I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13350

The CPS Citizenship Question and Survey Refusals: Causal and Semi-Causal Evidence Featuring a Two-Stage Regression Discontinuity Design

Robert Bernhardt
Phanindra V. Wunnava

JUNE 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13350

The CPS Citizenship Question and Survey Refusals: Causal and Semi-Causal Evidence Featuring a Two-Stage Regression Discontinuity Design

Robert Bernhardt
Federal Reserve Bank of Chicago

Phanindra V. Wunnava
Middlebury College and IZA

JUNE 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The CPS Citizenship Question and Survey Refusals: Causal and Semi-Causal Evidence Featuring a Two-Stage Regression Discontinuity Design

The unsuccessful attempt to add a citizenship question to the Census has drawn attention to citizenship questions on other surveys. Simultaneously, researchers have noted a recent increase in Current Population Survey non-response. We combine these topics, studying the effect of the CPS citizenship question on refusals. We use the question's sudden introduction in 1994 as a natural experiment and obtain causal estimates via a regression discontinuity design (RDD). In January 1994, we find an immediate and sustained 20-50% jump in refusals. However, this cannot be attributed to the question alone, as numerous other survey characteristics were revised. We employ a two-stage RDD to relate state-specific refusal discontinuities to state characteristics. Discontinuity size is positively related to non-citizen and Hispanic populations, and a proxy for citizenship question item non-response. An 8% increase in refusals is potentially attributable to the question. Moreover, at the threshold, there is weak evidence of a discrete decrease in states' reported Hispanic populations. When non-citizenship is observable, state non-citizen population is positively related with refusals. These results imply the question makes non-citizens and Hispanics reluctant to take the survey. We recommend there be a trial to precisely estimate the question's effects, and decide if it merits continuation.

JEL Classification: C23, C83, J15

Keywords: current population survey, non-response, survey refusal, citizenship status, immigration, Regression Discontinuity, panel data

Corresponding author:
Phanindra V. Wunnava
Department of Economics
Middlebury College
Warner Hall 502F
Middlebury, VT 05753
USA

E-mail: wunnava@middlebury.edu

I. Introduction

In the leadup to the 2020 Decennial Census, the Trump Administration attempted to add a citizenship question to the questionnaire. Critics of the proposal argued that this question would reduce response rates, particularly among vulnerable immigrant and Hispanic communities. A lengthy and contentious legal battle was resolved in 2019 when the Supreme Court ruled that the question could not be added.[1] However, this ruling has not settled all debate over the matter. The administration has issued an executive order compelling the Census Bureau to investigate the issue using administrative data, such as state drivers' license records,[2] arguing that some states might desire to draw representative districts based on the location of eligible voters, not the population at large. Unsurprisingly, this also has been challenged in court by civil rights groups such as the Mexican American Legal Defense and Educational Fund, who view it as a "racially discriminatory scheme" to increase the political power of white people at the expense of minorities, particularly Hispanics.[3] [4]

In response to the newfound prominence of the issue, increasing attention has been drawn to citizenship questions in other surveys. For instance, the 2019 Census Test Form, the American Community Survey (ACS), and the Current Population Survey (CPS) all feature citizenship questions.[5] These questions are often highly similar. For instance, the CPS question, the Census Test Form question, and the proposed Census citizenship question share the same five possible answers.[6] [7] [8]

In the case of the CPS, the citizenship question was first added as part of the redesigned survey that was developed in the late 1980's and early 1990's, and debuted as the official revised Current Population Survey in January 1994.[9] Save for numerous smaller adjustments, this version of the CPS is the one that persists as of 2020. The redesign made existing questions more clear and consistent, allowed increased computerization for quality control, and implemented new questions. One of the new questions was the aforementioned citizenship question, which has remained unchanged since its introduction.[10] [11]

Because the CPS is an optional survey of housing units, certain households will not participate for various reasons. These may include temporary occupant absence, language barrier, and survey refusal. In recent years, the percentage of households not completing the survey has increased significantly. Researchers have noted that almost all of the increase in survey non-response can be attributed to non-response from occupied households, which in turn is almost entirely driven by survey refusals.[12] Refusals are a form of self-selection, and like any form of self-selection, they have the

---

[1] Berman, Ari. "Trump Backs Down on the Census Citizenship Question." *Mother Jones* (blog), July 11, 2019.
[2] Narea, Nicole. "Trump Is Still Trying to Collect Citizenship Data for Redistricting." Vox, October 17, 2019.
[3] Berman, Ari. 2019. "Civil Rights Groups Challenge Trump's 'Racially Discriminatory Scheme' to Skew Redistricting." Mother Jones. September 14, 2019.
[4] Wang, Hansi Lo. "Latinx Advocacy Groups Sue To Block Citizenship Data Release By Trump Officials." NPR.org. Accessed September 17, 2019.
[5] Wang, Hansi Lo. 2019. "Why Is The Census Bureau Still Asking A Citizenship Question On Forms?" NPR.Org. August 9, 2019.
[6] Ibid
[7] Determined from official CPS Microdata Documentation (https://data.nber.org/data/cps_basic.html)
[8] Wang, Hansi Lo. "Skipping The 2020 Census Citizenship Question? You'll Still Be Counted." NPR.org, April 19, 2018.
[9] Cohany, Sharon R, Anne E Polivka, and Jennifer M Rothgeb. "Revisions in the Current Population Survey Effective January 1994." *Employment and Earnings*, February 1994, 25.
[10] Ibid
[11] CPS Microdata Documentation (https://data.nber.org/data/cps_basic.html)
[12] Bernhardt, Robert. "Rising CPS Non-Response: A Potential Source of Bias?" presented at the Presentation to UIUC Ph.D. students, Champaign, IL, August 20, 2018.

potential to be distinctly non-random, and therefore make the sample of responders non-representative of the population.

Citizenship status is strongly related to numerous factors such as labor force participation, years of education, and household structure.[13] Potentially, it may also be associated with attributes such as trust in government and perception of social cohesion. This could make non-citizens differentially likely or unlikely to participate in government surveys such as the CPS, particularly if immigration-related questions are included. In light of the increased salience of both survey citizenship questions and CPS non-response, it has become more important than ever to understand the impact that this question has on response and the accuracy of CPS labor market, demographic and educational statistics. If it is determined that the citizenship question has a significant deleterious effect on survey participation, then we recommend the Census Bureau conduct randomized trial and cost-benefit analysis to evaluate if the question ought to be included in future iterations of the survey.

This paper proceeds as follows: section II reviews the literature relating to survey and Census non-response, the effects of non-response, and immigration status. Section III describes in detail the data used, and the methodology employed in order to identify trends in non-response behavior and test if they are related to the citizenship question. Immediately following in section IV are the empirical results. The first wave of results reports the causal effect of the 1994 Redesign as a whole on survey refusal, non-response, and observed demographic characteristics via a regression discontinuity model. The following series uses a two-stage procedure designed to relate this causal effect to attributes of each state, in order to provide evidence of a relationship between the redesign and state demographic characteristics. We provide further supporting correlational evidence relating survey refusal to citizenship status using aggregate panel data regressions, and the final part of the results discusses the policy implications. Immediately after the results section, in section V, the major findings are summarized, and limitations as well as possible extensions are discussed.

II. Insights from Related Literature

In this section, we begin by providing background and defining key terms, including non-response and refusals, relating to the Current Population Survey and studies of survey methodology. We then proceed to review the literature on these topics. We discuss two primary branches of the literature. First, we discuss the previously uncovered relationships between non-response, citizenship, and Hispanic ethnicity in a variety of surveys and censuses. We then review the literature analyzing the effects of non-response, and ways to address this issue.

1) Current Population Survey Background & Critical Definitions

The CPS is one of the most well-known and important surveys conducted by the United States government. Per its own technical documentation, the CPS is the preeminent source of labor market information for the United States, and includes invaluable information pertaining to wages, education, and employment. Frequently discussed and theoretically crucial statistics, such as the unemployment rate and poverty rate, are derived from the CPS.[14] It is therefore essential that the CPS be made as accurate as possible, so that our measures of these vital statistics are as precise and accurate as possible.

---

[13] Passel, Jeffrey S, and D'Vera Cohn. "A Portrait of Unauthorized Immigrants in the United States." PewResearchCenter Project. Pew Hispanic Center, April 14, 2009.

[14] "Design and Methodology: Current Population Survey." Technical Paper 66. Washington, D.C.: U.S. Bureau of Labor Statistics, October 2006.

While the CPS has undergone many changes, including the major redesign implemented in 1994, many aspects of the survey have remained stable, such as the monthly frequency, approximate sample size, and 8-month rotating design featuring a 4-8-4 interview pattern.[15] [16] These factors result in a large amount of intertemporal consistency, and enable the utilization of the time series and panel data approaches that we employ, further discussed in the data and methodology section.

The CPS is not *legally* required, and those surveyed may choose not to answer certain questions (*item* non-response), or to refuse the survey in its entirety (*unit* non-response). Our focus, and that of the majority of the literature, is on unit non-response. Any type of non-response is, of course, highly problematic. In addition to increasing costs, the principle concern arising here is that those who do not respond are systematically different from those that do respond. If they are different in unexpected or difficult to measure ways, then this could introduce bias to CPS statistics that may not be corrected even after the use of demographic-based weighting.[17]

In the CPS, unit non-response occurs at the household level, and is categorized into 3 types. A *Type-A* non-response[18] indicates that a person, persons, or family is present at a housing unit, but an interview was not completed. *Type-B* non-response indicates a housing unit was suitable for occupation, but temporarily unoccupied. *Type-C* non-response indicates that a household is not fit for residential living, typically because it has been converted, destroyed, or moved (such as for a trailer or mobile home).

Within each category, there are several sub-categories. Of particular interest is Type-A Category-3: refused. These housing units are occupied, but an interview is not completed, and their listed reason for Type-A non-interview/non-response is survey refusal. Accordingly, this corresponds to households where the occupiers explicitly refuse the survey, as opposed to other Type-A non-interview possibilities such as *temporary absence*, *language barrier*, or *other occupied*. Because a refusal is the result of an explicit choice, in which households directly self-select out of the sample, it is particularly worrisome to researchers.

2) Non-Response, Citizenship & Hispanic Ethnicity

Despite the importance of understanding bias in the CPS, as well as the increasing prominence of survey citizenship questions, as far as we have found, there have not been any high-quality academic papers analyzing the citizenship question's effect on response in the Current Population Survey. However, while those three topics have not been studied together, there is a large literature combining various aspects of these issues.

The most directly applicable papers are two recent studies analyzing the effects of the proposed citizenship question on the 2020 Census. The first of these is a 2019 study by researchers from the Shorenstein Center within the Harvard Kennedy School. They created a mock-census form and asked representative samples of both Hispanics and non-Hispanics to complete it, randomizing which individuals received the form including the citizenship question. Their results show the citizenship question caused a significant increase in item non-response, and discouraged individuals from reporting

---

[15] Bureau of Labor Statistics Staff. "Briefing Materials on the Redesigned Current Population Survey." Bureau of Labor Statistics, January 1994.

[16] Additional information about the CPS Structure is included in Appendix A: Notes on the Current Population Survey.

[17] "Design and Methodology: Current Population Survey." Technical Paper 66. Washington, D.C.: U.S. Bureau of Labor Statistics, October 2006.

[18] The official microdata documentation uses the terms *non-interview* and *noninterview* (e.g. Type-A non-interview). We typically use the phrase *non-response*, which we find more intuitive and useful, to indicate the same concept. However, the two are interchangeable in our paper.

