Kim Soffen
PPPA 8022
Final Paper

Replication of "Health Consequences of the US Deferred Action for Childhood Arrivals (DACA) immigration programme: a quasi-experimental study"

I used data posted by the author: no
My substantive addition is: N/A

Introduction

In their paper, "Health Consequences of the US Deferred Action for Childhood Arrivals (DACA) immigration programme: a quasi-experimental study," Atheendar Venkataramani et al examined whether the implementation of DACA in the U.S. improved the overall or mental health of eligible people. By way of background, DACA, implemented in 2012, allowed undocumented immigrants brought to the U.S. as children to be temporarily protected from deportation and obtain a work permit (USCIS, 2021). Being undocumented has been found to correlate with poor health outcomes, theorized to be because of mental stress and lack of access to health care and insurance (Martinez, 2016). There's a few papers with evidence of a causal relationship, but these are few and incur methodological limitations (Ibid). The Venkataramani paper strived to add to this causal link by asking whether DACA, a reprieve from the near-constant threat of deportation or job loss, could yield health benefits.

Examining this with a pure OLS approach – just measuring how DACA eligible people's health changed over time – would run into an endogeneity problem. There are many other reasons DACA eligible people's health could change over the course of the study from 2008 to 2015. For instance, this time period covers the Great Recession and the slow recovery afterwards; it's plausible, even likely, that these massive macroeconomic shifts affected the physical and mental health of undocumented people, and that this rather than DACA prompted a change in health outcomes. Another major confounder could be the implementation of the ACA's provisions, much of which occurred between 2010 and 2014. These changes improved the quality of health insurance and allowed more people (though primarily those with legal presence) to obtain it; this could also drive changes in health for the DACA-eligible population. As a final example, the culture around admitting mental health problems and seeking help for them changed over the study's time period. It's plausible this shift could make people more willing to admit mental health issues in a survey, or alternatively less likely to experience these issues because they have sought help. This confounder could also drive the change in a traditional OLS analysis.

To address this endogeneity problem, Venkataramani et al performed a difference-in-differences analysis to compare the health of DACA eligible people and ineligible people before and after the policy's implementation. To be eligible for DACA, an immigrant had to:

1. Be under 31 years old at the date of implementation (June 15, 2012).
2. Immigrate when under 16 years old.
3. Have continuously resided in the U.S. since June 15, 2007 (five years before implementation).
4. Be undocumented (either through unlawful entry or visa overstay) since the date of implementation (June 15, 2012).
5. Be physically present in the U.S. on the implementation date and at the time of their application for DACA status.
6. Be a student, have a high school diploma or GED, or have been honorably discharged from the military.

7. Have no felony or "significant misdemeanor" conviction, and no more than 3 misdemeanor convictions.
8. Not pose a threat to national security or public safety (USCIS, 2021).

The paper's 'ineligible' control group consisted of immigrants in the same general age range, ethnicity, and education level as the eligible group, but who are not eligible for DACA because they immigrated too late in life or are slightly over the age cutoff at the time of implementation (ie: fail criteria #1 or #2 above). The study looked at four health metrics as outcome variables: self-reported overall health on a 1-5 scale, the proportion of people who rate themselves 'fair' or 'poor' on that scale, score on the K6 test which measures mental distress from 0 to 24, and the proportion of people who do poorly on that test (score of 5 or higher).

The authors found that DACA-eligible people had a statistically significant improvement on both mental health metrics, but no improvement on their overall health metrics, after the program's implementation. The authors then showed this finding is robust to different specifications of the treatment and control groups. In this paper, I attempt to replicate this finding, but am ultimately unable to find any statistically significant results.

Data

The original authors and I both obtained data from the National Health Interview Survey (NHIS), which is run annually by the CDC to get an understanding of Americans' health problems and related behaviors. The NHIS includes both a shorter version of the survey, given to a large sample of people, and a longer version, given to a smaller sample (NHIS, n.d.). The original study and my replication utilize both components of the NHIS over eight years, 2008 to 2015, using 16 NHIS data files in total.

One of the major weaknesses of the dataset in the context of this study is that there is no way to use the NHIS data to perfectly determine who is DACA eligible (or 'nearly eligible', as in the control group). In response, the authors used several variables to create what they contend is a close proxy. From the entire dataset, they restricted their sample to include individuals who are 19-50 years old, Hispanic (since the vast majority of DACA recipients are Hispanic), non-citizens, have at least a high school education (as a proxy for DACA criteria #6 listed above), and have lived in the U.S. for at least 5 years (as they contend newer immigrants are not a good control group, and to approximate DACA criteria #3). From there, they divided this group into DACA eligible and ineligible based on the individual's age at DACA implementation per DACA criteria #1 (which can be precisely determined using their current age and survey year), and their approximate age at immigration, per DACA criteria #2.

