NBER WORKING PAPER SERIES

HEALTH, EMPLOYMENT, AND DISABILITY:
IMPLICATIONS FROM THE UNDOCUMENTED POPULATION

George J. Borjas
David J.G. Slusky

Working Paper 24504
http://www.nber.org/papers/w24504

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2018, Revised April 2021

We are grateful to Dayanand Manoli, Delia Furtado, Ninez Ponce, Kirk Doran, Heather Koball, Catia Nicodemo, Nicolas Ziebarth, conference participants at the 2017 International Health Economics Association, the 2017 Association for Public Policy Analysis & Management, and the 2018 III Workshop on Migration, Health and Wellbeing meetings, and seminar participants at West Virginia University, Cornell University, the University of Kansas, the Federal Reserve Bank of Kansas City, and the University of Delaware for their help and comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Health, Employment, and Disability: Implications from the Undocumented Population
George J. Borjas and David J.G. Slusky
NBER Working Paper No. 24504
April 2018, Revised April 2021
JEL No. I12,I18,J61

ABSTRACT

The number of disability beneficiaries doubled in the past two decades. It is difficult to determine how much is explained by changes in health, as we lack a counterfactual. We use undocumented immigrants to form the counterfactual, as they cannot claim benefits. Using NHIS data, we show that the relationship between health and disability is stronger for the legal population than for the undocumented. Much of the difference in disability rates between the populations is due to different labor supply responses to underlying health impairments and demographic differences, rather than to differences in the impairments or demographic variables themselves.

George J. Borjas
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138
and NBER
gborjas@harvard.edu

David J.G. Slusky
Department of Economics
University of Kansas
1460 Jayhawk Boulevard
Lawrence, KS 66045
and NBER
david.slusky@ku.edu

Health, Employment, and Disability:
Implications from the Undocumented Population

I. Introduction

Disability beneficiaries have nearly doubled in the past two decades (Social Security 2017a), even though the size of the working age (16+) population only increased by 25 percent and the size of the population aged 55+ increased by 67 percent (Bureau of Labor Statistics 2017). There are two explanations for the sizable increase in the size of the disability rolls (Autor and Duggan, 2003; Autor and Duggan, 2006; Duggan and Imberman 2009; Liebman 2015): 1) it is the product of both an aging population and decreasing overall health (i.e., a change in observable covariates); and/or 2) it is the result of lowering the minimum threshold of health limitations required for individuals to claim and be awarded disability benefits (i.e., a change in the coefficients applied to those covariates). The latter hypothesis, of course, encompasses both the increased use of the program by those who are somewhat disabled but still able to engage in productive employment,[1] as well as overuse of the program by the nondisabled. There are other explanations, including rising inequality and lower earnings opportunities among less skilled workers (Autor and Duggan, 2003; Liebman, 2015) which, while relevant, are less addressable with the data and approach of this paper.

To distinguish between these two hypotheses, we need to establish what the disability rolls would have looked like in a counterfactual world. This counterfactual scenario would help document what those persons who now receive disability benefits would have done had the disability program not been an option. Would they still be unable to work due to their poor

[1] The long history of investigation into moral hazard in the disability program goes back at least as far Parsons (1980; 1982).

health? Or would the lack of disability benefits persuade them to take a job despite their physical limitations?

In this paper, we propose a novel technique to distinguish between the two possibilities. In particular, we use the foreign-born undocumented population residing in the United States to create a counterfactual sample of physically disabled persons who, by law, do not qualify for disability benefits.[2]

The Department of Homeland Security (DHS) estimates that 12.1 million undocumented persons lived in the United States in January 2012 (DHS 2018).[3] These individuals reside in many of the same labor markets as the persons who have legal status (including, of course, the native-born, "green card" holders, and naturalized citizens), yet they are unable to claim public disability benefits. The sample of undocumented persons allows us to observe if a person with specific health limitations works in the absence of social insurance programs. We can then use the behavior of the undocumented to establish if the "exodus" of persons from the labor force to the disability rolls was the result of decreasing health in the population or of the lowering requirements needed to qualify for disability benefits.

In addition to providing a new way of examining the longstanding question of why the disability rolls have increased dramatically, our analysis also provides the first credible documentation of the health status of the undocumented population. Past research on immigrants

[2] We make the comparison using both a broad sample of all Americans of working age, and also a narrow sample of only Hispanic, non-veteran, high school dropouts.

[3] The DHS summarizes its approach as: "Two populations are estimated in order to derive the unauthorized population estimates: 1) the total foreign-born population living in the United States on January 1, 2014, and 2) the legally resident foreign-born population on the same date. The unauthorized population estimate is the residual when (2) is subtracted from (1).…Data on the foreign-born population…were obtained from the 2013 ACS [American Community Survey]….Data on persons who obtained LPR [Legal Permanent Resident] status…were obtained from DHS administrative records." (DHS 2018)

(which typically include both legal immigrants as well as the undocumented) concludes that they tend to have lower disability rates and use fewer disability services than natives (Benjamin et al. 2000), but are more likely to receive disability payments when they live near others of their ethnic group who have higher take-up rates (Furtado and Theodoropoulos, 2016). The existing research has not examined the difference in disability rates between documented and undocumented immigrants because of the inherent difficulties associated with identifying undocumented status in microdata.

In recent years, however, there has been progress in developing methods that impute the undocumented status of foreign-born persons in micro data sets, such as the Current Population Surveys. These attempts build on the "residual method" first developed by Warren and Passel (1987), and since adopted by the Department of Homeland Security, to estimate the size of the undocumented population. In particular, Passel and Cohn (2014) develop an algorithm that identifies foreign-born persons in the micro surveys who are likely to be legal immigrants (e.g., naturalized citizens, refugees, persons who are married to either citizens or permanent residents, etc.), and define the residual group as "likely undocumented." Borjas (2017a) applied this algorithm to examine differences in labor supply among the various populations in the post-1994 CPS files that contain the requisite background information for foreign-born persons.

Much of the existing literature on the health and disability of the immigrant population (Akbulut-Yuksel and Kugler, 2016; Giuntella and Stella, 2017) does not differentiate between the legal and undocumented groups. For example, Xiang et al. (2010) find that immigrants with disabilities are more often employed than the native-born, without investigating whether there's a disparity between legal and undocumented immigrants.

A handful of papers do explore the difference. Goldman, Smith, and Sood (2006) use an algorithm where noncitizen foreign-born survey respondents who did not reply affirmatively to having at least a permanent resident card, a green card, or a document allowing them to stay in the U.S. for a limited time were classified as “undocumented”, and find that undocumented immigrants use substantially less health care. This analysis, however, uses the 2000 Los Angeles Family and Neighbor Survey, which although having detailed information on respondents’ legal and visa status, covers only one city and has a relatively small sample size. Giuntella and Lonsky (2020) use the arbitrary eligibility rules for the 2012 Deferred Action for Childhood Arrivals (DACA) to study its impact on the health of eligible undocumented immigrants, finding that DACA increased health insurance coverage but did not have a statistically significant impact on health care utilization. Giuntella et al. (2021) also find improvements in immigrants’ sleep from DACA.[4]

The other few papers in the literature use variations of the Passel-Cohn residual method, albeit with fewer variables and reasons for excluding a foreign-born person from the undocumented population. Stimpson, Wilson, and Su (2013) use matched National Health Interview Survey (NHIS)-Medical Expenditure Panel Survey data to study the per-capita health spending of undocumented immigrants, and find that it is an order of magnitude smaller than that of the native born. Similarly, Pourat et al. (2014), use the 2009 California Health Interview Survey (CHIS) to examine health care consumption among undocumented immigrants, and find that undocumented immigrants consume substantially less health care than either natives or legal immigrants. Finally, Cohen and Schpero (2018) use the American Community Survey (ACS) to

[4] As DACA was first promulgated in late 2012, we are not concerned about it the policy shift affecting our results, which are obtained mostly from before 2012 and consistent when focusing on the years before DACA.

study the impact of the Affordable Care Act's Medicaid Expansion on undocumented immigrants. None of these studies, however, examine the propensity of being disabled (as measured by an inability to work for health-related reasons) in the undocumented immigrant population.

Additionally, none of the existing studies introduces the perspective of viewing the undocumented immigrant population as a counterfactual for the legal immigrant and native-born populations (i.e., the population eligible for benefits, hereafter "eligibles"). This is a key contribution of our study and distinguishes it from other studies that use administrative sources of variation like judges and examiners (e.g, von Wachter, Song, and Manchester 2011; Maestas, Mullen, and Strand 2013; French and Song 2014). It also enables us to avoid the obvious and well-documented issues of selection (including based on health) of who migrates to the U.S., because we are focusing only on the individuals who are already in the country and not comparing them to those who did not migrate.

More broadly, our paper, therefore, is part of the methodological approach started by Bound (1989), which used disability benefit applicants that failed to pass the medical screening as a control group. Many other studies have exploited other variation in receipt of benefits: (Gruber and Kubik 1997; Kostøl and Mogstad 2015; Mullen and Staubli 2016; Autor et al. 2019; Low and Pistaferri 2019), interactions with other welfare programs (Low and Pistaferri 2015), variation in benefits generosity (Gruber 2000; Campolieti 2004; Kostøl and Mogstad 2014; Gelber, Moore, and Strand 2017; Milligan and Schirle 2019), variations in benefit durations of other programs, such as unemployment insurance (Mueller, Rothstein, and von Wachter 2016), ease of application (Foote, Grosz, and Rennane 2018), and macroeconomic variation (Black, Daniel, and Sanders 2002; Maestas, Mullen, and Strand 2015; Jiménez-Martín, Juanmarti, and Vall 2018; Roberts and Taylor 2019)

We extend the literature by applying the residual method of identifying undocumented status to the National Health Insurance Survey (NHIS) and address three related issues: 1) we compare the health and disability status (i.e., being out of work due to health or disability) of undocumented immigrants to the eligible population; (2) we exploit the available information on disability, employment, and health to determine what share of disabled workers would actually be employed if the disability benefits were not available; and (3) we estimate the cost to the disability program of an "amnesty" that would regularize the status of undocumented immigrants and give them full access to disability benefits. These latter two questions, while seemingly only tangentially related, are actually the empirical converses of each other, and together provide substantial new insight into the relationship between disability benefits and work over the past two decades.

II. Conceptual Framework

It is instructive to begin by outlining a simple conceptual framework that illustrates how those eligible and ineligible for benefits might have a different mapping from health conditions to work-preventing disability. Specifically, consider the labor supply decision faced by an individual with a generic standard utility function. The individual faces a binary decision: work, or stay out of work due to health limitations. An improvement in the health of an individual (assumed to be exogenous) has two effects: it raises the individual's market wage and it reduces the probability an individual (if eligible) will receive disability benefits if not working. An individual will choose to work if the additional utility from wage income over expected disability benefits is greater than the lost utility from consuming less leisure.

Working becomes more likely the healthier the individual is, as wages increase and the probability of receiving disability benefits (and therefore the expected disability benefits) falls as health rises. In contrast, an unhealthy and eligible individual will likely not work because the available market wage is low, the expected benefits due to disability are high, and not working allows more time for leisure. In short, there will be a strong relationship between health and work.

Now imagine an individual who is ineligible for disability benefits. This individual will also work if the utility of doing so is greater than the utility of not working, but an ineligible individual will receive zero disability benefits. There will still be an extremely low level of health such that the individual does not work, as the available wage is so low that any utility from it is outweighed by increased utility from additional leisure time. At levels of health above this minimal threshold, however, the individual is much more likely to work, since without the possibility of disability benefits even a small wage may outweigh the increased utility from more leisure. Overall, the relationship between health and work will be much weaker, and substantially different from the health-work locus in the eligible population.[5]

III. Data and methods

We use publicly available microdata from the National Health Interview Survey for the post-1997 period. The NHIS is an annual, bilingual (English and Spanish), repeated cross-

[5] Please see Appendix D for a more detailed conceptual framework with accompanying equations.

section, household-level survey of about 40,000 households, containing 100,000 individuals per year. For most households, a sample adult and a sample child are interviewed in greater depth, and the questions asked for this subsample contain the information needed to determine both immigration status (through the "residual" imputation procedure described below) and specific health conditions. These sample adults and children also report scaled-up survey weights so that they can be used to produce nationally representative estimates of the entire population.[6] It is worth noting that the NHIS samples are sufficiently large to allow a statistically reliable estimate of the undocumented population.

Our analysis of the link between health conditions and disability status (as measured by an inability to work for health-related reasons) focuses on a set of specific health problems: heart disease, cancer, diabetes, hypertension, asthma, emphysema, liver disease, joint pain, back pain, neck pain, face pain, ulcers, and bronchitis. We focus on this subset because these health impairments are used by the Social Security Administration to determine whether an individual is disabled (Social Security 2017b). Additionally, NHIS has a variable for each condition corresponding to a question beginning, "Have you EVER been told by a doctor or other health professional that you had…"[7] One important caveat is that all the health diagnoses in the NHIS microdata are self-reported, and self-reported health issues may not be unbiased measures of the

[6] The NHIS adjusts for nonresponders and undersampling. See CDC (2014).

[7] This is essential because disability must be documented by medical evidence. However, the NHIS does also have variables for a wider range of whether a "condition or health problem causes you to have difficulty with" comment mental and physical tasks. While these variables may bias our results as they incorporate consequences of the conditions (in addition to just the presence of them) into our independent variables, we nevertheless in Table B-14 we incorporate these variables into our analysis, and find consistent results. NHIS also does have mental health variables (e.g., bipolar disorder, autism) as diagnosed by a provider, but only in 2007 (and very sparsely in 2012). In Table B-15, we also repeat our analysis using these variables, and also find consistent results where a majority of the difference is due to coefficients and not endowments.

actual underlying health conditions (Johnston, Propper, and Shields 2009). While using a data set such as the National Health and Nutrition Survey (NHANES), which provides objective measures of health status, would correct for the self-reporting bias, the NHANES lacks the variables that are necessary to identify undocumented immigrants. In addition, the smaller sample size in the NHANES would make it nearly impossible to conduct our empirical analysis. (Appendix B addresses the potential concern raised by the self-reporting of health issues by conducting several robustness checks, including incorporating self-assessed variables of functional limitations, as opposed to those diagnosed by a healthcare provider, and only using the subsample of respondents who had seen a physician in the past year. In both cases we find comparable results).

Our measure of a person's disability status is based on the NHIS variable that reports information for why an individual did not work in the week before the interview. While the specific response categories are not entirely consistent over the survey years, our initial strategy is to classify a person as disabled if he or she lists one of the following as the main reason for not working in the reference week: "unable to work for health reasons", "temporarily unable to work for health reasons",[8] or "disabled". We use this variable to define disability status, instead of the variables for receipt of disability benefits, because undocumented immigrants do not qualify for such benefits. We will instead use the benefit information as part of the algorithm that helps to differentiate legal immigrants from undocumented immigrants.

Our imputation of undocumented status applies the methods developed by Passel and Cohn (2014), as adapted by Borjas (2017a) and Borjas and Cassidy (2019) to the 1994-2015

[8] Given that SSDI eligibility requires a permanent disability, we alternatively define disability to be only those "unable to work for health reasons" or "disabled" and show in the Appendix in Table B-9 that our results are robust.

Current Population Surveys. In rough terms, we use a set of characteristics that suggest that a foreign-born person in the survey is likely to be a legal immigrant. Such “signals” include whether the person works in an occupation that requires licensing, whether the person receives specific types of public assistance, or whether the person has a family member (who in our data must also live in the same household) that grants them legal status (e.g., married to a US citizen). The residual sample of foreign-born persons then composes the sample of undocumented immigrants.

The NHIS was substantially redesigned in 1997, so that our empirical analysis uses only the data drawn from the post-1997 surveys. In addition, two of the annual surveys lack some of the information required to impute undocumented status at the micro level. In particular, the 1997 survey does not report if the person is a naturalized citizen, and the 2004 lacks a variable reporting a person’s Hispanic ethnicity, which is necessary to identify immigrants from Cuba (who are all legal because they are typically admitted as refugees).[9] As a result, our analysis uses the 1998-2003 and 2005-2015 NHIS cross-sections.

For illustrative purposes, we can use the self-reported measures for the various medical conditions in the NHIS to construct a variable that summarizes the overall health status of the undocumented and the eligible populations. In particular, we aggregate across the various medical conditions by using a modified Charlson Index (Charlson et al. 1987), which is essentially a weighted sum across conditions.[10]

We then estimate a generic regression model (separately by eligibility, pooling the native born and legal immigrants) that relates the probability that a person is disabled (as defined by

---

[9] We unfortunately lack broader information on country of origin and so cannot incorporate relevant information like pre-immigration smoking rates (as in Christopoulou and Lillard 2015).

[10] Please see Appendix E for more details.

whether he or she did not work in the past week due to health-related reasons) to self-reported medical conditions and various socioeconomic characteristics. The model is given by:

$$(\Pr y_{iaeqy} = 1) = F(\alpha + \boldsymbol{\gamma}\mathbf{D}_{iaeqy} + \mathbf{age}_a + \mathbf{education}_e + \mathbf{quarter}_q + \mathbf{year}_y + gender_i + \varepsilon_{iaeqy})$$

where *y* is a dummy variable indicating if individual *i*, in age bracket *a*, with educational attainment *e*, surveyed in year *y* and quarter *q*, is disabled. The term α is constant and in the linear model corresponds to a common intercept. The vector **D** contains dummy variables giving the medical conditions used by the Social Security Administration to evaluate being disabled: heart disease, cancer, diabetes, hypertension, asthma, emphysema, liver disease, joint pain, back pain, neck pain, face pain, ulcer, and bronchitis (Social Security 2017b). As described above, the variables for these from the NHIS are for physician diagnosed conditions. Finally, the **age, education, quarter,** and **year** variables are vectors of fixed effects for 10-year age brackets, educational attainment brackets, survey quarter, survey year, and gender, respectively.[11]

It is important to note that the educational attainment variables may be measuring different quality of education for immigrants and non-immigrants. This is a limitation of the entire literature and is not unique to our paper. Appendix B includes a robustness check that only

[11] Ideally, we would include state-level controls, including fixed effects and the time varying presence and generosity of relevant public programs (e.g., Secure Communities, E-Verify, expansions of drivers' licenses and health insurance). Unfortunately, the publicly available NHIS microdata does not contain state identifiers. We attempted to apply to the National Center for Health Statistics (NCHS) at the CDC for access to the restricted version of the data which does contain these variables. We were denied, and told by email that:

> We do not allow projects that try to infer anything about legal or documented status. We do not collect data on documentation or legal status. It is inappropriate to use the data that is collected to make inferences about status. We do allow comparisons of immigrants vs non-immigrants or other distinctions based on what NCHS surveys actually collect. You should remove any language that suggests legal status.

uses persons with less than a high school diploma (which, for most immigrants would have been obtained prior to migration), which should mitigate much of this concern.

We limit much of the empirical analysis reported below to persons aged 18–64. There are extremely few individuals aged 64+ in the NHIS sample that our algorithm identifies as undocumented, and therefore we lack the statistical power to draw robust conclusions for the elderly sample. Second, substantial government benefits (i.e., Medicare and Social Security) phase in for the vast majority of legal immigrants at age 65. This would exacerbate differences between the two groups in reporting being disabled as there is a substantial break in the types of benefits available to the two elderly groups.

To summarize the implications of the two regression models we perform an Oaxaca-Blinder decomposition (Oaxaca 1973; Blinder 1973). This exercise decomposes differences in an outcome between two groups into what can be explained by differences in the levels of a set of common covariates as opposed to differences in the coefficients on those covariates.[12] It complements the decomposition in Liebman (2015), which does not make use of a contemporaneously existing ineligible population.

An equally interesting application of our regression models is to use the regression model for one group to predict the trend in the disability rate of the other group. In other words, what would the secular trend in the disability rate of the eligible population look like if they responded to medical conditions in the same way as observationally equivalent undocumented immigrants? Or what would be the trend in the disability rate of undocumented workers if they responded to adverse medical conditions in the same way as observationally equivalent eligible individuals?

[12] See Appendix C for the mathematical details in the linear case. The non-linear decomposition follows Yun (2004).

