain fit with the number of HCCs reduced to 40, less than half of the current number. Figure 1 shows the diminishing returns of adding HCC variables to the risk adjustment formula with respect to PSF compared to an OLS regression with the same number of variables. The penalized regressions tracks the OLS regression and is better able to maintain overall fit with smaller numbers of HCCs (e.g., 49.5% PSF with 18 HCCs) but group fit, displayed in Table 2, consequently suffers, most notably for the mental health group (TPR of 0.65 vs. 0.81 in the Baseline Formula). Conversely, the constrained regressions with TPRs set to 1 have constant performance across the groups, as they must, with individual fit quickly approaching penalized regression once there are 28 HCCs. The constrained regression with 49 HCCs has a PSF of 50.8%, which also compares favorably to the constrained regression with all 94 HCC-based variables (52.7% PSF). Thus, the relative efficiency of the constrained regression that drops over 60% of the HCC-based variables is 96%. However, fit for the dropped HCCs can suffer; for example, in the 18 HCC regression, the TPR for hemophilia (not included in the regression) drops from 1.0 to less than 0.3 for both penalized and constrained regression. In Figure 2, we further depict individual fit gains among the constrained regressions as variable selection and reinsurance methods vary. We see the largest improvement when moving from 18 to 30 HCCs. The gain from using residual as opposed to conventional reinsurance is small given that all HCCs are included in the model.

TPRs set to 1.0: Results for Other CCS Groups and for Illnesses Associated with Drug Variable

Our methods optimize overall fit and constrain group fit for four specified groups: those with cancer, diabetes, heart disease, and mental health and substance use disorders. Results indicate that constraining the TPR for these four groups to 1.0, fully eliminating over- and undercompensation for these groups, also improves group fit for 15 of the 17 broad CCS categories.

This subsection reports on additional fit properties of this most promising formula from Table 2 in relation to the Baseline Formula. Specifically, we consider how predicted costs change for groups defined by multilevel diagnosis-based CCS categories, including those falling into no group at all, and for people most directly affected by the removal of the drug variable categories. We find that the constrained regression with TPRs set to 1.0 far outperforms the Baseline Formula by maintaining and often improving fit for these groups.

Figure 3 compares TPRs for constrained regression with TPRs set to 1.0 to the Baseline Formula for all 17 CCS groups, ranked by the TPR from the Baseline Formula. All 17 CCS groups are undercompensated in the Baseline Formula (recall, many of the conditions placing an individual in the CCS group are not explicitly recognized in the risk adjustment formula). Introducing reinsurance and the four constraints generally moves the TPRs towards 1.0, decreasing underpayment. Not surprisingly, we find that the CCS groups with the largest improvement have more overlap with the set of four chronic conditions considered in our regressions. For example, individuals with circulatory disease had the largest absolute improvement of 12% (from 0.87 to 0.99) and more than half of these individuals have one of the four chronic conditions. Similarly, the only CCS groups that did not improve have the smallest overlap (less than 1/3 of individuals in these categories had at least one of the four conditions). The TPR for pregnancy-related complications (5.3% of the sample) dropped slightly from 0.87 to 0.86 and those with perinatal-related conditions (0.1% of the sample) dropped from 0.86 to 0.85. CCS group overlap with our 4 chronic conditions can be found in Appendix C.

We also consider the impact of our new formulas on the 724,981 people in our sample (about 14.5% of the total) who do not fall into a CCS group. These enrollees have very little (if any) interaction with the health care system and so their costs are low. The baseline risk adjustment formula overestimates their costs by $2,258 on average. This overcompensation is more than cut in

half, to $855, in the constrained regression with TPRs set to 1.0 for the four chronic illness groups.[39] The cut in overpayment for those with no illness indicator is thus financing part of the additional compensation to previously undercompensated groups. We regard this transfer to be a particularly attractive feature of the constrained regression formula paired with high-cost risk sharing.

Those who would seem to be at the greatest risk for a reduction in payment by our constrained regression in relation to the Baseline Formula are individuals with specific illnesses associated with the drug indicators. For example, RXC_07 is an indicator for anti-diabetic agents (excluding insulin and metformin) used to treat diabetes. Figure 4 compares the TPRs for the set of HCC-defined conditions that are affected by the drug variables we dropped from the model. These HCCs were selected because they are relevant clinical groups for the drug variables and thus directly interacted with the drug variables in the Baseline Formula. See Appendix D for the list of HCCs considered, and the corresponding drug variables. The Baseline Formula leads to TPRs above 1.0 for all of these groups, indicating that in the presence of the RX variables they tend to be overcompensated. Removing the drug variables tends to reduce compensation but the TPR for every group is maintained above 1.0.[40] For the largest group (with respect to sample size), those with diabetes, the payment is increased slightly by 2% for the constrained regression with TPRs set to 1.0 in relation to the Baseline Formula.

Overall, Figures 3 and 4 present a remarkable set of results related to fit. The payment formula (a) dropping drug variables, (b) using 1 percent of funds used for reinsurance, and (c) constraining TPRs to 1.0 for four chronic illnesses moves fit at the individual level to a PSF of 52.7%. It also improves fit for virtually all CCS groups and does so without creating underpayment

[39] By comparison, constrained regression set to baseline reduces the overcompensation to $2,030 and penalized regression reduces it to $1,945.

[40] This may not be surprising. If the model were simple OLS, inclusion of the HCCs in the model would lead to no over or undercompensation. Our constrained model with reinsurance does not have that guarantee.

for any of the illness groups related to the drug variables. Looking at how the coefficients change from the baseline formula to the constrained regression (with TPRs set to 1), we find that the largest coefficient changes occur in groups impacted by the dropped drug variables, indicating that, overall, the constrained regression re-routes funds to these conditions via the main effects of the diagnostic variables. The other large increases were correlated with our four chronic conditions. Notably, coefficients correlated with mental illness, such as personality disorder, major depressive disorder, and psychosis more than doubled in value. Most coefficient decreases were among the SEVERE flag and HCC interactions, although the coefficient on hemophilia (HCC066) also decreased significantly from $268,000 thousand dollars to $194,000.

5. Discussion

Reinsurance, constrained regression, and machine learning variable selection are powerful methods to improve the fit of health plan payments systems at the individual and group level. We show in the case of the complex payment formula applied in the Marketplaces that using these tools permits radical reduction in the number of variables used in the risk adjustment formula, while maintaining or improving performance in terms of fit. Design of risk adjustment systems involves an inherent tradeoff between the benefits of adding a variable, generally measured in terms of fit, and the costs, poorly measured, in terms of incentives to upcode, administrative burden, and other costs. Our paper shows that use of reinsurance, constrained regression, and machine learning variable selection change the terms of the tradeoff to permit more consideration of the difficult-to-quantify but real costs of variable addition.

