ONLINE APPENDIX

To

Is Great Information Good Enough? Evidence from Physicians as Patients

Michael Frakes, Jonathan Gruber and Anupam Jena

Additional Information on Low- and High-Value Care Samples and Outcome Measures

Low-Value Care: Labor and Delivery

The first instance of low-value care that we consider is cesarean delivery. The low-value label applied by the health care community in the cesarean context is not premised on the idea that cesareans should never be performed, but rather that they should be performed at lower rates than currently prevail—e.g., lower than the nearly 30 percent U.S. average rate. As stated in a recent report by the World Health Organization (2015), the international health care community has long considered that the ideal cesarean rate is less than half of the prevailing U.S. rate—i.e., roughly 10-15%.[[1]](#footnote-1) According to the WHO, cesarean rates higher than this level have not been shown to be associated with lower rates of maternal or neonatal mortality, but are associated with both higher medical expenditures and higher incidences of post-surgical complications.

To explore the potential impact of greater information disclosure on cesarean rates, we compare cesarean use between physician and non-physician mothers using the sub-sample of deliveries contained within the inpatient MHS sample (pooling both Direct and Purchased Care inpatient records). To ensure age-overlap between our physician mother group and our comparison officer group, we limit the sample to mothers that are at least 25 years old, since there are no physician mothers under the age of 25 in the MDR records. For the comparison of dependents of physicians with dependents of non-physician officers, however, we impose no age restrictions, as there is overlap throughout all ages.[[2]](#footnote-2)

We take two approaches to identifying cesarean deliveries. First, we focus on the sub-sample of all deliveries and flag the incidence of any cesarean delivery. Second, we follow the definition of “primary” cesarean delivery set forth by the Agency for Health Care Research and Quality (AHRQ). In essence, primary cesarean section rates reflect the incidence of cesareans after removing from the sample certain types of deliveries over which physicians have less discretion in the cesarean decision (i.e., for which cesarean delivery may be clinically indicated): breech presentation deliveries, multiple deliveries (e.g., delivery of twins), and previous cesarean deliveries.[[3]](#footnote-3)

Low Value Care: Pre-operative care for low-risk surgeries

While cesareans are low-value in the sense of being over-used, other forms of health care are “low value” in the sense that the relevant medical guidelines recommend that such services not be performed at all. We turn now to low-value measures of this nature. In setting forth a list of low-value services, Schwartz et al. (2014) draw from procedures and tests characterized as low-value by sources such as the American Board of Internal Medicine Foundation’s Choosing Wisely initiative and the US Preventive Services Task Force, and then narrow down this list by focusing on those services that can be readily identified using administrative claims and enrollment data. We turn to Schwartz et al. (2014) for guidance in selecting low value care measures given the comparability in structure between the Medicare data at focus in Schwartz et al. and the MHS data relevant for the present analysis. In particular, we focus on one of the key sets of measures identified by Schwartz et al.—pre-operative testing prior to low-risk surgeries—with respect to which there appears to be close to universal agreement in the literature over its low-value status (Gawande et al. 2014).

Schwartz et al. (2014) does set forth several alternative low-value care measures, including head imaging in the evaluation of syncope, EEGs for headaches, back imaging for patients with low back pain, etc. As above, we primarily focus on diagnostic testing prior to low-risk surgeries given that it appears to have the strongest consensus in the literature of low-value status among this list. There are other advantages of the low-risk surgery approach as well, beginning with large sample sizes. Moreover, and perhaps most importantly, one chief concern with these other measures is that it is possible that physician patients will present themselves to their treating physicians at different stages of their disease / condition relative to non-physician patients, creating comparability concerns that may not be fully resolved through observables. The low-risk surgery approach avoids or at least ameliorates these temporal / case-presentation concerns as we are—by design—focusing on a comparable period of time and a comparable medical scenario in each group---i.e., 30 days prior to a type of surgery that is common across both groups.

The low-risk surgeries outlined in Schwartz et al. (2014) include breast procedures, colectomy, cholecystectomy, transurethral resection of the prostate, hysterectomy, orthopedic surgical procedures other than hip and knee replacement, corneal transplant, cataract removal, retinal detachment, hernia repair, lithotripsy, arthroscopy, and cholecystectomy. As a robustness exercise, we also consider an alternative to the Schwartz et al. list and use the inpatient and outpatient MDR records to directly identify a set of surgeries with low mortality rates. To define surgeries as low-risk through this approach, we identify those surgeries with mortality rates less than 1 per 1,000 surgeries (within 30 days following the surgery/procedure) and that are performed at least 1,000 times over our records.[[4]](#footnote-4)

Within each of these low-risk surgery samples, we follow Schwartz et al. (2014) and flag the incidence of a low-value diagnostic test that was performed shortly before surgery using the incidence of any of the following: chest radiography, echocardiography, pulmonary function test (PFT), stress testing, complete blood count, coagulation panel or comprehensive metabolic panel. In each case, there is a strong medical consensus that such diagnostic testing is completely unnecessary for low-risk surgeries. We identify low-value pre-operative testing by only assessing whether these procedures were performed within 30 days of the relevant low-risk surgery. We estimate specifications that consider these tests individually (in the case of those with at least modest baseline incidence rates—i.e., chest radiography and each of the blood laboratory tests), while also estimating specifications that pool them.[[5]](#footnote-5)

High-Value Care Samples: Overview

Much of the high-value care contexts that we explore are inspired by the standardized performance measures set forth in the Healthcare Effectiveness Data and Information Set (HEDIS), measures which were created by the National Committee for Quality Assurance (NCQA) and which have been used heavily in recent years by the U.S. health care industry to evaluate the performance of health plans and systems. As with the case of low-value care, we focus on those effectiveness measures contained in HEDIS that are especially suitable for the administrative and enrollment data that we have obtained from the MHS and with respect to which the health care community has reached especially strong agreement.

High-Value Care: Diabetes

In our first high-value care analysis, we focus on a sample of patients with diabetes and follow them throughout the course of the relevant sample year to determine if they received “Comprehensive Diabetes Care” (CDC) over that year, as that term is specified by HEDIS.[[6]](#footnote-6) Using all of the available MDR data records, we identify compliance with CDC over the relevant year by receipt of all of the following by the relevant patient: (1) hemoglobin A1c (HbA1c) testing (the primary means of assessing blood sugar control), (2) retinal eye exam (since diabetes can lead to retinal damage), and (3) medical attention for nephropathy (since diabetes can lead to kidney damage). Following HEDIS, we also consider an alternative high-value care measure for the diabetes sample and assess whether the focal patient received statin therapy over the observation period (since patients with diabetes are at a higher risk for cardiovascular disease).[[7]](#footnote-7)

In order to create the diabetes sample, we determine whether MHS beneficiaries have had a diabetes diagnosis flagged in their medical records for at least two years prior to the relevant observation year. The results that we present below, however, are not sensitive to this precise two-year selection criteria.