This counterfactual exercise helps us address the two crucial questions posed in this paper: 1) How much would the reported disability rate drop if the native born and immigrants with legal status could not claim benefits? and 2) How much would the reported disability rate of undocumented persons rise if they could claim benefits?

IV. Results

Table 1 reports the number of observations affected by each subsequent restriction used to classify foreign-born persons into the two groups of legal and undocumented immigrants. Out of the 1.6 million observations in the pooled NHIS Sample Adult and Sample Child files over the years used in our study, 1.3 million are native born and another 100,000 are naturalized citizens. A sizable number of the remaining non-citizens receive government benefits (which are typically available only to legal immigrant)[13], or are married to US citizens, or are the children or grandchildren of someone with legal status.[14] Because of the family preference system that regulates U.S. immigration policy since 1965, these family connections imply that the NHIS respondent will likely be a legal immigrant. After imposing all the restrictions used by the imputation method, we are left with a population estimate of 12.7 million undocumented persons in the typical sample year of the NHIS (or roughly about 6,100 observations per year).

Figure 1 contrasts our estimates of the number of undocumented immigrants (i.e., the sum of the survey weights) with the official DHS estimates and estimates created from the CPS (through the same algorithm). Although the three estimates are reasonably close to each other,

[13] A person is considered to be a legal immigrant if he or she receives any of the following benefits: Social Security (including from Social Security Disability Insurance), Supplemental Security Income, Medicaid, Medicare, or military health insurance, welfare, public housing, or TANF.

[14] Note that the converse is not assumed; we do not assume that the parent or grandparent of someone with legal status has such status.

follow the same upward trend in the 2000-2007 period, and are all roughly constant in the 2007-2011 period,[15] it is notable that the imputation method in the NHIS leads to about 1 million more undocumented persons in any given year than the DHS estimates. Using the CPS, Passel and Cohn (2014, p. 48) report a similar tendency for the imputation method to "overcount" the number of undocumented persons. They then use a "probabilistic method" to correct for the overcount and reweigh the sample so that the weighted number of undocumented immigrants is, by construction, exactly equal to the DHS official statistic. To make our analysis transparent and fully reproducible, we do not make any adjustments to the sample weights in the NHIS and simply note that the trends illustrated in Figure 1 suggest that the sums of survey weights for the persons that we impute to be undocumented seem to correctly summarize key trends in the undocumented population.

Table 2 reports summary statistics for many of the variables used in our empirical analysis. The first row of the table reports the fraction of persons in each of the groups that is "disabled," as indicated by whether the person did not work in the past week due to health reasons. Note that very few undocumented persons (only 1.4 percent) report a health-related reason for idleness, as compared to 4.5 percent of legal immigrants and 7.4 percent of the native-born.

It is also evident that undocumented immigrants self-report themselves to be far healthier than eligible individuals. In particular, they are less likely to suffer from any of the dozen medical conditions that we use in our analysis. The probability that an undocumented immigrant suffers from any of the dozen ailments is only 25.2 percent, as compared to 40.9 percent for a

[15] The correlation between the 10 DHS January 1 observations and the corresponding NHIS estimates (averaged across two surveys to correspond to January 1) is 0.85.

legal immigrant and 53.5 percent for a native-born person. Undocumented immigrants are also 5 years younger and have far less education: 45.2 percent of the undocumented immigrants lack a high school diploma, as compared to only 21.4 percent of the legal immigrants and 10.6 percent of the native-born.

Figure 2 shows the weighted average Charlson Index for each age (in 5-year brackets by legal status). Note that the Charlson Index is larger (indicating worse health) for the eligible population at every age. Not surprisingly, the index for the eligible population rises rapidly after about age 45. Interestingly, the overall health of undocumented persons also worsens as the population ages, but the rate at which the medical conditions worsen is not as steep for the undocumented. It seems, therefore that the undocumented are healthier (relative to the eligible population) particularly as the groups approach retirement age.[16]

It is instructive to begin our analysis of the link between employment and disability status by contrasting the trends in the number of disabled persons (as we have defined them in the NHIS) and the number of persons receiving Social Security Disability benefits (SSDI) or Supplemental Security Income (SSI). Figure 3 illustrates several trends, revealing that all measures have been increasing rapidly

The NHIS data, where disability status is defined by the number of persons who did not work in the past week due to health reasons, typically indicates about twice as many disabled persons as the number of persons who actually receive either type of disability benefits, whether from the NHIS data or from the official Social Security Administration (SSA) data. In 2010, for example, our definition of disability in the NHIS data implies a count of 16 million persons

[16] It is important to emphasize that the Charlson Index is only for descriptive purposes and will not be used in the more formal empirical analysis below.

disabled. This contrasts with the 8 million or the 7 million that the official SSA data or the NHIS, respectively, report as receiving Social Security disability benefits.

The "excess" number of disabled persons given by our definition is not surprising. Our count includes not only the persons receiving disability benefits, but also the eligible population who are unable to work for health-related reasons but do not receive benefits, as well as the undocumented persons who are ineligible for benefits. Note also that the NHIS estimates of the number of persons receiving benefits are of the same order of magnitude as the estimates from the SSA data, although the NHIS estimates are somewhat lower.

We now turn to our regression results. We use three alternative functional forms for the distribution function *F*: a linear probability model, a probit function, and a logit function. Our results are not sensitive to the choice of the distribution function. Table 3 reports the marginal effects (*dy*/*dx*) for each medical condition across the alternative statistical specifications when we estimate the regression model using the pooled sample of legal immigrants and native born as the "eligible" baseline. It is evident that all medical conditions increase the probability that a person did not work in the reference week due to health reasons, and all of the effects are statistically significant.

We re-estimated the regression model using the sample of undocumented persons, and Table 4 reports the relevant coefficients. Table 4 again shows that all of the coefficients are positive and statistically significant. The OLS results in column 1 are somewhat less significant than in Table 3, but this is probably because the linear probability model is misspecified (after all, the mean disability rate for undocumented persons is only 1.4 percent).[17]

---

[17] The coefficients on the year fixed effects are in Table A-3 and Table A-4, respectively.

Table 5 summarizes the results from the Oaxaca-Blinder decompositions. In all cases, the difference in the regression coefficients (i.e., how much each condition increases the propensity of an individual to report being disabled according to legal status) explain about 80 percent of the difference in the mean disability rate, whereas the differences in endowments (i.e., that the undocumented population is younger and healthier) only explains about 20%., The interaction term, which explains how differences in the coefficients (i.e., how health affects disability) differ across the distribution of values for the endowments (i.e., health differences), is relatively small in magnitude, implying that it does not factor into our interpretation of the results. This small interaction term suggests that the magnitude of the endowment effect does not differ between groups, or equivalently the magnitude of the coefficient effect does not differ between groups (Etezady et al. 2020).

In short, the different disability rates between the two groups are mostly attributable to the fact that adverse medical conditions and the values of the demographic variables are far less likely to lead to withdrawal from the labor force in the undocumented sample than in the eligible sample.[18,19,20] We can then break these decomposition results further into variables for health conditions (e.g., diabetes, asthma) and demographic variables (i.e., sex, education, age). We see that the difference in endowments is driven mainly by the health conditions (i.e., the levels of

[18] This result is consistent with Borjas (2017a), which finds the labor supply curve of undocumented workers is inelastic.

[19] One may also be concerned that the native born and legal immigrants are not a valid comparison group for undocumented immigrants. We address this by repeating our analysis using only Hispanic, non-veteran, high school drop outs, about half of whom are undocumented and about half are not. Table B-2 then shows the corresponding Oaxaca-Blinder decomposition, with a similar 25-75 split between endowments and coefficients.

[20] It is also possible that those who migrate have a different average relationship between health characteristics and labor supply than those who do not. We are not concerned about this as those who do not migrate are not in our sample.

these conditions), whereas the difference in coefficients is driven mainly by demographic variables (i.e., the mapping from these variables to disability).

Figure 4 shows the actual and predicted disability rates for the pooled sample of "eligible" persons (the legal immigrant and native-born born populations). The figure illustrates two alternative measures of the predicted disability rate. First, the disability rate as predicted by the regression model fitted on data from the eligible population. Second, the disability rate as predicted by the model fitted using the sample of undocumented persons.

It is visually obvious that the two trend lines corresponding to the actual disability rates and those predicted from the "own" regression model are very close to each other, and show the substantial upward trend in disability rates described earlier and first documented in Figure 3. In contrast, the trend predicted from the regression model estimated in the sample of undocumented persons shows both a lower overall disability level and no noticeable time trend. In other words, if the eligible population behaved as if they were undocumented workers (and lacked access to disability benefits), they would be far less likely to be absent from work due to health reasons, *and* we would not have witnessed the substantial increase in the disability rate of this population.

We repeated this exercise to illustrate the actual and predicted disability for the undocumented population. Figure 5 shows that the actual level of the disability rate for undocumented immigrants is quite low, has no time trend, and is very well predicted by our regression model. In contrast, when we use the regression model fitted in the eligible population, the predicted disability rate for undocumented persons is markedly higher and shows a noticeable upward time trend. Put differently, if the undocumented workers behaved as if they were eligible for disability benefits, their disability rate would increase by about 6 percentage points, and that disability rate would have almost doubled from about 4 percent to 8 percent between 1997 and 2015.

V. Falsification tests

There are three possible mechanisms that could be preventing undocumented immigrants from collecting disability benefits. One is the official ineligibility as described above. A second is that undocumented immigrants may be culturally different and therefore less likely to report that they are disabled given the same underlying health conditions; see Woodland and Yoshida (2006); Kapteyn, Smith, and van Soest (2007); Burkhauser, Daly, and Ziebarth (2016); and McVicar, Wilkins, and Ziebarth (2018) for evidence of cultural differences across countries in disparities in rates of receipt disability benefits. A third is that the interaction between undocumented immigrants with any official system is fundamentally different than for legal immigrants and the native born, even when their technical access to resources is the same. For example, given that any official interaction carries the risk of deportation, undocumented immigrants are less likely to report domestic abuse (Engelbrecht, 2018) and are more likely to be victims of wage theft (Theodore, 2017).

In other words, the disparity in disability rates documented in earlier sections may be a manifestation of either of these two other mechanisms and not directly attributable to the difference in disability benefit eligibility. This section performs two falsification tests with other outcomes to see if the respective Oaxaca-Blinder decompositions show a similar overwhelming majority of the difference being due to coefficients. If that is the case, it would support concern about our results being due to global differences. Alternatively, if the falsification tests show that substantial variation in the other outcomes can explained by observables, it would support our identification strategy that differences in eligibility for disability benefits is a valid identification strategy.

Table 6 illustrates the results for two alternative outcomes: self-reported health status and not having seen a general physician in the past year. The underlying regression model is the same as that used in our analysis of disability rates. We would expect the share of self-reported health status explained by underlying health conditions to be systematically differently only depending on cultural differences, and not any kind of benefit eligibility or fear of deportation. The decomposition reported in column (1) supports this conjecture, where the overwhelming majority of the difference is due to endowments, and not coefficients.

Columns (2)-(4) examine the outcome of having seen a general physician in the past year. As expected, there is a substantial difference between the eligible population (i.e, natives and legal immigrants) and the undocumented population. But when we decompose this difference into endowments and coefficients, we see a much more even split. We would still expect to see some of the difference be due to coefficients, as there may be cultural differences, a substantial fear of deportation, and because undocumented immigrants are generally not eligible for Medicaid. Still, there is a reasonable availability of healthcare due to charity or cash clinics, compared to minimal if any availability of disability benefits. Given this, we see a much more balanced split of 45-55 as compared to 20-80 above.

VI. Robustness checks

We now address the sensitivity of the evidence by including additional health conditions in the analysis, examining the results in sub-populations of immigrants, and replicating the analysis in an alternative data set: the California Health Insurance Survey (CHIS). These sensitivity tests show that our evidence is indeed robust. Undocumented immigrants are healthier than the legal population at every age, and disability rates would be far lower today, with no upward trend in the past two decades, had the Social Security disability program not existed. For the sake of brevity, the presentation of the results will often be relegated to tables or figures in the Appendix.

We first replicated our Oaxaca-Blinder decomposition using many more measures of health status beyond those used by the Social Security Administration (2017b) in determining disability. Table A-1 shows the results. The first column replicates the evidence from our earlier analysis. The second column adds the following severe health conditions to the vector of health variables: heart attack, angina, other heart disease, stroke, and kidney disease. Finally, the last column of the table adds indicators for different types of common cancers, including breast, cervical, colon, kidney, leukemia, lung, lymphoma, thyroid, and uterine cancers. The evidence from the most complete specification shows that the share of the difference explained by coefficients declines only from 83.2 to 80.6 percent. In short, the difference in disability rates between undocumented persons and the eligible population, is explained mostly by differences in the coefficients that determine disability. In other words, the undocumented have lower disability rates not because they tend to be healthier on average, but because they respond differently to the underlying health conditions.

We now conduct several placebo comparisons to again demonstrate the robustness of the key conclusion. For the first two, we leverage the fact that one needs a certain number of work credits to qualify for federal disability benefits.[21] In particular, we first compare two groups who should not have any difference in the ability to claim disability benefits: the native-born and

[21] https://www.ssa.gov/planners/credits.html

legal immigrants who came into the country as children, as both are approximately equally likely to have sufficient work credits.[22,23] We estimated the disability regressions in each of these two groups, and then predicted what the disability rate would have been had natives (or legal immigrants) responded to health conditions as did the legal immigrants (or natives). As Figures A-1a and A-1b show, the trends in disability rates are essentially similar, so that the status of being native versus being a legal immigrant who entered the country as a child provides no information whatsoever about disability rates. We also compared two alternative groups who should *not* qualify for benefits: legal immigrants who entered the country recently (up to 5 years prior to the survey), and therefore likely lack sufficient work credits, and undocumented immigrants. As Figures A-2a and A-2b show, the trends in disability rates in these two groups are again quite similar.

In short, the analysis of alternative placebos—in one case, both groups can claim benefits, and in the second case, neither group can claim benefits—shows that the evidence reported in the previous section arise specifically because we are comparing two populations that have different access to the Social Security disability system.

Next, following Pourat, Wallace, Hadler, and Ponce (2014), we re-estimated our

[22] This is plausible as those under 31 can qualify for benefits with a reduced number of credits (e.g., those under 24 can qualify with as few as 6 credits, of which 4 can be earned in a single year). See https://www.ssa.gov/benefits/retirement/planner/credits.html. It is also consistent with recent literature showing that immigrants who migrated as children are more similar to natives in terms of earnings than other immigrants (e.g., Hermansen 2017; Gustafsson, Innes, & Österberg 2017).

[23] The categories for the years-in-the-US variable in the NHIS are: <1, 1-4, 5-9, 10-14, and 15 or more. To be conservative as to whether an immigrant came as a minor, we subtracted the lower bound of each category from the individual’s age. We categorized an individual as immigrating as a minor if this result was less than 18. Additionally, given that the NHIS variable for years spent in the US topcodes at 15 years, we cannot determine whether a legal immigrant came as a minor not if that individual is older than 32.

regression models using the CHIS (2017) data. It is much more difficult to apply the residual method that imputes undocumented status in microdata in the CHIS data, as there is no information on the rest of a respondent's household (and so immigrants with legal spouses, parents, or grandparents cannot be classified as having legal status) and there are only extremely broad occupation and industry codes (limiting the exclusion of persons employed in licensed occupations). Additionally, many of the variables for medical conditions are entirely missing or only exist in certain years of the data.

We address this data problem by including two dummy variables for each condition: one for whether the individual has it (as in our analysis of the NHIS data) and one for whether there is no information available for that condition for that individual. This causes the model to be more unstable and not converge for a logit or probit specification. Nevertheless, Figure A-3 shows that our age/health profile result holds (where the legal and native-born population is less healthy at every age). Further, Table A-2 shows that our key Oaxaca-Blinder decomposition result (that the difference is coming from differences in the coefficients coefficients) also holds. In Figures A-4a and A-4b we show that as above predicting for the legal population using the undocumented linear probability model reduced the level and removes the trend, and vice versa increases the level and introduces a trend. We also conducted many other robustness checks. For example, in Table B-11 we show that our results are robust to excluding receipt of Social Security payments, including SSI and SSDI for disability, from the remainder method of assigning likely documentation status. This is because including these could potentially bias our analysis, given that our ultimate outcome variable is being out of work for reasons of health or disability.[24]

[24] Please see Appendix A and Appendix B for a full list.

VII. Discussion and Implications

We can try to use the estimates from our analysis of the NHIS data to attempt to quantify the answers to our questions: (1) What would be the cost savings if disability rates were reduced to the risk-adjusted levels that would be seen if the disability benefits were not available? And (2) what would the cost to the disability program of an “amnesty” that would regularize the status of undocumented immigrants and give them full access to disability benefits? We recognize that this analysis relies on strong assumptions about the external validity of our results.

Table 7 shows each element of the calculation required to begin to answer these questions. In 2015 (the last year of NHIS data used in our analysis), the sum of the survey weights corresponds to a population of 184 million eligible persons (i.e., the native-born plus the legal immigrants) aged 18–64. Figure 4 shows the disability rate dropping from the measured 8.1 percent (or roughly 14.9 million individuals) to only 2.4 percent (or 4.3 million individuals) when the model fitted on the undocumented population is used. Looking in the NHIS data at the disabled legal and native-born population aged 18–64, 40.6 percent of those who report being out of work for health or disability reasons receive SSDI and 29.5 percent receive SSI.[25] In January 2017, the average monthly benefits for SSDI were $1,171.25 (Social Security 2017a) and for SSI $542.5 (Social Security 2017c). A corresponding drop in payouts would potentially save $6.7 billion per month (or $81 billion per year). In January 2017, approximately $10.3 billion was paid in SSDI (Social Security 2017a) and $4.7 billion in SSI (Social Security 2017c). This potential decline thus represents a 45 percent drop in payouts.

[25] Specifically, these individuals said yes when asked if they received each of Social Security and SSI due to a disability.

Another way to summarize the evidence is that there is no trend in the disability rate for the eligible population when predicted from the undocumented model. This suggests that the entire rise that we've seen in the past two decades – from 5.8 to 8.1 percent - can be mostly explained by the differences in coefficients, and not by a population that is getting older and sicker.

The second exercise, relevant from the current policy discussion about regularizing the status of undocumented immigrants, is to calculate the increase in payouts if undocumented individuals were granted legal status. Table 8 shows each element of the required calculation. The most recent DHS estimate is that there are 12.1 million undocumented immigrants (DHS 2018), which closely matches the sum of survey weights from our analysis (11.7 million) and which we use above for consistency. In January 2017, the average monthly benefits for SSDI were $1,171.25 (Social Security 2017a) and for SSI $542.5 (Social Security 2017c).[26] The predicted increase in the share of undocumented immigrants who are disabled if they were "treated like" legal immigrants would be from 1.3 percent to 6.7 percent. Allowing all of the these persons to claim benefits (as even the ones who previously reported disability can now claim) would lead to an increase in federal liabilities of $6.0 billion per year, which represents an increase of 3.3 percent in total expenditures.[27] Note, however, that many undocumented immigrants may already be paying taxes to the disability system but currently are not qualified for benefits (Goss et al. 2013; Social Security 2015; Gee et al. 2017). Additionally, many newly

[26] Earnings histories may be different for the undocumented immigrant population which could to different expected disability benefits.

[27] This analysis is only the direct cash expenditures of the program, and does not incorporate potential changes other government outlays, such as providing Medicare to those with disability benefits, reductions in ACA or Medicaid insurance subsidies or TANF or SNAP eligible due to disability benefits, or changes in EITC payments.

authorized immigrants may not have sufficient official work history to qualify immediately for benefits (both for disability and from other programs).

Finally, we can try to use our empirical results to answer the question that motivated much of our analysis: how much of the rise in disability rates can be explained by an aging population? A straightforward way to answer this would be to use the 2015 age distribution (say in 5-year brackets) of the population but the 1998 disability rates for each of those brackets. Unadjusted, the disability rate for the 18–64 population (of any immigration status) was 5.6 percent in 1998 and 7.7 percent in 2015. If the disability-by-age rates had remained constant but the population had aged, the predicted rate would have been only 6.2 percent. In other words, the aging of the population may only explain 29 percent of the increase. The rest may be due to changes in other factors such as the impact of medical conditions increasing the probability that a person did not work in the reference week due to health reasons.