There are several reasons to consider excluding specific variables from the risk adjustment formula. Certain variables are vulnerable to the increased incentive to overuse and therefore good candidates for removal. In our analysis, we removed the drug variables for several reasons.

Accessing the drug claims file for purposes of risk adjustment classification adds significantly to demands on data. In the specification used in the Marketplaces, just one prescription for a designated drug triggers thousands of dollars of additional payments. Furthermore, coverage of specific drugs can vary across health insurance formularies, and the technology of drug treatment is changing constantly, increasing the cost of maintaining proper drug classifications.

Drug variables are not the only variables whose contribution to fit may not be worth their cost in terms of gaming and incentives. An important task for research is to identify which variable are most responsible for elevation in risk scores so that the costs and benefits of their inclusion can be assessed. Variables that are more difficult or time-consuming to document should also be considered for removal.

That said, dropping variables inevitably reduces overall fit measured by R^2 in the risk adjustment formula. Reinsurance is an effective antidote to loss of fit measured by R^2. Group fit is not generally included as part of the objectives in choice of risk adjustment weights, but it can be. Constrained or penalized regression can ensure that group fit is maintained (or improved). Incorporating group fit into the loss function means that the variable weights selected by a least squares procedure attain the target group fit as efficiently as possible in terms of the R^2 measure of individual fit. This improvement in the efficiency in which variables are used permits, if desired, a reduction in the number of variables used to hit fit targets. In our analysis, we constrained average predicted costs for only four large and well-known chronic condition groups. We found that imposing constraints on these four groups improved predictions for many other condition groups because it transferred money from healthier to sicker individuals. A particularly attractive feature of our four-group constrained formula is that it substantially *reduced the overcompensation* for those with no indicator of any CCS group. In effect, our formula reduces undercompensation for a

comprehensive set of illness groups, financed by a reduction in overcompensation by those with claims for no illnesses.

An obvious concern with the strategy of dropping variables is possible introduction of underpayment for the people with the particular risk adjustment flag. Without an indicator for Ischemic or Unspecified Stroke (HCC146), a variable that was removed when we went from 49 to 38 HCCs, persons with that condition may be underpaid. This argument can be made for *any* potential risk adjustor and cannot therefore be dispositive in terms of inclusion or exclusion. Some specific diagnoses may merit inclusion for fear of plan/provider response to underpayment, but if it is implausible that a plan or provider can do anything in response to this particular form of underpayment, inclusion of the variable has no positive effect on the outcomes induced by the payment formula.

Optimizing towards a payment system that considers a larger set of goals remains important. While adding additional risk adjusters can be beneficial, it is costly too. Our application focused on the U.S. health insurance Marketplaces, but these methods can be implemented in other individual health insurance markets such as Switzerland, which shares many features of the Marketplaces, or Germany, which currently has a fixed limit on the number of risk adjusters in the payment model. Utilizing modern statistical tools and modifying the conventional approach to risk adjustment, we find that the risk adjustment formula can be simplified at no evident cost to fit at the individual and group level.

References

Beck K, Kauer L, McGuire TG, and Schmid C. 2019 "Making Use of Canton Contributions to Health Care Costs in Switzerland," under review.

Bergquist S, Layton TJ, McGuire TG, and Rose S. 2019. "Data Transformations to Improve the Performance of Health Plan Payment Models," *Journal of Health Economics* 66: 195-207.

Buchner F, Wasem J, Schillo S. (2015) "Regression Trees Identify Relevant Interactions: Can this Improve the Predictive Performance of Risk Adjustment?" *Health Economics*, 26 OCT 2015, DOI: 10.1002/hec.3277

Carey C. 2017a. "Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D." *American Economic Journal: Economic Policy* 9, no. 1: 38-73.

Carey C. 2017b. "Time to Harvest: Evidence on Consumer Choice Frictions from a Payment Revision in Medicare Part D." Working Paper. Retrieved from https://drive.google.com/file/d/0B2TeS7lispKBQ21RLV9PckE2eWs/view.

Eijkenaar F, van Vliet RCJA, and van Kleef RC. 2017. "Diagnosis-based Cost Groups in the Dutch Risk-Equalization Model - Effects of Clustering Diagnoses and of Allowing Patients to be Classified into Multiple Risk-Classes." *Medical Care*.

Ellis RP, Martins B, and Rose S. 2018. "Risk Adjustment for Health Plan Payment," In McGuire. T.G., & Van Kleef. R., *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice* (pp. 105-131). London: Academic Press (Elsevier).

Ellis RP and McGuire TG. 2007. "Predictability and Predictiveness in Health Care Spending." *Journal of Health Economics* 26, no. 1: 25-48.

Geruso, M and Layton TJ. 2020. "Upcoding: Evidence from Medicare on Squishy Risk Adjustment." *Journal of Political Economy*.

Geruso M, Layton TJ, and Prinz D. 2019. "Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges." *American Economic Journal: Economic Policy* 11, no. 2: 64-107.

Geruso M and McGuire TG. 2016. "Tradeoffs in the Design of Health Plan Payment Systems: Fit, Power and Balance." *Journal of Health Economics* 47: 1-19.

Han T and Lavetti K. 2017. "Does Part D Abet Advantageous Selection in Medicare Advantage?" *Journal of Health Economics* (December).

Jacobs PD and Kronick R. 2018. "Getting What We Pay For: How do Risk-Based Payments to Medicare Advantage Plans Compare with Alternative Measures of Beneficiary Health Risk?" *Health Services Research* 56 (3) Part I, 4097-5015.

James G, Witten D, Hastie T, Tibshirani R. 2013. *An Introduction to Statistical Learning: with Applications in R*. New York: Springer.

Kautter J, Pope GC, Ingber M, Freeman S, Patterson L, Cohen M, and Keenan P. 2014. "The HHS-HCC Risk Adjustment Model for Individual and Small Group Markets under the Affordable Care Act." *Medicare & Medicaid Research Review* 4, no. 3: E1-E11.

Lamers LM and Vliet RCJA. 2003. "Health Based Risk Adjustment: Improving the Pharmacy Group Based Model to Reduce Gaming Possibilities." *European Journal of Health Economics.* 4: 107-114.

Layton TJ and McGuire TG. 2017. "Options for Dealing with High-Cost Cases in Health Insurance Marketplaces," *American Journal of Health Economics,* 3(2): 1-27.

Layton TJ, Ellis RP, McGuire TG, and Van Kleef RC. 2017. "Measuring Efficiency of Health Plan Payment Systems in Managed Competition Health Insurance Markets." *Journal of Health Economics* (December), 56: 237-255.

