How do mental health treatment delays impact long term mortality?

Sydney Costantini*

December 2024

Abstract

With a growing mental health crisis and a shortage of behavioral health specialists, those seeking mental health treatment often face long wait times to obtain care. I study how clinic congestion affects mortality for veterans experiencing mental health emergencies. I find that longer waiting times make it more likely that patients miss their follow up mental health visit, consequently increasing the probability that they permanently disengage from care. A one standard deviation increase in wait time between the ED visit and follow up appointment date (11.7 days) increases two-year mortality by about 1.5%.

The mental health crisis is increasingly a point of concern in public discourse. Those suffering from severe mental illness die an average of 10 to 20 years earlier than the general population (Liu et al. 2017). Mental health treatment has been increasingly destigmatized in recent years, with the number of patients seeking new therapy continuing to rise (APA 2021). Medical research has shown that behavioral therapy and medication are successful at achieving their primary goal – reducing depressive symptoms and improving general mental health (Cuijpers et al. 2016, Gartlehner et al. 2017, Holmes et al. 2018) – but what if patients most in need have difficulty accessing care?

There continue to be reports of a major shortage of mental healthcare providers across geographic regions in the US, often resulting in long appointment delays for patients seeking care (Weiner 2022).

*UC Berkeley, sydney_costantini@berkeley.edu. I would like to thank David Card and David Chan for incredible advising on this project, Sydnee Caldwell, Kaveh Danesh, Amy Finkelstein, Hilary Hoynes, Jonathan Kolstad, Erzo Luttmer, Ziad Obermeyer, Christopher Walters, and Danny Yagan for helpful comments and feedback, and Michael Finneran for his clinical expertise. I gratefully acknowlege financial support from the Center for Labor Economics at UC Berkeley. This project was approved by the Stanford IRB and the Department of Veterans Affairs, under the supervision of David Chan.

Provider shortages are exacerbated by a healthcare system that has traditionally put physical health before mental health. In 2008, the US government passed the Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act, intended to address treatment disparities between mental and physical health conditions, especially related to insurance coverage. Nevertheless, patients continue to report more difficulty accessing in network providers for mental health treatment, resulting in either higher copayments or longer wait times (Xu et al. 2019, Caron 2021). Given this context, waiting times for mental health treatment are of notable policy concern. The National Committee for Quality Assurance (NCQA) collects data on mental health wait times following ED visits – the setting of my analysis – considering them to be a key quality metric.

Issues surrounding access to and engagement with mental health services are particularly pertinent to veterans, who suffer in high rates from PTSD, substance use disorder (SUD), and other related issues. Of all government sponsored healthcare programs, the Department of Veterans Affairs (VA) provides the most comprehensive mental health services; treatment is free for the majority of low-income veterans. Despite comprehensive insurance covering mental health, veterans still report difficulty accessing behavioral therapy, notably due to long wait times (NASEM 2018). With its detailed demographic data and large patient population vulnerable to mental health crises, VA data provide a unique opportunity to study the impact of mental health treatment delays on long term mortality.

In this paper, I study patients who arrive at any Emergency Department (ED) across the VA system with a mental health-related condition. I collect wait times for these patients, defined as the difference between their ED visit date and their first follow-up psychology or psychiatry appointment. A patient's actual wait time is likely to be correlated with other characteristics affecting survival – notably, the patient's severity level and willingness to schedule future appointments. Thus, in order to isolate the causal effect of treatment delays, I do not use a patient's own wait time directly. Instead, I construct an *exogenous* measure for each patient: the "leave-out" mean wait time – the average wait time of all other patients who arrive in same two-week period and hospital as a given patient and were assigned the same type of follow up appointment (psychiatry, psychology, substance use disorder clinic, etc.).[1] As mean wait time is assigned conditionally based on hospital, time, and appointment type, I control for time in two-week intervals, mental health clinic type, and hospital $\times$ year-month fixed effects in all analyses.

Conceptually, one can think of mean wait time as a measure of mental health clinic congestion, which affects both providers' and patients' ability to schedule timely follow up visits. I first show that mean wait time is highly correlated with a patient's actual wait time, even after controlling for the design controls. Moreover, using leave-out mean wait time rather than a patient's own wait time eliminates the potential concerns that either 1) patients play some role in appointment scheduling, or 2) physicians schedule follow up appointments based on a patient's risk level, rather than just appointment availability. I find that mean wait time has a very small and statistically insignificant correlation with predicted patient mortality. In other words, patients who have higher and lower predicted mortality are equally likely to be assigned a high mean wait time, indicating quasi-random assignment.

As mean wait time is uncorrelated with patient risk, my measure of congestion enables me to isolate the causal impact of mental health treatment delays on long term mortality. I show that a one standard deviation increase in mean wait time increases two-year mortality by 0.113 percentage points (s.e. = 0.0041), or about 1.5% of baseline mortality risk. This effect is highly robust to controlling for a detailed range of patient characteristics – including demographics, ICD-9 codes from the ED visit, and prior mental health utilization, among others. Moreover, the mortality effect is long-lasting, persisting for at least five years after the index ED visit. The results are almost entirely driven by

[1]The VA often has separate clinics and providers for different types of specialty mental health appointments. Therefore, patient-specific wait times are highly dependent on appointment type.

patients with substance use disorders (SUD) or psychosis, rather than those with anxiety, depression, or PTSD.

How do relatively small treatment delays have a impact on long term mortality? Prior research has shown that those who are placed into timely mental health care are more likely to stay in consistent treatment (Kreyenbuhl et al. 2009). Similarly, I find that my measure of congestion, mean wait time, strongly predicts whether a patient attends their follow up mental health visit: a one standard deviation increase in mean wait time (11.7 days) leads to a 2.0 percentage point reduction in the probability of attending one's follow up appointment (t=24). Thus, I argue that the mortality effect is primarily driven by patients' engagement with mental health treatment.

However, I also consider competing explanations. I first show that mean (mental health) wait time is not correlated with other hospital wait times or mortality of non-mental health patients. These results imply that my measure of mental health clinic congestion only affects health outcomes via mental health care itself, rather than being correlated with overal hospital quality or crowding. Additionally, I find that mental health clinic congestion does not predict the length of follow up appointments, indicating that patients do not receive more rushed care when appointments are delayed.

Nevertheless, I do find some changes in care patterns on the supply side, which may also contribute to the mortality effect. First, patients who arrive when mental health clinics are congested are assigned providers with slightly lower experience. Moreover, physicians do appear to be aware of and respond to clinic congestion. They compensate for appointment delays by prescribing more mental health drugs at the ED visit itself; a one standard deviation increase in leave-out wait time (11.7 days) increases the probability of being prescribed a drug at the ED visit by 1.2 percentage points (s.e.=0.08) or 4.0%. As prescriptions mitigate mortality risk, prescribing patterns at the ED partially counteract the impact of missing future behavioral therapy appointments.

Rather than being the result of a single missed appointment, the mortality effect is likely a product of disengagement from the mental health care system as whole. In fact, patients with higher mean wait times are far more likely to permanently disengage from both outpatient and inpatient mental health care in the long run. Overall, my results suggest that shorter hospital wait times can be vital in nudging patients into appointment attendance, and that missing a single mental health visit can influence long run engagement with therapy and consequently, mortality.

This study contributes to the literature on the effects of mental health treatment. Previous research has considered outcomes such as employment and career earnings (Biasi et al. 2020, Bhat et al. 2022), high school completion and human capital development (Cuellar and Dave 2015, Angelucci and Bennett 2022), crime (Blattman et al. 2022), suicide (Lang et al. 2013), and mortality (Holliday 2018). Most similar to this paper, a recent medical study finds that quasi-random decreases in mental health staffing increases suicide rates among VA patients (Feyman et al. 2022). These results hint at the issues congestion may cause; however, the authors cannot identify the downstream effects of low provider supply. I show that delays to care have a direct impact on whether patients stay engaged with mental health services. Additionally, in investigating all-cause mortality, rather than suicide alone, this study takes a broader look at the way mental health treatment delays can influence one's overall health status.

Moreover, this study contributes to work addressing the efficiency effects of longer wait times in the health care system. In exploiting quasi-random variation in wait times, I build on a similar empirical strategy as that of Williams and Bretteville-Jensen (2022) and Gødoy et al. (forthcoming), who study wait times in SUD treatment and orthopedic surgery, respectively. Wait times are a common rationing mechanism for healthcare; unsurprisingly, there is evidence that delays have negative impacts on health outcomes (e.g., Reichert and Jacob 2018, Gruber et al. 2018, Singh and Venkataramani 2022, Turner et al. 2022). Concerns over rationing care may be particularly relevant in cases where cost-sharing is low, as reduced cost barriers to accessing healthcare may tip demand for healthcare services to exceed supply. Low cost sharing is both relevant to the VA, where care is essentially free for low income veterans, as well as for public, single payer healthcare systems, such as the National Health Service (NHS) in the United Kingdom. Cutler (2002) describes how this creates a trade off between efficiency and equity: although removing financial incentives likely improves access for low income patients, it may lead to longer waiting times and more inefficient healthcare delivery. This study provides additional evidence that longer wait times as a result of clinic crowding worsens health outcomes.

Related to the use of wait time as a rationing tool in healthcare is the broader idea that non-price ordeals affect targeting. In the healthcare system, both longer wait times and higher copays can both potentially act as a demand-reducing mechanism; however, these mechanisms likely deter different types of patients from seeking care. In in a preventative health context, Dupas et al. (2016) actually show a case where increasing nonmonetary costs actually improves targeting. In contrast, Deshpande and Li (2023) find that closing Social Security field offices, thus increasing the nonmonetary costs of applying for disability, has negative effects on targeting. Analagously, in the VA setting, clinic congestion deters those on the margin from seeking care, even though these patients may see large health benefits from attending mental health treatment.

The rest of this paper proceeds as follows: Section I describes the setting and data. Section II describes the empirical approach. Section III presents the results, including examining the validity of the wait time measure, as well as exploring heterogeneity and mechanisms. Section IV concludes.

I. Setting and Data

A. The Department of Veterans Affairs

The US Department of Veterans Affairs (VA) is the only public, fully integrated healthcare system in the US, acting as both an insurer and provider of healthcare. The agency provides low-cost medical care to veterans meeting certain criteria, including those with low incomes and/or a a service-connected disability. Veterans are a unique patient population with specific needs. Notable for this study, veterans suffer in high rates from substance use disorder, post-traumatic stress disorder (PTSD), and other mental health conditions. Moreover, veterans face a suicide risk more than 50% higher than that of the general population (VA Office of Mental Health and Suicide Prevention 2022). Veterans are also more likely to suffer from chronic diseases such obesity, diabetes, and heart disease compared to non-veterans, even when controlling for age (Betancourt 2023).

Mental health treatment is free for all veterans with a behavioral health condition connected to their service; all other veterans pay small copays for treatment. Thus, fees generally are not a barrier for this population. However, wait times for mental health treatment are perceived as a signficant obstacle, with less than half of veterans requesting urgent appointments reporting being able to get appointments when they requested. A 2015 VA policy aims to get all new patients seeking behavioral therapy an appointment within a 30-day time frame (NASEM 2018).

B. Patient encounter, characteristics, and mortality data

The Corporate Data Warehouse, the VA's detailed health record system, contains patient information on all inpatient and outpatient encounters, International Classification of Diseases Clinical Modification, 9th Revision (ICD-9) diagnosis codes connected with each appointment, and demographic characteristics such as date of birth, race, ethnicity, and gender. Relevant to my study design, the VA also records all appointment scheduling requests, and whether these appointments were ultimately cancelled or no-showed.

The primary outcome measure for this study is two-year mortality. Data on mortality from the VA is particularly reliable in that it combines data from the Veterans Benefits Administration (VBA), Centers for Medicare and Medicaid Services (CMS), and Social Security Administration (SSA) for a combined estimated sensitivity of 98%. Since the effect of treatment may take longer to appear than that of other medical interventions, using a longer time horizon to assess mortality is typical of mental health studies (Appleby et al. 1999, Bjorkenstam et al. 2011). Matching on Social Security Numbers, I also link patient data from the VA with cause of death data, coded by ICD-9 or ICD-10 code, from the National Death Index. As intermediate outcomes, I assess prescriptions for mental health drugs as well long term engagement with both mental and physical health services, as falling out of care is a major concern for this population of interest (Nosé et al. 2003, Kreyenbuhl et al. 2009).

II. Empirical approach

A. Sample selection

The primary sample consists of all patients with a VA Emergency Department (ED) visit for a mental health condition between January 1, 2001 and February 28, 2015, allowing me to follow all patients up to five years prior to the start of the COVID-19 pandemic. I define a mental health ED visit as any ED visit assigned a mental health diagnosis code (ICD-9 code). Reasons for visiting the emergency room range from thoughts of hurting oneself to an upper GI bleed due to alcohol use. The most common mental health diagnoses at the ED in my sample are general depressive disorder (19% of cases), general anxiety (16% of cases), alcohol use (14% of cases), and adjustment reaction (10% of cases). Additionally, 39.9% of the sample also has a *physical* health diagnosis connected to their ED visit, the most common of which are hypertension, sleep disturbances, and respiratory distress. Some of these may be directly connected to mental health – for instance, those with anxiety may experience breathing issues or difficulty sleeping.

In order to follow patients over a long time horizon, I use only the first ED mental health visit for each patient. Additionally, I drop patients who have no scheduled mental health visit within 90 days of their initial ED visit, my definition of a follow up appointment.[2] These patients are dropped as a result of the empirical design, where mean wait time (described in Section II.B) is constructed according to follow up appointment type (psychiatry, psychology, etc.). Appendix Table A1 describes patients with no scheduled follow up appointments in more detail, who are have higher rates of SUD and have less prior attachment to VA mental health services. In Table 2 and Appendix Section A3.1, I also perform robustness checks where I add these patients back to the sample using a slightly modified version of the main treatment variable and controls.

Figure 1 describes the sample selection process. The baseline analytical sample consists of 621,289 patients across 128 hospitals.

B. Measuring congestion

To measure mental health clinic congestion, I collect wait times for all patients with a mental health ED visit. I define a patient's wait time as the time between their initial ED visit and their follow up mental health appointment. I define a follow up appointment as the patient's first scheduled encounter with a mental health specialist following their initial ED visit. All appointments within a 90 day window are considered follow up appointments.[3]

While there is likely variation in scheduling practices across clinics, ED doctors work directly

[2]In VA data, outpatient encounters with mental health services are not explicitly linked to ED visits as follow ups. Thus, follow ups are likely measured with some error using this time window definition.

[3]In other words, if the patient's first mental health encounter following their ED visit occurs after 90 days, I consider this visit to not be a true follow up appointment.

with scheduling staff to coordinate follow up appointments. These appointments may be scheduled on site before a patient is discharged from the ED, or they may be scheduled with outpatient scheduling staff via phone following their index visit. The length of time between a patient's ED visit and their follow up appointment date is likely a product of various factors, including provider availability, patient scheduling preferences, and patient severity level. Thus, to construct a measure of clinic congestion orthogonal to a given patient's risk level, I do not use a patient's own wait time. Instead, I group together patients at the same hospital who arrive during the same two-week period and are assigned the same type of follow up appointment (psychology, psychiatry, substance use disorder clinic, etc.). Wait times are appointment specific in that different specialty mental health clinics treat different types of patients, and thus, congestion is specific to appointment type. I assign each patient a mean wait time value, W_{-i}, according to the mean leave-out wait time of all patients in their respective hospital-time-appointment type group, excluding the patient's own wait time. More formally, for patient i and hospital-time-appointment group j, W_{-i} is defined as follows:

$$W_{-i} = \frac{1}{|I_{j(i)}| - 1} \sum_{i' \in I_{j(i)}} \mathbf{1}(i' \neq i) W_{i'} \tag{1}$$

where W_i is a patient's wait time and $I_{j(i)}$ is the set of patients in the same hospital-time-appointment type group as patient i. W_{-i} is a measure of mental health clinic congestion at the time of a patient's arrival to the ED. To more accurately reflect congestion, rather than restricting to the first visit for each patient (as is the case for the primary sample; see Figure 1), I use all mental health ED visits and their corresponding wait times to form W_{-i}. The average cell size of $j(i)$ is 22 patients; mean W_{-i} is 17.7 days (standard dev. = 11.7 days). I standardize W_{-i} in regression analyses so that coefficients can be interpreted as the impact of a one-standard deviation increase in mean wait time.[4]

While a patient's own wait time may be influenced by patient specific-scheduling factors that are correlated with risk level, mean wait time captures potential appointment delays purely due to congestion. Mean wait time is assigned conditionally based on hospital, time, and mental health clinic type. Thus, I control for two-week period, year-month $\times$ hospital fixed effects, and appointment type in all analyses. Additionally, I add four-year age buckets to the baseline controls, necessary to achieve balance (see Section III.A). With these baseline controls, the standard deviation of residualized mean wait time is approximately 7.2 days. Across analyses, I cluster standard errors by hospital-month.

Ultimately, the study design relies on the assumption that conditional on the design controls, particularly the detailed time $\times$ hospital controls, the level of congestion a given mental health clinic experiences is as good as random. More specifically, for a given patient that arrives to the ED when mental health clinic congestion is particularly high, I compare outcomes for this patient with other patients who arrive to the *same* ED in the *same* month (e.g., the Palo Alto VA in November 2013), accounting for patient selection to certain hospitals in particular months. Moreover, I use two-week period fixed effects to adjust for the fact that particular time periods may be systematically busier. Finally, certain appointment types (such as psychiatry) may tend to have both sicker patients and longer wait times. Thus, I control for appointment type indicators to adjust for any systematic correlation between patient selection and type. I further reassure the reader that selection into certain appointment types does not drive the results in Section III.C.i.