One of the biggest issues here is that the NHIS provides no way to know whether an immigrant is documented or undocumented. The study narrowed survey subjects to Hispanic immigrants, but of this group, nationwide, only about one-third are undocumented (Gamboa, 2021). Since undocumented immigrants have reason to avoid making themselves known to the government, it's reasonable to believe the proportion of undocumented immigrants in the NHIS is even lower. That means the *bulk* of the study's sample is made up of documented immigrants whom DACA would not directly benefit. This could mean the effect found by the study is understated, as the documented people in the treatment group mute the effect seen by the undocumented people. But if there are meaningfully more documented people in the control group than the treatment group, it could undermine the comparability of the two groups and make the study's results overall difficult to interpret.

Another data issue is that the NHIS does not ask precisely when someone immigrated, but rather asks them to group themselves into buckets of immigrating within the past year, 1-5 years ago, 5-10 years ago, 10-15 years ago, or 15+ years ago. This means we can only get an approximate age of when someone immigrated; for instance, if someone is 30 at the time of the survey and fell into the 10-15 bucket, they could be as young as 15 years old at immigration, meeting DACA criteria, or as old as 20, which would exclude them from DACA. The authors said that they used respondents' answer to this question to get an "approximate" age of immigration that is "subject to classic measurement error," but they do not specify exactly how they come to this approximation. For my approximation, I determine age at immigration by assuming each individuals' actual years since immigration fell in the middle of the bucket they selected; for instance, I assumed someone who fell in the "at least 5 but less than 10" years bucket has immigrated 7 years ago. As will be detailed later, I also tested other specifications for robustness.

A final issue, affecting only my replication, is that the authors were able to access the survey participant's exact birth date, whereas the public file I used included only month and year. DACA eligibility criteria #1 requires knowing whether someone's birthday comes before or after June 15, 1981. With only month and date information, this precision is not possible. I assume all people born in June are born after that cutoff date (ie: eligible for DACA). In my sample, 32 people are born in June 1981. Assuming uniform distribution of birthdays, this means 16 people are miscategorized out of a total sample size of 11,067 (0.1%).

The authors' descriptive statistics of their final sample can be seen in Table 1 in the Appendix. I was unable to replicate these sample characteristics. In Figure 1 in the Appendix, the authors show their sample size after applying each sample restriction criterium. I was able to match their sample for the first two steps, limiting the sample age and ethnicity, but on the later steps, it becomes mathematically impossible to match their sample size, regardless of how "don't know," "refused," and other ambiguous responses are categorized. I ultimately followed the authors' written specification, rather than trying to match the numbers as closely as possible. The descriptive statistics for my sample can also be found in Table 1 in the Appendix. My sample has 3,906 (or 26%) fewer people than the original authors'.

I was also unable to replicate, and a bit puzzled by, some of the types of measurements included in the authors' descriptive statistics table. The authors specified that all measures in the table, aside from the number of respondents, used survey weights. However, they listed the number of people who report poor overall or mental health (along with other binary variables), rather than just the proportion. This is impossible to do while using survey weights, since one is using weighted units rather than individuals. It's unclear where the authors got those numbers from, so I just reported proportions in those rows for my descriptive statistics table. Despite these challenges to replication, the descriptive statistics for my sample and the original authors' are relatively similar, except for my sample appearing to have somewhat better mental health.

Methods:

This study used a difference-in-differences model to see if DACA implementation affected the overall or mental health of eligible individuals. Their model was as follows:

$$H_{it} = g(\beta_0 + \beta_1 \times Eligible_i \times DACA_t + \beta_2 \times Eligible_i + \beta_3 \times DACA_t + \beta \times X_{it} + \varepsilon_{it})$$

Here, subscript i represents the individual survey respondent, and subscript t represents the year and month the survey was administered. *Eligible* is a binary variable for whether that individual is eligible for DACA, and *DACA* is a binary variable for whether DACA was implemented at the time of the survey. X_{it} is a vector of covariates including age at DACA implementation, estimated age at immigration, region of residence (Northeast, Midwest, South, or West), gender, and year and month of interview. Because of the limited data availability in the public use file, my replication uses year and quarter of interview instead of year and month. In this equation, β_1 is the difference-in-differences variable of interest.

This equation was used to measure four different health outcome variables H_{it}:

1. Self-reported overall health on a scale of 1-5, 5 being the best.
2. A binary indicator of whether someone scored themselves as having fair or poor health (1 or 2)
3. Score on a K6 test, which measures mental health on a scale of 0 to 24, 0 being the best.
4. A binary indicator of whether someone has moderate or poor mental health (a K6 score of 5 or greater).

Outcome #1 used an ordinary least squares regression, outcomes #2 and #4 were logistic regressions, and outcome #3 was a poisson regression. These varying models are represented by the g in the equation above.

The authors also performed several robustness checks. First, they restricted the sample to interviews conducted between 2010 and 2015 to limit how the Great Recession would vary people's answers over time. Second they kept this 2010-2015 restriction and also restricted the sample to people up to 40 years old, instead of the original specification of up to 50 years old, to weed out any health problems caused by middle age. Finally, they performed a falsification test by restricting the sample to immigrants with less than a high school education. Since the vast majority of this group would not be eligible for DACA because of DACA criteria #6, DACA's implementation should not have affected their health and there should be no significant difference-in-differences between those who met the other DACA criteria and those who did not.

