**Supplemental Materials for**

“Association between Residential Neighborhood Social Conditions

and Healthcare Utilization and Costs”

[Appendix 1 Selection of the study sample 2](#_Toc30334770)

[Appendix 2 Linking Medicare claims and Social Determinants of Health Data 3](#_Toc30334771)

[Appendix 3 Model specifications 4](#_Toc30334772)

[Appendix 4 Geography of Study Sample and Variation in Area Deprivation Index in the Catchment Area of New York and New Jersey 5](#_Toc30334773)

[Appendix 5 Variation in Area Deprivation Index in the Catchment Area of New York City 6](#_Toc30334774)

[Appendix 6 Unadjusted Association of ADI Quintiles with Total Medicare Costs and Costs by Care Settings 7](#_Toc30334775)

[Appendix 7 Unadjusted Association of ADI Quintiles and Preventable Costs and Rates of Preventable Utilization by ADI Quintiles 8](#_Toc30334776)

[Appendix 8 Results After Adding Variables of Primary Care Resources and Utilization Pattern 9](#_Toc30334777)

[Appendix 9 Full regression results 11](#_Toc30334779)

# Appendix 1 Selection of the study sample

763,162 patients who were active Medicare patients when they visited the seven health systems affiliated with INSIGHT in 2013

Excluded 603,883 patients without 9-digit zip codes (without any zip code, invalid zip codes, or only 5-digit zip codes)

159,279 patients with 9-digit zip codes who were active Medicare patients when they visited the seven health systems affiliated with INSIGHT in 2013

100,089 Medicare FFS or dual-eligible patients who were continuously enrolled in Medicare Parts A and B in 2013

109,336 Medicare FFS or dual-eligible patients with 9-digit zip codes in New York or New Jersey identified from seven health systems affiliated with INSIGHT

93,557 Medicare FFS or dual-eligible patients who were continuously enrolled in Medicare Parts A and B in 2013 and 2014

Excluded 128 patients that the Area Deprivation Index rank was not available in the Neighborhood Atlas data (e.g., the area is too small to have a precise estimate on neighborhood conditions) or zip codes are PO boxes instead of residential address

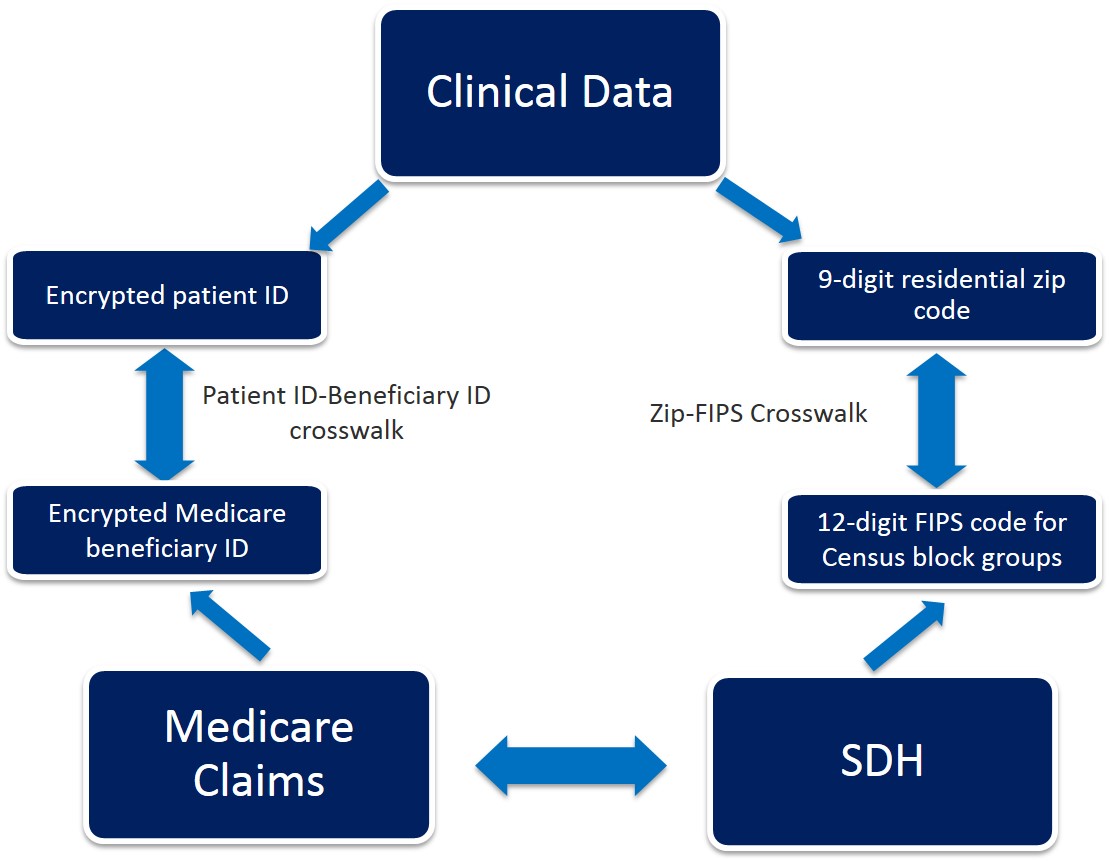
Excluded 49,943 patients not from New Jersey or New York (i.e., from other states or other U.S. territories)

