

*Trends in Mortality, Morbidity and  
Healthcare Utilization: Does it Make Sense to  
Use Healthcare as a Proxy for Health?*

*Brian Tysinger, Seth Seabury, Henu Zhao,  
Gwyn Pauley, Yifan Wei, Dana Goldman*

*Paper No: 2019-001*

**CESR-SCHAEFFER  
WORKING PAPER SERIES**

*The Working Papers in this series have not undergone peer review or been edited by USC. The series is intended to make results of CESR and Schaeffer Center research widely available, in preliminary form, to encourage discussion and input from the research community before publication in a formal, peer-reviewed journal. CESR-Schaeffer working papers can be cited without permission of the author so long as the source is clearly referred to as a CESR-Schaeffer working paper.*

# Trends in Mortality, Morbidity and Healthcare Utilization: Does it Make Sense to Use Healthcare as a Proxy for Health?

Bryan Tysinger<sup>1</sup>, Seth Seabury<sup>1</sup>, Henu Zhao<sup>1</sup>, Gwyn Pauley<sup>2</sup>, Yifan Wei<sup>1</sup>, and Dana Goldman<sup>1</sup>

<sup>1</sup> University of Southern California

<sup>2</sup> University of Wisconsin

## Abstract

Medical utilization, such as a count of hospitalizations, is routinely used as a health proxy for both policy and research purposes. Over time, trends in how medicine is practiced has impacted this relationship, as technological improvements or policy changes have altered hospitalizations at the same level of health. In this work, we document how hospitalizations have evolved as a marker for subsequent health (mortality and morbidity) and subsequent utilization (hospitalizations, emergency department visits, physician visits, and others). We assess these trends in the DI population (both fee-for-service and managed care), the aged Medicare fee-for-service population, and a commercial managed care population. We find that hospitalizations continue to predict subsequent mortality well, but that the relationship is more nuanced for morbidity. We find that hospitalizations are declining over time, but that they decline less for those with a prior hospitalization. Other types of utilization are increasing with time, with mixed evidence of the relationship with prior hospitalizations. We conclude that the relationship between hospitalizations and health continues to evolve and that the utility of this proxy depends on what measures of health one is assessing.

JEL: I100 I180

## Acknowledgments

Research reported in this publication was supported by the National Institute On Aging of the National Institutes of Health under Award Number P30AG024968. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The Centers for Medicare and Medicaid datasets on drug and medical claims and enrollment data for Medicare beneficiaries. A 20% random sample of Medicare beneficiaries' claims are available through the National Bureau of Economic Research, including Parts A, B, and D claims.

Optum claims data are from De-identified Clinformatics® Data Mart (OptumInsight, Eden Prairie, MN)

## Background and Introduction

Whether or not healthcare utilization predicts future health is an important question. Although it may seem obvious that those who are sicker are more likely to have a hospitalization, once trends in insurance, healthcare delivery, and economic factors are considered, this probability becomes less clear. The two outcomes have a complex relationship. For an initial illustration of the link between utilization and hospitalization, we used a nationally representative sample of Medicare enrollees to show age-adjusted trends over time for mortality, inpatient days, and a count of the number of chronic conditions

(Figure 1). A striking pattern is shown: while people in Medicare are becoming sicker over time, as measured by the number of chronic conditions for which they have been treated, they are also living longer and consuming less health care, as measured by the number of inpatient days. In fact, from 2002 to 2012, the average number of chronic conditions has increased by 14%, while the number of inpatient days has fallen by almost 25% and the mortality rate has experienced a general downward trend of about 4% between 2002 and 2012. As the likelihood of being hospitalized decreases, one might suspect that being hospitalized is a stronger measure of poor health than it was before, because it indicates a more serious health issue. However, we show that the relationship between previous hospitalizations and health has changed differently over time for different groups.

In this paper, we aim to answer three main questions related to the stark divergence in trends between chronic conditions and mortality/inpatient use described above.

1. How has health care utilization as a marker of disease severity changed over time?
2. What impact has this trend had on disease severity?
3. How well does utilization predict subsequent utilization?

Specifically, we are interested in the population of individuals who receive disability insurance (SSDI) and are enrolled in Medicare. This DI population is especially important because the process of qualifying for SSDI requires medical evidence, including hospitalizations. Better understanding the relationship between health care utilization and health motivates our study because it would clarify whether medical evidence standards requiring prior utilization remain good indicators of health status. We compare the SSDI population to two groups: those that are similar in age, 18-64, and those that have more complex healthcare needs, the aged (65+).

To approach this set of questions, we rely on two unique data sources. The first is a 20% sample of fee for service (FFS) Medicare claims. We include individuals who were enrolled in FFS Medicare plans—either because of disability or because of age. Second, we utilize claims data from Optum, which is a nationally available commercial insurance company. Optum provides private health insurance and Medicare Advantage (MA) plans. Similar to the FFS sample, we have detailed claim level data for this HMO population. Each data set is longitudinal, meaning that we are able to identify previous year health utilization, which is crucial to answering our questions. Previous work documenting trends in the SSDI population has noted the need for longitudinal data [1]. In summary, we have two samples: FFS and

Optum. Within each sample, we study four different populations: those who are/are not enrolled in their insurance plan through SSDI & those who did/did not have an inpatient visit in the previous year.

The dual data sets have several advantages. First, they allow us to study both the DI population enrolled in fee for service (FFS) and Medicare Advantage (MA). Rather than relying on a single year of data, the longitudinal structure allows us to account for the fact that there has been a substantial shift from FFS towards MA over the past fifteen years [2]. Second, it allows us to compare the DI population to individuals enrolled in commercial insurance who are of similar age (18-64) and to those who are enrolled in Medicare, who are on average sicker. Third, and most importantly, the longitudinal data of both the FFS claims and Optum allows us to study prior health care utilization. We use hospitalization in the previous year as a marker for health utilization.

Our analysis shows a complicated relationship between hospitalizations and health outcomes that has evolved over time differently depending on past healthcare utilization. Over time, patterns of utilization have not been stable across groups. In the following section, we show a simple model highlighting the challenges of using hospitalization as a proxy of health. We then give a more complete picture of the data used, present results, and discuss our findings.

## Conceptual Framework

Here we describe a simple conceptual framework that highlights how changes in utilization patterns could affect the relationship between observed hospitalizations (or other types of healthcare utilization) and unobserved health. To set notation, suppose  $h^*$  represents true health, which is unobservable. We simplify by assuming that there are two possible health states, high (H) or low (L). Formally,  $h^* \in [\theta_L, \theta_H]$  with  $\theta_L < \theta_H$ . The unconditional probability that a person is low health is given by  $\lambda$ .

Now suppose that  $y$  is an indicator variable representing whether a hospitalization occurred. Hospitalizations occur based on patient health and other characteristics as well as provider financial incentives, according to the relationship:

$$\Pr(y = 1) = \Pr(X\beta + \pi + \varepsilon > h^*)$$

where  $X$  represents other patient characteristics (e.g., demographics or insurance status),  $\pi$  represents the profitability of a hospitalization for the provider and  $\varepsilon$  is a random component with the distribution

function  $F(\cdot)$ . In this model, the probability of a hospitalization is increasing in provider financial incentives and decreasing in health.

While the unconditional probability that a person has low health is unknown, additional information about health can be inferred from the presence of a hospitalization.

When using  $y$  as a proxy for health, we are assuming that we can infer low health for a person based on the presence of a hospitalization. From Bayes' Theorem, the conditional probability that a person is hospitalized conditional on low health is equal to:

$$\Pr(y = 1|h^* = \theta_L) = \Pr(h^* = \theta_L|y = 1) \frac{\Pr(y = 1)}{\Pr(h^* = \theta_L)}$$

Using the fact that the unconditional probability of a hospitalization is the weighted average of the conditional probabilities with high or low health, this condition can be rearranged to:

$$\Pr(h^* = \theta_L|y = 1) = \lambda \left\{ \frac{\Pr(y = 1|h^* = \theta_L)}{\lambda \Pr(y = 1|h^* = \theta_L) + (1 - \lambda) \Pr(y = 1|h^* = \theta_H)} \right\}$$

Because lower health increases the probability of a hospitalization, the term in brackets on the right is greater than 1, so the conditional probability that a hospitalized person is low health is greater than the unconditional probability. Note that based on the assumptions above about the probability of a hospitalization, this expression can be further modified to:

$$\Pr(h^* = \theta_L|y = 1) = \lambda \left\{ \frac{1 - F(\theta_L - X\beta - \pi)}{(1 - F(\theta_H - X\beta - \pi)) - \lambda(F(\theta_L - X\beta - \pi) - F(\theta_H - X\beta - \pi))} \right\}$$

This illustrates how the impact of a hospitalization on the conditional probability that a person is low health depends on other factors in the model, including individual characteristics and provider financial incentives.

Now, suppose that growing public concern over healthcare costs leads to growing pressure to reduce the number of hospitalizations, which manifests as a decline in  $\pi$ . The change in the unconditional probability of a hospitalization with respect to a change in  $\pi$  is given by:

$$\lambda f(\theta_L - X\beta - \pi) + (1 - \lambda)f(\theta_H - X\beta - \pi) > 0$$

where  $f(\cdot)$  is the probability density function of  $\varepsilon$ . So, a decline in financial incentives unambiguously leads to a decline in the probability of a hospitalization, *ceteris paribus*.

A change in financial incentives for a hospitalization has no direct effect on the conditional probability that a person is low health (i.e.,  $\frac{d\lambda}{d\pi} = 0$ ). However, because hospitalizations are less likely with reduced financial incentives, this does change the information content of a hospitalization. In other words, a change in  $\pi$  will change the bracketed term in our expression for  $\Pr(h^* = \theta_L | y = 1)$  above. While the full expression is cumbersome, we can simplify with some additional notation by letting  $f(\theta_j)$  and  $F(\theta_j)$  represent the pdf and cdf, respectively, of a person with health type  $\theta_j$ . Using this simplified notation, we can express:

$$\frac{d\Pr(h^* = \theta_L | y = 1)}{d\pi} = \lambda \left\{ \frac{f(\theta_L)((1 - F(\theta_H)) - \lambda(F(\theta_L) - F(\theta_H))) - (1 - F(\theta_L))(\lambda f(\theta_L) + (1 - \lambda)f(\theta_H))}{((1 - F(\theta_H)) - \lambda(F(\theta_L) - F(\theta_H)))^2} \right\}$$

It can be shown that this term is negative as long as the share of sick people in the population is not too small,<sup>1</sup> which means that decreasing financial incentives for a hospitalization will make it more likely that a hospitalized person is sick.

This simple framework provides several insights into the challenge of using hospitalizations as a proxy for health. First, there is information provided in a hospitalization, so we do expect that someone who has a prior hospitalization will be less healthy on average. On the surface, we might expect that the recent trend towards fewer hospitalizations makes them more useful as a predictor of health, because only sicker patients will be hospitalized. Or, put another way, as the difference between healthy and sick patients grows – that is, as  $\theta_H - \theta_L$  grows – the marginal impact of financial incentives on admission decisions falls.

## Data

We briefly describe the data sources and measures used in the analysis. Data codebooks and programs are provided in the Appendix.

### Medicare Fee For Service

Medicare claims data is a data resource offered by the Centers for Medicare and Medicaid Services (CMS), documenting claims for services that Medicare pays for. In this analysis, we use the Master Beneficiary Summary File (MBSF) of Medicare claims data from 2002 to 2013, which provides Medicare

---

<sup>1</sup> The sufficient (but not necessary) condition for  $\frac{d\Pr(h^* = \theta_L | y = 1)}{d\pi} < 0$  is  $1 - \lambda < (1 + \lambda)F(\theta_H) + 2\lambda F(\theta_L)$ , which holds automatically as  $\lambda \rightarrow 1$ .

beneficiaries' enrollment and summary information. We only include the Medicare fee-for-service enrollees in our analysis. The sample size for regression analysis is 11 million observations. The MBSF Base file includes beneficiaries' demographic information (age, gender and death date) and plan enrollment information (enrollment for Part A/B/C/D). The MBSF Chronic Conditions Segment and Other Chronic or Potentially Disabling Conditions Segment provides each beneficiary's presence flag of a list of 27 chronic conditions and 35 chronic or potentially disabling conditions, we use these two files to derive our chronic conditions measure in this analysis. We also use the MBSF Cost and Utilization Segment to derive measures of utilization like inpatient service, emergency room services, prescription drug fills and Medicare payment.

### Optum Claims

To be complementary to Medicare claims, we also analyze health claims data for members of a large national managed care company affiliated with Optum (Optum claims). Optum claims data is provided by the Clinformatics<sup>®</sup> Data Mart (formerly called InVision Data Mart or "LabRx"). These claims have been verified, adjusted and de-identified before they were delivered to us. We use data from 2007 to 2014 and restrict the sample to include only enrollees aged between 18 and 64, with both medical and drug coverage and enrollment for the full year. The final sample includes 819,648 unique individuals covered by Medicare Advantage plans and 17,878,803 individuals covered by commercial health plans. By applying the age restriction, we believe the sampled individuals covered by Medicare Advantage plan are mainly on SSDI. The data tables used in the analysis include member eligibility data, medical claims, pharmacy claims, date of death file and inpatient confinements.

### Measures

In Medicare and Optum, we analyze the same measures, including age, gender, mortality, CCW count, and health care utilization measures on inpatient, emergency room use, physician visit and health care costs. We rely on the Medicare MBSF file for the measures, as described in the last section. In Optum, we derive the same set of flags for chronic conditions as in Medicare, according to the algorithm by the Chronic Condition Data Warehouse. We identify emergency room visits and physician visits from the type of service variable in the medical claims. Number of drugs fills is derived from the pharmacy claims, with a standard 30-day supply defined as one fill. Number of inpatient days is counted from the inpatient confinements. Total medical cost is derived from standardized costs, which is a value assigned by Optum by applying standard pricing algorithms. Given data availability and difference in data structure, measures in Optum are not exactly the same as in Medicare. The major difference in the two data sources include: Medicare includes days in other years when counting inpatient days, while we

exclude them in Optum. Medicare only counts physician office visits by applying a restriction using Berenson-Eggers Type of Service (BETOS) code, which is not available in Optum. Medicare provides actual costs while Optum only includes standardized costs.

Table 1 shows the number of observations in each year, and the mean and standard deviation of key measures, by group. In Medicare, compared with the aged group, the DI group is younger on average, and contains more females. The DI group has a lower mortality rate. However, they have a comparable number of chronic conditions with the aged, and they consume more medical care, in terms of inpatient days, emergency room visits and drug fills, resulting in higher medical expenditures. In Optum, the DI group is older than the commercial group. They have similar gender composition. The DI group has a higher mortality rate and more chronic conditions than the commercial group. They also use more medical care, reflected in all utilization measures.

Because of the differences in samples and measures we do not directly compare the DI group in Medicare and Optum data in any analysis.

## Hospitalization as a predictor of subsequent health

We now turn to estimating trends in hospitalization on subsequent health outcomes for the groups of interest in the two samples.

### Methods

The parsimonious model for subsequent health we explore is

$$Health_{t+1} = \alpha \cdot hosp_{i,t} + \gamma \cdot year_t + \delta \cdot hosp_{i,t} \cdot year_t + \mathbf{X}\boldsymbol{\beta} + \epsilon_{i,t}$$

The key parameter of interest is  $\delta$ , which indicates if hospitalizations in a particular year have a stronger or weaker impact on subsequent health relative to the base year. In this analysis, we consider two health outcomes: subsequent mortality and the count of chronic conditions in the next year, as measure of morbidity. We estimate the above regression separately for the SSDI population and the non-SSDI population in each sample.

### Results

For each data source, FFS and Optum, we present a table that shows the predicted level of each outcome for 2013 and either 2002 (FFS) or 2007 (Optum) (Tables 2 and 3). We also show the log difference between the two years and the difference in the log differences between those with and without a prior year hospitalization. The log difference shows the approximate percent change in each outcome over time. The difference in difference is the difference in changes between those with and



without a prior year hospitalization. For each outcome, we statistically test that this joint difference equals 0. We discuss the relative difference between the two groups below. Table 2 shows findings for the FFS sample and Table 3 shows findings for the Optum sample.

### Mortality

We present mortality results measured as deaths per 100,000 (Tables 1, 2, and 3). For both populations enrolled in Medicare FFS, disabled and aged, mortality was substantially larger for those with a hospitalization in the previous year. Across utilization, the aged group had higher rates of mortality than the disabled population.

Between 2002 and 2013, there was a downward trend in mortality for three of the four FFS groups. The exception is that SSDI individuals who were hospitalized in the past year had a small increase in mortality of about 0.9%. Mortality for those who did not have a prior year hospitalization decreased by 7.2%, which amounts to a difference of 8.1 percentage points across healthcare utilization for the disabled population. Similar to the disabled population, the decrease in mortality was larger for those with a hospitalization in the previous year than for those without a hospitalization for the aged population (13.6% decrease vs 2.4% decrease, an 11.2 percentage point difference).

When we consider individuals carrying insurance through Optum, we also observe a downward trend in mortality for three of the four groups. However, in the Optum sample, the DI population that did not have a hospitalization in the previous year experienced the increase in mortality (3.7%). The DI sample that did have a hospitalization experienced a decrease in mortality of 11.1%. This amounts to a statistically significant difference across utilization for the disabled group of almost 15 percentage points. For the commercially insured sample, both individuals with and without a hospitalization in the previous year saw large decreases in mortality, but the decrease was larger for those without a hospitalization (54.6% and 71.4% respectively). It should be noted however that mortality is a very rare occurrence for this sample. In 2013, it was just 51 per 100,000 for those with no hospitalization the previous year and 598 per 100,000 for those with a hospitalization.

