**Appendix**

*Baseline Risk Adjustment Formula*

As described in the main text, the baseline risk adjustment regression we implement is a similar specification to the Version 21 CMS-HCC formula for the aged, community-dwelling subgroup. We simplify our formula and do not include interaction terms between diseases, which are sparsely populated, and we also do not modify coefficients post-estimation as CMS does (Pope, et al. 2011). To calculate risk scores, CMS converts the dollar risk adjustment coefficients into relative risk factors by dividing by the average per person predicted spending for a designated year. In our analyses we use the coefficients as dollar weights. With an analysis for one year, these methods are equivalent.

Appendix Table 1: Baseline Risk Adjustment Formula HCC Indicators

| **HCC Indicator** | **Description** |
|---|---|
| HCC1 | HIV/AIDS |
| HCC2 | Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock |
| HCC6 | Opportunistic Infections |
| HCC8 | Metastatic Cancer and Acute Leukemia |
| HCC9 | Lung and Other Severe Cancers |
| HCC10 | Lymphoma and Other Cancers |
| HCC11 | Colorectal, Bladder, and Other Cancers |
| HCC12 | Breast, Prostate, and Other Cancers and Tumors |
| HCC17 | Diabetes with Acute Complications |
| HCC18 | Diabetes with Chronic Complications |
| HCC19 | Diabetes without Complication |
| HCC21 | Protein-Calorie Malnutrition |
| HCC22 | Morbid Obesity |
| HCC23 | Other Significant Endocrine and Metabolic Disorders |
| HCC27 | End-Stage Liver Disease |
| HCC28 | Cirrhosis of Liver |
| HCC29 | Chronic Hepatitis |
| HCC33 | Intestinal Obstruction/Perforation |
| HCC34 | Chronic Pancreatitis |
| HCC35 | Inflammatory Bowel Disease |

| | |
|---|---|
| HCC39 | Bone/Joint/Muscle Infections/Necrosis |
| HCC40 | Rheumatoid Arthritis and Inflammatory Connective Tissue Disease |
| HCC46 | Severe Hematological Disorders |
| HCC47 | Disorders of Immunity |
| HCC48 | Coagulation Defects and Other Specified Hematological Disorders |
| HCC51 | Dementia with Complications |
| HCC52 | Dementia without Complication |
| HCC54 | Drug/Alcohol Psychosis |
| HCC55 | Drug/Alcohol Dependence |
| HCC57 | Schizophrenia |
| HCC58 | Major Depressive, Bipolar, and Paranoid Disorders |
| HCC70 | Quadriplegia |
| HCC71 | Paraplegia |
| HCC72 | Spinal Cord Disorders/Injuries |
| HCC73 | Amyotrophic Lateral Sclerosis and Other Motor Neuron Disease |
| HCC74 | Cerebral Palsy |
| HCC75 | Polyneuropathy |
| HCC76 | Muscular Dystrophy |
| HCC77 | Multiple Sclerosis |
| HCC78 | Parkinson's and Huntington's Diseases |
| HCC79 | Seizure Disorders and Convulsions |
| HCC80 | Coma, Brain Compression/Anoxic Damage |
| HCC82 | Respirator Dependence/Tracheostomy Status |
| HCC83 | Respiratory Arrest |
| HCC84 | Cardio-Respiratory Failure and Shock |
| HCC85 | Congestive Heart Failure |
| HCC86 | Acute Myocardial Infarction |
| HCC87 | Unstable Angina and Other Acute Ischemic Heart Disease |
| HCC88 | Angina Pectoris |
| HCC96 | Specified Heart Arrhythmias |
| HCC99 | Cerebral Hemorrhage |
| HCC100 | Ischemic or Unspecified Stroke |
| HCC103 | Hemiplegia/Hemiparesis |
| HCC104 | Monoplegia, Other Paralytic Syndromes |
| HCC106 | Atherosclerosis of the Extremities with Ulceration or Gangrene |
| HCC107 | Vascular Disease with Complications |
| HCC108 | Vascular Disease |
| HCC110 | Cystic Fibrosis |
| HCC111 | Chronic Obstructive Pulmonary Disease (COPD) |
| HCC112 | Fibrosis of Lung and Other Chronic Lung Disorders |
| HCC114 | Aspiration and Specified Bacterial Pneumonias |
| HCC115 | Pneumococcal Pneumonia, Empyema, Lung Abscess |