Results:

The original author's main results can be found in Table 2 in the Appendix. They found two statistically significant results: the score on the K6 mental health test, and the likelihood of having moderate or poor mental health. However, when I ran these same models, I found no significant results. The coefficients of interest in each of the regressions can be interpreted as follows. According to the original authors' results, DACA implementation caused an increase in 0.056 points on the 1-5 scale for overall health; (my result was 0.058). The authors also found that DACA implementation decreased the odds of an eligible person having poor or fair health by 2 percent; my result was 16 percent. On the mental health front, the authors found that after DACA implementation, eligible people were 78% as likely as ineligible people to have a 1-point increase in their K6 score; my corresponding result was 91%. (As a reminder, a higher K6 score represents worse mental health). Finally, the authors found that DACA implementation decreased the odds of having moderate or poor mental health by 38% for eligible people as compared to ineligible people; my corresponding figure was 22%. As you can see, the authors found significant results where I did not both because their coefficients had greater magnitudes, but also because they had smaller standard errors, thanks in part to their larger sample size.

None of the covariates specified by the authors were ambiguous, so I believe the difference in our results is driven by the difference in our samples, explained above in the Data section. In an attempt to replicate the authors' results, I tried each of the following:

- Threw out all survey weights and use only the raw data. This yielded some statistically significant results on mental health, but the numbers were nowhere near a match for the original study.
- Changed how Stata weights single unit strata. Did not make a difference.
- Artificially increased the size of the sample by combinations of assuming people who refused to answer the citizenship question were non-citizens and pretending people who dropped out in 12th grade had actually finished high school. Did not make a difference.
- Increased and decreased the threshold for what counted as a "moderate or severe" psychological distress score. Did not make a difference.
- Lowered the minimum age to 18 from 19, in line with what I believe to be a typo in one sentence of the paper, as all other references say 19. Did not make a difference.
- Changed how I calculated age at immigration. If you recall, the years someone has been in the U.S. are in 5-year buckets. In my main specification, I assumed the true value was in the middle of the bucket, ie: someone in the 11-15 bucket had been in the country 13 years. I also checked the effect of assuming the minimum (ie: 11 years in the 11-15 year bucket) and maximum (ie: 15 years in the 11-15 year bucket) values. Using the maximum value slightly lowered mental health p-values to 0.10-0.15, but they remained insignificant.

In their sensitivity analyses, the authors found their results were largely robust to both constraining their sample to 2010-2015, and to that year restriction plus restricting the sample to people under 40. These results can be seen in Table 3 in the Appendix. Both of those alternate models yielded similar coefficients to the main specification. With the exception of the 2nd model's mental health test score outcome, all the results that were significant in the main model remained significant. Additionally, the authors' falsification test, constraining the sample to people without high school diplomas, was successful in that none of the results were statistically significant, making it more compelling that the main model's results were due to DACA eligibility.

Naturally, since my main model did not yield any significant results, none of these sensitivity analyses did either. So in a sense, the lack of relationship I found between DACA eligibility and each outcome variable was robust to these alternate specifications. Similar to the main specification, none of my attempts to produce a significant result (in the bullets above) were successful.

Extension:

Alternative Regression #1: Dropping Ambiguous Cases

As stated above, one of the issues with the data was that people's number of years since immigration was in buckets, rather than a precise number. I ran an additional set of regressions that dropped all the cases made ambiguous by this bucketing. An ambiguous case would be, for instance, someone who immigrated between 14 and 19 years old (ie: DACA eligible and ineligible) depending on where they fall within the bucket. There were 851 ambiguous cases in total, or 7.7% of the total sample. These regressions were otherwise unchanged from their original specifications.

My hope was that this change to the sample would eliminate some mis-categorization between the treatment and control groups, and therefore make the results more significant. In fact, the opposite occurred. Across all four outcome variables, the p value of the coefficient of interest increased, and it became even less likely that DACA implementation caused an improvement in health. The coefficients and p values from these regressions can be found in Table 4 in the Appendix.

Alternative Regression #2: Regression Discontinuity

As an alternative to the author's difference-in-differences design, I created a regression discontinuity design to see if this would yield a significant result. As you'll recall, one of DACA's requirements is that someone must be under age 31 on the date of implementation, which creates a sharp discontinuity for eligibility. My regression was as follows:

$$H_{it} = g(\beta_0 + \beta_1 \times Age\_at\_Impl_i + \beta_2 \times Over\_Age_i + \beta_3 \times Age\_at\_Impl_i \times Over\_Age_i + \beta \times X_{it} + \varepsilon_{it})$$

Similar to the original regression, subscript *i* represents the individual survey respondent, and subscript *t* represents the year and month the survey was administered. *Age_at_Impl* is the running variable, the individual's age at the time of DACA's implementation. *Over_Age* is a binary variable for whether they were 31 or older at the time of implementation (ie: the discontinuity). X_{it} is a vector of covariates, including everything from the difference-in-differences equation that's not collinear with one of these new measures: estimated age at immigration, region of residence (Northeast, Midwest, South, or West), gender, and year and quarter of interview. In this equation, β_2 is the regression discontinuity variable of interest.