### Morbidity

For both FFS and HMOs, as one would expect, those who were hospitalized in the previous year were treated for more chronic conditions than those who were not, regardless of their disability status. Within the FFS enrollees, conditional on prior healthcare utilization, the disabled and aged populations were quite similar in terms of chronic condition counts in both 2002 and 2013. For example, in 2013,

disabled individuals with a prior hospitalization had 7.2 chronic conditions, while their aged counterparts had 6.6. However, within Optum, the disabled population had a substantially higher chronic condition count than their working aged counterparts.

Individuals who were enrolled in FFS, regardless if they were hospitalized in the previous year or if they were enrolled through SSDI, experienced an increase of about 35% in the number of chronic conditions for which they were treated. The increase was slightly larger for those without a hospitalization, although the difference is statistically significant. Those who were enrolled in Medicare because of their age also experienced an increase in the number of chronic conditions. However, in contrast to the DI population, the aged individuals experienced an increase in the number of chronic conditions of 29.8% and 24.6% for those with and without a hospitalization respectively, a difference of 5.2 percentage points.

In line with the FFS population, between 2007 and 2013, all four groups of Optum enrollees experienced an increase in the number of chronic conditions they were treated for. The increase in the number of chronic conditions was largest for those without a previous hospitalization for both the DI population as well as the commercially insured population. For the DI population, the increase was about 44.8% for those who were not hospitalized compared to 24.8% for those who were hospitalized the previous year, which amounts to a statistically significant difference of 20 percentage points. The commercially insured population also experienced large increases in the number of chronic conditions- 30.5% for those who were not hospitalized the previous year and 18.9% for those with a previous hospitalization, a difference of 11.5 percentage points.

## Hospitalizations as a predictor of subsequent utilization

Now we turn to assessing the trend in hospitalizations as a predictor of subsequent utilization, examining subsequent hospitalizations, emergency department utilization, physician encounters, expenditures, and prescription drug fills.

### Methods

The parsimonious model for subsequent utilization we explore is

$$Utilization_{t+1} = \alpha \cdot hosp_{i,t} + \gamma \cdot year_t + \delta \cdot hosp_{i,t} \cdot year_t + \mathbf{X}\boldsymbol{\beta} + \epsilon_{i,t}$$

As in the previous model, the key parameter of interest is  $\delta$ , which indicates if hospitalizations in a particular year have a stronger or weaker impact on subsequent utilization relative to the base year. In this analysis, we consider a broad set of claims-based utilization measures, as described in the Data

section. Briefly, we assess inpatient hospitalizations (any, and how many), emergency department use (any and how many), physician visits, prescription drug fills, and an expenditure measure. This analysis is designed to help characterize the changing role of hospitalizations as a marker for health.

## Results

### Inpatient Utilization

Within the FFS sample, the probability of having any inpatient visit and the number of inpatient days fell for all four groups across time, consistent with our findings in *Figure 1*. Relatedly, hospitalization is persistent across time. Individuals who were hospitalized in year  $t-1$  were more likely to be hospitalized in year  $t$  for both the disabled and aged groups. Comparing the disabled to aged population in FFS, we see that the size and magnitude of differences between the two groups, in terms of both inpatient use and number of inpatient days, depends on prior year healthcare utilization. The aged population without previous year hospitalization was slightly more likely than their disabled counterparts to have an inpatient visit. However, for individuals with a previous year hospitalization, the opposite holds—disabled individuals were slightly more likely to have an inpatient visit. The number of inpatient visits was similar for the disabled and aged groups if they had a prior hospitalization, but if they did not, the disabled population had a higher number of inpatient visits.

For both measures, in both the DI and aged populations, the fall in inpatient visits was larger for individuals who did not have a prior year hospitalization. For the DI population, the fall in the probability of having an inpatient day was 4.9% and 14.2% or a 9.2 percentage point difference, for individuals with and without a hospitalization in the previous year respectively. For the number of inpatient days, the decrease was 7.2% and 19.9%, a 12.7 percentage point difference.

Turning to the aged population enrolled in FFS plans, we find that the decrease in the probability of an inpatient visit was 20.3% for those without a prior hospitalization and 9.2% for those with a prior hospitalization, for a difference of 11.1 percentage points. The decrease in the number of inpatient visits was 15.8% and 33.5% for those with/without a hospitalization in the previous year, a difference of 17.7 percentage points.

For individuals enrolled in Optum managed care plans we find that the disabled population was both more likely to have an inpatient visit and had a higher number of inpatient visits compared to their working-age counterparts, regardless of prior year hospitalizations. Similar to the FFS population, hospitalization is persistent. However, counter to what we see in the FFS population and *Figure 1*, we find that inpatient use and the number of inpatient days increased between 2007 and 2013 for the

disabled population, regardless of prior healthcare utilization. For both outcomes, the difference in increases was not statistically significant between those with and without a previous year hospitalization. For the working-aged commercially insured individuals, we find that the probability of having an inpatient visit decreased by 17.7% for those without a prior hospitalization and 3.9% for those with, a statistically significant difference of 13.7 percentage points. Commercially insured individuals with a hospitalization the previous year experienced an increase in the number of inpatient visits of 10.9%, while those without a prior hospitalization saw a decrease of 14.3%, a difference of 25.3 percentage points.

### Other Measurers of Utilization

For those enrolled in a FFS plan, all alternative healthcare utilization measures increased between 2002 and 2013 for both the disabled and aged population. This includes any ER visits, the number of ER visits, physician visits, drug fills, and costs.

Within the disabled FFS population there is not a clear pattern with respect to which group utilization increased for more, those with a prior hospitalization or those without. However, with the exception of costs, the differences in growth between those with and without a hospitalization the prior year are substantially smaller than the differences in inpatient utilization and days. For example, although the probability of an emergency department visit increased for those with and without a prior hospitalization by 13.9% and 17.3% respectively, a 3.4 percentage point difference, the number of emergency department visits increased by 29.8% and 26.6% respectively, a 3.2 percentage point difference. Physician visits increased by about 20% both for those with and without a prior hospitalization (21% and 18.6% respectively, a difference of 2.4 percentage points.) Costs increased by 51.1% for those without a prior hospitalization, but by 29.7% for those with, a difference of 21.4 percentage points.

Turning to the aged FFS population, a similar finding emerges: increases occurred between 2002 and 2013 for all measures but no clear patterns with regards to whether the increase was larger for those with or without a hospitalization in the previous year. Also similar to the disabled population, we find that the differences between the two groups were substantially smaller than the differences in inpatient care for the two groups, although for the aged population this holds for all alternative utilization outcomes. The probability of an emergency room visit increased by 16.7% for those without a hospitalization in the previous year and 18.2% for those with a prior hospitalization, a difference of 1.5 percentage points. During the same time, the number of emergency visits increased by 13.2% and 12.3%

for those with and without a hospitalization in the previous year. The only measure for which those without a hospitalization the previous year increased less than for those with a hospitalization for the aged group was drug fills, 7.1% vs 8.2% or a 1.1 percentage point difference. Contrary to the large increase in costs experienced by the SSDI population, costs increased by 13.5% for the aged without a hospitalization in the previous year and 16.7% for those with, a difference of 3.1 percentage points.

Like FFS, the disabled population enrolled in a HMO saw increases in all alternative measures of healthcare utilization. The largest increase was in the number of drug fills, 29.6% for those without a prior hospitalization and 17.6% for those with a prior hospitalization, a difference of 12 percentage points. The probability of an emergency department visit increased by 9.4% without hospitalization the previous year and 8.0% for those with, a 1.5 percentage point difference. The number of ED visits increased by 28.2% for those with a prior hospitalization and just 8.4% for those without a prior hospitalization, a difference of almost 20 percentage points. Costs increased by a relatively small amount compared to the FFS disabled population- just 5.9% for those without a prior hospitalization and 3.7% for those with, a difference of 2.2 percentage points.

Working aged adults enrolled in an HMO did not experience the universal increase in alternative healthcare utilization measures experienced by the disabled population. In fact, all measures of utilization fell for this population. The largest decrease was in emergency department encounters, which fell by 20.8% for those with a previous hospitalization and 16.4% for those without, a difference of 4.4 percentage points. The probability of an ED visit decreased by 8.3% for those with a hospitalization the prior year and 7.4% for those without, a statistically significant 0.9 percentage point difference. Costs also fell- by 23.8% for those without a prior hospitalization and 12.3% for those with, an 11.5 percentage point difference.

## Limitations

Administrative claims are an excellent resource for assessing health outcomes, but are not without limitations. The first set of limitations are related to enrollment. Over the last fifteen years there has been a marked shift from fee-for-service to managed care within the Medicare population, affecting both the aged and disabled populations. In 2004, 13% of Medicare enrollees were on Medicare Advantage plans, gradually increasing to 34% in 2018 [2]. The reasons for this migration are complex and beyond the scope of this paper, but all of the subsequent outcomes we assess (with the exception of mortality for those shifting from FFS to managed care) are estimated only on those who remained in either FFS or within the Optum plans.

Following individuals who seek to enroll in SSDI, qualify for SSDI, and then enroll in Medicare is challenging for a myriad of reasons. Since SSDI enrollment rules limit the amount of work an applicant can do, it is possible that individuals disenroll from their employer-based insurance plan. Once an individual qualifies for SSDI they typically are required to wait five months before receiving SSDI benefits and then an additional 24 months before qualifying for Medicare benefits. Thus, it can be up to 29 months before they enroll in a Medicare plan. Consequently, hospitalizations observed in the Medicare DI population are fairly distant from the time of qualifying for SSDI and should be interpreted accordingly.

We do not observe the qualifying reason in the Medicare DI populations we examine. As a result, we do not attempt to assess trends within particular qualifying conditions. With appropriate data linkage, this would be feasible, but was beyond the scope of this analysis.

## Discussion

Although trends in mortality, morbidity, and healthcare utilization are complex, a few key patterns emerge. First, when studying those enrolled in fee for service plans, the disabled and aged populations generally moved in the same direction. For example, the number of chronic conditions, probability of an ER visit, number of ER visits, physician visits, and drug fills increased for all four groups. Mortality generally fell across the board; the exception being that disabled individuals with a prior hospitalization did experience a slight increase. The probability of having any inpatient visit and the number of inpatient visits decreased for all four groups. This suggests some substitution between inpatient visits and alternative ways to consume healthcare. This substitution came at higher costs for everyone except for the commercially insured working aged individuals.

The second pattern that emerges in the FFS population is that, for most outcomes, the change across time was larger in magnitude for those without a hospitalization in the previous year than it was for those who did have a prior hospitalization. This pattern holds for mortality, morbidity, inpatient utilization, and drug fills. It is only for the probability of having an emergency room visit and the number of ER visits that this does not hold.

The patterns in trends over time for mortality, morbidity, and healthcare utilization are even murkier when studying individuals enrolled in MA plans. For most outcomes, the exception being morbidity, the disabled population moved in a different direction than their working-aged counterparts.

Taken together, the evidence suggests that trends driving hospitalizations are not uniform. Although people getting hospitalized are sicker, as evidenced by the general decrease in hospitalizations and increase in morbidity, the quality of hospitalizations as a marker of health did not uniformly change over time.

## Conclusions

This paper documents that hospitalization as a measure of health is complicated and trends in each outcome across time and utilization are not uniform. We find some evidence of substitution from hospitalizations towards other forms of healthcare utilization within the FFS population. We also show evidence that hospitalization is persistent. Individuals who were hospitalized in the prior year are more likely to be hospitalized in the current year. Further, changes in healthcare utilization were generally larger for those without prior year hospitalizations.

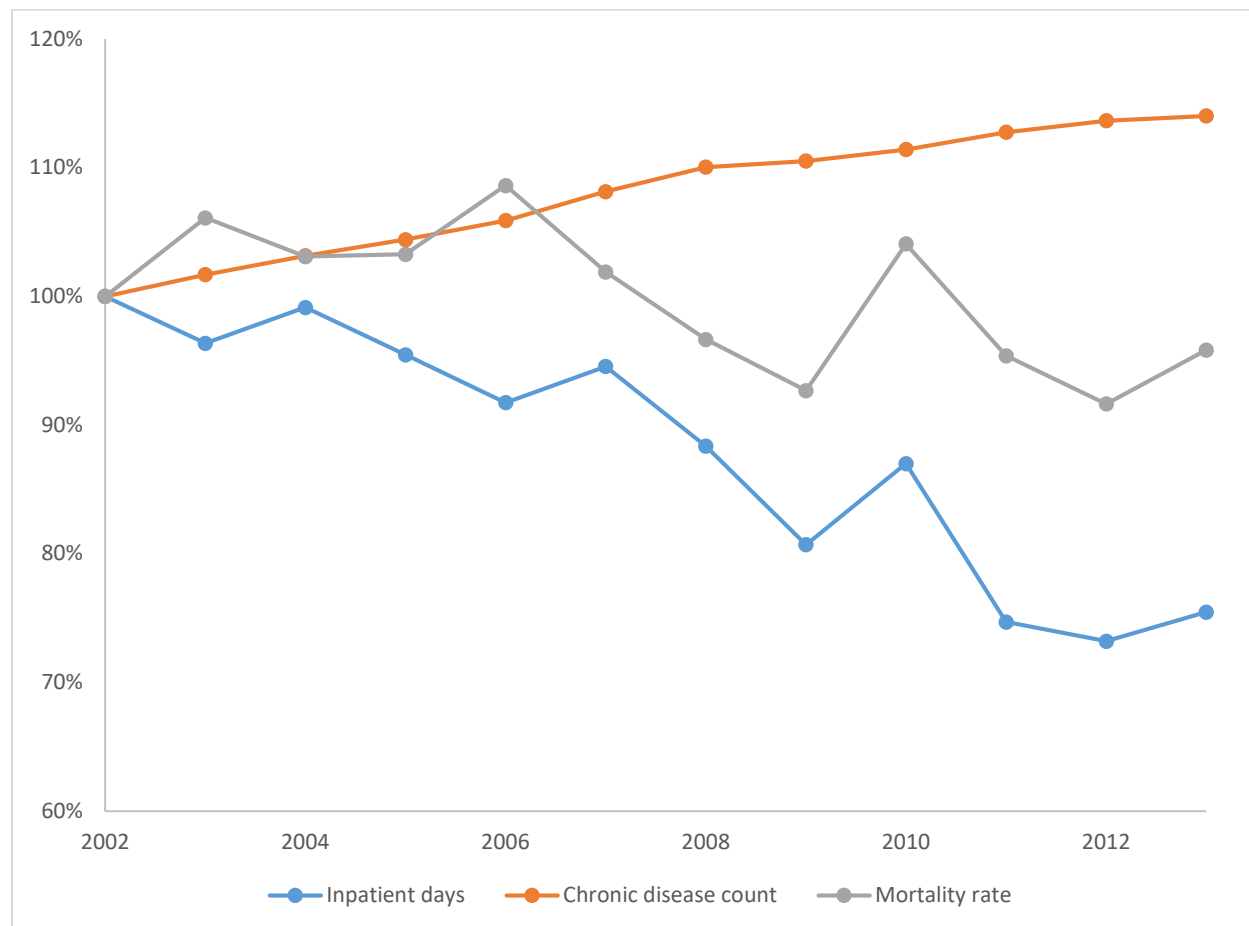
When considering SSDI eligibility, it is important to remember that part of the criteria is prior healthcare utilization. For example, consider eligibility due to a mental disorder, specifically neurocognitive disorders. One way to become eligible for SSDI requires medical documentation of a decline and a “medically documented history of the existence of the disorder over a period of at least 2 years”. Because of the general downward trend in inpatient utilization over time, one might worry that hospitalizations as a marker of health would only be flagging the sickest patients, meaning that the population of SSDI recipients is also becoming sicker over time. While our analysis does not refute this, it does suggest that hospitalizations as a marker of health is complicated. Further expanding listing criteria to require or include more healthcare consumption may be unnecessary or even counterintuitive.

Finally, it is worth noting that changing hospitalization patterns are likely second order to changing economic conditions when considering enrollment in SSDI. For example, there is evidence that poor local labor market conditions contribute to higher receipt of SSDI and SSI [3]. Prior literature has also documented the importance of women entering the workforce [4].

## Figures and Tables

### Figures

Figure 1: Normalized rates of inpatient days, mortality rates, and count of chronic conditions



Notes: Authors calculations using the MCBS 2002-2013. Yearly predictions from a model adjusted for age and race\*sex. Each year is shown as a relative change to 2002.



