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

We are grateful to Dayanand Manoli, Delia Furtado, Ninez Ponce, Kirk Doran, and Heather Koball 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
JEL No. I12,I18,J61

ABSTRACT

Disability benefit recipients in the United States have nearly doubled in the past two decades, growing substantially faster than the population. It is difficult to estimate how much of this increase is explained by changes in population health, as we often lack a valid counterfactual. We propose using undocumented immigrants as the counterfactual, as they cannot currently claim benefits. Using NHIS microdata, we estimate models of disability as a function of medical conditions for both the legal and undocumented populations. The relationship between health and disability is far stronger for those with legal status than it is for those who are undocumented. We find that almost all of the difference in disability trends between the two populations can be explained by different responses to underlying health impairments.

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
335 Snow Hall
1460 Jayhawk Boulevard
Lawrence, KS 66045
david.slusky@ku.edu

I. Introduction

Disability claims 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% and the size of the population aged 55+ increased by 67% (Bureau of Labor Statistics 2017).

There are two explanations for the sizable increase in the size of the disability rolls: 1) it is the product of both an aging population and decreasing overall health; 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 (Autor and Duggan 2006; Liebman 2015; Mueller, Rothstein, and von Wachter 2016; Gelber, Moore, and Stand 2017; Milligan and Schirle 2017). This 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, as well as outright fraud by the nondisabled.

To unambiguously 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 us 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 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 that allows us 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. This eligibility provision has been enforced strictly in recent

years, especially since the Deficit Reduction Act of 2005 requires proof of citizenship when applying (Sommers 2010).

The Department of Homeland Security estimates that 11.4 million undocumented persons lived in the United States in January 2012 (DHS, 2017). 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 estimate a model that can be used to predict if a disabled person would work in the absence of social insurance programs. We can then use this model 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 the lowering requirements needed to qualify for disability benefits.

In addition to providing a new way of examining the long-standing question of why the disability rolls have increased dramatically, our analysis also provides the first credible measurement of the health status of the large undocumented population in the United States. Past research on immigrants (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). This 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 framework 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 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 of foreign-born persons 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.

Although much of the existing literature on the health of the immigrant population (Giuntella and Stella, 2016; Akbulut-Yuksel and Kugler, 2016) does not differentiate between the legal and undocumented groups, we know of three papers where the difference is explored. Goldman, Smith, and Sood (2006) use an algorithm where noncitizen foreign-born survey respondents who did not answer yes to having at least one of a permanent resident card, green card, or document allowing them to stay in the U.S. for a limited time were classified as "undocumented". 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.

The other two papers 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), uses 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, Wallace, Hadler, and Ponce (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. Neither of these studies, however, examine disability status in the undocumented immigrant population.

Further, none of the existing studies use the perspective of viewing the undocumented immigrant population as a counterfactual for the legal immigrant and native-born populations. This is a key contribution of our study, as many of the other studies on disability benefits cited above have lacked a counterfactual for what disability rates would have been in the complete absence of benefits. Some studies have instead exploited kinks in the Disability Insurance benefits formula (Gelber, Moore, and Strand 2017) or variations in benefit durations of other programs, such as Unemployment Insurance (Mueller, Rothstein, and von Wachter 2016). Both of these studies provide less opportunity to measure the extensive margin, i.e., how does eligibility affect whether an individual claims benefits at all?

Our analysis, therefore, extends the literature by applying the residual method of identifying undocumented status to the National Health Insurance Survey (NHIS) and addresses three related issues: 1) we compare the health and disability status of undocumented immigrants to both legal immigrants and the native 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.

II. Data

We use publicly available microdata from the National Health Interview Survey for the post-1997 period. The NHIS is an annual, repeated cross-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 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. 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 focuses on a set of specific health problems: heart disease, cancer, diabetes, hypertension, asthma, emphysema, liver disease, joint pain, ulcers, and bronchitis. We focus on this subset because these are the health impairments used by the Social Security Administration to determine whether an individual is disabled (Social Security 2017b). One important caveat is that all the health conditions in the NHIS microdata are self-reported, and self-reported health issues may not be unbiased measures of the 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.

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", 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 disability benefits. We will use the benefit information 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; 2017b) to the 1994-2015 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" would include whether the person entered the country as a refugee, 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 relationship 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 necessary to identify immigrants from Cuba (who are all legal because they are admitted as refugees). As a result, our analysis uses the 1998-2003 and 2005-2015 NHIS cross-sections.

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 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)[1], or are married to US citizens, or are the children or grandchildren of someone with legal status.[2] 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).