| | |
|---|---|
| HCC122 | Proliferative Diabetic Retinopathy and Vitreous Hemorrhage |
| HCC124 | Exudative Macular Degeneration |
| HCC134 | Dialysis Status |
| HCC135 | Acute Renal Failure |
| HCC136 | Chronic Kidney Disease, Stage 5 |
| HCC137 | Chronic Kidney Disease, Severe (Stage 4) |
| HCC138 | Chronic Kidney Disease, Moderate (Stage 3) |
| HCC139 | Chronic Kidney Disease, Mild or Unspecified (Stages 1-2 or Unspecified) |
| HCC140 | Unspecified Renal Failure |
| HCC141 | Nephritis |
| HCC157 | Pressure Ulcer of Skin with Necrosis through to Muscle, Tendon, or Bone |
| HCC158 | Pressure Ulcer of Skin with Full Thickness Skin Loss |
| HCC159 | Pressure Ulcer of Skin with Partial Thickness Skin Loss |
| HCC160 | Pressure Pre-Ulcer of Skin Changes or Unspecified Stage |
| HCC161 | Chronic Ulcer of Skin, Except Pressure |
| HCC162 | Severe Skin Burn or Condition |
| HCC166 | Severe Head Injury |
| HCC167 | Major Head Injury |
| HCC169 | Vertebral Fractures without Spinal Cord Injury |
| HCC170 | Hip Fracture/Dislocation |
| HCC173 | Traumatic Amputations and Complications |
| HCC176 | Complications of Specified Implanted Device or Graft |
| HCC186 | Major Organ Transplant or Replacement Status |
| HCC188 | Artificial Openings for Feeding or Elimination |
| HCC189 | Amputation Status, Lower Limb/Amputation Complications |

Note: The Version 21 CMS-HCC Formula includes the following interaction terms that are omitted from our baseline risk adjustment regression formula: cancer (HCC8, HCC9, HCC10, HCC12) with immune disorders (HCC47); sepsis (HCC2) with cardio-respiratory failure (HCC82, HCC83, HCC84); COPD (HCC110, HCC111) with cardio-respiratory failure; COPD with chronic heart failure (HCC85); chronic heart failure with diabetes (HCC17, HCC18, HCC19); and chronic heart failure with renal disease (HCC 134, HCC135, HCC136, HCC137, HCC138, HCC139, HCC140, HCC141).

*Super Learner*

Machine learning is well-suited to estimating disparities because our definition is formulated as a prediction exercise and the functional form of any differences is not known in advance. We therefore adjust for health status in a general and flexible way, using ensemble machine learning methods. We use the super learner framework, which combines many algorithms into a weighted average prediction function (van der Laan, Polley and Hubbard 2007, van der Laan and Rose 2011). While it has not, to our knowledge, been used in disparities work, it has been

deployed for prediction or classification in other applications, such as health outcomes (e.g., Petersen, et al. 2015 and Pirracchio, et al. 2015) and health spending (e.g., Rose 2016; Rose, Bergquist and Layton 2017; Park and Basu 2018).

The super learner takes as input: (1) the data and (2) a collection of algorithms. The first step involves performing V-fold cross-validation for each algorithm in the collection. An optimal weight vector is then constructed by regressing the outcome $Y_i$ on the cross-validated predicted spending values generated for each algorithm. The estimated coefficient for each column in the matrix of cross-validated predicted spending values is the weight for the algorithm that produced those values. One can show that this regression selects the optimal weight vector, and that by restricting the family of possible weighted combinations to convex combinations (i.e., nonnegative weights that sum to one) performance can be improved (van der Laan, Polley and Hubbard 2007). The collection of algorithms the super learner ensembles can include not only different classes of algorithms (e.g., decision trees and penalized regressions), but also multiple similar algorithms with alternative tuning parameters (e.g., random forests with 250 observations in terminal nodes vs. 500 observations in terminal nodes). Algorithms, such as penalized regressions, that may select tuning parameters (e.g., λ), via cross-validation will still do so within the super learner procedure, leading to nested layers of cross-validation. The final super learner algorithm fits each algorithm on the full data and combines those functions with the estimated optimal weights to produce a final predicted value. Our super learner considered six algorithms: a main terms logistic regression, three penalized regressions, and a random forest with 500 trees and a minimum node size of 250 observations. We implement our analyses using the SuperLearner package in R, which calls the randomForest and glmnet packages (Liaw and Wiener 2002; Friedman, Hastie and Tibshirani 2010).

