Doing More with Less: Predicting Primary Care Provider Effectiveness

Janet Currie, Princeton University

Jonathan Zhang,* McMaster University and Veterans Health Administration

September 2022

Abstract

We use data from the Veterans Administration to examine the efficacy of primary care providers (PCPs). Leveraging quasi-random assignment of veterans to PCPs, we measure effectiveness using ambulatory care sensitive conditions (ACSC) and hospitalizations/emergency department (ED) visits for mental health or circulatory conditions. PCPs variation along these dimensions predicts future outcomes. For example, a one standard deviation improvement in mental health effectiveness reduces patient risk of death by 3.8% and lowers costs by 4.4% over the next three years. More effective PCPs do more with less: their patients have fewer primary care visits, specialist referrals, lab panels, or imaging tests. JEL codes: I10; J24; I30

Critics of the U.S. health care system argue that it provides too much high-cost/low-value care, and not enough low-cost/high-value care (Chandra and Skinner, 2012). It has been suggested that providers should be compensated on the basis of the value, rather than the quantity, of care they provide, where high-value care is care that yields better health outcomes on average (Cutler, 2014). These arguments beg several questions: Are some providers more effective than others in promoting patient health and how can we measure that? Do patients whose providers are effective in one domain do better in other domains as well? And if some providers are generally more effective than others, what characteristics of providers predict effectiveness?

This paper investigates these questions using the unique setting of primary care providers (PCPs) in the Veterans Health Administration (VHA). The most important aspect of this setting is that veterans who enter the system seeking primary care are assigned to PCPs in a quasi-random, first-come, first-served basis which depends only on the patient's desired appointment time, location, and the PCP's availability. A second advantage is that the VHA was a pioneer in the use of electronic medical records so that we have detailed records of inpatient, outpatient, and pharmaceutical claims including rich information about referrals, screenings, tests and labs which allow us to investigate possible reasons for variations in provider effectiveness. A third advantage is that providers in the VHA system are salaried, so they have no financial incentive to provide low-value care, such as excessive screening.

Using data from 802,777 veterans assigned to 7,548 PCPs at 725 clinics we ask whether PCP assignment is predictive of three important future patient health outcomes: Hospitalizations and emergency department (ED) visits for mental health; hospitalizations and ED visits for circulatory problems, and hospitalizations for ambulatory care sensitive conditions (ACSC). We chose the first two measures because they are two of the most common types of serious health problems seen in the VHA. The third measure, ACSC, captures outcomes due to a broad range of

conditions that are amenable to primary care. Given quasi-random assignment, we characterize physicians who have better patient outcomes (leaving out the index patient) as more effective.

Past research provides considerable evidence of variations in provider effectiveness, beginning with the literature on geographical variations in care (e.g., Cutler et al., 2019; Finkelstein, Gentzkow, and Williams, 2016; Fisher et al., 2003a,b); continuing with studies of quasi-random assignment in ambulance referrals and emergency departments (e.g., Doyle et al, 2011, 2015, forthcoming; Gowrisankaran, Joiner, and Leger, 2017; Van Parys, 2016); and including attempts to quantify physician practice style and link it with patient outcomes (e.g., Abaluck et al., 2020; Currie and MacLeod 2016, 2020; Currie, MacLeod, and Van Parys, 2016; Epstein and Nicholson, 2009; Fadlon and Van Parys, 2020; Fletcher, Horwitz, and Bradley, 2014; Grytten and Sorensen, 2003, Kwok, 2019; Simeonova, Skipper, and Thingholm, 2020).

Consistent with these studies we find a significant range in our effectiveness measures across PCPs, and we find that patient outcomes differ significantly depending on the physician they are assigned to. For example, a one standard deviation improvement in our measure of mental health effectiveness predicts a 0.21 percentage point (3.8%) lower risk of patient death over the next three years and 4.4% lower costs.

Turning to the two more novel questions that we ask, we find that patients of doctors whom we judge to be more effective in one domain, tend to also be more effective in others. For example, doctors who are effective at preventing hospitalizations due to ambulatory care sensitive conditions are also more effective in preventing deaths from cancer, heart conditions, and possible suicides (external causes of death measured by suicides plus overdoses, poisonings, and accidents). The one exception to this generalization is that only mental health effectiveness predicts fewer patient visits for mental health. These results suggest that it may not be necessary to measure effectiveness in every possible dimension in order to identify more effective physicians.

Our most novel conclusion is that the most effective PCPs do more with less: Their patients have fewer primary care visits, referrals to specialists, lab panels or imaging tests. Effective PCPs are slightly more likely to comply with guidelines for mental health screenings, and slightly less likely to comply with guidelines for physical health screenings, but these differences in screening propensities are negligible in magnitude suggesting that adherence to screening guidelines is not the main determinant of differences in outcomes.

We also find that older PCPs, those who see more patients per day, and those who see more new patients over the period we observe them tend to be more effective. PCPs in some facilities at the VHA have the option to call in mental health professionals for immediate same-day patient consultations that are joint with the PCP rather than referring them for later appointments. Physicians who take advantage of this option for care coordination (conditional on its availability) also tend to be more effective.

Conditional on these measures, part-time physicians are more effective, which leads us to interpret part-time status as a marker for physicians who devote some of their time to research. We also find some evidence that nurse practitioners/physician assistants are more effective primary care providers than physicians in the VHA on average though this result could possibly reflect the type of nurse practitioner who is selected to be a PCP. All our results hold if PCPs who are nurse practitioners are excluded from the sample.

A few previous studies have shown results with a similar flavor to ours: Chan, Gentzkow, and Yu (2019) find that radiologists who are less skilled at diagnosing pneumonia compensate by treating marginal patients more aggressively. Currie and MacLeod (2016) find that obstetricians with better diagnostic skills perform fewer C-sections on low-risk women and have better patient outcomes. In contrast to these two studies focusing on specialist's use of particular procedures for specific conditions, we construct broader measures of effectiveness and consider a wide range of health inputs and outputs.

Our work is also related to Doyle, Ewer, and Wagner (2010) who compare physicians from two different medical schools who were employed at the same Veterans Affairs hospital. They find that physicians from the lower ranked school achieved similar patient outcomes but at a higher cost because they ordered more tests and took longer to perform each test. In contrast to their work, we do not use an external proxy for effectiveness (e.g., medical school ranking) but propose ways to construct and validate effectiveness measures from within the data.

The rest of this paper proceeds as follows. Section 2 provides an overview of the VHA setting and the data. Section 3 discusses our empirical strategy. Results appear in Section 4 and conclusions are in Section 5.

2. Setting, Data, and Sample

2.1 Assignment of Veterans to PCPs

Veterans entering primary care in the VHA system are assigned to patient aligned care teams (PACT) that coordinate care. Teams are led by a PCP, who can be a physician, nurse practitioner, or physician's assistant (all of whom have full diagnosing and prescribing authority in the VHA). Note that residents are not permitted to be PCPs, though they can serve as Associate Providers under the supervision of a PCP. PCPs are supported by an advanced nurse (e.g., a registered nurse care manager), a clinical associate (e.g., a licensed practical nurse, licensed vocational nurse, or certified nursing assistant, medical assistant, or health technician), and an administrative associate.

Assignment to a PCP is based on geographic location, scheduling availability, and team capacity.[1] The assignment of patients to PCPs is done via the mandatory Primary Care

---

[1] Per an email exchange with the National VA Office of Primary Care: "New enrollee appointment requests are reviewed for preferred clinic, panel capacity, and [scheduling]

Management Module (PCMM) software program. PCMM is managed by an assigned PCMM coordinator at each VA facility. The data is validated monthly for quality at the centralized Austin Corporate Franchise Datacenter and assessed for consistency with the PCMM protocols. In the system, each PCP is assigned a target number of patients which usually varies between 1000 and 1400 for a full-time equivalent PCP.

The software algorithm generates a lower target if the PCP's existing caseload is expected to take more time, and a higher one if the PCP's existing caseload is expected to take less time. PCPs who are part time (because they have administrative responsibilities, grant support, or because they may do specialist consulting as well as primary care) get pro-rated targets. New patients are assigned to the PCPs with the most unused capacity. Hence, whether a new patient is assigned to a particular PCP depends not on the characteristics of the new patient, but on the characteristics of the PCP's existing patients because that affects measured PCP capacity.

Generally, the assignment is done after a veteran completes Form 1010-EZ to enroll in VHA health benefits. See the appendix for the most recent Form 1010-EZ. The veteran lists basic demographic information, military history, their preferred outpatient clinic, and whether they would like to be contacted by the VA to set up their first appointment. If this last box is checked—as it is on roughly three quarters of all 1010-EZ forms, then a scheduling administrator contacts the veteran. At this point, the veteran explains the reason for their request and gives a desired appointment date[2] and the administrator schedules an appointment. The scheduling typically occurs within seven days after a request is made.

---

availability. If there is capacity and appointment availability at the patient's preferred clinic, an appointment is scheduled and [the patient is assigned to a primary care] team."

[2] The General Accountability Office (GAO) mandates that the VHA collects desired time to monitor wait times.

When the initial primary care visit takes place, the PCP is assigned to the veteran and the relationship is entered into the system.[3] Veterans can choose to switch PCPs, but this is not actively encouraged in the VHA and empirically we do not observe many switches in our sample. Hence, we focus on the first PCP assigned in an "intent-to-treat" framework though we also look at the length of a patient's relationship with the initial PCP as an outcome. In sum, new benefit enrollees seeking primary care services within the same clinic and the same window of time, for future appointments around the same time, are quasi-randomly assigned to PCPs.

2.2 Description of Data Sources

We analyze electronic health records data from the Veterans Health Administration's Corporate Data Warehouse (CDW) between 2004 and the end of February 2020. The standard outpatient, inpatient, and pharmacy data include fields such as hospital, patient, and physician identifiers, diagnoses, procedures performed, origin of prescriptions, prescriber, visit times and dates, etc. Form 1010-EZ and appointment data are available to identify when the patient first enrolled, their preferred clinic, and desired appointment time, which can be linked to the visit with their new PCP. Access to electronic health records gives us a fairly complete view of a patient's health and medical care. For example, we observe referrals to specialists, physician orders (e.g., orders for lab and imaging tests, vaccinations, prosthetics, etc.), patient surveys and questionnaires (e.g., wellness and depression screens); lab and imaging results (e.g., hemoglobin A1c levels which are used for diabetes screening); vital signs (e.g., blood pressure); and receipt of patient education

---

[3] Veterans who do not request an appointment on Form 1010-EZ when enrolling for health benefits get assigned to a PCP at a later point, whenever they request their first primary care appointment. We exclude these veterans because we do not observe the patient's desired appointment date.