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

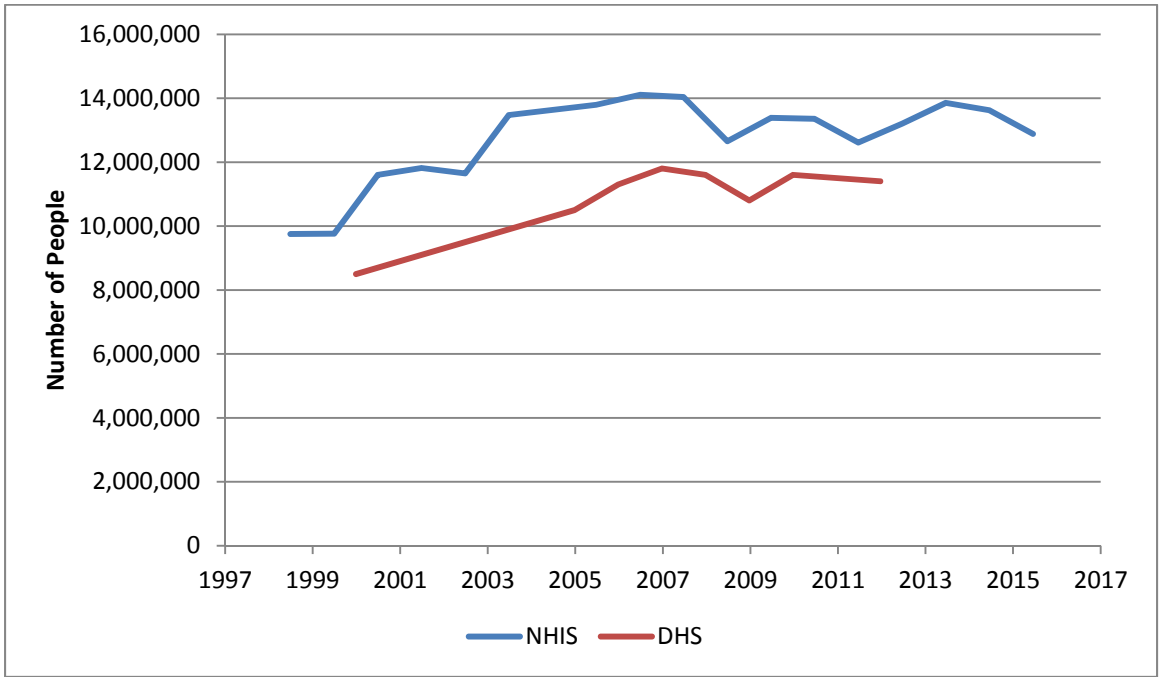
Notes: Data from NHIS Sample Adult and Sample Child files. Pooled for years 1998-2003 and 2005-2015. Each row represents the count of those excluded by that row but not the above rows.

[1] 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.

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

Figure 1 contrasts our estimates of the number of undocumented immigrants with the official DHS estimates. Although the two estimates are reasonably close to each other, follow the same upward trend in the 2000-2007 period, and are both roughly constant in the 2007-2011 period, it is notable that the imputation method in the NHIS leads to about 1 million more undocumented persons in a given year. 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 persons that we impute to be undocumented seem to correctly summarize key trends in the undocumented population.

Figure 1: Undocumented Population by NHIS and DHS



Note: Red line from DHS (2017).

We limit much of the regression reported below to persons aged 18-65. There are extremely few individuals aged 65+ in the NHIS sample that our algorithm labels 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 claiming disability status as there is a substantial break in the types of benefits available to the two elderly groups.

Table 2 reports summary statistics for many of the variables used in our empirical analysis. It is evident that undocumented immigrants self-report themselves to be far healthier than both legal immigrants and the native born. 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%, as compared to 41% percent for a legal immigrant and 54% for someone who is native born. Undocumented immigrants are also 5 years younger and have much less education. In particular, 45 percent of the undocumented immigrants lack a high school diploma, as compared to only about 12 percent of the legal immigrants and 11 percent of the native born.