Hispanic household members. They estimated that the Hispanic Population would be underestimated by 12% if the question was added to the 2020 Census.[19]

The second study comes from the Census Bureau itself, and was published in *Demography* in 2019. The authors compared response rates from the American Community Survey (which included a citizenship question) and the 2010 Census (which did not feature a citizenship question). They found that households with non-citizens were disproportionately less likely to respond to the ACS. Extrapolating this result, they estimated a citizenship question on the 2020 Census could cause a 2% reduction in overall response, increase costs, and decrease the quality of the population count.[20]

Other research has looked at the non-response and refusal in other surveys. Researchers studying the Michigan Panel Study of Income Dynamics (PSID) have found that the non-response rates for immigrants were significantly higher than the core group of native-born American citizens. This has persisted despite financial incentives for response, and other extensive measures designed to increase survey participation.[21] Studying a survey in Denmark, Deding et al. (2008) concluded immigrants, typically from Pakistan or Turkey in their sample, were disproportionately less likely to be successfully contacted by surveyors, and less likely to participate once contacted. Thus, the non-response rates of immigrants were significantly higher than those for native-born Danes.[22]

Conversely, Bachmier et al. (2014) obtained a null result, finding that a citizenship question did not reduce response rates in the most vulnerable group, unauthorized immigrants. Analyzing the Los Angeles Family and Neighborhood Survey and the Survey of Income and Public Participation, the authors found that upon the introduction of citizenship questions, there was no "appreciable 'chilling effect' on survey participation of unauthorized immigrant respondents." However, it is important to note that these surveys were not conducted by the Federal government, and that extensive legal guarantees were made that information in the survey would remain confidential.[23]

Further research has looked at these topics, but from different angles. Van Hook & Bachmeier (2013) analyzed the accuracy of immigrant's self-reporting of citizenship status in the American Community Survey. They replicated the methods and results from earlier studies and determined that immigrants may often misrepresent their citizenship status, particularly recent arrivals and immigrants from Mexico.[24] A 2017 paper by Pedraza et al. studied the salience of immigration on healthcare utilization. They found that Latino *citizens* are less likely to use healthcare services when topics like immigration status are mentioned. They also concluded that citizens may perceive information shared with healthcare providers might not be secure, particularly if they know someone who has been deported.[25] It is highly plausible such information insecurity anxiety could extend to government surveys as well.

---

[19] Baum, Matthew A., Bryce J. Dietrich, Rebecca Goldstein, and Maya Sen. 2019. "Estimating the Effect of Asking About Citizenship on the U.S. Census." *Shorenstein Center Discussion Paper*, March.

[20] Brown, J. David, Misty L. Heggeness, Suzanne M. Dorinski, Lawrence Warren, and Moises Yi. 2019. "Predicting the Effect of Adding a Citizenship Question to the 2020 Census." *Demography* 56 (4): 1173–94.

[21] Tourangeau, Roger, and Thomas J. Plewes, eds. 2013. *Nonresponse in Social Science Surveys: A Research Agenda*. Washington, D.C.: The National Academies Press.

[22] Deding, Mette, Torben Fridberg, and Vibeke Jakobsen. 2008. "Non-Response in a Survey among Immigrants in Denmark." *Survey Research Methods* 2 (3): 107–21.

[23] Bachmeier, James D., Jennifer Van Hook, and Frank D. Bean. 2014. "Can We Measure Immigrants' Legal Status? Lessons from Two U.S. Surveys." *The International Migration Review* 48 (2): 538–66.

[24] Van Hook, Jennifer, and James D. Bachmeier. "Citizenship Reporting in the American Community Survey." *Demographic Research* 29 (July 2, 2013): 1–32.

[25] Pedraza, Franciso I., Vanessa Cruz Nichols, and Alana M. W. LeBron. "Cautious Citizenship: The Deterring Effect of Immigration Issue Salience on Health Care Use and Bureaucratic Interactions among Latino US Citizens." *Journal of Health Politics, Policy and Law* 42, no. 5 (October 2017): 925–60.

3) Quantifying Non-Response and its Effects

The other branch of the literature has focused on measuring non-response, understanding its deleterious effects, and mitigating these effects. While this may not directly inform our understanding of the relationship between immigration-related questions and responses, it motivates our research question by quantifying the negative effects of non-response. From 1994 to 2018, approximately 75% of households selected in the CPS sample participated. Over this time period, however, there has been a marked increase in non-response, which has increased from approximately 22% to 28%. Particularly striking has been the increase in refusals. In the eight-year period from January 2010 to January 2018, survey refusals increased from less than 4% of households to over 10% of households. Not only is this trend large, but it is consistent and methodical: almost all 12-month intervals in this time frame experienced a discernible increase in refusals.[26]

Moreover, previous studies have also looked at the behavior of non-response during the period of the 1994 redesign. In a 2017 paper, Krueger et al. studied the effect of non-response on rotation group bias in the Current Population Survey. They found that rotation group bias increases discretely in January 1994, and attribute this increase to a simultaneous increase CPS non-reponse, particularly Type-A non-response, which they also found to have increased discretely at that time. Similar to Bernhardt (2018), they note that there has been a large recent increase in non-response driven primarily by survey refusals.[27]

Numerous studies have found that non-response biases some of the most important survey based statistics. A 2015 study from the University of Kentucky Poverty Research Center found that the official poverty rate, which is derived from the March supplement of the CPS, was underestimated by a percentage point due to the relationship between low socio-economic status and non-response.[28] Likewise, a 2018 working paper by Heffetz & Reeves found that non-response biases key statistics in a variety of major surveys. They utilized a design based on the number of attempts required to successfully contact a household, inferring that those who were harder to contact were relatively more similar to non-responders. They found that difficult to reach respondents are systematically different in their unemployment and labor force participation rates in the CPS, the obesity rate in the Behavioral Risk Factor Surveillance System (BRFSS), and household expenditures in the Consumer Expenditure Survey (CEX). This pattern was true even after controlling for the typical demographic factors used in corrective weighting schemes.[29]

Numerous methods have been proposed to address the problem of non-response. As discussed previously, financial incentives and ironclad guarantees of confidentiality have been used to encourage survey cooperation. Technology has also been used to help increase overall response. In the 2010 Census, for instance, new innovations were implemented to streamline data collection processes and increase communications with citizens.[30] Other studies have proposed *ex-post* correction methods for addressing non-response bias. A 2015 paper by Behagel et al., for instance, introduces a method for correction

---

[26] Bernhardt, Robert. "Rising CPS Non-Response: A Potential Source of Bias?" presented at the Presentation to UIUC Ph.D. students, Champaign, IL, August 20, 2018.

[27] Krueger, Alan B., Alexandre Mas, and Xiaotong Niu. "The Evolution of Rotation Group Bias: Will the Real Unemployment Rate Please Stand Up?" *The Review of Economics and Statistics* 99, no. 2 (May 2017): 258–64. https://doi.org/10.1162/REST_a_00630.

[28] Hokayem, C., Bollinger, C. & Ziliak,J., 2015. "The Role of CPS Nonresponse on the Level and Trend in Poverty." *University of Kentucky Center for Poverty Research Discussion Paper Series*, DP2014-05.

[29] Heffetz, Ori, and Daniel B Reeves. 2018. "Difficulty to Reach Respondents and Nonresponse Bias: Evidence from Large Government Surveys." *NBER Working Paper Series*, January 2018, 44.

[30] Desouza, Kevin C, and Akshay Bhagwatwar, 2012. "Leveraging Technologies in Public Agencies." *Public Administration Review*. 72 (4): 605–614.

similar to Heffetz & Reeves' identification strategy. In their correction procedure, the number of contacts required to reach a person is allowed to influence the sample weight assigned to that individual. Executed correctly, they showed that this model can correct for bias and result in more precise estimates.[31]

III. Data, Variables & Methodology

Naturally, in order to study the Current Population Survey, we use data from the CPS. Unlike some other approaches, we choose not augment our data with data from another survey. As previously discussed, the literature uses a variety of techniques to study the effects on or effects of non-response, including randomized controlled trial (Baum et al.), quasi-experimental design (Brown et al.), and differentiation of responders based on similarity to non-responders (Heffetz & Reeves, Behagel et al.) Our own approach is most thematically similar to Brown et al. in that we use a quasi-experimental design. However, instead of comparing two surveys, one with and one without a citizenship question, our variation occurs intertemporally: before and after the introduction of the citizenship question in the January 1994 CPS redesign. We employ a regression discontinuity design and a two-stage state-specific regression discontinuity design in order to estimate the effect of the citizenship question on refusal rates.

The remainder of this section is laid out as follows: the first part introduces the raw data, providing context and information about their origin and properties. Next, we discuss the methods by which we aggregate data into the state and national-level datasets which we use for our analysis. We then review the summary statistics and time series properties of the aggregated data. Lastly, we discuss our methodology, including the three types of model that we employ: regression discontinuity, two-stage regression discontinuity, and correlational multivariate panel regression.

1) Original Data

Our raw data are the CPS monthly microdata from January 1989 to December 2002, inclusive. This includes information on 25,051,781 individual responses and household non-responses, which are listed alongside each other. Each microdata observation comes with several hundred variables, indicating household and person identifiers, household interview characteristics, demographic and labor force information, relation to others in the household, and meta-variables known as allocation flags indicating whether or not answers for particular questions have been adjusted by survey researchers after completion.

In order to make a manageable data set, we import only selected variables of interest. For each observation, the imported variables include unique person and household identifiers, information about month in sample, interview type, response or non-response category, and a variety of labor force and demographic information. This includes people's age, sex, marital status, employment status, family income, race, and hours worked. For each month of the survey, we aggregate these data, using two different levels, creating two datasets: one aggregated at the state level, and one aggregated at the national level. Values in the aggregated data indicate such things as the household refusal frequency in a state (or the country), the average household income of a state (the country), or the portion of people in a state (the country) of a given race for that month of the survey. The advantage of this data management approach is that while individuals may have missing observations, entire states will have some observable information on all characteristics for all months.[32]

---

[31] Behagel, Luc., Bruno Crépon, Marc Gurgand, and Thomas Le Barbanchon, 2015. "Please Call Again: Correcting Nonresponse Bias in Treatment Effects Models." *The Review of Economics and Statistics*. December 2015, 97(5): 1070–1080.

[32] This excludes months in which a variable is not present at all in the CPS.

Starting in January 1994, the substantially redesigned version of the CPS was introduced. This version largely persists as of 2020, save for lesser modifications. The official listed goals of the redesign were to measure labor force characteristics more precisely, to increase the amount of data collected, to redefine variables for increased consistency and practicality, and to take advantage of computerization for correcting errors and smoothing data collection. For example, the reference week for labor force status questions was more clearly defined, increasing the consistency of answers, and new information was collected on multiple jobholders.[33]

These changes included new questions, such as the questions on citizenship status (microdata variable PRCITSHP), and Hispanic status (PEHSPNON), as well as information about Hispanic origin group (PRORIGIN[34]). These questions, particularly the citizenship question, could make immigrants and/or Hispanic people less likely to participate in the survey. Being of critical interest to this research, these variables are included for the entire sample period for which they are present, January 1994 to December 2002.

2) Data Transformation & Aggregation

The data used in the analysis are monthly aggregated state and national data derived from the CPS microdata. The aggregated variables are generated from the variables given by the Bureau of Labor Statistics in the microdata, typically by assigning indicators to indicate a certain status, and averaging said indicators over all relevant households or individuals in a state or the country. For instance, the microdata variables LFSR (Jan 1989 - Dec 1993) and PEMLR (Jan 1994 - Dec 2002) are used to generate the variable *laborforce_part* (indicating labor force participation). The microdata variables take on the values of -1 and 1 through 7 inclusive. A value of -1 is assigned to a non-responding household or an individual ineligible to receive the labor force question, such as a child under 15 years old. Values 1 and 2 indicate an individual is employed, while values 3 and 4 demonstrate unemployment, and values 5 through 7 correspond to an individual who is not in the labor force.

