NET WORK LISER

Discussion Paper Series

IZA DP No. 19038

October 2026

The 988 Suicide and Crisis Lifeline and Youth Suicide

Dave E. Marcotte
American University
and IZA@LISER

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA@LISER Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.

IZA DP No. 19038 October 2026

The 988 Suicide and Crisis Lifeline and Youth Suicide

Abstract

In July 2022 the ten-digit National Suicide Prevention Lifeline was replaced with a dedicated three-digit number, 988. This followed a precipitous rise in the suicide rate, especially among youth. Suicide is now the second leading cause of death among Americans aged 15 to 24. In this paper, I estimate the effect of 988 on youth and overall suicide rates. I show that the rollout of 988 increased the number of calls to suicide hotlines and the share of calls that were answered. Nonetheless, I find no evidence that 988 reduced suicide mortality. Using multiple identification strategies, I find no reduction following the national rollout in July 2022, nor do I find evidence that state-level improvements in call answer rates reduced suicide. These findings do not imply that 988 has been ineffective: the rollout clearly improved hotline access and response rates. However, I find no evidence that it reduced emergency-department visits for injuries caused by self-harm.

JEL classification

I10, I18

Keywords

suicide, prevention, evaluation

Corresponding author

Dave E. Marcotte
marcotte@american.edu

I. Introduction

Suicide is now the second leading cause of death among Americans aged 15 to 24, and the youth suicide rate has risen nearly 60 percent since the turn of the century. Between 2001 and 2024, the number of 15- to 24-year-olds who died in motor vehicle accidents fell from roughly 7,600 to 6,700, while the number who died by homicide rose modestly, from about 5,300 to 5,700. Over the same period, the number who died by suicide more than doubled, from 3,800 to 8,800. No other cause of death among young Americans has risen so precipitously, yet there is little credible causal evidence explaining the increase.

In July 2022 the United States replaced the ten-digit National Suicide Prevention Lifeline (1-800-273-8255) with a dedicated three-digit number, 988, styled after 911. The rollout was intended to make reaching suicide prevention services faster and easy to recall for those in crisis. It was also accompanied by public service advertising online and in public places. In states that already funded and staffed an effective crisis-line system, the transition amounted to a rebrand. Other states, however, had poorly funded response systems and the new 988 system included funding for expansion in staffing and call center capacity. Some states funded the 988 rollout with new per-line telecommunications fees or one-time appropriations; others took no legislative action at all. The federal government has already committed over $1.6 billion for the support and operation of the new 988 system, and states have allocated an additional $500-700 million.[1] Despite its public health and budgetary significance, there is little causal evidence linking crisis-line performance to suicide among the age group most likely to use 988 by call, text, or chat.

As we will see, 988 increased the number of calls to suicide hotlines and increased the share of calls answered, and these changes were more pronounced in some states than others. I

[1] https://www.gao.gov/assets/gao-26-107915.pdf

use this variation to estimate the effect of 988-driven improvements in crisis-line service quality on suicide. I focus on suicide mortality for young persons because they have seen the steepest rise in suicide and because there is evidence that youth are more likely to engage with help lines when in crisis. I also estimate impacts on suicide more broadly, a question central to assessing the efficacy of 988. The national rollout of 988 could have reduced suicide through two potential channels. First, the national launch brought immediate attention to an accessible, dedicated hotline. Second, the associated federal and state investments in suicide hotline infrastructure resulted in a larger share of calls being answered by response centers in the same state where calls originate. This meant calls were answered more quickly and were less likely to need re-routing to backup centers elsewhere, a process that increased chances that the caller would abandon the call.

In Figure 1, I illustrate the salience of these potential channels. The figure presents national call volume to emergency suicide hotlines and population-weighted answer rates for the year prior to the July 2022 launch of 988 and then for the subsequent three-plus years. In addition to the time series for both calls and answer rates, Figure 1 also includes linear trends for both before and after the national 988 launch. It is apparent that the switch to the three-digit number 988 resulted in a sharp jump in the rate at which calls to suicide hotlines were answered. At the national level the answer rate jumped in 2022 from 72 percent in June to 81 percent in July, when 988 was launched mid-month. Answer rates were 84% in the first fully operational month of August. There was also an increase in call volume, even if less abrupt. After July 2022, the rate of growth in calls to suicide hotlines increased, even if there was no break in trend. Because of this higher volume and improvement in answer rates, it is clear that more people were connected to suicide hotline help in the period following the 988 launch.

To assess the impact of 988 on suicide mortality, I adopt several estimation strategies. First, I estimate whether the shift apparent in Figure 1 translates into a discrete shift in suicide following the July 2022 launch. Since states vary in their call volume and answer rates before the national launch, I estimate the dose-response relationship separately at each event-time period using each state's overall gain in answer rate.

I also parameterize average reduced form effects using a two-way fixed effects difference-in-difference estimate for some insight into whether states with the largest and earliest improvements in 988 call and answer rates saw correspondingly larger declines in suicide mortality. However, interpreting a TWFE DiD model in this context is hazardous because it requires a strict generalized parallel-trends assumption within levels of treatment intensity. To limit this threat to validity, I then estimate dose-specific average treatment effects, *ATT(d)*, following the continuous-treatment difference-in-differences framework of Callaway, Goodman-Bacon, and Sant'Anna (2024). Each state serves as its own comparison at various levels of post-988 call center performance.

I also estimate treatment effects for answer rates using an instrumental variables estimator. I instrument changes in a state's answer rate with the timing of its 988 fee and appropriation legislation in a fuzzy DiD-IV framework following de Chaisemartin, D'Haultfœuille, Pasquier, Sow, and Vazquez-Bare (2024). This IV strategy leverages the portion of state policy responses due to budgetary capacity to support the rollout of 988, rather than state-specific youth mental-health trends.

Finally, I assess whether 988 has had an effect on suicidal behavior more broadly. Mortality is the outcome that matters most, but it is also rare with limited within-state variance. I supplement the mortality analysis using data on emergency-department (ED) visits from the

State Emergency Department Databases (SEDD) for Michigan, North Carolina, and Rhode Island. These states did not enact 988 fee or appropriation legislation during the panel and yet saw very different patterns in 988 call volume and response quality following the national rollout. This high vs. low intensity comparison provides insight on a higher-frequency, more prevalent outcome, visits to hospital emergency departments for self-harm and suicide attempts. Changes in suicide attempts in response to any suicide hotline quality changes induced by 988 are more likely to be detectable than any mortality effects.

My primary focus is on youth suicide, since teens and young adults in their early twenties are at special risk and the heaviest users of 988 by call, text, and chat. But restricting the analysis to youth has a cost: suicide counts at the state-month level are already small, and narrowing further to a ten-year age band produces more zero-count cells and widens confidence intervals. While I present the youth estimate as the primary result, I also include an all-age suicide-count estimate. The all-age estimate is better powered and relevant to understanding whether crisis-line capacity reduced suicide mortality overall. The youth-specific estimate assessed whether any effect is concentrated in the population that generates more calls to 988.

Finally, I compare results here to Hansen et al. (2026), who study the same policy but use a different identification strategy. They measure each state's exposure using its pre-988 baseline capacity gap as measured by routed-but-unanswered calls per 1,000 residents in the months just before national rollout, rather than the state's realized, contemporaneous change in answer rates. Using an event-study design around this fixed baseline measure, they find a strong first stage, with a 24 percent increase in in-state answer rates. Their reduced-form mortality effect, however, is small and statistically indistinguishable from zero. Hansen et al. decline to treat their 2SLS

estimates as their preferred causal parameter, citing a demanding exclusion restriction and a noisy first stage. This is an issue I discuss, below.

The remainder of the paper proceeds as follows. Section II reviews trends in youth suicide, the institutional history of crisis lines before and after 988, and the state legislative efforts to support it. Section III describes the data. Section IV lays out the empirical strategy: the reduced-form continuous DiD, the event-study specification, the IV-DiD robustness check, and the SEDD case-study extension. Results and discussion of their implications follow in sections V and VI.

II. Background

A. Youth Suicide

The U.S. suicide rate has risen by roughly 40 percent since 2000, reversing decades of decline. The increase has been sharpest among the young: the suicide rate for 15- to 24-year-olds has risen almost 60 percent over the same period, and suicide rates among people in their twenties are now as high as those for any other age group. This is a pattern with no recent precedent. In 2023 the U.S. Surgeon General, Vivek Murthy, identified declining adolescent and young-adult mental health as "the defining public health issue of our time.[2]" Rising youth suicide is the starkest manifestation of that crisis.