Table 2: Summary Statistics

	(1) Native Born	(2) Legal Immi-grants	(3) Pooled Native Born and Legal Immigrants[3]	(4) Undocu-mented	(5) Differ-ence between (3) and (4)	(6) Standard Error
Disabled	0.074	0.045	0.070	0.014	-0.057***	(0.002)
Male	0.488	0.477	0.487	0.559	0.072***	(0.003)
Heart Disease	0.023	0.016	0.023	0.008	-0.015***	(0.001)
Cancer	0.053	0.026	0.050	0.008	-0.042***	(0.001)
Diabetes	0.060	0.065	0.060	0.038	-0.022***	(0.002)
Hypertension	0.222	0.184	0.217	0.100	-0.112***	(0.003)
Asthma	0.128	0.068	0.121	0.036	-0.085***	(0.002)
Emphysema	0.011	0.003	0.010	0.002	-0.008***	(0.0007)
Liver Disease	0.013	0.014	0.013	0.008	-0.005***	(0.0007)
Joint Pain	0.302	0.192	0.290	0.111	-0.179***	(0.003)
Ulcer	0.067	0.046	0.065	0.029	-0.036***	(0.002)
Bronchitis	0.043	0.018	0.040	0.009	-0.031***	(0.001)
Any Ailment	0.535	0.409	0.521	0.252	-0.269***	(0.003)
Age (years)	40.5	41.8	40.7	35.4	-5.2***	(0.088)
High School Drop Out	0.106	0.214	0.118	0.452	0.334***	(0.002)
High School Graduate	0.284	0.214	0.276	0.210	-0.066***	(0.003)
Some College	0.333	0.244	0.323	0.140	-0.183***	(0.003)
College Graduate	0.277	0.327	0.283	0.197	-0.085***	(0.003)
N	333,938	46,725	380,663	30,106		

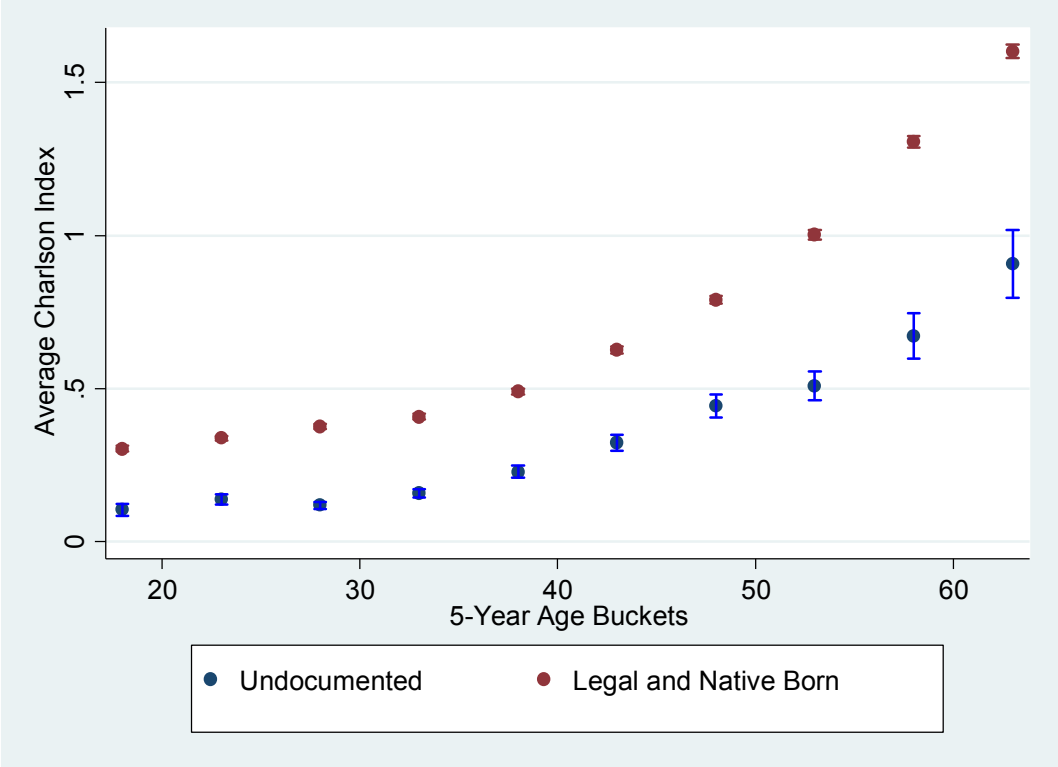
Notes: NHIS Sample Adults, 18-65. Weighted. *** p<0.01, ** p<0.05, * p<0.1

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 pooled native born and legal immigrant populations. In particular, we aggregate across the

[3] Throughout this paper, we pool 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.

various medical conditions by using a modified Charlson Index (Charlson et al. 1987), which is essentially a weighted sum across conditions.[4]

Figure 2: Charlson Index by Age for Undocumented Immigrants and Pooled Legal Immigrants/Native Born



Notes: NHIS Sample Adults, 18-65. Weighted. 95% confidence interval shown in whiskers around each point.