Excluded 9,247 patients who had Medicare Advantage enrollment or without 12-month Part A or Part B enrollment in 2013 (including those who deceased in 2013)

Excluded 6,532 patients who had Medicare Advantage enrollment or without 12-month Part A or Part B enrollment in 2014 (including those who deceased in 2014)

93,429 patients with 12-month Part A or Part B enrollment in 2014 were included in the study

# Appendix 2 Linking Medicare claims and Social Determinants of Health Data



# Appendix 3 Model specifications

**Generalized linear model:**

The distribution of cost data usually is heavily skewed to the right as a small percentage of patients incur extremely high costs compared to other patients. In addition, the assumption of homoscedasticity (constant variance) is often violated because cost data exhibit variability that increases as the mean cost increases. Generalized linear model (GLM) is commonly used to address the skewness in the distribution and heteroscedasticity. Instead of modeling means for cost as in an ordinal least square (OLS) linear regression or means of log-transformed costs as in a log-OLS: , the GLM models the natural log of the mean cost (). The gamma distribution of variance was specified. Specifically, we modeled:

+

is the total costs or preventable costs for each patient in each year; are indicators for Area Deprivation Quintiles (Quintile 3 is the omitted reference group). is a vector for control variables, including age categories (i.e., <=65 (ref.), 65-74, 75-84, >=85); gender (i.e., male and female), race (i.e., white, black, Hispanic, or other), indicator for dual-eligible, indicator for Part D coverage, indicator for ESRD, indicator for enrollment because of disability, indicators for number of chronic conditions (i.e., 0, 1-2, 3-5, 6-8, >=9), indicator for frailty, indicator for mental illness, indicator for serious illness, HCC score, indicators for years, and indicators for hospital referral regions. The error term is assumed to follow a gamma distribution, i.e. ~ Г(k, θ), where k is a shape parameter, and θ is a scale parameter, both positive.

The coefficients of GLM are interpreted as percent changes of the mean outcome. We also calculated the marginal effects to present the cost changes in dollar amount.

**Logistic regression model:** For the binary outcomes, we used the logistic regression to examine the association between

+

is the probability of having a type of preventable utilization (i.e., 30-day readmission, preventable ED visit, and preventable hospitalization) for patient i in year t; are indicators for Area Deprivation Quintiles (Quintile 3 is the omitted reference group). is a vector of control variables, including age categories (i.e., <=65 (ref.), 65-74, 75-84, >=85); gender (i.e., male and female), race (i.e., white, black, Hispanic, or other), indicator for dual-eligible, indicator for Part D coverage, indicator for ESRD, indicator for enrollment because of disability, indicators for number of chronic conditions (i.e., 0, 1-2, 3-5, 6-8, >=9), indicator for frailty, indicator for mental illness, indicator for serious illness, HCC score, indicator for years, and indicators for hospital referral regions. is the error term.

# Appendix 4 Geography of Study Sample and Variation in Area Deprivation Index in the Catchment Area of New York and New Jersey

C:\Users\yoz2009\Dropbox\WCM\Project\High Utilizer\manuscript\new papers\manuscript 4\draft\Submission\Medical Care\revision\resubmission\Final_editor comments\map_final\Appendix4_R2.tif

*Notes: ADI: Area Deprivation Index; Quintile 1 represent the least disadvantaged area and quintile 5 present the most disadvantaged area.*

# Appendix 5 Variation in Area Deprivation Index in the Catchment Area of New York City



*Notes: ADI: Area Deprivation Index; Quintile 1 represent the least disadvantaged area and quintile 5 present the most disadvantaged area.*