Marcotte and Hansen (2024) examined broader trends in suicide in the U.S. and their likely causes. They concluded that economic hardship associated with the Great Recession explains much of the variation this century among prime-age adults. For those under 25, the bulk of the increase in suicide mortality since 2012 is better explained by a rise in the prevalence of untreated depression. Bullying victimization does not appear to drive the aggregate trend, though

[2] https://www.hhs.gov/sites/default/files/surgeon-general-youth-mental-health-advisory.pdf

it may account for part of the rise among LGBTQ youth. The roles of firearm access and the opioid epidemic in recent youth suicide trends remain less clear.

B. Crisis Lines Before 988

Because many people contemplating suicide communicate their intent or reach out to others while in crisis, telephone hotlines have long been considered a first line of defense in suicide prevention (Gould, Kalafat, Harris-Munfakh, and Kleinman, 2007). There is a well-developed clinical literature on the characteristics and short-term efficacy of crisis hotlines (Gould and Kalafat, 2009). The goals of a suicide hotline are to de-escalate the caller's immediate state of crisis and to connect them to emergency and/or follow-up mental-health care.

The system that would become 988 began in 2005 with the launch of the National Suicide Prevention Lifeline (NSPL), a ten-digit number that routed callers to a network of state and local crisis centers. That network grew from 109 call centers in 2005 to 170 by 2018, when Congress passed the National Suicide Hotline Improvement Act directing the Federal Communications Commission and the Department of Health and Human Services to study a three-digit alternative. Staffing and funding for individual call centers varied enormously, and calls that a local center could not answer were either abandoned by the caller or transferred to the national backup network, where hold times were typically longer. Because a slow or failed local response raised the odds that a caller disconnected before reaching a counselor, the in-state answer rate is a key measure of crisis-line system quality and response capacity and the outcome around which the entire empirical strategy in this paper is built.

C. The 988 Transition and State Policy Variation

The NSPL migrated to the 988 dialing code over a three-year period beginning in 2019 and culminating in nationwide activation on July 16, 2022. The 988 Lifeline can also be reached

by SMS text and online chat. While the nationwide activation date was common to every state, its practical consequences were not. States that had already invested in local call center capacity experienced 988 as a rebranding of an already-functional system. States that had not saw their answer rates climb (in some cases from under 60 percent to over 90 percent) as federal SAMHSA capacity grants and state-funded supplements took effect. In Figure 2, I provide some sense of differences in improvements in answer rates across states. The figure is a population-weighted scatterplot of changes in answer rates over the panel by the average per-capita call volume in the year prior to the launch of 988. Some states saw large increases in the percent of calls answered by in-state emergency hotlines. For example, Illinois saw its answer rate increase by more than 60 percent, and Texas by more than 40 percent. Clearly, these states had substantial room for improvement. States with no change in answer rates like California and Rhode Island had answer rates above 90 percent even before 988.

To shed light on some of the differences in the changing quality of service among suicide hotlines, I compiled a state-by-state database of 988-related legislation covering 26 states that took some legislative action to bolster their systems between 2021 and 2025. The most common strategy was to implement a per-line telecommunications fee dedicated to fund the system, modeled on the existing 911 surcharge. Virginia was the first to implement a surcharge, enacting a $0.12-per-line fee effective July 2021, followed by Colorado, Nevada, Washington, California, and a dozen other states through 2025, with per-line amounts ranging from $0.08 to $0.60. Several states also adjusted their per-line fees up or down after initial enactment. Other states relied instead on one-time or biennial general-fund appropriations. For example, Ohio allocated $46.5 million for 988 support in their FY2024–2025 budget, the largest single allocation in my data. A small number of states created oversight commissions or *ad hoc* study groups without

committing new funding. Finally, many states took no legislative action of any kind during the study period.

This legislative variation is the source of the instrument I use below. Because most of these fee and appropriation bills were introduced and passed for ordinary state budgetary reasons, their timing is potentially unrelated to state-specific trends in youth mental health, conditional on state and calendar-month fixed effects. I discuss the plausibility of this exclusion restriction further in Section IV.

D. Related Literature

There is a vast clinical and public-health literature on hotline efficacy, relying mainly on within-call survey evidence of reduced caller distress (Gould, Kalafat, Harris-Munfakh, and Kleinman, 2007; Gould and Kalafat, 2009). This literature suggests that crisis calls reduce acute suicidality at the moment of contact. However, it is necessarily limited and provides no evidence about whether call centers reduce suicide attempts or deaths at the population level.

A second, much thinner literature evaluates 988 itself. The most widely known paper is Patel, Liu, and Jena (2026), a research letter in *JAMA* that has shaped much of the public narrative around the lifeline's effectiveness. Using National Vital Statistics System death-certificate data from 1999 through 2022, the authors project a counterfactual quarterly suicide-mortality series for 15- to 34-year-olds absent 988's launch and compare it with observed mortality from July 2022 through December 2024. They report 35,529 observed suicide deaths compared to their projection of 39,901 absent 988. This would be an 11.0 percent reduction, or roughly 4,372 fewer deaths than expected subsequent to the launch of 988.

The Patel et al. (2026) design has real limitations that this paper attempts to address. Their counterfactual is a historical-trend extrapolation rather than a comparison to an observed

control group, leaving the estimate vulnerable to any other national-level shock to youth mental health coincident with the 2022 launch of 988. During this time, post-pandemic mental-health trends and opioid-related mortality were also shifting. Their analyses are at the aggregate, national level. They do estimate differences between the 10 states with the largest increases in call answer rates and the 10 states with the smallest increases and estimate more averted suicides in the former. However, this dose-response comparison also makes no use of within-state variation, nor do they assess parallel-trends between these groups. The authors are themselves candid about these limitations and interpret their findings as an early signal, not a causal estimate.

The paper most directly relevant to my paper is Hansen et al. (2026). Their design measures treatment exposure with each state compared to a pre-988 baseline capacity gap measured over the 6 months before 988. They estimate a reduced-form effect on suicide mortality in an event-study framework, finding no detectable mortality effect. A major difference between this paper and theirs is that they do not focus on realized, within-panel changes in answer rates for identifying variation. I use the panel's realized variation in answer rates as the primary treatment measure, and I also instrument changes in answer rates with legislative timing rather than relying on a static baseline measure of resource capacity. Further, Hansen et al. do not examine emergency-department visits for injuries caused by suicide attempts.

III. Data

I use data from the CDC cause of death mortality series to measure the count of suicide deaths among 15- to 24-year-olds by state and month, and suicide deaths for ages 15-74 by state and month. Below, I refer to the broader group ranging from 15-74 years old as the all age group. Because the CDC Wonder data suppresses cells with fewer than 10 deaths, I use restricted-use

NCHS mortality microdata provisional multiple-cause-of-death files (ICD-10 underlying cause codes X60–X84, Y87.0, and U03). State-month cell counts can be small or zero for the youth series and occasionally for all ages in low-population states. So, I use Poisson pseudo-maximum-likelihood (PPML) as the primary estimator, but supplement with OLS models of suicide rates.

I obtain data on 988 call volume and performance from SAMHSA, which publishes state-month administrative data on 988 Lifeline call volume, the share of calls answered in-state, the share routed to the national backup network, average wait time, and the abandonment rate. These series begin in 2021 for most states, providing about a year of pre-national-rollout data with which to characterize each state's starting position, and extending through the end of the panel in 2024. I use in-state answer rate as the treatment of interest and include call volume as a measure of exposure.

I collected the legislative database from state legislative records, the FCC's annual 988 Fee Accountability Reports, NAMI's state legislative tracker, and contemporaneous news coverage, cross-checking each entry against at least two independent sources. For each of the 26 states with legislative action, the database records the policy type (fee, appropriation, trust fund, or study commission), the enacting bill, enactment and effective dates, the fee amount per line where applicable, the appropriation amount where applicable, and whether the funding mechanism is recurring. I use these measures as an instrument for call answer rates.

I combine these data into a state-by-month panel spanning January 2018 through December 2024, covering the 50 states and the District of Columbia. I include data on population by age group at the state level from the Census Bureau's population estimates program and state by month unemployment rates from the Local Area Unemployment Statistics program of the Bureau of Labor Statistics.