Figure 2 shows the weighted average Charlson Index for each age by legal status. Note that the Charlson Index is greater (indicating worse health) for the pooled native born and legal immigrants at every age. Not surprisingly, the index for the pooled native born and legal

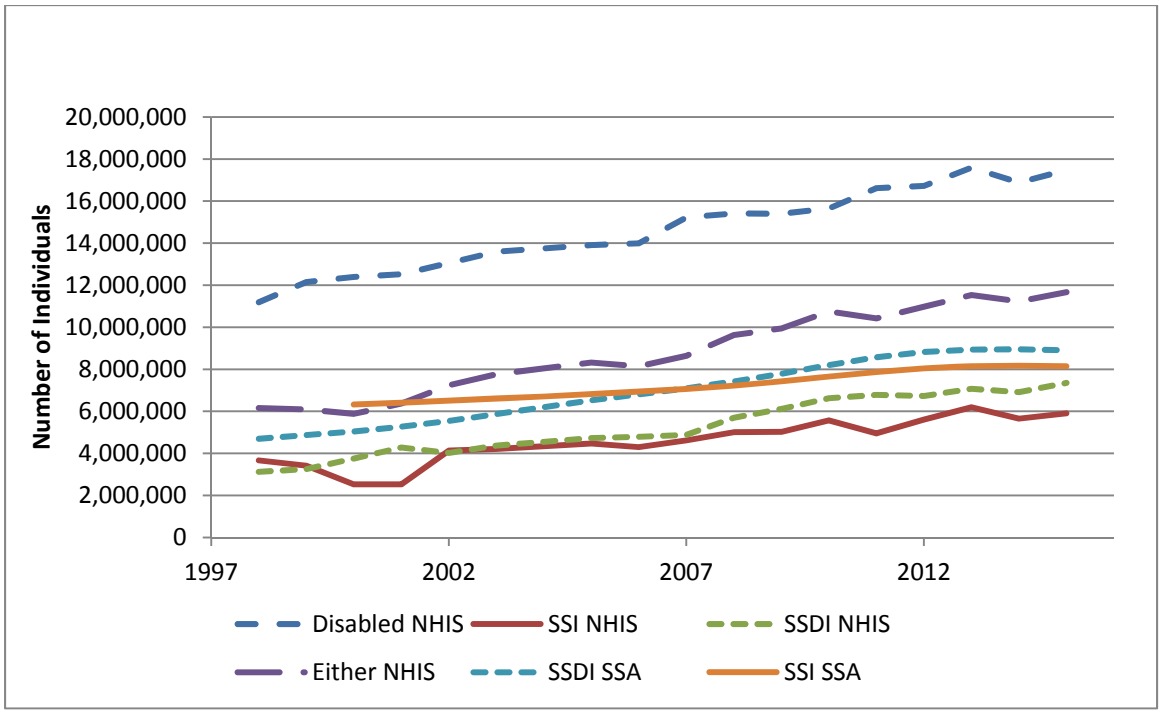
[4] The modification is due to a lack of data on all of the component diagnoses. This is necessary because the index was designed to work with hospital discharge data that contains ICD-9 or ICD-10 codes, as opposed to survey data regarding broad categories. As with the original index, we assign one point for 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.

immigrants 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 the pooled native born and legal immigrants) particularly as the groups approach retirement age.

IV. Results

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 disabled persons receiving Social Security Disability benefits (SSDI) or Supplemental Security Income (SSI) (or at least one). Figure 3 illustrates several trends, revealing that all measures have been increasing rapidly

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.

The NHIS data, where disability is defined by the number of persons who did not work in the reference 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 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 legal immigrants and natives 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.

Our regression model predicts the probability that a person is disabled based on his or her self-reported medical conditions and on a vector of 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 vector **D** contains dummy variables giving the medical conditions used by the Social Security Administration to evaluate disability status: heart disease, cancer, diabetes, hypertension, asthma, emphysema, liver disease, joint pain, ulcer, and bronchitis (Social Security 2017b). Finally, the **age, education, quarter,**

and **year** variables and are vectors of fixed effects for 10-year age brackets, educational attainment brackets, survey quarter, survey year, and gender, respectively.

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. 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.