In this regression, I also further narrowed the sample from the difference-in-differences estimation. Because this is a regression discontinuity design, I dropped any interviews that were done before DACA's implementation. I also dropped people who were too old at immigration to be DACA eligible. Therefore, the only thing separating my "eligible" and "ineligible" samples was their age at DACA implementation – the running variable.

Narrowing the sample did prevent one of my logistic regressions from converging due to limited variation within a subgroup of the sample, so at the suggestion of the professor, I switched the regression for the binary 'poor overall health' outcome variable from a logistic regression to an OLS that included a polynomial term for the running variable.

The results from these regressions can be seen in Table 5 in the Appendix. There was no effect of the discontinuity on any of the four outcome variables. These results were much further from statistical significance than the difference-in-difference results. Therefore, the regression discontinuity design was unable to provide any support for the hypothesis that DACA's implementation improved the health of eligible people.

Checking Parallel Trends:

In their paper, Venkataramani et al did not give any indication that they checked parallel trends before performing their difference-in-differences analysis. As part of my extension, I check parallel trends for all four outcome variables. Trend graphs can be seen in Figure 2 of the Appendix. To begin with the two 'overall health' variables, the trends are a bit noisy, but look largely parallel. There's not any visual change in the post-implementation period, which aligns with the authors finding no effect of DACA implementation on those variables.

For the two mental health variables, the parallel trends assumption is arguably not met. Before DACA implementation, it appears that both measures are increasing over time for the treatment group but flat for the control group. This alone makes the difference-in-differences analysis invalid, as the control group is not truly comparable to the treatment group. Further, it seems plausible that the 2012 measurement for both mental health outcome variables is an outlier, and that the drop from 2012 to 2013 is the bulk of what's driving the authors' ostensibly significant results. This means that, even if the parallel trends assumption held, the authors' results may be more of a fluke than actual evidence of an effect.

In addition to this visual examination, I ran additional regressions to determine if the pre-implementation slope differed between the treatment and control groups. These regressions were constructed as follows:

$$H_{it} = g(\beta_0 + \beta_1 \times Survey\_Year_t + \beta_2 \times DACA\_Eligible_i + \beta_3 \times Survey\_Year_t \times DACA\_Eligible_t + \beta \times X_{it} + \varepsilon_{it})$$

Here, the *i* subscript represents an individual and the *t* subscript represents a time. *Survey_Year* is the year the survey was taken, and *DACA_Eligible* is a binary variable for whether the person is eligible for DACA. The coefficient of interest is β_3, because this will show if the change in the outcome variable over time varies based on whether someone is DACA eligible or not. If this difference is minimal, that means the trends are parallel. This regression was run over all the data where the survey was taken prior to DACA implementation, so only the pre-trend would be captured. Additionally, it includes all the covariates from the original regression that are not duplicative of the variables of interest: gender, age at immigration, age at DACA implementation, and geographical region.

The results of these regressions can be found in Table 6 in the Appendix. None of the β_3 coefficients achieve statistical significance, but three of them (the two mental health outcomes plus the 'poor or fair health' binary indicator) have p values below 0.35. In other words, for these three, there's a two-in-three chance that there's a real difference in trends between the treatment and control group, not just random chance. While this doesn't rule out the possibility of parallel trends, it does prompt further investigation. Between this result and the graphical evidence of non-parallel trends stated above, I'm highly skeptical of the validity of the difference-in-differences methodology used in this study.

Conclusion:

Venkataramani et al created a difference-in-differences study that found DACA's implementation had a statistically significant positive effect on eligible people's mental health. I find this result dubious at best. First, given that we don't have information on individuals' documentation status, a small minority of the treatment group was *actually* able to receive the treatment, making it possible that something other than DACA implementation could be driving their result. Second, despite numerous efforts, I was unable to replicate the authors' results, and am frankly dumbfounded how they got the numbers they did. This held true when I switched to a regression discontinuity design, which was also unable to yield significant results. Finally, upon examining the trends for the outcome variables in the treatment and control group, it seems likely the parallel trend assumption doesn't even hold, rendering the entire difference-in-differences design invalid. Therefore, I do not believe Venkataramani's study was able to draw a compelling causal conclusion.