# Appendix 6 Unadjusted Association of ADI Quintiles with Total Medicare Costs and Costs by Care Settings

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Average total annual Medicare costs ($) | Physician ($) | Outpatient ($) | Inpatient ($) | Post-acute care and hospice ($) | DME ($) | Part D ($) |
| ADI, quintile 1 | 17,661 | 5,928 | 2,460 | 3,744 | 1,732 | 230 | 3,566 |
| ADI, quintile 2 | 22,965 | 6,880 | 3,210 | 5,260 | 2,400 | 325 | 4,890 |
| ADI, quintile 3 | 23,841 | 6,867 | 3,311 | 5,305 | 2,665 | 336 | 5,356 |
| ADI, quintile 4 | 21,564 | 6,174 | 2,965 | 4,830 | 2,567 | 312 | 4,716 |
| ADI, quintile 5 | 26,815 | 6,174 | 4,154 | 6,001 | 2,969 | 432 | 6,860 |
| Average | 20,670 | 6,398 | 2,901 | 4,514 | 2,147 | 285 | 4,425 |

*Notes: ADI: Area Deprivation Index. We compared the differences in cost between quintile 1 and each other quintile (quintiles 2-5) using t-test and the difference are all statistically significant (P<0.001).*

# Appendix 7 Unadjusted Association of ADI Quintiles and Preventable Costs and Rates of Preventable Utilization by ADI Quintiles

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Potentially preventable Medicare costs ($) | Preventable ED visits (%) | Preventable hospitalizations (%) | 30-day readmissions (%) |
| ADI, quintile 1 | 1,343 | 8.7 | 3.7 | 13.7 |
| ADI, quintile 2 | 2,260 | 11.9 | 5.4 | 16.8 |
| ADI, quintile 3 | 2,568 | 13.3 | 6.1 | 18.3 |
| ADI, quintile 4 | 2,232 | 13.5 | 5.6 | 18.9 |
| ADI, quintile 5 | 3,361 | 18.5 | 8.1 | 20.0 |
| Average | 1,918 | 11.1 | 4.8 | 3.8 |