## Tables

Table 1: Summary Statistics

	Medicare		Optum	
	DI	Aged	DI	Commercial (18-64)
All years	11255795	58472804	2382218	49500523
2002	835159	5068386		
2003	876079	5140648		
2004	906702	5122297		
2005	893074	4963123		
2006	907672	4830337		
2007	920413	4737309	265019	6702298
2008	933220	4710375	255536	6582789
2009	956433	4700604	272786	6398815
2010	991874	4770619	293546	5982947
2011	1011032	4789259	311818	6080616
2012	1021169	4820102	329308	6069210
2013	1002968	4819745	341667	6054123
2014			312538	5629725
<b>Predictors</b>				
Age	50.51 (10.26)	76.12 (7.59)	56.78 (7.32)	41.43 (12.71)
Male	0.47 (0.50)	0.58 (0.49)	0.46 (0.50)	0.49 (0.50)
<b>Outcomes</b>				
Mortality rate	0.023 (0.15)	0.052 (0.22)	0.014 (0.12)	0.003 (0.06)
Chronic Conditions count	3.00 (2.90)	3.26 (2.56)	5.27 (3.62)	2.12 (2.35)
Any inpatient	0.20 (0.40)	0.19 (0.39)	0.13 (0.34)	0.05 (0.22)
Inpatient Days	11.83 (17.86)	8.66 (11.77)	1.22 (6.45)	0.24 (2.25)
Any Emergency department	0.34 (0.47)	0.21 (0.41)	0.27 (0.45)	0.18 (0.38)
ED visits	0.86 (2.56)	0.33 (0.91)	1.47 (4.68)	0.66 (2.50)
Physician events	6.34 (7.86)	7.34 (7.46)	15.89 (17.71)	8.11 (11.14)
Expenditures	11892.42 (25609.13)	8492.78 (17193.85)	15075.42 (39978.49)	5065.78 (17930.55)
Rx drug fills (30 day equivalents)	60.19 (48.68)	54.63 (39.46)	36.10 (38.61)	10.72 (17.97)

Table 2: Trends in Health and Healthcare Utilization for Individuals with FFS Medicare

Previous Year Hospitalization		FFS-DI					FFS-Aged				
		2002	2013	Log difference	Diff in diff	p-value	2002	2013	Log difference	Diff in diff	p-value
<b>Panel A: Health Outcomes</b>											
Mortality	No	1379	1283	-7.2			3180	2774	-13.6		
	Yes	5989	6041	0.9	8.1	0.00	11414	11140	-2.4	11.2	0.00
Morbidity	No	2.3	3.2	35.7			2.7	3.4	24.6		
	Yes	5.1	7.2	34.5	-1.1	0.00	4.9	6.5	29.8	5.2	0.00
<b>Panel B: Utilization Outcomes</b>											
Any inpatient	No	0.15	0.13	-14.2			0.18	0.15	-20.3		
	Yes	0.49	0.47	-4.9	9.2	0.00	0.42	0.39	-9.2	11.1	0.00
Inpatient days	No	1.42	1.16	-19.9			1.54	1.10	-33.5		
	Yes	8.37	7.79	-7.2	12.7	0.00	5.58	4.76	-15.8	17.7	0.00
Any emergency department	No	0.26	0.31	17.3			0.17	0.20	16.7		
	Yes	0.48	0.56	13.9	-3.4	0.00	0.30	0.37	18.2	1.5	0.00
Emergency encounters	No	0.50	0.65	26.6			0.24	0.30	22.7		
	Yes	1.42	1.91	29.8	3.2	0.00	0.53	0.70	27.9	5.3	0.00
Physician visits	No	5.15	6.35	21.0			6.29	7.11	12.3		
	Yes	8.63	10.40	18.6	-2.4	0.00	8.93	10.19	13.2	0.9	0.00
Drug fills	No	55.29	60.53	9.1			49.61	53.87	8.2		
	Yes	77.42	82.45	6.3	-2.8	0.00	68.29	73.31	7.1	-1.1	0.00
Costs	No	7508	12512	51.1			8380	9596	13.5		
	Yes	27175	36573	29.7	-21.4	0.00	20816	24590	16.7	3.1	0.00

Notes: Mortality is measured as deaths per 100,000. Morbidity is measured as the number of chronic conditions. The log difference is multiplied by 100.

Table 3: Trends in Health and Healthcare Utilization for Individuals Enrolled in Optum

Previous Year Hospitalization	HMO-DI					HMO-Commercial				
	2007	2013	Log difference	Diff in diff	p-value	2007	2013	Log difference	Diff in diff	p-value
<b>Panel A: Health Outcomes</b>										
Mortality	No	621	644	3.7			104	51	-71.4	
	Yes	3723	3333	-11.1	-14.8	0.00	1033	598	-54.6	16.8
Morbidity	No	3.9	6.1	44.8			1.7	2.4	30.5	
	Yes	7.6	9.8	24.8	-20.0	0.00	3.5	4.3	18.9	-11.5
<b>Panel B: Utilization Outcomes</b>										
Any inpatient	No	0.08	0.11	25.4			0.05	0.04	-17.7	
	Yes	0.32	0.38	16.5	-8.9	0.21	0.15	0.14	-3.9	13.7
Inpatient days	No	0.51	0.88	54.0			0.19	0.16	-14.3	
	Yes	3.63	5.58	43.0	-11.0	0.13	1.18	1.32	10.9	25.3
Any emergency department	No	0.23	0.25	9.4			0.18	0.17	-7.4	
	Yes	0.47	0.51	8.0	-1.5	0.00	0.30	0.28	-8.3	-0.9
Emergency encounters	No	1.05	1.14	8.4			0.68	0.57	-16.4	
	Yes	3.15	4.17	28.2	19.8	0.00	1.39	1.13	-20.8	-4.4
Physician visits	No	13.70	15.07	9.5			8.04	7.48	-7.3	
	Yes	26.98	27.44	1.7	-7.8	0.00	15.66	14.89	-5.0	2.2
Drug fills	No	30.61	41.14	29.6			11.49	11.17	-2.8	
	Yes	51.90	61.86	17.6	-12.0	0.00	22.19	21.42	-3.5	-0.7
Costs	No	12252	12999	5.9			5432	4281	-23.8	
	Yes	39240	40734	3.7	-2.2	0.00	17007	15040	-12.3	11.5

Notes: Mortality is measured as deaths per 100,000. Morbidity is measured as the number of chronic conditions. The log difference is multiplied by 100.

## Appendices

Appendix A – Comorbidities selection

Appendix B - Medicare FFS and Optum program descriptions

Appendix C1 – Medicare FFS codebook

Appendix C2 – Optum codebook

Appendix D1 – Medicare FFS regression results

Appendix D2 – Optum regression results

## Bibliography

1. National Academies of Sciences, E. and Medicine, *Health-care utilization as a proxy in disability determination*. 2018: National Academies Press.
2. Jacobson, G., A. Damico, and T. Neuman, *A dozen facts about Medicare Advantage*. 2018, Henry J. Kaiser Family Foundation San Francisco (CA).
3. Charles, K.K., Y. Li, and M. Stephens Jr, *Disability benefit take-up and local labor market conditions*. *Review of Economics and Statistics*, 2018. **100**(3): p. 416-423.
4. Liebman, J.B., *Understanding the increase in disability insurance benefit receipt in the United States*. *Journal of Economic Perspectives*, 2015. **29**(2): p. 123-50.

## Appendix A: Chronic Condition Identification and Selection

We rely on algorithms from the Chronic Conditions Data Warehouse to identify chronic conditions in administrative claims data. These algorithms search the administrative claims data for specific ICD-9 diagnosis codes, Medicare Severity Diagnosis Related Groups (MS-DRG) codes, or procedure codes. The algorithms were developed to facilitate identifying cohorts of individuals with particular conditions in claims data. We implemented these algorithms on the Optum data and used the pre-populated codes in the Medicare FFS data. The code for generating these condition codes is included in a separate appendix.

### Selected Conditions

We selected conditions from both the 27 CCW Chronic Conditions and the 39 Other Chronic or Potentially Disabling Conditions sets. Ultimately, we selected a set of 24 conditions from these 66, requiring that each condition have a prevalence over 5% for the 2014 Medicare FFS DI population and that a condition not be double-counted. This resulted in the set of conditions listed in the Table below. We then sum the total number of conditions for each individual, which can range from 0 to 24.

Chronic Condition	Prevalence (DI)	Prevalence (Aged)
Acquired hypothyroidism	10.53%	16.01%
Anemia	19.66%	23.51%
Anxiety disorder	25.46%	11.57%
Asthma	8.02%	4.63%
Bipolar disorder	12.94%	1.50%
Cataract	5.95%	20.45%
Chronic kidney disease	13.92%	18.86%
Chronic obstructive pulmonary disease	11.72%	11.45%
Diabetes	26.63%	28.54%
Epilepsy	7.26%	1.71%
Fibromyalgia chronic pain and fatigue	23.56%	10.23%
Heart failure	10.12%	16.03%

Hyperlipidemia	31.58%	49.45%
Hypertension	41.41%	60.35%
Intellectual disabilities and related conditions	5.44%	0.26%
Ischemic heart disease	17.35%	31.69%
Liver disease cirrhosis and other liver conditions (excluding Hepatitis)	5.61%	3.21%
Major depressive affective disorder	28.40%	14.02%
Migraine and other chronic headache	5.85%	1.39%
Obesity	18.65%	9.56%
Peripheral vascular disease	7.28%	13.54%
Rheumatoid arthritis/osteoarthritis	25.38%	33.14%
Schizophrenia and other psychotic disorders	10.61%	2.84%
Tobacco use disorders	22.36%	5.42%

## Appendix B: Program Descriptions

### *Code description*

This appendix describes SAS and Stata programs for Medicare and Optum claims data analysis. The programs are sent in a separate package.

### *Medicare programs:*

SAS programs for data cleaning:

- `clean_raw_data.sas`: this is the SAS code used to clean and select variables from the raw data
- `partd.sas`: this is the SAS code used to determine the beneficiaries' Part D enrollment status for each year
- `analysis_dataset.sas`: this is the SAS code used to generate analysis dataset, output in Excel format. This needs to run after `clean_raw_data.sas` and `partd.sas`

Stata programs for analysis, which need to run after the SAS programs:

- `analysis_measure.do`: this is the Stata code used to read in the Excel format analysis dataset into Stata format and also generate some measures used in analysis. The output is in Stata (.dta) format and ready for analysis
- `model_di.do`: this is the regression analysis Stata code for DI population
- `model_aged.do`: this is the regression analysis Stata code for AGED population

### *Optum programs:*

`master.sas` is a master program to call all the other macros. This is the only program that needs to run.

Macros are recorded in the following SAS programs:

- `select_vars.sas`: this program contains macro to select variables needed to develop disease measures using CCW algorithm
- `ccw.sas`: this program contains macros to apply CCW algorithm for multiple chronic conditions
- `enrollment.sas`: this program contains macros to process Optum enrollment file and identify individuals who enrolled for the entire year
- `cost_and_use.sas`: this program contains macros to process Optum inpatient files, counting yearly number of inpatient stays, inpatient days and cost
- `er_visit.sas`: this program contains macros to process Optum medical files, counting yearly number of emergency room stays and cost
- `physician_visit.sas`: this program contains macros to process OPTUM medical files, counting yearly number of physician visits
- `drug_cost.sas`: this program contains macros to process Optum prescription drug files, counting yearly drug cost
- `drug_fill.sas`: this program contains macros to process Optum prescription drug files, counting yearly drug fills and standardize the fills to 30-day supply
- `dod.sas`: this program processes Optum date of death file, assigning flags for the deceased
- `join.sas`: this program joins different datasets together and creates a summary dataset

Stata code to run the models, which needs to run after the SAS programs:

- `model_di.do`: this is the regression analysis Stata code for DI population
- `model_commercial.do`: this is the regression analysis Stata code for commercial population



## Appendix C1: Medicare Codebook

-----  
bene\_id (unlabeled)  
-----

type: string (str15)  
unique values: 12,262,531 missing "": 0/82,627,412  
examples: "mmmmmmDssaaaUm"  
"mmmmmmUGXWsJXJ"  
"mmmmmmWWaDmmfD"  
"mmmmmmWDaJaGG"

-----  
age65p (unlabeled)  
-----

type: numeric (byte)  
range: [0,1] units: 1  
unique values: 2 missing .: 0/82,627,412  
tabulation: Freq. Value  
13,331,539 0  
69,295,873 1

-----  
age (unlabeled)  
-----

type: numeric (int)  
range: [0,199] units: 1  
unique values: 165 missing .: 26/82,627,412  
mean: 72.2117  
std. dev: 12.5536  
percentiles: 10% 25% 50% 75% 90%  
56 67 73 80 87

-----  
male (unlabeled)  
-----

type: numeric (byte)  
range: [0,1] units: 1  
unique values: 2 missing .: 29/82,627,412  
tabulation: Freq. Value  
46,303,472 0  
36,323,911 1  
29 .

-----  
race\_bg (unlabeled)  
-----

type: numeric (byte)  
range: [1,6] units: 1  
unique values: 6 missing .: 188,951/82,627,41  
> 2  
tabulation: Freq. Value  
67,244,652 1

7,759,394 2  
974,783 3  
1,546,165 4  
4,554,893 5  
358,574 6  
188,951 .

-----  
dead\_thisyear (unlabeled)  
-----

type: numeric (byte)  
range: [0,1] units: 1  
unique values: 2 missing .: 0/82,627,412  
tabulation: Freq. Value  
78,684,144 0  
3,943,268 1

-----  
dead\_ny (unlabeled)  
-----

type: numeric (byte)  
range: [0,1] units: 1  
unique values: 2 missing .: 0/82,627,412  
tabulation: Freq. Value  
79,105,516 0  
3,521,896 1

-----  
enroll (unlabeled)  
-----

type: numeric (byte)  
range: [1,1] units: 1  
unique values: 1 missing .: 0/82,627,412  
tabulation: Freq. Value  
82,627,412 1

-----  
allcc\_count (unlabeled)  
-----

type: numeric (byte)  
range: [0,23] units: 1  
unique values: 24 missing .: 0/82,627,412  
mean: 3.32205  
std. dev: 2.69842  
percentiles: 10% 25% 50% 75% 90%  
0 1 3 5 7

-----  
any\_ip (unlabeled)  
-----

type: numeric (byte)  
range: [0,1] units: 1  
unique values: 2 missing .: 0/82,627,412  
tabulation: Freq. Value  
65,198,482 0

17,428,930 1

-----  
year (unlabeled)  
-----

type: numeric (int)  
range: [2002,2014] units: 1  
unique values: 13 missing .: 0/82,627,412  
mean: 2008  
std. dev: 3.75648  
percentiles: 10% 25% 50% 75% 90%  
2003 2005 2008 2011 2013

-----  
ip\_days (unlabeled)  
-----

type: numeric (int)  
range: [0,1494] units: 1  
unique values: 399 missing .: 0/82,627,412  
mean: 2.15223  
std. dev: 7.75332  
percentiles: 10% 25% 50% 75% 90%  
0 0 0 0 6

-----  
any\_er (unlabeled)  
-----

type: numeric (byte)  
range: [0,1] units: 1  
unique values: 2 missing .: 0/82,627,412  
tabulation: Freq. Value  
62,549,206 0  
20,078,206 1

-----  
er\_visits (unlabeled)  
-----

type: numeric (int)  
range: [0,422] units: 1  
unique values: 234 missing .: 0/82,627,412  
mean: .447085  
std. dev: 1.3877  
percentiles: 10% 25% 50% 75% 90%  
0 0 0 0 1

-----  
phys\_events (unlabeled)  
-----

type: numeric (int)  
range: [0,400] units: 1  
unique values: 299 missing .: 0/82,627,412  
mean: 7.03  
std. dev: 7.53369

```

      percentiles:      10%      25%      50%      75%      90%
                       0         2         5         10        16
-----
mocr_pmt (unlabeled)
-----
      type: numeric (float)
      range: [0,16756456]          units: .01
unique values: 7,285,615          missing .: 5,122,596/82,627,
> 412

      mean: 10267.4
      std. dev: 21149.4

      percentiles:      10%      25%      50%      75%      90%
                       267.66   909.65  2943.87  9739.68  28495.7
-----
mocr_pmt_ad (unlabeled)
-----
      type: numeric (float)
      range: [0,19932866]          units: 1.000e-10
unique values: 21,763,345          missing .: 5,122,596/82,627,
> 412

      mean: 12573.1
      std. dev: 25628.8

      percentiles:      10%      25%      50%      75%      90%
                       344.228  1154.11 3621.61  11943.3  35130.2
-----
pdt_mocr_pmt (unlabeled)
-----
      type: numeric (float)
      range: [0,5695328.5]          units: .01
unique values: 1,992,645          missing .: 53,082,183/82,627
> ,412

      mean: 2653.19
      std. dev: 6339.71

      percentiles:      10%      25%      50%      75%      90%
                       25.13   344.75  1280.38  2315.47  6227.61
-----
pdt_fill_cnt (unlabeled)
-----
      type: numeric (int)
      range: [0,1223]              units: 1
unique values: 799                missing .: 53,082,183/82,627
> ,412

      mean: 55.2928
      std. dev: 41.8256

      percentiles:      10%      25%      50%      75%      90%
                       11         25         47         76        109

```

## Appendix C2: Optum Codebook

-----  
-----  
-----

```
name: <unnamed>
log: /schaeffer-a/sch-projects/dua-data-projects/OPTUM/SSA/pgm/codebook_optum.log
log type: text
opened on: 14 May 2019, 15:16:47
```

```
.
. * DI
. use ../data/analytic0408, clear

. codebook age male dead ccw_count any_ip ip_days any_er n_er n_car fills total_cost
```

-----  
-----  
-----

age  
(unlabeled)

-----  
-----  
-----

```
type: numeric (byte)

range: [18,64]          units: 1
unique values: 47      missing .: 0/2,412,358

mean: 56.7794
std. dev: 7.31446

percentiles:          10%      25%      50%      75%      90%
                   47        54        59        62        64
```

-----  
-----  
-----

male  
(unlabeled)

-----  
-----  
-----

type: numeric (byte)

range: [0,1] units: 1  
unique values: 2 missing .: 0/2,412,358

tabulation:	Freq.	Value
	1,299,222	0
	1,113,136	1

-----  
-----  
-----

dead  
(unlabeled)