VIII. Conclusion

This paper applies newly developed methods that can be used to impute undocumented status to the foreign-born population to the NHIS micro data. The imputation allows us to investigate the health of undocumented immigrants, compare their health status to legal immigrants and the native born, and calculate counterfactuals that help us understand how being unable to work due to a health impairment responds to legal constraints on the availability of benefits.

Our empirical analysis reveals that undocumented immigrants are healthier than those with legal status (either native- or foreign-born) at every age and are less likely to be disabled (in the sense that an existing health condition limits work). We also found that the differences in the disability rates among the various groups can mostly be explained by differences in how medical conditions, age, and education affect disability and not by differences in the mean values of those variables for the groups. In other words, undocumented immigrants are less likely to be disabled not because they are younger and healthier, but because their labor supply is far less responsive to those characteristics than they are for persons legally in the country. Put differently, the relationship between health and disability is stronger for those with legal status than it is for those who are undocumented.

We used those insights to construct two counterfactual scenarios: one where the legal population could not claim disability benefits and one where the undocumented population could. In the first case, the disability rate for the legal population drops substantially and there is no longer the upward sloping time trend in disability observed over the past two decades. In the second, the level of the disability rate increases substantially and an upward sloping time trend appears.

These results suggest that there may be substantial moral hazard in the current disability benefits system and that there may exist numerous situations where an individual with some health limitations could find work. Crafting policy around both of these outcomes could substantially reduce federal outlays and mitigate the upward-sloping trend in disability rates. The results also indicate that legalizing the undocumented population could be accompanied by a modest increase in fiscal outlays without a corresponding increase in revenue, as many undocumented immigrants may be already paying taxes.

References

Akbulut-Yuksel, Mevlude and Adriana D. Kugler. 2016. Intergenerational Persistence of Health in the U.S.: Do Immigrants Get Healthier as they Assimilate? Economics & Human Biology 23: 136–148.

Autor, David H. and Mark G. Duggan. 2003. The Rise In The Disability Rolls And The Decline In Unemployment. Quarterly Journal of Economics 118(1): 157–206.

Autor, David H. and Mark G. Duggan. 2006. The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding. Journal of Economic Perspectives 20(3): 71–96.

Autor, David, Andreas Ravndal Kostøl Magne Mogstad, and Bradley Setzler. 2019. Disability Benefits, Consumption Insurance, and Household Labor Supply. American Economic Review 109(7): 2613–2654

Benjamin, A.E., Steven P. Wallace, Valentine Villa, Kathy Mccarthy. 2000. Disability and Access to Health & Support Services among California's Immigrant Populations, UCLA Center for Health Policy Research.

Black, Dan, Kermit Daniel, Seth Sanders. 2002. The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust American Economic Review 92(1): 27–50.

Blinder, Alan S.: 1973. Wage discrimination: reduced form and structural estimates. Journal of Human Resources 8, 436–455.

Borjas, George J. 2017a. The Labor Supply of Undocumented Immigrants. Labour Economics 46:1–13.

Borjas, George J. 2017b. The Earnings of Undocumented Immigrants NBER Working Paper No. 23236.

Borjas, George J. and Hugh Cassidy. 2019. The wage penalty to undocumented immigration. Labour Economics 61: 101757.

Bound, John. 1989. The Health and Earnings of Rejected Disability Insurance Applicants The American Economic Review 79(3): 482–503.

Bureau of Labor Statistics. 2017. Civilian Noninstitutional Population. https://fred.stlouisfed.org/series/CNP16OV

Burkhauser, Richard. V, Mary C. Daly, and Nicholas. R. Ziebarth. 2016. Protecting Working-Age People with Disabilities: Experiences of Four Industrialized Nations. Journal of Labour Market Research 49(4): 367–386.

California Health Interview Survey. 2017. CHIS 2001-2014. Adult Public Use File. Los Angeles, CA: UCLA Center for Health Policy Research. http://healthpolicy.ucla.edu/chis/Pages/default.aspx.

Campolieti, Michele. 2004. Disability Insurance Benefits and Labor Supply: Some Additional Evidence. Journal of Labor Economics 22(4): 863–889.

CDC. 2014. “Design and Estimation for the National Health Interview Survey, 2006–2015.” Vital and Health Statistics 2(165): 1–43. https://www.cdc.gov/nchs/data/series/sr_02/sr02_165.pdf

Charlson, Mary E.; Pompei, Peter; Ales, Kathy L.; MacKenzie, C. Ronald. 1987. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Chronic Diseases. 40 (5): 373–83.

Christopoulou, Rebekka and Dean R Lillard. 2015. Is smoking behavior culturally determined? Evidence from British immigrants. Journal of economic behavior & organization 110: 78–90.

Cohen, Michael S. and William L. Schpero. 2018. Household Immigration Status Had Differential Impact On Medicaid Enrollment In Expansion And Nonexpansion States. Health Affairs 37(3): 394–402.

Department of Homeland Security (DHS). 2018. Estimates of the Unauthorized Immigrant Population Residing in the United States. https://www.dhs.gov/immigration-statistics/population-estimates/unauthorized-resident

Duggan, Mark and Scott A. Imberman. 2009. Why Are the Disability Rolls Skyrocketing? The Contribution of Population Characteristics, Economic Conditions, and Program Generosity. In Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly, National Bureau of Economic Research, 337–379.

Engelbrecht, Cora. 2018. Fewer Immigrants Are Reporting Domestic Abuse. Police Blame Fear of Deportation. New York Times. Available at https://www.nytimes.com/2018/06/03/us/immigrants-houston-domestic-violence.html

Etezady, A., Shaw, F.A., Mokhtarian, P.L., and G. Circella. 2020. What drives the gap? Applying the Blinder–Oaxaca decomposition method to examine generational differences in transportation-related attitudes. Transportation.

Flood, Sarah Miriam King, Steven Ruggles, and J. Robert Warren. 2017. Integrated Public Use Microdata Series, Current Population Survey: Version 5.0. CPS ASEC. Minneapolis, MN: University of Minnesota. https://doi.org/10.18128/D030.V5.0

Foote, Andrew, Michel Grosz Stephanie Rennane. 2019. The Effect of Lower Transaction Costs on Social Security Disability Insurance Application Rates and Participation. Journal of Policy Analysis and Management 38(1): 99–123.

French, Eric and Jae Song. 2014. The Effect of Disability Insurance Receipt on Labor Supply. American Economic Journal: Economic Policy 6(2): 291–337.

Furtado, Delia and Nikolaos Theodoropoulos. 2016. Immigrant Networks and The Take-Up Of Disability Programs: Evidence From The United States, Economic Inquiry 54(1): 247–267.

Gee, Lisa Christensen, Matthew Gardner, Misha E. Hill, and Meg Wiehe. 2017. Undocumented Immigrants' State & Local Tax Contributions. The Institute on Taxation & Economic Policy. https://itep.org/wp-content/uploads/ITEP-2017-Undocumented-Immigrants-State-and-Local-Contributions.pdf

Gelber, Alexander, Timothy Moore, and Alexander Strand. 2017. The Effect of Disability Insurance Payments on Beneficiaries' Earnings. American Economic Journal: Economic Policy 9(3): 229–261.

Giuntella, Osea and Luca Stella. 2017. The Acceleration of Immigrant Unhealthy Assimilation. Health Economics 26(4): 511–518.

Giuntella, Osea and Jakub Lonsky. 2020. The Effects of DACA on Health Insurance, Access to Care, and Health Outcomes. Forthcoming, Journal of Health Economics.

Giuntella, Osea, Jakub Lonsky, Fabrizio Mazzonnac, and Luca Stella. 2021. Immigration policy and immigrants' sleep. Evidence from DACA. Journal of Economic Behavior & Organization 182: 1–12.

Goldman, Dana P., James P. Smith and Neeraj Sood. 2006. Immigrants And The Cost Of Medical Care. Health Affairs 25(6):1700–1711.

Goss, Stephen, Alice Wade, J. Patrick Skirvin, Michael Morris, K. Mark Bye, and Danielle Huston. 2013. Effects of Unauthorized Immigration on the Actuarial Status of the Social Security Trust Funds. Social Security Administration Actuarial Note Number 151.

Gruber, Jonathan and Jeffrey D. Kubik. 1997. Disability insurance rejection rates and the labor supply of older workers. Journal of Public Economics 64: 1–23.

Gruber, Jonathan. 2000. Disability Insurance Benefits and Labor Supply. Journal of Political Economy 108(6): 1162–1183.

Gustafsson, B.A., Mac Innes, H. & Österberg, T. 2017. Age at immigration matters for labor market integration—the Swedish example. IZA Journal of Development and Migration 7(1).

Hermansen, A.S. 2017. Age at Arrival and Life Chances Among Childhood Immigrants. Demography 54: 201–229.

Jiménez-Martín, Sergi. Arnau Juanmarti Mestres, Judit Vall Castelló. Great Recession and disability insurance in Spain. Forthcoming in Empirical Economics.

Johnston, David W., Carol Propper, and Michael A. Shields. 2009. Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. Journal of Health Economics 28: 540–552.

Kapteyn, Arie, James P. Smith, and Arthur Van Soest. 2007. Vignettes and Self-Reports of Work Disability in the United States and the Netherlands. American Economic Review, 97(1): 461–473.

Kostøl, Andreas Ravndal and Magne Mogstad. 2014. How Financial Incentives Induce Disability Insurance Recipients to Return to Work. American Economic Review 104(2): 624–655.

Kostøl, Andreas Ravndal and Magne Mogstad. 2015. Earnings, Disposable Income, and Consumption of Allowed and Rejected Disability Insurance Applicants. American Economic Review: Papers & Proceedings 105(5): 137–141.

Liebman, Jeffrey B. 2015. Understanding the Increase in Disability Insurance Benefit Receipt in the United States. Journal of Economic Perspectives 29(2): 123–150.

Low, Hamish and Luigi Pistaferri. 2015. Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off. American Economic Review 105(10): 2986–3029.

Low, Hamish and Luigi Pistaferri. 2019. Disability Insurance: Error Rates and Gender Differences. NBER Working Paper No. 26513.

National Health Interview Survey. 2017. https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2013. Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. American Economic Review 103(5): 1797–1829.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2015. Disability Insurance and the Great Recession. American Economic Review: Papers & Proceedings 105(5): 177–182.

McVicar, Duncan, Roger Wilkins, and Nicholas .R. Ziebarth. 2018. Four Decades of Disability Benefit Policies and the Rise and Fall of Disability Recipiency Rates in Five OECD Countries. In: Besharov, D. and D. Call, Labor Activation in a Time of High

Unemployment: Encouraging Work while Preserving the Social Safety Net, Oxford University Press.

Milligan, Kevin and Tammy Schirle. 2019. Push and Pull: Disability Insurance, Regional Labor Markets, and Benefit Generosity in Canada and the United States. Journal of Labor Economics 37(S2): S289–S323.

Mueller, Andreas I., Jesse Rothstein, and Till M. von Wachter, 2016. Unemployment Insurance and Disability Insurance in the Great Recession. Journal of Labor Economics 34(S1): S445–S475.

Oaxaca, Ronald. 1973. Male-female wage differentials in urban labor markets. International Economic Review 14: 693–709

Parsons, Donald. 1980. The decline of male labor force participation. Journal of Political Economy 88: 117–134.

Parsons, Donald. 1982. The male labour force participation decision: Health, reported health, and economic incentives. Economica 49: 81–91.

Passel, Jeffrey S. and D'Vera Cohn. 2014. Unauthorized Immigrant Totals Rise in 7 States, Fall in 14 States: Decline in Those From Mexico Fuels Most State Decreases. Washington, DC: Pew Research Center.

Pourat, Nadereh, Steven P. Wallace, Max W. Hadler, and Ninez Ponce. 2014.Assessing Health Care Services Used By California's Undocumented Immigrant Population In 2010. Health Affairs 33(5): 840–847.

Roberts, Jennifer and Karl Taylor. 2019. New Evidence on Disability Benefit Claims in the UK: The Role of Health and the Local Labour Market. IZA Discussion Paper No. 12825.

Social Security. 2015. Status of the Social Security Administration's Earnings Suspense File. http://oig.ssa.gov/audits-and-investigations/audit-reports/A-03-15-50058

Social Security. 2017a. Selected Data From Social Security's Disability Program. https://www.ssa.gov/oact/STATS/dibStat.html, https://www.ssa.gov/oact/progdata/icp.html .

Social Security. 2017b https://www.ssa.gov/disability/professionals/bluebook/AdultListings.htm

Social Security. 2017c. Monthly Statistical Snapshot, January 2017. https://www.ssa.gov/policy/docs/quickfacts/stat_snapshot/2017-01.html https://www.ssa.gov/policy/docs/statcomps/supplement/index.html

Stimpson, Jim P., Fernando A. Wilson and Dejun Su. 2013. Unauthorized Immigrants Spend Less Than Other Immigrants And US Natives On Health Care. Health Affairs 32(7): 1313–1318.

Theodore, Nik. 2017. After the Storm: Houston's Day Labor Markets in the Aftermath of Hurricane Harvey. Available at https://greatcities.uic.edu/2017/11/21/after-the-storm-houstons-day-labor-markets-in-the-aftermath-of-hurricane-harvey/

Von Wachter, Till, Jae Song, and Joyce Manchester. 2011. Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program. American Economic Review 101(7): 3308–29.

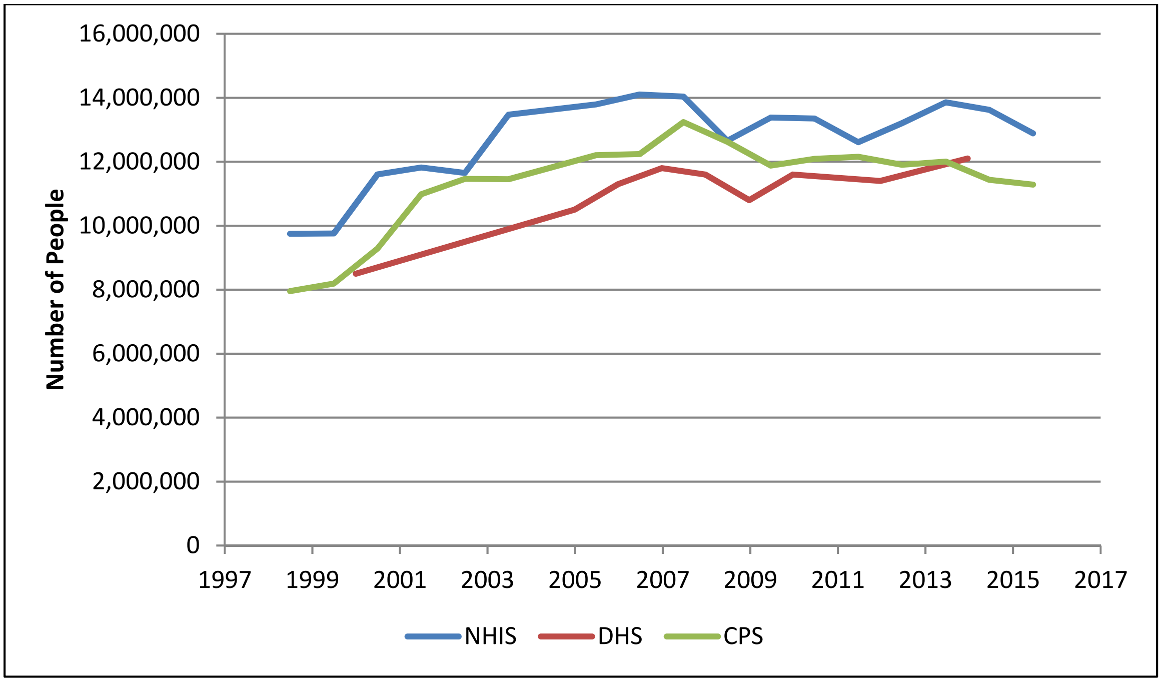
Warren, Robert E. and Jeffrey S. Passel. 1987. A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 United States Census. Demography 24(3): 375–393.

Woodland, Alan D., and Chisato Yoshida. 2006. Risk preference, immigration policy and illegal immigration. Journal of Development Economics 81(2): 500–513.

Xiang, H., Shi, J., Wheeler, K. and Wilkins, J., III. 2010. Disability and employment among U.S. working-age immigrants. American Journal of Industrial Medicine 53: 425–434.

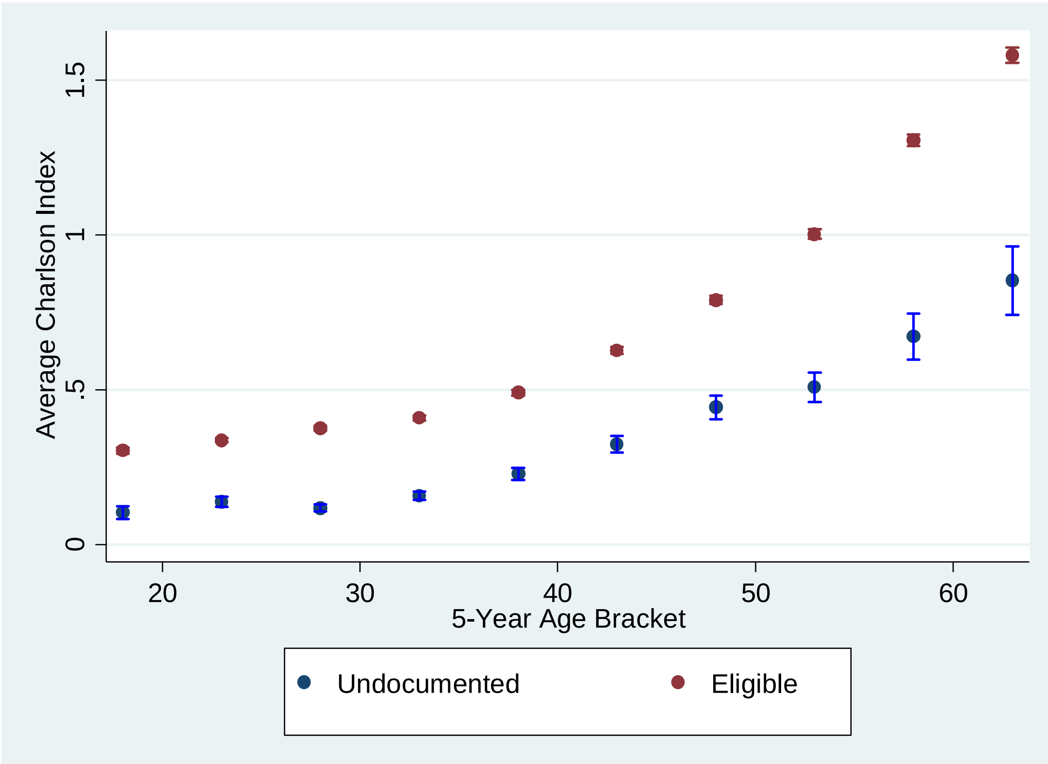
Yun, Myeong-Su. 2004. Decomposing differences in the first moment. Economics Letters 82: 275–280.