To aggregate, we generate an indicator of individual labor force participation equal to 1 if an individual is in the labor force (LSFR/PEMLR = 1, 2, 3, or 4), 0 if an individual is not in the labor force (LSFR/PEMLR = 5, 6, or 7), and missing otherwise, such as for ineligible individuals. During the aggregation process, the mean of non-missing observations of this variable is taken, indicating the average value of the indicator, and therefore the portion of eligible individuals at the level of aggregation (state or national) who are in the labor force in a particular time period.

Excluding indicators for fixed effects and interaction terms, the state aggregated panel data contain 35 variables, while the nationally aggregated data contain a subset of these.[35] Since the majority of the analysis depends on the state aggregated data, this section will henceforth be discussing those. The nationally aggregated data are defined identically except for each variable being calculated at the national level, and for the omission of spatial indicators for states and regions. The variable *gestfips* corresponds to the Federal Information Processing Standards (FIPS) code. It uniquely identifies states, and takes on a value between 1 and 56 for each of the 50 American States and Washington D.C. This variable is later complemented by *state_number*, which contains the same information but takes on the values 1 to 51.[36] Additionally, the variable *division* indicates which of the nine Census Geographic

[33] Cohany, Sharon R, Anne E Polivka, and Jennifer M Rothgeb. "Revisions in the Current Population Survey Effective January 1994." *Employment and Earnings*, February 1994, 25.

[34] This variable was later removed from the CPS from January 2003 onwards.

[35] For example, the variable indicating state is excluded in the national data.

[36] Despite not being included in the CPS, territories such as American Samoa and Guam are included in the official FIPS code listing. This makes the values of gestfips found in the CPS non-consecutive. For instance, Alaska has FIPS number 2, and Arizona has FIPS value 4, but American Samoa has value 3, meaning there is a

Regions or Divisions a particular state belongs to. The variables *date* and *mdate* encode the survey month in two different formats, but contain equivalent information.[37]

The variables *non_response*, *typeA*, and *refusal* indicate the response behavior of households in each state. They provide the portion of households that didn't respond, had a Type-A non-response, and refused the CPS, respectively, for each state and survey month. Variables *avg_hh_members* & *faminc* record the average number of individuals in responding households in a particular state and month, and the average annual family income.[38] Additionally, the variable *log_faminc* is generated from the variable faminc. This variable is the natural log of each state's average last-year total family income for each time period.[39]

Additional demographic variables report the average of responding individuals in each month. The variable *avg_age* indicates the mean age of all responders in a given state, while variables *HS*, *grade12*, *BACH* and *ACA_Assoc* indicate educational attainment. Variables HS and BACH evaluate the portion of those who have graduated from high school and 4-year college respectively, while grade12 and ACA_Assoc are slightly more lenient measures that count individuals who have attended grade 12 or completed several years of college education but may not necessarily have a diploma or Bachelor's degree. While BACH & HS are more commonly used measures of educational achievement, they are not necessarily ideal in this context due to changes in microdata variable definitions over time. In contrast, grade12 and ACA_Assoc are consistent over time, and are used whenever a regression involves a time period where the educational attainment variable definitions are changed. The variables *female*, *white*, *white_and_other*, and *black* indicate the portion of responding individuals in each state who reported being female, listed their race as white, listed their race as white or listed their race as other, and listed their race as black, respectively.

Likewise, variables *hispanic_origin* and *hispanic* indicate if an individual has listed a Hispanic ethnic origin, and reports being Hispanic, respectively.[40] The variable *entry_year* approximates the average entry year reported by all immigrants in that state in that month, using an interval conversion method similar to the one used for faminc. Critically, the variable *non_citizen* indicates the portion of responding individuals who reported not being a United States citizen. Economic variables *unemployment* and *laborforce_part* indicate the unemployment and labor force participation rates in each state and month. They are calculated using the same procedure that the BLS uses to calculate official national unemployment and labor force participation rates, but without using the Census Bureau and the Bureau of Labor Statistics' sample adjustment weights.

---

gap in values present in the CPS data. By reassigning the state-identifiers to be consecutive numbers, the technical aspects of the analysis were streamlined.

[37] The variable date records values as a number where the first four digits indicate year, and the last two indicate month (e.g. January 1989 is 198901). Conversely, mdate numbers the months, giving the even 1-unit month to month spacing necessary for a time trend control. It is also formatted to be recognized by Stata as a month for time series functionality.

[38] Specifically, the variable microdata variables (H$FAMINC & HUFAMINC) indicate total family income in the last twelve months from all sources in fourteen brackets, which are then converted to numerical dollar values. The conversion assigns each household the midpoint of its listed bracket (except the top bracket) before taking the average value across households in nominal dollars.

[39] The variable log_faminc is desirable for the following reasons: first, it accounts for what is known as the "tapering off effect," the tendency for marginal utility to decrease for each successive unit of income received. Secondly, the Levin-Lin-Chu test for panel unit roots indicates that log_faminc is stationary, whereas faminc is not.

[40] The first variable is derived from the detailed Hispanic Origin question (microdata variable PRORIGIN) found in all iterations of the CPS in our sample. The second is derived from a binary variable for Hispanic ethnicity (PEHSPNON) present in the documentation starting in January 1994. Where both exist, the information contained is identical.

Additionally, there are several variables recording item non-response. Variables *nc_nr*, *hispanic_origin_nr*, and *lfq_nr* indicate the portion of responders eligible for the citizenship question, Hispanic origin question, and labor force question, respectively, for which the variable is missing in the microdata. Somewhat similarly, the variable *nc_revision* indicates revisions to citizenship question response, and acts as a rough proxy for item non-response. This variable is derived from the allocation flag indicating how the value of the citizenship question microdata variable is adjusted in the final data, which is frequently due to it being left unanswered.[41] The value of this variable found in the aggregated data indicates the portion of individuals whose responses were altered in a given state and month.

In the course of analysis, a number of additional variables are created. The variable *post* is an indicator, taking on a value of 1 if an observation occurs after the redesign threshold (i.e. date ≥ 199401) and a value of 0 before the January 1994 redesign. The variables *month_* and *year_* indicate the month and year of that particular round of the survey, separating the information contained in mdate. A pair of indicator variables *critical_state1* and *critical_state2* are generated based on variables hispanic_origin and non_citizen.[42] Additionally, variables are generated corresponding to regression coefficient estimates, so that the size of the estimates can be compared to demographic features in the data for the two-stage regression.

3) Summary Statistics & Time Series Properties

The summary statistics of the state-level variables are listed below in Table 0.1, while detailed summary statistics including the maximum and minimum values are included in Table 0.2 in Appendix A. Since national level data come from the same source, the national data have virtually identical characteristics. Sample sizes are 8568 for variables present across the entire time period (51 spatial categories and 168 time periods), 3060 for variables only present before the redesign (51 × 60) and 5508 for variables introduced after the redesign (51 × 108). Both the mean values and time trends are consistent with external estimates.

Table 0.1: Variable Summary Statistics (State-Level Aggregated Data)

Definition	Variable	Sample Size	Mean	Standard Deviation
FIPS State Identifier	gestfips	8568	28.96078	15.67775
Total Non-Response	non_response	8568	0.213435	0.046162
Type-A Non-Response Rate	typeA	8568	0.04676	0.020211
Survey Refusals	refusal	8568	0.027231	0.011955
Average Household Members	avg_hh_members	8568	2.563158	0.152287
Average Family Income	faminc	8568	39557.73	9529.812
Survey Month	date	8568	199556.5	403.1512

[41] The Census Bureau microdata basic informational guide discusses the use of allocation flags for evaluating item non-response: https://www2.census.gov/programs-surveys/cps/methodology/PublicUseDocumentation_final.pdf

[42] Critical States are so designated because they rank in the top-10 in non-citizen or Hispanic population (critical_state1) or top-20 in the same categories (critical_state2). Further details are provided in Appendix A (Extended Notes on Critical States).

Survey Month (time series formatted)	mdate	8568	431.5	48.49939
Average State Age	avg_age	8568	35.30712	1.800807
High School Graduate Rate	HS	8568	0.783647	0.044944
Grade 12 Completion Rate	grade12	8568	0.794084	0.04502
Bachelor's Degree Attainment Rate	BACH	8568	0.205941	0.046159
Academic Associates Degree/Some College Rate	ACA_Assoc	8568	0.228912	0.051884
Portion Female	female	8568	0.517648	0.012376
Unemployment Rate	unemployment	8568	0.052316	0.016899
Labor Force Participation Rate	laborforce_part	8568	0.661002	0.041458
Labor Force Question Non-Response	lfq_nr	8568	0.005514	0.006589
Portion White	white	8568	0.842958	0.140682
Portion White or Other	white_and_other	8568	0.845845	0.140275
Portion Black	black	8568	0.105245	0.120373
Portion of Hispanic Origin	hispanic_origin	8568	0.057103	0.080394
Hispanic Origin Question Non-Response	hispanic_origin_nr	3060	0.001762	0.001522
Portion Hispanic	hispanic	5508	0.063193	0.084284
Portion Non-Citizens	non_citizen	5508	0.0381	0.03515
Average Immigrant Entry Year	entry_year	5508	1980.257	4.181538
Citizenship Question Non-Response[43]	nc_nr	5508	0	0
Citizenship Question Revision	nc_revision	5508	0.007547	0.006581
Census Bureau Division	division	8568	5.117647	2.510104

[43] The citizenship question variable (PRCITSHP) is never missing in the final published microdata for responding individuals, hence this variable has a uniform value of 0. It is therefore not used in any regression models.

Critical State 1 Designation	critical_state1	8568	0.254902	0.435832
Critical State 2 Designation	critical_state2	8568	0.45098	0.49762
Month	month_	8568	6.5	3.452254
Year	year_	8568	1995.5	4.031364
Natural Log of Family Income	log_faminc	8568	10.55669	0.240785
Threshold Indicator	post	8568	0.642857	0.479185
Alternative State Identifier	state_number	8568	26	14.72046

In order to preempt issues of spurious correlation, we tested all regressors for non-stationarity. For the national-level data, we used the Dickey-Fuller unit root test. At the 1% confidence level, we were able to reject the null hypothesis of a unit root for the variables laborforce_part, and black. Additionally, we were able to reject the null at the 15% level for the variables refusal, typeA, female, and unemployment. While many of the control variables tested positive for unit roots, the dependent variables were shown to be reasonably stationary. Moreover, given that the hypothesis from the literature is that there is a discrete level increase associated with the 1994 redesign, the fact that certain variables might potentially be non-stationary is neither surprising nor especially worrisome.

In the state-level aggregated data, we used the Levin-Lin-Chu unit root test. For all dependent and independent variables except for faminc, we were able to reject the null hypothesis of a unit root. While we were unable to do so for faminc, the variable log_faminc was determined to not have a unit root, allowing us to substitute that variable in and avoid issues of non-stationarity in the state-level data.

4) Methodology

In order to address whether or not the citizenship question had an effect on non-response, we use the sudden introduction of the question as a natural experiment. While response rates do have time trends, these are typically slow moving on a monthly-basis, and one can expect that all else equal, response, non-response, and refusal rates ought to be consistent across consecutive months, particularly after controlling for more stable yearly and long-term trends. Similarly, we expect that demographic and economic trends such as a state's racial makeup, average household size, and family income ought to be approximately similar month-to-month. This fact allows us to employ a regression discontinuity design for our causal identification strategy. This design models non-response using a linear time trend[44] before and after January 1994, and allows for a discontinuity at that point. Depending on specification, the model may estimate the size of the discontinuity in refusal or Type-A non-response rates for the nation at large, or for individual states. The outcome of interest is therefore a discontinuous increase in the proportion of surveys that are refused, for the country at large, or the individual states. If the identifying assumptions of the regression discontinuity are satisfied, then this increase represents the causal effect of the treatment applied in January 1994: the survey redesign.