References:

"Consideration of Deferred Action for Childhood Arrivals." 2021. *United States Citizenship and Immigration Services.* https://www.uscis.gov/DACA

Gamboa, S. 2021. "Americans way off on number of Latinos they think are undocumented, poll finds." *NBC News.* https://www.nbcnews.com/news/latino/americans-way-number-latinos-think-are-undocumented-poll-finds-rcna2464

Martinez, O., Wu, E., Sandfort, T., Dodge, B., Carballo-Dieguez, A., Pinto, R., Rhodes, S., Moya, E., & Chavez-Baray, S. 2016. "Evaluating the Impact of Immigration Policies on Health Status Among Undocumented Immigrants: A Systematic Review." *Journal of Immigrant and Minority Health* 18(1): 288.

"National Health Interview Survey (NHIS)." n.d. *United States Census Bureau.* https://www.census.gov/programs-surveys/nhis.html#:~:text=The%20National%20Health%20Interview%20Survey%20(NHIS)%20is%20the%20principal%20source,the%20U.S.%20Centers%20for%20Disease

Venkataramani, A.S., Shaha, S.J., O'Brien, R., Kawachi, I., & Tsai, A.C. 2017. "Health Consequences of the US Deferred Action for Childhood Arrivals (DACA) immigration programme: a quasi-experimental study." *Lancet Public Health* (2): 175-181.

Appendix:

Table 1, Descriptive statistics of study population

	ORIGINAL AUTHOR'S RESULTS				MY RESULTS			
	Eligible for DACA		**Not eligible for DACA**		**Eligible for DACA**		**Not eligible for DACA**	
	Pre-DACA	Post-DACA	Pre-DACA	Post-DACA	Pre-DACA	Post-DACA	Pre-DACA	Post-DACA
# Respondents								
Self-reported health	2188	1784	6331	4670	1257	1326	4769	3715
Mental health	598	540	2217	1680	390	450	1914	1497
Self-reported overall health (Likert scale 1-5)	3.99 (0.91)	4.00 (0.94)	3.83 (0.98)	3.81 (0.98)	4.01 (0.03)	4.05 (0.03)	3.87 (0.02)	3.84 (0.02)
Fair or poor health	95 (4%)	101 (6%)	523 (8%)	408 (9%)	4.4% (0.007)	3.8% (.006)	7.4% (0.004)	7.7% (0.005)
K6 score (0-24)	3.06 (4.49)	2.66 (4.3)	2.72 (4.57)	2.70 (4.38)	2.29 (0.21)	2.02 (0.20)	2.05 (0.09)	2.04 (0.10)
Moderate or worse psychological (K6 score >= 5)	168 (28%)	133 (25%)	554 (25%)	423 (25%)	19.1% (0.02)	17.0% (0.02)	15.3% (0.01)	16.9% (0.01)
Gender, female	1116 (51%)	906 (51%)	3270 (52%)	2428 (52%)	48.3% (0.02)	46.9% (0.01)	45.4% (0.01)	47.4% (0.01)
Age (years)	23.0 (3.32)	25.39 (4.02)	36.9 (6.73)	38.27 (6.71)	22.95 (0.09)	24.41 (0.13)	36.54 (0.14)	37.97 (0.15)
Age at immigration (years)	9.6 (4.19)	10.6 (3.81)	24.2 (6.48)	24.9 (6.01)	9.26 (0.13)	9.96 (0.10)	24.19 (0.14)	24.89 (0.13)
Census region								
Northeast	240 (11%)	168 (9%)	884 (14%)	536 (11%)	10.4% (0.01)	10.7% (0.01)	14.4% (0.01)	12.3% (0.01)
Midwest	161 (7%)	199 (11%)	536 (8%)	430 (9%)	8.5% (0.02)	11.1% (0.01)	8.5% (0.01)	9.6% (0.01)
South	728 (38%)	587 (33%)	2216 (35%)	1750 (33%)	38.3% (0.02)	36.8% (0.02)	39.9% (0.01)	41.6% (0.02)
West	1059 (42%)	830 (47%)	2695 (43%)	1954 (48%)	42.8% (0.02)	41.4% (0.02)	37.3% (0.01)	36.5% (0.02)

Original author's results for binary outcomes are formatted as: raw # (percentage). Mine are: percentage (standard error). The reason for this difference is discussed in the text of the paper.

Table 2, Difference-in-Differences Estimate

	Self-reported health (Likert scale score 1-5)	Poor or fair health	K6 score (0-24)	Moderate or worse psychological distress (K6 score >= 5)
Regression method (estimate)	Least squares (b)	Logistic (adjusted OR)	Poisson (adjusted IRR)	Logistic (adjusted OR)
ORIGINAL AUTHORS				
Difference-in-differences estimate (95% CI)	0.056 (-0.024 to 0.14)	0.98 (0.66 to 1.44)	0.78 (0.56 to 0.95)*	0.62 (0.41 to 0.93)*
P value	0.17	0.91	0.020	0.022
Number	14973	14973	5035	5035
MY ESTIMATES				
Difference-in-differences estimate (95% CI)	0.058 (-0.032 to 0.148)	0.84 (0.51 to 1.39)	0.91 (0.70 to 1.19)	0.78 (0.50 to 1.22)
P value	0.205	0.50	0.48	0.27
Number	11067	11067	4189	4189