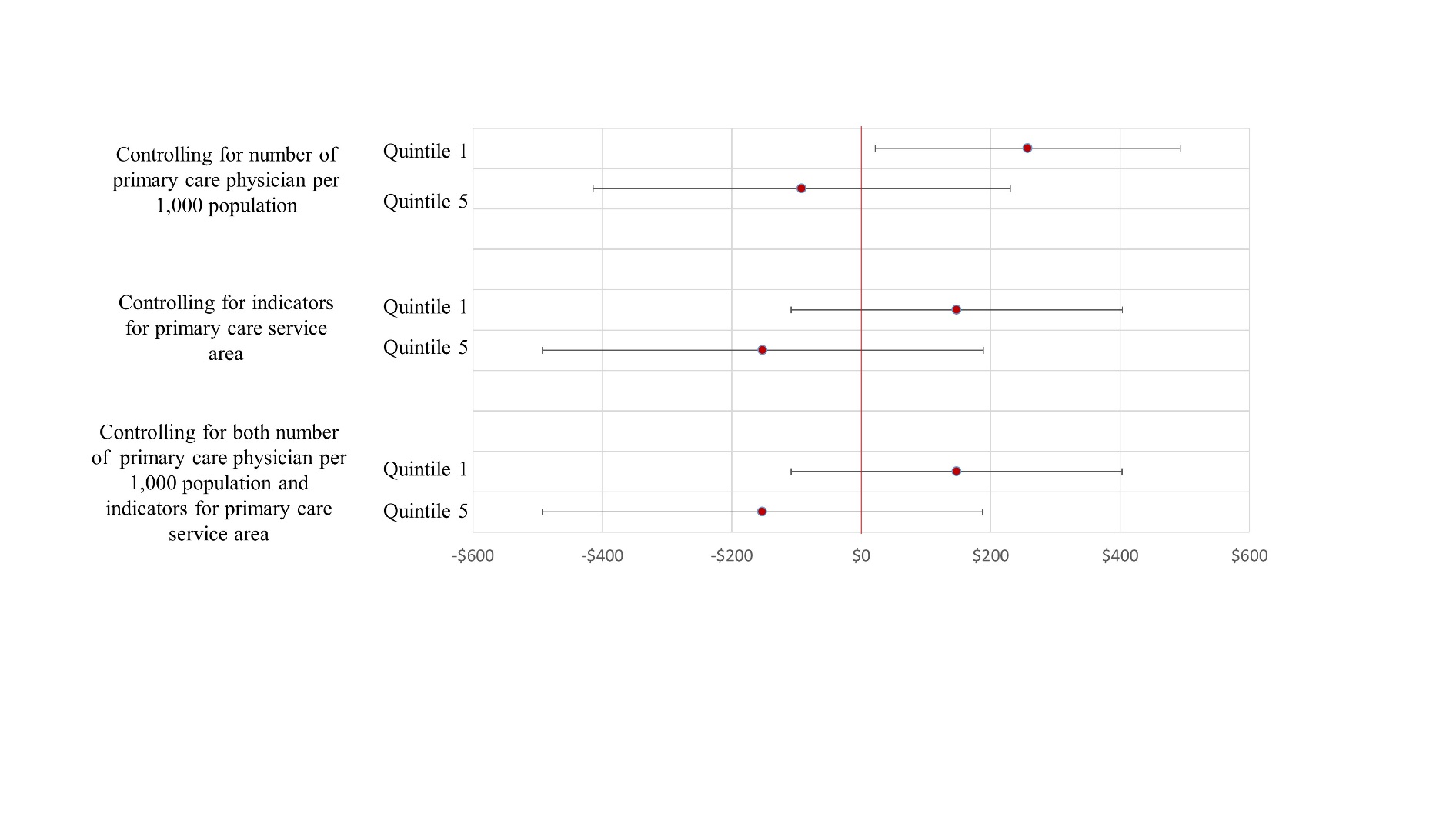
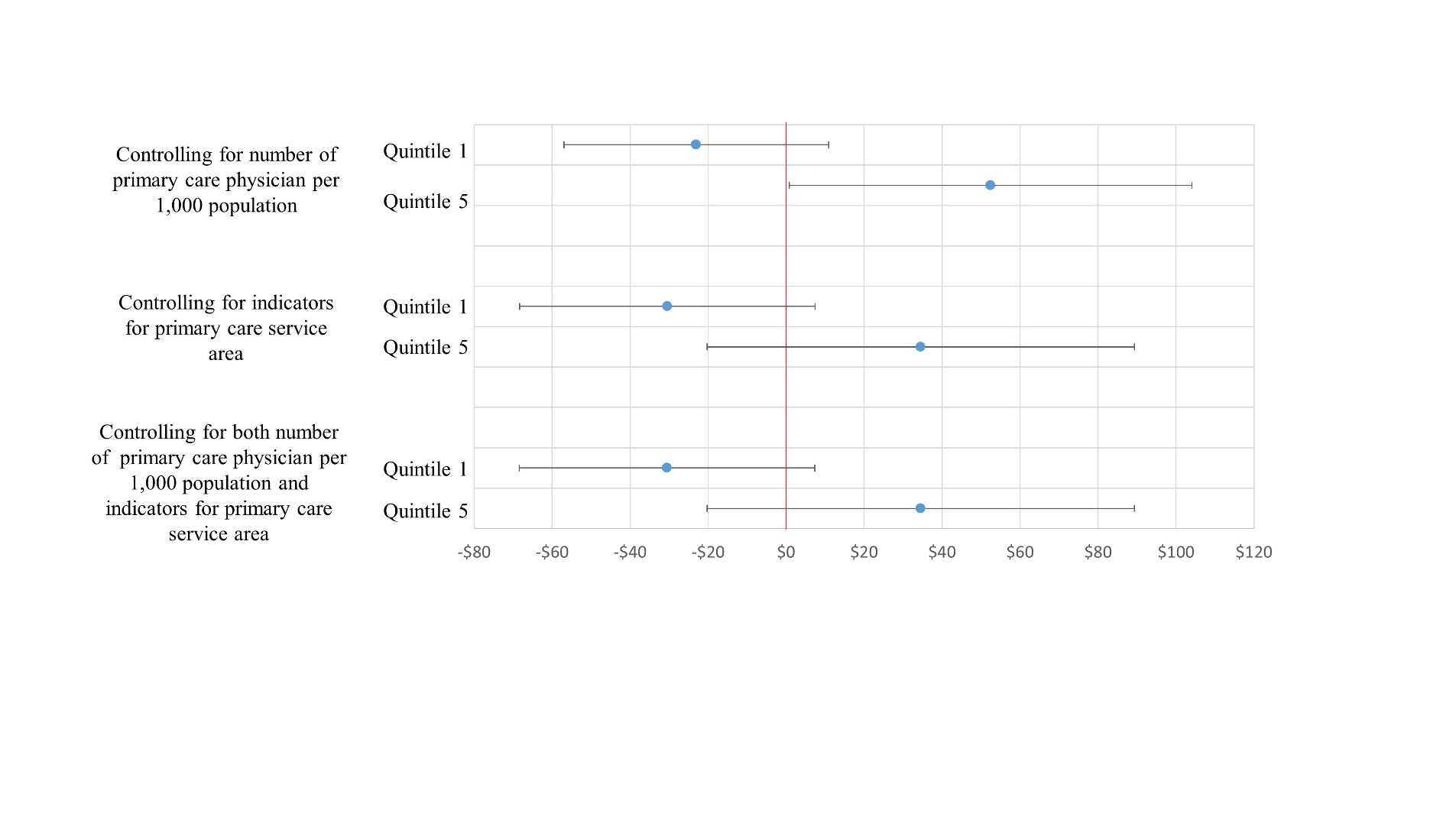
*Notes: ADI: Area Deprivation Index. We compared the differences in cost between quintile 1 and each other quintile (quintiles 2-5) using t-test and the difference are all statistically significant (P<0.001).*