I supplement the analyses of suicide mortality using data on hospital emergency-department visits for self-harm and suicide attempts from the Agency for Healthcare Research and Quality's State Emergency Department Databases (SEDD), part of the Healthcare Cost and Utilization Project (HCUP). SEDD records nearly universal, all-payer emergency room encounters within a state, with information on patient demographics and ICD-10-CM external-cause-of-injury codes. I use these data to identify self-harm visits to emergency departments (ICD10 codes X71–X83 and T14.91) by patient age group.

I focus this extension on three states, Michigan, North Carolina, and Rhode Island for several reasons. First, all participate in SEDD across the panel period with consistent ICD-10-CM coding. This is not the case for every state. Second, they each present different trajectories in in-state answer rates over the panel, with Michigan seeing a substantial rise in answer rates, North Carolina in call volume, and Rhode Island seeing no real change in either. Finally, SEDD data are expensive, and it was practically impossible to add additional states.

IV. Empirical Strategy

There are a variety of possible strategies for estimating the effect of 988 and subsequent improvements in crisis-line service quality on suicide. One possibility is to estimate dose-response type relationships between observed changes in call answer rates and suicide rates. Another is to rely only on changes in suicide within states with different baseline capacities for improvement after rollout of 988, treating the observed post launch as endogenous to other state-level factors shaping suicide after July 2022. A third is to instrument for state-level 988 service improvements, using federal and state investments in call center systems. Each of these presents its own challenges for identification and has its own limitations. I employ all of these strategies and discuss the strengths and limitations of the relevant estimators, below.

A. Dynamic Event-Study

To test whether 988 induced a discrete shift in suicide rates following its launch in July 2022, I estimate a dose-response relationship between answer rate changes and suicide for each month after launch. The dose is a measure of each state's realized gain in answer rate over its pre-launch baseline. I define $AnswerRateGain_s$ as the mean answer rate for a state in all months after July 2022, net of the mean answer rate for that state before that month. This time-invariant measure captures the total improvement in state s's call-center performance after the national launch. The event study then estimates whether suicide rates evolved differently in states with larger improvements. Let k index event time in months relative to national launch (July 2022), with $k = -1$ (June 2022) as the omitted reference period, and define $Dose_{s,k} = AnswerRateGain_s \cdot 1\{EventTime_t = k\}$. I estimate the PPML event-study specification as:

$$(1)\quad E[Deaths_{st}] = Pop_{st} \cdot exp(\Sigma\beta_k Dose_{s,k} + \delta \cdot CallsPer1000_{st} + \theta \cdot Unemployment_{st} + \alpha_s + \gamma_t)$$

I also estimate a linear analogue on the rate per 100,000 as a companion specification. I plot dynamic coefficients from July 2022 through the end of the panel, testing formally for the absence of differential pre-trends across high- and low-exposure states in the pre-988 period (Abadie, Angrist, Frandsen, and Pischke, 2026).

B. Changes by Baseline Capacity Gap

One concern about the event study specification is that a state's improvement in call center performance may be endogenous to changes in need or other changes that may affect mental health or suicide in a state. For this reason, I estimate a specification based only on each state's pre-988 baseline capacity gap, rather than the $AnswerRateGain_s$ realized after 988 launched. States with a lower baseline answer rate have more room to improve, so this measure

is fixed in the pre-period and does not depend on state patterns after 988 launched. This is a specification similar to Hansen et al.'s (2026) dose measure.

Here, I let $Dose_{s,k} = (1 - AnswerRate_pre_s) \cdot 1\{EventTime_t = k\}$, where $AnswerRate_pre_s$ is state s's pre-launch baseline answer rate. Unlike the realized, post-period dose $AnswerRateGain_s$ that would be confounded by any state level investments in mental-health infrastructure, $AnswerRate_pre_s$ is measured before 988 launched and therefore cannot be affected by any subsequent investments. As in equation (1), I estimate both the PPML and linear versions of this specification, with standard errors clustered at the state level.

C. Discrete-Dose Difference-in-Differences

Because states see different patterns of growth in answer rates, the event-study and baseline gap models above ignore within-state variation. A standard TWFE DiD model would identify off of variation in the in-state 988 answer rate for state s in month t over the full period, rather than from a binary treated/untreated indicator or by a static or dynamic parameter fixed to a state's actual or potential improvement. A TWFE DiD specification has a number of limitations, however. First, there is no clear test of the effect of 988, *per se*, just of changes in a continuous measure of answer rates that could even precede 988. Second, in such a model estimates of the effects of a continuous treatment include weighted averages of treatment effects across different levels of treatment and treatment years, as well as different levels of selection bias at those treatment margins (Callaway, Goodman-Bacon, and Sant'Anna, 2024, hereafter CGS). Unlike the dichotomous treatment effect case, there is not a fixed number of underlying 2x2 estimands.

As a discrete-dose counterpart to provide clearer insight into the effects of 988, I bin each state's realized answer-rate gain into quartiles and interact the quartile indicators with a post-launch indicator. The specification is:

$$(2)\quad SuicideRate_{st} = \Sigma_{q=1}^{4} \beta_q \cdot (Quartile_{q,s} \times Post988_t) + \delta \cdot CallsPer1000_{st} + \rho \cdot log(Pop_{st}) + \theta \cdot Unemployment_{st} + \alpha_s + \gamma_t + \varepsilon_{st}$$

where $Quartile_{q,s} = 1\{\text{state } s \text{ falls in dose-gain quartile } q\}$ and $Post988_t = 1\{t \geq \text{July } 2022\}$. Here, identification comes directly from 988's launch, differencing off realized differences in treatment dose across states. I estimate this specification separately for the youth and all-age samples, and standard errors are clustered by state. I estimate the linear specification only. Equation (2) is reported as a summary of the quartile-level dose-response pattern but has the usual TWFE caveat under staggered or heterogeneous timing of changes in the level of answer rates within states (Goodman-Bacon 2021; de Chaisemartin and D'Haultfœuille 2020).

D. Dose-Specific ATT(d)

Because of the limitations of the TWFE estimates in equation (2), I adapt Callaway and Sant'Anna's (2021) group-time treatment effect estimator following the extension of CGS (2024) for continuous treatment. I use the same dose quartiles defined in equation (2), $d \in \{1,2,3,4\}$ (Q1 = smallest realized gain, Q4 = largest), with the same high-baseline comparison group as the omitted reference. For each dose quartile d the estimand is:

$$(3)\quad ATT(g,t;d) = E[Y_t(g) - Y_{pre}(g) \mid G = g] - E[Y_t(0) - Y_{pre}(0) \mid \textit{not yet treated}]$$

To estimate average post-treatment effects periods by dose quartile, I use the aggregation of Callaway and Sant'Anna (2021):

$$(4)\quad ATT(d) = \Sigma_{g,t}\, w(g,t) \cdot ATT(g,t;d)$$

Because 988 launched nationally on a single date, every treated state is in the same treatment-timing group. The usual Callaway-Sant'Anna group dimension is replaced with the dose quartile, and the estimator tests whether states with larger realized answer-rate gains show

larger mortality effects than states with the smallest gains, rather than whether treatment timing itself matters. This estimator does not require the homogeneous-effects restriction of a TWFE DiD specification. It does, however, use $SuicideRate_{st}$ directly rather than the death count with population as the offset, and the Stata csdid package does not support population weighting. So $ATT(d)$ is an unweighted, state-average effect rather than a population-average one.

E. IV-DiD

The designs above require that conditional on state, month, and year fixed effects, states with larger realized changes in answer rates were not on different suicide trajectories before 988. This is a strong assumption since some state legislatures were making investments in suicide hotline infrastructure. I assess the value of state policies for empirical leverage in understanding changes in crisis-line quality using a design to isolate the component of a state's answer rate in a given month due to implementation of state fee and appropriation legislation. The instrument is coded from the policy database described above. I estimate the corresponding first stage, reduced form, and two-stage least squares specification, absorbing state, year and month fixed effects and clustering standard errors by state.

Identification requires two conditions. The instruments must satisfy parallel trends on their own, meaning that states adopting fee or appropriation legislation would have otherwise trended similarly to states that did not. Further, legislative adoption needs to affect suicide mortality only through effects on a state's 988 response, with no direct channel. This assumption would be questionable in cases where broader mental-health appropriations are bundled into the same bill.