[4]Note that I use only patients with an ED mental health visit to form W_{-i}. Although there are many patients with scheduled mental health visits who do not have an ED visit, my definition of wait time relies on calculating the number of days between the ED visit and follow up visit date. Moreover, timing of follow ups for emergent patients are more likely to reflect congestion, whereas visits for non-emergent patients engaged with behavioral therapy are more likely to follow a recurrent schedule.

III. Results

A. Design validity

i. Does mean wait time predict actual wait time?

In order for W_{-i} to be a useful measure of mental health treatment delays, it must be highly correlated with a patient's actual wait time. To assess the relationship between mean wait time and patient-specific wait time, in Figure 2 Panel A, I show a binned scatterplot of the relationship between W_i and W_{-i}, conditional on the baseline controls. I run the following regression to assess the predictive power of the treatment variable:

$$W_i = \pi_1 W_{-i} + \zeta_{1,t(i)} + \phi_{1,a(i)} + \gamma_{1,h(i)m(i)} + \mathbf{X}_i \delta_1 + \epsilon_{1,i} \tag{2}$$

where W_i is actual wait time, W_{-i} is mean wait time, $\zeta_{1,t(i)}$ are two-week period fixed effects,[5] $\phi_{1,a(i)}$ are appointment type fixed effects, $\gamma_{1,m(i)h(i)}$ are hospital $\times$ year-month fixed effects, and $\mathbf{X}_i$ are additional patient demographic controls, which include age buckets (in four-year intervals) in the baseline specification. Appointment types are categorized as follows: general mental health visit (individual, group, or phone visit), psychiatry, psychology, SUD specialty clinic (individual or group), PTSD specialty clinic (individual or group), mental health integrated care, or other specialty clinic type.[6] Excluding reference categories, there are 19,137 hospital-year-month fixed effects, 368 two-week period fixed effects, 10 appointment type fixed effects, and 20 age buckets.[7]

Even with the design controls, which should eliminate non-random correlation between W_i and W_{-i} due to shared hospital, appointment type, and time period, I find that W_i and W_{-i} are highly

[5] These refer to exact two-week period, e.g. the first two weeks of January, 2011.

[6] In Section III.C.i, for simplicity, I group these into general, psychiatry, PTSD, SUD, or other. My baseline results hold using both the more and less granular versions of the appointment type categories (for both the construction of W_{-i} and controls); see Section III.C.i for details.

[7] There are no observations in some hospital-year-month bins due to emergency departments not operating in those locations at particular times.

correlated: a one standard deviation increase in mean wait time (11.7 days) increases actual wait time by about 1.44 days (s.e.=0.046), with a t=32. A patient's own wait time depends on various factors – including patient scheduling preferences and health severity, in addition to congestion levels. Thus, one should not expect mean wait time to perfectly predict patient-specific wait time. Nevertheless, the above analysis demonstrates that W_{-i} is able to pick up mental health crowding that is highly predictive of true patient treatment delays.

ii. Is mean wait time as good as random?

Next, I assess independence, or whether W_{-i} appears to be quasi-randomly assigned, conditional on the baseline controls. Importantly, since the main outcome is mortality, high and low risk patients must be equally likely to be assigned a high mean wait time. Following Chan et al. (2022), I construct a predicted mortality for each patient, $\hat{Y}_i$, using detailed patient characteristics and encounter information available in the VA data. Specifically, $\hat{Y}_i$ is formed using a logistic regression with the following covariates: age, gender, race, ethnicity, 3-digit ICD-9 diagnosis codes associated with the ED visit, high risk for suicide flag,[8] priority status,[9] whether the patient arrived by ambulance, prior Elixhauser comorbidities,[10] mental health utilization in the prior year, including prior inpatient and outpatient visits, no shows, and cancellations, and physical health utilization in the prior year, including prior primary care visits, prior specialty appointments, and prior hospitalizations for physical disease. To assess the quality of my mortality prediction, in Figure 3, I plot coefficients and standard errors when I run a logistic regression of two-year mortality on these characteristics.[11] I report a chi-squared statistic of 72,008 and a psuedo-R^2 of 0.210. As age is controlled for in my baseline analysis,

[8]Suicide coordinators at the VA can assign patients this flag if patients are screened as being high risk for suicide. Patients deemed high risk for suicide receive specialized review and care at the VA (Veterans Health Administration 2008).

[9]Priority status is a system the VA uses to determine eligibility for VA healthcare and copay responsibilities, based on both disability status and income.

[10]This a standard way of grouping health comorbidities (see Elixhauser et al. 1998).

[11]One interesting note here is that Black patients actually have a lower mortality risk than white patients in this population. This finding has been replicated in other work on mortality for patients with severe mental illness (Das-Munshi et al. 2017).

I also report a partial R^2 (after conditioning on age) of 0.0704.[12] Secondly, in Appendix Figure A1, I show a binscatter plot of observed versus predicted mortality. The model is able to predict a mortality risk of over 40% for the top 5% of observations and a mortality risk of over 20% for the subsequent 5%.

I then run the following regression to assess balance:

$$\hat{Y}_i = \pi_2 W_{-i} + \zeta_{2,t(i)} + \phi_{2,a(i)} + \gamma_{2,h(i)m(i)} + \mathbf{X}_i \delta_2 + \epsilon_{2,i} \tag{3}$$

I find no relationship between $\hat{Y}_i$ and Z_i, with $\hat{\pi}_2$ = 0.00014 (0.00011) (t = 1.22). In other words, leave-out wait time is uncorrelated with a patient's risk level, indicating quasi-random assignment.

As the VA flags patients deemed high risk for suicide, a natural alternative balance test is whether this high risk flag is correlated with W_{-i}. Using this flag as the left hand side variable, I again find no relationship with W_{-i}. Additionally, another, broader metric indicative of patient severity is whether a patient has a prior hospitalization. Let I_i be an indicator that a patient was hospitalized in the year prior to the index ED visit (for any cause). With I_i as the dependent variable, I again find that the coefficient on W_{-i} is reverse-signed and insignficant; Appendix Section A2 and Figure A2 provide further details.

B. Mortality effect

i. Baseline effect

I run the following reduced form regression to determine the relationship between W_{-i} and Y_i, or two-year mortality:

[12] Note that one would not expect a very high psuedo-R^2 here, as 2-year mortality is a binary 0/1 variable and mortality predictions are probabilities between 0 and 1.

$$Y_i = \pi_3 W_{-i} + \zeta_{3,t(i)} + \phi_{3,a(i)} + \gamma_{3,h(i)m(i)} + \mathbf{X}_i \delta_3 + \epsilon_{3,i} \tag{4}$$

I find a strong positive relationship, with $\hat{\pi}_3$ = 0.00113 (0.00041). In other words, a one standard deviation increase in mean mental health wait time increases mortality by 0.11 percentage points, or 1.5% of baseline risk. In Figure 2 Panel B, I show binned scatterplots of the balance and reduced form relationships (corresponding to Equations 3 and 4), residualizing by the baseline controls. Not only is the balance coefficient insignficant, but the mortality effect is an order of magnitude larger.

Table 1 Panel A displays the baseline mortality results. In column 1, I show mortality effect estimate only using baseline controls. In columns 2 through 4, I show the estimate when I add additional patient controls, including gender, race, mental health diagnosis codes associated with the ED visit, and prior mental health utilization. Finally, column 5 shows the estimate when I control for characteristics of other patients who form a given patient's value of W_{-i}, including (leave-out) mean age, mean prior visits, and mean predicted mortality. The rationale behind these latter controls is as follows: leave-out mean wait time should be uncorrelated with patient characteristics, which could be violated if similar types of patients arrive at the hospital at similar times. The baseline results are highly stable across a wide range of controls; I discuss additional robustness checks in more detail in Section III.D.

Additionally, in Panel B, I show heterogeneity results by mental health condition, using mental health ICD-9 diagnosis codes attached to the index ED visit.[13] The effects are by and large driven by patients with SUD or psychosis, such as schizophrenia – the two groups that also have the highest baseline mortality risk.

Finally, in Figure 4, I plot the mortality effect over the five years following the ED visit. Given

[13] Visits at the ED are always assigned ICD-9 diagnosis codes. These codes represent the reason for the ED visit, rather than necessarily being a new diagnosis for the patient.

that a course of therapy often lasts several months, intuitively, the coefficient at 30 days is essentially 0 (coeff. = 0.00013 (0.00010)). However, the effect accrues over time, reaching 0.00141 percentage points (s.e.=0.00055) five years after the initial ED visit.

C. Endogeneity concerns

i. Is follow up appointment type endogenous?

As described in Section II.B, W_{-i} conditions on what type of follow up appointment the patient is assigned. Visit types include general mental health, which can be a visit with a social worker or psychologist, psychiatry, and specialized SUD and PTSD clinics. This conditioning is necessary to capture true queue lengths, which can vary substantially by appointment type: the mean wait time for a psychiatry appointment is 24.5 days, while the mean wait time for a SUD specialty appointment is 12.0 days. However, one potential concern with this specification is that appointment type is endogenous in cases where a patient may be appropriate for various types of care.

Since I control for appointment type in the baseline specification, my analysis accounts for the fact that certain types of patients select into certain appointment types, on average. However, the correlation between appointment type and patient severity could still present an issue for the analysis if ED staff adjust referral decisions based on queue length–in other words, if selection into a given appointment type varies with residualized W_{-i}. In this case, quasi-random variation in wait times could also induce variation in the types of patients assigned appointment type a, leading to a correlation between W_{-i} and a patient's health status.

More concretely, if wait times are particularly high for appointment type a in a given hospital-time period, are ED staff more or less likely to assign high risk patients to that appointment type? It is not clear which direction the bias may go: on the one hand, patients with higher severity may be most appropriate for appointment types with high wait times (notably, psychiatry). If this match effect drives the selection of appointment type, high risk patients may be more likely to assigned a longer wait time. On the other hand, if ED staff know that a patient has particularly urgent needs, they may decide to send the patient to a provider with shorter wait times, regardless of that provider's specialty.

To directly test whether there is time-varying selection into appointment type that also varies with W_{-i}, I run the following specification:

$$\mathbf{1}(a_i = k) = \beta_1 \tilde{W}_{k,-i} + \beta_2 \hat{Y}_i + \beta_3 \tilde{W}_{k,-i} \times \hat{Y}_i + \epsilon_i \tag{5}$$

where $\mathbf{1}(a_i = k)$ is an indicator for patient i choosing appointment type k, $\tilde{W}_{k,-i}$ is the residualized mean wait time for appointment type k (for all patients arriving in the same hospital-two-week period as patient i),[14] and $\hat{Y}_i$ is predicted mortality. The key coefficient is β_3, which answers the question: if wait times for a given appointment type are (randomly) higher, do high risk patients select into or out of appointment type k?

In Appendix Table A2, I show results for β_1 and β_3 across appointment type categories. Coefficients on β_1 are consistently negative, implying that when the queue is particularly high for a specific appointment type, patients are less likely to be assigned that appointment category. This may either be due to physicians responding to wait times by rushing patients to providers with higher availability, or it may be mechanical: if a patient is referred to multiple providers, they will mechanically be assigned the appointment type he attends first. More importantly, β_3 is either negative or statistically insignificant for all appointment types. These negative coefficients suggest some "reverse-selection": if the queue is especially long for a general mental health visit, high risk patients are assigned to other appointment types. Since I find that high wait times increase mortality, this reverse-sorting suggests

[14] I residualize $W_{k,-i}$ by all baseline controls, excluding appointment type.

that if anything, my baseline results are an underestimate of the true effect of treatment delays.[15]

Nevertheless, to ensure the reader that selection into appointment type is not an important driver of my results, in Table 2 Panel A, I show results for two alternative versions of the baseline specification, both of which are plausibly more exogenous. First – in the "saturated" specification – I control for hospital × appointment type and appointment type × two-week period, accounting for more granular selection into appointment type within hospital and time period. Second, as a truly exogenous measure, I create a new, predicted measure of congestion – $\hat{W}_{-i}$ – which does not rely on assigned appointment type at all. Instead, I predict appointment type for these patients using a multinomial logit model. I run predictions hospital-by-hospital, as different hospitals are likely to have different assignment mechanisms to appointment types. Within each hospital, I use the following characteristics to predict appointment type: an indicator for the patient having a prior mental health visit, indicators for most common prior mental health visit type (e.g., if the patient saw the psychiatrist twice in the last year and no other mental health specialist, he would receive a psychiatrist indicator), and indicators for ICD-9 diagnosis code. Using this model, I extract $\mathbf{p}_i$, a vector of the probabilities of being assigned each appointment type.

Then, I use a given patient's vector of predicted appointment types to construct a weighted average predicted leave-out wait time, $\hat{W}_{-i}$, as:

$$\hat{W}_{-i} = \sum_a p_{a,i} \bar{W}_{a,h(i)t(i)} \tag{6}$$

where $p_{a,i}$ is the probability that patient i is assigned appointment type a and $\bar{W}_{a,h(i)t(i)}$ is the (leave-out) mean wait time for patients arriving at the same hospital and in the same two-week period as patient i who are assigned appointment type a.

[15]For simplicity, Table A2 shows results separated into five major appointment categories: general, psychiatry, PTSD, SUD, and other. These results also hold qualitatively with the more granular appointment type categories used in the main specification, as described in Section III.A.

Both the aforementioned specifications remove some of the predictive power of the congestion measure. Thus, to improve power in these specifications, in lieu of hospital × year-month controls, I use slightly less granular hospital × year × quarter controls. Additionally, to lessen the computational intensity of the prediction, I group appointment types into major categories – general, psychiatry, SUD, PTSD, and other. In Table 2 column 1, I show the baseline mortality result. In column 2, I show the baseline result when I use the slightly less granular hospital × time and appointment type categories; the coefficient is nearly identical.[16] Finally, in columns 3 and 4, I show the saturated and predicted specifications, respectively. Although these specifications are slightly noisier, they yield coefficients highly aligned with the baseline results.

ii. Do patients have discretion over which VA hospital they visit?

Another potential concern with the main regression specification is that patients may use discretion in which hospital they visit. Therefore, conditioning on hospital in the construction of W_{-i} may be an issue if hospital choice is also endogenous to crowding. However, in constructing W_{-i}, hospitals are only divided into broad geographic regions (such as the Cleveland VA), known as VA stations. These stations are usually comprised of a main medical center with an ED, as well as several affiliated outpatient clinics. Patients are highly unlikely to travel to an entirely different geographic region for medical appointments, especially for an emergency visit. Thus, one should think of assigned hospital as the general geographic catchment area for the patient (e.g., Los Angeles, Chicago, etc.), for which there is an associated ED and outpatient clinics; W_{-i} captures the mean mental health clinic congestion in that region. Secondly, I construct W_{-i} based on ED location. Therefore, if initial ED facility choice is unrelated to congestion, even if patients adjust which outpatient clinic they attend in response to wait times, W_{-i} will still be orthogonal to this future optimization.

[16] I also reconstruct W_{-i} using these less granular appointment type categories in this specification.

D. Additional robustness

In Table 2 Panel B, I consider whether sample selection choices are driving my results. In constructing the sample (described in Section II.A), I exclude patients with no scheduled follow up visit, as W_{-i} depends on type of appointment scheduled. To eliminate concern that dropping these patients has implications for my findings, in column 2, I add these patients back to the sample, increasing the sample size to 935,843. This larger sample has a somewhat higher mean two-year mortality of 9.6%. I provide more details on sample characteristics in comparison to the baseline sample in Appendix Section A1.2 and Table A1.

Since patients with no scheduled follow up appointment do not have an assigned appointment type, I let $W_{-i} = \hat{W}_{-i}$ for these patients, as described in Section III.C.i.[17] One can think of $\hat{W}_{-i}$ as a weighted average mean wait time, weighted by the probability of the patient being assigned a particular appointment type. As appointment type is also a baseline control in my main analysis, in this specification, I control for either appointment type directly (for the baseline sample) or the predicted probability of being assigned each appointment type (for patients with no appointment type). In Table A4, I also show additional robustness checks using this larger sample; results are very similar to those of the baseline analysis sample.

Moving to column 3, about 34% of the sample is immediately admitted as an mental health inpatient following the ED visit. These patients are likely to see a mental health specialist during their inpatient stay (which is considered to be their follow up appointment in the main sample specification). However, due to inconsistencies in the definition of follow up between these patients and others, I modify the definition of a follow up appointment (in both the definition of W_{-i} and regressions) to only include true outpatient encounters. In this specification, I essentially consider one's inpatient stay to be a "long" ED visit. Following this logic, I define a patient's first outpatient mental health

[17] W_{-i} for observations also in the baseline sample are held fixed in these analyses.

visit following discharge as their follow up appointment.[18]

Relevant to inpatient stays, a small literature touches on the implications of the decline in inpatient psychiatric beds overtime (Gibbons et al. 2017, McBain et al. 2022). In hospitals across the US, including at the VA, mental health care has primarily shifted from inpatient to outpatient care, and the consequent decline in beds has been of some policy concern (Raphelson 2017). Relatedly, in this analysis, higher psychiatric capacity could potentially shift patients from outpatient to inpatient care. However, I find that excluding inpatient mental health stays has little effect on the reduced form coefficient, implying that this mechanism is not a primary driver of my results.