High-Value Care: Cardiovascular Care

In our next high-value care analysis, we assess whether patients comply with the HEDIS protocols for cardiovascular care, focusing on the subsample of patients with a previous atherosclerotic cardiovascular disease diagnosis (CD sample).[[8]](#footnote-8) Within this sub-sample, we follow patients over a year-long observation period and identify high value care by observing (a) whether the affected patient received statin therapy (of high or moderate intensity) at least once over this period or, alternatively, (b) the number of days over the year in which patients filled a prescription for statin therapy, in addition to an indicator, following HEDIS guidelines, for whether they received statin therapy for at least 80% of the observation year.[[9]](#footnote-9)

High-Value Care: Medication Adherence

Inspired by drug adherence investigations from the medical literature (for example, Briesacher et al. 2008), we next endeavor to explore the degree to which patients adhere to a medication regimen during the first year in which the patient was prescribed the indicated therapy. This analysis builds on the adherence inquiries set forth above in the diabetes and cardiovascular HEDIS analyses. For this final exercise, we focus on two additional sub-samples: (1) patients with a new diagnosis for hypertension and (2) patients with a new diagnosis of hypercholesterolemia.[[10]](#footnote-10) We exclude the first sample year—2003—for these purposes to ensure that we follow patients only with a new diagnosis. In both cases, we follow the affected patients over the first year following diagnosis of this new condition. Our key outcome variable is the days supply of the relevant medication filled by the affected patient during this one-year period. In the alternative, we use the Medication Possession Ratio as the outcome variable, which equals the total number of medication supply days divided by the total number of days indicated for the therapy during that observation year. For the hypertension sample, we explore adherence of the following therapeutic classes: beta blockers, ACE inhibitors, angiotensin II receptor, calcium channel blockers, and diuretics. For the hypercholesterolemia sample, we explore adherence of the following classes: statins, antilipemic agents, and bile acid sequestrants.

High-Value Care: Vaccination / Immunization

In a final high value care analysis, we explore the extent to which children of physicians (and children of non-physician officers) receive by 2 years of age the following vaccinations (and does frequencies) recommended by HEDIS: (1) four vaccines for diphtheria, tetanus and acellular pertussis (DTaP), (2) three polio vaccines (IPV), (3) one measles, mumps and rubella (MMR), (4) three vaccines for haemophilus influenza type B (HiB), (5) one vaccine for hepatitis A, (6) three vaccines for hepatitis B, (7) one vaccine for chicken pox, (8) four vaccines for pneumococcal conjugate (PCV), and (9) two or three vaccines for rotavirus. For these purposes, we focus on a sample of children (to officer sponsors) 2 years of age that we can track all the way from birth (and otherwise subject to our general sample restrictions).

One concern with this analysis respects possible under-reporting / recording of vaccinations in the MDR database. Consider for instance compliance with recommended DTaP vaccinations. According to the NCQA, the mean rate of compliance among the Medicaid population ranges from 72 to 80 percent over our sample period. The mean in our raw sample, however, is only 58 percent. To the extent this lower rate is indeed attributable to under-reporting / recording, such recording problems would not appear to be spread uniformly throughout the MHS. Rather, it simply appears that some bases record vaccinations at very low rates—for instance as low as 30% in the case of DTaP.[[11]](#footnote-11) In light of this possible under-reporting by some bases, our primary specifications exclude low-vaccination-reporting bases. Specifically, we determine the average composition compliance rate—across all immunization types—for each base and then restrict our primary sample by dropping the bottom quartile of the sample based on the composite immunization rate of their corresponding base. In this Online Appendix, we show results without any such sample restrictions and demonstrate that the key takeaway from this analysis is robust to these sample selection choices.

Comparative High-Value and Low-Value Care Rates

*HEDIS high-value care measures*

On the National Committee for Quality Assurance’s (NCQA’s) website, the NCQA reports the compliance rates for: (1) [Comprehensive Diabetes Care](https://www.ncqa.org/hedis/measures/comprehensive-diabetes-care/), (2) [statin therapy](https://www.ncqa.org/hedis/measures/statin-therapy-for-patients-with-cardiovascular-disease-and-diabetes/) for diabetes and/or cardiovascular disease, and (3) [childhood](https://www.ncqa.org/hedis/measures/childhood-immunization-status/) vaccination. The NCQA provides rates for years going back as far as 2000 in some cases. Rates are provided for different types of insurance as well. In Table A1, we show the reported commercial HMO rates from 2005-2017.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A1  Compliance Rates among Commercial HMOs for Various HEDIS Measures | | | | | | | | | | | | | | |
| **Measure** | **Our 2013**  **Rate[[12]](#footnote-12)** | **2005** | **2006** | **2007** | **2008** | **2009** | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** |
| *Diabetes* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Eye Exam | .860 | .548 | .546 | .550 | .565 | .565 | .577 | .569 | .568 | .557 | .562 | .537 | .536 | .550 |
| HbA1c Test | .798 | .875 | .875 | .881 | .890 | .892 | .899 | .900 | .901 | .899 | .905 | .901 | .906 | .912 |
| Nephropathy Attention[[13]](#footnote-13) | .858 | .551 | .797 | .806 | .824 | .829 | .836 | .838 | .843 | .845 | .854 | .904 | .902 | .904 |
| Statin Evidence | .686 | . | . | . | . | . | . | . | . | . | . | . | .602 | .615 |
| Statin Adherence (80%) | .535 | . | . | . | . | . | . | . | . | . | . | . | .665 | .665 |
| *Cardiovascular Disease* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Statin Evidence | .670 | . | . | . | . | . | . | . | . | . | . | . | .792 | .804 |
| Statin Adherence (80%) | .611 | . | . | . | . | . | . | . | . | . | . | . | .698 | .735 |
| *Cancer Screening* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Breast Cancer Screen | .713 | .720 | .689 | .691 | .702 | .713 | .708 | .705 | .703 | .743 | .737 | .732 | .727 | .727 |
| Cervical Cancer Screen | .685 | .818 | .810 | .817 | .807 | .773 | .770 | .765 | .755 | . | .763 | .747 | .743 | .743 |
| *Childhood Immunization* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DtaP/Dt | .704 | .861 | .872 | .869 | .872 | .854 | .863 | .865 | .872 | .867 | .873 | .859 | .853 | .853 |
| IPV | .674 | .903 | .914 | .915 | .921 | .911 | .918 | .924 | .928 | .922 | .923 | .907 | .904 | .906 |
| MMR | .826 | .930 | .936 | .935 | .935 | .906 | .908 | .915 | .918 | .915 | .921 | .930 | .928 | .925 |
| HiB | .796 | .929 | .934 | .931 | .948 | .948 | .943 | .941 | .943 | .935 | .937 | .919 | .913 | .916 |
| HepB | .773 | .900 | .910 | .913 | .918 | .901 | .902 | .879 | .892 | .881 | .889 | .865 | .854 | .863 |
| VZV | .813 | .899 | .909 | .919 | .920 | .906 | .908 | .913 | .916 | .915 | .920 | .919 | .920 | .920 |
| PCV | .839 | . | .728 | .836 | .848 | .846 | .856 | .870 | .867 | .870 | .875 | .859 | .854 | .857 |
| HepA | .912 | . | . | . | . | . | .354 | .390 | .655 | .825 | .837 | .855 | .868 | .878 |
| Rotavirus | .515 | . | . | . | . | . | .635 | .751 | .767 | .799 | .808 | .802 | .805 | .818 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