(e.g., interventions to promote smoking cessation).[4] Finally, we have data on veteran deaths from

the VHA Vital Status files through early 2020, and from the CDC National Death Index (NDI)

Plus files which gives us both date and cause of death through the end of 2017.

2.3 Sample and Variable Construction

We analyze male veterans between the ages of 20 and 90 who enrolled in VA benefits and first

requested a primary care appointment between January 2005 and February 2017. Starting with

2005 gives us a one year "look back" window to see the patient's previous health history, while

ending in 2017 allows us to follow all patients for three years after enrolling in the VHA. We

focus on male veterans because female veterans are often assigned to Women's Health PACT

teams (Leung et al., 2020). Often there is only one such team in a given clinic so there is no

possibility of random assignment within a clinic and we have little power to conduct a within-

clinic analysis for female veterans.

We begin with 1.02 million Form 1010-EZs representing new VHA enrollees who

a) requested a primary care appointment; b) submitted the form between January 2005 and the

end of February 2017; and c) had at least one completed appointment with a PCP.[5] We restrict

---

[4] Our main analysis focuses on care provided by the VA (i.e., VA medical clinics and

community-based outpatient clinics that are VA-staffed or contracted). For some years we also

have VA data linked to Medicare claims (2011–2016) and Medicaid (2011–2014) and we observe

some care that is paid for by the VA when the VA does not have capacity, or if the veteran lives

far away from a VA clinic. Such care may include emergency care, nursing homes, and various

types of specialty care. Appendix Table A3 shows that including available Medicare, Medicaid

and non-VA data on hospitalizations and ED visits has little impact on our main findings.

[5] We exclude patients whose first visits were connected to an application for disability

compensation or a referral to social work or occupational health. We also excluded patients

our attention to veterans seen at clinics with at least two PCPs in each year (which results in a loss of 40,000 patients) and to PCPs with at least 20 new patients over our study period. The purpose of this latter restriction is to focus on PCPs with enough patients that we can identify their practice style. As discussed further below, we use Bayesian shrinkage methods to compensate for the additional error involved in measuring practice style among doctors with few patients. We lose 3,819 PCPs at this stage. The final baseline sample covers 802,777 veterans who are assigned to 7,548 PCPs at 725 clinics.

We measure PCP effectiveness in the three years following the veteran's initial assignment using hospitalizations and emergency department (ED) visits for mental health/substance abuse and circulatory conditions,[6] and hospitalizations for ambulatory care sensitive conditions.

An alternative approach to measuring effectiveness in a health care setting would focus on what the provider does rather than on patient outcomes. For example, effectiveness may be

---

whose first visit was not to a PCP but to a specialist such as an audiologist. Some veterans with private health insurance rely on the VA to provide services that are not covered by their private plans. The fact that patients who need an immediate referral to a specialist for mental health are not in our sample strengthens the case that the remaining patients are quasi-randomly assigned.

[6] Grouping ED visits and hospitalizations together allows us to look at all patients who arrive at the hospital, whether they are admitted or not. We have also constructed separate ED and inpatient hospital effectiveness measures for mental health and circulatory conditions. The two measures are highly correlated: 0.41 for mental health; and 0.53 for circulatory conditions. Appendix Table A8 reports our main results for the separate ED and inpatient effectiveness measures. They are quite similar.

assessed by how closely the provider adheres to a checklist. However, providers faced with checklists may focus on "checking the boxes" and neglect other important aspects of patient care.[7] Moreover, dealing with checklist can take time away from direct patient care and communication between providers and patients. Hence, many analysts have argued that health systems should put greater weight on health outcomes rather than solely relying on process-based measures in the evaluation of health system quality (see for example, Cutler 2014).

Mental health conditions are among the most common conditions affecting veterans: Over a quarter of primary care veterans have at least one diagnosis of depression, post-traumatic stress disorder (PTSD), substance use disorder, anxiety disorder, or other serious mental illness (Trivedi et al., 2015). Improving the quality of these services has been a VA focus in recent years.[8] VHA guidelines for primary care now recommend annual mental health screenings for depression, post-traumatic stress disorder (PTSD), and alcohol/substance abuse for all new enrollees.

Diseases of the circulatory system are also among the most common health issues among veterans; veterans are twice as likely as non-veterans to have heart disease (Assari, 2014).[9] Earlier and correct management of heart disease in a primary care setting is thought to lead to fewer hospitalizations and better patient health outcomes (Anderson et al., 2020).

---

[7] For example, Medicare's Physician Quality Reporting System included 194 separate quality metrics (Centers for Medicare & Medicaid Services 2012).

[8] Mental health conditions include but are not limited to, psychotic conditions, psychoses and episodic mood episodes, depression, substance use disorders, and suicide attempts/ideation.

[9] This category includes International Classification of Disease (ICD-10) codes beginning with "I", including rheumatic fever and heart diseases, hypertension, ischemic heart disease, pulmonary heart disease, cerebrovascular diseases, and other diseases affecting the arteries, veins, and lymphatic vessels etc.

ACSC hospitalizations are hospitalizations due to conditions such as diabetes, asthma, hypertension, and pneumonia that can largely be avoided with timely, effective, and continued primary care (Barker et al., 2017). The VA does not track hospitalizations for ACSC at the individual PCP level as we do here, but they do track ACSC at the clinic and geographic region level as an indicator of the quality of care and as a cost driver which indicates that the VA is concerned about this outcome.[10]

For all three metrics, we construct an indicator for whether the patient experiences an adverse outcome within three years of requesting an initial PCP appointment. As discussed above, veterans are quasi-randomly assigned to PCPs and PCPs are broadly responsible for managing a patient's care. Hence, we interpret any significant differences in average patient outcomes (leaving out the index patient) as an indicator of PCP effectiveness. The VHA also computes the cost for each patient in each fiscal year.[11] We study average costs both one year and three years after the initial appointment request.

As an early adopter of electronic health records, the VHA has rich data across multiple sources which allows us to go beyond studying differences in outcomes to examine processes of care. We study PCP adherence to VHA clinical guidelines on mental and physical health

---

[10] We construct ACSC hospitalizations using a VHA-modified version of the measure used by the Agency for Health Care Research and Quality (AHRQ 2018).

[11] Average cost is constructed using non-VHA relative value weights (a CMS resource-based relative value scale) to distribute aggregate, national-level costs to each individual inpatient and outpatient encounter (Wagner et al., 2003) and allow dollar-for-dollar comparisons of costs across geographic areas and clinics.

screenings. The VHA has clinical guidelines on mental health screenings,[12] colorectal cancer (CRC), hepatitis C (HCV), HIV, influenza immunization, and tobacco use. Depending on the specific screen, we use outpatient procedure codes, chemical labs, radiology tests, referrals, and orderable request items to identify the performance of these screenings (e.g., the PCP can place an order/request for a technician to conduct a blood test). All screening metrics are restricted to suitable populations; for example, guidelines for CRC recommend annual fecal occult blood testing for adults between ages of 50 and 75 but not for younger or older veterans. We construct indicators for whether the veteran got each recommended screening in the first year after the initial primary care appointment.

Finally, we examine management of diabetes, high cholesterol, and hypertension in patients who have been diagnosed with those conditions. Appropriate management of these conditions could, in the medium to longer run, greatly improve health and reduce health care costs.

For all our metrics, we do not require the screen or outcome to be linked to the PCP. The VHA's primary care philosophy is one where the PCP team is responsible for coordinating a patient's care which could well be rendered by other practitioners.

3. Empirical Strategy

3.1 Constructing measures of PCP effectiveness

Measures of PCP effectiveness are constructed using an empirical Bayes jackknifed value-added measure (Kane and Staiger, 2008, Chetty, Friedman, and Rockoff, 2014, Jackson et

---

[12] Specifically, mental health VHA guidelines recommend all new patients receive a Patient Health Questionnaire (PHQ; two item or nine item), Primary Care PTSD Screen for DSM-5 (PC-PTSD-5), and Alcohol Use Disorders Identification Test-Concise (AUDIT-C).

al., 2020).[13] This approach improves on using the raw "leave-out" probability that a doctor's patients experience an adverse outcome which would be calculated simply by taking the mean after leaving out the index patient. These raw probabilities may be very noisy for PCPs with few patients. Instead, probabilities are weighted using the number of new patients assigned to each PCP each year. Given the way that patients are assigned to PCPs, we need to account for the specific appointment month and year, the primary care clinic, the day of the week of the initial visit, and for the number of days the patient waited for an appointment. The effectiveness measure is thus constructed as a weighted average of the residualized probabilities that a PCP's patients suffered adverse outcomes, where the weights depend on the number of observations in each period.

---

[13] In this literature, researchers first construct a measure of teacher value-added using Bayesian shrinkage methods, and then explore the effects of the value-added measure on student outcomes. Value-added measures are estimated with error but are "Best Linear Unbiased Predictors of a teacher's impacts on average student achievement" (Kane and Staiger, 2008). Hyslop and Imbens (2001) prove that when the reported variable (value-added) is an optimal predictor, then the measurement error in the reported variable (the prediction error) is uncorrelated with the predictor and the model can be consistently estimated by OLS. Measurement error will cause the estimated standard errors to be larger than with a perfectly measured regressor. The key issue is whether the measurement error in the value-added measure is correlated with the estimated measure. We rely on the random assignment of patients to PCPs to break any link between unobserved patient health and PCP value added. The approach is also similar to studies using a "judge instrument" such as Doyle et al. 2015, Dobbie et al. 2018, and Eichmeyer and Zhang 2022.

In order to calculate the measure for each PCP we first estimate the following equation for patient i, PCP j, and year t:

$$Y_{ijt} = \theta X_{it} + \gamma_{ym} + \gamma_{clinic} + \gamma_{day} + \gamma_{desired} + \varepsilon_{ijt}, \tag{1}$$

where Y_{ijt} is an indicator variable for an ED or hospital encounter for mental health, circulatory condition, or an ACSC within three years of assignment to a new PCP. This outcome variable is regressed on indicators for year by month, γ_{ym}; primary care clinic, γ_{clinic}; day of week of the initial visit, γ_{day}; and bins for the number of days between the veteran's desired date for a first appointment and the date of the actual appointment, $\gamma_{desired}$ (0, 1–7 days, 8–14 days, 15–21 days, 22–30 days, 30–60, and 60+ days). These are the only controls required for unbiased estimation of PCP effectiveness; veterans are quasi-randomly assigned to PCPs controlling for clinic, year-month, day of week, and a vector of days to desired appointment date fixed effects.