Layton TJ, Montz E, and Shepard M. 2018. "Health Plan Payment in US Marketplaces: Regulated Competition with a Weak Mandate." In *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice*, edited by Thomas G. McGuire and Richard C. van Kleef. Elsevier.

McGuire TG, Newhouse JP, Normand SL, Shi J, and Zuvekas S. 2014. "Assessing Incentives for Service-Level Selection in Private Health Insurance Exchanges." *Journal of Health Economics* 35: 47-63.

McGuire TG and van Kleef RC (eds.) 2018. *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice*, Elsevier.

McGuire TG and van Kleef R. 2018. Risk Sharing. In McGuire TG & Van Kleef R, *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice* (pp. 105-131). London: Academic Press (Elsevier).

McGuire TG, Schillo S, and van Kleef R. 2020. Two-Sided Reinsurance and Risk Adjustment in Individual Health Insurance: Germany, the Netherlands and the US Marketplaces. *American Journal of Health Economics*.

Montz E, Layton TJ, Busch AB, Ellis RP, Rose S, and McGuire TG**.** 2016. "Risk Adjustment Simulation: Health Plans May Have Incentives to Distort Mental Health and Substance Abuse Coverage." *Health Affairs*, 35(6): 1022-1028, 2016.

Pope GC, Kautter J, Ingber JJ, Freeman S, Sekar R, and Newhart C. 2011. "Evaluation of the CMS-HCC Risk Adjustment Model." In *Final Report*, RTI Project Number 0209853.006, RTI International, March.

Rose S. 2016. "A Machine Learning Framework for Plan Payment Risk Adjustment." *Health Services Research*, 51(6): 2358-74.

Rose S, Bergquist SL, and Layton TJ. 2017. "Computational Health Economics for Identification of Unprofitable Health Care Enrollees." *Biostatistics*, 18(4): 682-694.

Rose S and McGuire TG. 2019. "Limitations of p-values and R-squared for Stepwise Regression Building: A Fairness Demonstration in Health Policy and Risk Adjustment," *The American Statistician*, 73(S1): 152-6.

Schillo S, Lux G, Wasem J, Buchner F. 2016. "High-cost pool or high-cost groups? how to handle the high(est) cost cases in a risk adjustment mechanism." *Health Policy.* (120): 141-47. DOI: 10.1016/j.healthpol.2016.01.003.

Shepard M. 2016. "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange." Harvard University. Unpublished.

Van Kleef R, Eijkenaar F, van Vliet RCJA, van de Ven WPMM. 2018. "Health Plan Payment in the Netherlands." In *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice*, edited by McGuire TG and Van Kleef RC. Elsevier.

Van Kleef RC, McGuire TG, van Vliet RCJA, van de Ven WPMM. 2017. "Improving Risk Equalization with Constrained Regression." *The European Journal of Health Economics*, 18, no. 9: 1137-1156.

Van Kleef RC, van Vliet RCJA, and van de Ven WPMM. 2013. "Risk Equalization in The Netherlands: An Empirical Evaluation." *Expert Review of Pharmacoeconomics & Outcomes Research* 13, no. 6: 829-839.

Van Veen SHCM, van Kleef RC, van de Ven WPMM, van Vliet RCJA. 2015. "Is There One Measure of Fit That Fits All? A Taxonomy and Review of Measures-Of-Fit for Risk-Equalization Models." *Medical Care Research and Review* 72, no. 2: 220-243.

Xu J, Trish E, and Joyce G. 2009. "Incorporating Prescription Drug Utilization Information into the Marketplace Risk Adjustment Model Improves Payment Accuracy and Reduces Adverse Selection Incentives." *Medical Care Research and Review* (August).

Zink A and Rose S. 2019. "Fair Regression for Health Care Spending." *Biometrics*, in press. DOI: 10.1111/biom.13206.

Table 1: Descriptive Statistics for 2017 Sample (N=5,000,000)

Variable	Proportion (%)	Mean Spending ($)
All	100.0	6,683
Sex		
Female	52.1	7,459
Male	47.9	5,839
Age		
21-29	18.0	3,705
30-39	20.6	4,940
40-49	23.7	6,135
50-64	37.7	9,397
Any HCC		
Yes	21.1	20,306
No	78.9	3,035
Any RXC		
Yes	6.3	28,206
No	93.7	5,230
CCS Chronic Group		
Cancer	6.9	19,673
Diabetes	9.0	15,106
Heart Disease	9.3	20,154
Mental Health	11.7	12,293

Table 2: Baseline Risk Adjustment Formula and Alternative Methods Using Reinsurance

Method	R^2 (%)	PSF (%)	Total Payment Ratio (TPR) Cancer	Diabetes	Heart	Mental	Improved CCS (%)*
Baseline	36.8	36.8	0.91	1.00	0.83	0.81	NA
Removing Drug Variables							
Constrained Regression (TPRs set to baseline)	40.3	55.6	0.93	1.02	0.84	0.81	23.5
Penalized Regression†	40.0	55.4	0.92	0.98	0.83	0.90	47.1
Constrained Regression (TPRs set to 1)	36.4	52.7	1.02	1.02	1.02	1.01	88.2
Constrained Regression (TPRs set to 1)‡	37.8	53,2	1.02	1.02	1.02	1.01	88.2
Removing Drug Variables and HCCs							
Constrained Regression (TPRs set to 1)							
Age × Gender + 18 HCCs	-25.6	6.7	1.02	1.02	1.02	1.01	76.5
Age × Gender + 30 HCCs	15.1	36.9	1.03	1.02	1.03	1.02	94.1
Age × Gender + 38 HCCs	31.8	49.3	1.02	1.02	1.02	1.01	82.4
Age × Gender + 49 HCCs	33.8	50.8	1.02	1.02	1.02	1.01	88.2
Penalized Regression†							
Age × Gender + 18 HCCs	32.1	49.5	0.72	0.96	0.75	0.65	0.0
Age × Gender + 30 HCCs	35.5	52.0	0.86	0.97	0.77	0.72	0.0
Age × Gender + 38 HCCs	37.0	53.2	0.88	0.97	0.80	0.86	5.9
Age × Gender + 49 HCCs	38.4	54.2	0.90	0.97	0.82	0.86	5.9

*The share of CCS groups (out of 17 total) whose TPRs moved closer to 1 compared to the Baseline Formula (see Appendix C for the list of CCS groups).
†Mental health penalty is 500,000, cancer and heart penalties are 50,000, and no penalty is set for diabetes.
‡This method uses residual reinsurance, all other alternative methods use conventional reinsurance.

Note: All methods estimated using a sample of 5 million observations from the full sample of 10,043,052 observations. Measures calculated with 5-fold cross-validated predicted values.