An alternative approach is to treat legislative adoption as an instrument for the continuous treatment dose in a dynamic fuzzy DiD-IV framework suggested by de Chaisemartin,

D'Haultfœuille, Pasquier, Sow, and Vazquez-Bare (2024, hereafter dCDH). This strategy compares "switcher" states that implement a fee or allocate money between consecutive periods to "stayer" states with no such changes. An advantage of the dCDH approach is that it constructs a separate Wald-type estimate at each event-time period using only switcher/stayer comparisons within that period. This is akin to a series of separate 2x2s avoiding the bias that staggered, heterogeneous timing of instrument adoption can introduce.

However, restricting identification at each period to a subset of about a dozen switcher states severely taxes an already weak instrumental variables design. Given this, I do not report the dCDH estimator as a primary or robustness specification in the main text. For transparency, I report the youth dynamic estimates in Appendix Table A1. I omit the all-age results because the placebo test rejects parallel trends for the instrument ($p = .018$), which is not the case for the youth sample ($p = .154$).[3]

In the results below, I estimate all models for the population aged 15-24, and again for the population aged 15-74, which I refer to as the all-age count. The measures of 988 call volume and response, the instruments, and the state and calendar-month fixed-effect structure are identical across the two specifications. The only differences are the outcome and the exposure population used to scale rates.

F. Emergency-Department Visits

Finally, to explore the relationship between 988 and non-fatal suicidal behavior I construct a state-month panel of ED visits among 15- to 24-year-olds and for all-ages in

[3] In Appendix Table A1, the joint test of the twelve post-adoption event-time coefficients is significant ($p < 0.001$) and every point estimate is negative, consistent with a protective effect. However, five of the six placebo coefficients estimated in the pre-adoption periods are also negative, at a magnitude comparable to the post-adoption estimates. The similar patterns before and after adoption are more consistent with switcher states drifting differently from stayer states for reasons unrelated to legislative timing than with a treatment effect.

Michigan, North Carolina, and Rhode Island and estimate differences in visits for self-harm and suicide-related injuries. Since there are only three states, I do not attempt a formal cross-state comparison, but report within-state event-study estimates around each state's own 988 exposure path, using the pre-period as the counterfactual trend and the same state, month, and year fixed-effect structure as the main specification. This analysis provides descriptive evidence on a more proximate and higher-frequency outcome of suicidal behavior than mortality.

V. Results

In Table 1, I present summary statistics for the full state-year panel from 2018-2024. The average (annualized) suicide rate for youth 15-24 years old was 14 per 100,000, while the rate for those 15-74 was 17.3 per 100,000.[4] For the average state/month, there were 5.757 calls to 988 per 1,000 residents, with 82.4 percent of those calls answered in-state. Prior to the launch of 988, about 73.9 percent of calls were answered in-state, and on average answer rates rose by 12 percent after the rollout of 988. Some form of policy to support 988 was in place in 7 percent of state/month observations, and the average monthly charge dedicated to 988 was 1.2 cents per phone line.

A. Dynamic Event Study

In Figure 3, I present the event study estimates of the effects of realized gains in state answer rates in each month leading into and out of the 988 rollout in July of 2022. The figure has four panels, showing results from the regression estimates of suicide rates and then the PPML estimates of suicide counts for youth and all ages separately. As is clear in Figure 3, there is no

[4] The rates for 15–24-year-olds match those reported in CDC summary reports for the period https://www.cdc.gov/nchs/products/databriefs/db541.htm. The rates for 15-74 year olds are slightly higher than CDC-reported rates of "all-age" suicide. The CDC all-age rates include those under 15 (for whom suicide is very rare and who constitute about 17 percent of the population) and those over 75, for whom suicide rates are relatively low (https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-detail.html).

clean discontinuity in the dose-response relationship at national launch and no evidence of a relationship between the size of states' improvements in answer rates or a monotonic buildup over the post-launch period in any of the four panels. In all cases, coefficients before and after t=0 are similarly scattered around zero, rather than falling at launch and remaining below zero. Formally testing the pre-period coefficients for joint significance, three of the four specifications show no evidence against parallel trends prior to launch (PPML youth: $\chi^2(11)=11.67$, p=.389; linear youth: $F(11,50)=0.39$, p=.953; linear all-age: $F(11,50)=0.49$, p=.901). The exception is for the PPML all-age results, where the pre-period coefficients are jointly significant ($\chi^2(11)=22.86$, p=.019). Nonetheless, the null results across both samples and estimators are inconsistent with any interpretation of reduced suicide mortality.

B. Baseline Capacity Gap Models

In Table 2 I present estimates of suicide mortality before and after the national 988 launch by states' pre-988 baseline capacity gaps. The left-hand columns report incidence rate ratios from the PPML model, and the right-hand columns report marginal effects from OLS estimates. In all specifications, the baseline capacity gap ranges from 0, when all calls were answered, to 1, when no calls were answered before the national launch. Table 1 reports a mean baseline answer rate of 0.739, with a standard deviation of 0.137. If states with the lowest initial answer rates (and therefore the largest gaps to close) experienced the largest reductions in youth or all-age suicide, the post-launch PPML incidence rate ratios would be less than one. Before the launch, the incidence rate ratios scatter around one. After the launch, three estimates are statistically significant, but they indicate higher rather than lower suicide rates in the states most likely to be affected by 988. These estimates imply increases of roughly 3 to 5 percent at the launch and again 10 months later. They represent only three of the 12 estimates of differential

suicide trends by state baseline gap, however, and the OLS estimates are uniformly statistically insignificant. Neither specification shows a discontinuity at launch or a sustained post-launch pattern, consistent with Hansen et al. (2026), who use a similar specification in their primary analyses.

C. Discrete-Dose Difference-in-Differences

The results in Table 2 leverage differences by baseline gaps in suicide hotline answer rates, not actual improvement in system performance. These reduced form estimates could reflect an absence of a treatment effect of 988 on suicide mortality or be driven by a lack of differential improvement (i.e. a null intent-to-treat effect). In Table 3 I report the discrete-dose TWFE specification, interacting each state's actual change in answer rates with a post-launch indicator. These interaction terms compare each group of states by quartile of improvement in answer rates. For youth, two of the four quartile coefficients are positive and statistically significant at standard levels. The pattern is qualitatively similar for the all-age sample, though smaller in magnitude and significant only for Q3.

If changes in the quality of states' suicide hotline response rates were reducing suicide, we should expect to see a monotonic negative relationship with dose. It is unlikely that better suicide hotline service increased suicide mortality. Instead, the positive and uneven change in suicide rates for youth as 988 systems are improving is more likely due to unobserved heterogeneity between states by baseline suicide hotline performance, or endogeneity of the response to the national 988 driven by different underlying changes in other factors affecting the well-being of youth. I return to this pattern, and its relationship to the Callaway-Sant'Anna estimates below, in the discussion.

D. Dose-Specific ATT(d)

In Table 4 I report the Callaway-Sant'Anna *ATT(d)* estimates for the same quartile-vs-high-baseline comparisons. In contrast to the discrete TWFE results in Table 3, none of the eight estimates are statistically distinguishable from zero, and the point estimates are both smaller in magnitude and mixed in sign. There are different ways to interpret the divergence between the positive, non-monotonic pattern in Table 3 and the null effects in Table 4 for the identical underlying comparison. At the very least, they call into question a naïve read of the discrete TWFE specification as evidence that 988 raised suicide risk. The Callaway-Sant'Anna estimates in Table 4 provide a more credible basis for estimating effects of a continuous treatment varying over time differently by state. Nonetheless, the *ATT(d)* estimate rests on the assumption that conditional on state and month fixed effects, states with larger realized changes in answer rates were not on different suicide trajectories before 988.

E. IV-DiD

In Table 5 I report the IV-DiD estimates using state fee and appropriation legislation as an instrument for the in-state answer rate. In columns 1 and 2 I report the reduced form estimates of an indicator of whether the state had implemented legislation to support 988 services, and a measure of the amount of any monthly fee charged per phone line as a dedicated 988 revenue source. In both instances, the reduced form coefficients are in the expected direction, with any state action and the magnitude of monthly fees negatively associated with suicide rates, although none are statistically significant.

In columns 3 and 4 I report the first stage for both samples. The coefficients on the instruments are smaller than their standard errors. The corresponding 2SLS estimates are subsequently imprecise. The F-statistics on the instruments are low. While the Hansen-J test

allays some concern about their validity, the instruments are weak and the 2SLS estimates of the impact of answer rates on suicide are not distinguishable from zero for youth or all ages. Nonetheless, like the *ATT(d)*, the sign of the 2SLS coefficients further weakens any read of the TWFE model as suggesting 988 *increased* suicide mortality.