For a regression discontinuity model, the identifying assumptions are that observations are equivalent before and after the threshold, except for the effects of the treatment applied at the threshold itself, that nothing besides the treatment changes at the threshold, and that the trend is correctly modelled

[44] This linear time trend is later augmented with month fixed effects.

before and after the intercept. In this context, the first assumption requires that the CPS and those interviewed be highly similar in December 1993 and January 1994, particularly in ways that would relate to immigration. Since changes in demographic factors and response rates are typically slow-moving, we then expect that this assumption is fulfilled.

The second assumption is that the only change that occurred was the treatment. If the treatment is defined to be the citizenship question in particular, this assumption is certainly not satisfied. The survey went through a major redesign, and any number of those changes could have potentially influenced the rate of refusals in the CPS. However, if the entire redesign is taken to be the treatment, we argue the assumption is satisfied, with the caveat being that those results apply to the redesign as a whole. In January 1994, there was no unusual political turmoil, no outbreak of disease or declaration of war in the United States, and no major legal (particularly immigration-related), economic, or technological shift.[45] Thus, the assumption is met. Lastly, we must assume that the trend is modelled correctly before and after the intercept. We will argue in the Empirical Results section that this is indeed the case.

The first model is a basic regression discontinuity design. It is designed to capture the cumulative causal effect of the redesign, without attributing it to any one factor, including the addition of the citizenship question. Since this model looks at the cumulative effect of the redesign, then our second identifying assumption holds, since we only seek the aggregate effect of all changes to the CPS in this instance. The primary purpose of this model is to establish that there was indeed a discontinuity at the threshold: without a discontinuity whatsoever, it cannot be argued that there is a discontinuity associated with specific questions.

$$REF_{it} = \beta_0 + \beta_1 \underset{(+)}{Month\_Year_t} + \beta_2 \underset{(+)}{Post_t} + \boldsymbol{\beta} \boldsymbol{X} + \boldsymbol{\lambda} + \varepsilon_{it}$$

The variables are defined as follows:

REF_{it}	The refusal rate in state *i* and survey month *t*
$Month\_Year_t$	The current survey month, allows for a linear time trend (the *running variable*)
$Post_t$	A dummy variable indicating the threshold at January 1994
X	A vector of demographic controls indicating the average value or rate of each characteristic in the state *i* and survey month *t*
λ	Month-specific fixed effects
ε_{it}	A classical stochastic error term

Based on the findings from Bernhardt (2018) and Krueger et al. (2017), our prior is that the time trend will have a positive coefficient as refusals have been increasing over time. Additionally, we expect that the inclusion of the citizenship question in the redesign will cause an increase in refusals, making the coefficient for the January 1994 threshold positive as well. This result would also align with the finding in the literature noted by Krueger et al. The coefficient expectations for the demographic controls are mixed. This model may also be run with other dependent variables. Most simply, one can substitute in broader measures of response behavior, such as Type-A non-response. Additionally, by substituting

[45] In order to investigate this assumption, we examined the Wikipedia category pages for 1993 in American Law, 1994 in American Law, December 1993 Events in the United States, and January 1994 Events in the United States. While not a thoroughly rigorous approach, none of these pages detailed any law or event that might have had a plausible nation-wide effect on CPS refusal in January 1994.

in the variable for Hispanic ethnicity, we may estimate the causal effect of the survey redesign on the portion of responding individuals who are Hispanic in each state and month.

While instructive, the primary drawback of the above model is that it is not sufficiently sophisticated to allow us to discern if an increase in refusals may be attributed to any particular aspect of the redesign. The solution to this limitation is the following (possibly novel) two-stage regression discontinuity model.[46] In the first stage, we estimate multiple intercepts and discontinuities, one for each state. It is then possible to extract the estimated results from the state-specific regression discontinuity model and relate the magnitude of these discontinuities to the characteristics of each state.

i. First Stage:

$$REF_{it} = \Sigma\, \alpha_i \;+ \underset{(+)}{\beta_1}\, Month\_Year_t + \Sigma\, \underset{(+)}{\delta_i}\, State_i \times Post_t + \boldsymbol{\beta}\, \boldsymbol{X} + \varepsilon_{it}$$

The variables are defined as before. Likewise, we expect that the linear time trend ought to have a positive coefficient, and the threshold coefficients δ_i ought to be positive as well. The coefficients of interest are the 51 state (and Washington D.C.)-specific baseline intercepts, α_1 through α_{51}, and particularly the 51 state-specific jump discontinuity coefficients, δ_1 though δ_{51}. For each state i, α_i indicates the baseline level of refusal in that state in the pre-1994 period, and $\alpha_i + \delta_i$ indicates the baseline state level in the post-1994 period. If these coefficients are significant and correlated in magnitude with the level of non-citizenship in that state, it would imply that states with higher levels of non-citizenship have higher levels of survey refusal. Most critically, for each state, δ_i represents the *causal* effect of the entire 1994 redesign on that state's refusal rate, per the primary identifying assumptions of the regression discontinuity design. Because there is variation in this causal effect estimated for each state, we can relate this variation in the second stage with variation in states' key immigration-related demographic characteristics, as follows:

ii. Second Stage:

$$Discontinuity_i = \beta_0 + \underset{(+)}{\beta_1}\, Characteristic_i + \varepsilon_i$$

The independent variable $Characteristic_i$ can be non-citizen population, Hispanic population, or citizenship question adjustment frequency. Characteristics are extracted from January 1994 in order to match discontinuities to current state attributes.[47] If non-citizens or Hispanic people are refusing the survey due to the citizenship question, states with greater numbers of non-citizens or Hispanics will have larger discontinuities. Likewise, higher rates of refusals are thought to be associated with rates of the proxy for citizenship question item non-response. Hence β_1 is expected to be positive.

The constant term β_0 estimates discontinuity that a state with negligible non-citizen or Hispanic populations would experience. It may be loosely interpreted as the component of the change in refusals independent of those groups, and therefore presumably unrelated to the citizenship question. Thus, it functions as an estimate of the effect of all other aspects of the redesign on refusals. It could therefore be either negative or positive.

[46] We are not currently aware of any other paper which has used the specific multiple-intercept two-stage regression discontinuity design that we employ.

[47] Later, January 1995 characteristics are used as a robustness check.

A relationship between non-citizen and Hispanic populations and the causal effect of the redesign would imply that the causal effect itself is driven by those factors. It would be plausible, therefore, to assume that if non-citizen populations are related to causal increases in refusals, that the citizenship question would be a proximate cause of that aversion to survey participation. This claim would be strengthened further by a positive β_1 in the specification with citizenship question revisions as the second stage regressor, which would imply the question *itself* is related to the causal increase.

The subscript in the second stage is *i*, not *it*: while the microdata begin with more than 25 million observations, and the aggregated data includes 8568 state-month observations, one may only estimate a single jump discontinuity for each state. Thus, the major limitation of the previous model is that it has relatively little statistical power. Therefore, in order to build additional supporting evidence, we run additional correlational regressions of the following form:

$$REF_{it} = \beta_0 + \boldsymbol{\beta}\, \boldsymbol{X} + \beta_1\, Month\_Year_t + \underset{(+)}{\beta_2}\, \underset{(+)}{Non\_Citizenship_{it}} + \varepsilon_{it}$$

In this regression, the variable $Non\text{-}Citizenship_{it}$ indicates the percentage of individuals in each state in each month who reported not being a citizen of the United States. Our prior is that it will be positively related with the refusal rate for each state and month, because states with more non-citizens will have a larger population that is theorized to be differentially unlikely to participate in the survey. While correlational, this model would indicate that controlling on other factors, states with higher levels of non-citizens have higher rates of refusal. This would be consistent with the citizenship question causing non-response.

Because no state or demographic group has complete non-response, then even though many non-citizens will not respond, one would expect to observe enough non-citizens responding in each state to observe that the relative levels of non-citizenship would be related to the level of non-response driven by the presence of non-citizens. Because of the non-observation noise generated by the particularly incomplete observation of a group that is (theoretically) particularly likely to not respond, we expect that any relationship discovered would be revealed to be much stronger if it became possible to later observe non-responders.

IV. Empirical Results

The empirical results replicate the finding from Krueger et al. (2017). Regardless of model, specification, restriction, or correction, there was a statistically significant discontinuous increase in refusals and Type-A non-response in January 1994, which is attributable to the redesign by the regression discontinuity design. Supporting regressions find that states with high fractions of non-citizens, Hispanic people, and citizenship question adjustments often appear to be affected most strongly, and that refusal rate is related to non-citizen populations more generally.

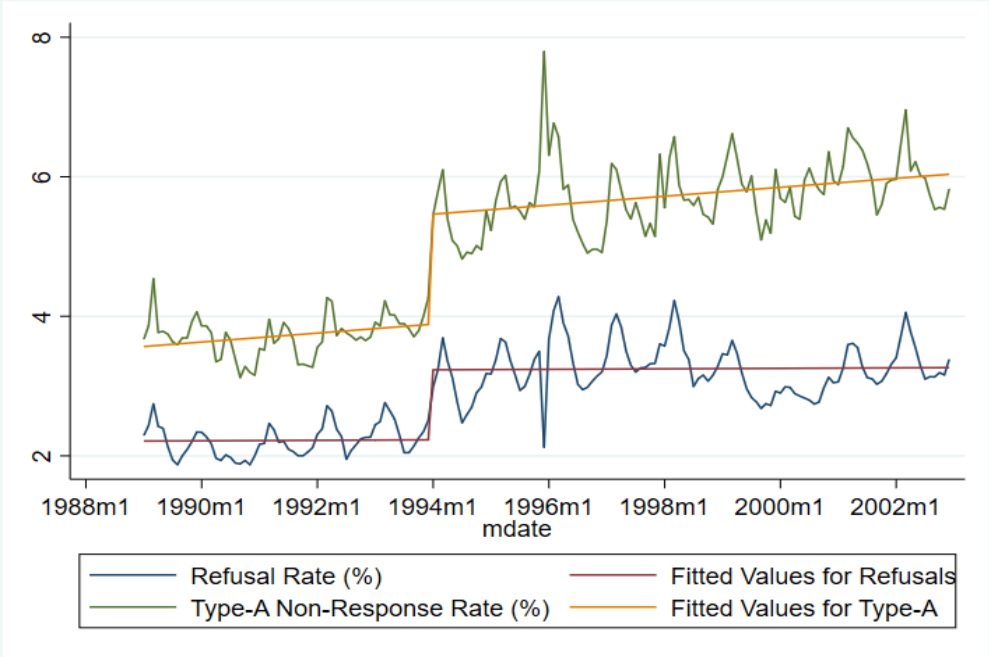
1) Basic Regression Discontinuity Models

i. National-Level Data Results

The basic regression discontinuity results feature a large instantaneous increase in the national refusal rate at the point of the redesign in 1994. Moreover, not only is the increase abrupt, but it is sustained across the remainder of the sample period. Figure 1 shows the evolution of the national refusal rate and Type-A non-response rate over time, as well as predicted values obtained from regressing each

dependent variable with a linear time trend and a threshold indicator. The result is striking: precisely in January 1994, there are large increases in both series. While both variables display cyclical behavior, their mean values after January 1994 are distinctly and persistently higher than before the redesign. The regression results were highly statistically significant, particularly the coefficient for the threshold term. These results were robust in both magnitude and significance to the inclusion of month-fixed effects and a vector of controls, including variables for each state's demographic makeup, average age, family income, and educational attainment.[48]

Figure 1: Type-A Non-Response & Refusals with Fitted Values (National Data)



To assess the validity of these results, we tested for the presence of serial correlation and heteroskedasticity, finding that all models displayed in Table 1 (located below) were afflicted with those problems.[49] In order to correct for these issues, we re-executed all six regressions using robust standard errors to address the artificial deflation caused by both statistical issues.[50] Invariably, the adjusted results revealed that the coefficient on the variable of interest (post) remained extremely statistically significant. Moreover, the goodness of fit indicators indicate an extremely high level of overall fit. Table 1 reports the results of the corrected regressions featuring robust standard errors.