*Additional Risk Adjustors*

We add three risk adjustors from the Medicare Part D risk adjustment formula (CMS-RxHCC), which has been shown to recognize more individuals with mental health conditions in the ACA Marketplace setting (Montz, et al. 2016). The goal is to choose additional risk adjustors to capture individuals in the mental illness group (as defined by CCS groups) not already well represented by the included HCCs. Appendix Table 2 shows the overlap in underlying ICD-9 codes for each mental illness CCS group and the corresponding HCCs and RxHCCs. We select indicators for bipolar disorders, major depression, and depression to include in our regression. To pick these three risk adjustors, we first examined RxHCCs with the fewest ICD-9 codes in common with the mental illness HCCs, and then we chose the RxHCCs that covered the largest number of individuals in the CCS mental illness group.

Appendix Table 2: CCS ICD9 Codes cross walk with HCCs and RxHCCs

| **ICD9 Codes** | **CCS Group** | **HCC** | **RxHCC** |
|---|---|---|---|
| 3090 3091 30922 30923 30924 30928 30929 3093 3094 30982 30983 30989 3099 | Adjustment disorders (650) | N/A | Depression (62) |
| 29384 30000 30001 30002 30009 30010 30020 30021 30022 30023 30029 3003 3005 30089 3009 3080 3081 3082 3083 3084 3089 30981 3130 3131 31321 31322 3133 31382 31383 | Anxiety Disorders (651) | N/A | Anxiety Disorders (63) Specified Anxiety, Personality, and Behavior Disorders (61) |
| 31200 31201 31202 31203 31210 31211 31212 31213 31220 31221 31222 31223 3124 3128 31281 31282 31289 3129 31381 31400 31401 3141 3142 3148 3149 | ADHD, conduct, and disruptive behavior disorders (652) | N/A | Specified Anxiety, Personality, and Behavior Disorders (61) |
| 2900 29010 29011 29012 29013 29020 29021 2903 29040 29041 29042 29043 2908 2909 2930 2931 2940 2941 29410 29411 29420 29421 2948 2949 | Delirium, dementia, and amnestic and other cognitive disorders (653) | Dementia without Complication (52) Dementia with Complications (51) | Dementia, Except Alzheimer's Disease (55) |

| | | | |
|---|---|---|---|
| 3100 3102 3108 31081 31089 3109 3310 3311 33111 33119 3312 33182 797 | | | Alzheimer's Disease (54) |
| 3070 3079 31500 31501 31502 31509 3151 3152 31531 31532 31534 31535 31539 3154 3155 3158 3159 317 3180 3181 3182 319 V400 V401 | Developmental disorders (654) | N/A | Mild or Unspecified Mental Retardation/ Developmental Disability (68)<br>Profound or Severe Mental Retardation/ Developmental Disability (66)<br>Moderate Mental Retardation/ Developmental Disability (67) |
| 29900 29901 29910 29911 29980 29981 29990 29991 30720 30721 30722 30723 3073 3076 3077 30921 31323 31389 3139 | Disorders usually diagnosed in infancy, childhood, or adolescence (655) | N/A | Specified Anxiety, Personality, and Behavior Disorders (61)<br>Autism (65) |
| 31230 31231 31232 31233 31234 31235 31239 | Impulse control disorders, NEC (656) | N/A | Specified Anxiety, Personality, and Behavior Disorders (61) |
| 29383 29600 29601 29602 29603 29604 29605 29606 29610 29611 29612 29613 29614 29615 29616 29620 29621 29622 29623 29624 29625 29626 29630 29631 29632 29633 29634 29635 29636 29640 29641 29642 29643 29644 29645 29646 29650 29651 29652 29653 29654 29655 29656 29660 | Mood disorders (657) | Major Depressive, Bipolar, and Paranoid Disorders (58) | Bipolar Disorders (59)<br>Major Depression (60)<br>Depression (62) |