There are a number of other sources that provide a source of comparison with respect to our diabetes results. The CDC reported compliance rates for diabetes treatments similar to HEDIS measures. In 2010, they found 62.8% of adults with diagnosed diabetes reported receiving a dilated-eye examination within the last year. This is higher than the rate reported by HEDIS shown in Table A1 despite the fact that the HEDIS measure includes the last two years. The same source indicates 68.5% of adults with diabetes received two or more HbA1c tests in the last year. This is lower than the rate shown in Table A1, but this difference could be due to the fact the CDC rate requires two or more tests while the HEDIS rate requires just one[[14]](#footnote-14). The U.S. Department of Veteran Affairs website shows that the 2016 diabetes eye exam rate is nearly 50% higher for VA medical centers (89%) when compared to non-VA medical centers (60%)[[15]](#footnote-15) which may suggest the MHS rate should be higher than non-MHS rates for this measure. An article looking at the TRUVen MarketScan Commercial Claims data set documents attention to nephropathy rates consistent with those reported by HEDIS[[16]](#footnote-16) and our own. Another article from *Military Medicine* shows the HbA1c test rate for diabetes patients in the MHS during 2008 is 83.4%, more than 5% lower than the rate shown in Table A1[[17]](#footnote-17), but slightly higher than our own. A letter in the *Journal of The American College of Cardiology* found approximately 62% of patients with diabetes had a statin prescription based on a sample from May 2008 to October 2013[[18]](#footnote-18); our rate is higher, but within a reasonable distance.

A study from 2013, looking at two cardiovascular disease patient trials found a statin therapy rates between 62% and 76%[[19]](#footnote-19), a range in which our rate falls.

We discuss above that vaccination rates for most vaccines are lower in our sample than those reported by HEDIS. Others have noted that childhood vaccination rates may be lower for the MHS relative to other types of insurance[[20]](#footnote-20). Another source for comparative vaccination rates is the National Immunization Survey[[21]](#footnote-21). NIS rates are fairly similar to the HEDIS rates.

*Drug adherence*

Briesacher et al. (2008) report the follow adherence rates for patients with new diagnoses for hypertension and hypercholesterolemia. The mean Medication Possession Ratio (MPR) among the sample of patients with a new diagnosis of hypertension in their study was 82.1%, slightly above that of our study. For those patients with a new diagnosis of hypercholesterolemia, Briesacher et al. (2008) do not report the exact MPR, but their graphical analysis suggests that it is between 70 and 75%, which is also slightly above that of our study. They do report that 54.6% of the hypercholesterolemia sample had a better than 80% adherence rate.

*Low-value care: Cesarean*

The average U.S. cesarean rate according to the Centers for Disease Control in 2017 was 32%.[[22]](#footnote-22) This is slightly above the mean rate found in our sample.

*Low-value care: Pre-operative testing prior to low-risk surgery*

Schwartz et al. (2014)—though inspiring our particular construction of the preoperative testing analysis—does not provide helpful comparison statistics. Central to our purpose is to treat each low-risk surgery as an event and then to compare utilization of preoperative testing between physicians and non-physicians within this event. Schwartz et al., on the other hand, reported statistics bearing on the number of instances of preoperative testing prior to low-risk surgeries per Medicare beneficiary. Other sources, however, do analyze and report preoperative testing rates out of a sample of low-risk surgeries. For instance, Kirkham et al. (2015) report that chest radiography is used 10.8% of the time prior to the sample of low-risk surgeries that they construct, which includes endoscopy and opthamology procedures (the full list of procedures they consider can be found at <http://www.cmaj.ca/content/suppl/2015/06/01/cmaj.150174.DC1>). Though their list differs slightly from the Schwartz et al. list and from our derived list using MHS data records, the incidence of preoperative chest radiography is comparable among these samples.

Preoperative Testing: Separate Visit Analysis

Surgeons frequently recommend diagnostic testing—including both blood testing and radiological testing—prior to low-risk surgeries. In our paper, we explore whether superior patient information contributes to lower rates of utilization of this arguably unnecessary testing. We find that information may contribute to only a minor reduction in the use of preoperative testing prior to low-risk surgeries. A concern arises however that this modest finding may be in part due to the low cost-sharing environment in which MHS beneficiaries exist—e.g., a patient may know that they are unnecessary but since there is little cost, they may get the test anyway. One response to this concern may be that the costs of preoperative testing go beyond the financial costs of the tests themselves. Importantly, the testing may require a separate trip on a separate day to a medical office or laboratory in order to receive the test, which may require additional child-care expenses and other expenses, not to mention general inconvenience. In this section of the Online Appendix, we explore the likelihood that preoperative tests indeed require separate trips to medical offices, as distinct from simply combining such tests with the initial surgery consultation.

To begin, we note that, in their investigation into propriety of preoperative testing, Benarrock-Gample et al. (2012) find that the majority of tests occur within two weeks of surgery. Though they do not report the average time span between the date in which surgery was recommended / scheduled and the date of the surgery, one might expect that, on average, surgeries are planned more than two weeks in advance, in which case, this statistic alone might tend to suggest that the tests are being performed on separate days.

To explore matters further, we turn to our data. Specifically, we turn to cataract surgery as a case study—one of the most common low-risk surgeries targeted by the medical literature on over-testing prior to low-risk surgeries (and a surgery that is especially low risk and high in volume). Further, we focus on cataract surgeries performed on the base at Military Treatment Facilities.

Our key sample for this analysis is all cataract surgeries performed on the base, where we attempt to collect all claims for the cataract surgery patients in our sample around the time of the surgery. For each cataract surgery, we create a 121-day panel (90 days prior to surgery, day of surgery, and 30 days after) which will enable us to see trends in care leading up to the cataract surgery. 1 The panel consists of 55,293 cataract surgeries.

We focus on direct care because of the presence of a variable corresponding to *appointment type*: (1) routine appointment, (2) acute appointment, (3) telephone consult, (4) initial specialist appointment, (5) established follow-up, (6) ambulatory procedure appointment, (7) procedure appointment and (8) other appointment type. The variable seems less useful than we had hoped due to treating physicians possibly using different indicators for the same type of appointment. That is, some may register a yearly eye appointment as an established followup and some as a routine appointment. Some may record cataract surgery as an ambulatory procedure, while some may record it as a general procedure appointment. At the outset, we note that it is difficult to say for certain which appointment is the scheduling appointment. It may not necessarily be the “initial specialist appointment” considering that the surgery recommendation may have come after a series of previous appointments addressing the cataracts medical condition. Perhaps more likely this occurs during one of the “established follow-up” visits.