# Appendix 8 Results After Adding Variables of Primary Care Resources and Utilization Pattern

To explore the potential pathway between SDoH and healthcare costs and utilization, we added the following additional variables in the model of the primary analysis: (1) number of primary care physicians per 1,000 population in the primary care service areas (PCSAs) for the availability of primary care resources; (2) indicators for PCSAs that reflect the overall primary care resources and utilization patterns; (3) indicators for PCSAs plus the number of primary care physicians per 1,000 population. If the availability of primary care resources and utilization pattern could explain the relationships between SDoH and outcomes, we expect that associations between ADI quintiles and outcomes will attenuated (e.g., the coefficients will be smaller or insignificant) after adding these additional variables.

As we presented in the following figures, the difference in total Medicare costs between quintile 1 and quintile 3 decreased (from $427 to $256) after adding the number of primary care physicians per 1,000 population, but it was still statistically significant. The difference was insignificant after adding PCSA indicators or both number of primary care physicians and PCSA indicators. The difference between quintile 5 and quintile 3 did not change significantly after adding the number of primary care physicians per 1,000 population.

For the preventable costs, adding the number of primary care physicians per 1,000 population did not change the difference between quintiles 5 and 3. However, the difference in preventable costs between quintiles 5 and 3 was not statistically significant after adding PCSA indicators or both number of primary care physicians and PCSA indicators.



Panel A: Total Medicare costs

Panel B: Preventable costs

# Appendix 9 Full regression results

Appendix 9a Full regression results for total Medicare costs and costs by care settings