* p < .05, ** p < .01, *** p <.001

Table 3, Sensitivity Analyses

	Self-reported health (Likert scale score 1-5)	Poor or fair health	K6 score (0-24)	Moderate or worse psychological distress (K6 score >= 5)
Sample restricted to 2010-2015				
Regression method (estimate)	Least squares (b)	Logistic (adjusted OR)	Poisson (adjusted IRR)	Logistic (adjusted OR)
ORIGINAL AUTHORS				
Difference-in-differences estimate (95% CI)	0.017 (-0.072 to 0.11)	1.00 (0.65 to 1.54)	0.69 (0.52 to 0.92)	0.56 (0.36 to 0.87)
P value	0.71	0.99	0.010	0.011
Number	11672	11672	4008	4008
MY RESULTS				
Difference-in-differences estimate (95% CI)	0.02 (-0.08 to 0.12)	0.85 (0.50 to 1.44)	0.83 (0.62 to 1.10)	0.68 (0.43 to 1.10)
P value	0.73	0.55	0.20	0.12
Number	8816	8816	3338	3338
Sample restricted to 2010-2015 and <40 years old				
Regression method (estimate)	Least squares (b)	Logistic (adjusted OR)	Poisson (adjusted IRR)	Logistic (adjusted OR)
ORIGINAL AUTHORS				
Difference-in-differences estimate (95% CI)	-0.013 (-0.11 to 0.08)	1.16 (0.73 to 1.83)	0.76 (0.56 to 1.02)	0.61 (0.38 to 0.99)
P value	0.78	0.53	0.073	0.044
Number	8715	8715	2963	2963

MY RESULTS				
Difference-in-differences estimate (95% CI)	0.01 (-0.09 to 0.12)	0.96 (-0.53 to 1.72)	0.88 (0.65 to 1.20)	0.74 (-.45 to 1.22)
P value	0.81	0.89	0.42	0.23
Number	6099	6099	2281	2281
Sample restricted to less than high-school education (falsification test)				
Regression method (estimate)	Least squares (b)	Logistic (adjusted OR)	Poisson (adjusted IRR)	Logistic (adjusted OR)
ORIGINAL AUTHORS				
Difference-in-differences estimate (95% CI)	0.043 (-0.06 to 0.14)	0.72 (0.49 to 1.06)	1.07 (0.76 to 1.49)	1.38 (0.89 to 2.15)
P value	0.40	0.11	0.67	0.15
Number	16552	16552	5696	5696
MY RESULTS				
Difference-in-differences estimate (95% CI)	0.002 (-0.10 to 0.10)	1.02 (0.63 to 1.65)	1.22 (0.83 to 1.80)	1.11 (0.61 to 2.04)
P value	0.97	0.95	0.31	0.73
Number	15851	15851	5777	5777

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4, Difference-in-Differences without Ambiguous Cases

	Self-reported health (Likert scale score 1-5)	Poor or fair health	K6 score (0-24)	Moderate or worse psychological distress (K6 score >= 5)
Regression method (estimate)	Least squares (b)	Logistic (adjusted OR)	Poisson (adjusted IRR)	Logistic (adjusted OR)
ORIGINAL AUTHORS' RESULTS				
Difference-in-differences estimate (95% CI)	0.056 (-0.024 to 0.14)	0.98 (0.66 to 1.44)	0.78 (0.56 to 0.95)*	0.62 (0.41 to 0.93)*
P value	0.17	0.91	0.020	0.022
Number	14973	14973	5035	5035
MY ORIGINAL ESTIMATES				
Difference-in-differences estimate (95% CI)	0.058 (-0.032 to 0.148)	0.84 (0.51 to 1.39)	0.91 (0.70 to 1.19)	0.78 (0.50 to 1.22)
P value	0.21	0.50	0.48	0.27
Number	11067	11067	4189	4189
MY ESTIMATES WITH NEW SPECIFICATION				
Difference-in-differences estimate (95% CI)	0.41 (-0.06 to 0.14)	1.01 (0.58 to 1.74)	0.94 (0.70 to 1.26)	0.77 (0.47 to 1.25)
P value	0.41	0.98	0.68	0.29
Number	11067	11067	4189	4189

* p < .05, ** p < .01, *** p <.001

Table 5, Regression Discontinuity Estimates

	Self-reported health (Likert scale score 1-5)	Poor or fair health	K6 score (0-24)	Moderate or worse psychological distress (K6 score >= 5)
Regression method (estimate)	Least squares (b)	Least squares (b)	Poisson (adjusted IRR)	Logistic (adjusted OR)
Over-age at implementation (RD estimate)	2.71	1.58	0.00	0.00
P value	0.81	0.30	0.72	0.81
Age at implementation (running variable)	-0.02	0.01	1.04	1.17*
Age at implementation ^ 2		0.00		
Age at implementation * over-age at implementation	-0.09	-0.05	1.39	1.45