-----  
-----  
-----

type: numeric (byte)

range: [0,1] units: 1  
unique values: 2 missing .: 0/2,412,358

tabulation:	Freq.	Value
	2,378,853	0
	33,505	1

-----  
-----  
-----

ccw\_count  
(unlabeled)

-----  
-----  
-----

type: numeric (byte)

range: [0,23] units: 1  
unique values: 24 missing .: 874,626/2,412,358

mean: 5.27249  
std. dev: 3.6217

percentiles:	10%	25%	50%	75%	90%
	1	3	5	8	10

-----  
-----  
-----  
any\_ip  
(unlabeled)  
-----  
-----  
-----

type: numeric (byte)

range: [0,1] units: 1  
unique values: 2 missing .: 0/2,412,358

tabulation:	Freq.	Value
	2,079,229	0
	333,129	1

-----  
-----  
-----  
ip\_days  
(unlabeled)  
-----  
-----  
-----

type: numeric (int)

range: [0,365] units: 1  
unique values: 280 missing .: 0/2,412,358

mean: 1.30544  
std. dev: 6.72091

percentiles:	10%	25%	50%	75%	90%
	0	0	0	0	2

-----  
-----  
-----  
  
any\_er  
(unlabeled)  
  
-----  
-----  
-----

type: numeric (byte)

range: [0,1] units: 1  
unique values: 2 missing .: 0/2,412,358

tabulation:	Freq.	Value
	1,745,233	0
	667,125	1

-----  
-----  
-----  
  
n\_er  
(unlabeled)  
  
-----  
-----  
-----

type: numeric (int)

range: [0,508] units: 1  
unique values: 234 missing .: 0/2,412,358

mean: 1.49711  
std. dev: 4.71532



percentiles:	10%	25%	50%	75%	90%
	0	0	0	1	4

n\_car  
(unlabeled)

type: numeric (int)

range:	[0,1400]	units:	1
unique values:	360	missing .:	0/2,412,358

mean:	15.8834
std. dev:	17.834

percentiles:	10%	25%	50%	75%	90%
	0	4	11	22	36

fills  
(unlabeled)

type: numeric (double)

range:	[-2.2666667,670.2]	units:	1.000e-10
unique values:	53,207	missing .:	0/2,412,358

mean:	35.8881
std. dev:	38.5487

percentiles:	10%	25%	50%	75%	90%
--------------	-----	-----	-----	-----	-----

0 4 25.8333 54.0333 86.3667

-----  
-----  
-----  
total\_cost  
(unlabeled)  
-----  
-----  
-----

type: numeric (double)

range: [-3066459.3,16811998] units: .00001

unique values: 2,132,329 missing .: 0/2,412,358

mean: 15468.1

std. dev: 41330.5

percentiles: 10% 25% 50% 75% 90%  
194.88 1423.27 4933.91 14127.3 36585.6

.

. \* commercial

. use ../data/analytic0408\_comm, clear

. codebook age male dead ccw\_count any\_ip ip\_days any\_er n\_er n\_car fills total\_cost

-----  
-----  
-----  
age  
(unlabeled)  
-----  
-----  
-----

type: numeric (byte)

range: [18,64] units: 1

unique values: 47 missing .: 0/49,652,115

mean: 41.4565  
std. dev: 12.7123

percentiles:      10%      25%      50%      75%      90%  
                  23      31      42      52      58

-----  
-----  
-----

male  
(unlabeled)

-----  
-----  
-----

type: numeric (byte)

range: [0,1]                      units: 1  
unique values: 2                    missing .: 6,478/49,652,115

tabulation: Freq. Value  
              25,217,669 0  
              24,427,968 1  
               6,478 .

-----  
-----  
-----

dead  
(unlabeled)

-----  
-----  
-----

type: numeric (byte)

range: [0,1]                      units: 1  
unique values: 2                    missing .: 0/49,652,115

tabulation: Freq. Value

49,487,638 0

164,477 1

-----  
-----  
-----  
ccw\_count  
(unlabeled)  
-----  
-----  
-----

type: numeric (byte)

range: [0,23]

units: 1

unique values: 24

missing .: 18,264,256/49,652,115

mean: 2.12277

std. dev: 2.34559

percentiles:	10%	25%	50%	75%	90%
	0	0	1	3	5

-----  
-----  
-----  
any\_ip  
(unlabeled)  
-----  
-----  
-----

type: numeric (byte)

range: [0,1]

units: 1

unique values: 2

missing .: 0/49,652,115

tabulation: Freq. Value

47,085,418 0

2,566,697 1

-----  
-----  
-----  
ip\_days  
(unlabeled)  
-----  
-----  
-----

type: numeric (int)

range: [0,365]

units: 1

unique values: 341

missing .: 0/49,652,115

mean: .251678

std. dev: 2.4238

percentiles:	10%	25%	50%	75%	90%
	0	0	0	0	0

-----  
-----  
-----  
any\_er  
(unlabeled)  
-----  
-----  
-----

type: numeric (byte)

range: [0,1]

units: 1

unique values: 2

missing .: 0/49,652,115

tabulation:	Freq.	Value
	40,743,805	0
	8,908,310	1

n\_er  
(unlabeled)

-----  
-----  
-----

type: numeric (int)

range: [0,472]

units: 1

unique values: 237

missing .: 0/49,652,115

mean: .66311

std. dev: 2.50255

percentiles:	10%	25%	50%	75%	90%
	0	0	0	0	2

-----  
-----  
-----

n\_car  
(unlabeled)

-----  
-----  
-----

type: numeric (int)

range: [0,1432]

units: 1

unique values: 484

missing .: 0/49,652,115

mean: 8.1107

std. dev: 11.1907

percentiles:	10%	25%	50%	75%	90%
	0	1	5	11	20

-----  
-----  
-----

fills  
(unlabeled)

-----  
-----  
-----

type: numeric (double)

range: [-5,531.33333]                   units: 1.000e-10  
unique values: 56,466                   missing .: 0/49,652,115

mean: 10.7097

std. dev: 17.9685

percentiles:	10%	25%	50%	75%	90%
	0	0	2.53333	13.6333	31.8667

-----  
-----  
-----

total\_cost  
(unlabeled)

-----  
-----  
-----

type: numeric (double)

range: [-1175348.7,23172305]           units: 1.000e-21  
unique values: 31,474,275               missing .: 0/49,652,115

mean: 5152.61

std. dev: 18673.6

percentiles:	10%	25%	50%	75%	90%
	0	201.025	1128.64	4087.08	11822.7

.  
. log close

name: <unnamed>

log: /schaeffer-a/sch-projects/dua-data-projects/OPTUM/SSA/pgm/codebook\_optum.log

log type: text

closed on: 14 May 2019, 15:45:03

-----  
-----  
-----



## Appendix D1: Regression tables – Medicare FFS

DI

```
. * next year mortality - DI
. logit dead_ny age male any_ip##i.yr_after_2001 if (enroll_ny==1|dead_ny==1),
> nolog cluster(bene_id2)
```

```
Logistic regression                Number of obs    = 11,255,795
                                   Wald chi2(25)     = 209169.83
                                   Prob > chi2         = 0.0000
Log pseudolikelihood = -1191465.4   Pseudo R2       = 0.0779
```

(Std. Err. adjusted for 2,130,226 clusters in bene\_id2)

dead_ny	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
age	.0470634	.000232	202.89	0.000	.0466087	.047518
male	.3639758	.0039234	92.77	0.000	.3562861	.3716656
1.any_ip	1.502641	.0140515	106.94	0.000	1.4751	1.530181
yr_afte~2001						
2	-.0240976	.0141696	-1.70	0.089	-.0518695	.0036743
3	-.0205777	.0140359	-1.47	0.143	-.0480875	.0069321
4	-.0568303	.0141926	-4.00	0.000	-.0846472	-.0290134
5	-.0559351	.0141131	-3.96	0.000	-.0835963	-.0282739
6	-.0536991	.0140363	-3.83	0.000	-.0812099	-.0261884
7	-.0850694	.0140846	-6.04	0.000	-.1126747	-.0574641
8	-.1290985	.0141306	-9.14	0.000	-.1567939	-.1014031
9	-.1030759	.0138925	-7.42	0.000	-.1303047	-.075847
10	-.116475	.0138565	-8.41	0.000	-.1436331	-.0893168
11	-.088694	.0137246	-6.46	0.000	-.1155937	-.0617943
12	-.073249	.0136908	-5.35	0.000	-.1000826	-.0464155
any_ip#						
yr_afte~2001						
1 2	-.0237954	.019773	-1.20	0.229	-.0625497	.0149589
1 3	-.0104036	.0195402	-0.53	0.594	-.0487018	.0278946
1 4	.0174952	.0196759	0.89	0.374	-.0210689	.0560593
1 5	.0268592	.0196048	1.37	0.171	-.0115654	.0652839
1 6	-.0044873	.0196094	-0.23	0.819	-.042921	.0339465
1 7	.0361133	.0196	1.84	0.065	-.0023019	.0745285
1 8	.0638301	.0196013	3.26	0.001	.0254122	.1022479
1 9	.0319264	.0193674	1.65	0.099	-.006033	.0698858
1 10	.0675876	.0192873	3.50	0.000	.0297852	.10539
1 11	.0453594	.0192045	2.36	0.018	.0077192	.0829997
1 12	.0825965	.0191926	4.30	0.000	.0449797	.1202132
_cons	-6.838013	.0162581	-420.59	0.000	-6.869879	-6.806148

```
. * next year CCW count - DI
. reg ccw_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, cluster(bene_id2)
> )
```

```
Linear regression                Number of obs    = 11255795
                                   F(25, 2130225)   = 54047.66
                                   Prob > F           = 0.0000
                                   R-squared            = 0.2602
                                   Root MSE         = 2.6428
```

(Std. Err. adjusted for 2,130,226 clusters in bene\_id2)

ccw_ny	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0531197	.0001402	378.88	0.000	.0528449	.0533945
male	-.7576005	.0034073	-222.35	0.000	-.7642787	-.7509223
1.any_ip	2.73309	.0072745	375.71	0.000	2.718833	2.747348

yr_afte~2001						
2	.1006878	.0026905	37.42	0.000	.0954145	.1059611
3	.1844959	.0029169	63.25	0.000	.1787789	.190213
4	.2553979	.0031389	81.36	0.000	.2492457	.2615501
5	.3399516	.003283	103.55	0.000	.3335171	.3463862
6	.4230878	.0034087	124.12	0.000	.4164068	.4297688
7	.5394511	.0035202	153.24	0.000	.5325515	.5463506
8	.6233156	.0036012	173.09	0.000	.6162574	.6303738
9	.7418544	.0037037	200.30	0.000	.7345952	.7491135
10	.8184053	.0037563	217.87	0.000	.8110431	.8257676
11	.9104064	.00382	238.33	0.000	.9029193	.9178934
12	.9731022	.0038994	249.55	0.000	.9654595	.980745
any_ip#						
yr_afte~2001						
1 2	.0823959	.0085415	9.65	0.000	.0656549	.0991369
1 3	.1518033	.0091909	16.52	0.000	.1337895	.1698171
1 4	.2184017	.0095587	22.85	0.000	.199667	.2371364
1 5	.3221573	.0098515	32.70	0.000	.3028488	.3414659
1 6	.4580183	.0100613	45.52	0.000	.4382984	.4777381
1 7	.5636266	.0102445	55.02	0.000	.5435477	.5837055
1 8	.6391949	.010339	61.82	0.000	.6189309	.6594589
1 9	.8315281	.0106218	78.28	0.000	.8107097	.8523465
1 10	1.043296	.0107907	96.68	0.000	1.022147	1.064446
1 11	1.073233	.010904	98.43	0.000	1.051861	1.094604
1 12	1.119957	.0111356	100.57	0.000	1.098132	1.141782
_cons	.0059621	.0075351	0.79	0.429	-.0088064	.0207306

```
. * next year inpatient use - DI
. logit anyip_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, nolog cluste
> r(bene_id2)
```

```
Logistic regression          Number of obs   = 11,255,795
                             Wald chi2(25)      = 690216.54
                             Prob > chi2        = 0.0000
Log pseudolikelihood = -5273733.5   Pseudo R2      = 0.0951
```

(Std. Err. adjusted for 2,130,226 clusters in bene\_id2)

anyip_ny	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
age	.0089253	.0000935	95.48	0.000	.0087421	.0091086
male	-.116415	.0018879	-61.67	0.000	-.1201151	-.1127149
1.any_ip	1.678461	.0059017	284.40	0.000	1.666894	1.690028
yr_afte~2001						
2	.0026326	.0047484	0.55	0.579	-.006674	.0119393
3	.0057168	.0045713	1.25	0.211	-.0032427	.0146764
4	-.0107383	.0046422	-2.31	0.021	-.0198368	-.0016398
5	-.0273195	.004653	-5.87	0.000	-.0364391	-.0181998
6	-.04258	.0046659	-9.13	0.000	-.051725	-.0334349
7	-.0455512	.0046719	-9.75	0.000	-.0547079	-.0363945
8	-.0712526	.0046729	-15.25	0.000	-.0804112	-.0620939
9	-.0881268	.0046555	-18.93	0.000	-.0972514	-.0790021
10	-.1133656	.00466	-24.33	0.000	-.122499	-.1042323
11	-.1495059	.004677	-31.97	0.000	-.1586727	-.1403391
12	-.165037	.0047058	-35.07	0.000	-.1742601	-.1558139
any_ip#						
yr_afte~2001						
1 2	-.003789	.0074157	-0.51	0.609	-.0183235	.0107455
1 3	-.0074395	.0079034	-0.94	0.347	-.0229299	.0080509
1 4	-.0102704	.0080127	-1.28	0.200	-.0259749	.0054341
1 5	-.0028298	.0080552	-0.35	0.725	-.0186177	.0129582
1 6	.0063007	.0080822	0.78	0.436	-.0095402	.0221416
1 7	.0096234	.0080902	1.19	0.234	-.0062332	.02548
1 8	.0195851	.0080706	2.43	0.015	.003767	.0354033
1 9	.0305413	.0080292	3.80	0.000	.0148044	.0462782

1 10		.0488736	.0080322	6.08	0.000	.0331307	.0646165
1 11		.0551366	.0080654	6.84	0.000	.0393288	.0709445
1 12		.0703297	.0081594	8.62	0.000	.0543375	.0863219
_cons		-2.110074	.0059025	-357.49	0.000	-2.121643	-2.098505

```
. * next year inpatient days - DI
. nbreg ipdays_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, nolog clust
> er(bene_id2)
```

```
Negative binomial regression      Number of obs      = 11,255,795
                                Wald chi2(25)       = 490932.99
Dispersion = mean                Prob > chi2        = 0.0000
Log pseudolikelihood = -13367806 Pseudo R2          = 0.0184
```

(Std. Err. adjusted for 2,130,226 clusters in bene\_id2)

ipdays_ny		Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
age		.0061202	.0001562	39.18	0.000	.005814 .0064264	
male		-.0023178	.0031143	-0.74	0.457	-.0084217 .0037862	
1.any_ip		1.770452	.0075849	233.42	0.000	1.755586 1.785318	
yr_afte~2001							
2		.0056507	.00795	0.71	0.477	-.0099311 .0212324	
3		-.0269469	.0077256	-3.49	0.000	-.0420888 -.0118051	
4		-.031888	.0078914	-4.04	0.000	-.0473549 -.0164211	
5		-.0528706	.0079616	-6.64	0.000	-.068475 -.0372662	
6		-.0726768	.0079196	-9.18	0.000	-.088199 -.0571546	
7		-.1019658	.0079172	-12.88	0.000	-.1174833 -.0864483	
8		-.1325195	.0079038	-16.77	0.000	-.1480107 -.1170283	
9		-.1336368	.0078574	-17.01	0.000	-.1490371 -.1182365	
10		-.1752447	.0078553	-22.31	0.000	-.1906409 -.1598485	
11		-.1937909	.0078921	-24.56	0.000	-.2092591 -.1783227	
12		-.1994271	.0079423	-25.11	0.000	-.2149938 -.1838604	
any_ip#							
yr_afte~2001							
1 2		-.0025624	.0101148	-0.25	0.800	-.022387 .0172623	
1 3		.0212998	.0102503	2.08	0.038	.0012097 .04139	
1 4		.0128696	.0105077	1.22	0.221	-.0077252 .0334644	
1 5		.035048	.0106141	3.30	0.001	.0142448 .0558513	
1 6		.0383834	.0106068	3.62	0.000	.0175945 .0591723	
1 7		.0606545	.0106155	5.71	0.000	.0398485 .0814605	
1 8		.0758241	.0105876	7.16	0.000	.0550728 .0965753	
1 9		.0670143	.0105151	6.37	0.000	.0464052 .0876235	
1 10		.1086007	.0105203	10.32	0.000	.0879813 .1292202	
1 11		.1106855	.0106014	10.44	0.000	.0899072 .1314639	
1 12		.1271168	.0107045	11.88	0.000	.1061362 .1480973	
_cons		.0431301	.0099719	4.33	0.000	.0235856 .0626746	
/lnalpha		2.658976	.0012505			2.656525 2.661427	
alpha		14.28166	.0178586			14.2467 14.31671	

```
. * next year ED encounters - DI
. logit anyer_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, nolog cluste
> r(bene_id2)
```

```
Logistic regression      Number of obs      = 11,255,795
                        Wald chi2(25)       = 408429.08
                        Prob > chi2        = 0.0000
Log pseudolikelihood = -6947752.4 Pseudo R2          = 0.0456
```