| | | | |
|---|---|---|---|
| 29661 29662 29663 29664 29665 29666 2967 29680 29681 29682 29689 29690 29699 3004 311 | | | |
| 3010 30110 30111 30112 30113 30120 30121 30122 3013 3014 30150 30151 30159 3016 3017 30181 30182 30183 30184 30189 3019 | Personality disorders (658) | N/A | Depression (62) Specified Anxiety, Personality, and Behavior Disorders (61) |
| 29381 29382 29500 29501 29502 29503 29504 29505 29510 29511 29512 29513 29514 29515 29520 29521 29522 29523 29524 29525 29530 29531 29532 29533 29534 29535 29540 29541 29542 29543 29544 29545 29550 29551 29552 29553 29554 29555 29560 29561 29562 29563 29564 29565 29570 29571 29572 29573 29574 29575 29580 29581 29582 29583 29584 29585 29590 29591 29592 29593 29594 29595 2970 2971 2972 2973 2978 2979 2980 2981 2982 2983 2984 2988 2989 | Schizophrenia and other psychotic disorders (659) | Schizophrenia (57) Major Depressive, Bipolar, and Paranoid Disorders (58) | Schizophrenia (58) |
| 2910 2911 2912 2913 2914 2915 2918 29181 29182 29189 2919 30300 30301 30302 30303 30390 30391 30392 30393 30500 30501 30502 30503 3575 4255 5353 53530 53531 5710 5711 5712 5713 76071 9800 | Alcohol-related disorders (660) | Drug/Alcohol Psychosis (54) Drug/Alcohol Dependence (55) Polyneuropathy (75) Congestive Heart Failure (85) Cirrhosis of Liver (28) | Polyneuropathy (74) Congestive Heart Failure (87) |
| 2920 29211 29212 2922 29281 29282 29283 29284 29285 29289 2929 30400 30401 30402 30403 30410 30411 30412 30413 30420 30421 30422 30423 30430 30431 30432 30433 30440 30441 30442 30443 30450 30451 30452 30453 30460 | Substance-related disorders (661) | Drug/Alcohol Psychosis (54) Drug/Alcohol Dependence (55) | N/A |

| | | | |
|---|---|---|---|
| 30461 30462 30463 30470 30471 30472 30473 30480 30481 30482 30483 30490 30491 30492 30493 30520 30521 30522 30523 30530 30531 30532 30533 30540 30541 30542 30543 30550 30551 30552 30553 30560 30561 30562 30563 30570 30571 30572 30573 30580 30581 30582 30583 30590 30591 30592 30593 64830 64831 64832 64833 64834 65550 65551 65553 76072 76073 76075 7795 96500 96501 96502 96509 V6542 | | | |
| E9500 E9501 E9502 E9503 E9504 E9505 E9506 E9507 E9508 E9509 E9510 E9511 E9518 E9520 E9521 E9528 E9529 E9530 E9531 E9538 E9539 E954 E9550 E9551 E9552 E9553 E9554 E9555 E9556 E9557 E9559 E956 E9570 E9571 E9572 E9579 E9580 E9581 E9582 E9583 E9584 E9585 E9586 E9587 E9588 E9589 E959 V6284 | Suicide and intentional self-inflicted injury (662) | N/A | N/A |
| 3051 30510 30511 30512 30513 33392 7903 V110 V111 V112 V113 V114 V118 V119 V154 V1541 V1542 V1549 V1582 V6285 V663 V701 V702 V7101 V7102 V7109 V790 V791 V792 V793 V798 V799 | Screening and history of mental health and substance abuse codes (663) | N/A | N/A |
| 29389 2939 30011 30012 30013 30014 30015 30016 30019 3006 3007 30081 30082 3021 3022 3023 3024 30250 30251 30252 30253 3026 30270 30271 30272 30273 30274 30275 30276 30279 30281 30282 30283 30284 30285 30289 3029 3060 3061 3062 3063 3064 30650 30651 30652 30653 30659 3066 | Miscellaneous mental health disorders (670) | N/A | Specified Anxiety, Personality, and Behavior Disorders (61) Anxiety Disorders (63) |

| | | | |
|---|---|---|---|
| 3067 3068 3069 3071<br>30740 30741 30742 30743<br>30744 30745 30746 30747<br>30748 30749 30750 30751<br>30752 30753 30754 30759<br>30780 30781 30789 3101<br>316 64840 64841 64842<br>64843 64844 V402 V403<br>V4031 V4039 V409 V673 | | | |

*Disparities Application Extension*

We expand on the primary disparities example by demonstrating an even more targeted intervention: once again spending is increased by 10% for the low-income group, but funds are taken from only the upper half of the high-income group (the fifth quintile), leaving individuals in the fourth income quintile unaffected. To accomplish this we transform the *Y* vector as before for the low-income group, imposing an overall budget constraint, and imposing a second constraint that maintains the spending levels of the fourth quintile, so that only individuals in the fifth quintile are used to counterbalance the spending for the low-income group.