In order to improve precision, variables that are pre-determined as of baseline assignment to a PCP are included. These baseline controls, X_{it}, include race (Asian/Pacific Islander, Black, Hispanic, White); five-year age bins, marital status, enrollment priority groups;[14] indicators for being a beneficiary of Medicare or Medicaid; whether the patient used the VHA in the previous year; whether they had any prior year mental health, circulatory condition, or ACSC hospitalizations; whether the veteran has any service-connected disability or is considered unemployable; indicators for era of service (e.g., Korean war, Vietnam war), and indicators for exposure to Agent Orange or radiation.

Yearly jackknife PCP propensities are calculated by averaging the residuals leaving out the own residual term corresponding to patient i, PCP j: $\widehat{W}_{jt} = \sum_{i \in K_{-ijt}} \hat{\varepsilon}_{ijt}$, where K_{-ijt} denotes

---

[14] Priority for enrollment in VHA benefits depends on the veteran's income, disability status, and combat history. We include an indicator for each.

the set of patients assigned to PCP j in year t, excluding the index patient i. The final step computes the empirical Bayes PCP effectiveness measure as a function of the vector of yearly effectiveness measures for that PCP, $\vec{W}_j$, and a vector of the number of newly assigned veterans for each PCP, $\vec{N}_j : \hat{Z}_j = Z(\vec{W}_j, \vec{N}_j)$. Multiple years are used to improve statistical power and the weights are determined semi-parametrically and estimated from the data. Specifically, we estimate the following equation:

$$Y_{ijt} = \sum_{t=2005}^{2017} \sum_k \beta_{kt} \, 1\{N_{jt} = k\} \times \widehat{W}_{jt} + \epsilon_{ijt}, \tag{2}$$

where N_{jt} denotes the number of new veterans assigned to PCP j in year t. Four bins are created for the number of new patients seen: 0–9, 10–24, 25–50, and over 50 new veterans. Empirically equation (2) places more weight on yearly jackknife probabilities that are estimated with more precision and less weight on probabilities that are estimated with more noise. The latter shrink towards zero, the expected value of $\widehat{W}_{jt}$.[15]

---

[15]Chetty et al. (2014) note that calculating the probabilities over multiple years could allow for the idea that professionals learn and change their behavior over time. The estimates reported below are based on a single measure for each PCP in order to reduce noise and increase statistical power. We have also constructed effectiveness measures over two periods (2005–2011 and 2012–2017) for PCPs who had at least 20 new patients in each period. There were 2,566 PCPs such physicians. The within provider correlations between the 2005–2011 and 2012–2017 measures are shown in Appendix Table A10. The correlations in mental health effectiveness measures, circulatory measures and ACSC were 0.81, 0.79, and .49 respectively. This somewhat lower correlation for ACSC across time periods might reflect the effort the VHA has put into reducing ACSC at the facility level.

Figure 1 plots histograms for each of the raw PCP effectiveness measures before standardization (described below). The value of each measure represents the percentage increase in the probability that a PCP's patient visits an ED or hospital within three years, relative to all other providers, conditional on the equation (1) controls. All three raw PCP effectiveness metrics are symmetric around mean zero by construction. Effectiveness with respect to circulatory conditions exhibit the largest variation across PCPs while the variance of mental health and ACSC effectiveness is lower. The standard deviation of the effectiveness measures for circulatory conditions, mental health conditions, and ACSC conditions are 0.024, 0.017, and 0.017, respectively. It is important to keep in mind that these measures capture within-clinic, and within-year and month variation in PCP effectiveness. Hence, regional differences in health or trends over time should not affect them.

[Figure 1 here]

Finally, we take the fitted predicted values from equation (2), $\hat{Y}_{ijt}$, standardize the variable, and take its negative to be able to interpret it as effectiveness (as opposed to being the propensity to have patients experience *adverse* outcomes). This effectiveness measure is denoted:

$$E_{ijt} = -(\hat{Y}_{ijt} - E(\hat{Y}_{ijt}))/\sqrt{Var(\hat{Y}_{ijt})}. \tag{3}$$

Table 1 shows how the mean characteristics of veterans in our sample vary across PCPs in different effectiveness bands. The first column shows means for the entire sample while columns two through four show means for patients divided into terciles of provider effectiveness for circulatory issues. Dividing the sample by mental health or ACSC metrics yield similar patterns.

The average veteran is a late-middle aged (55) white male; the sample is 74.3% non-Hispanic White, 5.8% Hispanic, 13.2% Black, and 1.7% Asian/Pacific Islander. About 58% are currently married. About 30% are on Medicare and 5.4% are on Medicaid at the time of

enrollment. Half have some form of service-connected disability. The average veteran's income

is $44,413 in 2019 dollars. For some veterans (13.2%), we observe their prior medical history if

they were previously treated at a VA hospital or emergency department without enrolling in VA

health benefits. Prior circulatory hospitalization or ED use was most common, followed by

mental health, and then an ACSC, but less than 1% of new enrollees had one of these events. As

alluded to earlier, patients do not often switch providers; the average PCP-patient relationship

over the three-year window in which we follow patients is 23 months (693 days). Table 1 also

includes some information about diagnosis at the veteran's first visit, which was not included in

equation (1) as it may not be pre-determined.

Looking across terciles of PCP circulatory care effectiveness measures (columns 2

through 4) supports the idea that patients are quasi-randomly assigned. There is little difference in

any of the measures across columns suggesting that veterans are distributed evenly across terciles

of PCP effectiveness in terms of their demographics, service history, and medical conditions.

[Table 1 here]

3.3 Assessing quasi-random assignment

Figure 2 provides another look at the assumption that veterans are randomly assigned to

PCPs of differing levels of effectiveness by showing a "balance" test. Figure 2a is constructed by

first regressing the PCP effectiveness measures on clinic, year-month, day of week, and the

number of days between the veteran's desired date and the date their appointment was made (in

bins). We then construct the effectiveness measure in equation (3) and regress it on the set of

observable patient characteristics. Note that the controls X_{it} in equation (1) are *not* included in

the minimal PCP effectiveness measure plotted in Figure 2a.

Figure 2a allows us to see whether, within a clinic, veterans who are assigned to PCPs

with higher levels of effectiveness differ in terms of pre-determined observable variables such as

demographics, military service history, eligibility category, prior year's medical history (when

that is available). We have also included diagnosis codes (MDC codes) measured at the first PCP

visit, though these could possibly reflect provider effectiveness in terms of diagnosis as well as

any underlying condition. Although a few coefficients are statistically significant (which is not

surprising given the large sample size) there is little indication that PCP effectiveness is

systematically related to characteristics or patient health. In some cases, the estimates imply that

sicker patients are assigned more effective doctors, which would bias estimates of the importance

of PCP effectiveness towards zero. Furthermore, all our specifications control for prior patient

health (MH, circulatory condition, and ACSC) when it is available.

[Figure 2 here]

Following Chetty et al. (2014), Figure 2b provides a second look at balance. We group

patients into 20 equally-sized bins and calculate average patient outcomes, residualizing for

clinic, year-month, day of week, and days to desired appointment date fixed effects. The red lines

show the relationship between PCP effectiveness and *actual* three-year patient outcomes (mental

health, circulatory ED/hospitalizations, and ACSC hospitalizations). These lines show that

patients assigned to more effective PCPs have better actual future outcomes. The black lines

show *predicted* patient outcomes, predicted using the full set of predetermined observable veteran

characteristics in addition to clinic, year-month, day of week, and days to desired appointment

date fixed effects.[16] These lines are virtually flat. They indicate that predicted outcomes based on

---

[16] The *R*-squareds on the regressions of mental health, circulatory, and ACSC outcomes on the

complete set of veteran characteristics shown in Figure 2 are 0.024, 0.032, and 0.020

respectively. The modest *R*-squared values reflect how difficult it is to predict these health

pre-determined information are not meaningfully correlated with the effectiveness measures. For example, the correlation between mental health effectiveness and predicted mental health outcomes is only 1.3% of the correlation between mental health effectiveness and actual mental health ED visits and hospitalizations.

Information about the prior year's health history is only available if the patient was seen somewhere in the VA system in the previous year. Most patients in our sample are being seen for the first time. Since it is conceivable that patients are treated differently if prior information is available in the system, we also construct effectiveness measures and replicate our analyses using only veterans who had no prior VHA utilization (as discussed further below). Similarly, Appendix Figure A2 replicates Figure 2 for veterans without any prior VHA utilization. The figure is very similar to that obtained using the full sample.

Appendix Figure A3 shows similar balance figures for two additional outcomes, total costs and mortality, both measured three years after PCP assignment. As can be seen, predicted costs are almost perfectly flat with respect to our three effectiveness measures, showing that there is no relationship between predicted cost and our measures of PCP effectiveness. Predicted mortality is also flat when plotted against mental health effectiveness, but has some downward slope for ACSC and cardiovascular effectiveness, although predicted mortality is still considerably flatter than actual mortality. We interpret these results to mean that it is possible that omitted variables bias drives some of the relationship between provider effectiveness and reduced mortality that we see. Hence, we explore the likely extent of bias from omitted variables using methods suggested in Oster (2019) below.

---

outcomes even with the detailed administrative data available to the VA. Regressions of 3-year mortality and 3-year log total costs have higher R-squareds of 0.077 and 0.058 respectively.

3.3 Correlating PCP effectiveness with other measures of PCP practice variation and PCP characteristics

Equipped with these measures of PCP effectiveness, we validate them by asking whether each individual effectiveness measure (i.e., mental health, circulatory, or ACSC) is individually predictive of other patient outcomes of interest besides ED visits and hospitalizations, notably mortality and health care costs. Importantly, mortality and health care costs were not used to construct the metrics. We estimate the impact of PCP effectiveness, E_{ijt}, on mortality and total costs for the 802,777 new patients assigned to PCPs over the sample period:

$$Y_{ijt} = \beta E_{ijt} + \theta X_{it} + \gamma_{ym} + \gamma_{clinic} + \gamma_{day} + \gamma_{desired} + \varepsilon_{ijt}. \tag{3}$$

The parameter of interest, β, represents the impact of a standard deviation increase in one of the measures of PCP effectiveness on a patient outcome (for example, death in the next three years). Equation (3) includes the same controls as in equation (1).

We next ask how these measures of PCP effectiveness are related to measures of practice style. Do more effective physicians achieve better results by ordering more tests, by making more referrals, or by encouraging more visits? Are they more likely to conduct screenings as recommended by the VHA? These questions are explored using models similar to equation (3) but using alternative outcome measures.

We also correlate PCP effectiveness with time-invariant provider characteristics such as the demographics of the provider. Instead of patient-level regressions, these models focus on a provider-level measure of effectiveness obtained by averaging the fitted values in equation (2) across each PCPs' patients: $E_j = \sum_{it} \hat{Y}_{ijt}$. This provider-level PCP effectiveness measure is then regressed on a provider's own characteristics Q_j, for the 7,548 PCPs:

$$E_j = \mu + \theta Q_j + \eta_j + \varepsilon_j. \tag{4}$$

A fixed effect for the PCPs home clinic, η_j is included so (4) identifies within-clinic provider differences.