Nevertheless, despite these concerns, we plot frequency distributions for the timing of each of the appointment types in Figures A1 and A2. There are some informative trends that can be observed. In Figure A1, we can see the ambiguity in ambulatory procedure and procedure appointment types mentioned previously—there is a spike in both types on the day of the cataract surgery. The increase in the number of established followup appointments and in procedure appointments leading up to cataract surgery is possibly related to preoperative care associated with the surgery or possibly reflective of the scheduling visit for the surgery in the case of surgeries executed very shortly after scheduling. As discussed further below, it is tough to say for sure based on these trends exactly when the scheduling visit occurs. The established followup category is also associated with postoperative appointments. In Figure A2 similar trends are found for other types of appointments. There is one other important thing to note about Figure A2. There is some baseline level of appointments not associated with cataract surgeries, which we observe being distributed uniformly over this cataract-surgery window.

In Figures A3 and A4, we plot the frequency distribution over this 121-day time window for preoperative testing events, separately across different types of tests. These figures clearly evidence a strong upward trend in preoperative tests in the time leading up to cataract surgery, with the increase beginning approximately one month prior to the surgery—the time period in which surgeons begin recommending that preoperative testing take place. There are notable spikes especially one week and one day prior to cataract surgery.

Do these testing visits truly represent separate visits? There may be very good reason to suspect that many of these are indeed arising from separate visits given the fact that the distribution of visits for these tests is more concentrated in the month prior to the cataract surgery relative to the distribution of all other visit types and given the fact that the ramp-up in utilization over this month is steeper for the diagnostic testing visits relative to the other visit types from Figures A1 and A2.

The above figures plot distributions for all medical visits for the cataract patients in the window of the cataract surgery. Another approach is to specially focus on eye-appointment visits during this window in which cataracts are indicated in the diagnosis field. With this in mind, a potential proxy for the scheduling of the cataract surgery may be the earliest eye appointment in the 90 days leading up to the surgery. For this analysis, we drop cataract surgeries where this first diagnosis is the day of the surgery (approx. 2000 cases) leaving 50,188 surgeries (for these, it is possible that the true scheduling appointment occured longer than 90 days prior). We additionally drop all appointments which are not associated with an eye specialist. On average during this period, patients have 3.88 appointments with eye care specialists and 1.78 eye exams. We then sort appointments into three categories: day surgery is scheduled, days between scheduling day and surgery, and day of surgery. From here, we show in Figures A5 and A6 how many preoperative tests occur during these appointment types. The number of tests is by far the highest for the period between the initial diagnosis and the surgery. These results are consistent with those presented in Figures A1-A4.

In conclusion, despite challenges identifying scheduling date, this analysis likely demonstrates that in the case of cataract surgery, most preoperative care is not occurring during the day the surgery is scheduled and is thus likely necessitating a separate visit.