(Std. Err. adjusted for 2,130,226 clusters in bene\_id2)

anyer_ny		Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
----------	--	-------	------------------	---	------	----------------------

age		-.0188388	.0000968	-194.52	0.000	-.0190286	-.018649
male		-.2908484	.0020846	-139.52	0.000	-.2949341	-.2867627
1.any_ip		.9789149	.0055658	175.88	0.000	.9680062	.9898236
yr_afte~2001							
2		.0219591	.0035029	6.27	0.000	.0150935	.0288248
3		.0752016	.0035302	21.30	0.000	.0682826	.0821207
4		.0888671	.0036036	24.66	0.000	.0818042	.09593
5		.1029481	.0036221	28.42	0.000	.0958489	.1100473
6		.1171665	.0036312	32.27	0.000	.1100494	.1242835
7		.1528466	.0036346	42.05	0.000	.1457229	.1599702
8		.1673181	.0036359	46.02	0.000	.1601919	.1744443
9		.1956566	.0036119	54.17	0.000	.1885774	.2027358
10		.2419307	.0035976	67.25	0.000	.2348797	.2489818
11		.4121134	.0035663	115.56	0.000	.4051236	.4191031
12		.4362	.0035831	121.74	0.000	.4291772	.4432227
any_ip#							
yr_afte~2001							
1 2		.0122618	.0073724	1.66	0.096	-.0021879	.0267115
1 3		.0023578	.0074252	0.32	0.751	-.0121952	.0169109
1 4		-.0043417	.0075145	-0.58	0.563	-.0190699	.0103864
1 5		.003143	.0075512	0.42	0.677	-.0116571	.017943
1 6		.0213054	.0075729	2.81	0.005	.0064628	.036148
1 7		.0291691	.00757	3.85	0.000	.014332	.0440061
1 8		.0337414	.0075476	4.47	0.000	.0189483	.0485344
1 9		.0390905	.007508	5.21	0.000	.0243751	.0538058
1 10		.0482281	.0075038	6.43	0.000	.033521	.0629352
1 11		.3308439	.0076521	43.24	0.000	.315846	.3458419
1 12		.3387787	.0077671	43.62	0.000	.3235555	.3540018
_cons		.0648907	.0056801	11.42	0.000	.053758	.0760234

```

. * next year ER visits - DI
. nbreg ervisits_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, nolog clu
> ster(bene_id2)

```

```

Negative binomial regression      Number of obs   = 11,255,795
                                Wald chi2(25)     = 246354.46
Dispersion = mean                Prob > chi2      = 0.0000
Log pseudolikelihood = -13266755 Pseudo R2       = 0.0356

```

(Std. Err. adjusted for 2,130,226 clusters in bene\_id2)

		Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
ervisits_ny							
age		-.0261729	.0001353	-193.44	0.000	-.0264381	-.0259077
male		-.2313269	.003015	-76.73	0.000	-.2372362	-.2254176
1.any_ip		1.057355	.0071135	148.64	0.000	1.043413	1.071298
yr_afte~2001							
2		.0235331	.0042569	5.53	0.000	.0151897	.0318765
3		.0723176	.0042733	16.92	0.000	.0639421	.0806931
4		.0967799	.0045181	21.42	0.000	.0879246	.1056352
5		.1169048	.0045679	25.59	0.000	.1079519	.1258577
6		.1318689	.0045566	28.94	0.000	.1229382	.1407997
7		.1673667	.0045996	36.39	0.000	.1583517	.1763818
8		.1917591	.0046308	41.41	0.000	.1826829	.2008353
9		.2193858	.0046092	47.60	0.000	.2103519	.2284198
10		.2661616	.0044741	59.49	0.000	.2573925	.2749306
11		.4429187	.0044165	100.29	0.000	.4342625	.4515748
12		.4730886	.004403	107.45	0.000	.4644589	.4817183
any_ip#							
yr_afte~2001							
1 2		.0093714	.0081963	1.14	0.253	-.0066931	.0254358
1 3		-.0101848	.0084881	-1.20	0.230	-.0268212	.0064517
1 4		-.0130298	.0088337	-1.48	0.140	-.0303436	.004284
1 5		-.0085886	.0089514	-0.96	0.337	-.0261329	.0089558
1 6		.0075201	.0091762	0.82	0.412	-.0104649	.0255051

1 7		.0118953	.0091119	1.31	0.192	-.0059638	.0297543
1 8		.0222813	.0091579	2.43	0.015	.0043321	.0402305
1 9		.0309283	.0090563	3.42	0.001	.0131783	.0486783
1 10		.0315334	.0089979	3.50	0.000	.0138979	.049169
1 11		.2104036	.008541	24.63	0.000	.1936636	.2271436
1 12		.2135686	.0086082	24.81	0.000	.1966969	.2304402
_cons		.7408678	.0077078	96.12	0.000	.7257607	.7559749
-----							
/lnalpha		.9401553	.0022359			.935773	.9445376
-----							
alpha		2.560379	.0057248			2.549183	2.571624

```
. * next year physician visits - DI
. nbreg phys_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, nolog cluster
> (bene_id2)
```

```
Negative binomial regression      Number of obs   = 11,255,795
                                Wald chi2(25)    = 280764.21
Dispersion = mean                Prob > chi2     = 0.0000
Log pseudolikelihood = -32479578 Pseudo R2      = 0.0115
```

(Std. Err. adjusted for 2,130,226 clusters in bene\_id2)

phys_ny		Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
age		.0131253	.0000779	168.60	0.000	.0129727 .0132778	
male		-.4002281	.0016027	-249.73	0.000	-.4033693 -.3970869	
1.any_ip		.4889064	.0031167	156.87	0.000	.4827978 .4950151	
yr_afte~2001							
2		.0082604	.0015769	5.24	0.000	.0051697 .0113511	
3		.0194082	.0017303	11.22	0.000	.0160168 .0227996	
4		.020172	.0018599	10.85	0.000	.0165266 .0238173	
5		.0163637	.0019269	8.49	0.000	.0125871 .0201403	
6		.0228406	.0019775	11.55	0.000	.0189648 .0267164	
7		.0547468	.0019977	27.40	0.000	.0508313 .0586623	
8		.1158343	.0020081	57.68	0.000	.1118985 .1197701	
9		.1083391	.0020174	53.70	0.000	.1043851 .1122931	
10		.1106263	.002025	54.63	0.000	.1066574 .1145952	
11		.217451	.0020074	108.32	0.000	.2135165 .2213854	
12		.2101055	.0020299	103.51	0.000	.206127 .214084	
any_ip#							
yr_afte~2001							
1 2		-.0091187	.0037695	-2.42	0.016	-.0165067 -.0017307	
1 3		-.0231682	.0039167	-5.92	0.000	-.0308447 -.0154917	
1 4		-.0199786	.0040465	-4.94	0.000	-.0279096 -.0120476	
1 5		-.0148238	.0040806	-3.63	0.000	-.0228216 -.0068261	
1 6		-.010061	.004109	-2.45	0.014	-.0181144 -.0020076	
1 7		-.0155218	.0040841	-3.80	0.000	-.0235265 -.0075171	
1 8		-.0055357	.0040499	-1.37	0.172	-.0134733 .0024019	
1 9		-.0076255	.0040437	-1.89	0.059	-.0155511 .0003001	
1 10		-.0088066	.0040378	-2.18	0.029	-.0167205 -.0008926	
1 11		-.0222319	.0040005	-5.56	0.000	-.0300727 -.0143911	
1 12		-.0237135	.004036	-5.88	0.000	-.0316238 -.0158031	
_cons		1.194777	.0043152	276.88	0.000	1.186319 1.203234	
-----							
/lnalpha		.2992077	.0011538			.2969463 .301469	
-----							
alpha		1.34879	.0015562			1.345743 1.351843	

```
. * next year drug fills - DI
.
. reg ptdfillcnt_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, cluster(b
> ene_id2)
```

```
Linear regression      Number of obs   = 5,787,001
                      F(19, 1324027) = 11294.37
```





(Std. Err. adjusted for 9,218,464 clusters in bene\_id2)

ccw_ny	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0382131	.0000934	409.33	0.000	.0380301	.0383961
male	-.2307993	.0014271	-161.72	0.000	-.2335964	-.2280021
1.any_ip	2.084549	.0027671	753.33	0.000	2.079126	2.089973
yr_afte~2001						
2	.0997575	.0010807	92.31	0.000	.0976394	.1018756
3	.1746402	.001158	150.81	0.000	.1723705	.1769099
4	.2583996	.0012409	208.24	0.000	.2559675	.2608316
5	.3315605	.0013067	253.75	0.000	.3289995	.3341215
6	.4049069	.0013637	296.92	0.000	.4022341	.4075797
7	.4775188	.0014065	339.50	0.000	.4747621	.4802756
8	.5374264	.0014437	372.27	0.000	.5345969	.5402559
9	.6149147	.0014944	411.48	0.000	.6119857	.6178437
10	.6487686	.0015227	426.06	0.000	.6457842	.651753
11	.7292695	.0015498	470.56	0.000	.726232	.7323071
12	.7464974	.0015705	475.33	0.000	.7434193	.7495754
any_ip#						
yr_afte~2001						
1 2	.0617971	.003489	17.71	0.000	.0549588	.0686354
1 3	.1095684	.0036977	29.63	0.000	.102321	.1168158
1 4	.1524826	.0038156	39.96	0.000	.1450041	.159961
1 5	.2352696	.0039238	59.96	0.000	.2275791	.2429602
1 6	.3114961	.0040314	77.27	0.000	.3035947	.3193974
1 7	.3871949	.0041055	94.31	0.000	.3791483	.3952414
1 8	.4566233	.0041926	108.91	0.000	.448406	.4648406
1 9	.6277782	.0043431	144.55	0.000	.6192658	.6362905
1 10	.8240434	.0044579	184.85	0.000	.815306	.8327809
1 11	.8808824	.0045403	194.01	0.000	.8719835	.8897813
1 12	.9387712	.0046345	202.56	0.000	.9296878	.9478547
_cons	-.1117534	.0070514	-15.85	0.000	-.1255739	-.0979329

```
. * next year inpatient use - AGED
. logit anyip_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, nolog cluste
> r(bene_id2)
```

```
Logistic regression          Number of obs      = 58,472,804
                             Wald chi2(25)         = 2812835.75
                             Prob > chi2            = 0.0000
Log pseudolikelihood = -28572185   Pseudo R2          = 0.0613
```

(Std. Err. adjusted for 9,218,464 clusters in bene\_id2)

anyip_ny	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
age	.0402552	.0000538	748.59	0.000	.0401498	.0403606
male	.0674796	.0008012	84.22	0.000	.0659092	.0690499
1.any_ip	1.130984	.0024418	463.18	0.000	1.126198	1.135769
yr_afte~2001						
2	-.0079804	.0018096	-4.41	0.000	-.011527	-.0044337
3	-.0051159	.0017656	-2.90	0.004	-.0085765	-.0016554
4	-.0227543	.0017954	-12.67	0.000	-.0262733	-.0192353
5	-.0467006	.0018217	-25.64	0.000	-.050271	-.0431303
6	-.057163	.0018371	-31.12	0.000	-.0607637	-.0535623
7	-.0973546	.0018581	-52.40	0.000	-.1009963	-.0937129
8	-.1084305	.0018634	-58.19	0.000	-.1120827	-.1047783
9	-.134788	.0018669	-72.20	0.000	-.138447	-.1311291
10	-.1673227	.0018783	-89.08	0.000	-.1710041	-.1636414
11	-.2088576	.0018893	-110.55	0.000	-.2125605	-.2051547
12	-.241985	.0019004	-127.33	0.000	-.2457098	-.2382602
any_ip#						
yr_afte~2001						







any_ip#							
yr_afte~2001							
1	2	.0038096	.0037053	1.03	0.304	-.0034527	.0110719
1	3	.0002801	.0038261	0.07	0.942	-.007219	.0077792
1	4	.0059803	.0038611	1.55	0.121	-.0015873	.013548
1	5	.0114799	.0038682	2.97	0.003	.0038983	.0190614
1	6	.019179	.003943	4.86	0.000	.0114509	.0269071
1	7	.0405529	.0039245	10.33	0.000	.0328611	.0482448
1	8	.0507483	.0039724	12.78	0.000	.0429626	.058534
1	9	.0528308	.0039482	13.38	0.000	.0450924	.0605692
1	10	.0528248	.0039182	13.48	0.000	.0451454	.0605043
1	11	.2513028	.0036234	69.35	0.000	.2442009	.2584046
1	12	.248009	.0036478	67.99	0.000	.2408594	.2551585
_cons		-3.122614	.0051484	-606.52	0.000	-3.132705	-3.112524
/lnalpha		.7680808	.0012898			.7655528	.7706088
alpha		2.155625	.0027804			2.150183	2.161082

```

. * next year physician visits - AGED
. nbreg phys_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, nolog cluster
> (bene_id2)

```

```

Negative binomial regression          Number of obs   = 58,472,804
Dispersion = mean                    Wald chi2(25)   = 784828.37
Log pseudolikelihood = -1.780e+08   Prob > chi2     = 0.0000
Pseudo R2 = 0.0034                   Pseudo R2      = 0.0034

```

(Std. Err. adjusted for 9,218,464 clusters in bene\_id2)

phys_ny		Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
age		-.0006696	.0000401	-16.69	0.000	-.0007483	-.000591
male		-.0403045	.0006189	-65.12	0.000	-.0415176	-.0390915
1.any_ip		.3520299	.0011173	315.07	0.000	.34984	.3542198
yr_afte~2001							
2		.007557	.0005119	14.76	0.000	.0065538	.0085603
3		.0208718	.0005518	37.82	0.000	.0197902	.0219534
4		.0402579	.0005904	68.19	0.000	.0391009	.041415
5		.0515439	.0006193	83.23	0.000	.0503302	.0527577
6		.0637266	.0006415	99.34	0.000	.0624693	.0649839
7		.081479	.0006569	124.04	0.000	.0801915	.0827664
8		.1448651	.0006677	216.96	0.000	.1435564	.1461738
9		.132496	.0006776	195.54	0.000	.131168	.1338241
10		.1282387	.0006853	187.14	0.000	.1268956	.1295817
11		.1341796	.000693	193.62	0.000	.1328214	.1355379
12		.1229249	.0006994	175.76	0.000	.1215541	.1242957
any_ip#							
yr_afte~2001							
1	2	-.0020542	.0014306	-1.44	0.151	-.004858	.0007497
1	3	-.0073187	.0014765	-4.96	0.000	-.0102126	-.0044248
1	4	-.01139	.0015075	-7.56	0.000	-.0143446	-.0084354
1	5	-.0111072	.0015349	-7.24	0.000	-.0141155	-.008099
1	6	-.0125872	.0015561	-8.09	0.000	-.015637	-.0095373
1	7	-.0119426	.0015661	-7.63	0.000	-.0150121	-.008873
1	8	-.010607	.0015742	-6.74	0.000	-.0136923	-.0075217
1	9	-.0046088	.001577	-2.92	0.003	-.0076996	-.001518
1	10	.0014587	.001582	0.92	0.356	-.001642	.0045594
1	11	.0059661	.0015952	3.74	0.000	.0028395	.0090926
1	12	.0089241	.0016122	5.54	0.000	.0057643	.012084
_cons		1.906007	.0030294	629.16	0.000	1.90007	1.911945
/lnalpha		-.0823118	.0005241			-.083339	-.0812846
alpha		.9209847	.0004827			.9200392	.9219313

```

. * next year drug fills - AGED
.
. reg ptdfillcnt_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, cluster(bene_id2)
> ene_id2)

```

```

Linear regression      Number of obs   =   21214772
                      F(19, 4827262)    =   34159.93
                      Prob > F        =   0.0000
                      R-squared        =   0.0490
                      Root MSE       =   38.629

```

(Std. Err. adjusted for 4,827,263 clusters in bene\_id2)

ptdfillcnt~y	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
age	.2442559	.0021316	114.59	0.000	.240078	.2484337
male	-4.186222	.036613	-114.34	0.000	-4.257982	-4.114462
1.any_ip	16.66565	.0672513	247.81	0.000	16.53384	16.79746
yr_afte~2001						
5	6.197664	.0264552	234.27	0.000	6.145813	6.249516
6	7.891274	.0294135	268.29	0.000	7.833624	7.948923
7	8.819378	.0312067	282.61	0.000	8.758214	8.880542
8	9.547243	.0324616	294.11	0.000	9.48362	9.610867
9	9.61915	.0328135	293.15	0.000	9.554837	9.683463
10	10.2606	.0332781	308.33	0.000	10.19538	10.32583
11	12.06548	.0335063	360.10	0.000	11.99981	12.13115
12	12.15066	.0337887	359.61	0.000	12.08444	12.21689
any_ip#						
yr_afte~2001						
1 5	.9882805	.0887483	11.14	0.000	.814337	1.162224
1 6	1.472781	.0932084	15.80	0.000	1.290096	1.655466
1 7	1.530985	.0947212	16.16	0.000	1.345335	1.716635
1 8	1.698408	.0966422	17.57	0.000	1.508993	1.887824
1 9	1.007702	.0949667	10.61	0.000	.8215704	1.193833
1 10	1.221114	.0948879	12.87	0.000	1.035137	1.40709
1 11	2.402554	.0956398	25.12	0.000	2.215103	2.590004
1 12	2.232149	.0966995	23.08	0.000	2.042622	2.421677
_cons	24.66124	.1642832	150.11	0.000	24.33925	24.98323

```

. * next year cost - AGED
. reg mdcprmt_ny age male any_ip##i.yr_after_2001 if enroll_ny==1, cluster(bene_id2)
> _id2)