Appendix Table 3 shows increasing spending by 10% for the low-income group but reducing spending for only the fifth income quintile generates minimal change in payments for the low-income group. Adding a risk adjustor for the low-income group does not allow us to specify what happens within the high-income group, and we see the fourth income quintile receives substantially higher payments. Combining adding a risk adjustor with data transformation is once again the most powerful approach, although the overall changes are much more modest (from $10,050 to $10,058 for the low-income group) than when we allowed spending for the entire high-income group to be reduced.

Appendix Table 3: Reducing Disparities by Targeting Neighborhood Income Subgroup – Comparing Group Payments and Spending (in U.S. Dollars, 2011)

| Income Group | Base OLS | Data Transformation | Low-Income Indicator | Combination | Mean Spending |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $R^2$ | 11.38 | 11.32 | 11.38 | 11.27 | |
| *Binary* | | | | | |
| Low (Quintiles 1-3) | 10,055 | 10,104 | 9,929 | 10,0922 | 9,928 |
| High (Quintiles 4-5) | 9,660 | 9,587 | 9,850 | 9,851 | 9,852 |
| *Quintiles* | | | | | |
| 1st | 10,195 | 10,310 | 10,075 | 11,080 | 10,235 |
| 2nd | 10,063 | 10,100 | 9,9936 | 10,928 | 9,847 |
| 3rd | 9,908 | 9,902 | 9,777 | 10,758 | 9,702 |
| 4th | 9,773 | 9,726 | 9,964 | 9,975 | 9,726 |
| 5th | 9,548 | 9,449 | 9,735 | 9,727 | 9,977 |

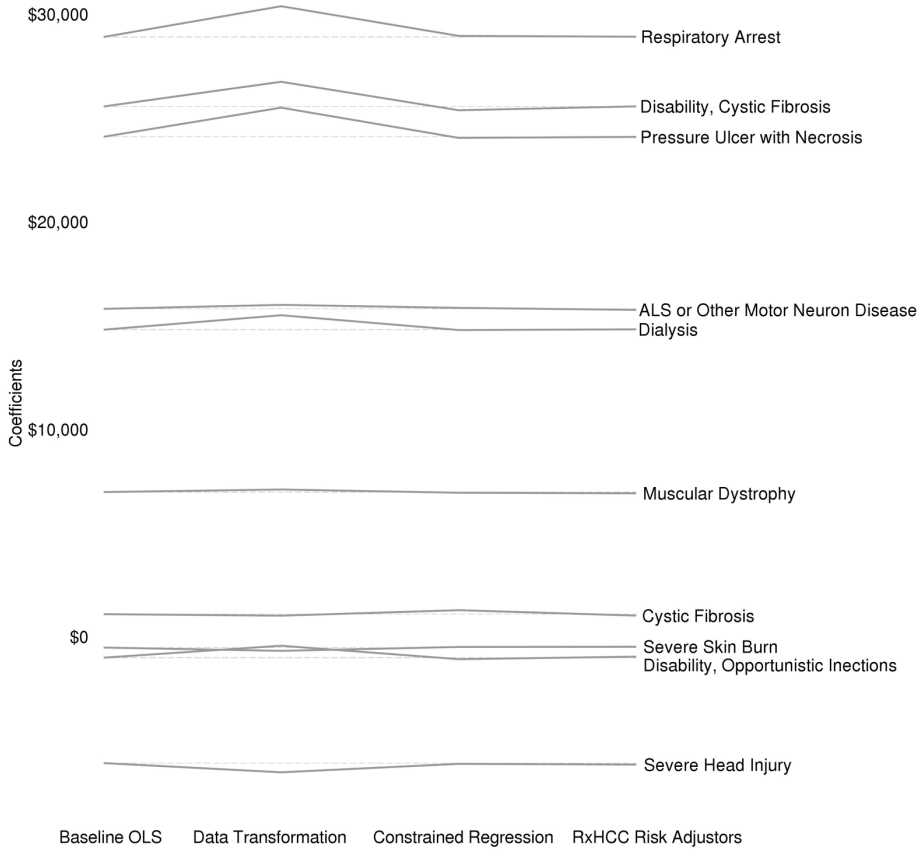
Notes: Column (1) contains the simplified CMS-HCC v21 risk adjustment regression. (2) contains the regression where we intervene to increase spending by 10% for the low-income group, but only reduce spending for the 5th income quintile. (3) is the baseline OLS regression with a low-income group indicator. (4) contains the baseline OLS regression with a low-income group indicator, combined with transforming the data to increase spending by 10% for the low-income group and reducing spending for the 5th income quintile. (5) shows mean group-level spending.