In column 4, I remove patients with non-VA insurance, which I describe in more detail in Section III.E.ii. Finally, in Appendix Figure A3 and Table A5 I perform various additional robustness checks of interest. These include adding additional control variables and considering alternative forms of clustering, among others. I find that the mortality effect is highly robust.

E. Follow up appointment attendance

i. Impact of treatment delays on follow up appointment attendance

Why do mental health treatment delays have a long term impact on mortality? One explanation is that patients with severe mental illness often have difficulty staying engaged with treatment, which can be exacerbated by treatment delays. In a meta-analysis, Nosé et al. (2003) find that an estimated 24% of patients with psychosis do not attend mental health visits as scheduled. In the VA setting, Fischer et al. (2008) show that in a cohort of veterans with schizophrenia or bipolar disorder, almost one quarter of the cohort drop out of care for over a year during a 5-year follow up window. Similarly, attrition from mental health care is a major concern among those in my sample. While all patients in my sample have a *scheduled* follow up visit, the no show or cancellation rate for these appointments

[18] In this specification, wait time is more naturally defined as the difference between the discharge date and follow up appointment date for those with inpatient stays.

is about 37%.

In a population who has difficulty staying engaged with care, small system-based nudges have been associated with improved attendance. Specifically, Klinkenberg and Calsyn (1996) find that minimizing wait time to the first appointment following ED discharge is associated with higher engagement with therapy – the exact setting of my analysis. This pattern appears to hold in my sample as well: in Figure 5 Panel A, I show the CDF of actual wait times, separating patients who do and do not attend their follow up appointment. Those who attend their appointments tend to have shorter wait times on average.

To formally evaluate whether appointment delays have a causal effect on appointment attendance, I assess whether mean wait time is correlated with the probability patients attend their follow up mental health visit. I run the following regression:

$$\mathbf{1}(V_i = 0) = \pi_4 W_{-i} + \zeta_{4,t(i)} + \phi_{4,a(i)} + \gamma_{4,h(i)m(i)} + \mathbf{X}_i \delta_4 + \epsilon_{4,i} \tag{7}$$

where V_i is an indicator for attending a follow up mental health appointment. I find a large and significant relationship, with $\hat{\pi}_1 = 0.0203$ (0.0008), suggesting that a one standard deviation appointment delay (11.7 days) increases the probability of missing one's follow up visit by 2.0 percentage points or 5.6% (t=24). Figure 5 Panel B shows a binned scatterplot of the probability of attending a follow up mental health visit versus mean wait time, or W_{-i}, residualizing by the baseline controls.

Overall, my analysis suggests that mental health treatment delays have a major impact on follow up appointment attendance, which I consider to be the primary mechanism driving the mortality effect. Thus, one way to think about these results is to consider treatment delays to be an instrument for attending mental health treatment. Conceptually, patients who arrive during busy periods may have longer delays to obtaining a follow up appointment, pushing those on the margin of returning for a follow up visit ("compliers") to disengage from mental health care. In turn, disengagement from mental health services impacts long term mortality. Specifically, in an instrumental variables framework, π_4 from Eq. 6 acts as the first stage, and π_3 from Eq. 4 – or the effect of mean wait time on mortality – acts as the reduced form. Then, uing the jacknifed-instrumental variables estimator, the causal effect of not attending a mental health follow up appointment on two-year mortality is given by $\hat{\beta}_{IV} = \hat{\pi}_3/\hat{\pi}_4$. With the baseline controls, I find that $\hat{\beta}_{IV} = 0.056$ (0.020); in other words, not attending a follow up mental health visit increases 2-year mortality by 5.6 percentage points.

In this setting, both providers and patients may adjust on other margins in response to treatment delays – in other words, there may be violations to the exclusion restriction that make using mean wait time as an instrument imperfect. Therefore, I consider this IV analysis to primarily function as a scaling exercise, giving the reader a sense of how detrimental congestion may be for compliers who disengage from mental health care. In Section III.G, I further evaluate how missing one's first follow up visit affects intermediate outcomes, particularly engagement with long run therapy. First, however, in Section III.F, I consider alternative explanations for the mortality effect.

ii. Could patients be attending mental health services outside the VA?

One important note in interpreting the prior results is that my data only captures patient encounters at VA hospitals. This fact leads to the natural question: could patients be attending mental health services outside the VA? If so, this complicates the counterfactual: patients who miss their appointment at the VA could still be seeing a therapist at another hospital.

Since healthcare in the US is incredibly expensive, it would be unusual and likely infeasible for veterans in my sample to get mental health treatment at another, private hospital without outside insurance. Overall, about 30.3% of the baseline sample has some kind of outside insurance, 55.4% of which is Medicare coverage (Appendix Table A1). Medicare is public insurance provided by the federal government for all those about 65, as well as individuals who qualify due to disability status (including mental illness).

To assess whether the 30.3% of patients with outside insurance could be driving result, I drop these patients from the baseline sample to create a no-insurance subsample. Described in Appendix Table A1, this sample is demographically similar to the main sample with a few differences: 1) the mean age is slightly lower, and 2) two-year mean mortality is lower. These latter two facts are consistent with these patients being non-Medicare recipients, so they are likely to be younger and suffer from less severe mental illness. I find that W_{-i} is also highly predictive of missing a follow up visit for this subsample, with a similarly-sized coefficient of 0.022 (0.001). Additionally, in Table 2 Panel B and Appendix Table A4, I rerun all primary specifications for this sample. The baseline mortality effect is 0.075 p.p. (s.e.=0.044) or 1.4% of the mean for this sample– highly consistent with the baseline effect. On the whole, these results suggest that substitution to the private sector is not an important channel in my analysis.

F. Alternative explanations

i. Are mental health wait times correlated with general crowding?

An alternative explanation for the mortality effect is that congestion is not specific to mental health visits; in other words, mental health appointment delays may also predict delays throughout the hospital or clinic, indicating poorer overall quality. Notably, Feyman et al. (2022) find significant correlation in waiting times across specialties at the VA. Thus, one hypothesis is that those with high W_{-i} receive poorer overall medical care, and this is driving the mortality effect. Note that while certain VA hospitals are likely always more crowded (for example, VAs in dense metropolitan areas) and that certain time periods may also be associated with higher crowding across specialties, my baseline controls guarantee that all congestion comparisons are done *within* hospital-year-month. In other words, I control for the fact that some hospitals are more crowded overall at certain times, and instead isolate congestion not explained by these controls, which I argue reflect some randomness.

Beyond controlling for these baseline differences in crowding across hospital-time, I conduct two additional tests to assess whether there could still be correlation between W_{-i} and general delays in care across other specialties. First, I directly plot wait times for other common types of ED follow up visits versus residualized W_{-i}, residualizing by the baseline controls. This approach allows me to determine whether the "quasi-random" component of mental health clinic congestion is correlated with other clinic wait times. Second, as a more downstream test, I construct a placebo sample of patients who arrive to the ED with non mental health diagnoses and assess whether W_{-i} predicts mortality for this group. If W_{-i} can neither predict wait times directly nor mortality of non mental-health patients, these tests suggest W_{-i} has an impact on mortality only through mental health treatment itself, rather than through care delivered at the hospital in other specialties.

First, I document the five most common specialties visited following an ED visit: primary care, labs or pathology, podiatry, opthamology, and internal medicine.[19] Then, I calculate mean wait times (as the time between the ED visit date and the follow up visit date) for each of these specialties within hospital-two-week bins.[20] Next, I assign these wait times to W_{-i} by matching on hospital-two-week period. Finally, in Figure 6 Panel A, I plot a binscatter of these other wait times versus residualized W_{-i}. I find that all five non-mental health speciality wait times are uncorrelated with residualized W_{-i}, confirming the fact that my measure of mental health clinic congestion does not predict general crowding across specialties.

As a second test, I construct a new analytical sample of 1,377,834 patients who visit the ED

[19] ED visits are identified by an ED visit code plus an ICD-9 diagnosis code, while outpatient follow up visits may only be identified by a specialty visit code (since a patient only receives a diagnosis if they attend their appointment). Therefore, it is not straightforward to link all ED visits to follow up visits using a combination of specialty codes and diagnosis codes. Instead, I use a simpler definition of follow up here: the outpatient appointment must be scheduled on the same day as the ED visit.

[20] This level of variation is equivalent to the variation of W_{-i}, except that W_{-i} also varies by mental health clinic type.

during the same time period as the primary sample, but are diagnosed with a circulatory condition at their ED visit, such as stroke or heart attack.[21] I assign these patients a value of W_{-i} according to their time and hospital of arrival.[22] This alternative sample acts as a placebo test, as the mortality of stroke and heart attack patients should not be affected by mental health clinic congestion. However, if mental health clinic congestion indicates poorer overall hospital quality as a result of crowding, the leave-out wait time measure should also predict higher mortality for this group of patients.

In Figure 6 Panel B, I show a binned scatterplot of W_{-i} versus two-year mortality, residualizing by two week period, year-month $\times$ hospital fixed effects, and age buckets. I find no significant relationship between W_{-i} and mortality for this placebo sample of non-mental health ED patients. Overall, I conclude that W_{-i} is only a measure of mental health clinic congestion and is not correlated with general delays in care.

ii. Do longer wait times affect the quality of mental health care?

While the prior tests suggest that W_{-i} is not associated with non-mental health quality of care, what about quality of care within the mental health clinic itself? For example, if appointments are more delayed due to high congestion, it is possible that physicians are either more rushed or stack more appointments closer together in time. Thus, crowding could lead to poorer outcomes independent of appointment attendance.

To assess this hypothesis, I precisely calculate the length of follow up mental health visits using treating provider identifiers. Specifically, I document the psychologist or psychiatrist assigned to a given patient for their follow up appointment and the precise start time of their visit (up to the minute). Then, I follow the same physician to their next appointment with a new patient and again

[21] Patients in this sample must have a diagnosed mental health condition during the study period in order to be a part of the initial sample construction (in the VA data request). However, these patients arrive to the ED with a circulatory condition, rather than a mental health condition.

[22] As patients in this alternative sample do not have a follow up mental health visit, I group together all appointment types to assign them a mean wait time.

document this visit's start time. I define the difference between these two start times as the follow up appointment length. Using this method, I censor any appointment lengths that are greater than six hours, as I assume this captures the doctor ending and starting a new shift. With censoring, I capture 299,958 appointment lengths (of 394,292 patients who attended their follow up appointment).

The median appointment length is exactly 60 minutes, while the 5th and 95th percentile lengths are 7 minutes and 223 minutes, respectively. This method for calculating appointment lengths is likely noisy: a provider may see multiple patients in the same time window or could have a shift break of only a few hours. Therefore, I assume any extreme outliers are data errors and either 1) focus on the median appointment length within W_{-i} bins or 2) winsorize observations at the 5th and 95th percentiles. In Appendix Figure A4, I show a binscatter of either the winsorized mean or median follow up appointment length versus W_{-i}, residualizing by the baseline controls. I find that appointment length is uncorrelated with W_{-i} in both cases, mitigating some concern that congestion also leads to rushed care.

Finally, one concern with the above analysis is that I can only capture appointment lengths for those who attend their follow up visit. As demonstrated in Section III.E, visit attendance is non-random and is highly correlated with W_{-i}. Thus, one story is that those who did not attend their appointment due to high congestion also would have had a shorter appointment, and I cannot observe this fact in the prior regressions. To address this non-random attrition problem, I use a Lee bounds approach. Since those with high W_{-i} are essentially censored here, Lee (2009) recommends censoring either the left tail or right tail appointment lengths of patients assigned low W_{-i} to construct worst-case scenario bounds.[23] Using median appointment lengths and worst-case bounds, I can still reject appointment lengths shorter than two minutes for those with mean wait times longer by one standard deviation.

[23]Since Lee conceptualizes the problem using a binary control and treatment group, for simplicity, I convert W_{-i} to an indicator for above and below median to censor observations.

Although I do not find evidence of more rushed care, it is possible that crowding changes delivery of mental health treatment in more subtle ways. For instance, perhaps crowding leads to treatment by lower-skill providers or worsens match quality between provider and patient if certain providers fill appointment slots more quickly. I do find some evidence of this phenomenon: as a second test of the quality of mental health care, I match treating providers (for follow up visits) to number of years of experience that provider has at the VA. Using provider experience as the left-hand side variable, I find that a one stand. dev. increase in wait time decreases provider experience by about 0.16 years (s.e.=0.023), or about 1.5% of the mean.[24] Does provider experience contribute to the mortality effect? Although not a causal analysis, in an OLS regression, I find that two-year mortality is uncorrelated with provider experience, conditional on the baseline controls (coeff.=0.000059; s.e.=0.000044). Nevertheless, I cannot rule out that other components of mental health treatment are changing when clinics are more congested. In fact, in the next section, I find evidence that providers marginally adjust treatment decisions in response to crowding. Thus, while I argue that appointment delays primarily affect long term health outcomes via *patient* behavior – by increasing the likelihood of missing their follow up visits – supply-side adjustments may also play a role in contributing to the mortality effect.

iii. Do ED doctors anticipate delays with compensatory behaviors?

What if higher W_{-i} is actually correlated with better mental health care delivery? This could be the case if there is any compensatory behavior; in other words, physicians are aware that mental health treatment delays could lead to poor patient outcomes, and therefore, they are compensating for these delays by modifying other care practices. In this case, lower appointment attendance among those with high wait times may be mitigated by improved care delivery.

[24] Similar to the analysis on appointment durations, I cannot link visits that were no showed or cancelled to a specific provider. Using a conservative Lee bounds approach to correct for attrition, I cannot rule out that treating provider experience is lower or higher for those with high wait times (lower bound: -0.52; upper bound: 0.42).

One potential way to compensate for longer appointment delays is to prescribe mental health medications at the ED visit itself, rather than waiting for a follow up consultation. Overall, 30.2% of patients are prescribed a mental health drug on the day of the ED visit, 54.5% of which are antidepressants. Interestingly, a one standard deviation increase in mean wait time increases the probability of being prescribed a drug by 1.2 percentage points (s.e.=0.08) or 4.0% (Table 3). Specifically, lithium, antidepressants, analgesics, and anticonvulsants all appear to be marginal drugs, with physicians being more likely to prescribe them if one's follow up appointment is delayed.

About 64% of patients were prescribed a mental health drug in the year prior to their index ED visit. Thus, another interpretation of the correlation between mean wait time and increased precriptions are that patients come to the ED to refill or update their prescriptions, and may be especially likely to do so if it difficult to get an outpatient mental health appointment. Notably, among the 30% of patients prescribed a drug at the ED visit, about 61% of prescriptions are for repeat drug classes (e.g., the patient had been prescribed an antidepressant in the prior year and are again prescribed one at the ED). Thus, to shed light on this interpretation, in column 2, I show that around half of the effect comes from prescriptions in entirely new drug classes, which is inconsistent with a patient requesting a refill. Secondly, I drop patients who were prescribed a mental health drug in the prior year from the main reduced form analysis; the resulting mortality effect is 0.00130 (0.00063), highly consistent with the baseline results.

Given the compensatory behavior of physicians, an additional concern is that wait time and prescriptions are jointly determined; for example, patients who are prescribed a drug at the ED may be able to wait longer for a follow up appointment if they are already receiving treatment through their medication. Since I am not using a patient's wait time directly, W_{-i} is not influenced by whether patient i himself is actually prescribed a mental health drug. However, other patients' wait times, which are used to form a given patient's W_{-i}, may be affected by prescribing patterns. This in turn could bias results if the rate of prescriptions differs when the ED is crowded. To address this, I drop patients that are prescribed a drug at the ED visit from the construction of W_{-i}. The resulting mortality effect is similarly 0.00130 (0.00039).

These results provide some indication that physicians are of aware of and compensating for appointment delays with more immediate treatment at the ED. Drugs prescribed at the ED could potentially have long run effects: the average days' supply for a given medication is 26 days, with an average of one refill allowed. Beyond refills, even if patients do not return to mental health care, primary care providers may re-prescribe medications to these patients. Since medications should mitigate mortality risk, prescribing patterns at the ED likely partially counteract the impact of missing future behavioral therapy appointments. As evidence of this point, returning to the IV specification in Section III.E.i, I directly control for whether a mental health drug was prescribed at the ED visit itself. As expected, this increases the coefficient to 0.063 (0.020), making the effects of therapy attendance look even larger.

G. Intermediate outcomes

i. Engagement with mental health care

In Section III.E, I showed evidence that mental health treatment delays make patients more likely to miss their follow up appointment. What services are patients actually missing as a result of appointment non-attendance, and how does congestion then affect long term engagement with mental health care? First and foremost, in my sample, 61% of follow up appointments are for general behavioral therapy, 15% are for addiction-related treatment, 7% are psychiatry appointments, 6% are for specialized PTSD treatment, and 11% are other specialty appointment types (such as mental health integrated care). During the two year follow up window, patients engage in a range of treatments, from medication to group therapy to long-term residential stays.

In Table 4 Panel A, I first evaluate how congestion affects long run engagement with behavioral therapy. Not only are patients with high mean wait time far more likely to miss their first follow up visit (column 1), as described in Section III.E, but they are far more likely to disengage *entirely* from mental health care in the two years following the index ED visit. In column 2, I show that a one standard deviation increase in mean wait time following the ED visit makes patients 0.67 percentage points (s.e.=0.04) less likely to attend *any* outpatient mental health appointment up to two years following the ED visit. Moreover, in column 3, I show that higher mean wait time decreases the likelihood patients attend any inpatient form of mental health health treatment, including long term residential stays. One interpretation of this latter result is that outpatient therapy attendance may be an important mechanism in connecting patients with more intensive forms of treatment.