[48] Full List of Controls: avg_hh_members, log_faminc, avg_age, grade12, ACA_Assoc, female, unemployment, laborforce_part, white, black, hispanic_origin

[49] We used the Durbin-Watson d-statistic test and the Breusch-Godfrey LM test to test for serial correlation. For heteroskedasticity, we used the Breusch-Pagan Cook-Weisberg test.

[50] To correct for serial correlation alone, we employed the Cochrane Orcutt-Iterative Method, and the Newey-West method. These results invariably remained highly statistically and practically significant. Full results for these corrections are available on request.

Table 1: Regression Discontinuity Results for Refusals (National Data)

Dependent Variable	Time Trend (Per Month)	Threshold Coefficient	Month FE	Demographic Controls	R^2
Refusals	0.000289 (0.00099)	1.00*** (0.103)	No	No	0.685
Refusals	0.00103 (0.000711)	0.942*** (0.0731)	Yes	No	0.835
Refusals	0.0263*** (0.00615)	0.796*** (0.0918)	Yes	Yes	0.910
Type-A NR	0.00534*** (0.00129)	1.58*** (0.133)	No	No	0.846
Type-A NR	0.00607*** (0.0011)	1.52*** (0.116)	Yes	No	0.903
Type-A NR	0.0137 (0.0096)	0.953*** (0.171)	Yes	Yes	0.9244

N = 168 [51] *** p < 0.001 ** p < 0.01 * p < 0.05
Robust standard errors in parentheses. Coefficient estimates given in percentage points.

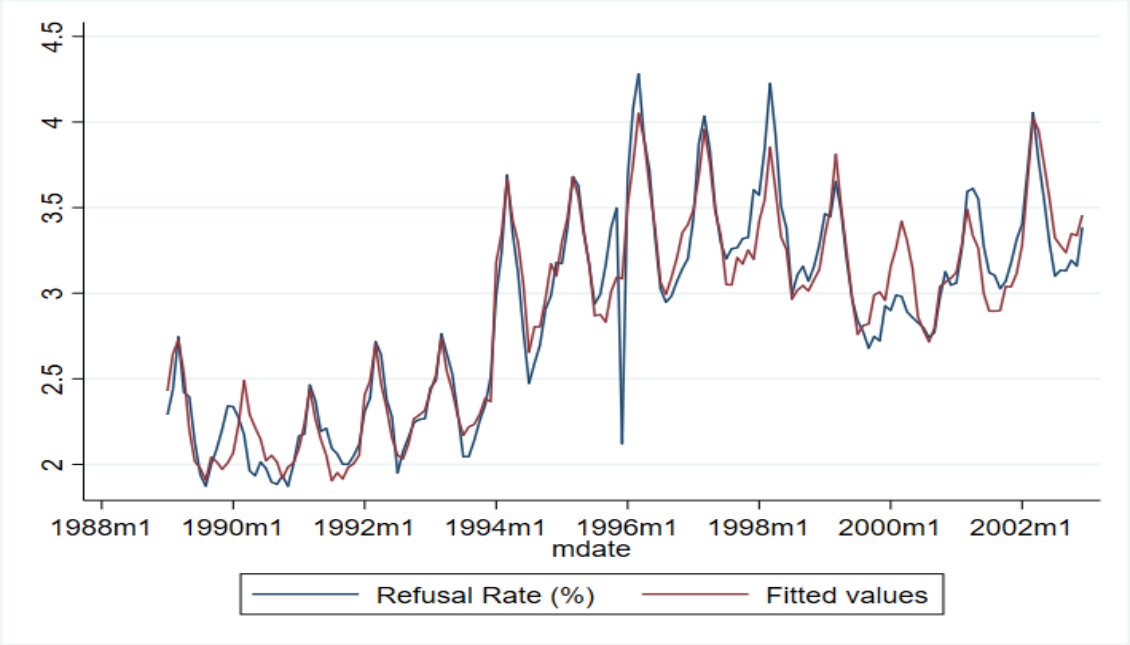
These estimated intercepts suggest an immediate overnight increase in the refusal rate in January 1994. Indeed, this is borne out in the refusal rate data itself, independent of regressions. The observed January 1994 refusal rate was 2.99%, 47.6 basis points higher than in December, which corresponds to an immediate observed increase of 19%. When using the regression models, this seismic shift is even more apparent. For each of the three refusal models above, the predicted values in July 1994 were 1.01, 0.954, and 0.486 percentage points higher than the predicted values for July 1993.[52] Those three increases correspond to year over year percentage increases of 45.2%, 48.4%, and 22.4%, respectively. Focusing on the threshold itself and applying the discontinuity point estimates to the observed Dec 1993 refusal rate of 2.51% suggests a 40.0%, 37.5%, or 31.7% increase in refusals is attributable to the threshold term. Because of the regression discontinuity design, this indicates that the causal effect of the redesign on the national refusal rate is estimated to be 30-40%.

Because of the high goodness of fit across models, we argue that the third RDD assumption, that the trend is modelled correctly, is met. Visually, this is supported by Figure 2, which shows actual and fitted values for national level data using a model that includes month fixed effects to account for response rate seasonality. The fitted values closely track the predicted values before and after the threshold, and do not deviate during the threshold transition. The only major deviation occurs in late 1995, which coincides with a historic government shutdown.

[51] The sample size of 168 for national level regressions arises from the 168 survey months, and thus 168 iterations of the CPS, from 1989 to 2002.

[52] Refusal rate is cyclical and highly dependent on month. In order to have an “apples to apples” comparison, we used predicted values from each model for the same month. Using the *actual* values of the refusal rate shows it increased from 2.045% to 2.473% over the same period, an increase of 20.9%.

Figure 2: Refusals with Fitted Values including Month FE & Dem. Controls (National Data)



ii. State-Level Data Results

Using the state-level aggregated data, we replicated the previous regressions. Again, the results showed a statistically significant increase in survey refusals at the threshold. The magnitude and statistical significance of the threshold increase was robust to the addition of a variety of controls. The results were qualitatively similar when using Type-A non-response as the dependent variable as well. Table 2 (below) features corrected[53] regression results for the state-level basic regression discontinuity models with refusals as the dependent variable.

Table 2: Regression Discontinuity Results for Refusals (State-Data)

Dependent Variable	Time Trend (Per Month)	Threshold Coefficient	Month FE	Division FE	Dem. Controls
Refusals	0.00394*** (0.000751)	0.591*** (0.063)	No	No	No
Refusals	0.00517*** (0.000724)	0.478*** (0.0616)	Yes	No	No
Refusals	0.00421*** (0.000691)	0.488*** (0.0603)	Yes	Yes	No
Refusals	-0.00276** (0.000831)	0.553*** (0.0543)	Yes	Yes	Yes

N = 8568 *** p < 0.001 ** p < 0.01 * p < 0.05

[53] We used the Wooldridge Test for panel data serial correlation and the Likelihood Ratio test for heteroskedasticity; all models tested positive for both problems. We corrected for these issues using a feasible Generalized Least Squares approach with heteroskedastic panels and both common and panel-specific AR-1 coefficients (these results are available on request).

Standard Errors in Parentheses. Coefficient estimates given in percentage points. Features PSAR corrections with heteroskedastic panels.

The state-level data regression discontinuity results largely mimicked the national-level results. The discontinuous jump in refusal rates was estimated to be between 45 and 60 basis points. These estimates were consistent across specifications as increasing numbers of controls were added, and were extremely significant, with corrected *z*-statistics ranging from 7.5-10.5.[54] The same four models were also estimated using Type-A non-response as the dependent variable. Like before, the results were qualitatively similar to the national-level discontinuity regressions. Across the same four specifications, the point estimates for the discontinuity were 1.27, 1.30, 1.27, and 1.30 percentage points, all of which were statistically significant at the 0.1% level. As before, the consistency of these estimates and their statistical significance affirms the robustness and strength of the results, emphatically demonstrating that the redesign caused a large increase in survey refusal and Type-A non-response more generally. Moreover, the results are similar in magnitude to the results obtained from the national-level data analysis.

To further test the robustness of these results, we re-estimated the PSAR-corrected regressions with sample restrictions which included only critical states. While restricting the sample to only critical states[55] reduced the sample size to 3864 and 2184, respectively, the point estimates for the coefficient remained statistically significant at better than the 0.1% level. As restrictions were implemented, the point estimates for the discontinuous increase in refusal rate became increasingly large and significant. For instance, in the model featuring controls, as well as month and Census Division fixed effects, the estimated discontinuity increased from 55 basis points in the unrestricted regression to 60 basis points, and finally 80 basis points in the most restricted model. The same trend held for Type-A non-response. As the sample was increasingly restricted, statistical significance remained high, and the point estimates increased. In the model including demographic controls, month fixed effects, and division effects, the point estimate for the discontinuity increased from 1.29 percentage points to 1.52, and then 1.74 pp in the most restricted model.

Lastly, we examined the effects of the redesign on states' reported Hispanic population. Using the same four regression discontinuity models as before, we obtained a consistent negative point estimate for the coefficient of the threshold of 40-62 basis points, suggesting that states' reported Hispanic populations experienced an average discontinuous decline of 0.4-0.6 percentage points. However, this result was only statistically significant in the model including controls, and both month and division fixed effects.[56] Moreover, after implementing the FGLS correction for serial correlation and heteroskedasticity, the results were statistically insignificant, though the coefficients remained negative.

Sample restrictions, however, improved the significance of the corrected results. When restricted to only critical-2 states, the coefficient became statistically significant and remained negative, with the point estimate increasing approximately seven-fold. When restricted to only the states with the highest number of Hispanic individuals and non-citizens, the estimate became even more negative and significant. The results of these regressions are detailed below in Table 3.

[54] Uncorrected *t*-statistics were as high as 24.91. Uncorrected point estimates were somewhat higher as well, ranging from 86-92 basis points.

[55] States where critical_state2 = 1 or critical_state1 = 1

[56] Controls excluded a state's Hispanic population in order to avoid collinearity issues.

Table 3: Regression Discontinuity Results for Hispanic Origin (State-Level Data)

Dependent Variable	Time Trend (Per Month)	Threshold Coefficient	Sample Size	Restriction
Hispanic Pop.	0.00934*** (0.00121)	-0.0431 (0.0667)	8568	None
Hispanic Pop.	0.032*** (0.00252)	-0.335* (0.14)	3864	critical_state2 = 1
Hispanic Pop.	0.073*** (0.00388)	-0.630*** (0.181)	2184	critical_state1 = 1

*** p < 0.001 ** p < 0.01 * p < 0.05

Standard Errors in Parentheses. Coefficient estimates given in percentage points. Features PSAR corrections with heteroskedastic panels. Demographic controls present in all models, but exclude dependent variable for each regression. All models include month & division fixed effects.

Thus, while mixed in significance, the evidence is roughly aligned with the notion that the Hispanic population observed in the CPS declined discontinuously after the redesign. As might be expected, this behavior is strongest and most noticeable in the states with the highest Hispanic and non-citizen populations. Overall, this behavior would be consistent with Hispanic individuals avoiding the CPS due to the citizenship question, resulting in fewer of them being detected in the pool of survey respondents.

2) Two-Stage Discontinuity Models

In order to relate the discontinuity to other characteristics, we employed state specific discontinuity models. These allowed us to connect the size of an estimated jump to that state's characteristics, drawing a connection between key factors and the causal effect of the redesign experienced by that state. We estimated six different specifications, each using the refusal rate as the dependent variable, and including a linear time trend, state-specific discontinuities,[57] and month fixed effects as regressors. Additional regressors varied on the inclusion of constant, state or division-specific baseline intercepts, and whether or not a vector of demographic controls was employed.[58]

To estimate each specification, we used the FGLS Panel-specific autocorrelation and heteroskedasticity correction as before. For the second stage, we used bivariate regressions of the form $Discontinuity_i = \beta_0 + \beta_1 \, Characteristic_i + \varepsilon_i$ to analyze the relationship between the size of each discontinuity and three factors of interest: a state's January 1994 non-citizen population, Hispanic population, and citizenship question adjustment rate. The regression results are displayed below in Table 4.1.