VI. Discussion

The question of whether the launch of 988 has saved lives is a policy question of first-order importance. The earliest work on this question is a research letter in JAMA (Patel, et al. 2026). They estimate that observed suicides after the launch of 988 were 11 percent lower than what would have been expected based on a seasonal autoregressive integrated moving average models of rates from 1999 to 2022. Patel et al. provide no formal test of differences in outcomes within states affected differently by the rollout of better suicide hotline service quality, and no explicit estimate of a treatment effect.

In this paper, I estimate treatment effects using within-state variation in numerous ways. One way to assess the impact of 988 is an intent-to-treat style estimate of changes in suicide after launch as population-weighted average state-level changes from pre-988 levels, regardless of how state 988 systems improved. Alternatively, the 988 effect on suicide can leverage observed improvement in suicide hotline services to identify a treatment effect from the dose-response relationship with suicide. While this latter strategy is intuitively appealing, the possibility that state-level patterns in service quality following 988 adoption are endogenous complicates identifying counterfactual changes by state. I have estimated treatment effects in many ways, from simple pre-post differences in differences, *ATT(d)* estimates by observed treatment response, dynamic event studies, and two distinct identification strategies (realized answer-rate change and estimates by baseline capacity gap). Taken together, across different models,

populations and identification strategies, the findings indicate that there has been no appreciable effect of 988 on suicide mortality.

One possibility for the null effect is that the analysis is focused on the wrong outcome. Callers to suicide hotlines may benefit in ways that do not show up in reduced mortality. Rather, 988 might limit non-fatal self-harm among those in despair. To assess if the launch of 988 reduced injuries caused by self-harm, I analyze monthly data on visits to emergency departments, by cause for select states from the SEDD from 2021 to 2024. Using the limited panel, I estimate separate models of changes in 988 system performance, the rate of visits (per 10,000) to emergency departments where the patient was treated for injuries derived from self-harm deemed to be a suicide attempt, and the number of total emergency department visits per month. For each outcome, I estimate difference-in-differences models interacting a post-988 indicator with a state dummy variable, including year and month fixed effects.

In Table 6 I summarize these analyses. The regression coefficients in each column measure the average change in the relevant outcome before versus after the national launch of 988 in July 2022, net of state, year, and month fixed effects. Each column also reports the state's outcome mean in the period before 988 launch. In the first column, there is little evidence of changes in total calls per 1000 population in Michigan or North Carolina after the July of 2022.[5] Rhode Island saw a small increase in calls.

Where changes in 988 service are apparent is in the substantial increase in answer rates in Michigan after July 2022. On average, answer rates increased by 20 percent in Michigan subsequent to the national launch (from a low baseline mean of 63.3 percent before 988). Rhode

[5] Since the age of the caller is not recorded, the higher rates of calls per 1,000 is not necessarily due to higher rates of calling by 15–24-year-olds. Rather, this arises because the denominator is the state population in panel A, and the 15–24-year-old population in panel B.

Island had an exceptionally high answer rate before 988, that declined slightly but remained very high. North Carolina had a baseline answer rate of 84.1 percent, which was unchanged by 988.

These differences are relevant for interpreting the changes in visits for self-harm by state, reported in column 3. Overall, the rate of visits for self-harm declined by 3.3 per month in Michigan. This was a decline of almost 10 percent of the mean of 36 per 10,000 before 988. This decline was even larger for 15-24 year olds: The rate of visits for injuries from self-harm among 15-24 year olds declined by an average of 11.5 per 10,000 visits after 988 launched, from a baseline mean of about 112. Since Michigan saw a substantial improvement in suicide hotline service quality, this could appear to be evidence consistent with a decline in suicide-related injuries after the rollout of 988. But reading the evidence this way raises the question of why suicide-related injuries declined even more in North Carolina, where call volume and answer rates were essentially unchanged after 988. The final column of Table 6 reports changes in visits to emergency departments for any reason (per 1,000 population). There is no change in the rate of total emergency room visits in any of the three states, for either age group.

Another interpretation of the null effect is that while the launch of 988 did result in more answered calls at suicide hotlines, much of the added call volume is from people in distress rather than at imminent risk of death or self-harm. About 3 percent of answered calls in 988 centers were classified as imminent risk, and about 2 percent led to emergency dispatch (Drapeau 2024). It is possible that the people who are in the process of attempting suicide and at imminent risk for death or injury are less likely to call. Suicide attempts can be impulsive and carried out quickly. Simon et al. (2001) reported that 24 percent of survivors of near-lethal attempts acted within five minutes of deciding to take their own lives, and nearly half of those

hospitalized after a suicide attempt reported that it was ten minutes or less between first thinking about a suicide attempt and then acting on it (Deisenhammer et al. 2009).

The reach of 988 also appears to be low, relative to need. Purtle (2025) estimates that only about 3 percent of people with a history of suicidality had used 988, and many adults report a fear of police involvement or forced hospitalization if they call 988.[6] None of this is to suggest that callers are low-risk. Among veterans, Veterans Health Administration patients who called the Veterans Crisis Line died by suicide at 298 per 100,000 over the following year (Hannemann et al. 2020), and more than half of suicidal Lifeline callers had prior attempts (Gould et al. 2007). Clearly, serious risk does reach suicide hotlines, but far more does not. And, we know little about how calling changed outcomes, and good evidence on mortality does not exist (Hoffberg et al. 2019).

One way to gain insights into the potential mechanisms through which 988 could have affected self-harm and suicide mortality in the population is to assess how estimates of suicidal behavior changed over the rollout period. In Table 7, I report estimates of the total number of Americans reporting seriously considering suicide in 2021 (the year before 988) and 2024 (the most recent year available). These estimates come from the National Survey of Drug Use and Health (NSDUH), a survey conducted annually to provide estimates of various health conditions and outcomes for the U.S. civilian, non-institutionalized population. The NSDUH estimates that 13.2 million adults reported seriously thinking about suicide at some point in the 12 months preceding the 2021 survey. That number was 13.9 million in 2024.

If 988 had protective effects at the population level, we would expect lower relative rates of those with suicidal thoughts to develop a suicide plan, or actively attempt suicide, or die by

[6] https://mindsitenews.org/2023/05/26/few-people-know-about-the-988-lifeline-and-many-who-do-fear-calling-it-will-lead-to-a-police-response/

suicide. In Table 7 I also report nationally representative estimates of the percent of those seriously considering suicide who made a suicide plan (row 2) and made an attempt (row 3). Between 2021 and 2024, the likelihood of having made a plan or attempt both *increased* among those considering suicide. Further, there was no change in the ratio of suicide deaths to the population considering suicide (row 4). While the launch of 988 in July 2022 did improve access to suicide prevention hotline services, at the population level there was no apparent diminution in the pathway from suicidal thinking to suicidal planning and attempts.

Despite the promising finding reported in Patel et al. (2026), the launch of 988 does not seem to have had any effect on death by suicide. While this is a sobering reminder of the difficulty in reducing suicide mortality, it is not a judgment on the value of the 988 suicide hotline. A service that reduces mental despair, or non-fatal suicidality can be effective on its own even without averting many deaths. Future studies that focus on potential impacts of access and calls to suicide hotlines on well-being, access to mental health treatment, and non-fatal injuries can provide a fuller accounting of any wider benefits to those in crisis and society more broadly.

References

Abadie, Alberto, Joshua Angrist, Brigham Frandsen, and Jörn-Steffen Pischke. 2026. "Harvesting Differences-in-Differences and Event-Study Evidence." NBER Working Paper 34550.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "Difference-in-Differences with a Continuous Treatment." NBER Working Paper 32117.

de Chaisemartin, Clément, Xavier D'Haultfœuille, Félix Pasquier, Doulo Sow, and Gonzalo Vazquez-Bare. 2024. "Difference-in-Differences for Continuous Treatments and Instruments with Stayers." Working paper.

Deisenhammer, Eberhard A., Chy-Meng Ing, Robert Strauss, Georg Kemmler, Hartmann Hinterhuber, and Elisabeth M. Weiss. 2009. "The Duration of the Suicidal Process: How Much Time Is Left for Intervention Between Consideration and Accomplishment of a Suicide Attempt?" *Journal of Clinical Psychiatry* 70 (1): 19–24.