Table 3: HCCs selected under 4 variable selection methods with increasing HCCs

18 HCCs

Diabetes (G01)
Immunodeficiencies and disorders of the Immune Mechanism (G08)
Respiratory Arrest, Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes (G13)
Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock (HCC002)
Metastatic Cancer (HCC008)
Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia (HCC009)
Protein-Calorie Malnutrition (HCC023)
Inflammatory Bowel Disease (HCC048)
Rheumatoid Arthritis and Specified Autoimmune Disorders (HCC056)
Multiple Sclerosis (HCC118)
Congestive Heart Failure (HCC130)
Acute Myocardial Infarction (HCC131)
Specified Heart Arrhythmias (HCC142)
End Stage Renal Disease (HCC184)
Artificial Openings for Feeding or Elimination (HCC253)
1 if adult high cost interaction, else 0 (INT_GROUP_H)
1 if adult medium cost interaction, else 0 (INT_GROUP_M)
Adult severe illness 0/1 marker (SEVERE_3)

+ 12 HCCs (30 in total)

Necrotizing Fasciitis, Bone/Joint/Muscle Infections/Necrosis (G03)
Drug psychosis and dependence (G09)
Chronic Obstructive Pulmonary Disease, Including Bronchiectasis, Asthma (G15)
Completed Pregnancy (G18)
HIV/AIDS (HCC001)
Colorectal, Breast (Age < 50), Kidney, and Other Cancers (HCC011)
Breast (Age 50+) and Prostate Cancer, Benign/Uncertain Brain Tumors, and Other Cancers and Tumors (HCC012)
Intestinal Obstruction (HCC045)
Hemophilia (HCC066)
Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy (HCC115)
Respirator Dependence/Tracheostomy Status (HCC125)
Stem Cell, Transplant Status/Complications (HCC251)

+ 8 HCCs (38 in total)

Metabolic and endocrine disorders (G02A)
Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis (HCC042)
Coagulation Defects and Other Specified Hematological Disorders (HCC075)
Major Depressive and Bipolar Disorders (HCC088)
Non-Traumatic Coma, Brain Compression/Anoxic Damage (HCC122)
Unstable Angina and Other Acute Ischemic Heart Disease (HCC132)
Hemiplegia/Hemiparesis (HCC150)
Kidney Transplant Status (HCC183)

+11 HCCs (49 in total)

Heart Assistive Device/Artificial Heart (G14)
Chronic Viral Hepatitis C (HCC037)
Non-Hodgkin's Lymphomas and Other Cancers and Tumors (HCC010)
Heart Infection/Inflammation, Except Rheumatic (HCC135)
Intracranial Hemorrhage (HCC145)
Ischemic or Unspecified Stroke (HCC146)
Atherosclerosis of the Extremities with Ulceration or Gangrene (HCC153)
Vascular Disease with Complications (HCC154)
Cystic Fibrosis (HCC159)
Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections (HCC163)
Hip Fractures and Pathological Vertebral or Humerus Fractures (HCC226)