[57] Created by interacting state fixed effects with the post-1994 indicator.

[58] Controls exclude hispanic_origin. This is done so the size of each state's intercept may be correlated with its January 1994 state Hispanic population.

Table 4.1: Regression Results for State Discontinuities & January 1994 Characteristics

Specification	Significant[59] & Positive Discontinuities	Regression with Non-Citizen Population	Regression with Hispanic Population	Regression with Citizenship Q. Revision
State Baselines & Discontinuities, No Controls	36	0.0202	0.000924	0.156
State Baselines & Discontinuities, Controls	36	0.0156	-0.000341	0.123
Division Baselines, State Discontinuities, No Controls	34	0.0932***	0.0306**	0.451***
Division Baselines, State Discontinuities, Controls	38	0.0609***	0.0250**	0.291*
Constant Baseline, State Discontinuities, No Controls	31	0.129***	0.0360***	0.786***
Constant Baseline, State Discontinuities, Controls	35	0.0806***	0.332***	0.393***

N = 51 *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Coefficient values derived from second stage regression of discontinuities and each state characteristic in January 1994. Statistical significance determined by a *1-tailed* hypothesis test on the regression $Discontinuity_i = \beta_0 + \beta_1 Characteristic_i + \varepsilon_i$, testing the null $H_0: \beta_1 \leq 0$.

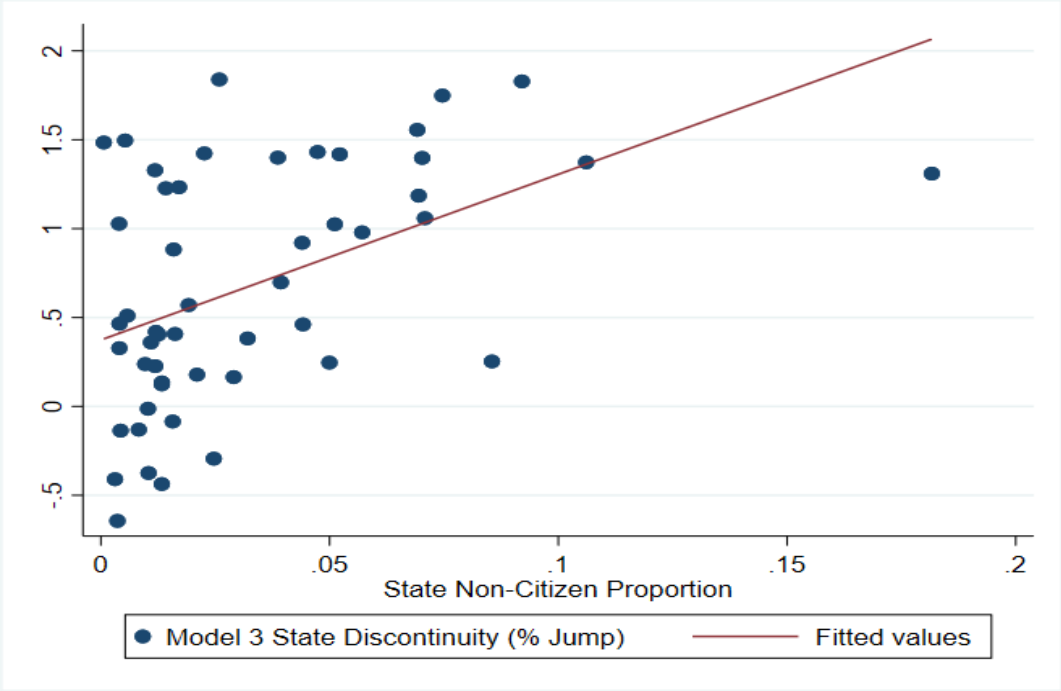
The results demonstrate that in all models, the vast majority of state-specific intercepts were statistically distinct from 0, and positive. Across all models, the size of the discontinuity was positively related with state non-citizen population, drawing a direct positive relationship between the causal effect the redesign had on refusal and the state's non-citizen population. Despite a small sample size of 51, these results were statistically significant in four of the six models presented. Figure 3, located below, depicts the positive relationship between the discontinuities from the third model, and non-citizen population. These findings suggest that on average, a percentage point change in state non-citizen population is associated with a 5-12 basis point change in the same direction in the discontinuity experienced by that state at the threshold. The average state non-citizen proportion in January 1994 was

[59] Significantly different from 0 at the 5% significance level based on a two-tailed hypothesis test.

3.23%. Applying this value to the significant model with the most conservative significant point estimate suggests non-citizenship can account for 20 basis points of the increase in refusals.[60]

Furthermore, there was a positive and frequently significant relationship between a state's discrete change in refusal rate and adjustments to the citizenship question. This evidence draws a direct parallel between adjustments of the citizenship question, typically caused by *item* non-response to that particular question, and the size of the jump in whole-survey refusals. Lastly, we analyzed the relationship between Hispanic population and discontinuity size. While this relationship was less strong and had smaller coefficients than that for non-citizen population, the coefficients were significant in four specifications, and all significant coefficients were positive. This shows that states with larger Hispanic population proportions tended to experience larger causal increases from the 1994 redesign.

Figure 3: Model-3 Discontinuity & Jan. 1994 Non-Citizen Pop. with Fitted Values



In order to test the robustness of the above relationships, we re-analyzed the intercepts by relating them to January *1995* state non-citizen population, Hispanic population, and citizenship question revision. By relating the intercepts to data from a different survey month, this suggests that any results obtained in Table 4.1 are not solely due to idiosyncrasies of the January 1994 survey. Because of our assumption that demographic changes are slow-moving, we expect that January 1995 demographics ought to be highly similar to 1994 demographics, and therefore the results should capture essentially the same underlying relationship.

The results using January 1995 data[61] were virtually identical for all models and independent variables. Both statistical significance and point estimates were highly similar to the models executed using January 1994 point estimates. This finding therefore affirms the robustness of those results.

Altogether, these regressions show that state discontinuities, causally attributable to the 1994 CPS redesign, are positively related with non-citizen population, Hispanic population, and citizenship question revision. The implication behind these findings is that states with more non-citizens and

[60] We have $0.0323 \times 0.0609 \approx 0.002$.

[61] These results are compiled in Table form in Table 4.2, located in Appendix A.

Hispanics are particularly affected by the redesign, hence causing them to experience larger increases in refusal rate. Moreover, the relationship with citizenship question revision suggests that avoidance of the citizenship question in particular is associated with the increase in whole-survey refusals.

3) Correlational Regressions

In this section, we sought to uncover a longer-term relationship between survey refusals, non-response, and non-citizenship status in the period after the question was introduced. We executed a variety of regressions with state survey refusals as the dependent variable and various controls, including year and month fixed effects, division fixed effects, and demographic characteristics.[62] As was the case previously, these regressions were plagued by the presence of heteroskedasticity and serial correlation. We corrected for these issues in the same manner, using the FGLS estimator and allowing for heteroskedastic panels and panel-specific AR-1 corrections.

Across a variety of specifications, greater portions of non-citizens in a state were consistently associated with higher levels of non-response. The PSAR corrected results were extremely significant, with *z*-values uniformly in excess of 8. Detailed regression results are provided in Table 5, located below. The coefficient point estimates ranged from 5-10 basis points. This is similar to the second stage state-specific discontinuity results, where a percentage point change in non-citizen population was estimated to cause a 5-12 basis point change in *the jump in* refusals at the threshold. Thus, the correlational results corroborate the results from the two stage regression discontinuity through a different design.

Based on the most conservative regression estimate, we estimate that a one percentage point change in the non-citizen population of a state would be associated with a 5.3 basis point change in the same direction in refusals, holding all else constant. This result corresponds to an elasticity of 0.0659. This suggests that, holding all else constant, a 10% change in the number of non-citizens would, on average, cause a 0.7% change in the same direction in refusals.

Equivalent regressions were run for Type-A non-response. Non-citizenship was strongly positively related to Type-A non-response. The non-citizenship coefficient had a *z*-statistic between 4 and 12, indicating an extremely strong statistical relationship. Again using the most conservative estimate, a percentage point change in a state's non-citizenship rate is expected to be associated with a 7.4 basis point change in the same direction in the Type-A non-response rate, holding all else constant. The coefficients for Type-A non-response were larger than those for refusals, which is expected given Type-A is a broader category of non-response. For the most conservative estimate, the elasticity with respect to non-citizen population was calculated to be 0.0526. This value was similar in magnitude to the elasticity of refusals, a component of Type-A non-response, but somewhat smaller. This suggests that Type-A non-response is moderately less responsive to non-citizen population than refusals, a finding that is consistent with the notion that the relationship is driven by the underlying relationship with deliberate refusals, which are only a subset of Type-A non-responses.

[62] The correlational controls are different from the discontinuity controls. The full list is as follows: avg_hh_members, log_faminc, avg_age, HS, BACH, female, unemployment, laborforce_part, white_and_other, black, hispanic_origin

Table 5: Regressions of Refusals & Type-A NR on Non-citizen & controls (1994-2002)

Dependent Variable	Non-Citizen Coefficient	Month FE	Long-Term Time Trend	Division FE	Dem. Controls
Refusals	0.103*** (0.00966)	Yes	Linear Time Trend	Yes	No
Refusals	0.108*** (0.00893)	Yes	Year FE	Yes	No
Refusals	0.0533*** (0.0111)	Yes	Year FE	Yes	Yes
Type-A NR	0.157*** (0.0141)	Yes	Linear Time Trend	Yes	No
Type-A NR	0.157*** (0.0138)	Yes	Year FE	Yes	No
Type-A NR	0.0740*** (0.0167)	Yes	Year FE	Yes	Yes

N = 5508 *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$
Standard Errors in Parentheses. Features PSAR corrections with heteroskedastic panels.

4) Policy Implications

The empirics thus show that (a) controlling on numerous factors, states with more non-citizens per capita have higher household response rates, (b) the 1994 redesign, of which the citizenship question was a component, caused a large, immediate, discrete increase in the household refusal rate, and (c) the magnitude of causal increase is related with non-citizenship, such that we estimate at least 20 basis points of the ~55 bp discrete jump in refusals is driven by the non-citizen populations. Thus, about 36% of the discrete change in refusals is potentially attributable to the citizenship question. The increase of 55 basis points corresponded to a roughly 22% overnight overall jump in refusals. If 36% of this jump is attributable to the citizenship question, this suggests conservatively that an 8% increase in refusals may be attributable to that question.[63]

By 2019, the refusal rate had become approximately four times higher than in the 1990's, having reached over 13%. Moreover, the size of the foreign born population (which includes both non-citizens, and also naturalized citizens) has continued to grow in both absolute size, and percentage of the American population.[64] These factors suggest that the total effect of the citizenship question may have grown substantially since 1994. If we take seriously the prospect that 8% of refusals are driven by the citizenship question, and apply it to the modern rate of 13%, it suggests that 1% of all households may be refusing due to the citizenship question specifically. This value is lower than but of similar magnitude

[63] $0.36 \times 0.22 = 0.0792 \approx 0.08$
[64] Lange, Jason, and Torbati Yeganeh. "U.S. Foreign-Born Population Swells to Highest in over a Century." *Reuters*. September 13, 2018.

to the 2% reduction in overall response found by Brown et al. for the projected effect of the same question on the Census.

While the potential for impact on total response has never been greater, the policy implications of these results are not straightforward. In adding the citizenship question, the Bureau of Labor Statistics has expressed an interest in understanding the labor force, educational, and other demographic characteristics of non-citizens. In order to distinguish this information, they must be able to identify individuals who are not citizens, hence the introduction of the question. However, the evidence discussed in this paper suggests that the introduction of sensitive immigration-related questions had a negative effect on the response rates of non-citizens.