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Total Medicare costs | Physician costs | Outpatient costs | Inpatient costs | Post-acute care costs | DME costs | Prescription costs |
|  | **Coefficient** | **Coefficient** | **Coefficient** | **Coefficient** | **Coefficient** | **Coefficient** | **Coefficient** |
| ADI quintiles |  |  |  |  |  |  |  |
| Quintile 1 | 0.03 (0.009) \*\*\* | 0.05 (0.008) \*\*\* | 0.02 (0.022) | 0.08 (0.029)  \*\* | -0.16 (0.053) \*\* | -0.04 (0.042) | 0.30 (0.26) |
| Quintile 2 | 0.02 (0.010)  \* | 0.02 (0.009)  \* | -0.08 (0.024)  \*\* | 0.10 (0.033)  \*\* | -0.03 (0.059) | -0.02 (0.047) | 0.34 (0.29) |
| Quintile 3 (ref.) | – | – | – | – | – | – | – |
| Quintile 4 | -0.009  (0.014) | -0.03  (0.014) | -0.02 (0.036) | 0.02 (0.048) | 0.09 (0.087) | -0.03 (0.068) | 0.54 (0.45) |
| Quintile 5 | -0.002  (0.012) | -0.04  (0.012)  \*\* | 0.03 (0.031) | -0.03 (0.042) | 0.02 (0.076) | 0.04 (0.059) | 0.10 (0.36) |
| Age categories |  |  |  |  |  |  |  |
| <65 (ref.) | – | – | – | – | – | – | – |
| 65-74 | -0.42 (0.013) \*\*\* | -0.19 (0.013) \*\*\* | -0.47 (0.033)  \*\*\* | -0.85 (0.045) \*\*\* | 1.03 (0.078) | -0.35 (0.062) | -0.02 (0.43) |
| 75-84 | -0.50 (0.014) \*\*\* | -0.23 (0.013) \*\*\* | -0.61 (0.035)  \*\*\* | -0.98 (0.048) \*\*\* | 0.58 (0.084) \*\*\* | -0.42 (0.067) \*\*\* | -0.31 (0.46) |
| >=85 | -0.62 (0.015) \*\*\* | -0.39 (0.015) \*\*\* | -0.98 (0.039)  \*\*\* | -1.15 (0.052) \*\*\* | 1.36 (0.091) \*\*\* | -0.53 (0.072) \*\*\* | -0.73 (0.49) |
| Gender |  |  |  |  |  |  |  |
| Male (ref.) | – | – | – | – | – | – | – |
| Female | -0.04 (0.006) \*\*\* | 0.06 (0.006) \*\*\* | -0.07 (0.015) \*\*\* | -0.17 (0.019) \*\*\* | 0.26 (0.036) \*\*\* | -0.16 (0.028) \*\*\* | 0.11 (0.18) |
| Race |  |  |  |  |  |  |  |
| White (ref.) | – | – | – | – | – | – | – |
| Black | -0.06 (0.010) \*\*\* | -0.22 (0.010) \*\*\* | 0.05 (0.026) | 0.15 (0.034) \*\*\* | 0.26 (0.062) \*\*\* | 0.17 (0.048) \*\*\* | -0.42 (0.30) |
| Hispanic | -0.07 (0.016) \*\*\* | -0.18 (0.016) \*\*\* | 0.10 (0.042)  \* | 0.13 (0.056) \* | -0.20 (0.102) | -0.15 (0.081) | -0.26 (0.48) |
| Other | -0.04 (0.011) \*\*\* | -0.09 (0.011) \*\*\* | -0.13 (0.027)  \*\*\* | -0.12 (0.037) \*\* | -0.05 (0.067) | -0.05 (0.052) | 0.42 (0.35) |
| Dual-eligible  (ref. not dual-eligible) |  |  |  |  |  |  |  |
| Yes | 0.18 (0.008) \*\*\* | -0.11 (0.008) \*\*\* | -0.13 (0.020)  \*\*\* | -0.0004 (0.028) | 0.50 (0.048) \*\*\* | 0.43 (0.038) \*\*\* | 0.69 (0.24)  \*\* |