4. Results

Table 2 shows regressions in the form of equation (1), where the dependent variables are the veteran's mental health, circulatory, or avoidable hospitalization outcomes and the independent variable of interest is that veteran's PCP's leave-out jackknife measure of effectiveness. As discussed above, this (standardized) effectiveness measure is the PCP's (residualized) average 3-year mental health ED or inpatient rate, excluding the index patient. It is important that the patient's own residual is left out of the model; otherwise, there would be a mechanical correlation.

The estimates—which are scaled by 100—suggest that PCPs with a one standard deviation (SD) higher mental health effectiveness are 1.56 percentage points (pp) less likely to have their patients visit an ED or be hospitalized for mental health, on a baseline of 4.97% (a 31% reduction; column 1). Similarly, a one SD higher circulatory condition effectiveness PCP is 1.96 pp less likely to have an adverse circulatory outcome (over a baseline of 7.37%; a 27% reduction) in column 4; and a one SD higher ACSC effectiveness PCP is 1.12 pp less likely to have their patient be hospitalized for ACSC (over a baseline of 2.52%; a 44% reduction) in column 7.

[Table 2 here]

Looking at the other columns in Table 2 suggests that physicians who are more effective in dealing with one type of health condition are also more effective at dealing with the other two. For example, patients whose PCPs have a one SD higher measure of mental health effectiveness are 0.58 pp less likely to have an ED visit or hospitalization for a circulatory condition (column

21

2) and 0.48 pp less likely to have an avoidable hospitalization over the next three years (column 3).

We have also estimated these measures and models using three mutually exclusive patient subsamples. Each effectiveness measure (e.g., ACSC hospitalizations) was computed using 1/3 of the sample patients, and then that measure was used in a regression estimated using the other 2/3 of the sample. Using separate subsamples addresses the concern that there could be within-patient correlations in the need for the three types of care. The estimated coefficients are very similar to those in Table 2 and can be found in Appendix Table A1.

Table 3 explores the relationship between being assigned to a PCP with a one standard deviation increment in PCP effectiveness, and 3-year mortality, 1-year costs, and 3-year costs. Each element of the table corresponds to a separate regression and only the coefficient of interest, β, is shown. The regressions are in the form of equation (3) and the standard errors are clustered at the PCP level.

Panel A shows that assignment to a PCP with a one standard deviation higher effectiveness measure is associated with a reduction of 0.20 to 0.23 percentage points in the risk of mortality in the next three years. Given the baseline three-year mortality risk of 5.5%, this estimate translates into a 3.6 to 4.2% reduction in mortality. Both one-year and three-year total costs also fall by between 2.5 and 5.4% depending on the measure, with the largest reductions in total costs for PCPs who are relatively more effective than others within their clinics at preventing ER visits and hospitalizations for circulatory conditions. Hence, these effectiveness measures are also predictive of important patient outcomes that were not used in their construction.

[Table 3 here]

These estimates are robust to several changes in sample design. Appendix Table A2 shows estimates after dropping veterans who appear in the VHA records prior to PCP assignment: Arguably, the VHA system could use such prior information to assign veterans to PCPs. The table includes circulatory ED visits and hospitalizations; three-year mortality; and three-year total costs as well as number of visits and referrals as discussed further below. Table A2 shows that in the subsample without any such information, the estimates are similar to those discussed above.

Appendix Table A3 uses a subset of veterans for whom information about utilization of care outside the VHA system is available and includes outside ED visits and hospitalizations in the construction of the outcome variables. Appendix Table A4 excludes veterans who waited more than two weeks for an initial appointment (since it is conceivable that they might be waiting to see a specific PCP). In all cases, the estimates are extremely similar to those reported in Table 3.

Appendix Table A5 reports an additional check in which we first regressed patient age on PCP fixed effects in addition to controls for year-month, clinic, day of week, and bins for the number of days between desired and actual date of appointment, separately for each clinic. We then tested to see whether we could reject the null that the PCP dummies were predictive of patient age. We found that 402 out of 725 clinics had a p-value greater than 0.10. Table A5 shows the results of repeating our main results in this subset of clinics where it is most plausible that there is random assignment. The results are similar to those in Table 3.

Appendix Table A6 reproduces the results from Table 3 excluding non-physician PCPs (i.e., teams led by nurse practitioners and physician assistants) in case there is some non-randomness in the assignment of patients across types of PCPs. The estimates are very similar to those in Table 3.

Panel B of Table 3 drills down on the mortality results by examining three-year mortality for the largest cause of death categories. It is reasonable to assume, for example, that PCPs who are effective in reducing ER visits and hospitalizations for circulatory conditions might be good at helping patients avoid deaths due to heart conditions. It is unclear though whether they would also be good at helping patients avoid deaths from other common causes such as cancer. The extent to which there are spillovers onto other causes of death depends on how correlated effectiveness is across domains of care. Table 3b suggests that there are some spillovers, but that these different measures also capture particular domains of PCP expertise.

For example, being assigned to a PCP with a one standard deviation higher measure of mental health effectiveness is associated with reductions of 13.3% in the probability of death from suicide and an 8.7% fall in the probability of death from external causes. This latter category includes confirmed suicides as well as deaths from overdoses, poisonings, and accidents, some of which may have actually been suicides.[17] A one standard deviation improvement in mental health effectiveness is also associated with a 0.050 percentage point reduction in the probability of a cancer death, on a baseline of 1.48%, a 3.4% reduction. The estimates also imply a 4.0% reduction in the probability of death from heart disease.

Patients assigned to PCPs with a one standard deviation higher measure of circulatory care effectiveness see similar reductions in the probability of death from cancer or heart disease,

---

[17] In Appendix Table A9 we look at non-poisoning accidents separately and show that PCP effectiveness does not have a statistically significant effect. The most frequent external causes are drug poisonings, accidents, and suicides but in the VHA accidents are the most common cause of death among patients who had been seen in the ED within the past month for suicidality, so that "death by car accident" may in fact be a type of suicide in some cases.

but no reduction in the probability of death from suicide, and only a 4.0% reduction in the probability of death from external causes. These results suggest that some PCPs who are effective at caring for patients with circulatory conditions may lack expertise in caring for patients with mental health risks.

Patients whose PCPs are one standard deviation higher in terms of effectiveness in preventing hospitalizations for ambulatory care sensitive conditions achieve the largest reductions in deaths from cancer (4.3%), and heart disease (4.5%), as well as a 6.3% reduction in external causes of death over the next three years, though there is no statistically significant effect for confirmed suicides.

None of the three measures predict reductions in deaths from lower respiratory conditions or cerebrovascular events suggesting either that these deaths may be harder to prevent, or that they represent another dimension of care effectiveness that may not be highly related to the measures we examine.

As discussed above, we found some evidence of a relationship between predicted mortality, cardiovascular effectiveness, and ACSC effectiveness suggesting that some of the estimated mortality effects associated with these two measures could be driven by patient sorting on omitted variables, a possibility we have explored using methods suggested by Oster (2019). The results are shown in Appendix Table A11. They suggest that the unobservables would have to be at least twice as important, and in some cases up to eight times as important as the observables, in order for our findings to be driven by unobservables. Since Oster (2019) suggests that a ratio of one to one is reasonable, these estimates are reassuring.

4.1 Effects on use of care

So far, we have seen that patients of PCPs with higher effectiveness scores face a lower risk of death and incur lower total costs over a one year or a three-year horizon. How are these positive results achieved? Is it the case, for example, that the patients consume more preventive

care and thus are spared expensive illnesses? These questions are explored in Tables 4 and 5 which estimate models in the form of equation (3), separately for each effectiveness measure.

Table 4 examines the relationship between PCP effectiveness and the number of medical encounters in the first year after assignment to a PCP (panel A). The first column shows that a one standard deviation in PCP effectiveness is associated with a reduction of 2 to 3% in the overall number of medical encounters (e.g., a one standard deviation improvement in mental health effectiveness reduces the total number of visits by 0.395 percentage points on a baseline of 13.4%). Some of this improvement is due to large reductions in the probability of any ED visits or inpatient hospitalizations as shown in columns 4 and 5. However, since the effectiveness measures were constructed with reference to ED visits and hospitalizations these significant relationships are not surprising.

[Table 4 here]

What is more surprising is that there are reductions in the number of primary care visits of 1.3% to 2.2%, as well as reductions in the number of mental health visits. It is striking that patients assigned to a PCP who is one standard more effective at treating mental health have 8.2% fewer mental health visits in the first year (a reduction of 0.106 on a baseline of 1.3 visits). Hence, it does not seem to be the case that more effective doctors are providing quantitatively more general primary care or more mental health care. Column 6 shows that in the subset of patients over 65 who also qualify for Medicare (and for whom we have Medicare records) there are no differences in the number of visits outside of the VA. Hence, the reduction in visits at the VHA is not offset by increases in visits elsewhere. One possibility is that more effective physicians see patients less but do more per visit. We investigated this hypothesis by examining

26

the relationship between effectiveness and Relative Value Units of care in Appendix Table A7 but found little evidence in support of this explanation.[18]

Another way that a PCP might achieve greater effectiveness is by referring patients to specialists when needed, or by conducting more lab and imaging tests. Panel B of Table 4 examines referrals, laboratory tests, and imaging. It suggests however that more effective PCPs are actually somewhat less likely to do any of these things.

While some of the differences in referrals are quite small, a one standard deviation increase in a PCP's mental health effectiveness is estimated to reduce referrals for mental health by 3.0% (0.63 on a baseline of 20.9%) and to reduce referrals to cardiology by 4.1% (0.29 on a baseline of 7.0%). A one standard deviation increase in circulatory condition effectiveness reduces referrals for mental health by 1.5% but reduces referrals to cardiology by 9.6% (0.67 on a baseline of 7.0%). The measure of effectiveness at preventing hospitalizations for ambulatory care sensitive conditions has little effect on referrals however.

All three measures of PCP effectiveness are negatively associated with ordering laboratory panels, with reductions ranging from 1.5% for a one standard deviation increase in mental health effectiveness to 3.2% for a one standard deviation increase in circulatory condition effectiveness. Similarly, for imaging there are reductions of 2.0% (for mental health effectiveness) to 4.0% (for circulatory condition effectiveness).

---

[18] One problem with examining RVUs is that because VA providers are paid a salary rather than fee-for-service, they do not always record procedures rendered. The average PCP visit in our data has an RVU of 0.63 compared to 1.59 for visits in the general population in 2016 (NACHC 2016) and 0.97 in a standard established Medicare visit (CPT 99213).