Figure 1: Diminishing returns by adding HCCs

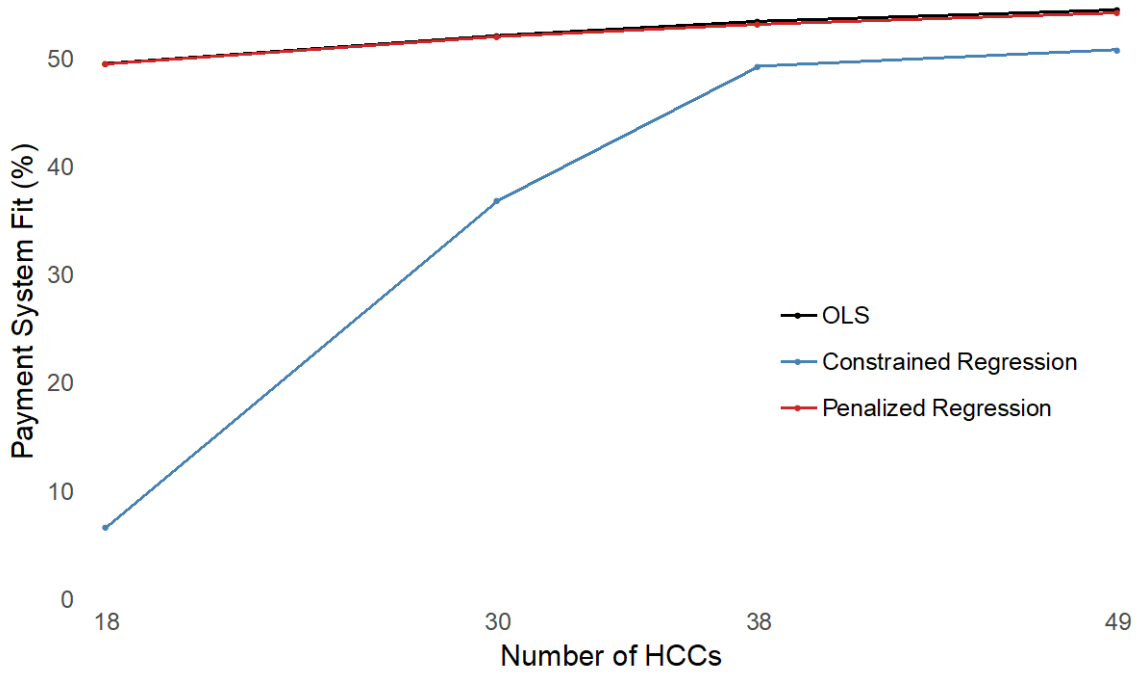
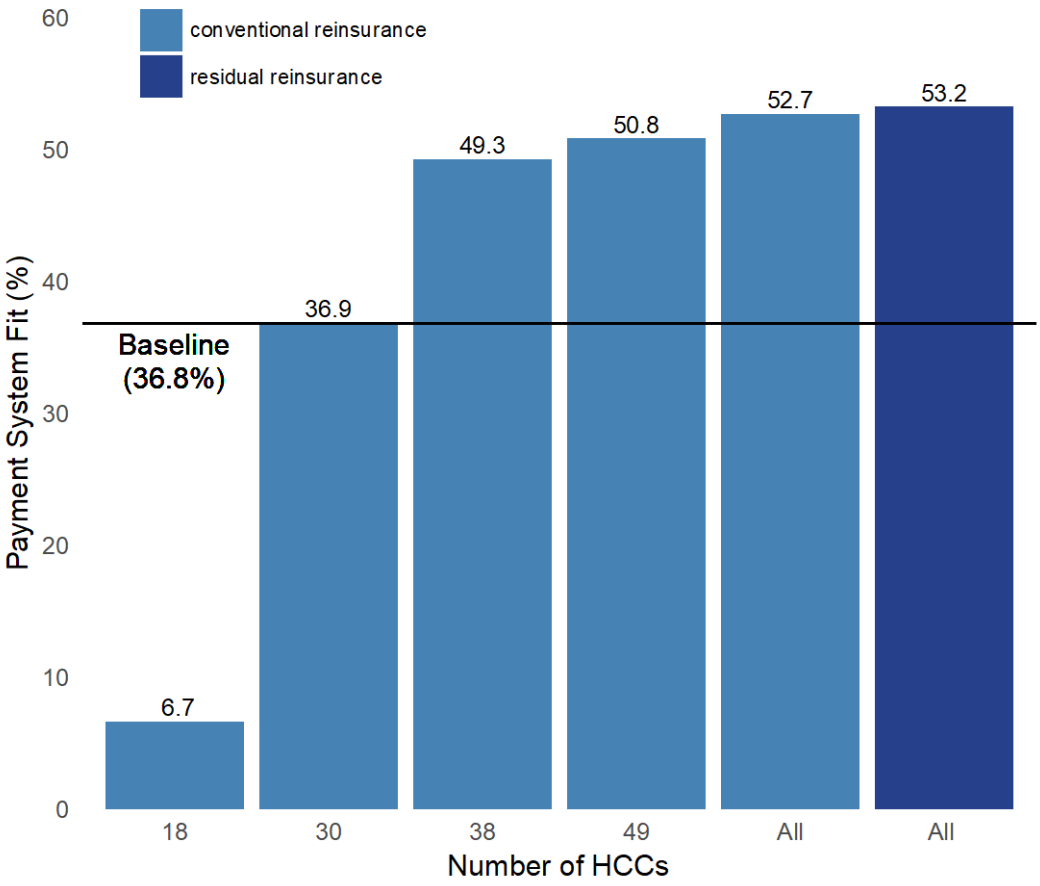
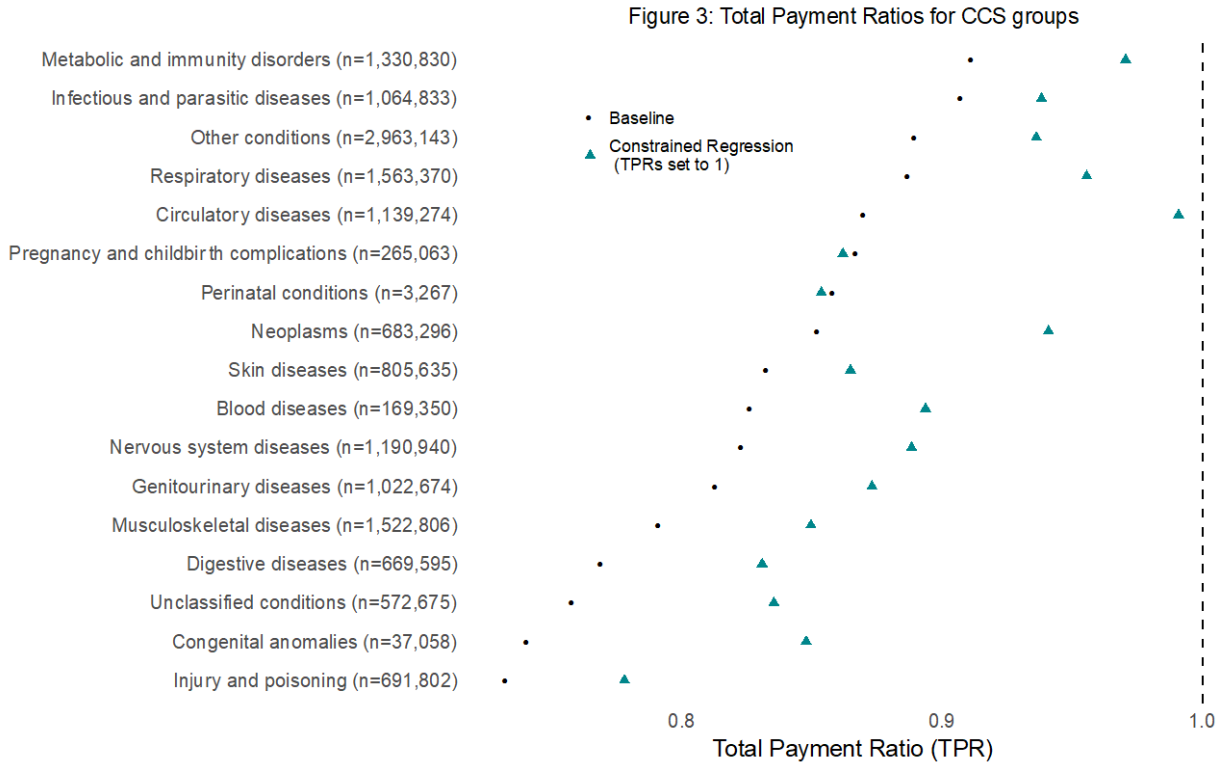


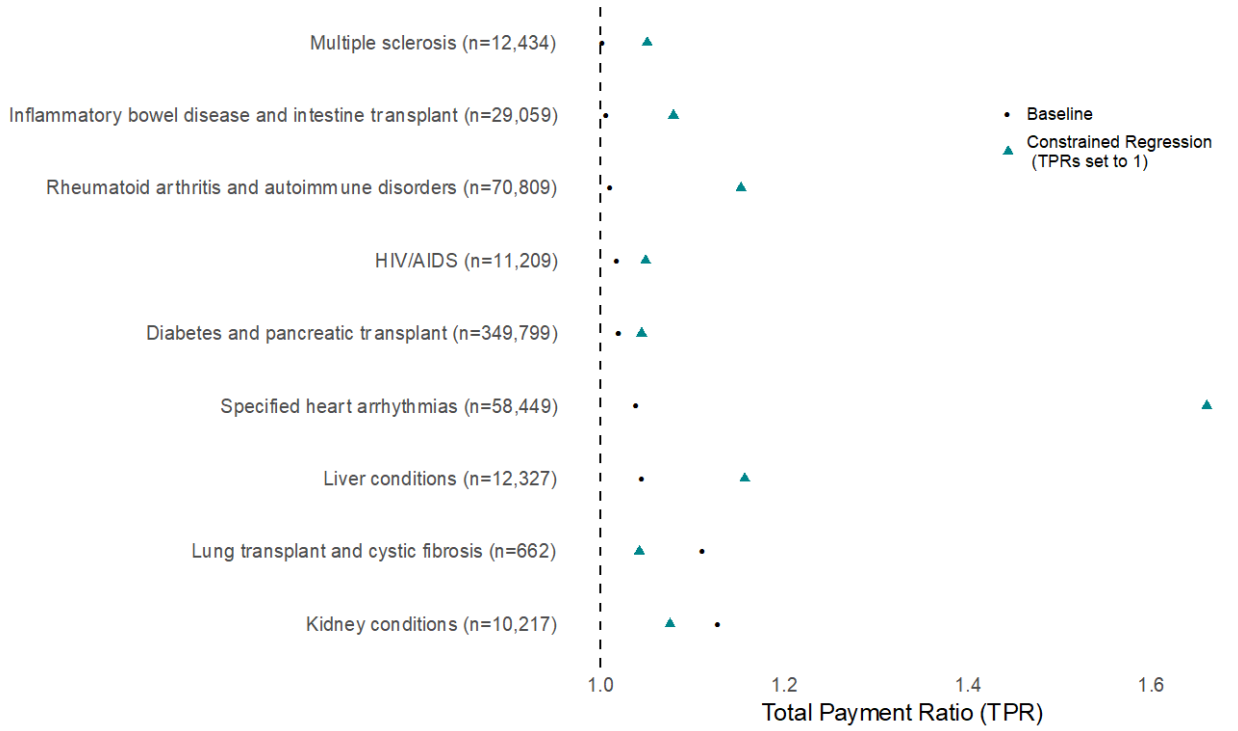
Figure 2: Constrained regression (TPRs set to 1)