The BLS also has a goal of understanding labor markets as accurately as possible, and particularly analyzing the entire United States labor market, including non-citizens and Hispanic individuals. If the citizenship question causes these groups to refuse the survey and participate less often, then they will be underrepresented, and there will be less information about their labor force characteristics. Ironically, while adding the citizenship question allows researchers to better identify information from non-citizens, they receive less overall information about them. Thus, as non-citizens and Hispanic people comprise a critical part of the labor force, the addition of the question means that information about the labor force is less accurate and vulnerable to potential biases.

The optimal policy therefore depends on the precise magnitude of the non-response bias, and the priorities of the Census Bureau and the Bureau of Labor Statistics in conducting the Current Population Survey. An important consideration is that this research shows that the question has caused an increase in survey refusals from particular marginalized groups, but does not prove labor market statistics are distorted. It is also known that these groups have distinctly different labor market characteristics, giving the *potential* for bias. However, we have not attempted to determine if labor market statistics are *actually* biased due to this decrease in non-response. Because the CPS is weighted to correct for non-response, it is possible that this procedure eliminates or drastically reduces bias in final labor market statistics.

If the objective is to have the most accurate and precise labor market statistics possible, then it seems to be best to eliminate the question. However, it might alternatively be judged that the information gained from the CPS about non-citizens' characteristics, labor market and otherwise, is highly valuable and not obtainable through other surveys such as the American Community Survey. Moreover, it might also be determined that the bias on final weighted statistics is negligible. In these cases, it may still be worthwhile for the citizenship question to remain part of the CPS. In light of these findings, however, we recommend that the Census Bureau take seriously the issue of non-response bias associated with the question. In particular, we would suggest a controlled trial be conducted to determine the exact causal effect of the question, evaluate the potential for bias in statistics, and make an informed cost-benefit analysis to determine if the citizenship question warrants continuation.

V. Conclusions & Extensions

1) Overview of Major Results

Reviewing the empirics, we find that they are consistent with decreased survey participation, in particular from Hispanic individuals and individuals who are not United States citizens, arising from the citizenship question. The regression discontinuity designs have provided strong statistical evidence showing that the 1994 redesign, including the citizenship question, caused a large, sharp, sustained, and statistically significant increase in refusals. They indicate that non-response increased immediately by approximately 20-50%, depending on specification. Additional RDD models have suggested that there

may have been a simultaneous reduction in the percentage of individuals in the survey reporting Hispanic ethnicity. We find this to be consistent with the theory that the redesign disproportionately affected Hispanics, likely due to the citizenship question, though the statistical significance of this last result is not robust.

In order to more directly attribute the discontinuity in refusals to non-citizenship in particular, we employ a two-stage RDD. In the first stage, we estimate individual intercepts for each state, which we then use as the dependent variable in the second stage. Second stage explanatory variables include non-citizen population, Hispanic population, and citizenship question adjustment rate. Consistently across specifications, the size of the jump discontinuity at January 1994 was positively related with the number of non-citizens in that state observed in January 1994. Despite small sample sizes, in many cases these results were (often highly) statistically significant. As noted before, we conservatively estimate that this relationship with non-citizen population appears to have driven approximately a 20 basis point discrete increase in refusals, corresponding to an 8% overall jump in the refusal rate.

Moreover, the state jump discontinuities were also positively related with state Hispanic population and citizenship question revision frequency, our proxy for citizenship question item non-response. In many specifications, these results were also statistically significant. All relationships were closely replicated when January 1995 state characteristics were used as well. These findings show that the discrete jump experienced by each state, causally attributable to the redesign due to the regression discontinuity design, is positively related to states' non-citizen population, Hispanic population, and citizenship question adjustment rate. The first two relationships show that the causal effect of the redesign appears to be driven by states' non-citizen and Hispanic populations; the third suggests that the citizenship question itself is a critical factor affecting overall non-response.

Lastly, we examined the correlational relationship between non-citizenship and refusal in the period after the redesign. While these regressions lacked a causal identification strategy, they illustrate the strong statistical relationship between non-citizenship rates and both refusals and Type-A non-response. While coefficients were reduced in magnitude as additional controls were added, there still remained a highly significant positive relationship between the variables. This relationship indicates that all else equal, states with higher levels of non-citizenship have higher levels of Type-A non-response and refusals. This suggests that the presence of non-citizens is driving the higher levels of non-response in those states. Notably, the point estimates for the effect of a change in non-citizen population on refusal rate (0.0533 to 0.108) closely mimicked the (significant) semi-causal point estimates (0.0609 to 0.129) from the state-specific discontinuity models. Thus, the results from these disparate approaches confirm each other.

Combined, these results paint a picture that draws a connection between the citizenship question and CPS survey refusals. The data show a strong correlational relationship between non-citizenship and non-response. In January 1994, when the citizenship question was added, there was a discrete increase in refusals, and a simultaneous decrease in the survey's reported Hispanic population. Moreover, the increases at the threshold are positively related with the portion of non-citizens, the portion of Hispanic individuals, and the rate of citizenship question revision in each state, suggesting a direct relationship between those factors. These results indicate that the redesign as a whole caused a large increase in refusals, and that the increase was positively associated with and driven by factors related to the citizenship question. Logically, it follows that the introduction of the citizenship question in the redesign likely had a direct contribution to the increase in CPS survey refusals.

2) Limitations & Extensions

One issue that we have encountered is the limits of the allocation flag for the citizenship variable. An allocation flag indicates how a variable was assigned to the value that appears in the microdata. However, during the entire sample period, this allocation flag never took on a value indicating that the question was refused, and is never missing for a responding individual. This would suggest that the question itself is never refused. Since it seems implausible that all individuals would provide this information in all cases, a worthwhile extension might involve further researching the exact methods used by surveyors in order to conduct personal interviews, elicit this information, and process missing values. Moreover, another aspect of analysis could focus on the accuracy of the information given. Indeed, many sources in the literature, such as Van Hook & Bachemier (2013) have documented potential misrepresentation in surveys. A potentially compelling future study might therefore estimate the rate of misrepresentation, and relate it to survey refusal, item non-response and other forms of survey nonparticipation.

As previously discussed, in order to make their labor market statistics more accurate, the Bureau of Labor Statistics uses weights to correct for measurable discrepancies between the survey sample and established national demographics. These weights are used to determine the final published labor force statistics, including the unemployment rate. It would be useful to see if the weighting used by the BLS is effective in diminishing some of the problematic findings that we have observed. However, non-responders do not receive weights, and the weighting system has evolved over the course of the sample period. This means the assignments of weights in a panel-data based research project such as this could prove to be arbitrary and impractical, which is why we chose not to take this route. However, in order to evaluate the accuracy of published labor force and demographic statistics in light of non-response and refusals, the demographic weights must be used and tested for their ability to correct for those issues. Future research could determine if non-response caused by the citizenship question biases labor market statistics after demographic weighting.

Lastly, this work has focused on the aggregated data. There are several advantages of working with the aggregated data, including a lack of missing observations (all states have at least some observations for all variables, except where a variable is not available for a certain year), improved code runtime and ease of interpretation. However, by working directly with the microdata, it may be possible to employ new econometric designs and obtain more precise empirical evidence of specific individuals refusing the survey. We think that future researchers can obtain more accurate causal estimates by employing the full versatility of the CPS microdata.

Acknowledgments

This paper has been shaped through the efforts of numerous individuals. We thank Ryan Moise, who suggested several important clarifications and revisions. Additionally, we thank Middlebury College faculty, David Munro and Erin Wolcott for introducing one of the authors to this version of the CPS microdata, and for providing substantial instruction in interpreting and reviewing this data. The usual caveats apply.

Bibliography

Bachmeier, James D., Jennifer Van Hook, and Frank D. Bean. 2014. "Can We Measure Immigrants' Legal Status? Lessons from Two U.S. Surveys." *The International Migration Review* 48 (2): 538–66. https://doi.org/10.1111/imre.12059.

Baum, Matthew A., Bryce J. Dietrich, Rebecca Goldstein, and Maya Sen. "Estimating the Effect of Asking About Citizenship on the U.S. Census." *Shorenstein Center Discussion Paper*, March 21, 2019. https://shorensteincenter.org/estimating-effect-asking-citizenship-u-s-census/.

Behagel, Luc., Bruno Crépon, Marc Gurgand, and Thomas Le Barbanchon, 2015. "Please Call Again: Correcting Nonresponse Bias in Treatment Effects Models." *The Review of Economics and Statistics*. December 2015, 97(5): 1070–1080.

Berman, Ari. "Civil Rights Groups Challenge Trump's 'Racially Discriminatory Scheme' to Skew Redistricting." Mother Jones, September 14, 2019. https://www.motherjones.com/politics/2019/09/citizenship-census-lawsuit/.

Berman, Ari. "Trump Backs Down on the Census Citizenship Question." Mother Jones (blog), July 11, 2019. https://www.motherjones.com/politics/2019/07/trump-backs-down-on-the-census-citizenship-question/.

Bernhardt, Robert. "Rising CPS Non-Response: A Potential Source of Bias?" presented at the Presentation to UIUC Ph.D. students, Champaign, IL, August 20, 2018.

Brown, J. David, Misty L. Heggeness, Suzanne M. Dorinski, Lawrence Warren, and Moises Yi. "Predicting the Effect of Adding a Citizenship Question to the 2020 Census." *Demography* 56, no. 4 (August 2019): 1173–94. https://doi.org/10.1007/s13524-019-00803-4.

Bureau of Labor Statistics Staff. "Briefing Materials on the Redesigned Current Population Survey." Bureau of Labor Statistics, January 1994. https://data.nber.org/cps-basic/brief.pdf.

Cohany, Sharon R, Anne E Polivka, and Jennifer M Rothgeb. "Revisions in the Current Population Survey Effective January 1994." *Employment and Earnings*, February 1994, 25.

Deding, Mette, Torben Fridberg, and Vibeke Jakobsen. 2008. "Non-Response in a Survey among Immigrants in Denmark." *Survey Research Methods* 2 (3): 107–21. https://doi.org/10.18148/srm/2008.v2i3.98.

Desouza, Kevin C, and Akshay Bhagwatwar, 2012. "Leveraging Technologies in Public Agencies." *Public Administration Review*. 72 (4): 605–614. https://doi.org/10.1111/j.1540-6210.2012.02592.x

"Design and Methodology: Current Population Survey." Technical Paper. Technical Paper 66. Washington, D.C.: U.S. Bureau of Labor Statistics, October 2006.

Heffetz, Ori, and Daniel B Reeves. 2018. “Difficulty to Reach Respondents and Nonresponse Bias: Evidence from Large Government Surveys.” *NBER Working Paper Series*, January 2018, 44.

Hokayem, C., Bollinger, C. & Ziliak, J., 2015. “The Role of CPS Nonresponse on the Level and Trend in Poverty.” *University of Kentucky Center for Poverty Research Discussion Paper Series*, DP2014-05. Retrieved from http://www.ukcpr.org/Publications/DP2014-05.pdf.

Krueger, Alan B., Alexandre Mas, and Xiaotong Niu. “The Evolution of Rotation Group Bias: Will the Real Unemployment Rate Please Stand Up?” *The Review of Economics and Statistics* 99, no. 2 (May 2017): 258–64. https://doi.org/10.1162/REST_a_00630.

Lange, Jason, and Torbati Yeganeh. “U.S. Foreign-Born Population Swells to Highest in over a Century.” *Reuters*. September 13, 2018. https://www.reuters.com/article/us-usa-immigration-data-idUSKCN1LT2HZ.

National Research Council. (2013). *Nonresponse in Social Science Surveys: A Research Agenda*. Roger Tourangeau and Thomas J. Plewes, Editors. Panel on a Research Agenda for the Future of Social Science Data Collection, Committee on National Statistics. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Narea, Nicole. “Trump Is Still Trying to Collect Citizenship Data for Redistricting.” Vox, October 17, 2019. https://www.vox.com/policy-and-politics/2019/10/17/20918989/trump-2020-census-citizenship-data-redistricting-drivers-license.