| Part D coverage  (ref. no Part D coverage) |  |  |  |  |  |  |  |
| Yes | 0.38 (0.007) \*\*\* | 0.11 (0.007) \*\*\* | 0.04 (0.018)  \* | -0.10 (0.023) \*\*\* | 0.13 (0.043)  \*\* | 0.16 (0.034)  \*\* | -- |
| ESRD  (ref. no ESRD) |  |  |  |  |  |  |  |
| Yes | 0.72 (0.02) \*\*\* | 0.60 (0.020) \*\*\* | 2.05 (0.052)  \*\*\* | 0.14 (0.069) \* | -0.13 (0.013) | -0.42 (0.099) \*\*\* | 1.35 (0.72)  \*\*\* |
| Originally enrolled in Medicare because of disability  (ref. not enrolled because of disability) |  |  |  |  |  |  |  |
| Yes | 0.11 (0.010) \*\*\* | 0.01  (0.010) | 0.06 (0.026)  \* | 0.03 (0.035) | 0.50 (0.063) \*\*\* | 0.54 (0.050) \*\*\* | 0.85 (0.37)  \* |
| Number of chronic conditions (%) |  |  |  |  |  |  |  |
| 0 (ref.) | – | – | – | – | – | – | – |
| 1-2 | 0.86 (0.018) \*\*\* | 1.02 (0.017) \*\*\* | 1.03 (0.046)  \*\*\* | 0.97 (0.061) \*\*\* | 0.51 (0.112) \*\*\* | 0.40 (0.088) \*\*\* | -0.65 (0.56) |
| 3-5 | 1.26 (0.017) \*\*\* | 1.48 (0.016) \*\*\* | 1.39 (0.043)  \*\*\* | 1.70 (0.059) \*\*\* | 1.03 (0.106) \*\*\* | 0.84 (0.082) \*\*\* | -0.16 (0.54) |
| 6-8 | 1.57 (0.017) \*\*\* | 1.82 (0.017) \*\*\* | 1.64 (0.044) \*\*\* | 2.29 (0.060) \*\*\* | 1.49 (0.107) \*\*\* | 1.17 (0.084) \*\*\* | 0.05 (0.55) |
| >=9 | 1.70 (0.018) \*\*\* | 2.02 (0.018) \*\*\* | 1.58 (0.046) \*\*\* | 2.44 (0.063) \*\*\* | 1.58 (0.111) \*\*\* | 1.23 (0.089) \*\*\* | -0.01 (0.58) |
| Frailty (ref. no frailty) |  |  |  |  |  |  |  |
| Yes | 0.47 (0.009) \*\*\* | 0.17 (0.009) \*\*\* | -0.08 (0.023)  \*\*\* | 1.13 (0.027) \*\*\* | 2.24 (0.047) \*\*\* | 0.67 (0.041) \*\*\* | 0.71 (0.30)  \* |
| Mental illness  (ref. no mental illness) |  |  |  |  |  |  |  |
| Yes | 0.15 (0.007) \*\*\* | 0.11 (0.007) \*\*\* | 0.07 (0.018) \*\*\* | 0.50 (0.024) \*\*\* | 0.52 (0.043) \*\*\* | -0.09 (0.034) \*\*\* | 0.33 (0.23) |
| Seriously ill  (ref. not seriously ill) |  |  |  |  |  |  |  |
| Yes | 0.15 (0.009) \*\*\* | 0.11 (0.010) \*\*\* | 0.10 (0.024) \*\*\* | 0.28 (0.030) \*\*\* | 0.73 (0.050) \*\*\* | 0.15 (0.044) \*\*\* | -0.22 (0.29) |
| HCC score, mean | 0.37 (0.003) \*\*\* | 0.24 (0.003) \*\*\* | 0.47 (0.008) \*\*\* | 0.66 (0.012) \*\*\* | 0.60 (0.200) \*\*\* | 0.45 (0.015) \*\*\* | 0.32 (0.099)  \*\* |
| Year |  |  |  |  |  |  |  |
| 2014 | 0.21 (0.006) \*\*\* | 0.12 (0.006) \*\*\* | 0.34 (0.014) \*\*\* | 0.12 (0.019) \*\*\* | 0.12 (0.036)  \*\* | 0.19 (0.027) \*\*\* | 0.37 (0.170)  \* |
| Number of observations | 186,858 | 186,858 | 186,858 | 186,858 | 186,858 | 186,858 | 186,858 |