Note: Constrained regression excludes drug variables but includes all HCCs from the baseline formula.

Figure 4: Total Payment Ratios for HCC Groups Impacted by Drug Variables



Note: Constrained regression excludes drug variables but includes all HCCs from the baseline formula.

Appendix A: Drug Variables

RXC	Definition	Note
RXC_01	Anti-HIV Agents	
RXC_02	Anti-Hepatitis C (HCV) Agents	
RXC_03	Antiarrhythmics	
RXC_04	Phosphate Binders	
RXC_05	Inflammatory Bowel Disease Agents	
RXC_06	Insulin	RXC_07 is set to 0 if a person has RXC_06
RXC_07	Anti-Diabetic Agents, Except Insulin and Metformin	
RXC_08	Multiple Sclerosis Agents	
RXC_09	Immune Suppressants and Immunomodulators	
RXC_10	Cystic Fibrosis Agents	
RXC_11	Ammonia Detoxicants	Not used directly in the model, for interactions only
RXC_12	Diuretics, Loop and Select Potassium-Sparing	
RXC_01_x_HCC001	Additional effect for enrollees with RXC 01 and HCC 001	
RXC_02_x_HCC037_1_036_035_034	Additional effect for enrollees with RXC 02 and (HCC 037_1 or 036 or 035 or 034)	
RXC_03_x_HCC142	Additional effect for enrollees with RXC 03 and HCC 142	
RXC_04_x_HCC184_183_187_188	Additional effect for enrollees with RXC 04 and (HCC 184 or 183 or 187 or 188)	
RXC_05_x_HCC048_041	Additional effect for enrollees with RXC 05 and (HCC 048 or 041)	
RXC_06_x_HCC018_019_020_021	Additional effect for enrollees with RXC 06 and (HCC 018 or 019 or 020 or 021)	
RXC_07_x_HCC018_019_020_021	Additional effect for enrollees with RXC 07 and (HCC 018 or 019 or 020 or 021)	
RXC_08_x_HCC118	Additional effect for enrollees with RXC 08 and HCC 118	
RXC_09_x_HCC056_057_and_048_041	Additional effect for enrollees with RXC 09 and (HCC 048 or 041) and (HCC 056 or 057)	
RXC_09_x_HCC056	Additional effect for enrollees with RXC 09 and HCC 056	
RXC_09_x_HCC057	Additional effect for enrollees with RXC 09 and HCC 057	
RXC_09_x_HCC048_041	Additional effect for enrollees with RXC 09 and (HCC 048 or 041)	
RXC_10_x_HCC159_158	Additional effect for enrollees with RXC 10 and (HCC 159 or 158)	
RXC_11_x_HCC036_035_034	Additional effect for enrollees with RXC 11 and (HCC 036 or 035 or 034)	Removed in V0519
RXC_12_x_HCC130_129_128	Additional effect for enrollees with RXC 12 and (HCC 130 or 129 or 128)	

Appendix B: Coefficients for the Baseline Formula

Variable	Description	Coefficient
MAGE_LAST_25_29	Male, Age 25-29	20
MAGE_LAST_30_34	Male, Age 30-34	87
MAGE_LAST_35_39	Male, Age 35-39	368
MAGE_LAST_40_44	Male, Age 40-44	646
MAGE_LAST_45_49	Male, Age 45-49	1033
MAGE_LAST_50_54	Male, Age 50-54	1608
MAGE_LAST_55_59	Male, Age 55-59	1969
MAGE_LAST_60_GT	Male, Age 60+	2375
FAGE_LAST_21_24	Female, Age 21-24	731
FAGE_LAST_25_29	Female, Age 25-29	922
FAGE_LAST_30_34	Female, Age 30-34	1341
FAGE_LAST_35_39	Female, Age 35-39	1804
FAGE_LAST_40_44	Female, Age 40-44	2032
FAGE_LAST_45_49	Female, Age 45-49	2174
FAGE_LAST_50_54	Female, Age 50-54	2549
FAGE_LAST_55_59	Female, Age 55-49	2523
FAGE_LAST_60_GT	Female, Age 60+	2830
G01	Diabetes with Acute Complications Diabetes with Chronic Complications Diabetes without Complication	2577
G02A	Mucopolysaccharidosis Lipidoses and Glycogenosis Amyloidosis, Porphyria, and Other Metabolic Disorders Adrenal, Pituitary, and Other Significant Endocrine Disorders	9461
G03	Necrotizing Fasciitis Bone/Joint/Muscle Infections/Necrosis	20142
G04	Osteogenesis Imperfecta and Other Osteodystrophies Congenital/Developmental Skeletal and Connective Tissue Disorders	12714
G06	Myelodysplastic Syndromes and Myelofibrosis Aplastic Anemia	34243
G07	Acquired Hemolytic Anemia, Including Hemolytic Disease of Newborn Sickle Cell Anemia (Hb-SS) Thalassemia Major	31918
G08	Combined and Other Severe Immunodeficiencies Disorders of the Immune Mechanism	17407
G09	Drug Psychosis Drug Dependence	17500
G10	Traumatic Complete Lesion Cervical Spinal Cord Quadriplegia	48754
G11	Traumatic Complete Lesion Dorsal Spinal Cord Paraplegia	30981
G12	Muscular Dystrophy Parkinson`s, Huntington`s, and Spinocerebellar Disease, and Other Neurodegenerative Disorders	8952
G13	Respiratory Arrest Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes	55039
G14	Heart Assistive Device/Artificial Heart Heart Transplant	102112
G15	Chronic Obstructive Pulmonary Disease, Including Bronchiectasis Asthma	4125
G16	Chronic Kidney Disease, Stage 5 Chronic Kidney Disease, Severe (Stage 4)	-1291
G17	Ectopic and Molar Pregnancy, Except with Renal Failure, Shock, or Embolism Miscarriage with Complications Miscarriage with No or Minor Complications	5264
G18	Completed Pregnancy with Major Complications Completed Pregnancy with Complications Completed Pregnancy with No or Minor Complications	15001
HHS_CC037_1	Chronic Viral Hepatitis C	3535
HHS_CC037_2	Chronic Hepatitis, Except Chronic Viral Hepatitis C	4792
HHS_HCC001	HIV/AIDS	2416
HHS_HCC002	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	50827
HHS_HCC003	Central Nervous System Infections, Except Viral Meningitis	25213
HHS_HCC004	Viral or Unspecified Meningitis	25449
HHS_HCC006	Opportunistic Infections	14788