Figure 1: Counts of the Undocumented Population, by Year



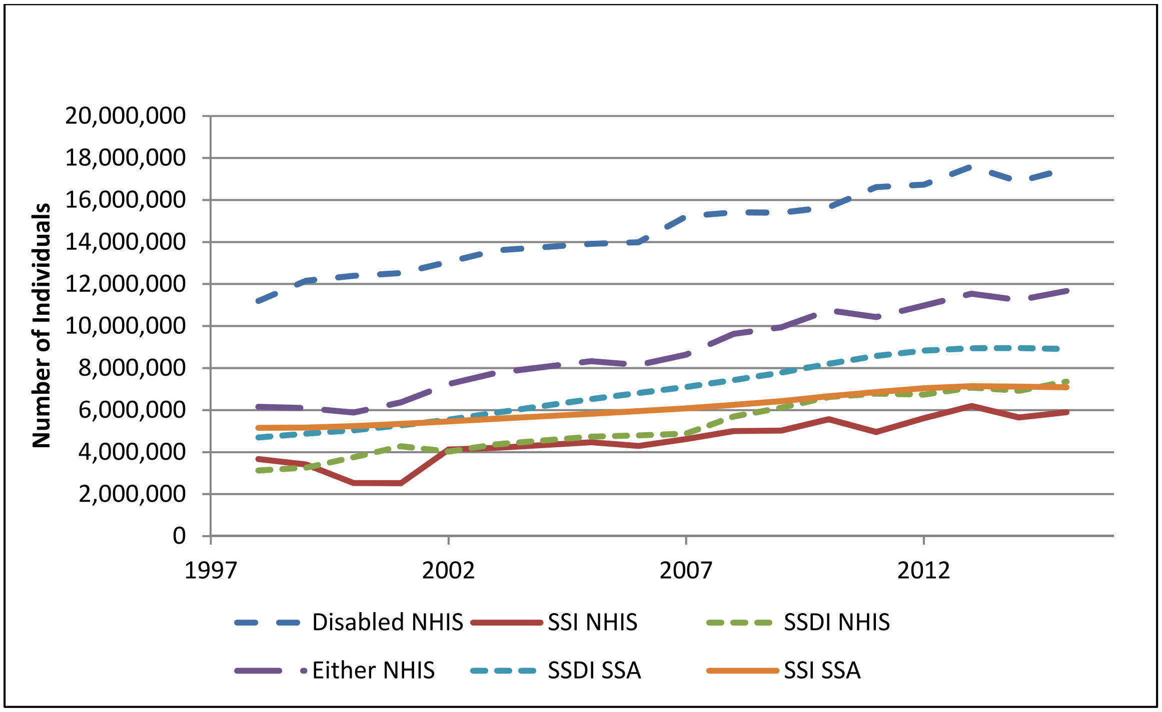
Note: The official count of undocumented persons is drawn from Department of Homeland Security (2018).

Figure 2: Charlson Index by Age for Undocumented Immigrants and Eligible Samples



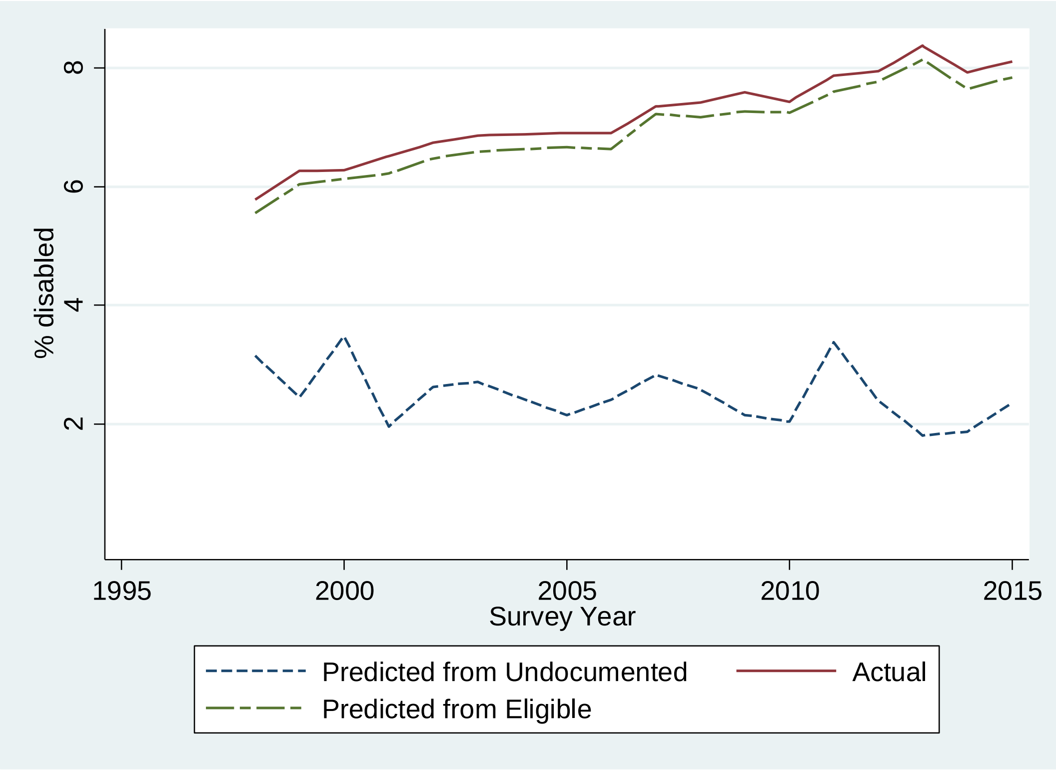
Notes: NHIS Sample Adults, 18–64. Weighted. 95% confidence interval shown in whiskers around each point.

Figure 3: Trend in Disability and Benefits, NHIS vs. Social Security



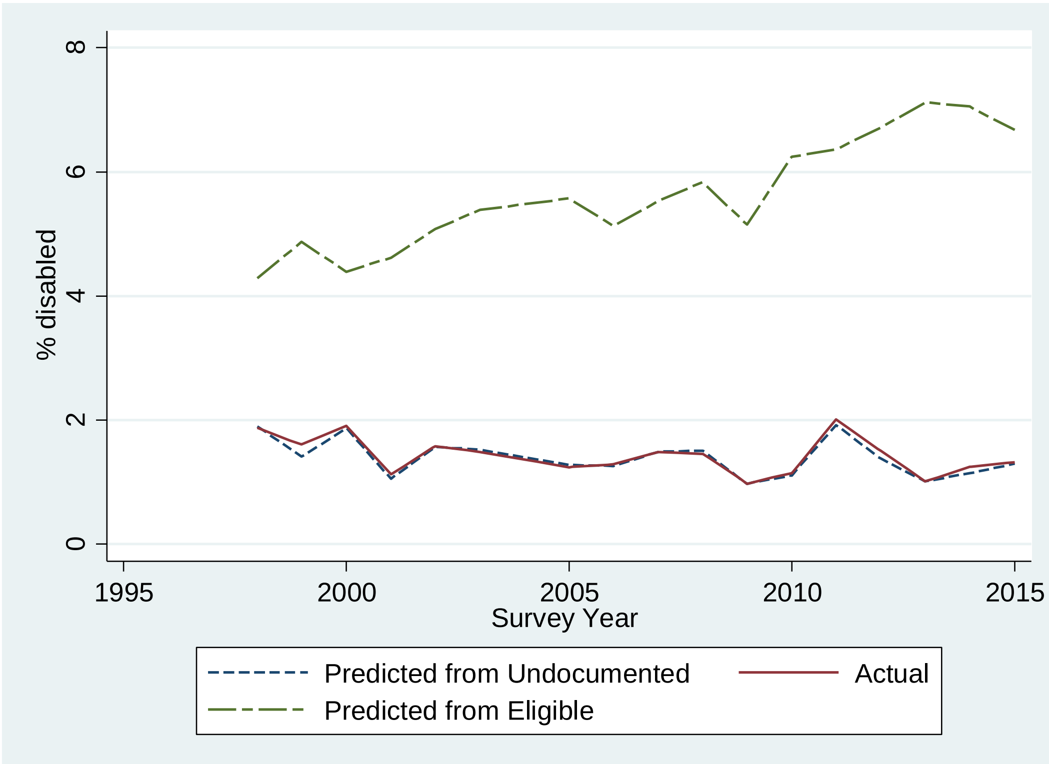
Notes: SSDI Data from Social Security (2017a). SSI Data from Social Security (2017c). Here we include adults of all ages to be consistent with the SSA data.

Figure 4: Predicted Trend in Disability Rates for the Eligible Sample



Notes: NHIS Sample Adult 18–64. Weighted. Uses Logit model from above.

Figure 5: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: NHIS Sample Adult 18–64. Weighted. Uses Logit model from above.

Table 1: Applying the imputation method to determine undocumented status

	Observations (17 years)	Sum of weights (17 years)	Sum of weights (annual average)
Total	1,615,911	4,996,834,913	293,931,465
Native Born	1,343,729	4,361,782,290	256,575,429
Citizens	112,550	293,346,825	17,255,696
Receive Government Benefits	23,902	49,432,561	2,907,798
In the Military	1,953	11,762,416	691,907
Veteran	374	851,672	50,098
Receives Welfare	677	1,502,691	88,394
Cubans	2,745	4,999,549	294,091
Works in a Licensed Occupation	1,177	7,964,862	468,521
Spouse Is a Citizen	7,186	17,496,593	1,029,211
Other Family Member Is a Citizen	16,613	32,141,844	1,890,697
Residual = undocumented	105,005	215,553,610	12,679,624

Notes: Data from NHIS Sample Adult and Sample Child files. Pooled for years 1998-2003 and 2005-2015. Each row represents the count of those excluded by that row but not the above rows.

Table 2: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)	
	Native Born	Legal Immi-grants	Eligible (Pooled Native Born and Legal Immigrants)[28]	Undocu -mented	Difference between (3) and (4)	Standard Error	
Disabled	0.073	0.044	0.070	0.014	-0.0561***	(0.0017)	
Male	0.488	0.477	0.487	0.559	0.0725***	(0.0033)	
Heart Disease	0.022	0.015	0.021	0.008	-0.0137***	(0.0009)	
Cancer	0.051	0.025	0.048	0.007	-0.0407***	(0.0014)	
Diabetes	0.058	0.063	0.058	0.038	-0.0206***	(0.0015)	
Hypertension	0.217	0.180	0.213	0.099	-0.114***	(0.0027)	
Asthma	0.128	0.068	0.121	0.036	-0.0855***	(0.0021)	
Emphysema	0.011	0.003	0.010	0.002	-0.00810***	(0.0006)	
Liver Disease	0.013	0.014	0.013	0.008	-0.00467***	(0.0007)	
Joint Pain	0.300	0.189	0.288	0.111	-0.177***	(0.0030)	
Back Pain	0.284	0.233	0.279	0.173	-0.105***	(0.0030)	
Neck Pain	0.154	0.125	0.151	0.084	-0.0667***	(0.00234)	
Face Pain	0.051	0.031	0.049	0.020	-0.0293***	(0.0014)	
Ulcer	0.066	0.045	0.064	0.029	-0.0352***	(0.0016)	
Bronchitis	0.043	0.018	0.040	0.009	-0.0311***	(0.0013)	
Any Ailment	0.631	0.515	0.618	0.365	-0.253***	(0.0032)	
Age (years)	40.2	41.5	40.3	35.4	-5.0***	(0.087)	
High School Dropout	0.105	0.213	0.117	0.452	0.335***	(0.0022)	
High School Graduate	0.284	0.214	0.276	0.211	-0.0655***	(0.0030)	
Some College	0.334	0.245	0.324	0.140	-0.183***	(0.0031)	
College Graduate	0.277	0.328	0.283	0.197	-0.0856***	(0.0030)	
					45,889	373,954	30,012
N	328,065	45,889	373,954	30,012			

Notes: NHIS Sample Adults, 18–64. Weighted. *** p<0.01, ** p<0.05, * p<0.1

[28] Throughout this paper, we pool those eligible for benefits (the native born and legal immigrants). In the appendix, we repeat our analysis comparing undocumented immigrants to native born and legal immigrants separately, and find broadly comparable results.

Table 3: Predicting Disability Status using Self-Reported Medical Conditions, for Eligibles

	(1)	(2)	(3)
	OLS	Probit	Logit
Heart Disease	0.154***	0.0548***	0.0490***
	(0.00621)	(1.97e-05)	(1.78e-05)
Cancer	0.0454***	0.0287***	0.0264***
	(0.00316)	(1.56e-05)	(1.48e-05)
Diabetes	0.102***	0.0436***	0.0394***
	(0.00338)	(1.34e-05)	(1.23e-05)
Hypertension	0.0379***	0.0255***	0.0250***
	(0.00152)	(9.72e-06)	(9.51e-06)
Asthma	0.0252***	0.0204***	0.0203***
	(0.00187)	(1.19e-05)	(1.15e-05)
Emphysema	0.229***	0.0612***	0.0524***
	(0.00949)	(2.76e-05)	(2.44e-05)
Liver disease	0.179***	0.0675***	0.0609***
	(0.00753)	(2.45e-05)	(2.20e-05)
Joint pain	0.0361***	0.0287***	0.0284***
	(0.00126)	(9.42e-06)	(9.46e-06)
Bain Pain	0.0379***	0.0304***	0.0312***
	(0.00134)	(9.69e-06)	(9.70e-06)
Neck Pain	0.0454***	0.0260***	0.0249***
	(0.00187)	(1.10e-05)	(1.05e-05)
Face Pain	0.0535***	0.0289***	0.0273***
	(0.00325)	(1.57e-05)	(1.47e-05)
	0.0472***	0.0220***	0.0203***
Ulcer	(0.00279)	(1.36e-05)	(1.27e-05)
	0.0535***	0.0218***	0.0198***
Bronchitis	(0.00364)	(1.69e-05)	(1.57e-05)
	0.154***	0.0548***	0.0490***
Observations	373,954	373,954	373,954
R-squared	0.162	0. 242	0.241

Notes: NHIS Sample Adult 18–64. Weighted. Columns 2 and 3 show marginal effects. Model also includes age category, education category, sex, and survey year and survey quarter fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Predicting Disability Status with Medical Conditions, for Undocumented

	(1)	(2)	(3)
	OLS	Probit	Logit
Heart Disease	0.0249	0.00733***	0.00653***
	(0.0218)	(5.88e-05)	(5.06e-05)
Cancer	0.0363*	0.0149***	0.0125***
	(0.0220)	(5.79e-05)	(5.12e-05)
Diabetes	0.0265***	0.00862***	0.00788***
	(0.00887)	(3.02e-05)	(2.68e-05)
Hypertension	0.00982*	0.00644***	0.00550***
	(0.00510)	(2.29e-05)	(2.19e-05)
Asthma	0.00238	0.00138***	0.00190***
	(0.00545)	(4.06e-05)	(3.80e-05)
Emphysema	0.0756	0.0115***	0.00887***
	(0.0603)	(0.000104)	(8.10e-05)
Liver disease	0.0417*	0.00958***	0.00903***
	(0.0225)	(5.71e-05)	(4.62e-05)
Joint pain	0.00744*	0.00384***	0.00329***
	(0.00392)	(2.24e-05)	(2.17e-05)
Back Pain	0.0249	0.00733***	0.00653***
	(0.0218)	(5.88e-05)	(5.06e-05)
Neck Pain	0.0363*	0.0149***	0.0125***
	(0.0220)	(5.79e-05)	(5.12e-05)
Face Pain	0.0265***	0.00862***	0.00788***
	(0.00887)	(3.02e-05)	(2.68e-05)
Ulcers	0.00361	0.00190***	0.00171***
	0.000566	-5.88e-05	-0.000254***
Bronchitis	(0.00714)	(3.96e-05)	(3.69e-05)
	0.0249	0.00559***	0.00534***
Observations	30,012	30,01	30,01
R-squared	0.028	0.124	0.123

Notes: NHIS Sample Adult 18–64. Weighted. Columns 2 and 3 show marginal effects. Model also includes age category, education category, sex, and survey year and survey quarter fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Oaxaca-Blinder Decomposition

	(1)	(2)	(3)
	OLS	Probit	Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699*** (0.00000476)	0.0697*** (0.00000446)	0.0699*** (0.00000445)
Undocumented	0.0138*** (0.00000865)	0.0138*** (0.00000854)	0.0138*** (0.0000085)
Difference in means	0.0561*** (0.00000987)	0.0560*** (0.00000963)	0.0561*** (0.0000096)
Share due to:			
Endowments	0.0114*** (0.0000109)	0.0112*** (0.000022)	0.0110*** (0.0000233)
Demographic variables:	12%	5%	6%
Health conditions:	87%	93%	92%
Coefficients	0.0502*** (0.0000121)	0.0445*** (0.0000115)	0.0435*** (0.0000113)
Demographic variables:	116%	84%	64%
Health conditions:	42%	6%	2%
Interaction	-0.00542*** (0.0000131)	0.000283*** (0.000023)	0.00158*** (0.0000242)
Observations	403,966	403,966	403,966

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. Demographic variables can have a share above 100% because the contribution of other variables can be of the opposite sign. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Oaxaca-Blinder Decompositions for Other Health Outcomes

	(1)	(2)	(3)	(4)
	Health Status (1=excellent, 5=poor)	Seen a general physician in past year		
	OLS	OLS	Probit	Logit
Means:				
Eligible (Legal Immigrants & Native Born)	2.133*** (0.0000193)	0.652*** (0.00000894)	0.652*** (0.000009)	0.652*** (0.00000901)
Undocumented	2.113*** (0.0000719)	0.402*** (0.0000366)	0.401*** (0.0000365)	0.402*** (0.0000365)
Difference in means	0.0197*** (0.0000745)	0.250*** (0.0000377)	0.251*** (0.0000376)	0.250*** (0.0000376)
Share due to:				
Endowments	0.106*** (0.0000887)	0.128*** (0.0000455)	0.127*** (0.0000417)	0.127*** (0.0000418)
Demographic variables:	-58%	52%	52%	52%
Health conditions:	159%	47%	47%	47%
Coefficients	-0.0447*** (0.0000718)	0.155*** (0.0000378)	0.150*** (0.0000381)	0.148*** (0.0000381)
Demographic variables:	8%	44%	38%	35%
Health conditions:	102%	-9%	-5%	-4%
Interaction	-0.0416*** (0.0000866)	-0.0322*** (0.0000455)	-0.0264*** (0.0000419)	-0.0246*** (0.000042)
Observations	403,784	399,545	399,545	399,545

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. Demographic variables can have a share above 100% because the contribution of other variables can be of the opposite sign. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Cost savings of disability reduction for eligible, 18–64

Total population	184 million (sum of survey weights)
Disabled population	14.9 million (sum of survey weights)
Disability rate	0.0811 (= 14.9 million / 184 million)
Counterfactual disability rate	0.0236 (using counterfactual prediction)
Counterfactual disabled population	4.34 million (= 184 million * 0.0236)
Change in population disabled	-10.6 million (=4.34 million – 14.9 million)
Share of disabled legal and native born receiving SSDI	0.406 (using survey response)
Population no longer receiving SSDI	-4.30 million (= -10.6 million * 0.406)
Average monthly benefits for SSDI	$1,171.25 (from Social Security)
Monthly savings from SSDI	-$5.04 billion (=-4.30 million * $1,171.25)
Share of disabled legal and native-born receiving SSI	0.295 (using survey response)
Population no longer receiving SSI	-3.13 million (= -10.6 million * 0.295)
Average monthly benefits for SSI	$542.5 (from Social Security)
Monthly savings from SSI	-$1.70 billion (=-3.13 million * $542.5)
Total monthly savings	=-$6.74 billion (=-$5.04 billion – $1.70 billion)
Total annual savings	=$81 billion ($6.74 billion * 12)

Table 8: Cost of providing disability benefits to undocumented immigrants

Total population	11.7 million (sum of survey weights)
Disabled population	156,000 (sum of survey weights)
Disability rate	0.0132 (= 156,000 / 11.7 million)
Counterfactual disability rate	0.067 (using counterfactual prediction)
Counterfactual disabled population	784,000 (= 11.7 million * 0.067)
Share of disabled legal and native born receiving SSDI	0.406 (using survey response)
Population now receiving SSDI	318,000 (= 784,000 * 0.406)
Average monthly benefits for SSDI	$1,171.25 (from Social Security)
Monthly cost from SSDI	$372 million (=318,000 * $1,171.25)
Share of disabled legal and native born receiving SSI	0.295 (using survey response)
Population now receiving SSI	231,000 (= 784,000 * 0.295)
Average monthly benefits for SSI	$542.5 (from Social Security)
Monthly cost from SSI	$125 million (=231,000 * $542.5)
Total monthly cost	=$497 million (=$372 million + $125 million)
Total annual cost	=$6.0 billion ($497 million * 12)