```

```

Linear regression      Number of obs   =   55576903
                      F(25, 8951401)    =   75402.15
                      Prob > F        =   0.0000
                      R-squared        =   0.0694
                      Root MSE       =   19349

```

(Std. Err. adjusted for 8,951,402 clusters in bene\_id2)

mdcprmt_ny	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
age	199.6369	.4780815	417.58	0.000	198.6998	200.5739
male	595.0367	7.656271	77.72	0.000	580.0306	610.0427
1.any_ip	7825.021	22.70292	344.67	0.000	7780.524	7869.518
yr_afte~2001						
2	453.1335	8.937297	50.70	0.000	435.6167	470.6502
3	834.7568	9.168538	91.05	0.000	816.7868	852.7268
4	1745.307	9.634551	181.15	0.000	1726.423	1764.19
5	2098.698	9.975417	210.39	0.000	2079.147	2118.25
6	2533.445	10.42391	243.04	0.000	2513.014	2553.875
7	2880.174	10.7757	267.28	0.000	2859.054	2901.294
8	3128.7	10.95134	285.69	0.000	3107.236	3150.164

9		3336.255	11.11974	300.03	0.000	3314.461	3358.049
10		3387.947	11.11799	304.73	0.000	3366.156	3409.738
11		3542.148	11.18359	316.73	0.000	3520.229	3564.068
12		3766.637	11.31881	332.78	0.000	3744.453	3788.822
any_ip#							
yr_afte~2001							
1 2		659.2063	30.17713	21.84	0.000	600.0602	718.3524
1 3		1282.853	32.82801	39.08	0.000	1218.512	1347.195
1 4		2377.619	34.81851	68.29	0.000	2309.376	2445.862
1 5		3122.634	36.62577	85.26	0.000	3050.848	3194.419
1 6		3960.728	41.09402	96.38	0.000	3880.185	4041.271
1 7		4812.255	40.61944	118.47	0.000	4732.642	4891.867
1 8		5351.928	42.03404	127.32	0.000	5269.543	5434.313
1 9		5761.906	42.67019	135.03	0.000	5678.274	5845.538
1 10		5853.664	43.53563	134.46	0.000	5768.335	5938.992
1 11		5985.646	44.03603	135.93	0.000	5899.336	6071.955
1 12		6340.104	45.42086	139.59	0.000	6251.081	6429.127
_cons		-9801.267	36.88437	-265.73	0.000	-9873.56	-9728.975

---

## Appendix D2: Optum regressions

---

```
name: <unnamed>
log: /schaeffer-a/sch-projects/dua-data-projects/OPTUM/SSA/pgm/0419run/
> model_di.log
log type: text
opened on: 19 Apr 2019, 15:48:08

.
. use ../../data/analytic0408, clear

. * this file contains DI only
.
. local title1 "mortality_DI_cluster"

. local title2 "CCW_DI_cluster"

. local title3 "anyIP_DI_cluster"

. local title4 "IPdays_DI_cluster"

. local title5 "anyER_DI_cluster"

. local title6 "ERencounters_DI_cluster"

. local title7 "physician_DI_cluster"

. local title8 "drug_DI_cluster"

. local title9 "cost_DI_cluster"

.
. local i=1

.
. * next year mortality
```

```
. logit dead_2yr age male any_ip##i.yr_after_2006 if (enroll_2yr==1|dead_2yr==1
> ), nolog cluster(Patid)
```

```
Logistic regression                Number of obs    = 1,552,638
                                   Wald chi2(15)     = 14766.95
                                   Prob > chi2        = 0.0000
Log pseudolikelihood = -83390.289   Pseudo R2       = 0.0707
```

(Std. Err. adjusted for 551,908 clusters in Patid)

-----						
		Robust				
dead_2yr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----						
age	.0059216	.001062	5.58	0.000	.0038402	.008003
male	.4830293	.0159774	30.23	0.000	.4517143	.5143444
1.any_ip	1.82538	.0466866	39.10	0.000	1.733876	1.916884
yr_afte~2006						
2	-.066407	.0425766	-1.56	0.119	-.1498557	.0170417
3	.1650918	.0399614	4.13	0.000	.0867688	.2434147
4	.0625176	.0405582	1.54	0.123	-.016975	.1420103
5	-.2090845	.0424628	-4.92	0.000	-.2923101	-.1258589
6	-.122383	.0409254	-2.99	0.003	-.2025953	-.0421707
7	.037496	.0407448	0.92	0.357	-.0423623	.1173543
any_ip#						
yr_afte~2006						
1 2	.1116649	.0654323	1.71	0.088	-.01658	.2399098
1 3	.0396144	.0616433	0.64	0.520	-.0812042	.160433
1 4	.1152493	.0613823	1.88	0.060	-.0050579	.2355565
1 5	.0436985	.0639481	0.68	0.494	-.0816375	.1690344
1 6	-.0027263	.0620282	-0.04	0.965	-.1242994	.1188468
1 7	-.1520976	.0618423	-2.46	0.014	-.2733062	-.030889
_cons	-5.63021	.0693849	-81.14	0.000	-5.766202	-5.494218
-----						

```

.      * save estimates
. estimates store m`i', title(`title`i'')

. estimates save `title`i'', replace
file mortality_DI_cluster.ster saved

.

. local i=`i'+1

.

. * next year CCW count
. reg ccw_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, cluster(Patid)

```

```

Linear regression              Number of obs   =   1,537,672
                               F(15, 546227)   =   19304.52
                               Prob > F           =    0.0000
                               R-squared          =    0.2089
                               Root MSE       =    3.2212

```

(Std. Err. adjusted for 546,228 clusters in Patid)

-----						
		Robust				
ccw_2yr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----						
age	.0048795	.0006585	7.41	0.000	.0035889	.0061701
male	-.6892644	.0096402	-71.50	0.000	-.7081589	-.6703698
1.any_ip	3.603372	.023146	155.68	0.000	3.558006	3.648737
yr_afte~2006						
2	.8614766	.0058865	146.35	0.000	.8499393	.8730139
3	1.436984	.0070353	204.25	0.000	1.423195	1.450772
4	1.817177	.0079182	229.49	0.000	1.801657	1.832696
5	2.168394	.0083534	259.58	0.000	2.152021	2.184766
6	2.61529	.0085718	305.10	0.000	2.598489	2.63209
7	3.054254	.009328	327.43	0.000	3.035972	3.072537



```

|
any_ip#|
yr_afte~2006 |
1 2 | .1340534 .0289015 4.64 0.000 .0774073 .1906994
1 3 | .2226163 .0308153 7.22 0.000 .1622193 .2830133
1 4 | .3303263 .0309729 10.67 0.000 .2696203 .3910322
1 5 | .3748588 .0306719 12.22 0.000 .3147429 .4349747
1 6 | .2871504 .030503 9.41 0.000 .2273655 .3469354
1 7 | .0838961 .0310382 2.70 0.007 .0230621 .14473
|
_cons | 3.062114 .0382324 80.09 0.000 2.98718 3.137049

```

```
-----
```

```

. estimates store m`i', title(`title`i'')

```

```

. estimates save `title`i'', replace
file CCW_DI_cluster.ster saved

```

```

. local i=`i'+1

```

```

. * next year inpatient use
. logit any_ip_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog clu
> ster(Patid)

```

```

Logistic regression           Number of obs   = 1,538,772
                             Wald chi2(15)     = 67879.45
                             Prob > chi2        = 0.0000
Log pseudolikelihood = -553897.17   Pseudo R2      = 0.0689

```

(Std. Err. adjusted for 546,867 clusters in Patid)

```
-----
```

```

|
|               Robust
any_ip_2yr |      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----

```

age		-.0101871	.0003435	-29.66	0.000	-.0108604	-.0095138
male		.0136474	.0052934	2.58	0.010	.0032725	.0240223
1.any_ip		1.636471	.0173179	94.50	0.000	1.602528	1.670413
yr_afte~2006							
2		.0773987	.0118041	6.56	0.000	.054263	.1005344
3		.1449345	.0113477	12.77	0.000	.1226934	.1671756
4		.1377674	.0113166	12.17	0.000	.1155873	.1599476
5		.1345901	.0111621	12.06	0.000	.1127128	.1564674
6		.2519399	.0108564	23.21	0.000	.2306617	.2732182
7		.2813687	.011148	25.24	0.000	.259519	.3032184
any_ip#							
yr_afte~2006							
1 2		-.0157711	.0220564	-0.72	0.475	-.059001	.0274587
1 3		-.0011896	.0227438	-0.05	0.958	-.0457667	.0433875
1 4		.0267138	.0225039	1.19	0.235	-.0173929	.0708206
1 5		.0064459	.0220792	0.29	0.770	-.0368285	.0497203
1 6		.0011443	.021599	0.05	0.958	-.041189	.0434775
1 7		-.0267996	.0218271	-1.23	0.220	-.06958	.0159809
_cons		-1.820044	.0214899	-84.69	0.000	-1.862163	-1.777924

```

-----
. estimates store m`i', title(`title`i')

. estimates save `title`i'', replace
file anyIP_DI_cluster.ster saved

.

. local i=`i'+1

.

. * next year inpatient days

. nbreg ip_days_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog cl
> uster(Patid)

```

```

Negative binomial regression          Number of obs   = 1,538,772
                                      Wald chi2(15)    = 44563.17
Dispersion          = mean           Prob > chi2     = 0.0000
Log pseudolikelihood = -1189145.4   Pseudo R2      = 0.0181

```

(Std. Err. adjusted for 546,867 clusters in Patid)

```

-----+-----
            |               Robust
ip_days_2yr |      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
      age |  -.0133278   .0006981   -19.09   0.000   -.0146959   -.0119596
      male |   .0768116   .0100593    7.64   0.000    .0570958    .0965274
  1.any_ip |   1.947995   .0269552   72.27   0.000    1.895164    2.000826
            |
yr_afte~2006 |
      2 |   .1137016   .0217632    5.22   0.000    .0710466    .1563567
      3 |   .1837706   .020852    8.81   0.000    .1429014    .2246397
      4 |   .2115975   .0212861    9.94   0.000    .1698775    .2533175
      5 |   .3182014   .0206678   15.40   0.000    .2776932    .3587096
      6 |   .5187058   .0206958   25.06   0.000    .4781428    .5592689
      7 |   .5399521   .0204414   26.41   0.000    .4998878    .5800164
            |
      any_ip# |
yr_afte~2006 |
      1 2 |  -.0675977   .0343383   -1.97   0.049   -.1348996   -.0002958
      1 3 |  -.0126008   .0346974   -0.36   0.716   -.0806064    .0554048
      1 4 |  -.0339047   .0349264   -0.97   0.332   -.1023591    .0345497
      1 5 |  -.0544865   .034087    -1.60   0.110   -.1212959    .0123228
      1 6 |  -.0773588   .0339211   -2.28   0.023   -.1438428   -.0108747
      1 7 |  -.1103132   .0333002   -3.31   0.001   -.1755804   -.045046
            |
      _cons |   .043551    .0425974    1.02   0.307   -.0399383    .1270403
-----+-----
      /lnalpha |   3.035786   .0036126                3.028705    3.042866
-----+-----

```



```

          7 | .1240219 .0075954 16.33 0.000 .1091351 .1389086
          |
    any_ip#|
yr_afte~2006 |
          1 2 | -.0101353 .0201518 -0.50 0.615 -.0496321 .0293616
          1 3 | .0129627 .0202016 0.64 0.521 -.0266318 .0525572
          1 4 | .047644 .0199467 2.39 0.017 .0085491 .0867389
          1 5 | .0695074 .0195434 3.56 0.000 .031203 .1078118
          1 6 | .0791477 .019284 4.10 0.000 .0413518 .1169436
          1 7 | .0326229 .019433 1.68 0.093 -.0054651 .0707109
          |
          _cons | .4209181 .0190829 22.06 0.000 .3835163 .4583199
-----

```

```
. estimates store m`i', title(`title`i'')
```

```
. estimates save `title`i'', replace
file anyER_DI_cluster.ster saved
```

```
.
```

```
. local i=`i'+1
```

```
.
```

```
. * next year ED encounters
```

```
. nbreg n_er_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog clust
> er(Patid)
```

```

Negative binomial regression          Number of obs   = 1,538,772
                                     Wald chi2(15)     = 34676.87
Dispersion = mean                    Prob > chi2      = 0.0000
Log pseudolikelihood = -1971959.4    Pseudo R2       = 0.0149

```

(Std. Err. adjusted for 546,867 clusters in Patid)

```

-----
          |               Robust
n_er_2yr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]

```

```

-----+-----
      age |  -.031994   .0004544  -70.41   0.000   -.0328845  -.0311035
      male |  -.1887403   .0071167  -26.52   0.000   -.2026888  -.1747919
    1.any_ip |  1.071538   .0188691   56.79   0.000   1.034555   1.108521
          |
yr_afte~2006 |
      2 |  .1059391   .009265   11.43   0.000   .0877801   .1240981
      3 |  .1081706   .0101898   10.62   0.000   .0881989   .1281422
      4 |  .0475578   .0103636    4.59   0.000   .0272456    .06787
      5 |  -.0485272   .010496   -4.62   0.000  -.0690989  -.0279555
      6 |  -.0679817   .0106825   -6.36   0.000  -.088919   -.0470444
      7 |  .0841887   .0109951    7.66   0.000   .0626387   .1057388
          |
      any_ip# |
yr_afte~2006 |
      1 2 |  .0691104   .0216919    3.19   0.001   .0265952   .1116257
      1 3 |  .1204599   .0225636    5.34   0.000   .0762362   .1646837
      1 4 |  .1605473   .0237388    6.76   0.000   .1140202   .2070745
      1 5 |  .2242774   .0237001    9.46   0.000   .1778262   .2707287
      1 6 |  .2283715   .0229541    9.95   0.000   .1833823   .2733607
      1 7 |  .1979701   .022975    8.62   0.000   .1529399   .2430003
          |
      _cons |  1.933777   .0267815   72.21   0.000   1.881286   1.986268
-----+-----
    /lnalpha |  1.916497   .0029767                1.910663   1.922331
-----+-----
      alpha |  6.797109   .0202327                6.757569   6.83688
-----+-----

```

```

. estimates store m`i', title(`title`i')

```

```

. estimates save `title`i'', replace
file ERencontres_DI_cluster.ster saved

```

```

.
. local i=`i'+1

```

```

.
. * next year physician visits
. nbreg n_car_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog clus
> ter(Patid)

```

```

Negative binomial regression      Number of obs   = 1,538,772
                                Wald chi2(15)     = 61051.32
Dispersion = mean                Prob > chi2      = 0.0000
Log pseudolikelihood = -5816926.1 Pseudo R2       = 0.0077

```

(Std. Err. adjusted for 546,867 clusters in Patid)

-----						
		Robust				
n_car_2yr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----						
age	-.0000109	.0001976	-0.06	0.956	-.0003982	.0003764
male	-.2425199	.002722	-89.10	0.000	-.2478549	-.2371848
1.any_ip	.6783047	.0069413	97.72	0.000	.6647001	.6919093
yr_after_2006						
2	.0559313	.0029843	18.74	0.000	.0500823	.0617803
3	.1202415	.0031558	38.10	0.000	.1140563	.1264267
4	.0703519	.0032654	21.54	0.000	.0639518	.0767521
5	.0042243	.0033544	1.26	0.208	-.0023503	.0107988
6	.0242712	.0033684	7.21	0.000	.0176693	.0308732
7	.095226	.003562	26.73	0.000	.0882446	.1022074
any_ip#yr_after_2006						
1 2	-.05305	.0088711	-5.98	0.000	-.070437	-.035663
1 3	-.0747493	.0089374	-8.36	0.000	-.0922664	-.0572323
1 4	-.0649658	.0089202	-7.28	0.000	-.0824491	-.0474825
1 5	-.0366158	.0088642	-4.13	0.000	-.0539894	-.0192422
1 6	-.0429562	.0088226	-4.87	0.000	-.0602481	-.0256643
1 7	-.078364	.0091446	-8.57	0.000	-.096287	-.0604409

_cons		2.728907	.0116098	235.05	0.000	2.706152	2.751662
-----+-----							
/lnalpha		.0271527	.0020911			.0230543	.0312512
-----+-----							
alpha		1.027525	.0021486			1.023322	1.031745
-----							

. estimates store m`i', title(`title`i`)

. estimates save `title`i`, replace  
file physician\_DI\_cluster.ster saved

.  
. local i=`i'+1

. \* next year drug fills

. reg fills\_2yr age male any\_ip##i.yr\_after\_2006 if enroll\_2yr==1, cluster(Patid)

Linear regression	Number of obs	=	1,538,772
	F(15, 546866)	=	2432.36
	Prob > F	=	0.0000
	R-squared	=	0.0451
	Root MSE	=	38.888

(Std. Err. adjusted for 546,867 clusters in Patid)

			Robust				
		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----+-----							
age		.0722649	.007475	9.67	0.000	.0576141	.0869157
male		-5.942968	.1155851	-51.42	0.000	-6.169511	-5.716425
1.any_ip		21.37437	.3265882	65.45	0.000	20.73427	22.01448
yr_after_2006							