Drapeau, Christopher W. 2024. What happens when people are actively suicidal? An in-depth analysis of 988 Suicide & Crisis Lifeline imminent risk data. Vibrant Emotional Health. https://www.vibrant.org/wp-content/uploads/2024/07/Imminent-Risk-Data-White-Paper-Final-Version.pdf

Gould, Madelyn S., and John Kalafat. 2009. "Crisis Hotlines." In Oxford Textbook of Suicidology and Suicide Prevention: A Global Perspective, edited by Danuta Wasserman and Camilla Wasserman, 459–462. New York: Oxford University Press.

Gould, Madelyn S., John Kalafat, Jimmie Lou Harris-Munfakh, and Marjorie Kleinman. 2007. "An Evaluation of Crisis Hotline Outcomes. Part 2: Suicidal Callers." Suicide and Life-Threatening Behavior 37 (3): 338–352.

Hannemann, Claire M., Ira R. Katz, Megan E. McCarthy, Gregory J. Hughes, Richard McKeon and John F. McCarthy. 2020. "Suicide mortality and related behavior following calls to the Veterans Crisis Line by Veterans Health Administration patients" *Suicide and Life-Threatening Behavior*. https://doi.org/10.1111/sltb.12722

Hansen, Benjamin, Niketana Kannan, and Jonathan Zhang. 2026. "The Impact of Suicide Hotlines on Health: Evidence from 988 Implementation." NBER Working Paper 35601.

Hoffberg Adam S., Kelly Stearns-Yoder Kelly and Lisa A. Brenner. 2019. "The Effectiveness of Crisis Line Services: A Systematic Review," *Frontiers in Public Health.* v. 7 10.3389/fpubh.2019.00399

Marcotte, Dave E., and Benjamin Hansen. 2024. "The Re-Emerging Suicide Crisis in the U.S.: Patterns, Causes and Solutions." *Journal of Policy Analysis and Management* 43 (2): 582–612.

Patel, Vishal R., Michael Liu, and Anupam B. Jena. 2026. "Suicide Mortality Among Adolescents and Young Adults After Launch of a Suicide and Crisis Lifeline." *JAMA*. https://doi.org/10.1001/jama.2026.5157.

Purtle J, Lindsey M. 2025. "The 988 Suicide and Crisis Lifeline in the US: status of evidence on implementation," *World Psychiatry*. Feb;24(1):135-136. doi: 10.1002/wps.21285. PMID: 39810681; PMCID: PMC11733462.

Simon, Thomas R., Alan C. Swann, Kenneth E. Powell, Lloyd B. Potter, Marcie-jo Kresnow, and Patrick W. O'Carroll. 2001. "Characteristics of Impulsive Suicide Attempts and Attempters." *Suicide and Life-Threatening Behavior* 32 (1 Suppl.): 49–59.

Results

Table 1 Summary Statistics, Full Sample

	Mean	SD	N
Youth (15–24) suicide rate (per 100,000)	14.04	6.898	4,218
All-age suicide rate (per 100,000)	17.3	5.342	4,284
Population, ages 15–24 (thousands)	855	968	4,284
Population, all ages (thousands)	4,281	5,421	4,897
In-state answer rate	0.824	0.115	2,652
Calls per 1,000 residents	5.757	3.143	2,142
Unemployment rate (%)	3.652	0.910	2,703
Baseline answer rate	0.739	0.137	2,652
Realized gain in answer-rate	0.120	0.138	2,652
State fee amount ($/line)	0.012	0.067	4,795
Any state 988 policy adopted	0.070	0.255	4,897
Observations (state-months)			**4,897**

Notes: Unit of analysis is state/month. Suicide rates are annualized. In-state answer rate, and calls per 1,000 are only observed starting in mid-2021, so N is below the full 4,897 state-months.

Table 2 Dynamic Estimates of Suicide Mortality by Pre-988 Capacity

	PPML		OLS	
Period	**Youth**	**All-age**	**Youth**	**All-age**
−12 mo	1.133	1.040	−12.42	−0.141
	(0.206)	*(0.143)*	*(10.79)*	*(0.475)*
−6 mo	1.105	1.023	−17.71	−0.164
	(0.251)	*(0.156)*	*(15.31)*	*(0.382)*
−3 mo	0.794	1.118	−14.68	−0.005
	(0.255)	*(0.176)*	*(16.61)*	*(0.484)*
0 mo	1.015	1.262**	8.677	−0.055
	(0.356)	*(0.146)*	*(7.117)*	*(0.567)*
+6 mo	0.724	1.341	−20.54	−0.233
	(0.187)	*(0.240)*	*(14.10)*	*(0.616)*
+10 mo	1.633**	1.341**	13.95	0.494
	(0.381)	*(0.193)*	*(9.457)*	*(0.388)*
+12 mo	0.932	1.058	−4.243	−0.102
	(0.307)	*(0.172)*	*(11.39)*	*(0.435)*
+18 mo	1.076	1.097	−3.014	0.029
	(0.265)	*(0.142)*	*(12.45)*	*(0.379)*
+24 mo	0.788	0.904	−13.80	−0.244
	(0.229)	*(0.072)*	*(12.30)*	*(0.338)*

*Notes: Dose = 1 − ar_pre (pre-988 baseline answer-rate shortfall) interacted with event time, mirroring Hansen, Kannan & Zhang (2026)'s identification strategy rather than this paper's own preferred realized-change dose. Standard errors in parentheses, clustered by state. *** p<0.01, ** p<0.05, * p<0.10. Consistent with Hansen et al.'s own reported result: scattered significance, no clean break at t=0, no sustained post-period pattern.*

Table 3 **TWFE Estimates of 988 Dose-Quartile × Post-Launch Interactions**

	Youth (15–24)	All-age
Q1 (smallest gain) × Post988	4.755***	0.631
	(1.736)	*(0.583)*
Q2 × Post988	3.104*	0.884
	(1.590)	*(0.542)*
Q3 × Post988	5.512***	1.322***
	(1.631)	*(0.441)*
Q4 (largest gain) × Post988	3.521	0.491
	(1.912)	*(0.614)*
Calls per 1,000	−0.092	−0.075***
	(0.098)	*(0.025)*
Outcome	Rate	Rate
State FE	Yes	Yes
Month & Year FE	Yes	Yes
Log population control	Yes	Yes
Unemployment control	Yes	Yes
Observations	2,104	2,142

*Notes: Dependent variable is the suicide rate per 100,000, annualized. Q1–Q4 are quartiles of each state's realized answer-rate. Standard errors clustered by state in parentheses. ** p<0.05, *** p<0.01.*

Table 4 Callaway-Sant'Anna Dose-Specific ATT(d) of 988 Volume and Answer Rate on Suicide

	Youth	All-age
Q1 (smallest gain)	−4.366	−0.136
	(4.078)	(0.193)
Q2	0.060	0.076
	(1.772)	(0.109)
Q3	0.588	0.111
	(3.753)	(0.159)
Q4 (largest gain)	−1.375	−0.048
	(2.420)	(0.150)

Notes: Aggregate ATT(d) by dose quartile (Q1 = smallest realized answer-rate gain, Q4 = largest). ATT(d) on suicide rate per 100,000, one quartile vs. high-baseline (grp_high==1) states at a time. Standard errors in parentheses.

Table 5 IV-DiD Estimates of 988 Volume and Answer Rate on Suicide: Youth and All-Ages

Fee/appropriation legislation as instrument for the in-state answer rate.

	Reduced form		First stage		2SLS	
	Youth	**All-age**	**Youth**	**All-age**	**Youth**	**All-age**
Fee amount	−1.673	−1.592	-0.106	-0.098		
	(2.916)	*(0.98)*	*(0.072)*	*(0.067)*		
Any 988 law?	−1.79	−0.814	0.040	0.041		
	(1.859)	*(0.896)*	*(0.032)*	*(0.033)*		
Calls per 1,000			−0.001	−0.001	−0.087	−0.073**
			(0.003)	*(0.002)*	*(0.111)*	*(0.028)*
Answer rate					−14.52	−5.21
					(33.22)	*(21.74)*
Cragg-Donald F					9.791	10.087
(Stock-Yogo weak ID crit. Value (10% max. LIML)					*8.68*	*8.68*
Hansen's J					0.606 *(p=0.436)*	1.99 *(p=0.158)*
Observations	2,104	2,142	2,104	2,142	2,104	2,142

Notes: Table reports reduced form, first stage and 2SLS estimates of 988 service quality on suicide rates for youth (15-24 year olds) and all ages. All models control for population, unemployment rates, state, year, and month fixed effects. Standard errors, clustered by state, in parentheses.