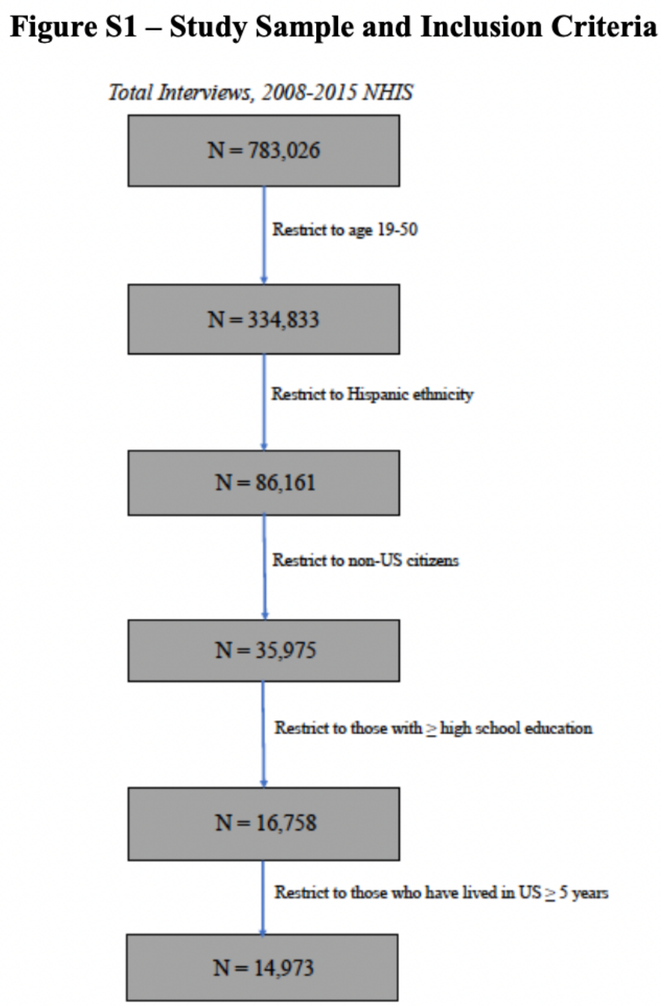
Gender, age at immigration, region, interview quarter-year covariates not shown.
* p < .05, ** p < .01, *** p <.001

Table 6, Parallel Trend Coefficients

	Self-reported health (Likert scale score 1-5)	Poor or fair health	K6 score (0-24)	Moderate or worse psychological distress (K6 score >= 5)
Regression method (estimate)	Least squares (b)	Logistic (adjusted OR)	Poisson (adjusted IRR)	Logistic (adjusted OR)
DACA eligibility * survey year (Var of interest)	0.02	0.89	0.09	1.14
P value	0.59	0.32	0.32	0.35
Survey year	-0.02*	1.09*	0.02	1.03
DACA eligibility	-33.11	$4x10^{100}$	-171.68	0.00

Gender, age at immigration, age at DACA implementation, and region covariates are not shown.
* p < .05, ** p < .01, *** p <.001

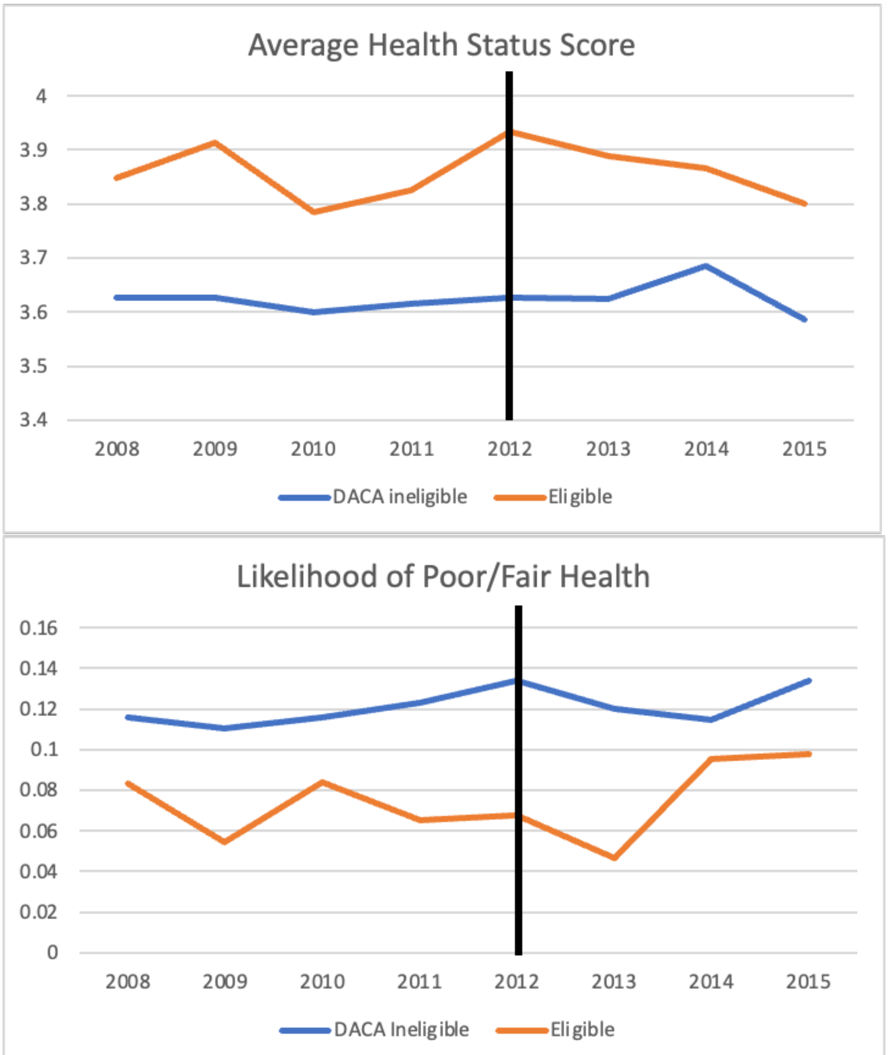
Figure 1, Original Author's Sample Specification