Another potential explanation for lack of future mental health engagement is a competing risks model, where patients die before they are able to attend therapy appointments. However, this concern is mitigated by the fact that the majority of the mortality benefit appears only over a long time horizon (see Figure 4). On the whole, my results imply that treatment delays following an ED visit have effects that go far beyond a single appointment – instead, they have a long run impact on future engagement with mental health care.

What about medications? The vast majority of patients – around 91% – take some kind of mental health drug during the two-year follow up window. Medications for substance use specifically are somewhat less common, as some disorders do not have FDA-approved medications – notably stimulant use disorder. However, for opioid and alcohol use, medications are recommended as first-line treatments. Among my substance use subsample, around 54% of patients are prescribed a drug related to alcohol or drug use specifically, such as suboxone or naltrexone. Prescriptions may not go hand in hand with therapy attendance – only physicians, not psychologist or social workers, can prescribe medications. While psychiatrists have traditionally prescribed mental health drugs, in the current mental health landscape, primary care providers are increasingly writing these prescriptions (Faghri et al. 2010).

In Appendix Figure A5, I show results on the effect of mean wait time on use of mental health drugs over time. As described in Section III.F.iii, those with high mean wait time are about 1.2 percentage points more likely to be prescribed a drug at the ED visit itself, which I argue is a compensatory response to treatment delays. While this effect lessens over time as more patients are prescribed drugs at follow up visits, those with high mean wait time are actually still marginally more likely to be using mental health drugs in the long term (including both prescriptions and refills).[25]

With these results, I can rule out that those with high wait times are less likely to take mental health drugs overall. Although patients with high W_{-i} are more likely to drop out of specialty mental health care, they are also more likely to attend a primary care appointment during the follow up period (see evidence in Section III.G.iii). Thus, these results suggest that those with high wait times are more likely to be prescribed drugs by non-mental health specialists (or primary care providers). I cannot rule out that medications play no role in the mortality effect: one potential story is that primary care providers are less skilled at prescribing appropriate drugs to patients with severe behavioral health needs. This explanation is in line with evidence from Currie and MacLeod (2020), who show wide variation in provider skill in prescribing antidepressants. I comment more extensively on the role of primary care doctors in mental health care in Section III.G.iv.

ii. Cause of death

A natural question is whether the mortality effect is being driven by mental health casues of death, such as suicide and overdose, or whether effects are driven by physical disease, which can be exarcerbated by poor mental health (APA 2014). Using standard ICD-10 code classifications for suicide

[25] The slight increase in medication use holds true for both all mental health drugs and substance use drugs specifically.

from the National Death Index, I find that only 4.8% of deaths in my sample are directly attributable to suicide (Cornes et al. 2021). If I include drug overdose and suspected suicide events, this percentage increases to 9.6%. There is evidence that suicide deaths may be highly undercounted in both the veteran population and in general, likely due to both stigma and financial incentives related to life insurance policies (Ellis 2022, Snowdon and Choi 2020).

However, to get at the distinction between physical diseases and suicide or overdose-related causes of death, in Appendix Table A6, I divide deaths into those caused by physical disease versus those caused by accident, suicide, or overdose. Within physical disease, I further divide diseases into major categories, such as heart disease, respiratory diseases, and cancer. Let M_i be an indicator for mortality due to a specific cause. To determine cause-specific death rates among those with high wait time, I run regressions with M_i as the left-hand side variable as follows:

$$M_i = \pi_5 W_{-i} + \zeta_{5,t(i)} + \phi_{5,a(i)} + \gamma_{5,h(i)m(i)} + \mathbf{X}_i \delta_5 + \epsilon_{5,i} \tag{8}$$

Then, $\bar{M}_i + \hat{\pi}_5$, where $\bar{M}_i$ is the mean death rate, yields the cause-specific death rate for those with a mean wait time one stand. dev. above the mean ("high wait time patients"). Overall, I find that 6.5% of the sample died from a physical health condition, while only 1.0% died from other causes. This decomposition is similar among those with high wait times, with over 87% of deaths being due to physical causes. Specifically, the primary cause of death among those with high mean wait time is heart disease; these patients have a 1.69% chance of dying from heart disease, about a 3.4% increase from the sample mean of 1.63%. Moreover, high mean wait time patients have increased risk of dying from respiratory diseases, liver disease, and infections.

While perhaps surprising, these results corroborates previously established relationships in the medical and psychology literature. Those with poorer mental health are likely to also suffer from poorer physical health, and severe mental illness can worsen the course of chronic disease (APA

2014). Notably, the fact that most preventable deaths among the severely mentally ill are due to physical rather than mental conditions is consistent with existing evidence (Kisely et al. 2013, Liu et al. 2017). The medical literature has specifically documented the association between psychiatric conditions and excess mortality due to heart disease, respiratory diseases, and infections (Lawrence et al 2003, Newcomer and Hennekens 2007, Liu et al. 2017).

Moreover, one might expect that death from physical disease is particularly likely for those suffering from substance use disorder. Prior research has shown that addiction treatment is associated with decreased incidence and progression of alcohol-associated liver disease (Vannier et al. 2022). In addition, illicit opioid use is associated with a range of physical diseases, including chronic obstructive pulmonary disease, HIV, endocarditis, and viral hepatitis, which can also lead to liver disease (Lewer et al. 2021, Nishimura et al. 2019).

In a similar vein to this study, Kisley et al. (2013) find that patients with orders to attend psychiatric treatment had an all-cause mortality risk nearly half that of the control group at two years, with the strongest effects being for death from physical diseases. Analagously, I find that higher clinic congestion leads to disengagement from mental health care, also resulting in higher death rates from physical disease.

iii. Engagement with non-mental health medical care

Increased mortality from physical causes could have various underlying mechanisms: first, those who do not seek mental health care may be more likely to become detached from the healthcare system at large, and lack of engagement with physical health may be driving the mortality result. For simplicity, call this the "physical health engagement" channel. Alternatively, improved mental health alone could result in improved physical health, without being driven specifically by use of other types of medical care. For simplicity call this the "mental health engagement" channel. Finally, an intermediary between these two mechanisms could be contributing to the mortality effect: those with untreated mental illness could engage differently, and perhaps less effectively, with physical health care, even if they continue attending appointments. For example, those with untreated mental illness may have more difficulty adhering to lifestyle changes associated with improved cardiovascular and respiratory health, such as diet, exercise routines, and smoking cessation. Alternatively, physical health providers may have more difficulty diagnosing and treating physical health conditions in patients with untreated mental illness (Liu et al. 2017). Call this the "physical-mental intersection" channel.

As a point of reference, my sample is actually comprised of a highly VA-attached population: 97% of the sample returns to a VA hospital for at least some type of appointment within the two-year follow up period, and 84% return for a primary care appointment. Therefore, very few individuals lose contact with the VA system following the index ED visit.

To directly test the physical health care engagement channel, in Table 4 Panel B, I show that those with high mean wait time are actually *more* likely to attend a primary care appointment during the follow up period. Why is this the case? There are two potential explanations: first, patients may be substituting between mental and physical health care. If it is hard to get a speedy mental health appointment, they may show up to the primary care doctor instead. In fact, the APA suggests that high usage of primary care due to mental distress may be common, with up to 70% of primary care visits being driven by behavioral health (APA 2014). As primary care doctors treat both physical and behavioral health conditions, those who attend primary care more often may be receiving additional physical or mental health care, though delivered by a generalist. I explore delivery of mental health care by primary care doctors in the subsequent section.

Secondly, since my results suggest that high mean wait time patients also exhibit deteriorating physical health, these patients may simply be in more need of medical care. To provide additional evidence that these patients have declining physical health, in column 2, I show suggestive, though underpowered evidence that those with high mean wait time are more likely to be admitted as an inpatient for a physical health condition. Moreover, in column 3, I show that those with a one stand. dev. higher mean wait time are 0.19 p.p. (s.e.=0.08) or 0.3% more likely to be diagnosed with a new physical Elixhauser comorbidity. As these comorbidities are long term, chronic conditions that can worsen over time, it is likely that existing comorbidities, in addition to new diagnoses, contribute to the mortality effect.

In Appendix Table A7, I also break down engagement with mental and physical health care for SUD patients, as they an important driver of death in my sample (see Table 1). For these patients, results on hospitalizations for physical health and new physical comorbidities are even stronger. These results are aligned with the medical evidence surrounding cause of death for SUD patients, where untreated addiction is highly associated with physical health decline.

Overall, I demonstrate that although patients stay attached to the medical system as a whole, they exhibit declining physical health. Thus, I can rule out the "physical health engagement" channel, in which mental health care helps patients stay connected to physical health providers. However, both the "mental health engagement" and "physical-mental intersection" channels may play a role in contributing to the mortality effect.

iv. The role of experts in healthcare

Those with high mean wait times, despite being far more likely to disengage from specialized mental health treatment, are likely to both continue taking mental health medications and continue attending primary care. So why are these patients so much worse off? One way to think about the results is that patients disengage from *specialty* mental health care. However, they may continue attending primary care for mental health conditions, and non-specialist providers may continue prescribing them mental health drugs.

To further elucidate delivery of mental health care by primary care providers, I define any primary care visit with an associated mental health ICD-9 diagnosis code as a primary care appointment for mental health.[26] In my sample, about 65.6% of patients attend at least one primary care visit for mental health. Of these patients, about 3.6% attend *only* primary care for their mental health needs; in other words, they drop out of mental health specialty care. Those with high wait times are far more likely to fully substitute to primary care: a one stand. dev. increase in mean wait time increases the probability of receiving mental health treatment only via primary care by 0.27 percentage points (s.e.=0.03), or about 11.4%.

How often do patients receive any type of mental health treatment, whether via specialists or not? To get at this question, I redefine "returning" to mental health care from Table 4 as attending 1) a mental health specialty clinic appointment or 2) attending a primary care appointment with a mental health diagnosis code. Using this broader definition of mental health care, I find that those with high wait times are still less likely to receive any type of mental health care during the two year follow up window. A one stand. dev. increase in mean wait time decreases the probability of returning to any type of mental health care by 0.39 percentage points (s.e.= 0.03), about half of the effect size as that of specialty care alone (shown in Table 4). These results imply that patients at least partially substitute to primary care to treat their mental health conditions.

Lastly, how much does disengagement from specialty mental health care, as opposed to all forms of mental health care, contribute to the mortality effect? To shed light on this question, I exploit variation in care practices across hospitals. Notably, some hospitals may utilize primary care providers for mental health treatment more often than others. Moreover, there may be systematic variation in patient engagement across hospitals, even conditional on wait times. For instance, different administrations may have better systems to remind patients to attend appointments, and transportation may

[26] This is a relatively broad definition of mental health care, and includes any visit where mental health is addressed. This could include prescribing a mental health drug, checking in on an existing prescription, or psychotherapy.

be more accessible in certain VA locations. To categorize hospitals, I first run the following set of regressions hospital by hospital:

$$D_{sp,i} = \alpha_{sp,h(i)} W_{-i} + \zeta_{1,h(i)t(i)} + \phi_{1,h(i)a(i)} + \mathbf{X}_i \delta_{1,h(i)} + \eta_{1,i} \tag{9}$$

$$D_{all,i} = \alpha_{all,h(i)} W_{-i} + \zeta_{2,h(i)t(i)} + \phi_{2,h(i)a(i)} + \mathbf{X}_i \delta_{2,h(i)} + \eta_{2,i} \tag{10}$$

where $D_{sp,i}$ is an indicator for disengaging from specialty mental health care (attending 0 visits) in the two years following the ED visit, and $D_{all,i}$ is an indicator for disengaging from all mental health treatment (or both primary and specialty care). For each hospital h, I collect estimated coefficients $\hat{\alpha}_{sp,h(i)}$ and $\hat{\alpha}_{all,h(i)}$. I then categorize values of $\hat{\alpha}_{sp,h(i)}$ and $\hat{\alpha}_{all,h(i)}$ above the 75th percentile as "high disengagement" hospitals. Conceptually, a high value of $\alpha_{h(i)}$ means that at that hospital, congestion makes patients particularly likely to disengage from mental health care. In particular, high $\alpha_{sp,h(i)}$ indicates high disengagement from specialty mental health care, while high $\alpha_{all,h(i)}$ indicates high disengagement from all forms of mental health care.

Finally, to determine what type of mental health care matters for the mortality effect, I run the following regression:

$$Y_i = \pi_H W_{-i} \times H_{h(i)} + \pi_L W_{-i} \times L_{h(i)} + \zeta_{t(i)} + \phi_{a(i)} + \gamma_{h(i)m(i)} + \mathbf{X}_i \delta + \epsilon_i \tag{11}$$

where $H_{h(i)}$ and $L_{h(i)}$ are indicators for high and low disengagement hospitals, respectively. I run this regression twice, first forming $H_{h(i)}$ and $L_{h(i)}$ using values of $\hat{\alpha}_{sp,h(i)}$, and then using values of $\hat{\alpha}_{all,h(i)}$. π_H and π_L represent the mortality effect for high and low disengagement hospitals, respectively. The intuition behind including these interactions is as follows: if disengagement from all forms of mental health care is an important driver of the mortality effect, π_H should be high when

$H_{h(i)}$ is formed using $\hat{\alpha}_{all,h(i)}$. Alternatively, if disengagement from specialty mental health care specifically is important, π_H should be high when $H_{h(i)}$ is formed using $\hat{\alpha}_{sp,h(i)}$.

I show regression results from Eq. 10 in Table 5. I find that hospitals with high disengagement from specialty care have very large mortality effects. In fact, $\hat{\pi}_H$ is about 10 times the magnitude of $\hat{\pi}_L$ when hospitals are categorized based on specialty care alone (column 1). Alternatively, I find little to no heterogeneity when hospitals are categorized based on all forms of mental health care, including primary care (column 2). On the whole, these results suggest that disengagement from specialty mental health care specifically is a key driver of the mortality effect.

Overall, my analysis provides suggestive evidence that primary care may not be sufficient to meet patients' mental health needs. In other words, mental health expertise is useful. These results may be particularly policy relevant in the current US mental health landscape: there is a growing shortage of mental health specialists, and primary care providers are increasingly responsible for treating patients' psychiatric conditions (Faghri et al. 2010).

IV. Discussion

This paper is the first to evaluate the causal impact of mental health treatment delays on long term, all-cause mortality. For patients with emergent mental health conditions, I find that delaying follow up appointments by just one standard deviation – about 11.7 days – increases 2-year mortality by 1.5%. Moreover, in analyzing the downstream effects of high wait times, I identify a key barrier to healthcare system engagement: I find that those whose appointments were initially delayed are not only more likely to miss their first follow up visit, but they are far more likely to never engage with mental health services at the VA again.

On the whole, my results suggest that reducing clinic wait times could have important downstream effects on patients' wellbeing. One way to illustrate this point is as follows: if mean wait times across VA clinics fell from 17.7 days (as in the baseline sample) to about 10 days, or about a 0.66 standard deviation reduction, the reduced form results suggest that mortality would decrease by 0.073 percentage points, or about 1% of mean mortality. This effect size may represent an upper bound relative to other patient settings: given the urgent nature of their conditions, these veterans have two-year mortality rates higher than other veterans with mental illness. For instance, Bowersox et al. (2012) finds that VA patients with psychosis have a two-year mortality rate of 4.6%; the mortality rate among psychosis patients in my sample is 11.3%. Moreover, compared to non-veterans, these patients have unique provider accessibility, with veterans often paying no copay for treatment.

Appointment cancellations and no shows are an important issue not only in this setting, where the combined non-attendance rate was 37%, but among low income and vulnerable populations in general (Sharp and Hamilton 2001). Remarkedly, this study demonstrates that modifying scheduling practices may not only have behavioral implications for a single appointment, but may affect how patients interact with the healthcare system long term. In areas such as mental health, where going lost to follow up from care is a major concern, determining practices that can nudge patients to attend appointments may be key. Preliminary research has shown that massed visits, or scheduling appointments closer together in time (such as triweekly instead of weekly), increases treatment completion rates (Yamokoski 2020). Additionally, incentives to attend therapy sessions, such as vouchers or monetary rewards, are already in use at some VA facilities. These types of interventions have been shown to be highly effective in improving treatment adherence, at least while incentives continue; a randomized-control trial from 2020 finds that financial rewards of just $2-7 per week increase adherence to antidepressents by 21% (Marcus et al. 2020). In comparison, about 63.5% of patients in my sample attend their follow up appointment, and a decrease in wait time of about 12 days increases this probability by 2.0 percentage points – about a 3% increase.

Finally, my results underscore the importance of expertise in medicine. I find that those with long wait times – who are far more likely to disengage from specialty mental health care – are actually more likely to have primary care appointments during the two year follow up period, suggesting that primary care cannot substitute for mental health care. With patients increasingly turning to primary care to treat their mental health conditions, these results further highlight the need for more mental health specialists.

While my analysis provides suggestive evidence of the downstream effects of provider shortages, further research is needed to confirm the mechanisms driving my results. First and foremost, this study did not randomize delivery of mental health care by specialists versus generalists, which would provide more concrete evidence on the effects of mental health experts. Specifically, why do specialists perform better – is it due to differential prescribing patterns, more effective therapy, or some combination of the above? Lastly, this study is limited to mental health patients; future research could investigate whether waiting times have similar effects for patients suffering from physical health conditions.