Table 6 988 and Self-Harm Injuries in Emergency Rooms

	Calls per 1,000 Pop.	Answer Rate	Self-Harm Rate	Total ED Visits per 1,000 Pop.
	(1)	(2)	(3)	(4)
		Panel A: All Ages		
Michigan	−0.031	0.206***	−3.307***	1.292
	(0.032)	(0.033)	(1.203)	(0.761)
Outcome mean (pre-988)	*0.710*	*0.633*	*36.145*	*38.242*
North Carolina	0.012	−0.034	−4.035***	1.561
	(0.030)	(0.041)	(1.362)	(1.154)
Outcome mean (pre-988)	*0.591*	*0.841*	*27.423*	*39.641*
Rhode Island	0.110***	−0.038**	−0.648	0.731
	(0.034)	(0.018)	(1.203)	(0.977)
Outcome mean (pre-988)	*0.418*	*0.982*	*22.731*	*36.922*
		Panel B: Ages 15–24		
Michigan	−0.053	0.208***	−11.498**	0.033
	(0.167)	(0.032)	(4.551)	(0.556)
Outcome mean (pre-988)	*4.095*	*0.633*	*111.633*	*29.448*
North Carolina	0.108	−0.034	−13.664**	0.057
	(0.155)	(0.040)	(5.580)	(0.938)
Outcome mean (pre-988)	*3.328*	*0.841*	*78.382*	*30.095*
Rhode Island	0.427**	−0.041**	5.339	−0.859
	(0.187)	(0.019)	(5.424)	(0.760)
Outcome mean (pre-988))	*2.277*	*0.982*	*66.362*	*26.032*
Month FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Population control	Yes	Yes	Yes	Yes
Observations (per panel)	126	126	126	126

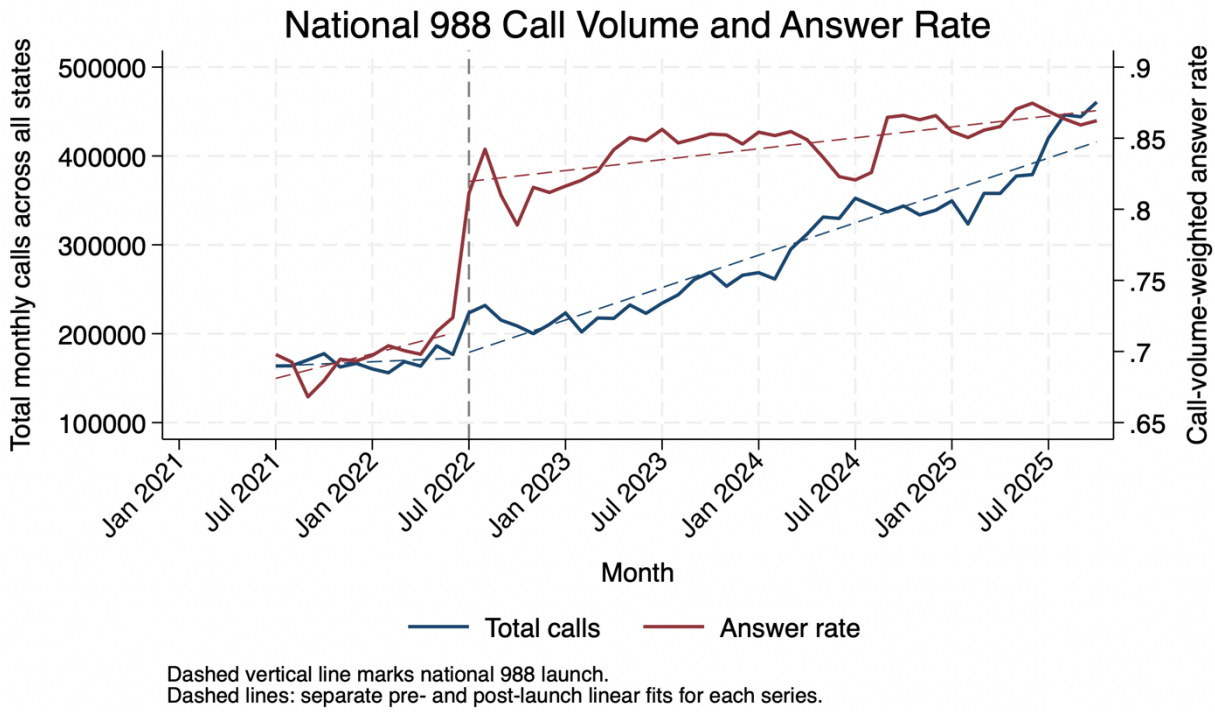
Notes: Each state's primary entry is a directly estimated regression coefficient: the model (column outcome on state fixed effects, state × post988 interactions, month and year fixed effects, and a population control) was run three times per panel, alternately changing reference state, so every state's post-988 change is a single coefficient with its own heteroskedasticity-robust SE rather than a derived linear combination. Standard errors are in parentheses beneath each estimate. *Outcome mean (pre-988)* rows report the unadjusted pre-period mean of the outcome for that state. **p<0.05, ***p<0.01.

Table 7 NSDUH Estimates of Suicidal Thoughts/Plans/Attempts and CDC Mortality Data: 2021 & 2024

	2021	2024
	Population Estimate	
Suicidal Thinking: "*At any time during the past 12 months, did you think seriously about trying to kill yourself?*"	13.2 million	13.9 million
	Plans/Attempts/Deaths as % of those with Suicidal Thinking	
Suicide Plans: *"During the past 12 months, did you make any plans to kill yourself?"*	28.80%	35.30%
Suicide Attempts: *"During the past 12 months, did you try to kill yourself?"*	12.10%	16.50%
Suicide Deaths (Total count)	0.33%	0.34%

First row presents estimates of number of persons 12-year and older reporting thinking about suicide in 2021 and 2024. Row 2 reports the percent of row 1 who made suicide plans, and row 3 the percent who reported making a suicide attempt. The final row is the total number of suicides from CDC mortality data, divided by the estimate in row 1, reported as a percent.

Figure 1



Dashed vertical line marks national 988 launch.
Dashed lines: separate pre- and post-launch linear fits for each series.