Figure A1

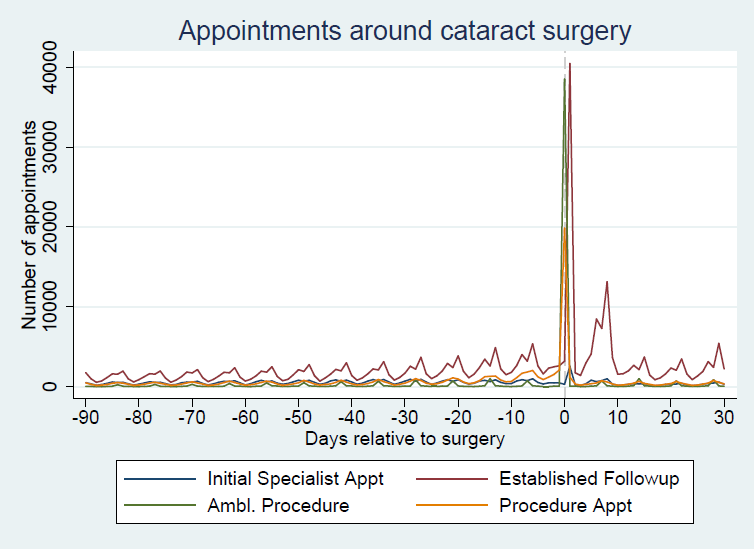


Figure A2

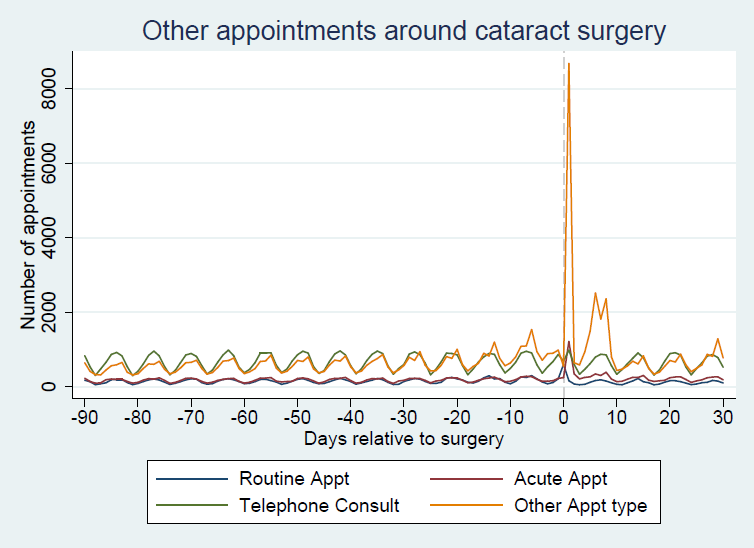


Figure A3

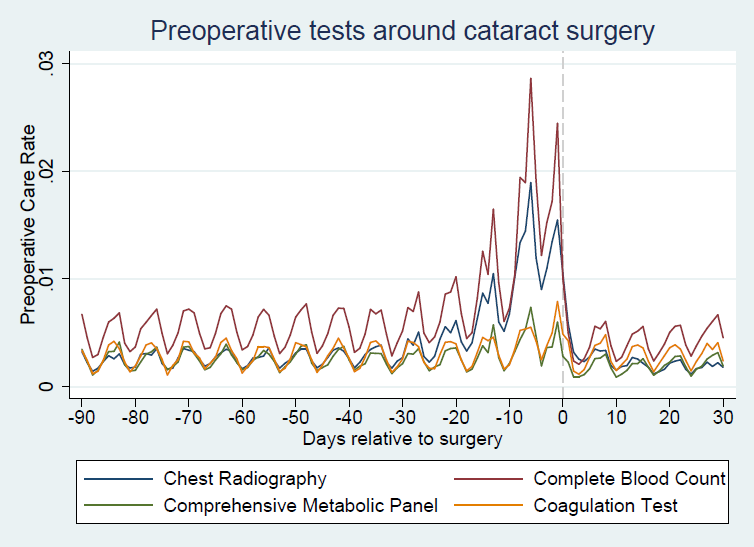


Figure A4

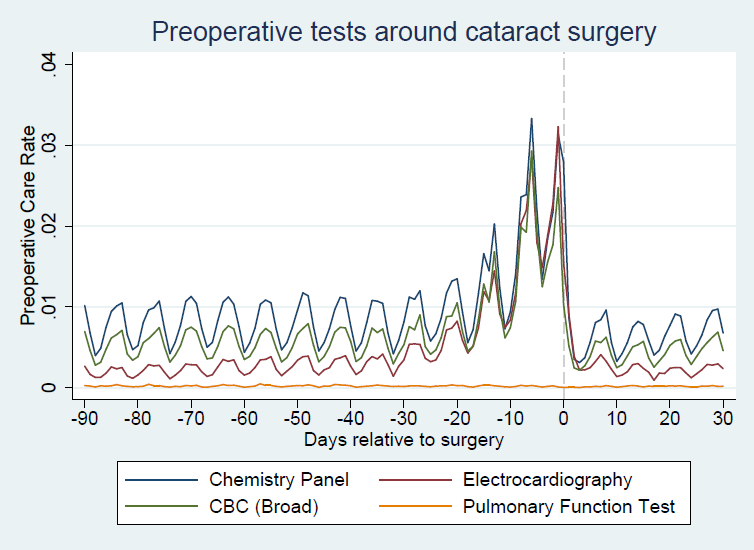


Figure A5

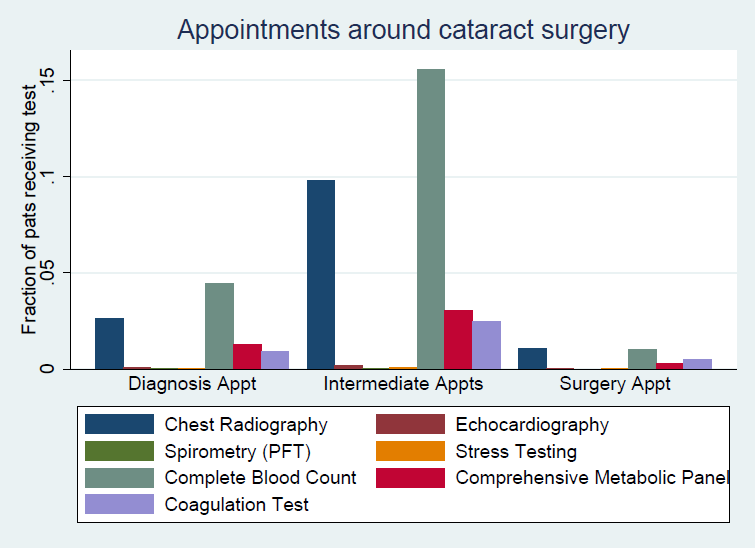


Figure A6

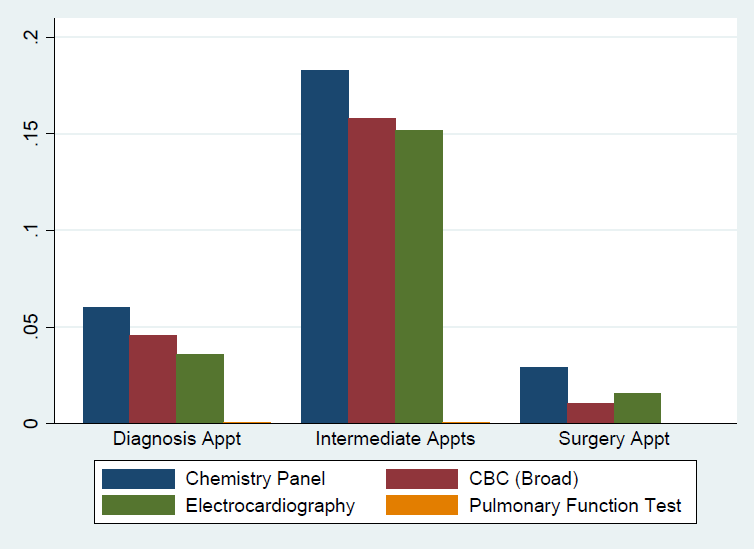


Table A2

Sample means and Uncontrolled Differences between Physician and Non-Physician Patients, by Relevant Sample