With mental health treatment making headlines from recent gun legislation to policies surrounding veteran suicide, improving access to and engagement with behavioral therapy is of notable policy relevance. Major new initiatives aim to make mental health care more accessible, both at the VA and at large: in 2022, the VA proposed to eliminate outpatient copayments for veterans flagged high risk for suicide. More broadly, in summer 2022, the US launched 988, an emergency number for those experiencing mental health emergencies. My results suggest that timely care, in addition to more affordable care, may be a key part of the puzzle.

References

Abed Faghri, Nahid M, Charles M Boisvert, and Sanaz Faghri. 2010. "Understanding the expanding role of primary care physicians (PCPs) to primary psychiatric care physicians (PPCPs): enhancing the assessment and treatment of psychiatric conditions." *Ment Health Fam Med*, 7(1): 17–25.

American Psychological Association (APA). 2014. "Briefing series on the role of psychology in health care: primary care."

American Psychological Association (APA). 2021. "Demand for mental health treatment continues to increase, say psychologists."

Angelucci, Manuela, and Daniel Bennett. 2022. "The Economic Impact of Depression Treatment in India: Evidence from Community-Based Provision of Pharmacotherapy." Job Market Paper.

Appleby, L., J. A. Dennehy, C. S. Thomas, E. B. Faragher, and G. Lewis. 1999. "Aftercare and clinical characteristics of people with mental illness who commit suicide: a case-control study." *Lancet*, 353(9162): 1397–1400.

Betancourt, Jose A., Diane M. Dolezel, Ramalingam Shanmugam, Gerardo J. Pacheco, Paula Stigler Granados, and Lawrence V. Fulton. 2023. "The Health Status of the US Veterans: A Longitudinal Analysis of Surveillance Data Prior to and during the COVID-19 Pandemic." *Healthcare (Basel)*, 11(14): 2049.

Bhat, Bhargav, Jonathan de Quidt, Johannes Haushofer, Vikram H. Patel, Gautam Rao, Frank Schilbach, and Pierre-Luc P. Vautrey. 2022. "The Long-Run Effects of Psychotherapy on Depression, Beliefs, and Economic Outcomes." NBER Working Paper 30011.

Biasi, Barbara, Michael S. Dahl, and Petra Moser. 2021. "Career Effects of Mental Health." NBER Working Paper 29031.

Bjorkenstam, Emma, Rickard Ljung, Bo Burstrom, Ellenor Mittendorfer-Rutz, Johan Hallqvist, and Gunilla Ringback Weitoft. 2012. "Quality of medical care and excess mortality in psychiatric patients–a nationwide register-based study in Sweden." *BMJ Open*, 2: e000778.

Blattman, Christopher, Sebastian Chaskel, Julian C. Jamison, and Margaret Sheridan. 2022. "Cognitive Behavior Therapy Reduces Crime and Violence over 10 Years: Experimental Evidence." NBER Working Paper 30049.

Bowersox, Nicholas W., Amy M. Kilbourne, Kristen M. Abraham, Brian H. Reck, Zongshan Lai, Amy S. B. Bohnert, David E. Goodrich, and Chester L. Davis. 2012. "Cause-specific mortality among Veterans with serious mental illness lost to follow-up." *Gen Hosp Psychiatry*, 34(6): 651–653.

Caron, Christina. 2021. "Mental Health Providers Struggle to Meet Pandemic Demand." *The New York Times*.

Chan, David C., David Card, and Lowell Taylor. 2023. "Is There a VA Advantage? Evidence from Dually Eligible Veterans." *American Economic Review*, 113(11): 3003–3043.

Cornes, Michelle, Robert W. Aldridge, Elizabeth Biswell, Richard Byng, Michael Clark, Graham Foster, James Fuller, Andrew Hayward, Nigel Hewett, Alan Kilmister, Jill Manthorpe, Joanne Neale, Michela Tinelli, and Martin Whiteford. 2021. "ICD-10 codes used to define underlying causes of death due to alcohol, drugs or suicide." In *Improving care transfers for homeless patients after hospital discharge: a realist evaluation*. NIHR Journals Library.

Cuellar, Alison Evans, Sara Markowitz, and Anne M. Libby. 2004. "Mental health and substance abuse treatment and juvenile crime." *J Ment Health Policy Econ*, 7(2): 59–68.

Cuijpers, Pim, Ioana A. Cristea, Eirini Karyotaki, Mirjam Reijnders, and Marcus J.H. Huibers. 2016. "How effective are cognitive behavior therapies for major depression and anxiety disorders?" *World Psychiatry*, 15(3): 245–258.

Currie, Janet M., and W. Bentley Macleod. 2020. "Understanding doctor decision making: the case of Depression treatment." *Econometrica*, 88(3): 847.

Cutler, David M. 2002. "Equality, Efficiency, and Market Fundamentals: The Dynamics of International Medical-Care Reform." *Journal of Economic Literature*, 40(3): 881–906.

Das-Munshi, Jayati, Chin-Kuo Chang, Rina Dutta, Craig Morgan, James Nazroo, Robert Stewart, and Martin J Prince. 2017. "Ethnicity and excess mortality in severe mental illness: a cohort study." *Lancet Psychiatry*, 4(5): 389–399.

Deshpande, Manasi, and Yue Li. 2019. "Who Is Screened Out? Application Costs and the Targeting of Disability Programs." *American Economic Journal: Economic Policy*, 11(4): 213–248.

Dupas, Pascaline, Vivian Hoffmann, Michael Kremer, and Alix Peterson Zwane. 2016. "Targeting health subsidies through a nonprice mechanism: A randomized controlled trial in Kenya." *Science*, 353(6302): 889–895.

Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey. 1998. "Comorbidity measures for use with administrative data." *Med Care*, 36(1): 8–27.

Ellis, Ralph. 2022. "Veterans Suicides May Be Highly Undercounted, Study Says."

Feyman, Yevgeniy, Daniel A. Asfaw, and Kevin N. Griffith. 2022. "Geographic Variation in Appointment Wait Times for US Military Veterans." *JAMA Netw Open*, 5(8): e2228783.

Feyman, Y., S. M. Figueroa, Y. Yuan, M. Price, A. Kabdiyeva, J. Nebeker, M. Ward, P. Shafer, S. D. Pizer, and K. L. Strombotne. 2022. “Effect of Mental Health Staffing Inputs on Suicide-Related Events.” *Health Serv Res*.

Fischer, Ellen P., John F. McCarthy, Rosalinda V. Ignacio, Frederic C. Blow, Kristen L. Barry, Teresa J. Hudson, Richard R. Owen, and Marcia Valenstein. 2008. “Longitudinal patterns of health system retention among veterans with schizophrenia or bipolar disorder.” *Community Ment Health J*, 44(5): 321–330.

Gartlehner, Gerald, Gernot Wagner, Nina Matyas, Viktoria Titscher, Judith Greimel, Linda Lux, Bradley N. Gaynes, Meera Viswanathan, Sheila Patel, and Kathleen N. Lohr. 2017. “Pharmacological and non-pharmacological treatments for major depressive disorder: review of systematic reviews.” *BMJ Open*, 7(6): e014912.

Gibbons, Robert D., Kwan Hur, and J. John Mann. 2017. “Suicide Rates and the Declining Psychiatric Hospital Bed Capacity in the United States.” *JAMA Psychiatry*, 74(8): 849–850.

Godoy, Anna, Venke F. Haaland, Ingrid Huitfeldt, and Mark Votruba. n.d.. “Hospital Queues, Patient Health and Labor Supply.” *American Economic Journal: Economic Policy*, forthcoming.

Gruber, Jonathan, Thomas Hoe, and George Stoye. 2021. “Saving Lives by Tying Hands: The Unexpected Effects of Constraining Health Care Providers.” *The Review of Economics and Statistics*.

Holliday, Stephanie. 2018. “The Relationship Between Mental Health Care Access and Suicide.” RAND Corporation.

Holmes, Emily A., Ata Ghaderi, Catherine J. Harmer, Paul G. Ramchandani, Pim Cuijpers, Anthony P. Morrison, Jonathan P. Roiser, Claudi L. H. Bockting, Rory C. O’Connor, Roz

Shafran, Michelle L. Moulds, and Michelle G. Craske. 2018. “The Lancet Psychiatry Commission on psychological treatments research in tomorrow’s science.” *The Lancet Psychiatry*, 5(3): 237–286. Publisher: Elsevier.

Kisely, Steve, Neil Preston, Jianguo Xiao, David Lawrence, Sandra Louise, and Elizabeth Crowe. 2013. “Reducing all-cause mortality among patients with psychiatric disorders: a population-based study.” *CMAJ*, 185(1): E50–E56.

Klinkenberg, W. D., and R. J. Calsyn. 1996. “Predictors of receipt of aftercare and recidivism among persons with severe mental illness: a review.” *Psychiatr Serv*, 47(5): 487–496.

Kreyenbuhl, Julie, Ilana R. Nossel, and Lisa B. Dixon. 2009. “Disengagement From Mental Health Treatment Among Individuals With Schizophrenia and Strategies for Facilitating Connections to Care: A Review of the Literature.” *Schizophr Bull*, 35(4): 696–703.

Lang, Matthew. 2013. “The Impact of Mental Health Insurance Laws on State Suicide Rates.” *Health Economics*, 22(1): 73–88.

Lawrence, David M., Cashel D’arcy J. Holman, Assen V. Jablensky, and Michael S. T. Hobbs. 2003. “Death rate from ischaemic heart disease in Western Australian psychiatric patients 1980-1998.” *Br J Psychiatry*, 182: 31–36.

Lee, David S. 2009. “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects.” *The Review of Economic Studies*, 76(3): 1071–1102.

Lewer, Dan, Thomas D. Brothers, Naomi Van Hest, Matthew Hickman, Adam Holland, Prianka Padmanathan, and Paola Zaninotto. 2022. “Causes of death among people who used illicit opioids in England, 2001-18: a matched cohort study.” *Lancet Public Health*, 7(2): e126–e135.

Liu, Nancy H., Gail L. Daumit, Tarun Dua, Ralph Aquila, Fiona Charlson, Pim Cuijpers, Benjamin Druss, Kenn Dudek, Melvyn Freeman, Chiyo Fujii, Wolfgang Gaebel, Ulrich Hegerl, Itzhak Levav, Thomas Munk Laursen, Hong Ma, Mario Maj, Maria Elena Medina Mora, Merete Nordentoft, Dorairaj Prabhakaran, Karen Pratt, Martin Prince, Thara Rangaswamy, David Shiers, Ezra Susser, Graham Thornicroft, Kristian Wahlbeck, Abe Fekadu Wassie, Harvey Whiteford, and Shekhar Saxena. 2017. "Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas." *World Psychiatry*, 16(1): 30–40.

Marcus, Steven C., Megan E. Reilly, Kelly Zentgraf, Kevin G. Volpp, and Mark Olfson. 2021. "Effect of Escalating and Deescalating Financial Incentives vs Usual Care to Improve Antidepressant Adherence: A Pilot Randomized Clinical Trial." *JAMA Psychiatry*, 78(2): 222–224.

McBain, Ryan K., Jonathan H. Cantor, and Nicole K. Eberhart. 2022. "Estimating Psychiatric Bed Shortages in the US." *JAMA Psychiatry*, 79(4).

National Academies of Sciences, Engineering, and Medicine. 2018. *Timely Access to Mental Health Care.* National Academies Press (US).

National Committee for Quality Assurance (NCQA). n.d.. "Follow-Up After Emergency Department Visit for Mental Illness."

Newcomer, John W., and Charles H. Hennekens. 2007. "Severe Mental Illness and Risk of Cardiovascular Disease." *JAMA*, 298(15): 1794–1796.

Nishimura, Marin, Harpreet Bhatia, Janet Ma, Stephen D. Dickson, Laith Alshawabkeh, Eric Adler, Alan Maisel, Michael H. Criqui, Barry Greenberg, and Isac C. Thomas. 2020. "The Impact of Substance Abuse on Heart Failure Hospitalizations." *Am J Med*, 133(2): 207–213.e1.

Nose, M., C. Barbui, and M. Tansella. 2003. “How often do patients with psychosis fail to adhere to treatment programmes? A systematic review.” *Psychol Med*, 33(7): 1149–1160.

Raphelson, Samantha. 2017. “How The Loss Of U.S. Psychiatric Hospitals Led To A Mental Health Crisis.” *NPR*.

Reichert, Anika, and Rowena Jacobs. 2018. “The impact of waiting time on patient outcomes: Evidence from early intervention in psychosis services in England.” *Health Econ*, 27(11): 1772–1787.

Sharp, D. J., and W. Hamilton. 2001. “Non-attendance at general practices and outpatient clinics.” *BMJ*, 323(7321): 1081–1082.

Singh, Manasvini, and Atheendar Venkataramani. 2022. “Capacity Strain and Racial Disparities in Hospital Mortality.” NBER Working Paper 30380.

Snowdon, John, and Namkee G. Choi. 2020. “Undercounting of suicides: Where suicide data lie hidden.” *Glob Public Health*, 15(12): 1894–1901.

Turner, Alex J, Igor Francetic, Ruth Watkinson, Stephanie Gillibrand, and Matt Sutton. 2022. “Socioeconomic inequality in access to timely and appropriate care in emergency departments.” *Journal of Health Economics*, 85: 102668.

Vannier, Augustin G. L., Jessica E. S. Shay, Vladislav Fomin, Suraj J. Patel, Esperance Schaefer, Russell P. Goodman, and Jay Luther. 2022. “Incidence and Progression of Alcohol-Associated Liver Disease After Medical Therapy for Alcohol Use Disorder.” *JAMA Network Open*, 5(5): e2213014.

VA Office of Mental Health and Suicide Prevention. 2022. “National Veteran Suicide Prevention: Annual Report.” Department of Veterans Affairs.

Veterans Health Administration. 2008. "VHA Directive 2008-036." Department of Veterans Affairs.

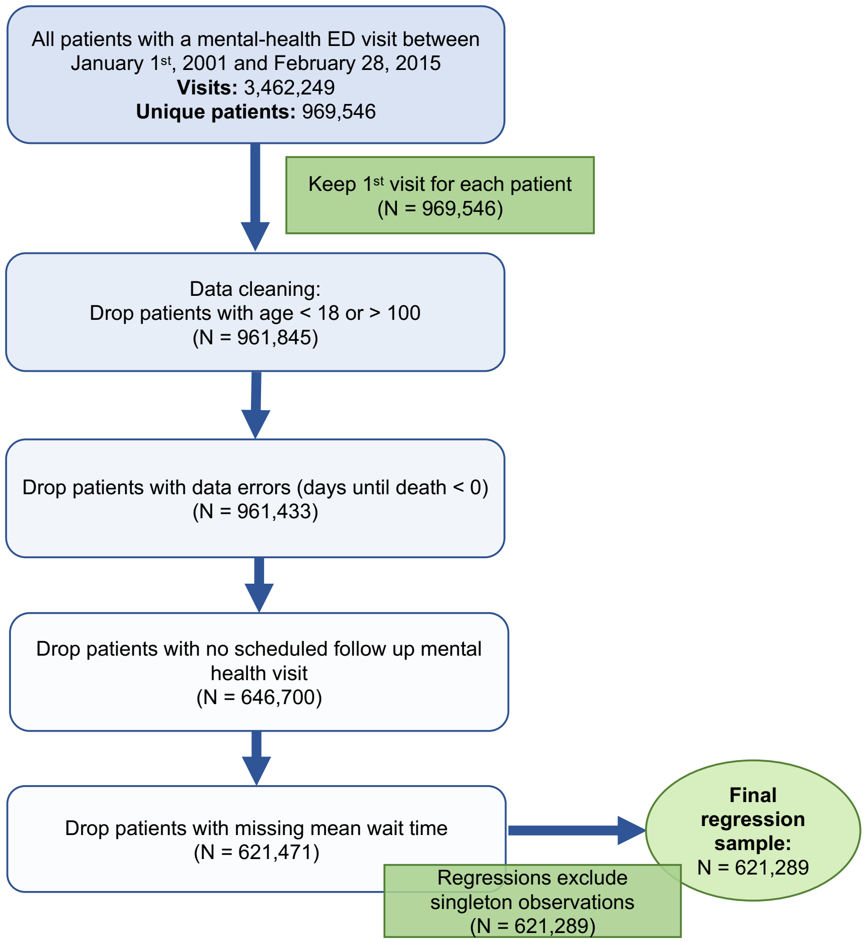
Weiner, Stacy. 2022. "A growing psychiatrist shortage and an enormous demand for mental health services."

Williams, Jenny, and Anne Bretteville-Jensen. 2022. "Whats Another Day? The Effects of Wait Time for Substance Abuse Treatment on Health-Care Utilization, Employment and Crime." *Institute of Labor Economics*.

Xu, Wendy Yi, Chi Song, Yiting Li, and Sheldon Michael Retchin. 2019. "Cost-Sharing Disparities for Out-of-Network Care for Adults With Behavioral Health Conditions." *JAMA Network Open*, 2(11): e1914554.

Yamokoski, Cynthia. 2020. "Massed Delivery of Evidence-Based Psychotherapy for PTSD." Department of Veterans Affairs.