2		2.356849	.0698187	33.76	0.000	2.220006	2.493691
3		4.560327	.0883193	51.63	0.000	4.387224	4.73343
4		4.611357	.0949574	48.56	0.000	4.425244	4.79747
5		4.879216	.0991798	49.20	0.000	4.684827	5.073605
6		5.353207	.1007928	53.11	0.000	5.155656	5.550757
7		10.52661	.1160097	90.74	0.000	10.29923	10.75398
any_ip#yr_after_2006							
1 2		.6329786	.413339	1.53	0.126	-.1771528	1.44311
1 3		1.700925	.4366242	3.90	0.000	.8451551	2.556694
1 4		.5116992	.4328996	1.18	0.237	-.3367703	1.360169
1 5		.7900928	.4235578	1.87	0.062	-.040067	1.620253
1 6		.8259512	.4132874	2.00	0.046	.015921	1.635981
1 7		-.5688175	.4387207	-1.30	0.195	-1.428696	.2910611
_cons		29.25924	.437423	66.89	0.000	28.40191	30.11658

```
-----
```

```
. estimates store m`i', title(`title`i'')
```

```
. estimates save `title`i'', replace
```

```
(note: file drug_DI_cluster.ster not found)
```

```
file drug_DI_cluster.ster saved
```

```
.
```

```
. local i=`i'+1
```

```
.
```

```
. * next year cost
```

```
. reg total_cost_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, cluster(Patid)
```

```
Linear regression
```

Number of obs	=	1,538,772
F(15, 546866)	=	969.74
Prob > F	=	0.0000
R-squared	=	0.0443
Root MSE	=	40942

(Std. Err. adjusted for 546,867 clusters in Patid)

```
-----
```

		Robust				[95% Conf. Interval]	
	total_cost_2yr	Coef.	Std. Err.	t	P> t		
	age	-172.0905	7.718978	-22.29	0.000	-187.2195	-156.9616
	male	-903.6148	94.29733	-9.58	0.000	-1088.435	-718.795
	1.any_ip	26838.37	536.1541	50.06	0.000	25787.52	27889.21
	yr_after_2006						
	2	-353.7241	90.26181	-3.92	0.000	-530.6344	-176.8139
	3	-108.2054	92.04729	-1.18	0.240	-288.6152	72.20434
	4	-439.9933	93.30005	-4.72	0.000	-622.8585	-257.1282
	5	-780.6948	93.16858	-8.38	0.000	-963.3023	-598.0874
	6	-732.7653	93.8105	-7.81	0.000	-916.6309	-548.8997
	7	747.8326	129.8764	5.76	0.000	493.279	1002.386
	any_ip#yr_after_2006						
	1 2	-1819.369	620.1172	-2.93	0.003	-3034.779	-603.9592
	1 3	-951.771	670.9866	-1.42	0.156	-2266.884	363.3415
	1 4	-1172.792	662.4456	-1.77	0.077	-2471.164	125.5807
	1 5	-1917.753	651.3843	-2.94	0.003	-3194.446	-641.0609
	1 6	-749.3726	675.6338	-1.11	0.267	-2073.593	574.8482
	1 7	746.4082	728.2576	1.02	0.305	-680.9535	2173.77
	_cons	22348.04	444.9962	50.22	0.000	21475.86	23220.22

```
-----
```

```
. estimates store m`i', title(`title`i`')
```

```
. estimates save `title`i`, replace
```

```
(note: file cost_DI_cluster.ster not found)
```

```
file cost_DI_cluster.ster saved
```

```
.
```

. esttab m1 m2 m3 m4 m5 m6 m7 m8 m9, replace wide label title(regression table-DI)

regression table-DI

{hline 281}

	(1)	(2)	(3)
(4)	(5)	(6)	(7)
(8)			
>	(9)		
	dead_2yr	ccw_2yr	any_ip_2yr
ip_days_2yr	any_er_2yr	n_er_2yr	
n_car_2yr	fills_2yr		
>	total_cost~r		
{hline 281}			
main			
>			
age	0.00592***	(5.58)	0.00488***
(-29.66)	-0.0133***	(-19.09)	-0.0276***
-0.0000109	(-0.06)	0.0723	(-86.24)
> ***	(9.67)	-172.1***	(7.41)
		(-22.29)	-0.0320***
			(-70.41)
male	0.483***	(30.23)	-0.689***
(2.58)	0.0768***	(7.64)	(-71.50)
-0.243***	(-89.10)	-5.943	-0.189***
> ***	(-51.42)	-903.6***	0.0136**
		(-9.58)	(-26.52)
any_ip=0	0	(.)	0
(.)	(.)	0	(.)
0	(.)	0	(.)
>	(.)	0	(.)
any_ip=1	1.825***	(39.10)	3.603***
(94.50)	1.948***	(72.27)	(155.68)
0.678***	(97.72)	21.37	1.072***
> ***	(65.45)	26838.4***	1.636***
		(50.06)	(56.79)
yr_after_2006=1	0	(.)	0
(.)	0	(.)	(.)
0	(.)	0	(.)
>	(.)	0	(.)
yr_after_2006=2	-0.0664	(-1.56)	0.861***
(6.56)	0.114***	(5.22)	(146.35)
0.0559***	(18.74)	2.357	0.106***
> ***	(33.76)	-353.7***	0.0774***
		(-3.92)	(11.43)
yr_after_2006=3	0.165***	(4.13)	1.437***
(12.77)	0.184***	(8.81)	(204.25)
0.120***	(38.10)	4.560	0.108***
> ***	(51.63)	-108.2	0.145***
		(-1.18)	(10.62)

```

yr_after_2006=4      0.0625      (1.54)      1.817***      (229.49)      0.138***
(12.17)      0.212***      (9.94)      0.105***      (14.12)      0.0476***      (4.59)
0.0704***      (21.54)      4.611
> ***      (48.56)      -440.0***      (-4.72)

yr_after_2006=5      -0.209***      (-4.92)      2.168***      (259.58)      0.135***
(12.06)      0.318***      (15.40)      0.0119      (1.60)      -0.0485***      (-4.62)
0.00422      (1.26)      4.879
> ***      (49.20)      -780.7***      (-8.38)

yr_after_2006=6      -0.122**      (-2.99)      2.615***      (305.10)      0.252***
(23.21)      0.519***      (25.06)      -0.0339***      (-4.55)      -0.0680***      (-6.36)
0.0243***      (7.21)      5.353
> ***      (53.11)      -732.8***      (-7.81)

yr_after_2006=7      0.0375      (0.92)      3.054***      (327.43)      0.281***
(25.24)      0.540***      (26.41)      0.124***      (16.33)      0.0842***      (7.66)
0.0952***      (26.73)      10.53
> ***      (90.74)      747.8***      (5.76)

any_ip=0 # yr_afte~1      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0
>      (.)      0      (.)

any_ip=0 # yr_afte~2      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0
>      (.)      0      (.)

any_ip=0 # yr_afte~3      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0
>      (.)      0      (.)

any_ip=0 # yr_afte~4      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0
>      (.)      0      (.)

any_ip=0 # yr_afte~5      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0
>      (.)      0      (.)

any_ip=0 # yr_afte~6      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0
>      (.)      0      (.)

any_ip=0 # yr_afte~7      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0
>      (.)      0      (.)

any_ip=1 # yr_afte~1      0      (.)      0      (.)      0
(.)      0      (.)      0      (.)      0
0      (.)      0

```

```

>          (.)          0          (.)
any_ip=1 # yr_afte~2      0.112      (1.71)      0.134***      (4.64)      -0.0158
(-0.72)      -0.0676*      (-1.97)      -0.0101      (-0.50)      0.0691**      (3.19)
-0.0531***      (-5.98)      0.633
>          (1.53)      -1819.4**      (-2.93)
any_ip=1 # yr_afte~3      0.0396      (0.64)      0.223***      (7.22)      -0.00119
(-0.05)      -0.0126      (-0.36)      0.0130      (0.64)      0.120***      (5.34)
-0.0747***      (-8.36)      1.701
> ***      (3.90)      -951.8      (-1.42)
any_ip=1 # yr_afte~4      0.115      (1.88)      0.330***      (10.67)      0.0267
(1.19)      -0.0339      (-0.97)      0.0476*      (2.39)      0.161***      (6.76)
-0.0650***      (-7.28)      0.512
>          (1.18)      -1172.8      (-1.77)
any_ip=1 # yr_afte~5      0.0437      (0.68)      0.375***      (12.22)      0.00645
(0.29)      -0.0545      (-1.60)      0.0695***      (3.56)      0.224***      (9.46)
-0.0366***      (-4.13)      0.790
>          (1.87)      -1917.8**      (-2.94)
any_ip=1 # yr_afte~6      -0.00273      (-0.04)      0.287***      (9.41)      0.00114
(0.05)      -0.0774*      (-2.28)      0.0791***      (4.10)      0.228***      (9.95)
-0.0430***      (-4.87)      0.826
> *      (2.00)      -749.4      (-1.11)
any_ip=1 # yr_afte~7      -0.152*      (-2.46)      0.0839**      (2.70)      -0.0268
(-1.23)      -0.110***      (-3.31)      0.0326      (1.68)      0.198***      (8.62)
-0.0784***      (-8.57)      -0.569
>          (-1.30)      746.4      (1.02)
Constant      -5.630***      (-81.14)      3.062***      (80.09)      -1.820***
(-84.69)      0.0436      (1.02)      0.421***      (22.06)      1.934***      (72.21)
2.729***      (235.05)      29.26
> ***      (66.89)      22348.0***      (50.22)

{hline 281}
/
>
lnalpha
3.036***      (840.33)      1.916***      (643.84)      0.0272***
(12.99)
>
{hline 281}
Observations      1552638      1537672      1538772
1538772      1538772      1538772
1538772
>          1538772
{hline 281}
t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

```

.

. log close

name: <unnamed>

log: /schaeffer-a/sch-projects/dua-data-projects/OPTUM/SSA/pgm/0419run/model\_di.log

log type: text

closed on: 19 Apr 2019, 15:52:21

---

```
-----  
name: <unnamed>  
log: /schaeffer-a/sch-projects/dua-data-projects/OPTUM/SSA/pgm/0419run/  
> model_comm.log  
log type: text  
opened on: 19 Apr 2019, 15:48:21  
  
.   
. use ../../data/analytic0408_comm, clear  
  
. xtset Patid year  
panel variable: Patid (unbalanced)  
time variable: year, 2007 to 2014, but with gaps  
delta: 1 unit  
  
. gen ccw_2yr=f1.ccw_count  
(18,947,739 missing values generated)  
  
. gen any_er_2yr = f1.any_er  
(18,898,255 missing values generated)  
  
. gen fills_2yr = f1.fills  
(18,898,255 missing values generated)  
  
. gen total_cost_2yr=f1.total_cost  
(18,898,255 missing values generated)  
  
. gen yr_after_2006 = year-2006  
  
. gen enroll_2yr=f1.enroll==2  
  
. gen dead_2yr=f1.dead==1  
  
. gen any_ip_2yr=f1.any_ip  
(18,898,255 missing values generated)
```

```
. gen ip_days_2yr=f1.ip_days
(18,898,255 missing values generated)

. gen n_car_2yr=f1.n_car
(18,898,255 missing values generated)

. gen n_er_2yr=f1.n_er
(18,898,255 missing values generated)

.
. * non-DI only
. keep if medicare==0
(0 observations deleted)

.
. local title10 "mortality_comm_cluster"

. local title11 "CCW_comm_cluster"

. local title12 "anyIP_comm_cluster"

. local title13 "IPdays_comm_cluster"

. local title14 "anyER_comm_cluster"

. local title15 "ERencounters_comm_cluster"

. local title16 "physician_comm_cluster"

. local title17 "drug_comm_cluster"

. local title18 "cost_comm_cluster"

.
. local i=10
```



```

.
. * next year mortality
. logit dead_2yr age male any_ip##i.yr_after_2006 if (enroll_2yr==1|dead_2yr==1
> ), nolog cluster(Patid)

```

```

Logistic regression                               Number of obs   = 30,749,311
                                                    Wald chi2(15)   = 117728.37
                                                    Prob > chi2     = 0.0000
Log pseudolikelihood = -363040.97                Pseudo R2      = 0.1257

```

(Std. Err. adjusted for 11,088,908 clusters in Patid)

-----						
		Robust				
dead_2yr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----						
age	.0756759	.0004859	155.75	0.000	.0747236	.0766283
male	.5183338	.008642	59.98	0.000	.5013958	.5352718
1.any_ip	2.378903	.020191	117.82	0.000	2.33933	2.418477
yr_afte~2006						
2	.0158466	.0172496	0.92	0.358	-.017962	.0496552
3	-.0104196	.017395	-0.60	0.549	-.0445131	.0236739
4	-.2027213	.0183252	-11.06	0.000	-.238638	-.1668045
5	-.5996692	.0205671	-29.16	0.000	-.6399799	-.5593585
6	-.6852453	.0210579	-32.54	0.000	-.726518	-.6439727
7	-.7144046	.0221156	-32.30	0.000	-.7577504	-.6710587
any_ip#						
yr_afte~2006						
1 2	.033878	.0283912	1.19	0.233	-.0217678	.0895237
1 3	.0140838	.0288396	0.49	0.625	-.0424408	.0706084
1 4	.1315741	.0298553	4.41	0.000	.0730588	.1900894
1 5	.1565894	.03349	4.68	0.000	.0909502	.2222287
1 6	.1122072	.0348237	3.22	0.001	.043954	.1804605
1 7	.1637813	.0365178	4.48	0.000	.0922077	.2353548

```
_cons | -10.28771      .02713  -379.20   0.000   -10.34089  -10.23454
```

```
-----  
.          * save estimates  
. estimates store m`i', title(`title`i')  
  
. estimates save `title`i'', replace  
(note: file mortality_comm_cluster.ster not found)  
file mortality_comm_cluster.ster saved  
  
.   
. local i=`i'+1  
  
.   
. * next year CCW count  
. reg ccw_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, cluster(Patid)
```

```
Linear regression                Number of obs    =    30699837  
                                F(15, 11070611)  >    99999.00  
                                Prob > F             =     0.0000  
                                R-squared              =     0.1605  
                                Root MSE           =     2.1449
```

(Std. Err. adjusted for 11,070,612 clusters in Patid)

```
-----  
                |                Robust  
ccw_2yr |                Coef.  Std. Err.      t    P>|t|     [95% Conf. Interval]  
-----+-----  
    age |    .0608093   .0000612   993.47   0.000   .0606893   .0609293  
    male |   -.4101459   .0015015  -273.15   0.000   -.4130889   -.407203  
1.any_ip |    1.657921   .0053074   312.38   0.000   1.647518   1.668323  
    |  
yr_afte~2006 |  
    2 |    .3542356   .0008029   441.22   0.000   .3526621   .3558092  
    3 |    .5749337   .0009832   584.74   0.000   .5730066   .5768608  
    4 |    .7489971   .0011066   676.87   0.000   .7468283   .7511659
```

5		.8625796	.0011835	728.83	0.000	.86026	.8648993
6		.9589954	.0012406	773.00	0.000	.9565638	.9614269
7		.9776318	.0013446	727.10	0.000	.9749965	.9802671
any_ip#							
yr_afte~2006							
1 2		.0689635	.0072317	9.54	0.000	.0547896	.0831374
1 3		.1145919	.0077933	14.70	0.000	.0993173	.1298666
1 4		.1607195	.0082245	19.54	0.000	.1445999	.1768392
1 5		.1774571	.0084918	20.90	0.000	.1608134	.1941008
1 6		.1935409	.0087345	22.16	0.000	.1764215	.2106602
1 7		.1852743	.0093208	19.88	0.000	.1670058	.2035429
_cons		-.9410492	.002581	-364.61	0.000	-.9461079	-.9359905

-----

```
. estimates store m`i', title(`title`i')
```

```
. estimates save `title`i'', replace
```

```
(note: file CCW_comm_cluster.ster not found)
```

```
file CCW_comm_cluster.ster saved
```

```
.
```

```
. local i=`i'+1
```

```
.
```

```
. * next year inpatient use
```

```
. logit any_ip_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog clu
```

```
> ster(Patid)
```

Logistic regression	Number of obs	=	30,700,196
	Wald chi2(15)	=	371770.39
	Prob > chi2	=	0.0000
Log pseudolikelihood = -5940512.2	Pseudo R2	=	0.0280

(Std. Err. adjusted for 11,070,770 clusters in Patid)

```
-----
```

		Robust				[95% Conf. Interval]	
any_ip_2yr	Coef.	Std. Err.	z	P> z			
age	.0021602	.000074	29.20	0.000	.0020152	.0023052	
male	-.5973621	.0019239	-310.49	0.000	-.6011329	-.5935912	
1.any_ip	1.154068	.0062584	184.40	0.000	1.141802	1.166334	
yr_afte~2006							
2	-.0188713	.0031954	-5.91	0.000	-.0251341	-.0126085	
3	-.0205417	.0031556	-6.51	0.000	-.0267265	-.0143569	
4	-.0533248	.003211	-16.61	0.000	-.0596183	-.0470312	
5	-.0875909	.0032428	-27.01	0.000	-.0939466	-.0812353	
6	-.1287557	.0032807	-39.25	0.000	-.1351858	-.1223257	
7	-.1843789	.0034482	-53.47	0.000	-.1911372	-.1776206	
any_ip#							
yr_afte~2006							
1 2	-.0097708	.0083685	-1.17	0.243	-.0261729	.0066312	
1 3	-.0057814	.0088189	-0.66	0.512	-.0230661	.0115033	
1 4	.040588	.0089674	4.53	0.000	.0230122	.0581637	
1 5	.0493877	.0090882	5.43	0.000	.0315751	.0672003	
1 6	.1398686	.0091008	15.37	0.000	.1220314	.1577058	
1 7	.1387514	.00964	14.39	0.000	.1198572	.1576455	
_cons	-2.817212	.0040234	-700.20	0.000	-2.825098	-2.809327	

```
-----
```

```
. estimates store m`i', title(`title`i')

. estimates save `title`i'', replace
(note: file anyIP_comm_cluster.ster not found)
file anyIP_comm_cluster.ster saved

.

. local i=`i'+1
```

```

.
. * next year inpatient days
. nbreg ip_days_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog cl
> uster(Patid)