For Online Publication,
"Health, Employment, and Disability:
Implications from the Undocumented Population"

Appendix A: Main Robustness Checks

Table A-1: Oaxaca-Blinder Decomposition using Additional Health Measures, Logit Specification

	(1) SSA Controls	(2) SSA Controls + Severity	(3) SSA Controls + Severity + Cancer Type
Means:			
Eligible (Legal Immigrants & Native Born)	0.0695*** (0.00000444)	0.0695*** (0.00000438)	0.0695*** (0.00000438)
Undocumented	0.0138*** (0.00000851)	0.0138*** (0.00000849)	0.0138*** (0.00000846)
Difference in means	0.0557*** (0.0000096)	0.0557*** (0.00000955)	0.0557*** (0.00000952)
Share due to:			
Endowments	0.0110*** (0.0000233)	0.0115*** (0.0000248)	0.0112*** (0.0000231)
Coefficients	0.0434*** (0.0000113)	0.0423*** (0.0000112)	0.0422*** (0.0000112)
Interaction	0.00128*** (0.0000242)	0.00191*** (0.0000256)	0.00228*** (0.000024)
Observations	403,350	403,350	403,350

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. Column (2) contains dummy variables for heart attack, angina, other heart disease, stroke, kidney disease. Column (3) additionally contains dummy variables for common types of cancer: breast, cervical, colon, kidney, leukemia, lung, lymphoma, thyroid, uterine. *** p<0.01, ** p<0.05, * p<0.1

Figure A-1a: Predicted trend in Disability Rates for Native Born

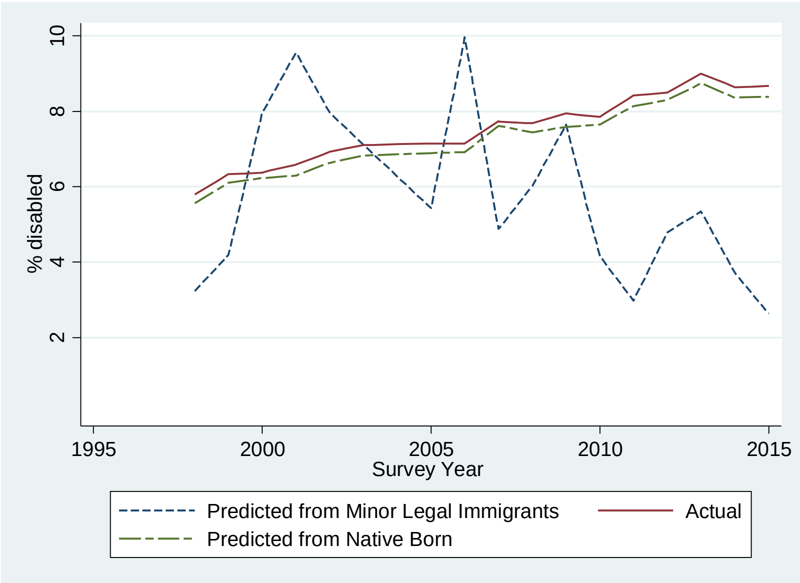
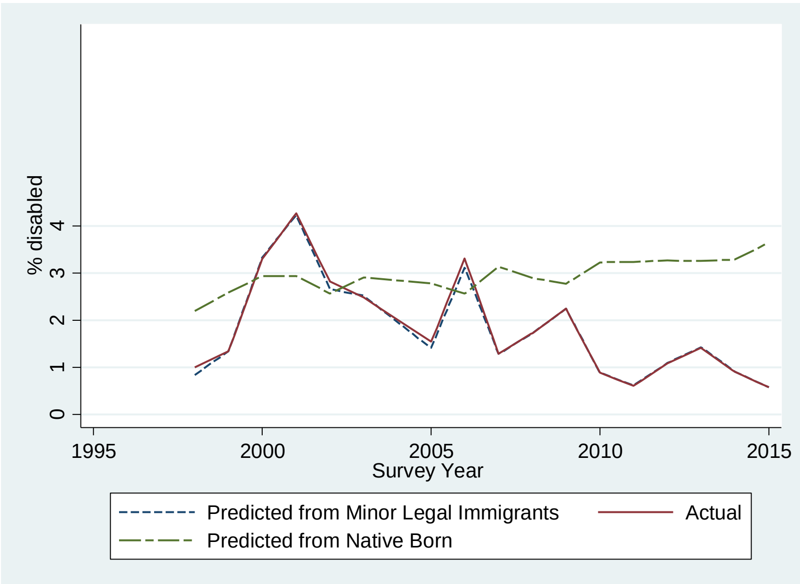


Figure A-1b: Predicted trend in Disability Rates for Minor Legal Immigrants



Notes: NHIS Sample Adult 18–64. Weighted. Uses Logit model.

Figure A-2a: Predicted trend in Disability Rates for Recent Legal Immigrants

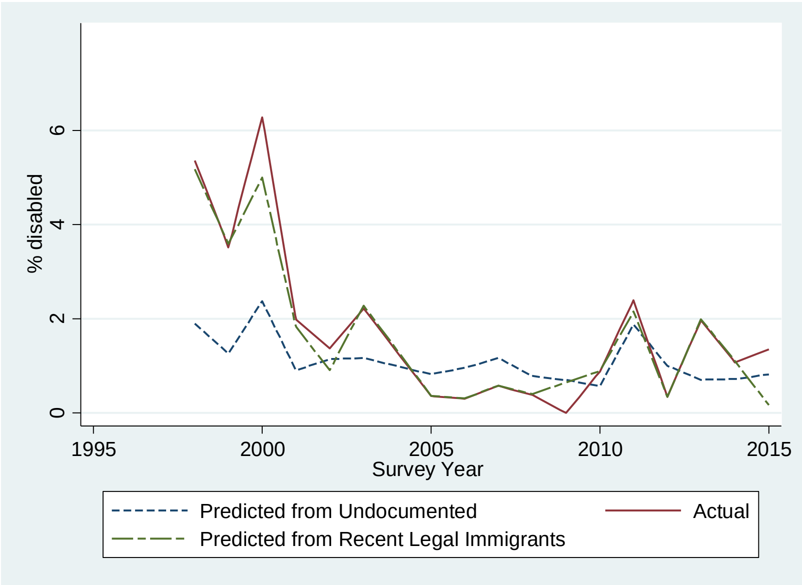
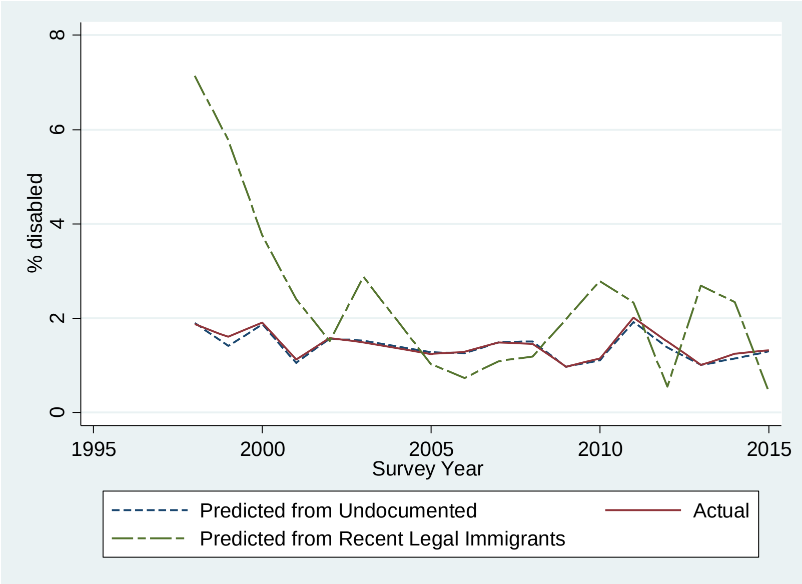
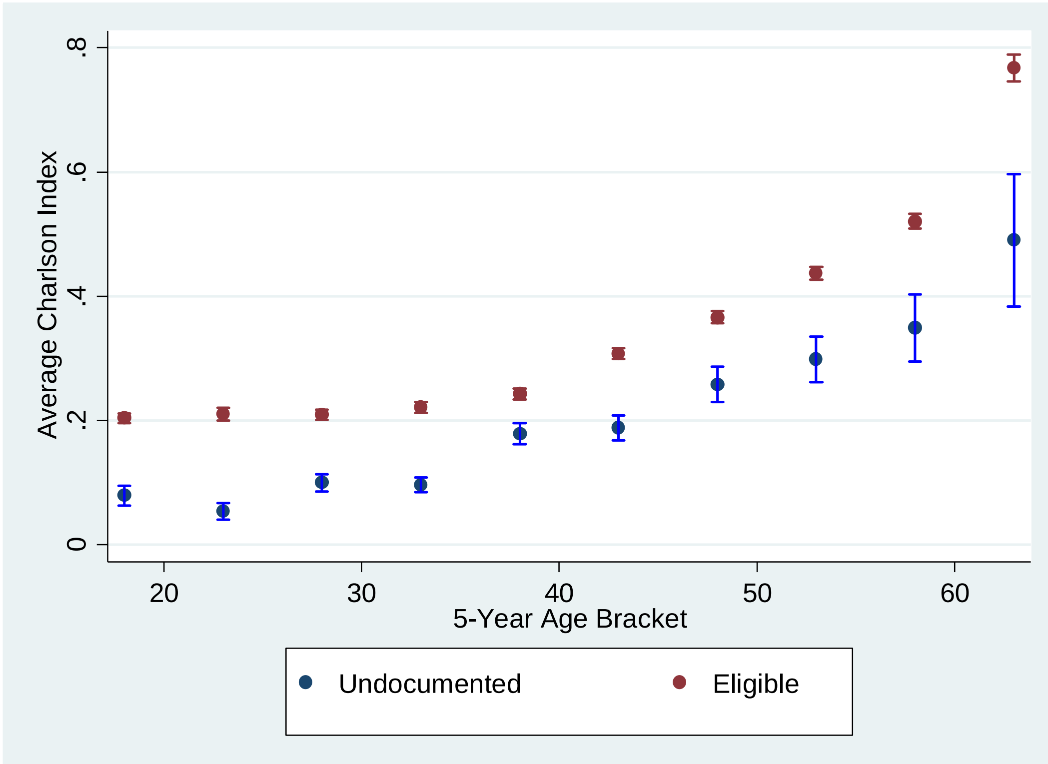


Figure A-2b: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: NHIS Sample Adult 18–64. Weighted. Uses Logit model.

Figure A-3: Charlson Index by Age for Undocumented and Legal Populations



Notes: CHIS Adult Sample, 18–64. Weighted.

Table A-2: Oaxaca Blinder

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0537***	0.0537***	0.0537***
	(1.69e-05)	(1.65e-05)	(1.65e-05)
Undocumented	0.0151***	0.0152***	0.0151***
	(2.36e-05)	(2.33e-05)	(2.32e-05)
Difference in means	0.0386***	0.0386***	0.0386***
	(2.90e-05)	(2.85e-05)	(2.85e-05)
Share due to:			
Endowments	0.00465***	0.00340***	0.00353***
	(2.30e-05)	(3.62e-05)	(3.80e-05)
Coefficients	0.0429***	0.0407***	0.0405***
	(3.42e-05)	(3.57e-05)	(3.54e-05)
Interaction	-0.00892***	-0.00555***	-0.00546***
	(2.98e-05)	(4.20e-05)	(4.34e-05)
Observations	-0.00892***	-0.00555***	-0.00546***

Notes: CHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure A-4a: Predictedtrend in Disability Rates for Eligible

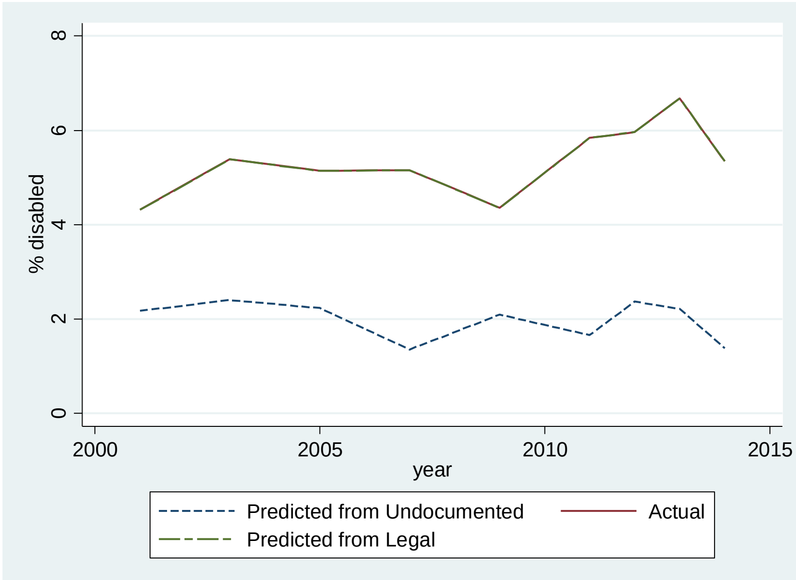
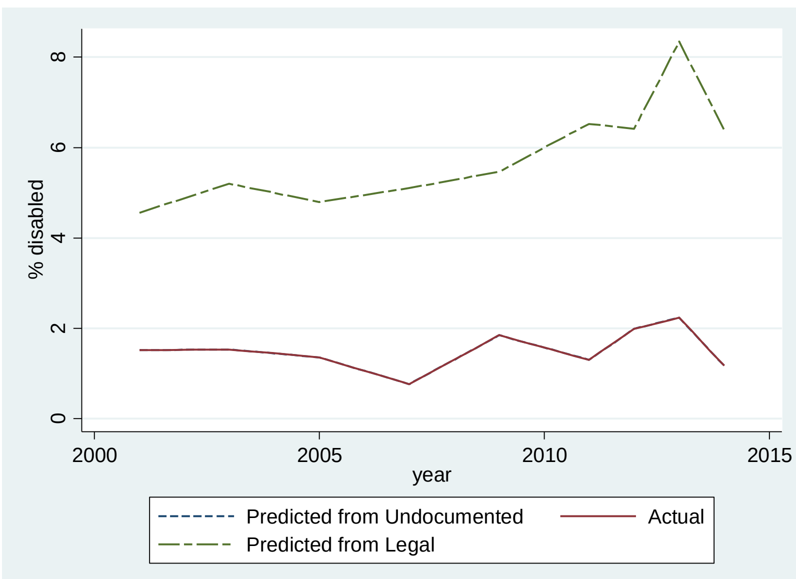


Figure A-4b: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: CHIS Adult Sample 18–64. Weighted. Uses Linear Probability (OLS) model.

Table A-3: Coefficients on Year Fixed Effects from Predicting Disability Status using Self-Reported Medical Conditions, for Eligibles

	(1)	(2)	(3)
	OLS	Probit	Logit
1999	0.00719***	0.00656***	0.00706***
	(0.00223)	(2.50e-05)	(2.56e-05)
2000	0.00563**	0.00463***	0.00477***
	(0.00219)	(2.47e-05)	(2.52e-05)
2001	-0.000872	0.000474***	0.000262***
	(0.00220)	(2.39e-05)	(2.43e-05)
2002	0.00887***	0.00775***	0.00764***
	(0.00228)	(2.49e-05)	(2.53e-05)
2003	0.00957***	0.00870***	0.00914***
	(0.00227)	(2.47e-05)	(2.52e-05)
2005	0.00756***	0.00699***	0.00709***
	(0.00230)	(2.43e-05)	(2.48e-05)
2006	0.00552**	0.00477***	0.00420***
	(0.00254)	(2.41e-05)	(2.44e-05)
2007	0.0164***	0.0150***	0.0151***
	(0.00269)	(2.51e-05)	(2.55e-05)
2008	0.0106***	0.0114***	0.0107***
	(0.00272)	(2.44e-05)	(2.47e-05)
2009	0.00793***	0.00744***	0.00732***
	(0.00257)	(2.40e-05)	(2.43e-05)
2010	0.00930***	0.00902***	0.00864***
	(0.00250)	(2.41e-05)	(2.44e-05)
2011	0.0139***	0.0115***	0.0116***
	(0.00238)	(2.43e-05)	(2.46e-05)
2012	0.0179***	0.0163***	0.0159***
	(0.00246)	(2.48e-05)	(2.51e-05)
2013	0.0221***	0.0201***	0.0201***
	(0.00254)	(2.51e-05)	(2.54e-05)
2014	0.0184***	0.0167***	0.0166***
	(0.00272)	(2.48e-05)	(2.52e-05)
2015	0.0200***	0.0176***	0.0179***
	(0.00257)	(2.49e-05)	(2.52e-05)
Observations	373,954	373,954	373,954
R-squared	0.162	0. 242	0.241

Notes: NHIS Sample Adult 18–64. 1998-2003 & 2005-2015. Weighted. Columns 2 and 3 show marginal effects. Model also includes health conditions, age category, education category, sex, and survey quarter fixed effects. Robust standard errors in parentheses. 1998 is the omitted category. *** p<0.01, ** p<0.05, * p<0.1

Table A-4: Coefficients on Year Fixed Effects from Predicting Disability Status using Self-Reported Medical Conditions, for Undocumented

	(1)	(2)	(3)
	OLS	Probit	Logit
1999	-0.00423	-0.00541***	-0.00451***
	(0.00577)	(6.79e-05)	(6.97e-05)
2000	0.000922	0.000610***	0.00164***
	(0.00589)	(7.05e-05)	(7.25e-05)
2001	-0.00969*	-0.0104***	-0.00995***
	(0.00534)	(6.10e-05)	(6.19e-05)
2002	-0.00311	-0.00624***	-0.00472***
	(0.00616)	(6.51e-05)	(6.66e-05)
2003	-0.00374	-0.00544***	-0.00456***
	(0.00556)	(6.32e-05)	(6.44e-05)
2005	-0.00745	-0.00889***	-0.00873***
	(0.00548)	(6.05e-05)	(6.09e-05)
2006	-0.00703	-0.00872***	-0.00785***
	(0.00571)	(6.05e-05)	(6.15e-05)
2007	-0.00325	-0.00459***	-0.00407***
	(0.00620)	(6.31e-05)	(6.41e-05)
2008	-0.00511	-0.00682***	-0.00629***
	(0.00675)	(6.20e-05)	(6.25e-05)
2009	-0.00918*	-0.0101***	-0.00989***
	(0.00535)	(5.96e-05)	(6.05e-05)
2010	-0.0104*	-0.0110***	-0.0108***
	(0.00536)	(5.82e-05)	(5.86e-05)
2011	-0.00222	-0.00334***	-0.00323***
	(0.00607)	(6.36e-05)	(6.37e-05)
2012	-0.00735	-0.00890***	-0.00844***
	(0.00550)	(5.96e-05)	(6.01e-05)
2013	-0.0109**	-0.0117***	-0.0117***
	(0.00544)	(5.75e-05)	(5.77e-05)
2014	-0.0105**	-0.0119***	-0.0111***
	(0.00531)	(5.73e-05)	(5.80e-05)
2015	-0.00836	-0.00807***	-0.00856***
	(0.00609)	(6.03e-05)	(6.03e-05)
Observations	30,012	30,01	30,01
R-squared	0.028	0.124	0.123

Notes: NHIS Sample Adult 18–64. 1998-2003 & 2005-2015. Weighted. Columns 2 and 3 show marginal effects. Model also includes health conditions, age category, education category, sex, and survey quarter fixed effects. Robust standard errors in parentheses. 1998 is the omitted category. *** p<0.01, ** p<0.05, * p<0.1

Appendix B: Additional Robustness Checks

This appendix contains numerous additional robustness checks, as described below.

Table B-1 repeats this analysis with a dummy variable for legal immigrants within the pooled sample and also estimating the model separated for native born and legal immigrants.

We next define a narrow comparison group: Hispanic, non-veteran, high school drop outs. This helps avoids our resulted being confounded by the Hispanic paradox of healthier of lower mortality despite lower socioeconomic status and cyclical migration, which may be selected on health (Teller and Clyburn 1974; Markides and Coreil 1986; Sorlie et al. 1993; Hayward et al. 2014; Lariscy, Hummer, and Hayward 2015; Beltrán-Sánchez et al. 2016; Antman, Duncan, and Trejo 2020). Figure B-1 (analogous to Figure 2 above) shows the average Charlson index for each age group, with the same pattern as above that undocumented immigrants are healthier are every age.