Table 5 looks at whether PCPs who are more effective according to our measures are more likely to follow VHA guidelines for screening veterans. For some types of screens, compliance is already very high in the VHA, leaving little room for within-clinic variation across PCPs. Panel A of Table 5 focuses on screenings for depression, PTSD, and substance use. Compliance with all these screens varies from 94.2% to 96.9% for new enrollees, in keeping with the strong emphasis the VHA places on mental health. Nevertheless, we do see some statistically significant, albeit small positive relationships between PCP effectiveness for circulatory conditions and ACSC and the probability of conducting these mental health screenings. The magnitudes vary from increases of 0.11% to 0.29%.

Panel B of Table 5 looks at whether patients received recommended screenings for colorectal cancer, hepatitis C, HIV, and tobacco use, and whether they received immunizations for influenza. Aside from screening for tobacco use, these physical health screenings have much lower average compliance rates. While most of the estimated coefficients are not statistically significant, those that are significant suggest a small negative relationship between PCP effectiveness and these screenings. For example, a one standard deviation increase in effectiveness for circulatory conditions is estimated to reduce the probability of screening for hepatitis C by 1.9% (0.9 on a baseline of 47.3%) while a one standard deviation increase in effectiveness for ACSC reduces it by 1.6%. A one standard deviation improvement in mental health effectiveness reduces the probability of screening for HIV by 1.4%. The only positive and significant coefficient in the table is for the effect of ACSC effectiveness on tobacco screening, but the magnitude is very small: 0.11%.

This section demonstrates that assignment to some PCPs generates better outcomes while leading to small reductions in the quantity of care consumed along most dimensions.

[Table 5 here]

4.3 Characteristics of effective PCPs and the patient-PCP match

We have argued that some PCPs appear to be more effective than others working within the same clinics in terms of avoiding negative health outcomes for their patients. How are our measures of PCP effectiveness related to observable PCP characteristics? This question is explored in Table 6 which shows estimates of equation (4). Because we are looking at within-clinic variations in PCP effectiveness, indicators for the PCPs home clinic, η_j, are included in the model to ensure that we are identifying within-clinic variation.

Unfortunately, we do not see information about the PCP's training, but we do know their gender, age, and whether or not they are a physician. PCP experience is proxied with the variables "New Patients Per Year," age, and to a certain extent, an indicator for "Part-Time." Given the limited information about not only the PCP but about the other team members, these data are not ideal for studying the effects of team composition on outcomes[19], and we keep the analysis at the level of the team as a whole with the focus on the PCP as the team leader.

Because age changes over time and this is a PCP-level regression, we take the weighted average of the PCP's age at the time each new patient is assigned. We can also generate information about the means of certain practice characteristics from the data. Here we look at the number of patients they see per day, the number of new patients they see per year, and whether they are a full-time equivalent (defined as seeing at least one patient on 250 days a year). While PCPs who work full-time may amass more relevant experience, in the VHA many research faculty hold part-time appointments so this flag may also be capturing that distinction.

[Table 6 here]

Table 6 suggests that effectiveness increases with age, number of patients per day and the number of new patients per year. A one standard deviation increase in patients per day (4.25 patients) is estimated to improve circulatory condition effectiveness and ACSC effectiveness by

---

[19] But see Chen (2021) and Agha et al. (2018) for interesting analyses of teams.

0.068 and 0.11 of a standard deviation, respectively. A one standard deviation in new patients per year (12.29 patients) would increase mental health, circulatory condition, and ACSC effectiveness by 0.17, 0.25, and 0.11 standard deviations, respectively. Physicians are slightly less effective (about 0.1 standard deviations) than nurse practitioners and physician assistants in terms of avoiding ED visits and hospitalizations.[20]

PCPs whose patients receive a larger proportion of mental health visits that utilize the embedded mental health team (in clinics where there is a licensed mental health specialist who can be called in for an immediate consultation[21]) achieve higher effectiveness along all three dimensions. This greater utilization of care coordination in the mental health sphere may help to explain the lower referral rate for mental health that was discussed above.

---

[20] There are some differences between the average characteristics of patients seen by non-physician PCP's and those seen by physicians. The former are slightly younger, less likely to be Black and less likely to be never-married. The two groups have the same probabilities of having a service-connected disability or of being deemed unemployable. The patients seen by physicians are slightly more likely to receive an initial diagnosis of cardiovascular disease or metabolic disorders after they have been seen by their new PCP, and slightly less likely to be diagnosed with mental health disorders, though it is unclear whether their underlying health status actually differs. In Appendix Table A6 below, we repeat the main analysis excluding non-physician PCPs and show that the results are robust to this change.

[21] Internally referred to in the VA as Primary Care-Mental Health Integration, PCMHI integrates mental health care with the veterans' primary care team in the same clinic, usually in the same day, rather than referring them to for a separate visit at a future date. The independent variable is defined as the fraction of all outpatient mental health visits that are integrated with primary care.

Providers who spend more of their time at the VHA as part-time workers also have higher effectiveness ratings. This may be because these PCPs are more likely to be researchers or in administrative leadership roles.

Table A12 seeks to address the question of whether patients are aware of provider effectiveness. As discussed above, patients are not encouraged to switch providers in the VHA, and switching is relatively rare. However, we do see variation in the length of time that a patient stays with a particular PCP after their initial assignment. Column 1 shows that there is a small positive relationship (a little over a week on a baseline of 693 days) between our measures of PCP effectiveness and the length of a patient's relationship with that PCP. Some of this could be mechanical since more effective PCPs were shown to reduce the patient's probability of death. Column (2) shows that if we exclude patients who die within three years, we see a very similar relationship between effectiveness and the length of the patient-PCP relationship.

5. Discussion

We address the following questions in the unique context of the VHA: Are some providers more effective than others in promoting patient health and how can we measure that? Do patients whose providers are effective in one domain do better in other domains as well? And if some providers are generally more effective than others, what characteristics of providers predict effectiveness?

These questions are hard to answer for the same reasons that make teacher "value-added" measures controversial. Teacher value-added models seek to assess teacher effectiveness by looking at student outcomes. Similarly, in health settings we may try to assess provider effectiveness using patient outcomes. In most settings, patients sort non-randomly across providers. If patients choose their providers, and if sicker patients are referred to more experienced providers, or if some patients do not have access to more skilled providers, then inferences based on patient outcomes may be biased. Researchers typically try to solve this

problem through risk adjustment, that is by correcting for observable differences in patient mix. But there may be important characteristics of patients that are observed by providers and not by the risk adjusters. The VHA's system of quasi-randomly assigning patients to PCPs within a clinic provides a solution to these measurement problems.

Our results suggest the following answers to the questions we posed: First, some PCPs are indeed more effective than others. While we constructed our measures with reference to future ER visits and hospitalizations, we were able to validate them by showing that these measures of PCP effectiveness predict future mortality from a variety of causes as well as health care costs.

Second, provider effectiveness is positively related across the three domains of effectiveness we examine (mental health, circulatory conditions, and hospitalizations for ambulatory care sensitive conditions). Patients of PCPs who are more effective in terms of one of our three measures also have better outcomes in the other two domains. These results suggest that it may not be necessary to measure effectiveness in every possible dimension in order to identify more effective PCPs. Still, since there are many other possible health domains and outcome measures that could be investigated, this finding should be regarded as a preliminary "proof of concept," to be validated by looking at additional domains of patient outcomes in order to find those that are most predictive of a range of patient outcomes.

Our third and most striking finding is that more effective PCPs do more with less. Patients of more effective providers have fewer primary care visits, fewer referrals, fewer lab and imaging tests, and even fewer preventive health screenings. This finding is consistent with previous work showing that physicians who do more do not necessarily achieve better patient outcomes (e.g., Currie and MacLeod, 2016; Chan, Gentzkow, and Yu, 2019). Doyle, Ewer, and Wagner's (2010) finding that physicians from lower ranked schools took more time and ordered more tests conditional on health outcomes seems particularly relevant. These results suggest that higher quality care does not necessarily involve more visits, tests, or procedures.

32

These results raise the question of mechanisms. One possibility is that more effective PCPs are good communicators. Several researchers suggest that better communication between patients and providers can improve take up of preventive care services (Alsan, Garrick, and Graziani, 2019; Koulayev, Simeonova, and Skipper, 2016; Simeonova, Skipper, and Thingholm, 2021). In our case, patients are actually consuming these services less frequently so a communication mechanism would have to operate in a somewhat different way, perhaps by allowing PCPs to obtain the information they need without ordering unnecessary tests.

The VA data have information about important health markers that an effective PCP might target, but in most cases, data is not available from the period before patients started seeing the PCP. Although random assignment implies that the distribution of patients with high blood pressure should be the same across providers at the time of assignment, we would like to know whether patients with initially high blood pressure are more likely to have it eventually brought under control when they see a more effective PCP. Appendix Figure A4 shows the estimated impacts of PCP effectiveness on the probability that patients who have been diagnosed with high blood pressure, diabetes, or high LDL cholesterol have these conditions under control. We also examine medication compliance defined as whether the average medication possession ratio of anti-hypertensives started that half-year is at least 80% (a threshold the literature uses). These figures use all the available data (i.e., an unbalanced panel) for six-month intervals ranging from one year before the first PCP visit to three years afterwards. The estimates are very noisy but do suggest a positive and slowly rising probability that these health conditions are brought under control after the patient begins seeing a more effective PCP, as well as improvements in medication compliance.

Another reason why physicians who do more do not necessarily achieve better outcomes has to do with the allocation of patients across providers. Chandra and Staiger (2007, 2020) discuss an example in which if one doctor is skilled in providing drug therapy to heart patients

while another is skilled at heart surgery, then outcomes will be better if the patients needing surgery are allocated to the skilled surgeon and vice-versa. Even if some PCPs are more effective along all relevant dimensions, the principal of comparative advantage suggests that there may still be gains from reallocating patients across physicians. Because we show that effectiveness measures across different domains are imperfectly related (e.g., mental health effectiveness is more predictive of mental health outcomes than of circulatory outcomes), there may be some potential gain from reallocating patients within the VA to PCPs who are relatively more skilled at dealing with their particular problems.

Currie and MacLeod (2016, 2020) focus instead on each individual physician's allocation of procedures across patients and show that some physicians are more skilled than others in terms of efficiently allocating procedures. In our context, this may mean that effective providers allocate resources to the patients who need them most, while less effective providers use resources more indiscriminately. They may, for example, order a lot of unnecessary tests. Our finding that more effective providers do not always follow guidelines suggests that a better targeting of resources may require providers to use their best judgment of when deviations from guidelines are warranted.