HHS_HCC008	Metastatic Cancer	90033
HHS_HCC009	Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia	51637
HHS_HCC010	Non-Hodgkin`s Lymphomas and Other Cancers and Tumors	21213
HHS_HCC011	Colorectal, Breast (Age < 50), Kidney, and Other Cancers	19147
HHS_HCC012	Breast (Age 50+) and Prostate Cancer, Benign/Uncertain Brain Tumors, and Other Cancers and Tumors	11751
HHS_HCC013	Thyroid Cancer, Melanoma, Neurofibromatosis, and Other Cancers and Tumors	4916
HHS_HCC018	Pancreas Transplant Status/Complications	8300
HHS_HCC023	Protein-Calorie Malnutrition	42840
HHS_HCC034	Liver Transplant Status/Complications	44407
HHS_HCC035	End-Stage Liver Disease	18472
HHS_HCC036	Cirrhosis of Liver	7147
HHS_HCC038	Acute Liver Failure/Disease, Including Neonatal Hepatitis	13754
HHS_HCC041	Intestine Transplant Status/Complications	61025
HHS_HCC042	Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis	27038
HHS_HCC045	Intestinal Obstruction	20611
HHS_HCC046	Chronic Pancreatitis	9284
HHS_HCC047	Acute Pancreatitis/Other Pancreatic Disorders and Intestinal Malabsorption	11488
HHS_HCC048	Inflammatory Bowel Disease	10129
HHS_HCC056	Rheumatoid Arthritis and Specified Autoimmune Disorders	4944
HHS_HCC057	Systemic Lupus Erythematosus and Other Autoimmune Disorders	5885
HHS_HCC063	Cleft Lip/Cleft Palate	268120
HHS_HCC066	Hemophilia	11053
HHS_HCC075	Coagulation Defects and Other Specified Hematological Disorders	13387
HHS_HCC087	Schizophrenia	7784
HHS_HCC088	Major Depressive and Bipolar Disorders	7537
HHS_HCC089	Reactive and Unspecified Psychosis, Delusional Disorders	6651
HHS_HCC090	Personality Disorders	8384
HHS_HCC094	Anorexia/Bulimia Nervosa	15636
HHS_HCC096	Prader-Willi, Patau, Edwards, and Autosomal Deletion Syndromes	2645
HHS_HCC097	Down Syndrome, Fragile X, Other Chromosomal Anomalies, and Congenital Malformation Syndromes	4169
HHS_HCC102	Autistic Disorder	3664
HHS_HCC103	Pervasive Developmental Disorders, Except Autistic Disorder	22847
HHS_HCC110	Spinal Cord Disorders/Injuries	17547
HHS_HCC111	Amyotrophic Lateral Sclerosis and Other Anterior Horn Cell Disease	-3214
HHS_HCC112	Quadriplegic Cerebral Palsy	2614
HHS_HCC113	Cerebral Palsy, Except Quadriplegic	4220
HHS_HCC114	Spina Bifida and Other Brain/Spinal/Nervous System Congenital Anomalies	24955
HHS_HCC115	Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy	21369
HHS_HCC118	Multiple Sclerosis	32041
HHS_HCC120	Seizure Disorders and Convulsions	27204
HHS_HCC121	Hydrocephalus	50831
HHS_HCC122	Non-Traumatic Coma, Brain Compression/Anoxic Damage	129922
HHS_HCC125	Respirator Dependence/Tracheostomy Status	11948
HHS_HCC130	Congestive Heart Failure	31179
HHS_HCC131	Acute Myocardial Infarction	24432
HHS_HCC132	Unstable Angina and Other Acute Ischemic Heart Disease	17394
HHS_HCC135	Heart Infection/Inflammation, Except Rheumatic	7897
HHS_HCC142	Specified Heart Arrhythmias	25450
HHS_HCC145	Intracranial Hemorrhage	12016
HHS_HCC146	Ischemic or Unspecified Stroke	12430
HHS_HCC149	Cerebral Aneurysm and Arteriovenous Malformation	20208
HHS_HCC150	Hemiplegia/Hemiparesis	8427
HHS_HCC151	Monoplegia, Other Paralytic Syndromes	31520
HHS_HCC153	Atherosclerosis of the Extremities with Ulceration or Gangrene	21349
HHS_HCC154	Vascular Disease with Complications	37919