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Physician Mean | Non-Physician Mean | Difference | Difference (Standard Error) |  |
| **Panel A. Delivery Sample, Sponsors Analysis** | | | | | | |
| Cesarean Incidence (within all deliveries) | | 0.2852 | 0.3138 | -0.0286\*\* | 0.0118 |  |
| Primary Cesarean Incidence (among  restricted delivery sample) | | 0.1485 | 0.1897 | -0.0412\*\*\* | 0.0095 |  |
| **Panel B. Delivery Sample, Dependents Analysis** | | | | | | |
| Cesarean Incidence (within all deliveries) | | 0.2520 | 0.2727 | -0.0208\*\*\* | 0.0073 |  |
| Primary Cesarean Incidence (among  restricted delivery sample) | | 0.1077 | 0.1261 | -0.0184\*\*\* | 0.0052 |  |
| **Panel C. Low-Risk Surgery Sample (Schwartz et al.), Sponsors Analysis** | | | | | |  |
| Chest Radiography | | 0.0511 | 0.1030 | -0.0519\*\*\* | 0.0036 |  |
| Complete Blood Count | | 0.3132 | 0.3135 | 0.0007 | 0.0082 |  |
| Comprehensive Metabolic Panel | | 0.0631 | 0.0996 | -0.0366\*\*\* | 0.0049 |  |
| Coagulation Panel | | 0.0791 | 0.1056 | -0.0265\*\*\* | 0.0044 |  |
| Any Preop Care | | 0.3384 | 0.3602 | -0.0218\*\*\* | 0.0087 |  |
| **Panel D. Low-Risk Surgery Sample (Schwartz et al.), Dependents Analysis** | | | | | |  |
| Chest Radiography | | 0.0345 | 0.0705 | -0.0360\*\*\* | 0.0000 |  |
| Complete Blood Count | | 0.3869 | 0.3521 | 0.0349\*\*\* | 0.0015 |  |
| Comprehensive Metabolic Panel | | 0.0586 | 0.0937 | -0.0351\*\*\* | 0.0049 |  |
| Coagulation Panel | | 0.0578 | 0.0707 | -0.0129\*\*\* | 0.0010 |  |
| Any Preop Care | | 0.4038 | 0.3866 | 0.0171\*\* | 0.0083 |  |
| **Panel E. Low-Risk Surgery Sample (Derived Low-Risk), Sponsors Analysis** | | | | | |  |
| Chest Radiography | | 0.0459 | 0.0456 | 0.0003 | 0.0012 |  |
| Complete Blood Count | | 0.1458 | 0.1369 | 0.0001 | 0.0026 |  |
| Comprehensive Metabolic Panel | | 0.0455 | 0.0592 | 0.0089\*\*\* | 0.0022 |  |
| Coagulation Panel | | 0.0224 | 0.0241 | -0.0017\* | 0.0009 |  |
| Any Preop Care | | 0.1779 | 0.1761 | 0.0017 | 0.0024 |  |
| **Panel F. Low-Risk Surgery Sample (Derived Low-Risk), Dependents Analysis** | | | | | |  |
| Chest Radiography | | 0.0356 | 0.0444 | -0.087\*\*\* | 0.0018 |  |
| Complete Blood Count | | 0.1778 | 0.1777 | 0.0009 | 0.0041 |  |
| Comprehensive Metabolic Panel | | 0.0594 | 0.0849 | -0.0256\*\*\* | 0.0025 |  |
| Coagulation Panel | | 0.0350 | 0.0365 | -0.0015 | 0.0020 |  |
| Any Preop Care | | 0.2067 | 0.2201 | -0.0134\*\*\* | 0.0045 |  |
| **Panel G. HEDIS Diabetes Sample, Sponsors Analysis** | | | | |  |  |
| hbA1c Testing |  | 0.7406 | 0.7635 | -0.0228 | 0.0191 |  |
| Eye Exam |  | 0.8493 | 0.7845 | 0.0648\*\*\* | 0.0195 |  |
| Attention to Nephropathy | | 0.8685 | 0.8548 | 0.0137 | 0.0155 |  |
| Comprehensive Diabetes Care | | 0.6219 | 0.5771 | 0.0448\* | 0.0230 |  |
| Statin Therapy |  | 0.7368 | 0.7079 | 0.0288 | 0.0334 |  |
| **Panel H. HEDIS Diabetes Sample, Dependents Analysis** | | | | |  |  |
| hbA1c Testing |  | 0.6878 | 0.7320 | -0.0442\*\* | 0.0196 |  |
| Eye Exam |  | 0.8059 | 0.7858 | 0.0201 | 0.0173 |  |
| Attention to Nephropathy | | 0.7701 | 0.8055 | -0.0354\*\* | 0.0174 |  |
| Comprehensive Diabetes Care | | 0.5239 | 0.5383 | -0.0144 | 0.0208 |  |
| Statin Therapy |  | 0.6337 | 0.6116 | 0.0221 | 0.0371 |  |
| **Panel I. HEDIS Cardiovascular Disease Sample, Sponsors Analysis** | | | | | |  |
| Any Statin Therapy | | 0.7781 | 0.7561 | 0.0221 | 0.0232 |  |
| Days Supplied of Statin Therapy | | 224.1438 | 225.4716 | -1.3278 | 7.9894 |  |
| Incidence of at least 80% Adherence | | 0.4289 | 0.4674 | -0.0384\* | 0.0231 |  |
| **Panel J. HEDIS Cardiovascular Disease Sample, Dependents Analysis** | | | | | | |
| Any Statin Therapy | | 0.6103 | 0.5808 | 0.0295 | 0.0460 |  |
| Days Supplied of Statin Therapy | | 171.5984 | 163.0764 | 8.5220 | 14.4706 |  |
| Incidence of at least 80% Adherence | | 0.3241 | 0.3193 | 0.0048 | 0.0351 |  |
| **Panel K. HEDIS Childhood Immunization Sample** | | | | |  |  |
| DTaP Immunization Compliance | | 0.66 | 0.67 | -0.01 | 0.01 |  |
| IPV Immunization Compliance | | 0.65 | 0.64 | 0.02 | 0.01 |  |
| MMR Immunization Compliance | | 0.82 | 0.80 | 0.02\*\* | 0.01 |  |
| HiB Immunization Compliance | | 0.71 | 0.69 | 0.02\* | 0.01 |  |
| Hepatitis A Immunization Compliance | | 0.87 | 0.85 | 0.02\*\* | 0.01 |  |
| Hepatitis B Immunization Compliance | | 0.76 | 0.73 | 0.03\*\*\* | 0.01 |  |
| Chicken Pox Immunization Compliance | | 0.81 | 0.79 | 0.02\*\* | 0.01 |  |
| PCV Immunization Compliance | | 0.81 | 0.79 | 0.02 | 0.01 |  |
| Rotavirus Immunization Compliance | | 0.34 | 0.32 | 0.02\* | 0.01 |  |
| **Panel L. First-Year Prescription Adherence (Medical Possession Ratio) Sample, Sponsor Analysis** | | | | | | |
| Hypertension | | 0.7296 | 0.7228 | 0.0069 | 0.0090 |  |
| Hypercholesterolemia | | 0.6239 | 0.6428 | -0.0189\*\* | 0.0083 |  |
| **Panel M. First-Year Prescription Adherence (Medical Possession Ratio) Sample, Dependent Analysis** | | | | | | |
| Hypertension | | 0.6887 | 0.6867 | 0.0020 | 0.0129 |  |
| Hypercholesterolemia | | 0.6512 | 0.6378 | 0.0135 | 0.0116 |  |
| *Notes*: standard errors are reported in Column 4 and clustered at the individual beneficiary level. Data are from the Military Health System Data Repository, 2003-2013. \*\*\* Significant at the 1 percent level; \*\* Significant at the 5 percent level; \* Significant at the 10 percent level. | | | | | | |

Covariate Balance

In Table 1 of the text, we presented a covariate balance analysis focused on predicted outcome measures, where predictions were formed from regressing the various outcome measures on the set of covariates. In this Online Appendix, we supplement that analysis by showing covariate balance across individual covariates. Given the large number of samples used in our analysis, we do not do this exercise across all samples. Though, we do so on a select few.

Table A3

Covariate Balance. Low-Value Care (Schwartz) Sample

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Physicians (as patients) | Officers | Difference | Dependents of Physicians | Dependents of Officers | Difference |
| Age | 39.49  (9.69) | 44.75  13.02) | -5.26\*\*\*  (0.22) | 31.81  (11.51) | 35.88  (15.57) | -4.07\*\*\*  (0.235) |
| White | 0.78  (0.41) | 0.75  (0.44) | 0.04\*\*\*  (0.00) | 0.81  (0.39) | 0.79  (0.41) | 0.02\*\*\*  (0.00) |
| Black | 0.06  (0.22) | 0.10  (0.30) | -0.04\*\*\*  (0.00) | 0.03  (0.17) | 0.07  (0.25) | -0.03\*\*\*  (0.00) |
| Male | 0.48  (0.49) | 0.69  (0.46) | -0.21\*\*\*  (0.01) | 0.10  (0.30) | 0.08  (0.28) | 0.02\*\*\*  (0.00) |
| Pay Grade (Junior Officer) | 0.33  (0.47) | 0.37  (0.48) | -0.05\*\*\*  (0.01) | 0.39  (0.49) | 0.44  (0.50) | -0.06\*\*\*  (0.01) |
| Previous Year’s RVU | 266.63  (291.38) | 297.61  (341.97) | -30.98\*\*\*  (5.87) | 221.93  (268.44) | 258.53  (293.58) | -36.60\*\*\*  (4.46) |
| Previous Year’s Inpatient Days | 1.42  (9.30) | 2.04  (14.41) | -0.62\*\*  (0.24) | 2.13  (13.08) | 2.15  (14.78) | -0.01  (0.22) |
| Charlson Comordity Index | 0.22  (0.74) | 0.45  (1.24) | -0.23\*\*\*  (0.02) | 0.20  (0.71) | 0.39  (1.10) | -0.20\*\*\*  (0.02) |
| *Notes*: standard deviations are reported in parentheses (Columns 1, 2, 4 and 5). Standard errors in the differencing columns (Columns 3 and 6) are clustered at the individual beneficiary level. Data are from the Military Health System Data Repository, 2003-2013. \*\*\* Significant at the 1 percent level; \*\* Significant at the 5 percent level; \* Significant at the 10 percent level. | | | | | | |