Determining the reasons why some PCPs are able to do more with less is an important avenue for future work. In the meantime, our results suggest that health administrators should be cautious in seeking to eliminate "unnecessary" referrals and tests: Given variations in provider effectiveness, some providers may need to use more resources to achieve the same patient health outcomes.

References
Abaluck, Jason, Leila Agha, David C. Chan, Daniel Singer, and Diana Zhu, "Fixing Misallocation with Guidelines: Awareness vs. Adherence," *NBER working paper 27467* (2020).

Agha, Leila, Keith M. Ericson, Kimberley H. Geissler, and James B. Rebitzer, "Team Relationships and Performance: Evidence from Healthcare Referral Networks," *NBER working paper #24338* (2018).

Anderson, Kim, Heather J. Ross, Peter C. Austin, Jiming Fang, and Douglas S. Lee, "Health Care Use Before First Heart Failure Hospitalization: Identifying Opportunities to Pre-Emptively Diagnose Impending Decompensation," *Journal of American College of Cardiology: Heart Failure*, 8:12, (2020), 1024–1034.

Assari, Shervin, "Veterans and Risk of Heart Disease in the United States: A Cohort with 20 Years of Follow Up," *International Journal of Preventive Medicine*, 5:6, (2014), 703–709.

Barker, Isaac, Adam Steventon, Sarah R. Deeny, "Association Between Continuity of Care in General Practice and Hospital Admissions for Ambulatory Care Sensitive Conditions: Cross Sectional Study of Routinely Collected, Person Level Data," *British Medical Journal*, 356:j84 (2017).

Bertsimas, Dimitris, Margrét V. Bjarnadóttir, Michael A. Kane, J. Christian Kryder, Rudra Pandey, Santosh Vempala and Grant Wang, "Algorithmic Prediction of Health-Care Costs," *Operations Research in Health Care*, 56:6 (2008), 1382–1392.

Chan, David C., Matthew Gentzkow and Chuan Yu, 2019. "Selection with Variation in Diagnostic Skill: Evidence from Radiologists," *NBER working paper #26467* (2019).

Chandra, Amitabh, and Jonathan Skinner, 2012 "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic Literature*, 50:3 (2012), 645–80.

Chen, Yiqun, "Team-Specific Human Capital and Team Performance: Evidence from Doctors," *American Economic Review*, 111:12 (2021), 3923–3962.
Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers I: Estimating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104:9 (2014), 2593–2632.

Currie, Janet, and W. Bentley MacLeod, "Diagnosing Expertise: Human Capital, Decision Making, and Performance Among Physicians," *Journal of Labor Economics*, 35:1 (2016), 1–43.

Currie, Janet, and W. Bentley MacLeod, "Understanding Doctor Decision Making: The Case of Depression Treatment," *Econometrica*, 88:3 (2020), 847–878.

Currie, Janet, W. Bentley MacLeod, and Jessica Van Parys, "Provider Practice Style and Patient Health Outcomes: The Case of Heart Attacks," *Journal of Health Economics*, 47 (2016), 64–80.

Cutler, David, *The Quality Cure: How Focusing on Health Care Quality Can Save Your Life and Lower Spending Too*, (Oakland: University of California Press, 2014).

Cutler, David M., Jonathan S. Skinner, Ariel Dora Stern, and David Wennberg, "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending," *American Economic Journal: Economic Policy*, 11:1 (2019), 192–221.

Dobbie, Will, Jacob Goldin, and Crystal S. Yang, "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, 108 (2018), 201–240.

Doyle, Joseph J., S. M. Ewer, and T. H. Wagner, "Returns to Physician Human Capital: Evidence from Patients Randomized to Physician Teams," *Journal of Health Economics*, 29, (2010), 866–882.

Doyle, Joseph J., 2011. "Returns to Local-Area Health Care Spending: Evidence from Health Shocks to Patients Far from Home," *American Economic Journal: Applied Economics*, 3:3 (2011), 221–43.

Doyle, Joseph J., John A. Graves, and Jonathan Gruber, "Evaluating Measures of Hospital Quality: Evidence from Ambulance Referral Patterns," *Review of Economics and Statistics,* (forthcoming).

Doyle, Joseph J., John A. Graves, Jonathan Gruber, and Samuel Kleiner, 2015. "Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns," *Journal of Political Economy*, 123:1 (2015), 170–214.

Eichmeyer, Sarah and Jonathan Zhang, "Pathways into Opioid Dependence: Evidence from Practice Variation in Emergency Departments," *American Economic Journal: Applied Economics*, 14:4, (2022).

Epstein, Andrew, and Sean Nicholson, "The Formation and Evolution of Physician Treatment Styles: An Application to Cesarean Sections," *Journal of Health Economics*, 28 (2009), 1126–1140.

Fadlon, Itzik and Jessica Van Parys, "Primary Care Physician Practice Styles and Patient Care: Evidence from Physician Exits in Medicare," *Journal of Health Economics*, 71 (2020) 1-18.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *Quarterly Journal of Economics*, 131:4 (2016), 1681–1726.

Fisher, Elliott S., David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F. L. Lucas, and Etoile L. Pinder, "The Implications of Regional Variations in Medicare Spending: The Content, Quality, and Accessibility of Care. Part 1," *Annals of Internal Medicine*, 138:4 (2003a), 273–287.

Fisher, Elliott S., David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F. L. Lucas, and Etoile L. Pinder, 2003b. "The Implications of Regional Variations in Medicare Spending: The Content, Quality, and Accessibility of Care. Part 2." Annals of Internal Medicine, 138 (4):288–298.

Fletcher, Jason, Leora Horwitz, and Elizabeth Bradley, "Estimating the Value Added of Attending Physicians on Patient Outcomes," *NBER working paper no. 20534*, (2014).

Gowrisankaran, Gautam, Keith Joiner, and Pierre-Thomas Leger, 2017. "Physician Practice Style and Healthcare Costs: Evidence from Emergency Departments," *NBER working paper no. 24155* (2017).

Grytten, Jostein and Rune Sorensen, "Practice Variation and Physician-Specific Effects," *Journal of Health Economics*, 22:3 (2003), 403–418.

Hyslop, Dean R. and Guido W. Imbens, 2001. "Bias from Classical and Other Forms of Measurement Error," *Journal of Business and Economic Statistics*, 19:4 (2001), 475–481.
Jackson, C. Kirabo, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastián Kiguel, "School Effects on Socioemotional Development, School-Based Arrests, and Education Attainment," *American Economic Review: Insights*, 2:4 (2020), 491–508.

Kane, Thomas J. and Douglas O. Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," *NBER working paper no. 14607* (2008).

Koulayev, Sergei, Emilia Simeonova, and Niels Skipper, "Can Physicians Affect Patient Adherence with Medication?," *Health Economics*, 26:6, (2016), 779–794.

Kwok, Jennifer H., 2019. "How Do Primary Care Physicians Influence Healthcare? Evidence on Practice Styles and Switching Costs from Medicare," *Working paper* (2019).

Leung, Lucinda B., Lisa V. Rubenstein, Edward P. Post, Ranak Trivedi, Alison Hamilton, Jean Yoon, Erin Jaske, Elizabeth Yano. "Association of Veterans Affairs Primary Care Mental Health Integration With Care Access Among Men and Women Veterans," *JAMA Network Open*, 3:10 (2020), e2020955.

Molitor, David, "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration," *American Economic Journal: Economic Policy*, 10 (2017), 326–356.

Ravi Prikh, Christopher Manz, Corey Chivers, Susan Regli, Jennifer Braun, Michael Draugelis, Lynn Schuchter, Lawrence Shulman, Amol Navathe, Mitesh Patel, Nina O'Connor. "Machine Learning Approaches to Predict 6-Month Mortality Among Patients with Cancer." *JAMA Netw Open.* 2(10):e1915997 (2019), doi: 10.1001/jamanetworkopen.2019.15997.

Simeonova, Emilia, Niels Skipper, and Peter Thingholm, "Physician Health Management Skills and Patient Outcomes," *NBER working paper # 26735*, (2020).

Trivedi, Ranak B., Edward P. Post, Haili Sun, Andrew Pomerantz. "Prevalence, Comorbidity, and Prognosis of Mental Health Among US Veterans," *American Journal of Public Health*, 105:12 (2015), 2564–2569.

Van Parys, Jessica, 2016. "Variation in Physician Practice Styles Within and Across Emergency Departments," *PloS One*, 11:8 (2016) 1-19.

Wagner, Todd H., Shuo Chen, and Paul G. Barnett, "Using Average Cost-Methods to Estimate Encounter-Level Costs for Medical-Surgical Stays in the VA," *Medical Care Research Review* 60:3 (2003), Suppl., 15S-36S.

Figures and Tables

Figure 1: Histogram of PCP Effectiveness Metrics
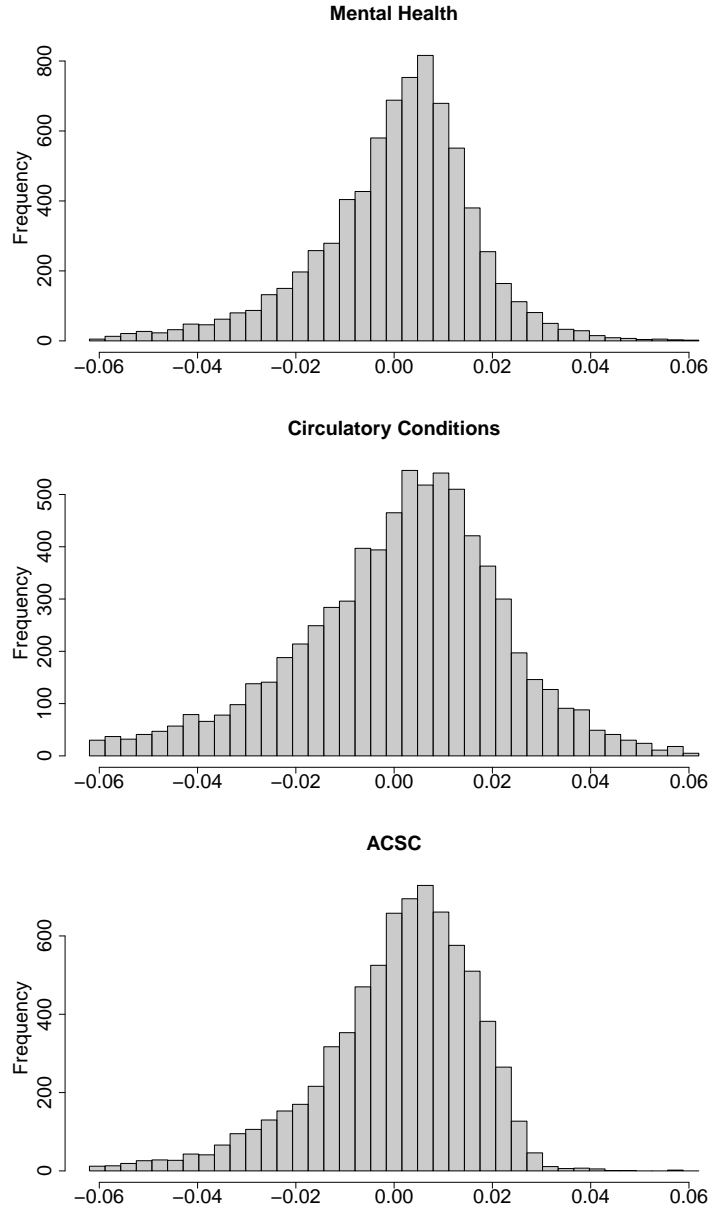
Notes: This figure plots the distribution of our 7,548 PCPs measured by the three dimensions of patient outcomes. The effectiveness measures are empirical Bayes jackknife value-add measures of mental health, circulatory condition, and ambulatory care sensitive conditions (ACSC). We construct these measures by first obtaining residualized (jackknife) value add measures for each provider-year, residualizing for year by month; primary care clinic; day of week of the initial visit; and bins for the number of days between the veteran's desired date for a first appointment and the date of the actual appointment along with controls for race, five-year age bins, marital status, priority groups, Medicare/Medicaid beneficiary status, prior year mental health, circulatory, and ACSC hospitalizations, disability/unemployable status, era of service, and exposure to Agent Orange or radiation. Next, we apply empirical Bayes to obtain a single provider value-add per patient. Finally, we average all the provider's cases to arrive at a effectiveness measure per provider.