*Notes: ADI: Area Deprivation Index; ESRD: end-stage renal disease; HCC: hierarchical condition categories. Results were estimated using generalized linear regression with a log link and gamma distribution. Coefficients of hospital referral region fixed effects are not presented in the table. The coefficients are the ratios of mean average costs associated with a category over the reference category. For example, in the “Total Medicare costs” column, the 0.03 of quintile 1 means that the average costs of quintile 1 is 3% ((exp(0.03)-1) \* 100) higher than the mean costs of quintile 3, holding other factors constant. \* P<0.05, \*\* P<0.01, \*\*\* P<0.001. Standard errors were reported in the parentheses.*

Appendix 9b Full regression results for preventable costs and preventable utilization

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Preventable costs | 30-day readmission\* | ACSC admission | Preventable ED visit |
|  | **Coefficient** | **OR** | **OR** | **OR** |
| ADI quintiles |  |  |  |  |
| Quintile 1 | -0.05 (0.039) | 0.95 (0.031) | 0.88 (0.031)  \*\*\* | 0.94 (0.022)  \*\* |
| Quintile 2 | 0.06 (0.044) | 1.01 (0.034) | 0.98 (0.038) | 0.99 (0.025) |
| Quintile 3 (ref.) | – | – | – | – |
| Quintile 4 | 0.05 (0.064) | 1.07 (0.053) | 1.01 (0.057) | 1.05 (0.038) |
| Quintile 5 | 0.11 (0.056) \* | 0.93 (0.037) | 1.12 (0.050)  \*\* | 1.15 (0.034) \*\*\* |
| Age categories |  |  |  |  |
| <65 (ref.) | – | – | – | – |
| 65-74 | -0.76 (0.059) \*\*\* | 0.72 (0.027) \*\*\* | 0.89 (0.044)  \* | 0.60 (0.018) \*\*\* |
| 75-84 | -0.72 (0.063) \*\*\* | 0.64 (0.028) \*\*\* | 1.10 (0.058) | 0.64 (0.021) \*\*\* |
| >=85 | -0.56 (0.069) \*\*\* | 0.63 (0.030) \*\*\* | 1.61 (0.090) \*\*\* | 0.67 (0.025) \*\*\* |
| Gender |  |  |  |  |
| Male (ref.) | – | – | – | – |
| Female | 0.06 (0.062)  \* | 0.92 (0.020) \*\*\* | 1.04 (0.026) | 1.13 (0.018) \*\*\* |
| Race |  |  |  |  |
| White (ref.) | – | – | – | – |
| Black | 0.47 (0.047) \*\*\* | 1.28 (0.040) \*\*\* | 1.59 (0.059) \*\*\* | 1.81 (0.042) \*\*\* |
| Hispanic | 0.32 (0.032) \*\*\* | 1.19 (0.059) \*\*\* | 1.43 (0.084) \*\*\* | 1.73 (0.061) \*\*\* |
| Other | -0.11 (0.049)  \* | 0.99 (0.045) | 1.07 (0.051) | 0.81 (0.026) \*\*\* |
| Dual-eligible  (ref. not Dual-eligible) |  |  |  |  |
| Yes | 0.32 (0.035) \*\*\* | 1.24 (0.033) \*\*\* | 1.23 (0.037) \*\*\* | 1.43 (0.028) \*\*\* |
| Part D coverage  (ref. no Part D coverage) |  |  |  |  |
| Yes | -0.04 (0.031) | 0.97 (0.031) | 0.93 (0.031)  \* | 1.00 (0.021) |
| ESRD  (ref. no ESRD) |  |  |  |  |
| Yes | 0.17 (0.093) | 1.13 (0.041)  \*\* | 0.82 (0.049)  \*\* | 0.91 (0.041)  \* |
| Originally enrolled in Medicare because of disability  (ref. not enrolled because of disability) |  |  |  |  |
| Yes | 0.19 (0.046) \*\*\* | 0.99 (0.031) | 1.24 (0.047) \*\*\* | 1.30 (0.033) \*\*\* |
| Number of chronic conditions (%) |  |  |  |  |
| 0 (ref.) | – | – | – | – |
| 1-2 | 0.71 (0.08)  \*\*\* | 1.01 (0.30) | 3.34 (1.32)  \*\* | 1.57 (0.104) \*\*\* |
| 3-5 | 1.30 (0.075) \*\*\* | 0.75 (0.22) | 8.26 (3.14)  \*\*\* | 1.88 (0.118) \*\*\* |
| 6-8 | 1.85 (0.077) \*\*\* | 0.76 (0.22) | 19.1 (7.26)  \*\*\* | 2.46 (0.154) \*\*\* |
| >=9 | 2.19 (0.083) \*\*\* | 0.85 (0.25) | 35.9 (13.63) \*\*\* | 3.06 (0.198) \*\*\* |
| Frailty (ref. no frailty) |  |  |  |  |
| Yes | 0.68 (0.037) \*\*\* | 0.98 (0.026) | 1.18 (0.038) \*\*\* | 1.19 (0.028) \*\*\* |
| Mental illness  (ref. no mental illness) |  |  |  |  |
| Yes | 0.48 (0.032) \*\*\* | 1.40 (0.032) \*\*\* | 1.18 (0.031) \*\*\* | 1.40 (0.025) \*\*\* |
| Seriously ill  (ref. not seriously ill) |  |  |  |  |
| Yes | 0.64 (0.040) \*\*\* | 1.93 (0.059) \*\*\* | 2.21 (0.075) \*\*\* | 1.21 (0.031) \*\*\* |
| HCC score, mean | 0.69 (0.015) \*\*\* | 1.23 (0.007) \*\*\* | 1.44 (0.011) \*\*\* | 1.10 (0.007) \*\*\* |
| Year |  |  |  |  |
| 2014 | 0.32 (0.026) \*\*\* | 1.05 (0.023)  \* | 1.07 (0.026)  \*\* | 1.09 (0.017) \*\*\* |
| Number of observations | 186,858 | 70,304 | 186,858 | 186,858 |

*Notes: ADI: Area Deprivation Index; ESRD: end-stage renal disease; HCC: hierarchical condition categories; OR: odds ratio. Results for preventable costs were estimated using generalized linear regression with a log link and gamma distribution. Coefficients of hospital referral region fixed effects are not presented in the table. Results for 30-day readmission, ACSC admission, preventable ED visit were estimated using logistic regression. Analysis for readmissions was at the admission level using hospitalized patients. \* P<0.05, \*\* P<0.01, \*\*\* P<0.001. Standard errors were reported in the parentheses.*