We then repeat our analysis using either of the components of the eligible sample (i.e., the native born or legal immigration population). Our results, as shown in Figures B-2, B-3, and B-4, and Table B-3, are broadly consistent.

Figure B-5 repeats Figure A-1 (native born and minor legal immigrants) for only those 18-32. This also results in us dropping the age bracket dummy variables from this entire analysis, with the exception of the one for 18-27 (vs. 28+).

We then check our results using the CPS ASEC data from IPUMS (Flood et al. 2017), as used in Borjas (2017a; 2017b). First we check the sums of survey weights for undocumented immigrants for California against the sums of survey weights for the CHIS data (Figure B-6). While the trends over time do not match (likely due the lack of sufficient family and occupation variables in the CHIS data), the levels are reasonably comparable. Repeating our results without information on specific health conditions forces us to use the single variable for self-reported health status, in addition to the demographic and time controls used above. We then used the coefficients from those two regressions to predict the two counterfactuals described above. Figure B-7 shows the results of this analysis, which are largely consistent with Figures 4 and 5 above.

Other robustness checks include: 1) Other controls including cancer type in the Oaxaca-Blinder decomposition, even though Social Security doesn't consider them when awarding disability benefits (Social Security 2017b). Table B-4 including smoking, as the Hispanic paradox of may be explained by lower smoking rates (Fenelon 2013), whereas Table B-5 includes BMI. 2) Alternative specifications and variable definitions: exclude those over 50 from the analysis in case undocumented immigrants migrate home to retire (Table B-6); interact the health condition dummies with gender (Table B-7); interact the health conditions dummies with gender, education, and age (Table B-8); use a more conservative definition of disability that excludes those only temporarily out of work (Table B-9); include those with borderline diabetes as having diabetes (Table B-10), exclude Social Security (SSI and SSDI) payments (Table B-11) or working in a licensed occupation (Table B-12) from implying legal status, include fixed effects for each age (Table B-13), incorporate the widest possible array of variables for health conditions and functional limitations, including drug and alcohol use (Table B-14), incorporate diagnosed mental health conditions variables when they become available in 2007 (Table B-15), limit the results to only those who have seen a doctor in the past year (Table B-16), and include controls for all possible martial statuses and relationships to the head of the household (Table B-17). The results are robust to all of these checks. Table B-18 shows the Oaxaca-Blinder results when estimated in three-year group and in single year. In every three-year group, a majority of the difference is explained by the coefficients, with on average 82% explained by differences in coefficients as opposed to endowments. In every single year regression, the majority of the difference is also explained by differences in coefficients, with an average of 75% explained by differences in coefficients as opposed to endowments. Finally, Table B-19 shows the results omitting spousal information from the undocumented algorithm, Table B-20 shows the results excluding individuals over age 62 (due to the potential for early claiming of Social Security), and Table B-21 shows the results controlling for whether a respondent is pregnant. The results are robust to all of these adjustments.

Additional References

Antman, Francisca M., Brian Duncan, and Stephen J. Trejo. 2020. Ethnic Attrition, Assimilation, and the Measured Health Outcomes of Mexican Americans NBER Working Paper No. 26742.

Beltrán-Sánchez, Hiram, Alberto Palloni, Fernando Riosmena, and Rebecca Wong. 2016. SES Gradients Among Mexicans in the United States and in Mexico: A New Twist to the Hispanic Paradox? Demography 53: 1555–1581.

Fenelon, Andrew. 2013. Revisiting the Hispanic mortality advantage in the United States: The role of smoking. Social Science & Medicine 82: 1–9.

Hayward, Mark D., Robert A. Hummer, Chi-Tsun Chiu, Cesar Gonzalez-Gonzalez, Rebeca Wong. 2014. Does the Hispanic Paradox in U.S. Adult Mortality Extend to Disability? Population Research and Policy Review 33:81–96.

Lariscy, Joseph T., Robert A. Hummer, and Mark D. Haywardet. 2015. Hispanic Older Adult Mortality in the United States: New Estimates and an Assessment of Factors Shaping the Hispanic Paradox. Demography 52:1–14.

Markides, Kyriakos S. and Jeannine Coreil. 1986. The Health of Hispanics in the Southwestern United States: an Epidemiological Paradox. Public Health Reports 101(3): 253–265.

Sorlie, Paul D., Eric Backlund, Norman J. Johnson, and Eugene Rogot. 1993. Mortality by Hispanic Status in the United States. Journal of the American Medical Association 270(20); 2464–2468.

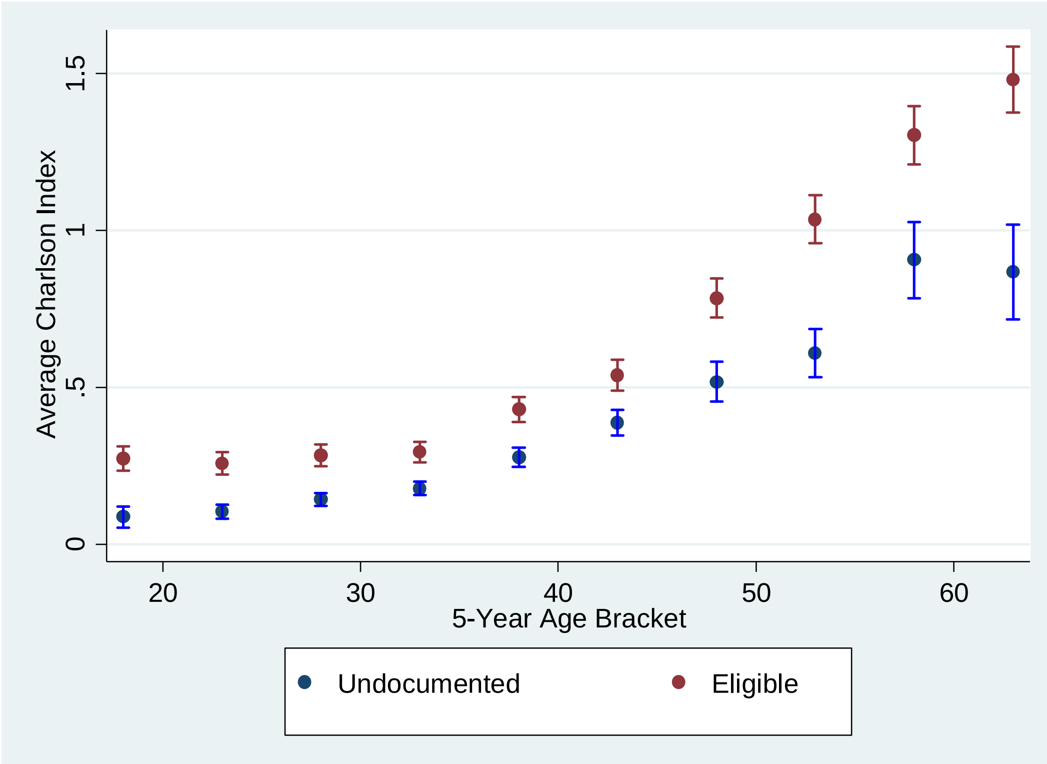
Teller, C. H. and S. Clyburn. 1974. Trends in infant mortality. Texas Business Review 29: 97–108

Table B-1: Predicting Disability Status with Medical Conditions

	(1) Native & Legal	(2) Native & Legal	(3) Native	(4) Legal	(5) Undocumented
Heart Disease	0.0490***	0.0485***	0.0510***	0.0281***	0.00653***
	(1.78e-05)	(1.77e-05)	(1.91e-05)	(4.70e-05)	(5.06e-05)
Cancer	0.0264***	0.0253***	0.0250***	0.0332***	0.0125***
	(1.48e-05)	(1.48e-05)	(1.58e-05)	(4.21e-05)	(5.12e-05)
Diabetes	0.0394***	0.0401***	0.0424***	0.0226***	0.00788***
	(1.23e-05)	(1.23e-05)	(1.34e-05)	(2.89e-05)	(2.68e-05)
Hypertension	0.0250***	0.0245***	0.0245***	0.0225***	0.00550***
	(9.51e-06)	(9.50e-06)	(1.02e-05)	(2.41e-05)	(2.19e-05)
Asthma	0.0203***	0.0193***	0.0202***	0.00832***	0.00190***
	(1.15e-05)	(1.15e-05)	(1.23e-05)	(3.58e-05)	(3.80e-05)
Emphysema	0.0524***	0.0511***	0.0534***	0.0237***	0.00887***
	(2.44e-05)	(2.44e-05)	(2.58e-05)	(9.40e-05)	(8.10e-05)
Liver disease	0.0609***	0.0614***	0.0653***	0.0299***	0.00903***
	(2.20e-05)	(2.20e-05)	(2.40e-05)	(5.18e-05)	(4.62e-05)
Joint pain	0.0284***	0.0272***	0.0275***	0.0245***	0.00329***
	(9.46e-06)	(9.46e-06)	(1.02e-05)	(2.41e-05)	(2.17e-05)
Back pain	0.0312***	0.0310***	0.0312***	0.0263***	0.0101***
	(9.70e-06)	(9.68e-06)	(1.04e-05)	(2.48e-05)	(2.11e-05)
Neck pain	0.0249***	0.0251***	0.0256***	0.0203***	0.00738***
	(1.05e-05)	(1.05e-05)	(1.13e-05)	(2.64e-05)	(2.36e-05)
Face pain	0.0273***	0.0272***	0.0286***	0.0151***	0.00864***
	(1.47e-05)	(1.46e-05)	(1.57e-05)	(3.91e-05)	(3.44e-05)
Ulcers	0.0203***	0.0199***	0.0212***	0.0101***	-0.000254***
	(1.27e-05)	(1.27e-05)	(1.36e-05)	(3.54e-05)	(3.69e-05)
Bronchitis	0.0198***	0.0190***	0.0193***	0.0153***	0.00534***
	(1.57e-05)	(1.57e-05)	(1.67e-05)	(5.10e-05)	(5.13e-05)
Legal immigrant		-0.0285***			
		(1.66e-05)			
Observations	373,954	373,954	328,065	45,889	30,012

Notes: NHIS Sample Adult 18–64. Weighted. Logit. Marginal effects. Model also includes age category, education category, sex, and survey year and survey quarter fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure B-1: Charlson Index for Narrow Comparison Group



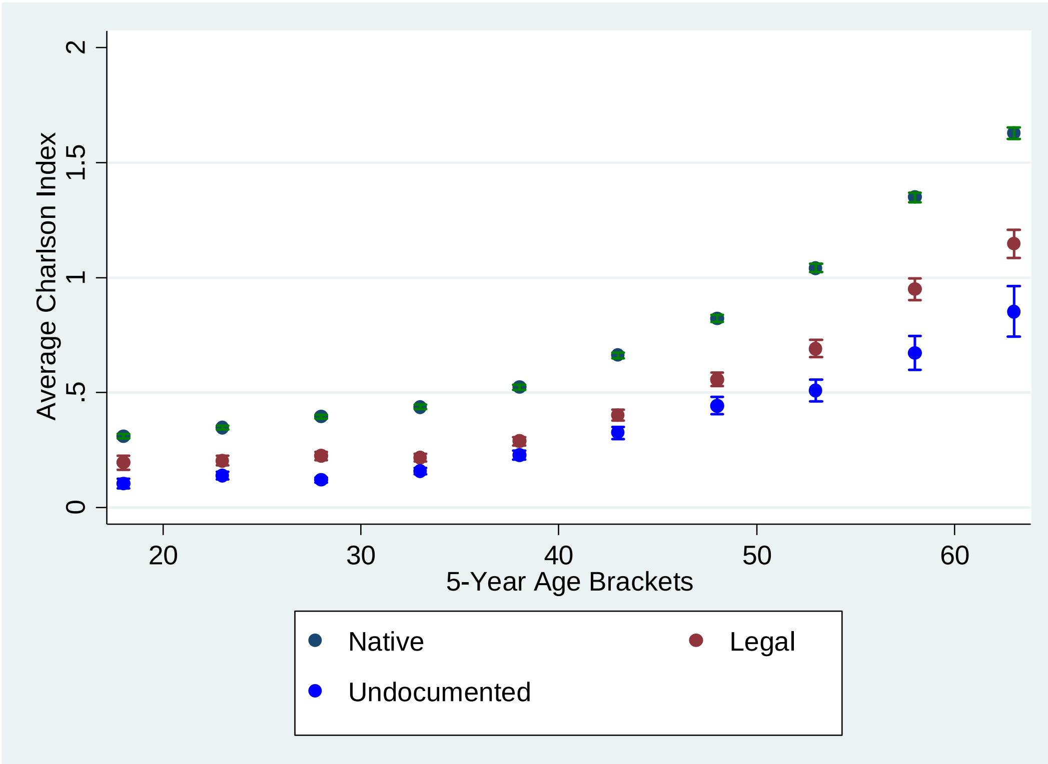
Notes: NHIS Sample Adults, 18–64, Hispanic, non-veteran, high school drop outs. Weighted. 95% confidence interval shown in whiskers around each point.

Table B-2: Oaxaca-Blinder for Narrow Comparison Group

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.103*** (0.0000328)	0.102*** (0.000031)	0.103*** (0.000031)
Undocumented	0.0188*** (0.0000156)	0.0188*** (0.0000153)	0.0188*** (0.0000152)
Difference in means	0.0842*** (0.0000364)	0.0836*** (0.0000345)	0.0842*** (0.0000345)
Share due to:			
Endowments	0.0140*** (0.0000124)	0.0170*** (0.0000257)	0.0177*** (0.0000271)
Coefficients	0.0394*** (0.0000373)	0.0418*** (0.000029)	0.0430*** (0.0000291)
Interaction	0.0308*** (0.0000218)	0.0248*** (0.000031)	0.0235*** (0.000032)
Observations	30,246	30,246	30,246

Notes: NHIS Sample Adult 18–64, Hispanic, non-veteran, high school drop outs. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure B-2: **Charlson Index by Age for Undocumented, Legal, and Native Born**



Notes: NHIS Adult Sample, 18–64. Weighted.

Table B-3a Oaxaca Blinder for Native Born vs. Undocumented Immigrants

	(1) OLS	(2) Probit	(3) Logit
Means:			
Native Born	0.0731***	0.0729***	0.0731***
	(0.00000515)	(0.00000482)	(0.00000481)
Undocumented	0.0138***	0.0138***	0.0138***
	(0.00000865)	(0.00000854)	(0.0000085)
Difference in means	0.0593***	0.0591***	0.0593***
	(0.0000101)	(0.0000098)	(0.00000977)
Share due to:			
Endowments	0.0119***	0.0117***	0.0116***
	(0.0000114)	(0.0000234)	(0.0000249)
Coefficients	0.0587***	0.0516***	0.0503***
	(0.0000129)	(0.0000127)	(0.0000123)
Interaction	-0.0113***	-0.00422***	-0.00256***
	(0.0000142)	(0.0000248)	(0.0000261)
Observations	358,077	358,077	358,077

Notes: NHIS Sample Adult 18–64. Native Born and Undocumented only. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-3b: Oaxaca Blinder for Legal vs. Undocumented Immigrants

	(1) OLS	(2) Probit	(3) Logit
Means:			
Legal Immigrants	0.0440***	0.0438***	0.0440***
	(0.0000116)	(0.0000109)	(0.0000109)
Undocumented	0.0138***	0.0138***	0.0138***
	(0.00000865)	(0.00000854)	(0.0000085)
Difference in means	0.0302***	0.0301***	0.0302***
	(0.0000144)	(0.0000139)	(0.0000138)
Share due to:			
Endowments	0.00671***	0.00650***	0.00645***
	(0.00000771)	(0.0000116)	(0.000012)
Coefficients	0.0201***	0.0180***	0.0177***
	(0.0000168)	(0.0000142)	(0.000014)
Interaction	0.00341***	0.00560***	0.00608***
	(0.0000124)	(0.0000142)	(0.0000143)
Observations	75,901	75,901	75,901

Notes: NHIS Sample Adult 18–64. Weighted. Legal immigrants and Undocumented only. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure B-3a: Predicted trend in Disability Rates for Native Born

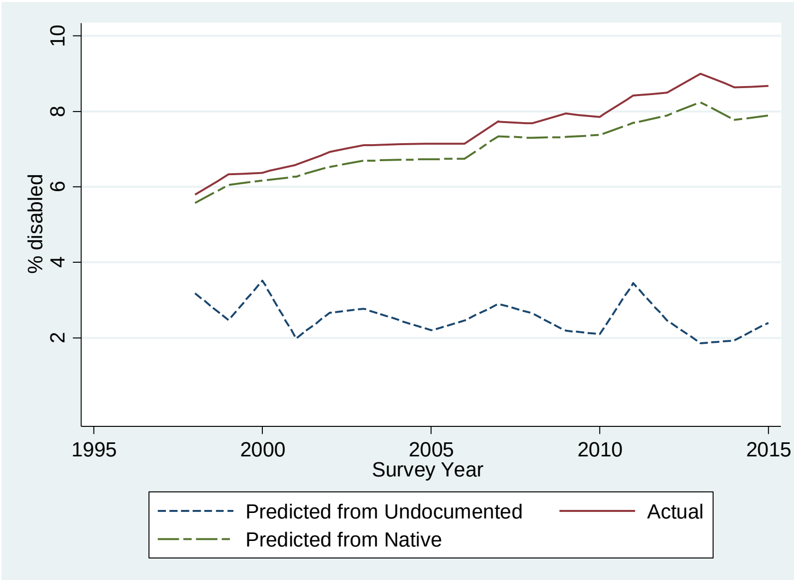
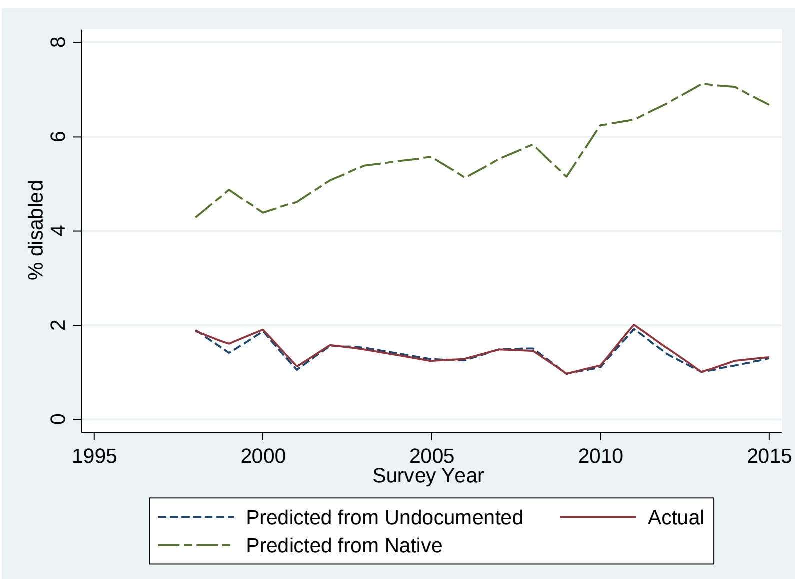


Figure B-3b: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: NHIS Sample Adult 18–64. Native Born and Undocumented only. Weighted. Uses Logit model.