Passel, Jeffrey S, and D’Vera Cohn. “A Portrait of Unauthorized Immigrants in the United States.” PewResearchCenter Project. Pew Hispanic Center, April 14, 2009.

Pedraza, Franciso I., Vanessa Cruz Nichols, and Alana M. W. LeBron. “Cautious Citizenship: The Deterring Effect of Immigration Issue.” *Journal of Health Politics, Policy and Law* 42, no. 5 (October 2017): 925–60.

Van Hook, Jennifer, and James D. Bachmeier. “Citizenship Reporting in the American Community Survey.” *Demographic Research* 29 (July 2, 2013): 1–32. https://doi.org/10.4054/DemRes.2013.29.1.

Wang, Hansi Lo. “Latinx Advocacy Groups Sue To Block Citizenship Data Release By Trump Officials.” NPR.org. Accessed September 17, 2019. https://www.npr.org/2019/09/13/760356414/latinx-advocacy-groups-sue-to-block-citizenship-data-release-by-trump-officials.

Wang, Hansi Lo. “Skipping The 2020 Census Citizenship Question? You’ll Still Be Counted.” NPR.org. Accessed October 26, 2019. https://www.npr.org/2018/04/19/603629576/skipping-the-2020-census-citizenship-question-youll-still-be-counted.

Wang, Hansi Lo. “Why Is The Census Bureau Still Asking A Citizenship Question On Forms?” NPR.org, August 9, 2019. https://www.npr.org/2019/08/09/743296249/why-is-the-census-bureau-still-asking-a-citizenship-question-on-forms.

Appendix A

1) Notes on the Current Population Survey

The Current Population Survey is described in its official technical documentation (Technical Paper 66) as “immensely important,” and one of the “oldest, largest, and most well-recognized surveys in the United States.” The CPS provides the data for the official unemployment rate, as well as valuable employment, earnings, industry, educational, and demographic data. While the CPS has been updated many times since its introduction in 1940, many of the key components and emphases have remained the same.

The survey is conducted on a monthly basis, and has always asked questions on labor force status and basic demographic information for the persons in approximately 60,000 households. The exact sample size has undergone some evolution as funding has been both increased and withdrawn at certain points. The survey has an 8-month rotating design, where each month, ~7500 new households are introduced to the survey, and ~7500 will complete their final interview. The survey uses a 4-8-4 interview pattern, where a household is interviewed for 4 consecutive months, not interviewed for 8 months, and then interviewed again 4 more times. Since the primary household respondent, known as the reference person, will provide information for all individuals in the household, typically 130,000-160,000 people and non-responding households are represented in the sample for a given survey month.

2) Variable Names, Definitions & Microdata Originations

(All variables apply to a particular time period, and a particular state. National level variables, where applicable, are defined and generated identically, but aggregated at the national level. Additional variables are created in the analysis at various points to represent fixed effects, interaction terms, and regression coefficient estimates.)

Name in Stata Code	Definition	Microdata Position (1989 - 1993) Microdata Position (1994 - 2002)
gestfips	FIPS State Code. Unique State Identifier	HG-FIPS (118 - 119) GESTFIPS (93 - 94)
non_response	Total Household Non-Response Rate	H$TYPINT (18 - 18) HRINTSTA (57 - 58)
typeA	Household Type-A Non-Response Rate	H-HHTYPE (69 - 69) HRINTSTA (57 - 58)
refusal	Household Refusal rate (Type-A Category 3)	H-AREASN (49 - 49) HUTYPEA (41 - 42)

avg_hh_members	Average Number of Individuals in Responding Households	Generated from household identifiers
faminc	Average total Family Income from all sources (past 12 months)	H$FAMINC (13 - 14) HUFAMINC (39 - 40)
date	Survey Month Date	Generated from Stata Macros
mdate	Survey Month Date (optimized for time series functionality)	Generated from date
avg_age	Average Age of Responding Individuals	A$AGEDG1 & A$AGEDG2 (169 - 170) PEEDUCA (137 - 138)
HS	Portion of individuals with High School Diploma or Equivalent	M-HGA (277 - 278) PEEDUCA (137 - 138)
grade12	Portion of individuals who have completed grade 12	M-HGA (277 - 278) PEEDUCA (137 - 138)
BACH	Portion of individuals with a Bachelor's Degree or higher	M-HGA (277 - 278) PEEDUCA (137 - 138)
ACA_Assoc	Portion of individuals with an Academic Associates Degree, some College or higher	M-HGA (277 - 278) PEEDUCA (137 - 138)
female	Portion of individuals who are female	A$SEX (175 - 175) PESEX (129 - 130)
unemployment	Portion of unemployed individuals in labor force	A-LFSR (348 - 348) PEMLR (180 - 181)
laborforce_part	Portion of eligible individuals participating in the labor force	A-LFSR (348 - 348) PEMLR (180 - 181)
lfq_nr	Labor force question item non-response rate among eligible individuals	A-LFSR (348 - 348) PEMLR (180 - 181)
white	Portion of individuals who are white	A$RACE (180 - 180) PERACE (139 - 140)
white_and_other	Portion of individuals listing race as white or "other"	A$RACE (180 - 180) PERACE (139 - 140)
black	Portion of individuals who are black	A$RACE (180 - 180) PERACE (139 - 140)
hispanic_origin	Portion of individuals of Hispanic	A-REORGN (344 - 345)

	origin	PRORIGIN (141 - 142)
hispanic_origin_nr	Hispanic origin question item non-response rate	A-REORGN (344 - 345) PRORIGIN (141 - 142)
hispanic	Portion of individuals who are Hispanic	Not in pre-1994 data PRHSPNON (157 - 158)
entry_year	Average immigrant year of entry among foreign-born individuals	Not in pre-1994 data PEINUSYR (176 - 177)
non_citizen	Portion of individuals who are not United States citizens	Not in pre-1994 data PRCITSHP (172 - 173)
nc_nr	Citizenship question rate of missing/left blank	Not in pre-1994 data PRCITSHP (172 - 173)
nc_revision	Citizenship question revision rate	Not in pre-1994 data PRCITFLG (174 - 175)
division	Census Division	Generated from gestfips
critical_state1	Equals 1 if state is top-10 in average Hispanic population or average non-citizen population	Generated from hispanic_origin & non_citizen
critical_state2	Equals 1 if state is top-20 in average Hispanic population or average non-citizen population	Generated from hispanic_origin & non_citizen
month_	Month (1 = Jan, 12 = Dec)	Generated from mdate
year_	Year	Generated from mdate
log_faminc	Natural log of faminc	Generated from faminc
post	Redesign Threshold Indicator (= 1 if occurring after Jan 1994)	Generated from date
state_number	State identifier, revised to include only consecutive numbers	Generated from gestfips

3) Extended Notes on Critical States

A state has a value of 1 for critical_state1 if it ranks in the top 10 for the largest Hispanic population proportion, averaged across months, or in the top 10 for largest non-citizen population, also averaged across months, and 0 otherwise. The same criteria are used for critical_state2, except a state need only rank in the top 20 to receive a value of 1. Due to high overlap between states with high Hispanic and non-citizen populations, 13 states are “critical 1” states, and 23 are “critical 2” states. These categories are based on the average over the entire applicable sample, and do not change over time: a state is considered a critical state for all periods, or a critical state is considered not a critical state for all periods.

4) Detailed Summary Statistics

Table 0.2

Definition	Variable	Sample Size	Mean	Standard Deviation	Minimum Value	Maximum Value
FIPS State Identifier	gestfips	8568	28.96078	15.67775	1	56
Total Non-Response	non_response	8568	0.213435	0.046162	0.088542	0.389873
Type-A Non-Response Rate	typeA	8568	0.04676	0.020211	0.002535	0.152821
Survey Refusals	refusal	8568	0.027231	0.011955	0.001267	0.08961
Average Household Members	avg_hh_members	8568	2.563158	0.152287	1.887218	3.250916
Average Family Income	faminc	8568	39557.73	9529.812	18674.36	71473.77
Survey Month	date	8568	199556.5	403.1512	198901	200212
Survey Month (time series formatted)	mdate	8568	431.5	48.49939	348	515
Average State Age	avg_age	8568	35.30712	1.800807	28.27547	40.74757
High School Graduate Rate	HS	8568	0.783647	0.044944	0.626087	0.886232
Grade 12 Completion Rate	grade12	8568	0.794084	0.04502	0.626087	0.886232
Bachelor's Degree Attainment Rate	BACH	8568	0.205941	0.046159	0.078853	0.429872
Academic Associates Degree/Some College Rate	ACA_Assoc	8568	0.228912	0.051884	0.087455	0.446266
Portion Female	female	8568	0.517648	0.012376	0.46646	0.580827
Unemployment Rate	unemployment	8568	0.052316	0.016899	0.010448	0.152742
Labor Force Participation Rate	laborforce_part	8568	0.661002	0.041458	0.498661	0.771972
Labor Force Question Non-	lfq_nr	8568	0.005514	0.006589	0	0.05174

Response						
Portion White	white	8568	0.842958	0.140682	0.220412	0.996875
Portion White or Other	white_and_other	8568	0.845845	0.140275	0.220412	0.998399
Portion Black	black	8568	0.105245	0.120373	0	0.715385
Portion of Hispanic Origin	hispanic_origin	8568	0.057103	0.080394	0	0.426352
Hispanic Origin Question Non-Response	hispanic_origin_nr	3060	0.001762	0.001522	0	0.018723
Portion Hispanic	hispanic	5508	0.063193	0.084284	0	0.426352
Portion Non-Citizens	non_citizen	5508	0.0381	0.03515	0	0.210097
Average Immigrant Entry Year	entry_year	5508	1980.257	4.181538	1962.949	1990.652
Citizenship Question Non-Response	nc_nr	5508	0	0	0	0
Citizenship Question Revision	nc_revision	5508	0.007547	0.006581	0	0.062234
Census Bureau Division	division	8568	5.117647	2.510104	1	9
Critical State 1 Designation	critical_state1	8568	0.254902	0.435832	0	1
Critical State 2 Designation	critical_state2	8568	0.45098	0.49762	0	1
Month	month_	8568	6.5	3.452254	1	12
Year	year_	8568	1995.5	4.031364	1989	2002
Natural Log of Family Income	log_faminc	8568	10.55669	0.240785	9.834907	11.17709
Threshold Indicator	post	8568	0.642857	0.479185	0	1
Alternative State Identifier	state_number	8568	26	14.72046	1	51

5) Second Stage Regression Discontinuity Results with January 1995 State Characteristics

Table 4.2: Regression Results for State Discontinuities & January 1995 Characteristics

Specification	Regression with Non-Citizen Population	Regression with Hispanic Population	Regression with Citizenship Q. Revision
State Baselines & Discontinuities, No Controls	0.0223	0.000563	0.221*
State Baselines & Discontinuities, Controls	0.179	-0.000575	0.188
Division Baselines, State Discontinuities, No Controls	0.0845***	0.0258**	0.462**
Division Baselines, State Discontinuities, Controls	0.0574***	0.0216**	0.279*
Constant Baseline, State Discontinuities, No Controls	0.116***	0.0304**	0.895***
Constant Baseline, State Discontinuities, Controls	0.0751***	0.0290***	0.487**

N = 51 *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Coefficient values derived from second stage regression of discontinuities and each state characteristic in January 1994. Statistical significance determined by a *1-tailed* hypothesis test on the regression $Discontinuity_i = \beta_0 + \beta_1 Characteristic_i + \varepsilon_i$, testing the null $H_0: \beta_1 \leq 0$.

6) Data Source

NBER: Current Population Survey Basic Monthly Data (https://www.nber.org/data/cps_basic.html Microdata documentation is available at the same website).