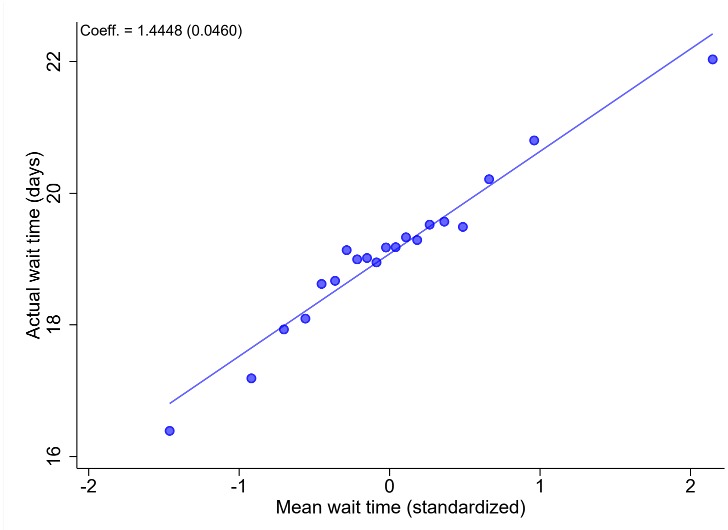
Figure 1: Sample selection



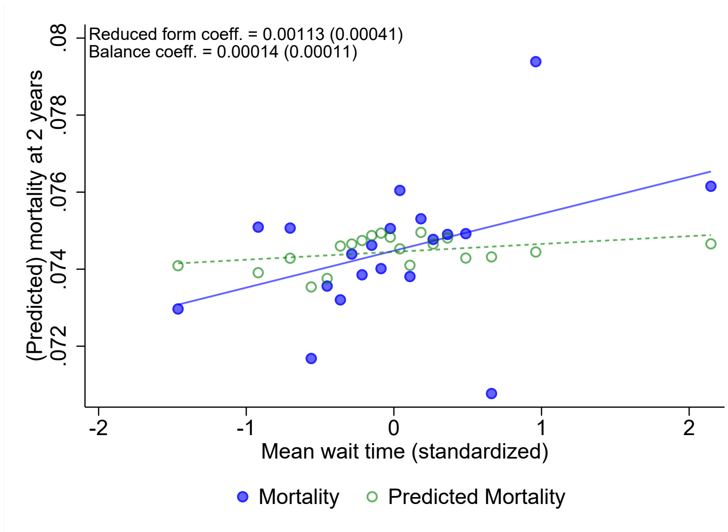
Note: This figure shows the sample selection process using encounter data and patient information from the VA. Patients have a missing mean wait time value if no other patients are in the same hospital-two-week period-appointment type group as the index patient.

Figure 2: Relevance, balance, and reduced form

A: Correlation between mean wait time and actual wait time

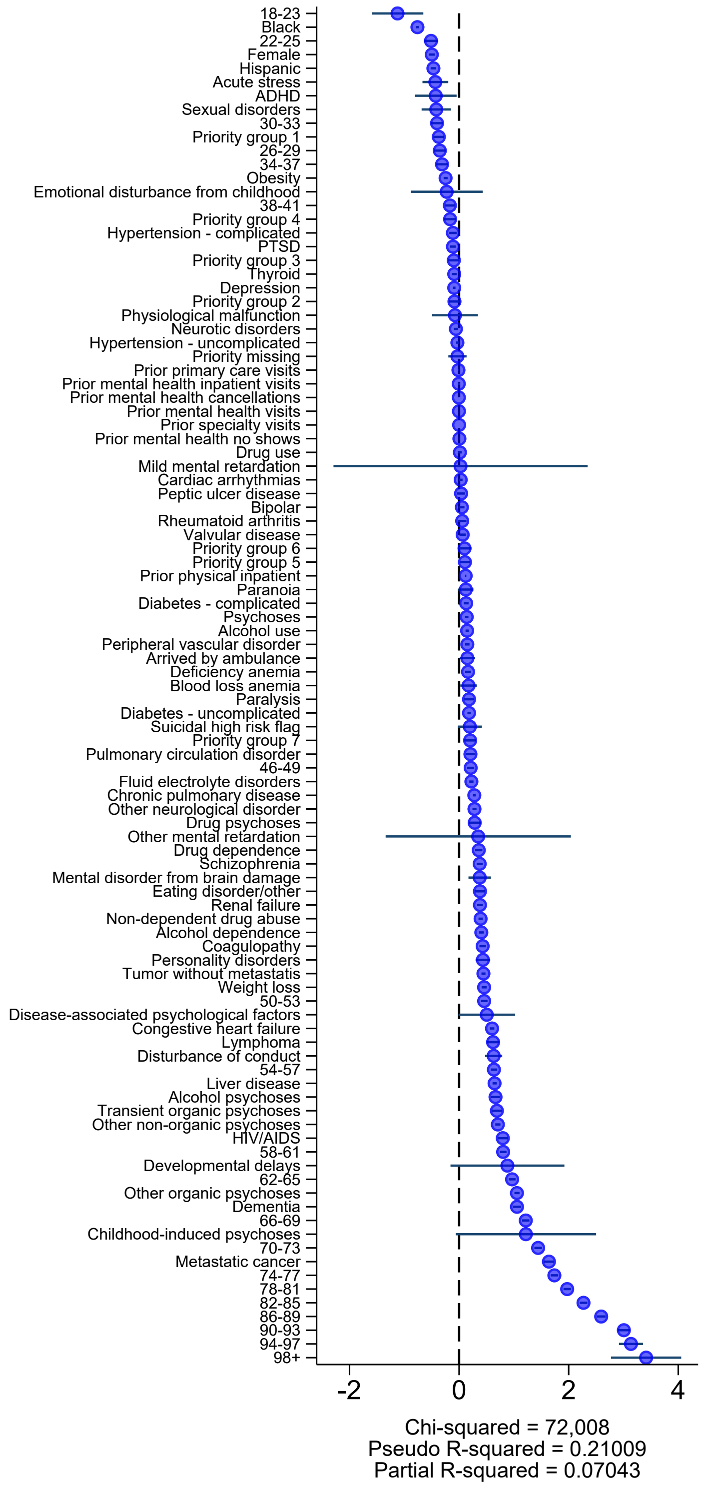


B: Balance and reduced form



Note: Panel A shows a binned scatterplot of actual wait time (in days) versus mean wait time (standardized), or W_{-i}, residualizing by the baseline controls. This represents relationship described in Equation 2. Panel B shows a binned scatterplot of two-year mortality or predicted mortality versus W_{-i}, residualizing by the baseline controls. These represent the balance and reduced form relationships described in Equations 3 and 4. Predicted mortality is plotted using green hollow circles, and mortality is plotted using solid blue circles. Baseline controls include two-week period, year-month × hospital, appointment type, and age buckets.

Figure 3: Mortality prediction



Note: This figure plots regression coefficients and standard errors for a logistic regression of two-year mortality on patient characteristics used to form $\hat{Y}_i$. Physical diagnoses are based on conditions patients were treated for at the VA in the prior year and are divided into Elixhauser comorbidity groups. The excluded group are white males aged 42 to 45 with depression in Priority group 8. The partial R-squared reports the marginal contribution of all other variables after removing the effects of age.

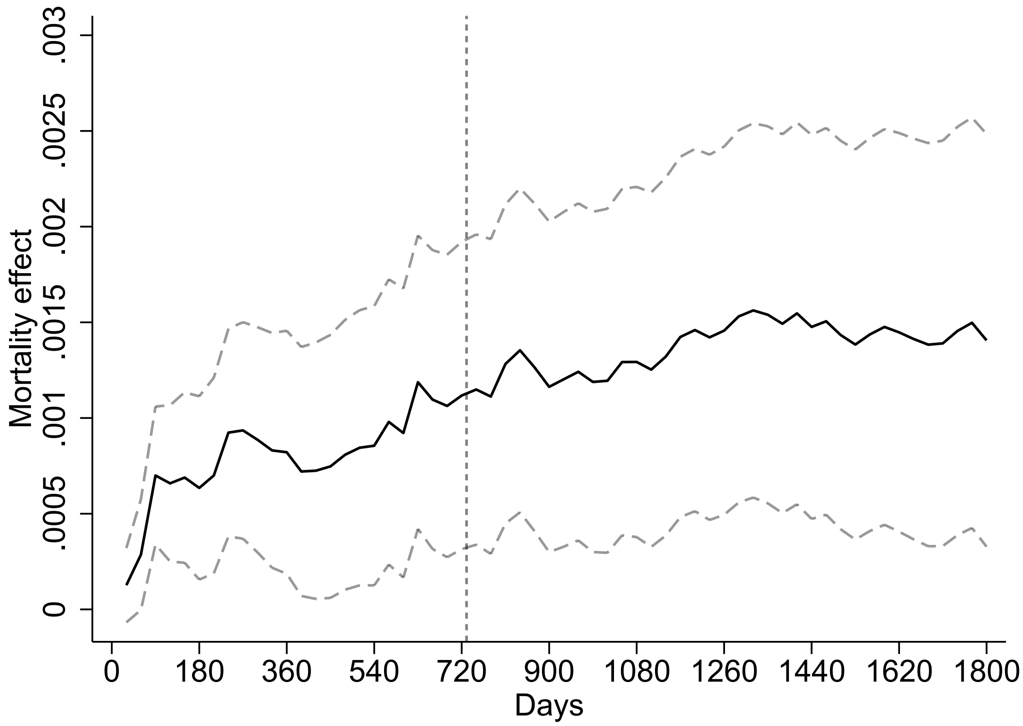
Table 1: Effect of mental health wait times on mortality

	A: Overall Sample				
	2-year mortality				
	(1)	(2)	(3)	(4)	(5)
Mean wait time	0.00113	0.00104	0.00107	0.00113	0.00093
	(0.00041)	(0.00041)	(0.00041)	(0.00041)	(0.00041)
Outcome mean	0.0746	0.0746	0.0746	0.0746	0.0746
Observations	621,289	621,289	621,289	621,289	621,289
Baseline controls	Yes	Yes	Yes	Yes	Yes
Demographics	No	Yes	Yes	Yes	Yes
Prior utilization	No	No	Yes	Yes	Yes
Diagnosis codes	No	No	No	Yes	Yes
Leave out controls	No	No	No	No	Yes

	B: Major Diagnosis Category				
	2-year mortality				
	Substance use	Psychosis	Mood disorders	Anxiety	PTSD
Mean wait time	0.00186	0.00360	0.00048	0.00030	0.00048
	(0.00081)	(0.00208)	(0.00074)	(0.00114)	(0.00107)
Outcome mean	0.0707	0.1130	0.0527	0.0527	0.0360
Observations	175,721	52,600	184,812	71,895	60,178

Note: In Panel A, I show the baseline mortality results. Column 1 shows the main estimate, only using baseline controls. Column 2 shows the estimate when I add demographic controls, including gender and race. Column 3 shows the estimate when I also add controls for prior mental health utilization, and column 4 adds 3-digit ICD-9 diagnosis codes (from the ED visit). Finally, column 5 shows the estimate when I control for characteristics of other patients who form a given patient's value of W_{-i}, including (leave-out) mean age, mean prior visits, and mean predicted mortality. In Panel B, I show mortality results in diagnosis group subsamples (based on ICD-9 codes from the ED visit). Baseline controls include month-year × hospital, appointment type, and age buckets. Standard errors are clustered by hospital-month.

Figure 4: Mortality effect over time



Note: This figure plots the mortality effect over time. Solid lines show point estimates, and dashed lines show 95% confidence intervals. The vertical dashed line indicates the main mortality estimate at 2 years.

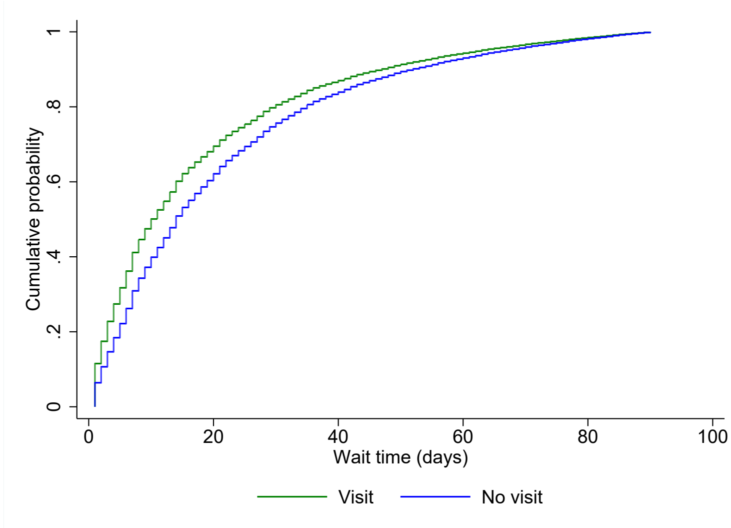
Table 2: Robustness checks

	A: Accounting for selection into appointment type			
	(1) Baseline	(2) Less granular controls	(3) Saturated	(4) Predicted wait time
Mean wait time	0.00113	0.00124	0.00090	0.00092
	(0.00041)	(0.00041)	(0.00043)	(0.00047)
Outcome mean	0.0746	0.0746	0.0746	0.0746
Observations	621,289	621,289	621,289	621,289
Appointment types	11	5	5	5
Hospital × time controls	Year × month	Year × quarter	Year × quarter	Year × quarter
	B: Alternative samples			
	(1) Baseline	(2) + Patients with no scheduled follow up	(3) Exclude inpatient visits	(4) Only VA insurance
Mean wait time	0.00113	0.00118	0.00129	0.00075
	(0.00041)	(0.00035)	(0.00041)	(0.00044)
Outcome mean	0.0746	0.0956	0.0746	0.0528
Observations	621,289	935,843	621,289	432,863

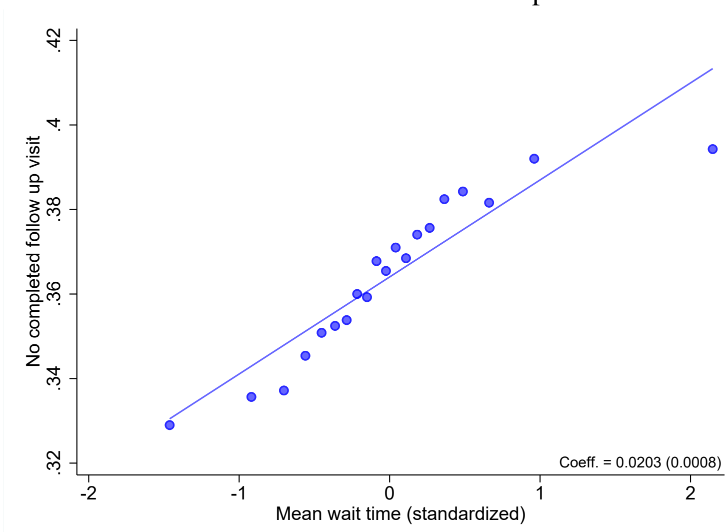
Note: Panel A shows regression results that account for appointment type endogeneity. Column 1 shows the baseline estimate from Table 1, column 1. Column 2 shows the estimate with less granular controls: hospital × year × quarter controls (in lieu of hospital × year × month controls) and less granular appointment type categories (general, psychiatry, PTSD, SUD, and other). In this specification, I also re-construct W_{-i} using these larger appointment type categories. Still using the less granular fixed effects, column 3 adds controls for hospital × appointment type and hospital × 2-week period. Finally, column 4 predicts appointment type instead of using assigned appointment type. Panel B shows the mortality effect using alternative versions of the sample. Column 2 adds back patients with no scheduled follow up visit, column 3 excludes inpatient mental health encounters from the universe of mental health follow up visits, and column 4 excludes patients with outside insurance.

Figure 5: Treatment delays and follow up appointment attendance

A: CDFs of actual wait times



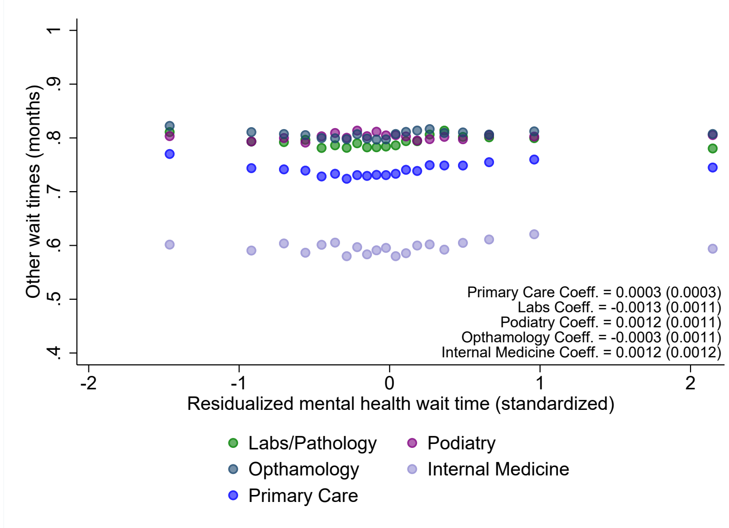
B: Effect of mean wait time on follow up attendance



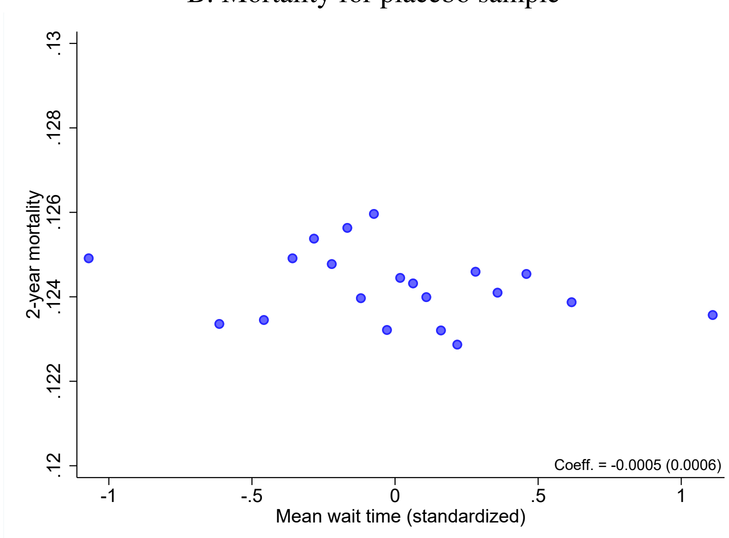
Note: Panel A plots the CDF of actual wait time (in days) for patients who did (green) and did not attend (blue) their follow up appointment. Panel B shows a binned scatterplot of the probability of not attending a follow up mental health visit versus leave-out wait time, or W_{-i}, residualizing by the baseline controls. This represents the relationship described in Equation 6.