Figure B-4a: Predicted trend in Disability Rates for Legal Immigrants

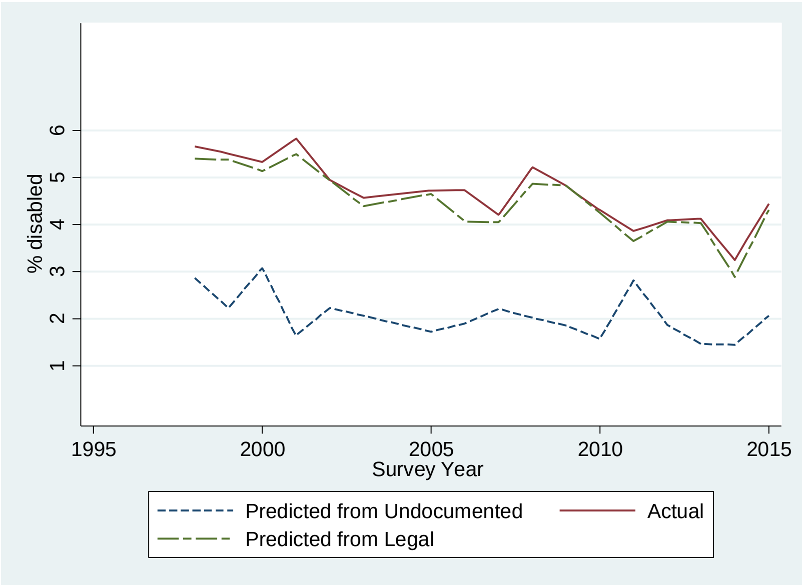
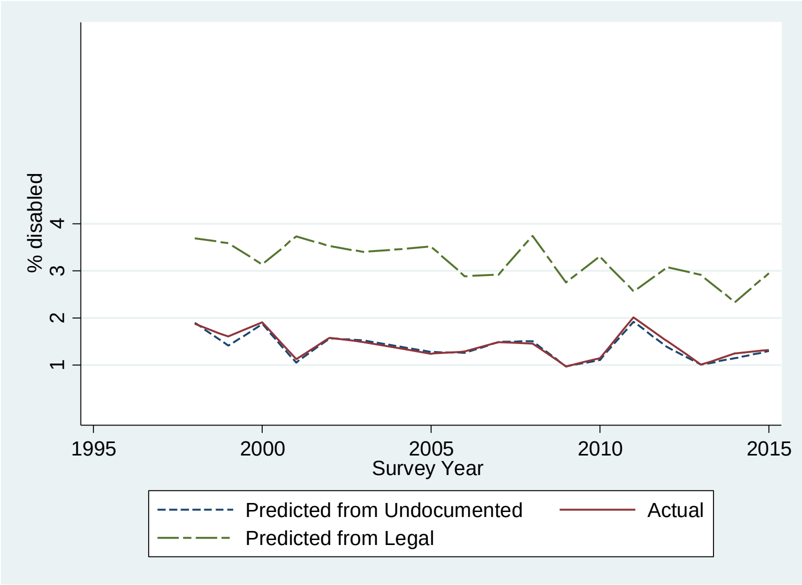


Figure B-4b: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: NHIS Sample Adult 18–64. Legal immigrants and Undocumented only. Weighted. Uses Logit model.

Figure B-5a: Predicted trend in Disability Rates for Native Born (Ages 18-32)

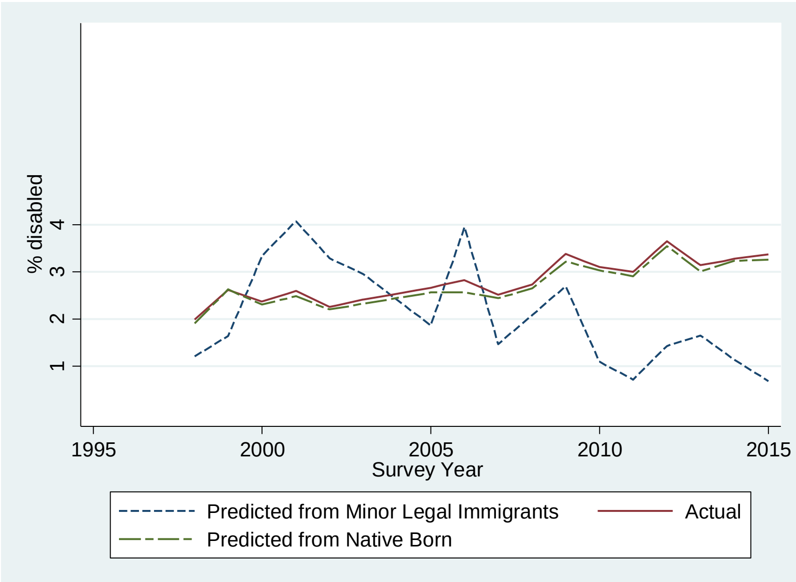
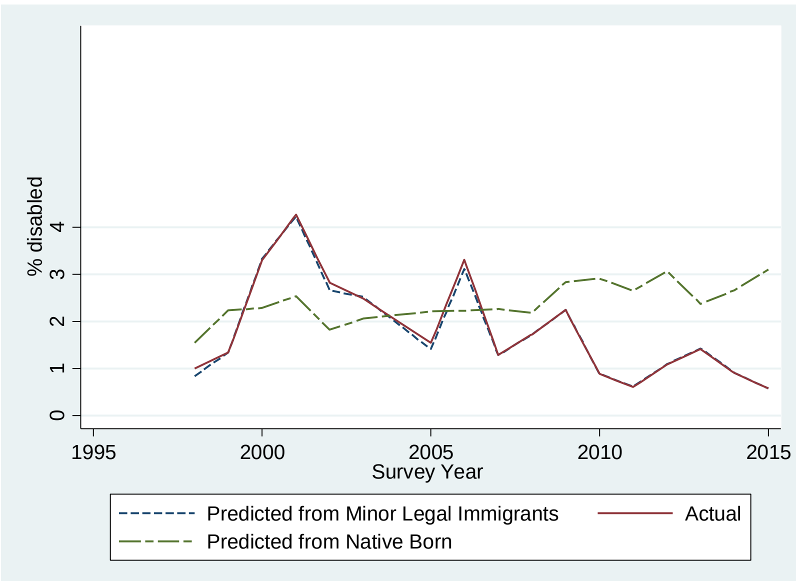


Figure B-5b: Predicted trend in Disability Rates for Minor Legal Immigrant (Ages 18-32)



Notes: NHIS Sample Adult 18-32. Weighted. Uses Logit model.

Figure B-6: California undocumented population from CPS and CHIS

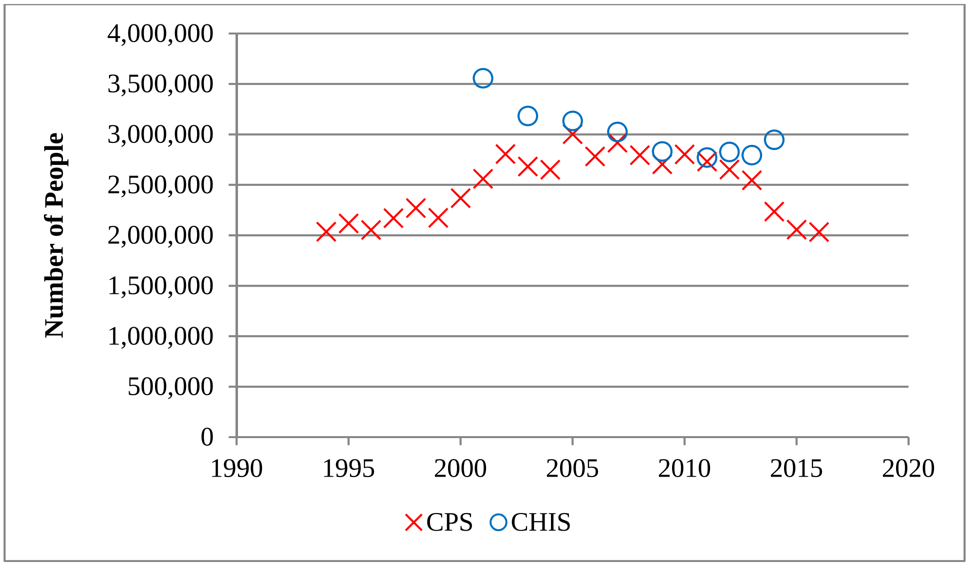


Figure B-7a: Predicted Trend in Disability Rates for Eligible

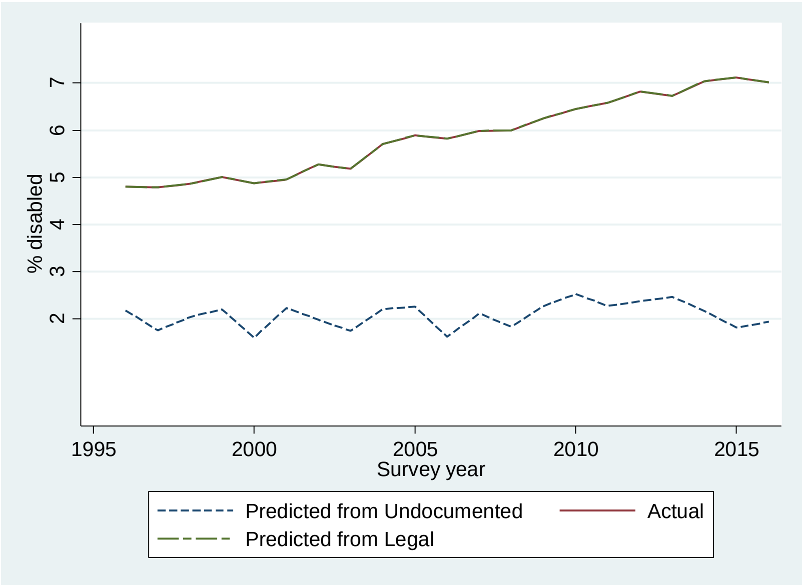
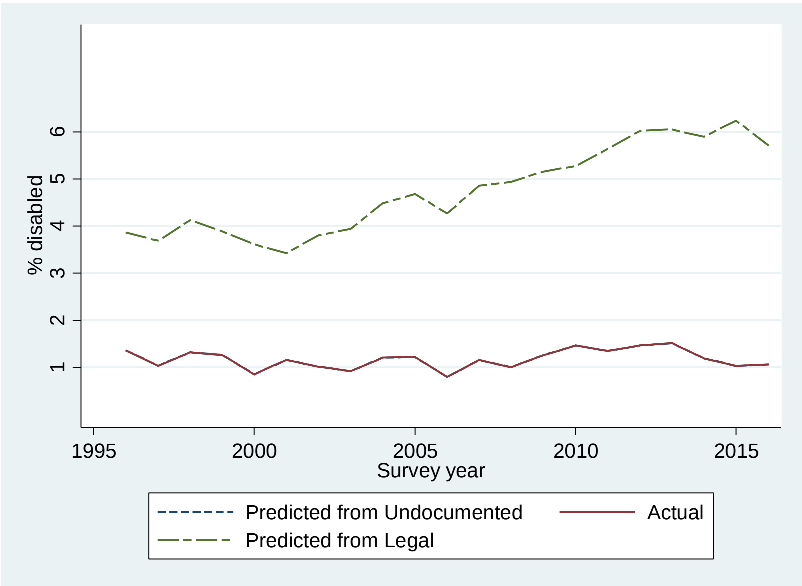


Figure B-7b: Predicted Trend in Disability Rates for Undocumented Immigrants



Notes: CPS 18–64. Weighted. Uses Logit model.

Table B-4: Oaxaca-Blinder Including Smoking Control

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0698*** (0.00000477)	0.0696*** (0.00000447)	0.0698*** (0.00000446)
Undocumented	0.0138*** (0.00000869)	0.0138*** (0.00000858)	0.0138*** (0.00000854)
Difference in means	0.0559*** (0.00000991)	0.0558*** (0.00000967)	0.0559*** (0.00000964)
Share due to:			
Endowments	0.0110*** (0.0000112)	0.0106*** (0.0000222)	0.0104*** (0.0000235)
Coefficients	0.0488*** (0.0000122)	0.0421*** (0.0000116)	0.0409*** (0.0000113)
Interaction	-0.00391*** (0.0000135)	0.00301*** (0.0000233)	0.00459*** (0.0000243)
Observations	401,538	401,538	401,538

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-5: Oaxaca-Blinder Including BMI Control

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699***	0.0697***	0.0699***
	(0.00000476)	(0.00000446)	(0.00000445)
Undocumented	0.0138***	0.0138***	0.0138***
	(0.00000865)	(0.00000854)	(0.0000085)
Difference in means	0.0561***	0.0560***	0.0561***
	(0.00000987)	(0.00000963)	(0.0000096)
Share due to:			
Endowments	0.0114***	0.0113***	0.0113***
	(0.0000109)	(0.0000222)	(0.0000237)
Coefficients	0.0505***	0.0449***	0.0439***
	(0.0000121)	(0.0000116)	(0.0000113)
Interaction	-0.00575***	-0.000251***	0.000952***
	(0.0000131)	(0.0000232)	(0.0000245)
Observations	403,966	403,966	403,966

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-6: Oaxaca-Blinder Excluding Those over 50

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0476*** (0.00000464)	0.0475*** (0.00000443)	0.0476*** (0.00000441)
Undocumented	0.0109*** (0.00000815)	0.0109*** (0.00000807)	0.0109*** (0.00000804)
Difference in means	0.0367*** (0.00000938)	0.0367*** (0.00000921)	0.0367*** (0.00000917)
Share due to:			
Endowments	0.00608*** (0.00000893)	0.00606*** (0.0000162)	0.00612*** (0.0000173)
Coefficients	0.0408*** (0.0000113)	0.0381*** (0.0000116)	0.0373*** (0.0000114)
Interaction	-0.0102*** (0.0000111)	-0.00745*** (0.0000177)	-0.00669*** (0.0000186)
Observations	296,561	296,561	296,561

Notes: NHIS Sample Adult 18-50. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-7: Oaxaca-Blinder Including Gender Interacted with Conditions

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699***	0.0697***	0.0699***
	(4.76e-06)	(4.46e-06)	(4.45e-06)
Undocumented	0.0138***	0.0138***	0.0138***
	(8.65e-06)	(8.51e-06)	(8.47e-06)
Difference in means	0.0561***	0.0560***	0.0561***
	(9.87e-06)	(9.61e-06)	(9.57e-06)
Share due to:			
Endowments	0.0103***	0.00962***	0.00935***
	(1.10e-05)	(2.07e-05)	(2.19e-05)
Coefficients	0.0504***	0.0445***	0.0436***
	(1.21e-05)	(1.16e-05)	(1.13e-05)
Interaction	-0.00457***	0.00179***	0.00322***
	(1.32e-05)	(2.18e-05)	(2.29e-05)
Observations	403,966	403,966	403,966

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-8: Oaxaca-Blinder with Gender, Age, and Education Interacted with Conditions

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699***	0.0694***	0.0699***
	(4.76e-06)	(4.35e-06)	(4.36e-06)
Undocumented	0.0138***	0.0139***	0.0138***
	(8.65e-06)	(8.46e-06)	(8.43e-06)
Difference in means	0.0561***	0.0555***	0.0561***
	(9.87e-06)	(9.51e-06)	(9.49e-06)
Share due to:			
Endowments	0.00994***	0.0118***	0.0124***
	(1.26e-05)	(3.16e-05)	(3.46e-05)
Coefficients	0.0305***	0.0295***	0.0303***
	(1.10e-05)	(9.60e-06)	(9.55e-06)
Interaction	0.0157***	0.0142***	0.0135***
	(1.37e-05)	(3.19e-05)	(3.47e-05)
Observations	403,966	403,791	403,791

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. A few observations in the undocumented population were dropped from Columns (2) and (3) as the interacted fixed effects perfectly predicted disability. *** p<0.01, ** p<0.05, * p<0.1

Table B-9: Oaxaca-Blinder Excluding Temporarily Out of Work (2001 Onward)

	(1)	(2)	(3)	(4)	(5)	(6)
		Disabled			Disabled (excluding temporary)	
	OLS	Probit	Logit	OLS	Probit	Logit
Means:						
Eligible (Legal Immigrants & Native Born)	0.0720***	0.0718***	0.0720***	0.0606***	0.0604***	0.0606***
	(0.00000527)	(0.00000494)	(0.00000493)	(0.00000487)	(0.00000454)	(0.00000452)
Undocumented	0.0132***	0.0132***	0.0132***	0.00699***	0.00700***	0.00699***
	(0.00000913)	(0.00000901)	(0.00000897)	(0.00000666)	(0.00000657)	(0.00000652)
Difference in means	0.0588***	0.0586***	0.0588***	0.0536***	0.0534***	0.0536***
	(0.0000105)	(0.0000103)	(0.0000102)	(0.00000825)	(0.00000799)	(0.00000793)
Share due to:						
Endowments	0.0115***	0.0113***	0.0113***	0.00804***	0.00727***	0.00749***
	(0.0000117)	(0.0000242)	(0.0000258)	(0.00000852)	(0.0000205)	(0.0000221)
Coefficients	0.0526***	0.0467***	0.0457***	0.0496***	0.0435***	0.0424***
	(0.0000131)	(0.0000125)	(0.0000122)	(0.0000111)	(0.0000104)	(0.00001)
Interaction	-0.00523***	0.000597***	0.00178***	-0.00403***	0.00266***	0.00378***
	(0.0000142)	(0.0000253)	(0.0000267)	(0.0000116)	(0.0000216)	(0.0000231)
Observations	328,536	328,536	328,536	328,536	328,536	328,536

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-10: Oaxaca-Blinder Considering Borderline Diabetes as Having Diabetes

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0709*** (0.00000477)	0.0707*** (0.00000447)	0.0709*** (0.00000446)
Undocumented	0.0138*** (0.00000864)	0.0138*** (0.00000853)	0.0138*** (0.0000085)
Difference in means	0.0571*** (0.00000987)	0.0569*** (0.00000963)	0.0571*** (0.0000096)
Share due to:			
Endowments	0.0116*** (0.0000108)	0.0114*** (0.000022)	0.0113*** (0.0000233)
Coefficients	0.0507*** (0.0000121)	0.0450*** (0.0000115)	0.0440*** (0.0000113)
Interaction	-0.00519*** (0.0000131)	0.000452*** (0.000023)	0.00173*** (0.0000242)
Observations	407,829	407,829	407,829

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-11: Oaxaca-Blinder Excluding Social Security from Implying Legal Status

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699*** (0.00000476)	0.0697*** (0.00000446)	0.0699*** (0.00000445)
Undocumented	0.0145*** (0.00000885)	0.0145*** (0.00000872)	0.0145*** (0.00000869)
Difference in means	0.0554*** (0.0000101)	0.0552*** (0.0000098)	0.0554*** (0.00000976)
Share due to:			
Endowments	0.0124*** (0.0000111)	0.0121*** (0.0000226)	0.0119*** (0.0000238)
Coefficients	0.0496*** (0.0000122)	0.0439*** (0.0000117)	0.0430*** (0.0000114)
Interaction	-0.00654*** (0.0000133)	-0.000826*** (0.0000236)	0.000559*** (0.0000247)
Observations	403,966	403,966	403,966

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-12: Oaxaca-Blinder Excluding Licensed Occupations from Implying Legal Status

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0701*** (0.00000477)	0.0699*** (0.00000447)	0.0701*** (0.00000446)
Undocumented	0.0134*** (0.00000841)	0.0134*** (0.0000083)	0.0134*** (0.00000827)
Difference in means	0.0566*** (0.00000967)	0.0564*** (0.00000943)	0.0566*** (0.00000939)
Share due to:			
Endowments	0.0111*** (0.0000104)	0.0109*** (0.0000211)	0.0107*** (0.0000224)
Coefficients	0.0488*** (0.0000118)	0.0437*** (0.0000112)	0.0428*** (0.000011)
Interaction	-0.00327*** (0.0000126)	0.00185*** (0.0000221)	0.00310*** (0.0000232)
Observations	403,966	403,966	403,966

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-13: Oaxaca-Blinder with Fixed Effects for Each Age

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699***	0.0697***	0.0699***
	(4.76e-06)	(4.46e-06)	(4.45e-06)
Undocumented	0.0138***	0.0138***	0.0138***
	(8.65e-06)	(8.53e-06)	(8.50e-06)
Difference in means	0.0561***	0.0560***	0.0561***
	(9.87e-06)	(9.63e-06)	(9.59e-06)
Share due to:			
Endowments	0.0110***	0.0106***	0.0104***
	(1.10e-05)	(2.18e-05)	(2.29e-05)
Coefficients	0.0518***	0.0453***	0.0441***
	(1.22e-05)	(1.17e-05)	(1.14e-05)
Interaction	-0.00656***	7.80e-05***	0.00169***
	(1.33e-05)	(2.29e-05)	(2.38e-05)
Observations	403,966	403,966	403,966

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-14: Oaxaca-Blinder Including Additional Health & Limitation Variables

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699*** (4.76e-06)	0.0682*** (4.15e-06)	0.0699*** (4.17e-06)
Undocumented	0.0138*** (8.65e-06)	0.0137*** (8.07e-06)	0.0138*** (8.09e-06)
Difference in means	0.0561*** (9.87e-06)	0.0545*** (9.07e-06)	0.0561*** (9.11e-06)
Share due to:			
Endowments	0.0178*** (1.13e-05)	0.0193*** (2.47e-05)	0.0187*** (2.55e-05)
Coefficients	0.0420*** (1.14e-05)	0.0347*** (1.06e-05)	0.0349*** (1.04e-05)
Interaction	-0.00364*** (1.30e-05)	0.000500*** (2.54e-05)	0.00255*** (2.61e-05)
Observations	403,966	403,937	403,937

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. A few observations in the undocumented population were dropped from Columns (2) and (3) as some of the additional variables perfectly predicted disability. *** p<0.01, ** p<0.05, * p<0.1.