HHS_HCC156	Pulmonary Embolism and Deep Vein Thrombosis	98109
HHS_HCC158	Lung Transplant Status/Complications	17645
HHS_HCC159	Cystic Fibrosis	7308
HHS_HCC162	Fibrosis of Lung and Other Lung Disorders	13066
HHS_HCC163	Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections	23984
HHS_HCC183	Kidney Transplant Status	98990
HHS_HCC184	End Stage Renal Disease	10450
HHS_HCC217	Chronic Ulcer of Skin, Except Pressure	41301
HHS_HCC226	Hip Fractures and Pathological Vertebral or Humerus Fractures	21778
HHS_HCC227	Pathological Fractures, Except of Vertebrae, Hip, or Humerus	83438
HHS_HCC251	Stem Cell, Including Bone Marrow, Transplant Status/Complications	31817
HHS_HCC253	Artificial Openings for Feeding or Elimination	15400
HHS_HCC254	Amputation Status, Lower Limb/Amputation Complications	2416
INT_GROUP_H	1 if adult high cost interaction, else 0	-38758
INT_GROUP_M	1 if adult medium cost interaction, else 0	-17561
RXC_01	Anti-HIV Agents	22491
RXC_02	Anti-Hepatitis C (HCV) Agents, Direct Acting Agents	73318
RXC_03	Antiarrhythmics	11405
RXC_04	Phosphate Binders	27610
RXC_05	Inflammatory Bowel Disease Agents	9739
RXC_06	Insulin	9103
RXC_07	Anti-Diabetic Agents, Except Insulin and Metformin Only	3198
RXC_08	Multiple Sclerosis Agents	61265
RXC_09	Immune Suppressants and Immunomodulators	30691
RXC_10	Cystic Fibrosis Agents	33444
RXC_01_x_HCC001	Additional effect for enrollees with RXC 01 and HCC001	7583
RXC_02_x_HCC037 _1_036_035_034	Additional effect for enrollees with RXC 02 and (HCC 037_1 or 036 or 035 or 034)	-2331
RXC_04_x_HCC184 _183_187_188	Additional effect for enrollees with RXC 04 and (HCC 184 or 183 or 187 or 188)	62833
RXC_05_x_HCC048_041	Additional effect for enrollees with RXC 05 and (HCC 048 or 041)	-8015
RXC_06_x_HCC018 _019_020_021	Additional effect for enrollees with RXC 06 and (HCC 018 or 019 or 020 or 021)	1287
RXC_07_x_HCC018 _019_020_021	Additional effect for enrollees with RXC 07 and (HCC 018 or 019 or 020 or 021)	-298
RXC_08_x_HCC118	Additional effect for enrollees with RXC 08 and HCC 118	-10036
RXC_09_x_HCC048_041	Additional effect for enrollees with RXC 09 and (HCC 048 or 041)	-1583
RXC_09_x_HCC056	Additional effect for enrollees with RXC 09 and HCC 056	-7208
RXC_09_x_HCC056_057 _and_048_041	Additional effect for enrollees with RXC 09 and (HCC 048 or 041) and (HCC 056 or 057)	-28236
RXC_09_x_HCC057	Additional effect for enrollees with RXC 09 and HCC 057	17662
RXC_10_x_HCC159_158	Additional effect for enrollees with RXC 10 and (HCC 159 or 158)	109855
SEVERE_V3	Adult severe illness 0/1 marker	-27696
SEVERE_V3_x_G03	1 if adult severe illness and group G03, else 0	27579
SEVERE_V3_x_G06	1 if adult severe illness and group G06, else 0	61638
SEVERE_V3_x_G08	1 if adult severe illness and group G08, else 0	61594
SEVERE_V3_x_HHS_HCC006	1 if adult severe illness and HHS_HCC006, else 0	58234
SEVERE_V3_x_HHS_HCC008	1 if adult severe illness and HHS_HCC008, else 0	64532
SEVERE_V3_x_HHS_HCC009	1 if adult severe illness and HHS_HCC009, else 0	46182
SEVERE_V3_x_HHS_HCC010	1 if adult severe illness and HHS_HCC010, else 0	62730
SEVERE_V3_x_HHS_HCC035	1 if adult severe illness and HHS_HCC035, else 0	23233
SEVERE_V3_x_HHS_HCC038	1 if adult severe illness and HHS_HCC038, else 0	33768
SEVERE_V3_x_HHS_HCC115	1 if adult severe illness and HHS_HCC115, else 0	41179
SEVERE_V3_x_HHS_HCC135	1 if adult severe illness and HHS_HCC135, else 0	69171
SEVERE_V3_x_HHS_HCC145	1 if adult severe illness and HHS_HCC145, else 0	62251
SEVERE_V3_x_HHS_HCC153	1 if adult severe illness and HHS_HCC153, else 0	33690
SEVERE_V3_x_HHS_HCC154	1 if adult severe illness and HHS_HCC154, else 0	38871
SEVERE_V3_x_HHS_HCC163	1 if adult severe illness and HHS_HCC163, else 0	34460

SEVERE_V3_x_HHS_HCC253	1 if adult severe illness and HHS_HCC253, else 0	32119

Appendix C: Multi-Level CCS Diagnosis-Based Groups (% Overlap)*

1 Infectious and parasitic diseases (40.2)
2 Neoplasms (67.5)
3 Metabolic and immunity disorders (57.9)
4 Blood diseases (53.6)
5 Mental Illness (100)**
6 Nervous system diseases (43.1)
7 Circulatory diseases (62.2)
8 Respiratory diseases (40.3)
9 Digestive diseases (47.6)
10 Genitourinary diseases (43.3)
11 Pregnancy and childbirth complications (29.5)
12 Skin diseases (46.0)
13 Musculoskeletal diseases (42.0)
14 Congenital anomalies (59.8)
15 Perinatal conditions (32.1)
16 Injury and poisoning (42.1)
17 Other conditions (38.5)
18 Unclassified codes (50.0)

* % of individuals in the CCS group who also have one of the four chronic conditions (mental illness, heart disease, cancer, and diabetes) considered in our penalized and constrained regressions.
***Mental illness is excluded from the group CCS measures since it is already represented separately in the table*

Appendix D: HCC used in RXC x HCC interactions and their corresponding RXC

HCC	HCC Definition	RXC
HCC001	HIV/AIDS	Anti-HIV Agents
HCC018	Pancreas Transplant Status/Complications	Insulin; Anti-Diabetic Agents, Except Insulin and Metformin Only
HCC019	Diabetes with Acute Complications	
HCC020	Diabetes with Chronic Complications	
HCC021	Diabetes without Complication	
HCC034	Liver Transplant Status/Complications	Anti-Hepatitis C (HCV) Agents, Direct Acting Agents
HCC035	End-Stage Liver Disease	
HCC036	Cirrhosis of Liver	
CC037_1	Chronic Viral Hepatitis C	
HCC041	Intestine Transplant Status/Complications	Inflammatory Bowel Disease Agents; Immune Suppressants and Immunomodulators
HCC048	Inflammatory Bowel Disease	
HCC056	Rheumatoid Arthritis and Specified Autoimmune Disorders	Immune Suppressants and Immunomodulators
HCC057	Systemic Lupus Erythematosus and Other Autoimmune Disorders	
HCC118	Multiple Sclerosis	Multiple Sclerosis Agents

HCC142	Specified Heart Arrhythmias	Antiarrhythmics
HCC158	Lung Transplant Status/Complications	Cystic Fibrosis Agents
HCC159	Cystic Fibrosis	
HCC183	Kidney Transplant Status	Phosphate Binders
HCC184	End Stage Renal Disease	
HCC187	Chronic Kidney Disease, Stage 5	
HCC188	Chronic Kidney Disease, Severe (Stage 4)	