Table A4

Covariate Balance. Diabetes Sample

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Physicians (as patients) | Officers | Difference | Dependents of Physicians | Dependents of Officers | Difference |
| Age | 54.76  (8.67) | 59.37  (8.05) | -4.61  (0.24) | 50.72  (10.67) | 55.56  (10.36) | -4.83\*\*\*  (0.29) |
| White | 0.66  (0.47) | 0.68  (0.46) | -0.02\*  (0.01) | 0.75  (0.44) | 0.67  (0.47) | 0.07\*\*\*  (0.13) |
| Black | 0.12  (0.33) | 0.12  (0.32) | 0.01  (0.01) | 0.07  (0.25) | 0.10  (0.30) | -0.03\*\*\*  (0.01) |
| Male | 0.81  (0.39) | 0.93  (0.26) | -0.13\*\*\*  (0.01) | 0.17  (0.37) | 0.07  (0.25) | 0.10\*\*\*  (0.01) |
| Pay Grade (Junior Officer) | 0.03  (0.16) | 0.15  (0.36) | -0.12\*\*\*  (0.01) | 0.08  (0.27) | 0.18  (0.38) | -0.10\*\*\*  (0.01) |
| Previous Year’s RVU | 57.52  (58.76) | 53.15  (52.34) | 4.36\*\*\*  (1.60) | 58.17  (54.46) | 61.01  (55.94) | -2.83\*  (1.54) |
| Previous Year’s Inpatient Days | 0.71  (10.60) | 0.62  (5.19) | 0.10  (0.16) | 0.61  (3.07) | 0.84  (6.90) | -0.23  (0.19) |
| Charlson Comordity Index | 1.34  (1.10) | 1.68  (1.42) | -0.05  (0.04) | 1.30  (1.31) | 1.48  (1.39) | -0.18\*\*\*  (0.04) |
| Diabetes Complication Severity Index | 0.52  (0.93) | 0.79  (1.24) | -0.33\*\*\*  (0.04) | 0.51  (1.05) | 0.65  (1.13) | -0.14\*\*\*  (0.03) |
| *Notes*: standard deviations are reported in parentheses (Columns 1, 2, 4 and 5). Standard errors in the differencing columns (Columns 3 and 6) are clustered at the individual beneficiary level. Data are from the Military Health System Data Repository, 2003-2013. \*\*\* Significant at the 1 percent level; \*\* Significant at the 5 percent level; \* Significant at the 10 percent level. | | | | | | |

Table A5

Covariate Balance. Cesarean Sample

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Physicians (as patients) | Officers | Difference | Dependents of Physicians | Dependents of Officers | Difference |
| Age | 33.09  (3.52) | 32.38  (4.53) | 0.71\*\*\*  (0.13) | 32.33  (3.97) | 30.53  (4.86) | 1.78\*\*\*  (0.07) |
| White | 0.67  (0.47) | 0.66  (0.47) | 0.01  (0.01) | 0.84  (0.36) | 0.78  (0.41) | -0.06\*\*\*  (0.01) |
| Black | 0.07  (0.26) | 0.17  (0.37) | -0.09\*\*\*  (0.01) | 0.02  (0.15) | 0.06  (0.24) | -0.04\*\*\*  (0.004) |
| Pay Grade (Junior Officer) | 0.50  (0.50) | 0.71  (0.45) | -0.21\*\*\*  (0.01) | 0.52  (0.50) | 0.68  (0.47) | -0.16\*\*\*  (0.01) |
| Previous Year’s RVU | 49.80  (39.69) | 46.97  (39.68) | 2.83\*\*  (1.15) | 41.32  (35.76) | 37.02  (33.74) | 4.30\*\*\*  (0.52) |
| Previous Year’s Inpatient Days | 0.17  (1.79) | 0.24  (2.73) | -0.06  (0.07) | 0.16  (1.94) | 0.16  (1.87) | -0.00  (0.03) |
| Charlson Comordity Index | 0.06  (0.25) | 0.05  (0.28) | 0.01  (0.01) | 0.06  (0.26) | 0.07  (0.29) | -0.01\*\*  (0.00) |
| *Notes*: standard deviations are reported in parentheses (Columns 1, 2, 4 and 5). Standard errors in the differencing columns (Columns 3 and 6) are clustered at the individual beneficiary level. Data are from the Military Health System Data Repository, 2003-2013. \*\*\* Significant at the 1 percent level; \*\* Significant at the 5 percent level; \* Significant at the 10 percent level. | | | | | | |
|  |  |  |  |  |  |  |

Table A6

Covariate Balance. Cardiovascular Disease Sample

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Physicians (as patients) | Officers | Difference | Dependents of Physicians | Dependents of Officers | Difference |
| Age | 57.41  (7.77) | 61.59  (7.93) | -4.18\*\*\*  (0.23) | 56.77  (7.10) | 61.20  (8.19) | -4.42\*\*\*  (0.37) |
| White | 0.81  (0.39) | 0.72  (0.45) | 0.10\*\*\*  (0.01) | 79.20  (40.62) | 60.20  (48.95) | 0.19\*\*\*  (0.02) |
| Black | 0.06  (0.24) | 0.05  (0.22) | 0.01  (0.01) | 0.08  (0.27) | 0.06  (0.24) | 0.02  (0.01) |
| Male | 0.92  (0.27) | 0.96  (0.18) | -0.05\*\*\*  (0.01) | 0.33  (0.47) | 0.09  (0.28) | 0.24\*\*\*  (0.01) |
| Pay Grade (Junior Officer) | 0.01  (0.11) | 0.11  (0.32) | -0.10\*\*\*  (0.01) | 0.01  (0.12) | 0.12  (0.33) | -0.11\*\*\*  (0.01) |
| Previous Year’s RVU | 64.02  (59.77) | 66.52  (58.77) | -2.50  (1.70) | 74.56  (60.08) | 80.94  (63.07) | -6.38\*\*  (2.90) |
| Previous Year’s Inpatient Days | 1.18  (10.88) | 1.28  (7.98) | -0.09  (0.23) | 1.63  (9.36) | 2.21  (14.47) | -0.58  (0.66) |
| Charlson Comordity Index | 0.71  (1.15) | 1.27  (1.69) | -0.55\*\*\*  (0.05) | 1.31  (1.75) | 1.49  (1.75) | -0.19\*\*  (0.08) |
| *Notes*: standard deviations are reported in parentheses (Columns 1, 2, 4 and 5). Standard errors in the differencing columns (Columns 3 and 6) are clustered at the individual beneficiary level. Data are from the Military Health System Data Repository, 2003-2013. \*\*\* Significant at the 1 percent level; \*\* Significant at the 5 percent level; \* Significant at the 10 percent level. | | | | | | | - |
|  | | | | | | |  |