Figure 6: Correlation between mean wait time and quality of non-mental health care

A: Other hospital wait times



B: Mortality for placebo sample



Note: Panel A shows a binned scatterplot of various other specialty wait times versus residualized W_{-i}. All wait times are calculated as the mean number of months between a patient's ED visit and their follow up appointment day for a given specialty. Panel B shows a binned scatterplot of two-year mortality versus W_{-i} for a placebo sample of patients with circulatory conditions, residualizing by the baseline controls. This is analagous with the reduced form relationship for the main sample described in Equation 4. Baseline controls for this sample include two-week period, year-month × hospital, and age buckets.

Table 3: Effect of mean wait time on prescriptions at the ED

	(1) Prescription (overall)	(2) New prescriptions	(3) Anticonvulsants	(4) Antidepressants
Mean wait time	0.0120	0.0058	0.0014	0.0088
	(0.0008)	(0.0005)	(0.0003)	(0.0006)
Percent increase	4.0%	5.0%	4.6%	5.3%
Outcome mean	0.3016	0.1162	0.0306	0.1643
Observations	621,289	621,289	621,289	621,289
	(5) Analgesics	(6) Sedatives or hypnotics	(7) Anti-psychotics	(8) Lithium
Mean wait time	0.0022	0.0022	0.0015	0.0003
	(0.0004)	(0.0005)	(0.0003)	(0.0001)
Percent increase	4.6%	2.1%	2.8%	7.4%
Outcome mean	0.0476	0.1033	0.0516	0.0042
Observations	621,289	621,289	621,289	621,289

Note: This table shows the effect of mean wait time on the likeliness of being prescribed a mental health drug at the ED visit. Baseline controls include two-week period, month-year × hospital, appointment type, and age buckets. Standard errors are clustered by hospital-month.

Table 4: Effect of mean wait time on intermediate outcomes

	A: Mental health outcomes		
	(1)	(2)	(3)
	No follow up visit	Returned to mental health care	Mental health inpatient treatment
Mean wait time	0.0203	-0.0067	-0.0099
	(0.0008)	(0.0004)	(0.0008)
Outcome mean	0.3652	0.9333	0.3153
Observations	621,289	621,289	621,289
	B: Physical health outcomes		
	(1)	(2)	(3)
	Primary care visit	Inpatient visit - non-mental health	New physical comorbidity
Mean wait time	0.0032	0.0012	0.0019
	(0.0006)	(0.0007)	(0.0008)
Outcome mean	0.8404	0.2114	0.6208
Observations	621,289	621,289	621,289

Note: This table presents reduced form results for various intermediate outcomes, following patients for 730 days after the ED visit. Panel A shows results for engagement with mental health care. Returned to mental health care refers to attending any oupatient appointment. Mental health inpatient treatment refers to any mental health inpatient visit, including residential stays (excluding admissions within one day of the initial ED visit). Panel B presents results for non-mental health outcomes. The first column looks at whether a patient attended a primary care appointment, the second column looks at non-mental health inpatient admissions, and the third column looks at whether the patient had any new Elixhauser comorbidity (only including physical conditions) during the two-year follow up period.

Table 5: Heterogeneity by level of disengagement from mental health care

	2-year mortality	
	(1)	(2)
	Specialty disengagement	Full disengagement
Mean wait time × high disengagement	0.0031	0.0013
	(0.0007)	(0.0008)
Mean wait time × low disengagement	(0.0003)	0.0011
	(0.0005)	(0.0005)
Outcome mean	0.0746	0.0746
Observations	621,289	621,289

Note: This table shows heterogeneity results for VA hospitals with high and low levels of disengagement from mental health care. "High disengagement" is an indicator for being in the top 25th percentile of hospitals, while "low disengagement" in an indicator for being in the bottom 75th percentile (based on values of $\hat{\alpha}_{h(i)}$ from Eqs. 8 and 9). In column 1, I show heterogeneity results for VA hospitals based on the probability of disengaging from specialty mental health care. In column 2, I show heterogeneity results for VA hospitals based on probabilities of disengaging from all mental health care, including specialty and primary care visits. Disengagement from mental health care is defined as not attending any mental health appointment in the two years following the ED visit. Mental health primary care is defined as a primary care visit assigned a mental health ICD-9 diagnosis code.

For Online Publication

Appendix for "How do mental health treatment and treatment delays impact long term mortality?"

Sydney Costantini

A1 Alternative Samples

A1.1 Patients with no outside insurance

In Section III.E.ii, I discuss a potential issue with the interpretability of my results: what if patients also receive mental health care outside of the VA? To address this issue, I create a subsample of patients who have no outside or non-VA insurance, making it highly likely that they receive care only at the VA. This subsample comprises 69.7% of patients in the primary sample. In Table A1, I show that across characteristics, the no insurance subsample is highly analagous to the main sample, with two exceptions. First, the mean age is slightly younger compared to that of the main sample. Second, two-year mortality in this subsample is 5.3%, as opposed to 7.5% in the main sample. Both of these facts likely reflect the fact that most patients with outside insurance have Medicare, which is only available for patients above 65 or for those who qualify due to disability status.

Mortality results for this sample are presented in Table A4.

A1.2 Patient without scheduled follow up visits

In the sample selection described in Section II.A, I exclude patients with no scheduled follow up visit, as W_{-i} depends on type of appointment scheduled. Since this group is comprised of a substantial 314,554 patients, in Table A1, I also compare characteristics of patients with no scheduled follow up appointments with those of the main sample. As expected, I find that these patients have much lower prior attachment to VA mental health services and are more likely to have Medicare, implying that more of these patients are using care outside of the VA. Additionally, patients with no follow up appointment are older and are much more likely to be diagnosed with SUD, suggesting that these types of patients may be particularly unlikely to return for mental health services.

Finally, these patients have slightly higher mean wait time than that of the main sample. This fact implies that those who arrive during more congested times may be less likely to have a follow up appointment scheduled at all. As this presents a potential non-random attrition problem that could bias results, I add these patients back to the main sample in Appendix Section A3.1.

A2 Additional balance checks

As the VA flags patients considered to be high risk for suicide, a natural additional balance test is to see whether these high risk patients are just as likely to be assigned low and high values of W_{-i}. I restrict to patients flagged as high risk before the ED visit, as being assigned this flag at the ED visit itself could potentially be endogeous to my experiment. Overall, only 0.3% of patients in my sample are assigned the high risk flag before the ED visit. In Figure A2 Panel A, I show a binscatter of percent of patients assigned the high risk flag versus W_{-i}, residualizing by the baseline controls. I find a small and statisitically insignificant coefficient, providing more evidence that mean wait time is approximately randomly assigned.

Secondly, another key metric indicative of patient severity is whether a patient has a prior hospitalization. Let I_i indicate that a patient was hospitalized in the year prior to the index ED visit (for any cause). In Figure A2 Panel B, I show a binscatter of I_i versus W_{-i}, residualizing by the baseline controls. I find that the coefficient on W_{-i} is both reverse-signed and insignificant, providing further evidence that W_{-i} is uncorrelated with patient severity.

A3 Additional robustness checks

A3.1 Patients without scheduled follow up appointments

To eliminate concern that dropping these patients has implications for my findings, I add patients with no scheduled follow up appointments back to the sample, increasing the sample size to 935,843. This combined sample has a somewhat higher mean two-year mortality of 9.6%.

As the primary construction of W_{-i} and controls rely on assigned appointment type, I make a few modifications to the design in order to run regressions using this larger sample. First, since patients with no scheduled follow up do not have an assigned appointment type, I use two alternative approaches to approximate congestion levels for these patients. I start with a simple average approach: I group all appointment types together during two-week-hospital periods to construct W_{-i} for these patients. Alternatively, I let $W_{-i} = \hat{W}_{-i}$, as described in Section III.C.i. of the main text.[1] In this latter approach, $\hat{W}_{-i}$ is essentially a weighted average mean wait time, weighted by the probability of the patient being assigned a particular appointment type. As appointment type is also a baseline control in my main analysis, in both specifications, I control for either appointment type directly (for the baseline sample) or the predicted probability of being assigned each appointment type (for patients with no appointment type).

In Table A3, I show regression results using both of the aforementioned methods. I show results both with the baseline controls and all additional controls used in Table 1; results are both robust and very similar to those of the baseline analysis sample.

A3.2 Miscellaneous robustness checks

In Figure A3, I show that the baseline results are stable when both 1) detailed patient controls are added, and 2) sensible modifications to the sample or treatment variable are made. Under "Additional controls," I first show the main result with the baseline controls– two-week periods, year-month × hospital, appointment type, and age buckets. Then, I incrementally add groups of additional controls (the next row adds demographics controls, the following row adds demographic and prior utilization controls, etc.).[2] Demographic controls include sex, race, and ethnicity. Prior utilization controls includes count of inpatient and outpatient mental health visits in the previous year, count of mental health appointment cancellations, and count of no shows. Diagnosis controls are 3-digit ICD-9 diagnosis codes associated with the ED visit. Priority status refers to a patient's eligibility for VA healthcare and responsibility for paying copays, which depends on both disability status and income. The second to last row in this section adds leave-out mean characteristics of patients who form a given patient's mean wait time. Leave-out characteristics include mean age, mean prior visits, and

[1] W_{-i} for observations also in the baseline sample are held fixed in these analyses.
[2] These controls overlap with the controls added in Table 1.

mean predicted mortality. Finally, the last row controls for first half versus second half of the month indicators. The motivation behind these final controls is that different types of patients could arrive to the hospital during the beginning or end of the month, perhaps driven by holidays, paychecks, or benefit schedules.[3]

Under "Sample changes", I consider some possible reasonable modifications to the sample and construction of W_{-i}.[4] First, the primary sample considers any appointment within a 90-day window post ED visit to be a follow up appointment. In contrast, in the first row, I only consider follow up appointments within a 60-day window; patients with a first appointment scheduled between days 61 and 90 are dropped. In the second row of this section, related to the concern that holidays may create non-idiosyncratic variation in crowding, I drop all ED visits in December.

Next, I present results using a slightly modified version of W_{-i}. One potential concern is that wait times of patients who arrive during the same time window as a given patient may be influenced by the patient's own wait time (for example, if a high risk patient is pushed to an earlier appointment slot, this may cause more delays for other patients who arrive at a similar time). To mitigate this concern, in this alternative version of W_{-i}, I leave out wait times of not only the patient themselves, but of all patients who arrive on the same day. This robustness check also accounts for the potential of correlated same-day patient shocks, such as multiple veterans ingesting the same contaminated drug.

In Appendix Table A5, I perform an additional series of robustness checks. To capture potential correlation in treatment assignment within a given hospital, in column 2, I use hospital level rather than hospital-month level clustering. In column 3, I add day of week fixed effects. In column 4, to capture potential seasonal variation in types of patients arriving, I interact month of the year with ICD-9 3-digit diagnosis group.

Additionally, one potential concern is that those with prior mental health utilization are differentially affected by congestion. Those with prior mental health utilization likely have an advantage in obtaining timely follow up appointments, as they may have pre-existing relationships with providers. However, W_{-i} does not use a patient's wait time directly; instead, it relies on the wait times of patients arriving at the same time as the index patient. Consequently, W_{-i} should exogenously capture clinic congestion, rather than capturing anything about a patient's prior utilization. However, as an

[3]Note that ability to pay is not a major concern here, since the vast majority of mentally ill patients receive free mental health services at the VA.

[4]When the main sample selection is modified, the construction of W_{-i} is held constant, unless otherwise noted.

additional test to ensure that W_{-i} only captures congestion, rather than variation in patients' prior utilization which may affect wait times, in column 5, I perform a robustness check where I construct W_{-i} only using patients with no prior mental health utilization.

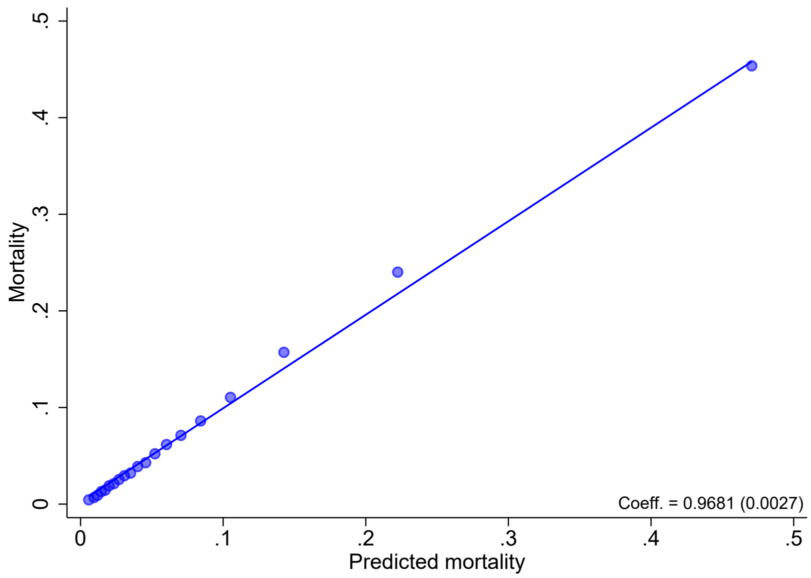
Finally, although all patients in my sample have a mental health diagnosis code connected with their visit, some patients who visit the ED may be experiencing both a mental and physical health emergency. Specifically, 39.9% of the sample also has a *physical* health diagnosis connected to their ED visit, the most common of which are hypertension, sleep disturbances, and respiratory distress. Some of these may be directly connected to mental health – for instance, those with anxiety may experience breathing issues or difficulty sleeping. In column 6, I add controls for physical health diagnoses assigned at the ED visit, including an indicator for no physical health diagnosis.

Table A1: Patient characteristics across samples

	Main sample	Patients with no other health insurance	Patients with no scheduled follow up visit
Male	0.908	0.905	0.927
Black	0.232	0.254	0.205
Hispanic	0.057	0.059	0.048
Age	51.4	47.9	57.8
Prior outpatient visit	0.573	0.547	0.188
Prior inpatient visit	0.096	0.103	0.027
Substance use disorder	0.285	0.321	0.446
Mood disorders	0.299	0.300	0.142
Anxiety	0.120	0.115	0.140
Psychosis	0.092	0.083	0.070
PTSD	0.103	0.101	0.048
Medicare	0.168	0.000	0.247
Other non-VA insurance	0.135	0.000	0.131
Mean wait time	17.73	17.55	18.30
Mortality	0.075	0.053	0.137
N	621,289	432,863	314,554

Note: This table shows characteristic means for the baseline sample compared to both the subsample of patients with no non-VA insurance and the dropped sample of those who have no scheduled follow up mental health visit. Prior oupatient and inpatient visit are indicators for attending a mental health appointment in the previous year. Mean wait time is equivalent to W_{-i} for the baseline sample. To form mean wait times for patients without a follow up mental health visit, I group all appointment types together, as I cannot observe which type of mental health clinic they would have been scheduled for.

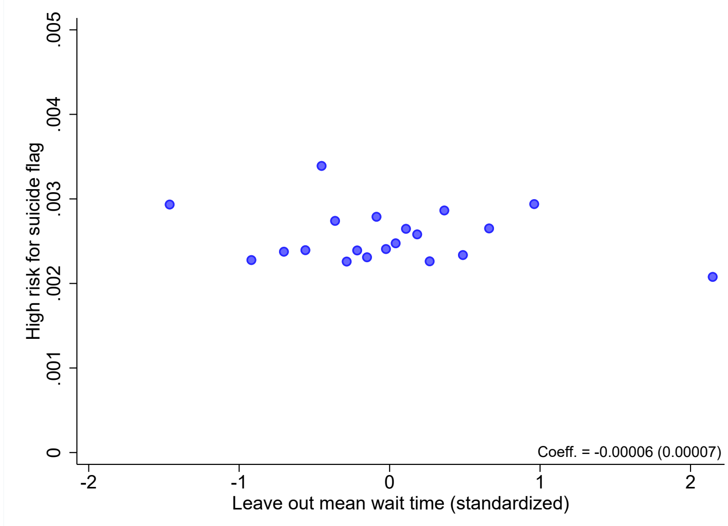
Figure A1: Predictive power



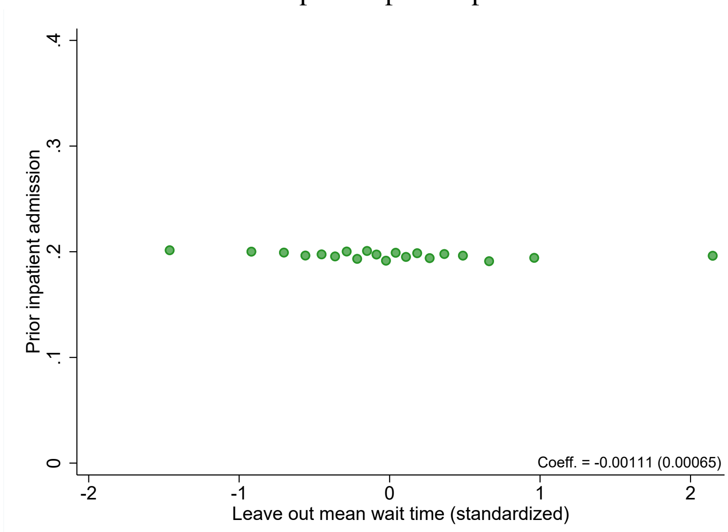
Note: This figure plots observed mortality, Y_i, versus predicted mortality, $\hat{Y}_i$, in twenty equal sized bins. The coefficient shows the regression fit of observed mortality on predicted mortality.