```

```

Negative binomial regression           Number of obs   = 30,700,196
                                     Wald chi2(15)    = 115986.98
Dispersion = mean                    Prob > chi2     = 0.0000
Log pseudolikelihood = -9673692.5    Pseudo R2      = 0.0115

```

(Std. Err. adjusted for 11,070,770 clusters in Patid)

```

-----+-----

```

		Robust				
ip_days_2yr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	.0144822	.0001547	93.63	0.000	.014179	.0147853
male	-.4584438	.0038415	-119.34	0.000	-.465973	-.4509145
1.any_ip	1.761078	.0128721	136.81	0.000	1.735849	1.786307
yr_afte~2006						
2	-.012337	.0064164	-1.92	0.055	-.0249129	.0002389
3	.000287	.0064821	0.04	0.965	-.0124177	.0129916
4	-.0311273	.0064322	-4.84	0.000	-.0437343	-.0185204
5	-.0455074	.0066248	-6.87	0.000	-.0584917	-.032523
6	-.083966	.0067116	-12.51	0.000	-.0971205	-.0708114
7	-.1434359	.0069004	-20.79	0.000	-.1569604	-.1299114
any_ip#						
yr_afte~2006						
1 2	-.0034437	.0169886	-0.20	0.839	-.0367408	.0298534
1 3	.0143282	.0181227	0.79	0.429	-.0211917	.0498481
1 4	.0633557	.0181288	3.49	0.000	.0278239	.0988876
1 5	.0867132	.0184656	4.70	0.000	.0505213	.1229052
1 6	.2200772	.0193439	11.38	0.000	.1821637	.2579906
1 7	.2525332	.0205172	12.31	0.000	.2123201	.2927462

_cons		-2.047953	.0076046	-269.30	0.000	-2.062858	-2.033049
-----+-----							
/lnalpha		3.791186	.0013702			3.7885	3.793871
-----+-----							
alpha		44.3089	.0607117			44.19007	44.42805
-----							

```
. estimates store m`i', title(`title`i')
```

```
. estimates save `title`i', replace
```

```
(note: file IPdays_comm_cluster.ster not found)
```

```
file IPdays_comm_cluster.ster saved
```

```
.
```

```
. local i=`i'+1
```

```
.
```

```
. * next year ER visit
```

```
. logit any_er_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog clu
```

```
> ster(Patid)
```

```
Logistic regression                               Number of obs   = 30,700,196
                                                    Wald chi2(15)   = 199711.38
                                                    Prob > chi2     = 0.0000
Log pseudolikelihood = -14391944                Pseudo R2      = 0.0086
```

(Std. Err. adjusted for 11,070,770 clusters in Patid)

		Robust				
any_er_2yr		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
-----+-----						
age		.0019565	.0000475	41.18	0.000	.0018634 .0020496
male		-.3024817	.0012119	-249.59	0.000	-.304857 -.3001064
1.any_ip		.6381619	.0046498	137.24	0.000	.6290484 .6472755
-----						



Dispersion = mean Prob > chi2 = 0.0000  
 Log pseudolikelihood = -26339517 Pseudo R2 = 0.0032

(Std. Err. adjusted for 11,070,770 clusters in Patid)

```

-----+-----
              |               Robust
n_er_2yr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      age |   .0043684   .0000722    60.53   0.000     .004227     .0045098
      male |  -.3347408   .0019567  -171.08   0.000    - .3385758    - .3309058
  1.any_ip |   .6650478   .0065912   100.90   0.000     .6521293     .6779664
      |
yr_afte~2006 |
      2 |   .0262587   .0021598    12.16   0.000     .0220256     .0304919
      3 |   .013351   .0023815     5.61   0.000     .0086833     .0180187
      4 |  -.0349637   .0024774   -14.11   0.000    - .0398194    - .030108
      5 |  -.1594079   .0025702   -62.02   0.000    - .1644455    - .1543703
      6 |  -.1979515   .0026042   -76.01   0.000    - .2030558    - .1928473
      7 |  -.164165   .002652   -61.90   0.000    - .1693628    - .1589672
      |
      any_ip#|
yr_afte~2006 |
      1 2 |  -.026613   .0085474    -3.11   0.002    - .0433656    - .0098604
      1 3 |  -.0572037   .0090845    -6.30   0.000    - .0750089    - .0393985
      1 4 |  -.0659046   .0091659    -7.19   0.000    - .0838694    - .0479398
      1 5 |  -.0577144   .0095536    -6.04   0.000    - .0764392    - .0389896
      1 6 |  -.0460923   .0095737    -4.81   0.000    - .0648564    - .0273282
      1 7 |  -.0436227   .009719     -4.49   0.000    - .0626715    - .0245738
      |
      _cons |  -.407701   .0035243  -115.68   0.000    - .4146085    - .4007935
-----+-----
      /lnalpha |   2.305311   .0007919                2.303758     2.306863
-----+-----
      alpha |   10.02729   .007941                10.01174     10.04287
-----+-----

```



```

. estimates store m`i', title(`title`i'')

. estimates save `title`i'', replace
(note: file EEncounters_comm_cluster.ster not found)
file EEncounters_comm_cluster.ster saved

.

. local i=`i'+1

.

. * next year physician visits
. nbreg n_car_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, nolog clus
> ter(Patid)

```

```

Negative binomial regression          Number of obs   = 30,700,196
                                     Wald chi2(15)    = 1236825.85
Dispersion = mean                    Prob > chi2      = 0.0000
Log pseudolikelihood = -95490974     Pseudo R2       = 0.0152

```

(Std. Err. adjusted for 11,070,770 clusters in Patid)

```

-----+-----
          |               Robust
n_car_2yr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      age |   .0213766   .0000293    729.67  0.000   .0213192   .0214341
      male |  -.5226766   .0007481   -698.70  0.000  -.5241428  -.5212104
  1.any_ip |   .576603    .0025475    226.34  0.000   .57161    .5815959
          |
yr_afte~2006 |
      2 |  -.021228    .0007262   -29.23  0.000  -.0226513  -.0198047
      3 |   .0023247   .0007897    2.94   0.003   .0007769   .0038724
      4 |  -.0225216   .0008172   -27.56  0.000  -.0241234  -.0209199
      5 |  -.0951922   .0008531  -111.58  0.000  -.0968644  -.0935201
      6 |  -.1604099   .000881    -182.09  0.000  -.1621365  -.1586833
      7 |  -.0725132   .0009348   -77.57  0.000  -.0743453  -.0706811
          |

```



(Std. Err. adjusted for 11,070,770 clusters in Patid)

```
-----
```

		Robust				
	fills_2yr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	age	.4991594	.0005143	970.60	0.000	.4981515 .5001674
	male	-3.042025	.0124339	-244.66	0.000	-3.066394 -3.017655
	1.any_ip	10.16078	.0567383	179.08	0.000	10.04958 10.27199
yr_afte~2006						
	2	.181636	.0073428	24.74	0.000	.1672443 .1960276
	3	.2414705	.008925	27.06	0.000	.2239779 .2589631
	4	-.3089754	.0095175	-32.46	0.000	-.3276293 -.2903216
	5	-.5931539	.009923	-59.78	0.000	-.6126025 -.5737052
	6	-1.051908	.010079	-104.37	0.000	-1.071663 -1.032154
	7	-.3226251	.0112342	-28.72	0.000	-.3446437 -.3006066
	any_ip#					
yr_afte~2006						
	1 2	-.046076	.0738124	-0.62	0.532	-.1907457 .0985937
	1 3	-.0256296	.077743	-0.33	0.742	-.1780031 .1267439
	1 4	-.3409326	.0789378	-4.32	0.000	-.495648 -.1862173
	1 5	-.6888356	.0790179	-8.72	0.000	-.8437078 -.5339633
	1 6	-1.149623	.0783561	-14.67	0.000	-1.303198 -.9960479
	1 7	-.4473101	.0862978	-5.18	0.000	-.6164506 -.2781695
	_cons	-7.848581	.020466	-383.49	0.000	-7.888694 -7.808469

```
-----
```

. estimates store m`i', title(`title`i`)

. estimates save `title`i`, replace

(note: file drug\_comm\_cluster.ster not found)

file drug\_comm\_cluster.ster saved

```

.
. local i=`i'+1
.
. * next year cost
. reg total_cost_2yr age male any_ip##i.yr_after_2006 if enroll_2yr==1, cluster
> (Patid)

```

```

Linear regression                Number of obs    =    30700196
                                F(15, 11070769)  =    15686.69
                                Prob > F              =     0.0000
                                R-squared              =     0.0288
                                Root MSE            =     18206

```

(Std. Err. adjusted for 11,070,770 clusters in Patid)

-----						
		Robust				
total_cost~r	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----						
age	156.1315	.3794795	411.44	0.000	155.3877	156.8752
male	-1116.178	9.440942	-118.23	0.000	-1134.682	-1097.674
1.any_ip	11379.53	92.53384	122.98	0.000	11198.17	11560.9
yr_afte~2006						
2	-309.1103	9.70448	-31.85	0.000	-328.1308	-290.0899
3	-364.3293	10.94715	-33.28	0.000	-385.7853	-342.8733
4	-627.145	11.01272	-56.95	0.000	-648.7296	-605.5605
5	-1101.174	10.61484	-103.74	0.000	-1121.979	-1080.37
6	-1185.23	10.2305	-115.85	0.000	-1205.281	-1165.178
7	-1151.388	10.72535	-107.35	0.000	-1172.409	-1130.367
any_ip#						
yr_afte~2006						
1 2	-705.661	115.6867	-6.10	0.000	-932.4029	-478.9192
1 3	-464.4167	135.9446	-3.42	0.001	-730.8633	-197.9701
1 4	-666.3258	131.1235	-5.08	0.000	-923.3231	-409.3285

1 5		-1651.152	127.4631	-12.95	0.000	-1900.975	-1401.329
1 6		-968.0793	130.4546	-7.42	0.000	-1223.766	-712.393
1 7		-816.008	138.9571	-5.87	0.000	-1088.359	-543.657
_cons		-536.0818	16.98503	-31.56	0.000	-569.3718	-502.7917

-----

```
. estimates store m`i', title(`title`i`)
```

```
. estimates save `title`i`, replace
```

```
(note: file cost_comm_cluster.ster not found)
```

```
file cost_comm_cluster.ster saved
```

```
.
```

```
. esttab m10 m11 m12 m13 m14 m15 m16 m17 m18, replace wide label title(regressi
```

```
> on table-commercial)
```

```
regression table-commercial
```

```
{hline 281}
```

```

(1) (2)
> (3) (4) (5)
> (6) (7)
> (8) (9)
dead_2yr ccw_2yr
> any_ip_2yr ip_days_2yr any_er_2yr
> n_er_2yr n_car_2yr fills_
> 2yr total_cost~r
```

```
{hline 281}
```

```
main
```

```
>
```

```
>
```

```
>
```

```

age 0.0757*** (155.75) 0.0608*** (993.47)
> 0.00216*** (29.20) 0.0145*** (93.63) 0.00196***
> (41.18) 0.00437*** (60.53) 0.0214*** (729.67) 0.
> 499*** (970.60) 156.1*** (411.44)
```

```

male                0.518***      (59.98)      -0.410***      (-273.15)
>      -0.597***      (-310.49)      -0.458***      (-119.34)      -0.302***
> (-249.59)      -0.335***      (-171.08)      -0.523***      (-698.70)      -3.
> 042***      (-244.66)      -1116.2***      (-118.23)
any_ip=0            0                (.)          0                (.)
>      0                (.)          0                (.)          0
>      (.)          0                (.)          0                (.)
> 0                (.)          0                (.)
any_ip=1            2.379***      (117.82)      1.658***      (312.38)
>      1.154***      (184.40)      1.761***      (136.81)      0.638***
> (137.24)      0.665***      (100.90)      0.577***      (226.34)      10
> .16***      (179.08)      11379.5***      (122.98)
yr_after_2006=1    0                (.)          0                (.)
>      0                (.)          0                (.)          0
>      (.)          0                (.)          0                (.)
> 0                (.)          0                (.)
yr_after_2006=2    0.0158          (0.92)      0.354***      (441.22)
>      -0.0189***      (-5.91)      -0.0123          (-1.92)      0.0414***
> (25.65)      0.0263***      (12.16)      -0.0212***      (-29.23)      0.
> 182***      (24.74)      -309.1***      (-31.85)
yr_after_2006=3    -0.0104          (-0.60)      0.575***      (584.74)
>      -0.0205***      (-6.51)      0.000287          (0.04)      0.0617***
> (36.64)      0.0134***      (5.61)      0.00232**          (2.94)      0.
> 241***      (27.06)      -364.3***      (-33.28)
yr_after_2006=4    -0.203***      (-11.06)      0.749***      (676.87)
>      -0.0533***      (-16.61)      -0.0311***      (-4.84)      0.0334***
> (19.31)      -0.0350***      (-14.11)      -0.0225***      (-27.56)      -0.
> 309***      (-32.46)      -627.1***      (-56.95)
yr_after_2006=5    -0.600***      (-29.16)      0.863***      (728.83)
>      -0.0876***      (-27.01)      -0.0455***      (-6.87)      -0.110***
> (-61.52)      -0.159***      (-62.02)      -0.0952***      (-111.58)      -0.
> 593***      (-59.78)      -1101.2***      (-103.74)
yr_after_2006=6    -0.685***      (-32.54)      0.959***      (773.00)
>      -0.129***      (-39.25)      -0.0840***      (-12.51)      -0.145***
> (-80.59)      -0.198***      (-76.01)      -0.160***      (-182.09)      -1.
> 052***      (-104.37)      -1185.2***      (-115.85)

```

```

yr_after_2006=7          -0.714***      (-32.30)          0.978***      (727.10)
>      -0.184***      (-53.47)          -0.143***      (-20.79)          -0.0887***
>      (-47.97)          -0.164***      (-61.90)          -0.0725***      (-77.57)          -0.
> 323***      (-28.72)          -1151.4***      (-107.35)
any_ip=0 # yr_afte~1          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)
any_ip=0 # yr_afte~2          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)
any_ip=0 # yr_afte~3          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)
any_ip=0 # yr_afte~4          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)
any_ip=0 # yr_afte~5          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)
any_ip=0 # yr_afte~6          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)
any_ip=0 # yr_afte~7          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)
any_ip=1 # yr_afte~1          0          (.)          0          (.)
>      0          (.)          0          (.)          0
>      (.)          0          (.)          0          (.)
> 0          (.)          0          (.)

```

```

any_ip=1 # yr_afte~2      0.0339      (1.19)      0.0690***      (9.54)
>   -0.00977      (-1.17)   -0.00344      (-0.20)   -0.0124
>   (-1.94)      -0.0266**   (-3.11)   -0.0283***   (-8.28)   -0.0
> 461      (-0.62)   -705.7***   (-6.10)
any_ip=1 # yr_afte~3      0.0141      (0.49)      0.115***      (14.70)
>   -0.00578      (-0.66)   0.0143      (0.79)   -0.0333***
>   (-5.08)      -0.0572***   (-6.30)   -0.0232***   (-6.53)   -0.0
> 256      (-0.33)   -464.4***   (-3.42)
any_ip=1 # yr_afte~4      0.132***      (4.41)      0.161***      (19.54)
>   0.0406***      (4.53)   0.0634***      (3.49)   -0.0408***
>   (-6.10)      -0.0659***   (-7.19)   -0.0190***   (-5.19)   -0.
> 341***      (-4.32)   -666.3***   (-5.08)
any_ip=1 # yr_afte~5      0.157***      (4.68)      0.177***      (20.90)
>   0.0494***      (5.43)   0.0867***      (4.70)   -0.0292***
>   (-4.27)      -0.0577***   (-6.04)   -0.00681      (-1.80)   -0.
> 689***      (-8.72)   -1651.2***   (-12.95)
any_ip=1 # yr_afte~6      0.112**      (3.22)      0.194***      (22.16)
>   0.140***      (15.37)   0.220***      (11.38)   -0.0151*
>   (-2.18)      -0.0461***   (-4.81)   0.0172***      (4.19)   -1.
> 150***      (-14.67)   -968.1***   (-7.42)
any_ip=1 # yr_afte~7      0.164***      (4.48)      0.185***      (19.88)
>   0.139***      (14.39)   0.253***      (12.31)   -0.0276***
>   (-3.83)      -0.0436***   (-4.49)   0.0221***      (4.57)   -0.
> 447***      (-5.18)   -816.0***   (-5.87)
Constant      -10.29***   (-379.20)   -0.941***   (-364.61)
>   -2.817***   (-700.20)   -2.048***   (-269.30)   -1.461***
>   (-609.76)   -0.408***   (-115.68)   1.452***   (1055.82)   -7.
> 849***      (-383.49)   -536.1***   (-31.56)
{hline 281}
/
>
>
>
lnalpha
>
>           3.791***   (2766.90)
>
>           2.305***   (2910.96)   0.317***   (665.80)

```



```
>
{hline 281}
Observations          30749311          30699837
>   30700196          30700196          30700196
>           30700196          30700196          30700
> 196           30700196

{hline 281}
t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

.
. log close
    name: <unnamed>
    log: /schaeffer-a/sch-projects/dua-data-projects/OPTUM/SSA/pgm/0419run/
> model_comm.log
    log type: text
closed on: 19 Apr 2019, 17:11:53
```

---