Figure 2

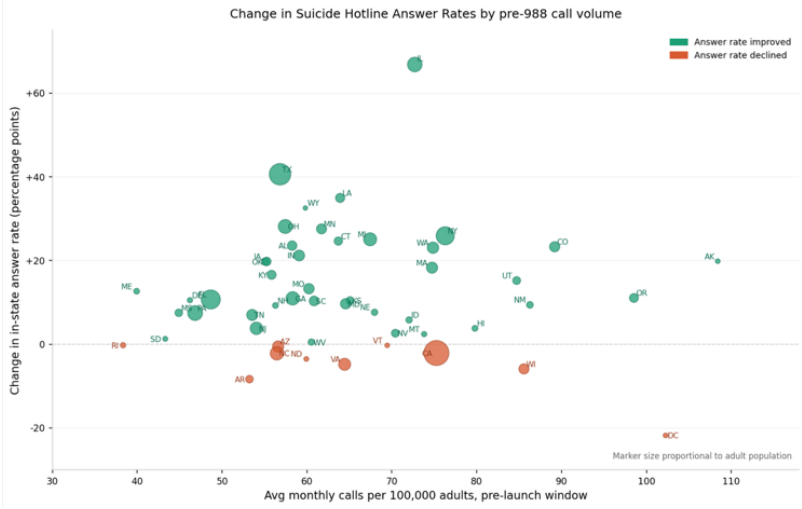
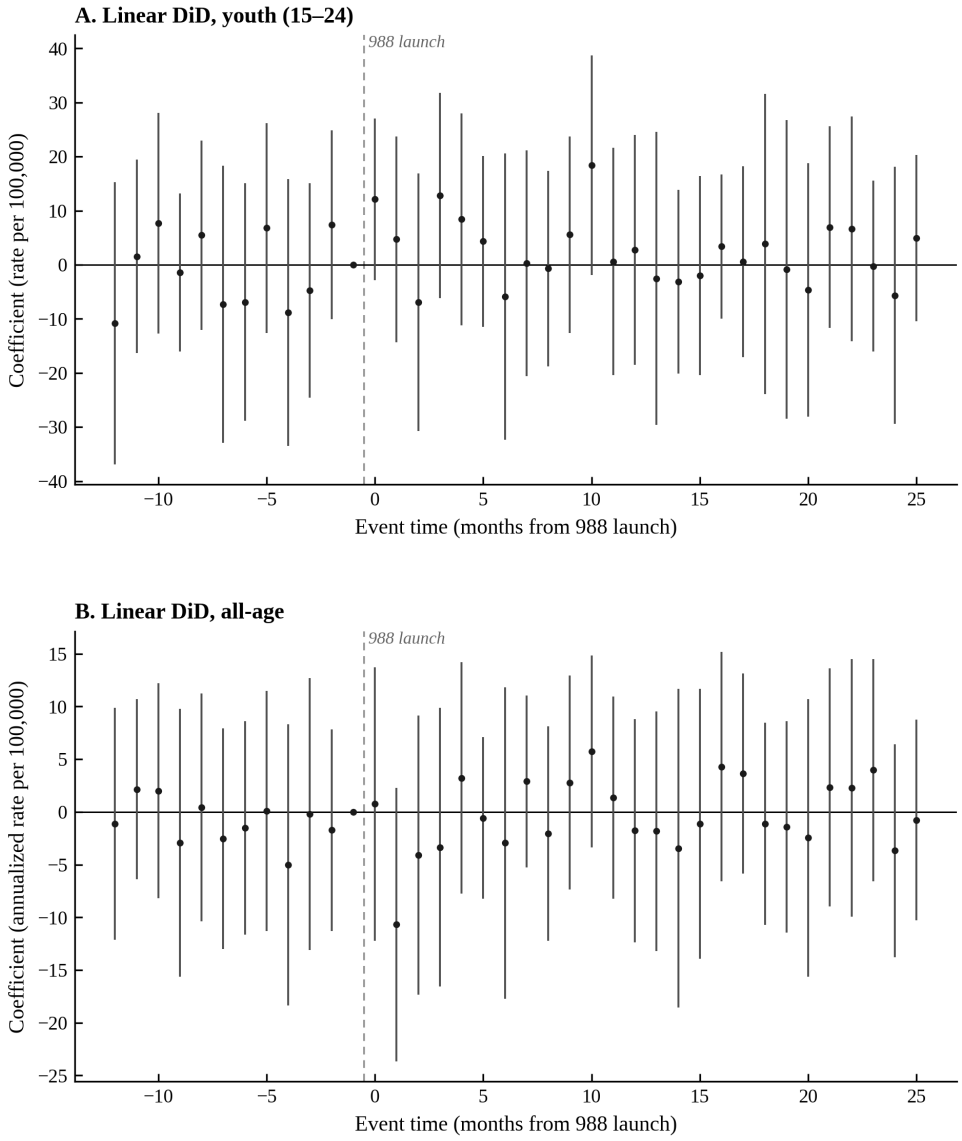
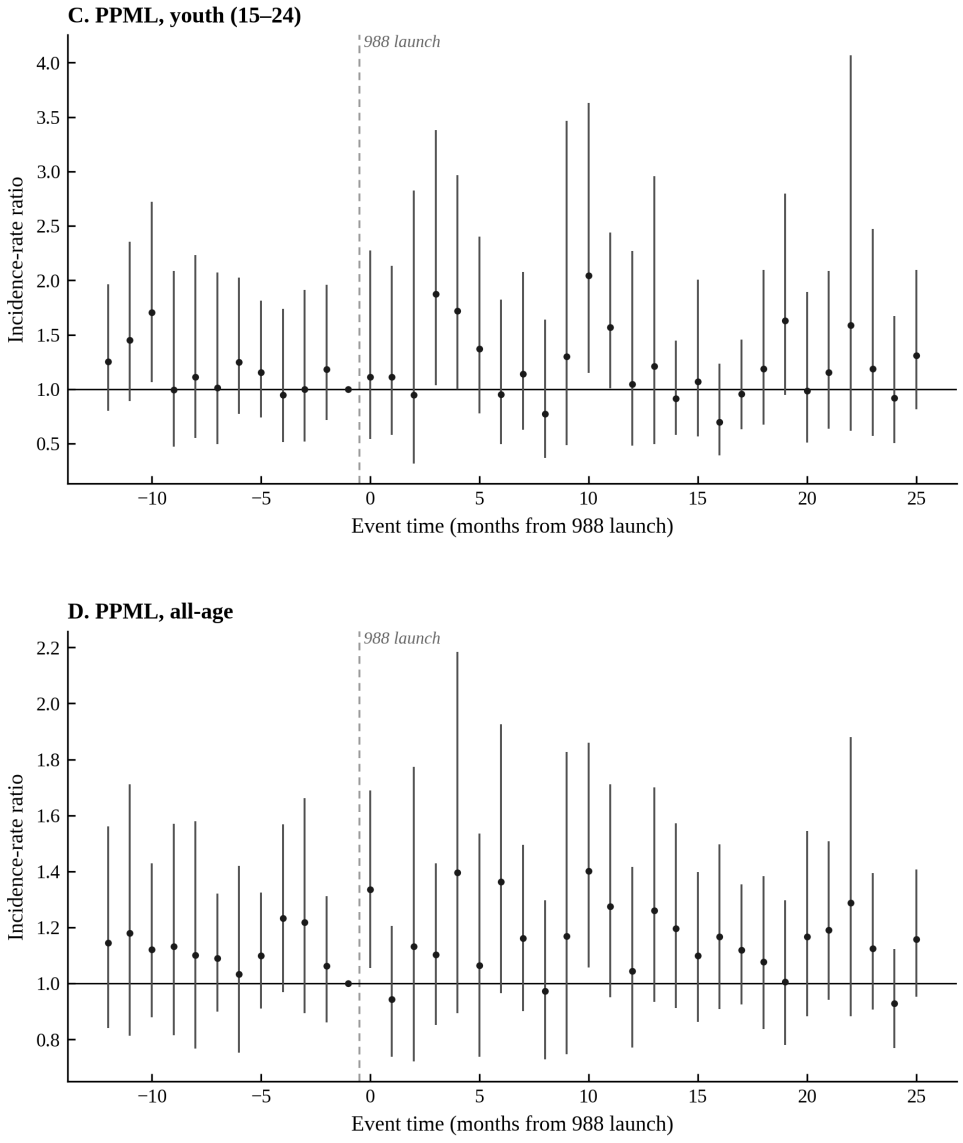


Figure 3. Event Study of Realized Change in Answer Rate × Event Time





Notes: Point estimates with 95% confidence intervals (±1.96 SE), reference period (t=−1) omitted by construction. Vertical line marks national 988 launch (July 2022).

Appendix Table A1 Dynamic Fuzzy DiD-IV, Youth (15–24) Sample

Event time	Estimate	SE	95% CI	N	Switchers
1 month post-adoption	-1.286	2.068	[-5.34, 2.77]	333	11
2 months post-adoption	-4.849	3.411	[-11.53, 1.84]	329	10
3 months post-adoption	-4.108	2.847	[-9.69, 1.47]	330	11
4 months post-adoption	-1.508	3.256	[-7.89, 4.87]	293	10
5 months post-adoption	-6.949	5.405	[-17.54, 3.65]	293	10
6 months post-adoption	-5.011	3.401	[-11.68, 1.65]	289	10
7 months post-adoption	-2.163	4.234	[-10.46, 6.14]	285	10
8 months post-adoption	-2.010	3.355	[-8.59, 4.56]	280	10
9 months post-adoption	-1.795	4.797	[-11.20, 7.61]	280	10
10 months post-adoption	-1.729	4.234	[-10.03, 6.57]	281	10
11 months post-adoption	-1.939	2.742	[-7.31, 3.43]	280	10
12 months post-adoption	-4.699	4.277	[-13.08, 3.68]	275	10

Joint nullity of post-adoption effects: $p < 0.01$

Event time	Estimate	SE	95% CI	N	Switchers
1 month pre-adoption (placebo)	-6.654	4.528	[-15.53, 2.22]	329	10
2 months pre-adoption (placebo)	-1.237	2.550	[-6.23, 3.76]	326	10
3 months pre-adoption (placebo)	-3.102	4.554	[-12.03, 5.82]	328	11
4 months pre-adoption (placebo)	-4.413	4.723	[-13.67, 4.84]	290	10
5 months pre-adoption (placebo)	0.298	3.349	[-6.27, 6.86]	288	10
6 months pre-adoption (placebo)	-4.037	4.647	[-13.15, 5.07]	238	9

Notes: Joint nullity of placebos: $p = .154$. Average cumulative effect per treatment unit: −3.163 (SE 3.292), 95% CI [−9.615, 3.290], N=1587, 122 switcher-periods, averaged over 6.35 periods.