Appendix Table 4: Reducing Disparities by Targeting Neighborhood Income Subgroup – Comparing Group Deviations from Mean Spending (in U.S. Dollars, 2011)

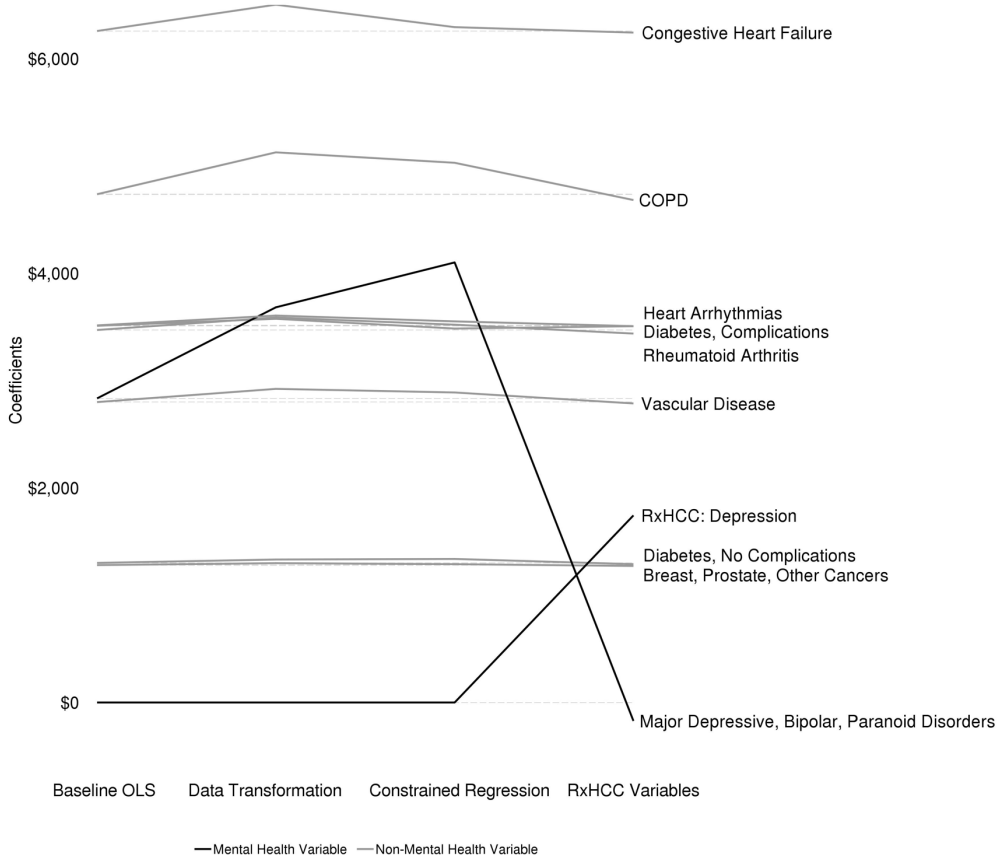
| Income Group | Base OLS | Data Transformation | Low-Income Indicator | Combination | Mean Spending |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| *Binary* | | | | | |
| Low (Quintiles 1-3) | 128 | 152 | 1 | 994 | 9,928 |
| High (Quintiles 4-5) | -191 | -228 | -2 | 0 | 9,852 |
| *Quintiles* | | | | | |
| 1st | -40 | 15 | -519 | 846 | 10,235 |
| 2nd | 216 | 236 | 88 | 1,081 | 9,847 |
| 3rd | 207 | 205 | 76 | 1,056 | 9,702 |
| 4th | 46 | 24 | 237 | 249 | 9,726 |
| 5th | -429 | -481 | -242 | -250 | 9,977 |

Notes: Deviation from mean spending is calculated as the difference between average predicted spending and average observed spending within each group; positive numbers represent predicted spending exceeding observed spending.

Appendix Figure 1: Coefficient Differences Across Risk Adjustment Alternatives Targeting Mental Illness Group – Ten Most Rare Conditions



Appendix Figure 2: Coefficient Differences Across Risk Adjustment Alternatives Targeting Mental Illness Group – Ten Most Common Conditions



Appendix Figure 3: Coefficient Differences Across Risk Adjustment Alternatives Targeting Mental Illness Group – Mental Health Conditions