Figure A2: Alternative balance tests

A: Balance with respect to high risk for suicide flag



B: Balance with respect to prior inpatient admission



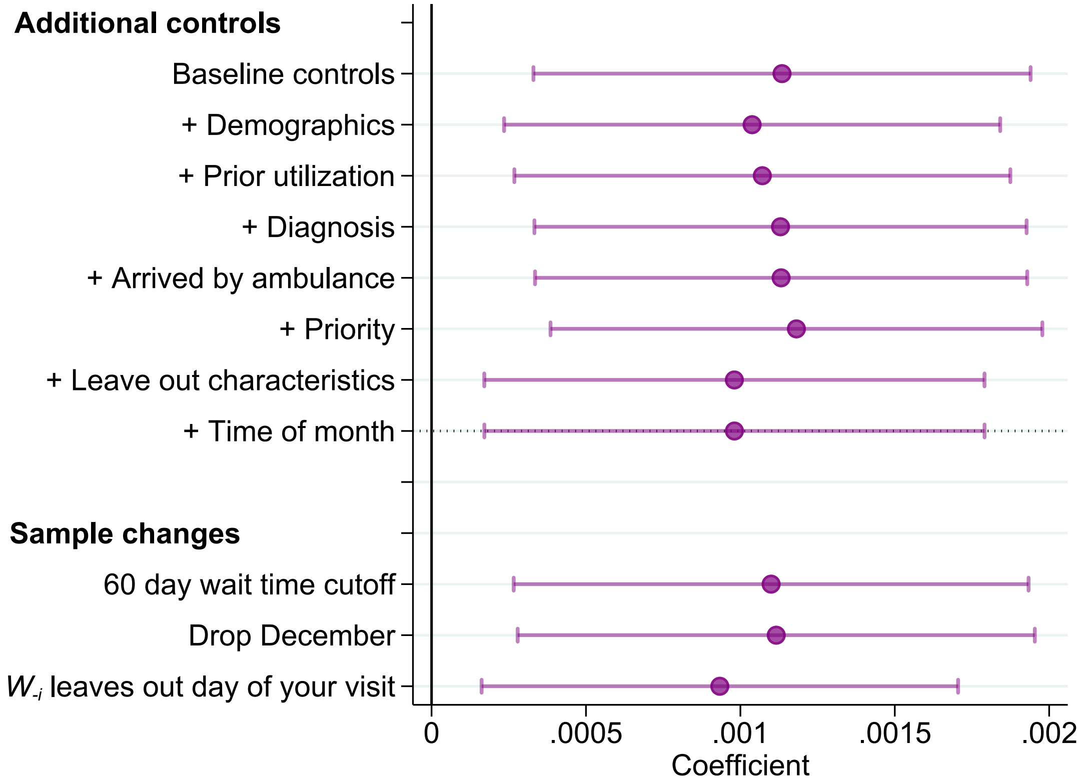
Note: Panel A shows a binscatter of the percent of patients assigned the high risk for suicide flag versus W_{-i}, residualized by the baseline controls. Panel B shows a binscatter of prior inpatient admission, I_i, versus W_{-i}, residualized by the baseline controls.

Table A2: Selection into appointment type

	General	Psychiatry	PTSD	Substance use	Other
Mean wait time$_k$	-0.0047	-0.0035	-0.0027	-0.0018	-0.0017
	(0.0011)	(0.0007)	(0.0011)	(0.0005)	(0.0007)
Mean wait time$_k \times \hat{Y}$	-0.0043	-0.0053	-0.0013	0.0002	-0.0013
	(0.0011)	(0.0007)	(0.0011)	(0.0006)	(0.0008)
Outcome mean	0.612	0.108	0.069	0.059	0.152
Observations	621,289	621,289	621,289	621,289	621,289

Note: This table shows regression results corresponding to Eq. 5 in the main text. The left hand side variable is an indicator for the patient being assigned a given appointment type (general, psychiatry, PTSD, substance use, or other specialty appointment). The first row shows β_1, the coefficient on residualized wait time for appointment type k, and the third row shows β_3, the coefficient on the interaction between residualized wait time and predicted mortality. Wait times are residualized by hospital $\times$ month, two week period, and age buckets. $\hat{Y}_i$ is standardized for ease of interpretation. General includes therapy visits with psychologists and social workers. All standard errors are clustered by hospital-month.

Figure A3: Robustness to additional controls and sample changes



Note: This figure shows the effect of not attending a mental health appointment on 2-year mortality when various changes to the controls, sample, or wait time measure are made. Dots are point estimates; bars show 95% confidence intervals. Under "Additional controls," I first show the main estimate using the baseline controls, two-week period, year-month × hospital, appointment type, and age buckets. I then incrementally add additional groups of controls. Demographics includes sex, ethnicity, and race. Prior utilization includes prior mental health outpatient visits, no shows, and cancellations, as well as prior inpatient mental health visits. Diagnosis refers to 3-digit ICD-9 diagnosis codes from the ED visit. Priority status, a designation between 1 and 8 assigned by the VA based on income and disability status, indicates eligibility for VA healthcare and copay responsibilities. Leave out characteristics include mean age, prior visits, and predicted mortality among patients in one's hospital-time-appointment group (the same group that forms a patient's W_{-i}). Time of month are indicators for first or second half of the month. Under "Sample changes," I make various modifications to the sample selection or construction of W_{-i}. The first row amends the definition of follow up appointment from 90 days to 60 days, the second drops ED visits in December, and the third row excludes all visits on the same day as the patient themselves from the construction of W_{-i}.

Table A3: Main results, including patients with no scheduled follow up

	2-year mortality			
	(1) Average wait time	(2) Average wait time	(3) Predicted wait time	(4) Predicted wait time
Mean wait time	0.00070	0.00095	0.00118	0.00124
	(0.00034)	(0.00035)	(0.00035)	(0.00035)
Outcome mean	0.0956	0.0956	0.0956	0.0956
Observations	935,843	935,843	935,843	935,843
Baseline controls	Yes	Yes	Yes	Yes
Additional patient controls	No	Yes	No	Yes
Leave out controls	No	Yes	No	Yes

Note: This table shows baseline mortality results, adding patients back to the sample who had no scheduled follow up appointment. "Average wait time" refers to forming W_{-i} by grouping all appointment types together (for those without a scheduled follow up). Predicted wait time refers to setting $W_{-i} = \hat{W}_{-i}$ for these patients. All standard errors are clustered by hospital-month.

Table A4: Effect of mental health wait times on mortality for no insurance subsample

	(1) No visit	(2) Mortality	(3) Mortality	(4) Mortality
Mean wait time	0.0218	0.00075	0.00082	0.00078
	(0.0010)	(0.00044)	(0.00043)	(0.00044)
Outcome mean	0.3624	0.0528	0.0528	0.0528
Observations	432,863	432,863	432,863	432,863
Baseline controls	Yes	Yes	Yes	Yes
Additional patient controls	No	No	Yes	Yes
Leave out controls	No	No	No	Yes

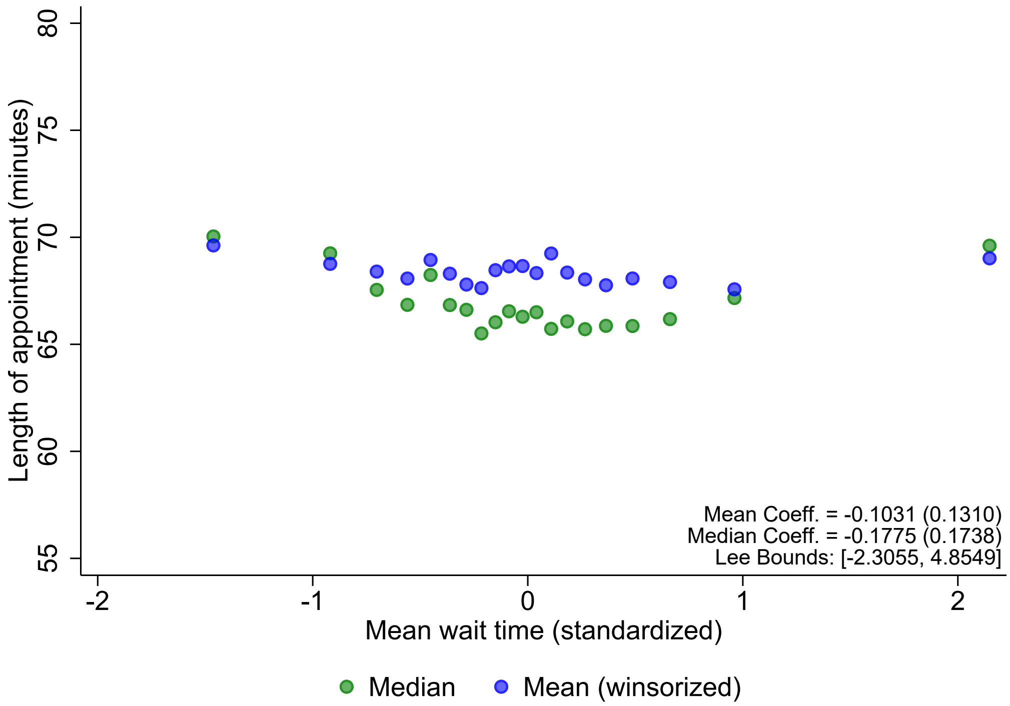
Note: This figure shows the main results for the subsample who have no other non-VA insurance. Column 1 shows the effect of mean wait time on the probability of not attending a follow up mental health visit. Column 2 shows the main mortality estimate, only using baseline controls. Columns 2 shows the estimate when I add additional patient controls, including gender, race, diagnosis, and prior utilization. Finally, column 4 shows the estimate when I control for characteristics of other patients who form a given patient's value of W_{-i}, including (leave-out) mean age, mean prior visits, and mean predicted mortality. Baseline controls include month-year $\times$ hospital, appointment type, and age buckets. Standard errors are clustered by hospital-month.

Table A5: Additional robustness checks

	(1) Baseline	(2) Hospital-level clustering	(3) Day of week controls
Mean wait time	0.00113	0.00113	0.00114
	(0.00041)	(0.00044)	(0.00041)
Outcome mean	0.0746	0.0746	0.0746
Observations	621,289	621,289	621,289
	(4) Month × disease type controls	(5) Mean wait time excludes those with prior utilization	(6) Physical health diagnosis controls
Mean wait time	0.00121	0.00096	0.00098
	(0.00041)	(0.00040)	(0.00041)
Outcome mean	0.0746	0.0751	0.0746
Observations	621,289	575,734	621,289

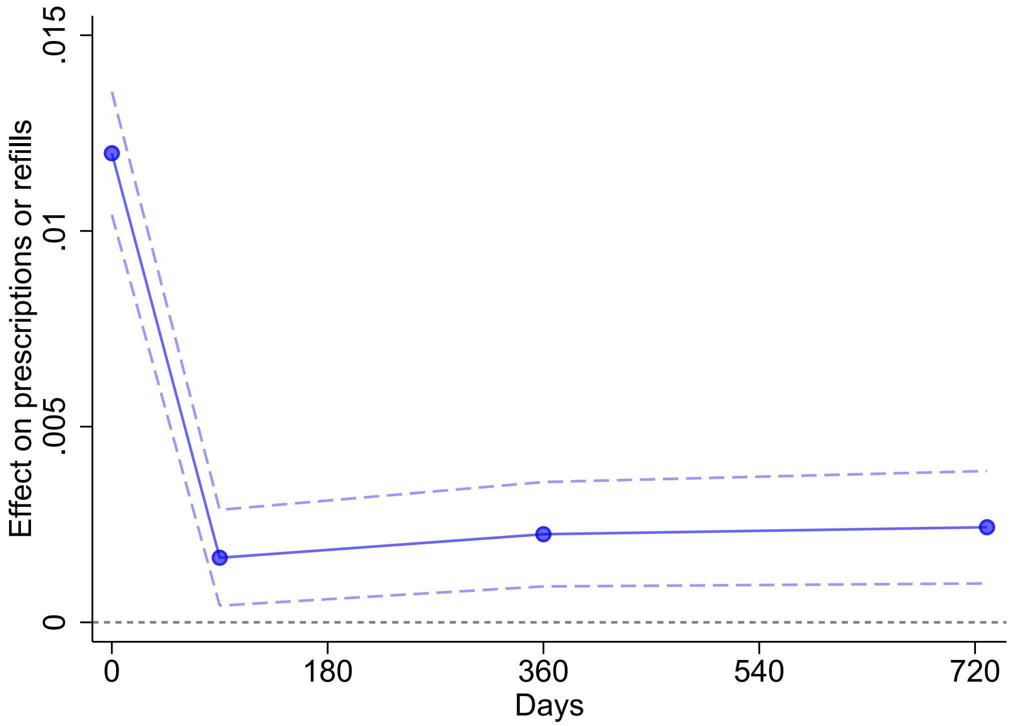
Note: This table shows additional robustness checks for the mortality estimate. The fourth column interacts month of the year with ICD-9 3-digit diagnosis group. The fifth column constructs W_{-i} only using patients with no prior mental health utilization. The sixth column controls for physical health ICD-9 codes assigned at the ED visit.

Figure A4: Length of follow up appointment vs. mean wait time



Note: This figure shows a binned scatterplot of the length of one's follow up appointment, in minutes, versus W_{-i}, residualized by the baseline controls. The mean reported coefficient winsorizes observations at the 5th and 95th percentiles, while median reported coefficient calculates the median wait time in hospital-two-week-appointment type bins. Lee Bounds reports worst case scenario bounds using median wait times.

Figure A5: Effect of mean wait time on prescriptions over time



Note: This figure shows the effect of mean wait time on mental health prescriptions or refills. The first dot shows the effect of mean wait on having any drug prescribed at the ED visit. The second dot shows the effect on having any drug prescribed between days 0 and 90 (including the day of the ED visit). Finally, the third and fourth dots show the effect on prescriptions or refills between days 90-360 and 360-730, respectively.

Table A6: Cause of death

	High wait time patients	Full sample	Risk ratio
All cause mortality	0.0756	0.0745	1.015
	(0.0004)		
Physical disease	0.0660	0.0648	1.018
	(0.0004)		
Accident, suicide, or overdose	0.0096	0.0096	0.993
	(0.0002)		
Heart disease	0.0169	0.0163	1.034
	(0.0002)		
Cancer	0.0131	0.0130	1.006
	(0.0002)		
Respiratory diseases	0.0071	0.0068	1.045
	(0.0001)		
Liver disease	0.0049	0.0047	1.025
	(0.0001)		
Other cardiovascular	0.0038	0.0036	1.034
	(0.0001)		
Infections	0.0030	0.0029	1.039
	(0.0001)		
Other physical disease	0.0203	0.0203	1.001
	(0.0002)		

Note: This table shows the probability of dying from a particular cause for both the full sample and for those with a mean wait time one standard deviation above the mean. The last row shows the relative risk of dying from that cause for those with high wait times versus the full sample.

Table A7: Effect of mean wait time on intermediate outcomes for substance use subsample

	A: Mental health outcomes		
	(1)	(2)	(3)
	No follow up visit	Returned to mental health care	Mental health inpatient treatment
Mean wait time	0.0239	-0.0104	-0.0088
	(0.0016)	(0.0009)	(0.0015)
Outcome mean	0.3385	0.9286	0.4151
Observations	175,721	175,721	175,721
	B: Physical health outcomes		
	(4)	(5)	(6)
	Primary care visit	Inpatient visit - non-mental health	New physical comorbidity
Mean wait time	0.0027	0.0044	0.0032
	(0.0011)	(0.0013)	(0.0014)
Outcome mean	0.8317	0.2323	0.6529
Observations	175,721	175,721	175,721

Note: This table shows intermediate outcomes for the SUD subsample. Panel A shows results for engagement with mental health care. Returned to mental health care refers to attending any oupatient appointment. Mental health inpatient treatment refers to any mental health inpatient visit, including residential stays (excluding admissions within one day of the initial ED visit). Panel B presents results for non-mental health outcomes. The first column looks at whether a patient attended a primary care appointment, the second column looks at non-mental health inpatient admissions, and the third column looks at whether the patient had any new Elixhauser comorbidity (only including physical conditions) during the two-year follow up period.