Figure 2: Balance

(a) Balance: Veteran Observables Do Not Predict Provider Effectiveness
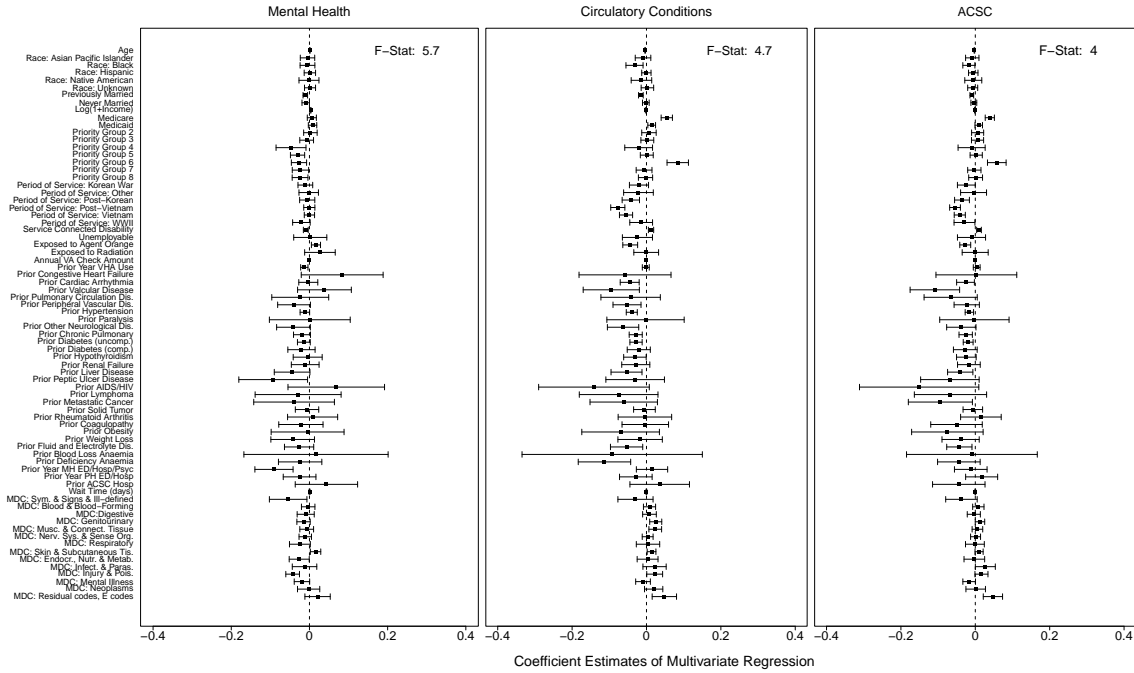
(b) Effect of Examiner IV on Actual and Predicted ED Visits and Hospitalization
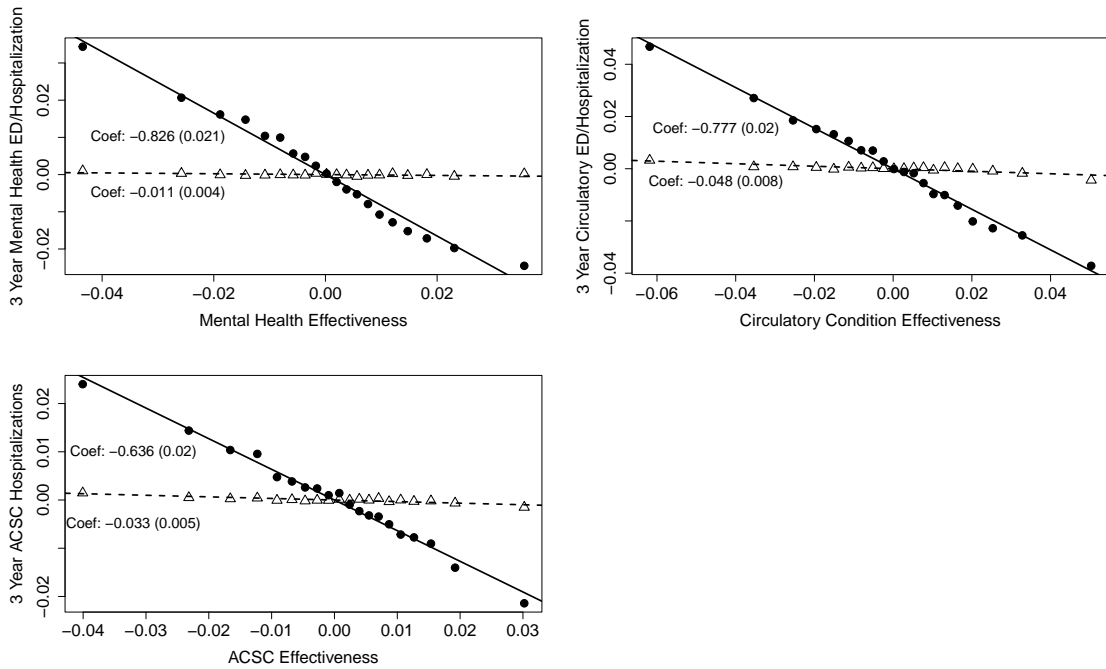
Notes: This figure tests for quasi-random assignment of new patients to PCPs. Panel a estimates regressions of our residualized (leave-out) effectiveness measure on a set of observables including (jointly) patient demographics, military history, Elixhauser comorbidities, prior year utilization, and adverse health outcomes and major diagnostic categories observed on their initial visit. Estimated regression coefficients and associated 95% confidence intervals (constructed from robust standard errors clustered at the clinic-level) are shown. All three metrics are standardized and constructed from residuals taken after controlling *only* for clinic, year-month, day of week, and days to desired appointment date fixed effects. The joint F-statistics are reported. The number of observations is 802,777 for all. Panel b plots actual and predicted 3 year mental health ED and hospitalizations; circulatory ED and hospitalizations; and ACSC hospitalizations against effectiveness ventiles. The solid circles represent actual outcomes, residualized *only* for clinic, year-month, day of week, and days to desired appointment date fixed effects against twenty equally-spaced effectiveness bins. The hollow triangles represent predicted outcomes using veteran characteristics (from the right-hand side of panel a), residualized *only* for clinic, year-month, day of week, and days to desired appointment date fixed effects against the same bins. The linear relationship (i.e., the simple linear regression coefficient and standard error) between the dependent variable and provider effectiveness using the underlying non-binned data are also displayed.

Table 1: Summary Statistics for Patients

	Mean	Assigned PCP: Circulatory Terciles		
		Bottom	Middle	Top
Age	55.3	55.6	55.4	54.8
Asian Pacific Islander	1.7	1.6	2.0	1.5
Black	13.2	13.4	12.8	13.2
Hispanic	5.8	5.6	6.1	5.8
Native American	0.7	0.7	0.8	0.7
White (non-Hispanic)	74.3	74.4	74.0	74.4
Currently Married	57.7	57.4	58.3	57.5
Previously Married	29.1	29.6	28.7	28.9
Never Married	13.2	13.1	13.0	13.6
Income	44,413	44,440	44,403	44,396
Medicare	29.7	29.7	30.5	28.8
Medicaid	5.4	5.5	5.3	5.4
Period of Service: Korean War (1950-55)	5.5	5.5	5.4	5.4
Period of Service: Vietnam War (1961-75	41.4	42.2	41.2	40.8
Period of Service: Gulf War Era (1990+)	30.9	29.6	31.3	31.9
Period of Service: Other	22.3	22.7	22.1	22.0
Any Service Connected Disability	50.7	49.9	51.4	50.8
Deemed Unemployable	0.3	0.3	0.3	0.3
Agent Orange Exposure	16.8	17.1	16.4	16.8
Other Radiation Exposure	0.3	0.3	0.3	0.4
Annual VA Check Amount	1,694	1,669	1,780	1,633
Any Prior Year VHA Care	13.2	13.5	13.1	13.0
Prior Year MH ED/Hosp	0.3	0.4	0.3	0.3
Prior Year Circulatory ED/Hosp	0.4	0.5	0.4	0.4
Prior Year ACSC Hosp	0.1	0.1	0.1	0.1
Wait Time (days)	5.6	5.6	5.6	5.6
Initial Diagnosis: Circulatory	25.5	26.4	25.1	25.0
Initial Diagnosis: Endocrine, Nutritional, & Metabolic	14.8	14.9	14.9	14.7
Initial Diagnosis: Musculoskeletal & Connective Tissue	13.6	13.3	13.5	14.0
Initial Diagnosis: Mental	7.5	7.6	7.4	7.5
Initial Diagnosis: Respiratory	3.8	3.8	3.8	3.8
Initial Diagnosis: Other	34.8	34.0	35.3	35.0
Relationship Length with PCP (days)	693	679	702	697
N=	802,777			

Notes: This table presents the baseline summary statistics for our baseline sample of new veteran health benefit enrollees, and those who are assigned to various PCPs, classified by their circulatory condition effectiveness.