Table A7

Estimated Physician Effect on Childhood Immunization Rates, Including all Bases

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | N | Physicians or Physician-years | Baseline Rate | Coefficient | Std. Err. |
| DTaP |  | 45955 | 2887 | 0.5760 | -0.0136 | 0.0093 |
| IPV |  | 45955 | 2887 | 0.5546 | 0.0064 | 0.0095 |
| MMR |  | 45955 | 2887 | 0.7010 | 0.0128 | 0.0082 |
| HiB |  | 45955 | 2887 | 0.6051 | 0.0094 | 0.0088 |
| Hepatitis A |  | 45955 | 2887 | 0.7893 | 0.0060 | 0.0075 |
| Hepatitis B |  | 45955 | 2887 | 0.6366 | 0.0073 | 0.0081 |
| Chicken Pox |  | 45955 | 2887 | 0.6904 | 0.0174\*\* | 0.0083 |
| PCV |  | 45955 | 2887 | 0.6967 | 0.0046 | 0.0081 |
| Rotavirus |  | 45955 | 2887 | 0.2836 | 0.0168\*\* | 0.0082 |
| *Notes*: standard errors are reported in parentheses and clustered at the individual beneficiary level. Each row represents results from a different specification using the specified dependent variable and the specified sample. All specifications control for patient age-by-sex dummies, patient race dummies, patient pay-grade dummies, previous year RVU, previous year RWP, previous year inpatient days, Charlson combordity index, and base-by-year-by-care-location fixed effects. Data are from the Military Health System Data Repository, 2003-2013. \*\*\* Significant at the 1 percent level; \*\* Significant at the 5 percent level; \* Significant at the 10 percent level. | | | | | | |

1. The WHO (2015) report further summarizes the relevant medical literature on cesareans. [↑](#footnote-ref-1)
2. We will take this approach throughout—that is, for the dependents analysis where there is age overlap, we will not impose age restrictions. In all of our contexts, however, starting with the cesarean context, the dependents results that we will present are nearly identical when also using the same age restrictions used in the sponsors analysis. [↑](#footnote-ref-2)
3. For more information on AHRQ’s formulation of primary cesarean rates, see <https://www.qualityindicators.ahrq.gov/Downloads/Modules/IQI/> V41/TechSpecs/IQI%2033%20Primary%20Cesarean%20Delivery%20Rate.pdf. [↑](#footnote-ref-3)
4. Surgeries and procedures in this alternative list include, among others, those with the following ICD descriptions: (1) “eye procedure – cataract removal/lens insertion,” (2) “eye procedure – retinal detachment,” (3) “major procedure –breast,” (4) “major procedure – explor/decompr/excis disc”, (5) “major procedure – other,” (6) “orthopedic – other,” (7) “ambulatory procedure – skin,” (8) “ambulatory procedure – musculoskeletal,” and (9) “ambulatory procedure – ground hernia repair.” [↑](#footnote-ref-4)
5. When comparing the use of preoperative testing between physician patients and non-physician officer patients, we limit our analysis to those between the ages of 25 and 75, considering that virtually no low-risk surgeries are performed on physician sponsors younger than 25 and considering the general risks associated with surgeries at older ages. As with the cesarean analysis, in the case of dependents, there is overlap between the physician and non-physician group at younger ages, in which event, we make no age restrictions at younger ages. [↑](#footnote-ref-5)
6. Consistent with the HEDIS protocol, we focus on patients between the ages of 18 and 75; however, for the sponsors analysis, we focus on patients between 27 and 75 to ensure overlap between the physician group and non-physician group as there are no diabetic physician patients in the sample at younger ages. [↑](#footnote-ref-6)
7. For statins, following HEDIS guidance, we focus on patients over 40 years of age and exclude patients that have received coronary artery bypass grafting, percutaneous coronary intervention or in-vitro fertilization or that have suffered from the following conditions: end-stage renal disease, myocardial infarction, ischemic vascular disease, pregnancy, cirrhosis, myalgia, myositis or rhabdemyolysis. [↑](#footnote-ref-7)
8. As with the diabetes sample, we determine inclusion in the CD sample by whether MHS beneficiaries have had a cardiovascular disease diagnosis flagged in their medical records for at least two years prior to the relevant observation year. [↑](#footnote-ref-8)
9. For this CD analysis, HEDIS recommends focusing on male patients between the ages of 21 and 75 and female patients between the ages 40 and 75. We follow these recommended ages fully for the dependents sample. For the male sponsors comparison, however, we focus on ages 26 to 75 given that there are no physician sponsors below the age of 26. [↑](#footnote-ref-9)
10. Briesacher et al. consider other new diagnoses that are not suitable candidates for our investigation. For instance, osteoporosis, seizure disorders, and gout suffer from sample size concerns. They also recommend looking at adherence to alpha-glucosidase inhibitors/ thiazolidinediones for new diagnoses of diabetes. However, as discussed above, for our diabetes analyses, we construct the relevant sample by focusing on those that have had been receiving care for diabetes for at least two years. We do this in order to better ensure comparability in the degree of diabetes across our comparison groups and avoid picking up instances of mild diabetes cases (where physicians may be or more less likely to seek treatment for mild case). For this reason, we elect not to focus on first year adherence to diabetes medications. [↑](#footnote-ref-10)
11. While vaccination rates do vary by regions (as reported by the Centers for Disease Control: <https://stacks.cdc.gov/view/cdc/59415>), this variation is generally not known to be as a wide-ranging as that reported across bases, suggesting indeed that some bases are perhaps not recording all vaccination events in the MDR records. [↑](#footnote-ref-11)
12. Based on the July samples and recent vaccination samples. [↑](#footnote-ref-12)
13. Nephropathy Attention compliance is defined as receiving a microalbumin exam, receiving ACE/ARB therapy, or receiving treatment for nephropathy in last measurement year. [↑](#footnote-ref-13)
14. <https://www.cdc.gov/diabetes/statistics/preventive/fallpractices.htm> [↑](#footnote-ref-14)
15. <https://www.va.gov/QUALITYOFCARE/initiatives/compare/diabetes-care-eye-exams.asp> [↑](#footnote-ref-15)
16. <https://www.managedcaremag.com/system/files/storypdfs/mc_1808_or_overstated_rates.pdf> [↑](#footnote-ref-16)
17. <https://academic.oup.com/milmed/article/178/2/142/4210873> [↑](#footnote-ref-17)
18. <http://www.onlinejacc.org/content/accj/68/12/1368.full.pdf> [↑](#footnote-ref-18)
19. <https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.112.000712> [↑](#footnote-ref-19)
20. <http://pediatrics.aappublications.org/content/135/5/e1148> [↑](#footnote-ref-20)
21. <https://www.cdc.gov/vaccines/imz-managers/nis/datasets.html> [↑](#footnote-ref-21)
22. <https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_08-508.pdf> [↑](#footnote-ref-22)