Table B-15: Oaxaca-Blinder Including Diagnosed Mental Conditions (2007-2015)

	(1)	(2)	(3)	(4)	(5)	(6)
	2007, Many Conditions			2007 & 2012; Only Bipolar		
	OLS	Probit	Logit	OLS	Probit	Logit
Means:						
Eligible (Legal Immigrants & Native Born)	0.0715***	0.0710***	0.0715***	0.0749***	0.0746***	0.0749***
	(1.97e-05)	(1.78e-05)	(1.78e-05)	(1.41e-05)	(1.30e-05)	(1.30e-05)
Undocumented	0.0150***	0.0157***	0.0156***	0.0144***	0.0144***	0.0144***
	(3.50e-05)	(3.55e-05)	(3.53e-05)	(2.45e-05)	(2.40e-05)	(2.39e-05)
Difference in means	0.0566***	0.0553***	0.0559***	0.0605***	0.0602***	0.0605***
	(4.02e-05)	(3.97e-05)	(3.95e-05)	(2.83e-05)	(2.73e-05)	(2.72e-05)
Share due to:						
Endowments	0.0148***	0.0239***	0.0230***	0.0232***	0.0263***	0.0290***
	(6.23e-05)	(0.000106)	(9.93e-05)	(3.84e-05)	(9.99e-05)	(0.000102)
Coefficients	0.0360***	0.0242***	0.0241***	0.0464***	0.0408***	0.0401***
	(4.77e-05)	(4.35e-05)	(4.25e-05)	(3.40e-05)	(3.21e-05)	(3.13e-05)
Interaction	0.00586***	0.00725***	0.00882***	-0.00907***	-0.00688***	-0.00862***
	(6.83e-05)	(0.000109)	(0.000102)	(4.33e-05)	(0.000102)	(0.000103)
Observations	18,359	18,292	18,292	-0.00907***	-0.00688***	-0.00862***

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. Columns (1)-(3) include controls for being diagnosed by a provider with phobias or fears, Attention Deficit Disorder or hyperactivity, autism, bipolar disorder, dementia, mania or psychosis, schizophrenia, or seizures. Columns (4)-(6) only include a control for bipolar disorder. A few observations in the undocumented population were dropped from Columns (2), (3), (5), and (6) as some of the additional variables perfectly predicted disability. *** p<0.01, ** p<0.05, * p<0.1

Table B-16: Oaxaca-Blinder for Those Who Have Seen a Doctor in the Past Year

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0883*** (0.0000066)	0.0881*** (0.0000062)	0.0883*** (0.00000618)
Undocumented	0.0200*** (0.0000165)	0.0200*** (0.0000163)	0.0200*** (0.0000162)
Difference in means	0.0683*** (0.0000178)	0.0681*** (0.0000174)	0.0683*** (0.0000173)
Share due to:			
Endowments	0.0105*** (0.0000188)	0.00861*** (0.0000316)	0.00818*** (0.0000328)
Coefficients	0.0639*** (0.0000214)	0.0594*** (0.0000204)	0.0584*** (0.0000202)
Interaction	-0.00609*** (0.0000225)	5.75e-05* (0.0000335)	0.00178*** (0.0000345)
Observations	251,386	251,386	251,386

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B-17: Oaxaca-Blinder Decomposition Including Family Controls

	(1)	(2)	(3)
	OLS	Probit	Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699***	0.0697***	0.0699***
	(4.76e-06)	(4.44e-06)	(4.42e-06)
Undocumented	0.0138***	0.0138***	0.0138***
	(8.65e-06)	(8.54e-06)	(8.50e-06)
Difference in means	0.0561***	0.0559***	0.0561***
	(9.87e-06)	(9.62e-06)	(9.58e-06)
Share due to:			
Endowments	0.0113***	0.0114***	0.0110***
	(1.22e-05)	(2.33e-05)	(2.46e-05)
Coefficients	0.0493***	0.0410***	0.0405***
	(1.26e-05)	(1.18e-05)	(1.16e-05)
Interaction	-0.00445***	0.00352***	0.00461***
	(1.47e-05)	(2.44e-05)	(2.55e-05)
Observations	403,966	403,939	403,939

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. All regressions include dummy variables for every possible value for marital status and relationship to the head of the household. A few observations in the undocumented population were dropped from Columns (2) and (3) as some of the additional variables perfectly predicted disability. *** p<0.01, ** p<0.05, * p<0.1

Table B-18: Oaxaca-Blinder Decompositions Over Time

Panel A: Three Years at a Time

	(1)	(2)	(3)	(4)	(5)	(6)
Years	1998-2000	2001-2003	2005-2007	2008-2010	2011-2013	2014-2015
Means:						
Eligible (Legal	0.0591***	0.0643***	0.0684***	0.0723***	0.0784***	0.0775***
Immigrants &	(1.02e-05)	(1.05e-05)	(1.05e-05)	(1.06e-05)	(1.08e-05)	(1.32e-05)
Native Born)						
Undocumented	0.0173***	0.0139***	0.0135***	0.0121***	0.0142***	0.0123***
	(2.55e-05)	(2.02e-05)	(1.87e-05)	(1.84e-05)	(1.99e-05)	(2.23e-05)
Difference in	0.0417***	0.0504***	0.0550***	0.0602***	0.0641***	0.0652***
means	(2.75e-05)	(2.28e-05)	(2.14e-05)	(2.12e-05)	(2.26e-05)	(2.59e-05)
Share due to:						
Endowments	0.00955***	0.0181***	0.0295***	0.000242***	0.00623***	0.00522***
	(5.74e-05)	(6.80e-05)	(6.88e-05)	(3.02e-05)	(4.81e-05)	(4.75e-05)
Coefficients	0.0289***	0.0369***	0.0408***	0.0445***	0.0522***	0.0487***
	(2.98e-05)	(2.60e-05)	(2.51e-05)	(2.55e-05)	(2.78e-05)	(3.11e-05)
Interaction	0.00334***	-0.00456***	-0.0153***	0.0154***	0.00573***	0.0113***
	(5.89e-05)	(6.94e-05)	(7.02e-05)	(3.40e-05)	(5.10e-05)	(5.10e-05)
	0.00334***	-0.00456***	-0.0153***	0.0154***	0.00573***	0.0113***
Observations	75,430	75,339	62,260	60,010	78,537	52,248

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. All columns use logit specification. *** p<0.01, ** p<0.05, * p<0.1.

Panel B: One Year at a Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years	1998	1999	2000	2001	2002	2003	2005	2006	2007
Means:									
Eligible (Legal Immigrants & Native Born)	0.0555***	0.0604***	0.0613***	0.0622***	0.0647***	0.0659***	0.0667***	0.0663***	0.0723***
	(0.0000173)	(0.0000179)	(0.0000179)	(0.000018)	(0.0000182)	(0.0000182)	(0.0000181)	(0.0000178)	(0.0000185)
Undocumented	0.0195***	0.0146***	0.0191***	0.0106***	0.0158***	0.0154***	0.0129***	0.0183***	0.0154***
	(0.0000482)	(0.0000513)	(0.0000436)	(0.0000331)	(0.0000353)	(0.0000352)	(0.0000378)	(0.0000432)	(0.000035)
Difference in means	0.0361***	0.0458***	0.0422***	0.0516***	0.0490***	0.0505***	0.0538***	0.0480***	0.0569***
	(0.0000513)	(0.0000543)	(0.0000471)	(0.0000377)	(0.0000397)	(0.0000396)	(0.0000419)	(0.0000468)	(0.0000396)
Share due to:									
Endowments	0.00743***	0.0127***	0.0133***	0.0160***	0.0166***	0.0235***	0.0189***	0.0463***	0.0236***
	(0.0000727)	(0.000106)	(0.000143)	(0.000105)	(0.0000951)	(0.000119)	(0.000126)	(0.000202)	(0.000102)
Coefficients	0.0221***	0.0308***	0.0230***	0.0380***	0.0347***	0.0363***	0.0410***	0.0457***	0.0332***
	(0.0000552)	(0.0000527)	(0.0000497)	(0.0000435)	(0.0000444)	(0.0000459)	(0.0000449)	(0.0000579)	(0.000044)
Interaction	0.00659***	0.00230***	0.00592***	-0.00237***	-0.00234***	-0.00927***	-0.00606***	-0.0440***	0.000055
	(0.0000763)	(0.000106)	(0.000144)	(0.000107)	(0.0000978)	(0.000122)	(0.000127)	(0.000205)	(0.000104)
Observations	25,468	24,231	25,599	26,488	24,477	24,352	24,711	18,652	18,350

	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Years	2008	2009	2010	2011	2012	2013	2014	2015
Means:								
Eligible (Legal	0.0717***	0.0727***	0.0725***	0.0760***	0.0777***	0.0814***	0.0765***	0.0784***
Immigrants &	(0.0000187)	(0.0000183)	(0.0000183)	(0.0000185)	(0.0000187)	(0.000019)	(0.0000186)	(0.0000185)
Native Born)	0.0185***	0.0105***	0.0114***	0.0193***	0.0139***	0.0103***	0.0115***	0.0161***
Undocumented	(0.000043)	(0.0000304)	(0.0000383)	(0.0000411)	(0.0000331)	(0.0000289)	(0.0000278)	(0.0000406)
Difference in	0.0532***	0.0622***	0.0611***	0.0567***	0.0639***	0.0711***	0.0650***	0.0624***
means	(0.0000469)	(0.0000355)	(0.0000424)	(0.000045)	(0.000038)	(0.0000346)	(0.0000335)	(0.0000446)
Share due to:								
Endowments	0.00698***	0.00356***	0.00699***	0.00407***	0.0137***	0.00508***	0.0178***	0.00103***
	(0.000103)	(0.0000645)	(0.0000881)	(0.0000766)	(0.00012)	(0.0000667)	(0.0000985)	(0.0000472)
Coefficients	0.0395***	0.0386***	0.0434***	0.0427***	0.0518***	0.0579***	0.0532***	0.0404***
	(0.000056)	(0.0000421)	(0.0000439)	(0.0000512)	(0.000047)	(0.0000455)	(0.0000427)	(0.0000511)
Interaction	0.00671***	0.0201***	0.0107***	0.00993***	-0.00171***	0.00809***	-0.00601***	0.0209***
	(0.000108)	(0.0000695)	(0.0000893)	(0.0000809)	(0.000123)	(0.0000736)	(0.000102)	(0.0000547)
Observations	16,749	21,689	21,227	25,566	26,641	26,271	27,466	24,470

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. All columns use logit specification. *** p<0.01, ** p<0.05, * p<0.1.

Table B-19: Oaxaca-Blinder Excluding Spousal Information from Implying Legal Status

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0707*** (0.00000482)	0.0705*** (0.00000452)	0.0707*** (0.0000045)
Undocumented	0.0143*** (0.000008)	0.0144*** (0.00000787)	0.0143*** (0.00000783)
Difference in means	0.0563*** (0.00000934)	0.0561*** (0.00000907)	0.0563*** (0.00000903)
Share due to:			
Endowments	0.0115*** (0.00000928)	0.0114*** (0.0000183)	0.0114*** (0.0000192)
Coefficients	0.0515*** (0.0000114)	0.0459*** (0.000011)	0.0449*** (0.0000107)
Interaction	-0.00675*** (0.0000116)	-0.00111*** (0.0000195)	4.50e-05** (0.0000202)
Observations	403,966	403,966	403,966

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-20: Oaxaca-Blinder Excluding Individuals Above Age 62

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0665*** (0.00000477)	0.0664*** (0.00000447)	0.0665*** (0.00000445)
Undocumented	0.0135*** (0.00000861)	0.0135*** (0.00000849)	0.0135*** (0.00000846)
Difference in means	0.0530*** (0.00000984)	0.0529*** (0.0000096)	0.0530*** (0.00000955)
Share due to:			
Endowments	0.0107*** (0.0000105)	0.0105*** (0.0000212)	0.0104*** (0.0000225)
Coefficients	0.0489*** (0.000012)	0.0437*** (0.0000116)	0.0427*** (0.0000113)
Interaction	-0.00655*** (0.0000127)	-0.00133*** (0.0000223)	-0.000144*** (0.0000234)
Observations	384,204	384,204	384,204

Notes: NHIS Sample Adult 18–62. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B-21: Oaxaca-Blinder Including a Variable for Pregnant

	(1) OLS	(2) Probit	(3) Logit
Means:			
Eligible (Legal Immigrants & Native Born)	0.0699*** (0.00000476)	0.0697*** (0.00000447)	0.0699*** (0.00000445)
Undocumented	0.0138*** (0.00000865)	0.0138*** (0.00000854)	0.0138*** (0.00000851)
Difference in means	0.0562*** (0.00000987)	0.0560*** (0.00000964)	0.0562*** (0.0000096)
Share due to:			
Endowments	0.0114*** (0.0000109)	0.0113*** (0.0000221)	0.0111*** (0.0000235)
Coefficients	0.0502*** (0.0000121)	0.0445*** (0.0000115)	0.0436*** (0.0000113)
Interaction	-0.00547*** (0.0000131)	0.000212*** (0.0000231)	0.00146*** (0.0000243)
Observations	403,490	403,490	403,490

Notes: NHIS Sample Adult 18–64. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix C: Mathematical Details for the Oaxaca-Blinder Decomposition

In the case of a linear model, if

$$\bar{y}_L = \boldsymbol{\beta}_L \bar{x}_L$$
$$\bar{y}_U = \boldsymbol{\beta}_U \bar{x}_U$$

meaning that average outcome variable for each group (legal/native born and undocumented) equals the coefficient vector for that group times the vector of the average of each covariate.[29] We can then subtract one equation from the other, add zero, rearrange, and factor:

$$\begin{aligned}
\bar{y}_L - \bar{y}_U &= \boldsymbol{\beta}_L \bar{x}_L - \boldsymbol{\beta}_U \bar{x}_U \\
\bar{y}_L - \bar{y}_U &= \boldsymbol{\beta}_L \bar{x}_L - \boldsymbol{\beta}_U \bar{x}_U + (\boldsymbol{\beta}_U \bar{x}_L - \boldsymbol{\beta}_U \bar{x}_L) + (\boldsymbol{\beta}_U \bar{x}_U - \boldsymbol{\beta}_U \bar{x}_U) + (\boldsymbol{\beta}_L \bar{x}_U - \boldsymbol{\beta}_L \bar{x}_U) \\
\bar{y}_L - \bar{y}_U &= \boldsymbol{\beta}_U \bar{x}_L - \boldsymbol{\beta}_U \bar{x}_U + \boldsymbol{\beta}_L \bar{x}_U - \boldsymbol{\beta}_U \bar{x}_U + \boldsymbol{\beta}_L \bar{x}_L - \boldsymbol{\beta}_L \bar{x}_U - \boldsymbol{\beta}_U \bar{x}_L + \boldsymbol{\beta}_U \bar{x}_U \\
\bar{y}_L - \bar{y}_U &= \boldsymbol{\beta}_U (\bar{x}_L - \bar{x}_U) + \bar{x}_U (\boldsymbol{\beta}_L - \boldsymbol{\beta}_U) + (\bar{x}_L - \bar{x}_U)(\boldsymbol{\beta}_L - \boldsymbol{\beta}_U)
\end{aligned}$$

The first set of terms on the right-hand side of the equation tells us how much of the difference is due to the difference in covariates. This is part of the difference that is explained by observables, i.e., the difference in disability rates we would expect based on how the groups' underlying difference. The second set of terms tells us how much of the difference is due to the difference in coefficients, given the actual covariate levels of the higher group. This is the part of the difference due to unobservable factors, i.e., the difference we would not expect based on the underlying differences. The third term represents the interaction of the two differences. The Oaxaca-Blinder decomposition, therefore, helps us determine the extent to which the difference in disability rates between the eligible and undocumented samples arises because the two groups

[29] The vector **β** can also include a group-specific constant term, the difference in which would then be part of the difference to coefficients (not observable covariates).

have different underlying health conditions, or because, for a given set of medical conditions, the two groups are behaving differently in terms of how they approach the work decision.

Appendix D: Detailed Conceptual Framework

Consider the labor supply decision faced by an individual with a generic standard utility function $U(c,l)$ where c is consumption, l is leisure. As usual, $U_c > 0$, $U_l > 0$, $U_{cc} < 0$, $U_{ll} < 0$.

The individual with health H, which is an exogenous continuous variable between 0 (poor health) and 1 (excellent health), faces a binary decision: work, or stay out of work due to health limitations. Health has two effects: it raises the individual's market wage by some function $w(H)$, where $w' > 0$, and it reduces the probability an individual (if eligible) will receive disability benefits if not working by the function $p(H)$, where $p' < 0$.

If the individual works, he or she gets $c = w(H)$ and $l = l_{min}$. If he or she does not work and files for disability benefits, the individual gets $c = p(H)d$, where d is the disability benefits (if eligible), and $l = l_{max}$, where $l_{max} > l_{min}$.

An individual will choose to work if the additional utility from wage income over expected disability benefits is greater than the lost utility from less leisure:

$$U(w(H), l_{min}) > p(H)U(d, l_{max}) + [1-p(H)]U(0, l_{max})$$

This inequality is more likely to hold the healthier the individual is, as wages increase and the probability of receiving disability benefits (and therefore the expected disability benefits) falls as H rises. In contrast, an unhealthy *and* eligible individual will likely not work because the available market wage is low, the expected benefits due to disability are high, and not working allows more time for leisure. Not surprisingly, there will be a strong relationship between health and work.

Now imagine an individual who is ineligible for disability benefits. This individual will also work if the utility of doing so is greater than the utility of not working, but an ineligible individual will receive zero disability benefits:

$$U(w(H), l_{min}) > U(0, l_{max})$$

There will still be an extremely low level of health such that the individual does not work, as the available wage is so low that any utility from it is outweighed by increased utility from additional leisure time. At levels of health above this minimal threshold, however, the individual is much more likely to be working, since without the possibility of disability benefits even a small wage may outweigh the increased utility from more leisure. Overall, the relationship between health and work will be much weaker, and substantially different from the health-work locus in the eligible population.

Appendix E: Modified Charlson Index Calculation

The modification is due to a lack of data on all of the component diagnoses. This is necessary because the index was designed to work with hospital discharge data that contains ICD-9 or ICD-10 codes, as opposed to survey data regarding broad categories. As with the original index, we assign 1 point for each of the following conditions: myocardial infarction, congestive heart failure, cerebrovascular disease, chronic lung disease, connective tissue disease, or ulcer. We assign 1.5 points for diabetes (since in the unmodified index diabetes is 1 point and whereas diabetes with end organ damage is 2 points). We assign 2 points for liver disease (chronic liver disease is 1 point in the unmodified index and moderate and severe liver disease is 3), or moderate or severe kidney disease. We assign 4 points for cancer (as cancer/tumor is 2 points in the original index and malignant tumor/metastatis is 6). We do not have any data on the other components of the index (peripheral vascular disease, dementia, AIDS, or Hemiplegia/paraplegia) and omit them from our modified index here.