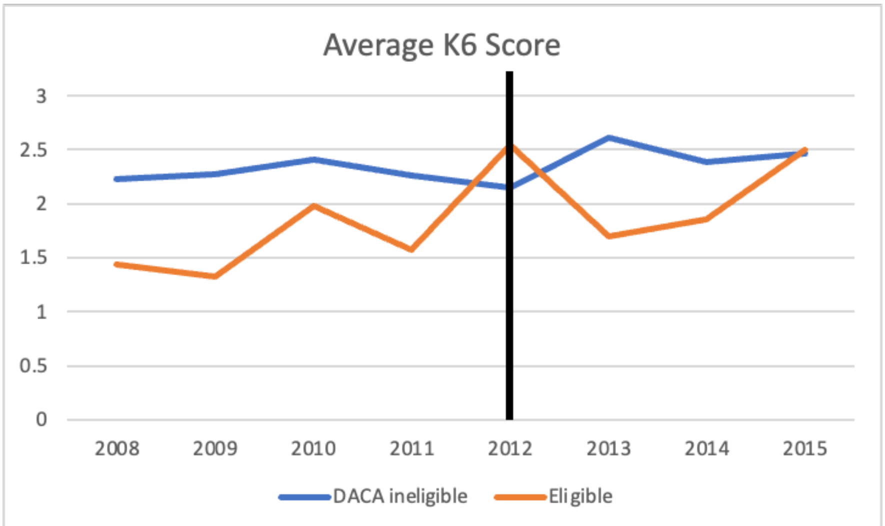


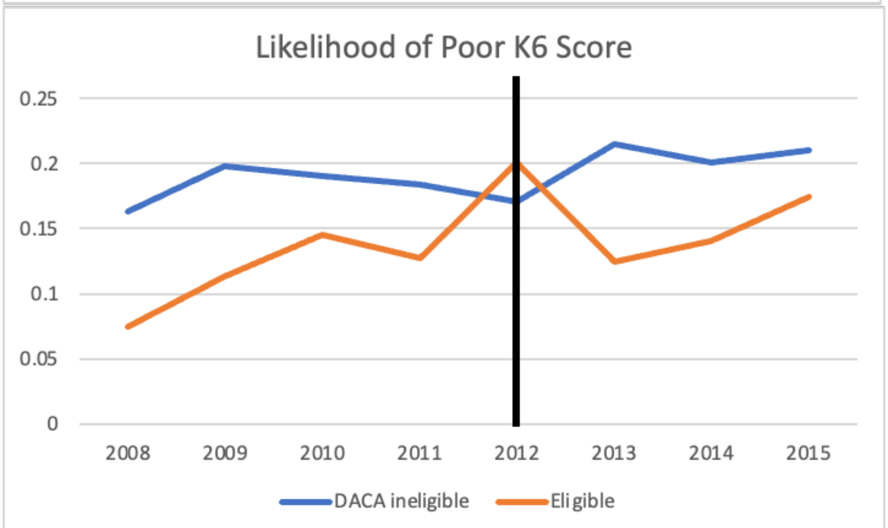
Notes: Figure displays changes in sample size after applying each restriction. The final sample size of 14,973 reflects all individuals in the analytic sample with measures of self-reported overall health. Of this group, a randomly selected third (5,035) were administered the K6.

Figure 2, Parallel Trend Analysis

The vertical black line represents the date of DACA implementation.




Average K6 Score
3
2.5
2
1.5
1
0.5
0
2008
2009
2010
2011
2012
2013
2014
2015
DACA ineligible
Eligible


Likelihood of Poor K6 Score
0.25
0.2
0.15
0.1
0.05
0
2008
2009
2010
2011
2012
2013
2014
2015
DACA ineligible
Eligible