Table 2: Effects of a One Standard Deviation Change in an Effectiveness Measure on Mental Health, Circulatory Condition and ACSC Outcomes

	Mental Health			Circulatory Conditions			ACSC		
One SD of..	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mental Health	−1.56***			−0.80***			−0.37***		
	(0.04)			(0.04)			(0.02)		
Circulatory Conditions		−0.58***			−1.96***			−0.67***	
		(0.04)			(0.04)			(0.03)	
ACSC			−0.48***			−1.15***			−1.12***
			(0.03)			(0.04)			(0.03)
FE + Controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (%)		4.97			7.37			2.52	
Observations		802,777			802,777			802,777	

Dependent variable: (×100)

Notes: This table reports the regression output of regressions of each effectiveness measure on mental health ED and hospitalizations, circulatory condition ED and hospitalizations, and ACSC hospitalizations. All regressions include clinic, year-by-month, day of week, bins for days between desired and actual appointment date, race, five-year age bins, marital status, priority groups, Medicare/Medicaid beneficiary status, prior year mental health, circulatory, and ACSC hospitalizations, disability/unemployable status, era of service, and exposure to Agent Orange or radiation. Coefficient estimates are scaled by 100, and robust standard errors are clustered at the clinic-level. *p<0.1; **p<0.05; ***p<0.01.

Table 3: Impacts of PCP Effectiveness Metrics on Mortality and Cost

Panel A. Mortality and Cost

One SD of..	*Dependent variable:* (×100)		
	3-Year Mortality	Log 1Y Avg Cost	Log 3Y Avg Cost
	(1)	(2)	(3)
Mental Health	−0.21***	−4.50***	−4.43***
	(0.03)	(0.46)	(0.39)
Circulatory Conditions	−0.20***	−5.40***	−5.23***
	(0.03)	(0.44)	(0.39)
ACSC	−0.23***	−2.85***	−2.48***
	(0.03)	(0.50)	(0.43)
Mean Dep. Var.	5.50%	$4,275	$12,120
Observations	802,777	788,743	758,655

Panel B. Causes of Death

One SD of..	*Dependent variable:* 3Y Mortality (×100)					
	Cancer	Heart	Suicide	External Causes	Lower Respiratory	Cerebro -vascular
	(1)	(2)	(3)	(4)	(5)	(6)
Mental Health	−0.050**	−0.043***	−0.012***	−0.026***	−0.006	0.006
	(0.014)	(0.012)	(0.004)	(0.007)	(0.007)	(0.006)
Circulatory Conditions	−0.046***	−0.042***	−0.005	−0.012*	−0.008	−0.006
	(0.016)	(0.013)	(0.004)	(0.008)	(0.006)	(0.006)
ACSC	−0.063***	−0.049***	−0.006	−0.019***	−0.008	−0.009
	(0.015)	(0.014)	(0.004)	(0.009)	(0.008)	(0.006)
Mean Dep. Var. (%)	1.48	1.08	0.09	0.30	0.31	0.20
Observations	802,777	802,777	802,777	802,777	802,777	802,777

Notes: This table reports the regression output of regressions of mortality and cost. Panel A displays 3-year all-cause mortality, log of one plus one-year average cost, and log of one plus three-year average cost outcomes and panel B displays 3-year cause of death specific mortality. Cancer, heart disease, lower respiratory, and cerebrovascular diseases are selected as the five most common causes of death among veterans (and the American population more generally). External causes of death include suicides, overdoses and poisonings, and accidents. All regressions include clinic, year-by-month, day of week, bins for days between desired and actual appointment date, race, five-year age bins, marital status, priority groups, Medicare/Medicaid beneficiary status, prior year mental health, circulatory, and ACSC hospitalizations, disability/unemployable status, era of service, and exposure to Agent Orange or radiation. Coefficient estimates are scaled by 100, and robust standard errors are clustered at the clinic-level. The sample in columns 2 and 3 are constrained such that the veteran is alive for the outcome period. *p<0.1; **p<0.05; ***p<0.01.

Table 4: Encounters, and Referrals and Testing

Panel A. Encounters

One SD of..	Dependent variable: Number of Encounter Days					
	All VA	Primary Care	MH	Emerg.	Inpat.	Medicare
	(1)	(2)	(3)	(4)	(5)	(6)
Mental Health	−0.395***	−0.108***	−0.106***	−0.029***	−0.016***	0.011
	(0.025)	(0.010)	(0.009)	(0.002)	(0.001)	(0.008)
Circulatory Conditions	−0.407***	−0.112***	−0.016***	−0.030***	−0.017***	0.008
	(0.024)	(0.010)	(0.008)	(0.002)	(0.001)	(0.009)
ACSC	−0.255***	−0.066***	−0.004	−0.022***	−0.015***	0.002
	(0.028)	(0.011)	(0.008)	(0.002)	(0.001)	(0.009)
Mean Dep. Var.	13.4	5.0	1.3	0.26	0.095	1.47
Observations	802,777	802,777	802,777	802,777	802,777	238,386

Panel B. Referrals and Testing

One SD of..	Dependent variable: (×100)				
	Referrals (indicator)			Testing (Counts)	
	Any	MH	Cardiology	Lab Panels	Imaging
	(1)	(2)	(3)	(4)	(5)
Mental Health	−0.54***	−0.63***	−0.29***	−0.13***	−0.03***
	(0.09)	(0.09)	(0.07)	(0.02)	(0.01)
Circulatory Conditions	−0.67***	−0.31***	−0.67***	−0.27***	−0.06***
	(0.08)	(0.09)	(0.08)	(0.02)	(0.01)
ACSC	−0.38***	−0.16***	−0.42***	−0.18***	−0.04***
	(0.10)	(0.10)	(0.07)	(0.03)	(0.01)
Mean Dep. Var.	74.8%	20.9%	7.0%	8.5	1.5
Observations	802,777	802,777	802,777	802,777	802,777

Notes: This table reports the regression output of the number of encounter (days) a veteran has in the first year and the number of regerrals and testing in the first year, on our PCP effectiveness metrics. In panel A, columns 1 to 5 report the number of encounter days for its respective type of care. Column 6 reports the number of Medicare encounter days (across all settings and modalities) for veterans over the age of 65 at assignment. Panel B reports regression output of referrals (any, MH referrals, and cardiology referrals) and testing (number of outpatient lab panels, and imaging) orders on our PCP effectiveness metrics. Referrals are indicators for whether the patient is ever referred in the first year and testing orders are the number of distinct orders in the first year. All regressions include clinic, year-by-month, day of week, bins for days between desired and actual appointment date, race, five-year age bins, marital status, priority groups, Medicare/Medicaid beneficiary status, prior year mental health, circulatory, and ACSC hospitalizations, disability/unemployable status, era of service, and exposure to Agent Orange or radiation. robust standard errors clustered at the clinic-level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Panel A. Mental Health Guidelines

	Dependent variable: (×100)		
	Depression	PTSD	SUD
One SD of..	(1)	(2)	(3)
Mental Health	0.04	−0.07	0.03
	(0.06)	(0.10)	(0.06)
Circulatory Conditions	0.14**	0.27***	0.11*
	(0.05)	(0.10)	(0.06)
ACSC	0.13**	0.26***	0.12*
	(0.06)	(0.10)	(0.07)
FE + Controls?	Yes	Yes	Yes
Mean Dep. Var. (%)	96.9	94.2	96.5
Observations	670,060	670,060	670,060

Panel B. Physical Health Guidelines

	Dependent variable:(×100)				
	CRC	HCV	HIV	Flu	Tobacco
One SD of..	(1)	(2)	(3)	(4)	(5)
Mental Health	0.005	−0.37	−0.31*	−0.12	−0.04
	(0.22)	(0.25)	(0.16)	(0.09)	(0.04)
Circulatory Conditions	−0.33	−0.90***	−0.21	−0.33***	0.06
	(0.21)	(0.23)	(0.18)	(0.10)	(0.04)
ACSC	0.03	−0.74***	−0.06	−0.07	0.11**
	(0.19)	(0.26)	(0.19)	(0.10)	(0.05)
FE + Controls?	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (%)	49.3	47.3	22.3	45.3	97.1
Observations	437,203	738,225	532,853	802,777	802,777

Notes: This table reports the relationship between adherence to annual physical and mental health guidelines set forth by the VHA and our PCP effectiveness metrics. Mental health screens are for depression, PTSD, alcohol and substance use disorder via mental health questionnaires and begin after 2008. The sample is restricted to new enrollees after 2008. Physical health adherence for colorectal cancer screens (for patients between the ages of 50 and 75), hepatitis C screens (patients under the age of 80), HIV screens (patients under the age of 65), flu immunizations, and tobacco screens are our physical health margins. All dependent variables are indicators for screenings in the first year and samples are restricted to age groups relevant to each guideline. See text for details on the construction of each. All regressions include clinic, year-by-month, day of week, bins for days between desired and actual appointment date, race, five-year age bins, marital status, priority groups, Medicare/Medicaid beneficiary status, prior year mental health, circulatory, and ACSC hospitalizations, disability/unemployable status, era of service, and exposure to Agent Orange or radiation. robust standard errors clustered at the clinic-level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table 6: Provider Demographics and Characteristics

		Dependent variable: One SD of		
	Weighted Mean	Mental Health	Circulatory Conditions	ACSC
	(1)	(2)	(3)	(4)
Female	0.46	0.022	−0.049	−0.048
		(0.036)	(0.038)	(0.030)
Age: 35-44	0.24	−0.063	0.060	0.122**
		(0.073)	(0.070)	(0.058)
Age: 45-54	0.35	−0.003	0.125*	0.174**
		(0.063)	(0.081)	(0.074)
Age: 55+	0.36	0.039	0.155*	0.158**
		(0.073)	(0.081)	(0.074)
Part-Time	0.32	0.128	0.240***	0.505***
		(0.079)	(0.083)	(0.079)
Primary Care & Mental Health Integration	0.095	1.211***	0.884***	1.055***
		(0.178)	(0.157)	(0.079)
Patients Per Day	12.1	0.009	0.016***	0.027***
		(0.007)	(0.005)	(0.006)
New Patients Per Year	23.6	0.014***	0.020***	0.009***
		(0.002)	(0.002)	(0.002)
Physician	0.76	0.004	−0.095**	−0.077
		(0.047)	(0.047)	(0.049)
Observations	-	7,544	7,544	7,544

Notes: This table reports the output of regressing each of our PCP effectiveness metrics on provider observables, controlling for clinic fixed effects. Regressions are weighted by each PCP's number of new patients and robust standard errors are clustered at the clinic-level. Age is a weighted average of age at each new patient assignment, part time indicator is the fraction of the years where the provider works fewer than 240 days during the calendar year, and primary care-MH integration is the fraction of each PCP's mental health outpatient visits that are joint with their primary care team in the same clinic. Missing characteristics are coded as a separate category within each variable and not displayed. *p<0.1; **p<0.05; ***p<0.01.