Table 3: Predicting Disability Status using Self-Reported Medical Conditions (for Pooled Sample of Legal Immigrants and Native Born)

	(1) OLS	(2) Probit	(3) Logit
Heart Disease	0.152***	0.0565***	0.0505***
	(0.00599)	(1.96e-05)	(1.76e-05)
Cancer	0.0495***	0.0318***	0.0296***
	(0.00312)	(1.56e-05)	(1.47e-05)
Diabetes	0.101***	0.0438***	0.0397***
	(0.00333)	(1.35e-05)	(1.23e-05)
Hypertension	0.0412***	0.0285***	0.0280***
	(0.00152)	(9.80e-06)	(9.63e-06)
Asthma	0.0313***	0.0252***	0.0252***
	(0.00187)	(1.20e-05)	(1.16e-05)
Emphysema	0.230***	0.0648***	0.0553***
	(0.00928)	(2.76e-05)	(2.43e-05)
Liver disease	0.191***	0.0758***	0.0686***
	(0.00751)	(2.47e-05)	(2.22e-05)
Joint pain	0.0573***	0.0466***	0.0464***
	(0.00125)	(9.04e-06)	(9.12e-06)
Ulcer	0.0598***	0.0314***	0.0295***
	(0.00278)	(1.36e-05)	(1.27e-05)
Bronchitis	0.0674***	0.0308***	0.0284***
	(0.00365)	(1.70e-05)	(1.58e-05)
Observations	380,663	380,663	380,663
R-squared	0.147	0. 219	0.217

Notes: NHIS Sample Adult 18-65. 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

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).

Table 4: Predicting Disability Status with Medical Conditions (for Undocumented)

	(1) OLS	(2) Probit	(3) Logit
Heart Disease	0.0384	0.0109***	0.0100***
	(0.0240)	(5.76e-05)	(4.73e-05)
Cancer	0.0358*	0.0155***	0.0136***
	(0.0211)	(5.85e-05)	(5.04e-05)
Diabetes	0.0281***	0.00894***	0.00839***
	(0.00895)	(3.05e-05)	(2.67e-05)
Hypertension	0.0125**	0.00795***	0.00716***
	(0.00509)	(2.32e-05)	(2.22e-05)
Asthma	0.00414	0.00306***	0.00325***
	(0.00553)	(4.07e-05)	(3.84e-05)
Emphysema	0.0706	0.0114***	0.00876***
	(0.0592)	(0.000106)	(8.15e-05)
Liver disease	0.0462**	0.0138***	0.0124***
	(0.0230)	(5.69e-05)	(4.60e-05)
Joint pain	0.0163***	0.0102***	0.00959***
	(0.00393)	(2.18e-05)	(2.08e-05)
Ulcers	0.00361	0.00190***	0.00171***
	(0.00704)	(4.02e-05)	(3.71e-05)
Bronchitis	0.0278	0.00865***	0.00809***
	(0.0202)	(6.15e-05)	(5.14e-05)
Observations	30,106	30,106	30,106
R-squared	0.116	0.100	0.101

Notes: NHIS Sample Adult 18-65. 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

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. For example, in the 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. We can then subtract one equation from the other, add zero, rearrange, and factor:

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

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, given the lower set of coefficients. 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. 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 pooled native born and legal immigrant and undocumented samples arises because the two groups 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.

Table 5 summarizes the 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 75 percent of the variation in the mean disability rate, whereas the differences in endowments (i.e., that the undocumented population is younger and healthier) only explains about 25%. In short, the different disability rates between the two groups is mostly attributable to the fact that adverse medical conditions are far less likely

to lead to withdrawal from the labor force in the undocumented sample than in the pooled native born and legal immigrant sample.[5]

Table 5: Oaxaca-Blinder Decomposition

	(1) OLS	(2) Probit	(3) Logit
Means:			
Legal Immigrants	0.0704***	0.0704***	0.0704***
& Native Born	(4.74e-06)	(4.49e-06)	(4.48e-06)
Undocumented	0.0139***	0.0139***	0.0139***
	(8.68e-06)	(8.61e-06)	(8.58e-06)
Difference in	0.0565***	0.0565***	0.0565***
means	(9.89e-06)	(9.71e-06)	(9.68e-06)
Share due to:			
Endowments	0.0107***	0.0107***	0.0108***
	(1.09e-05)	(2.21e-05)	(2.37e-05)
Coefficients	0.0538***	0.0475***	0.0468***
	(1.21e-05)	(1.18e-05)	(1.16e-05)
Interaction	-0.00796***	-0.00188***	-0.00113***
	(1.31e-05)	(2.32e-05)	(2.46e-05)
Observations	410,769	410,769	410,769

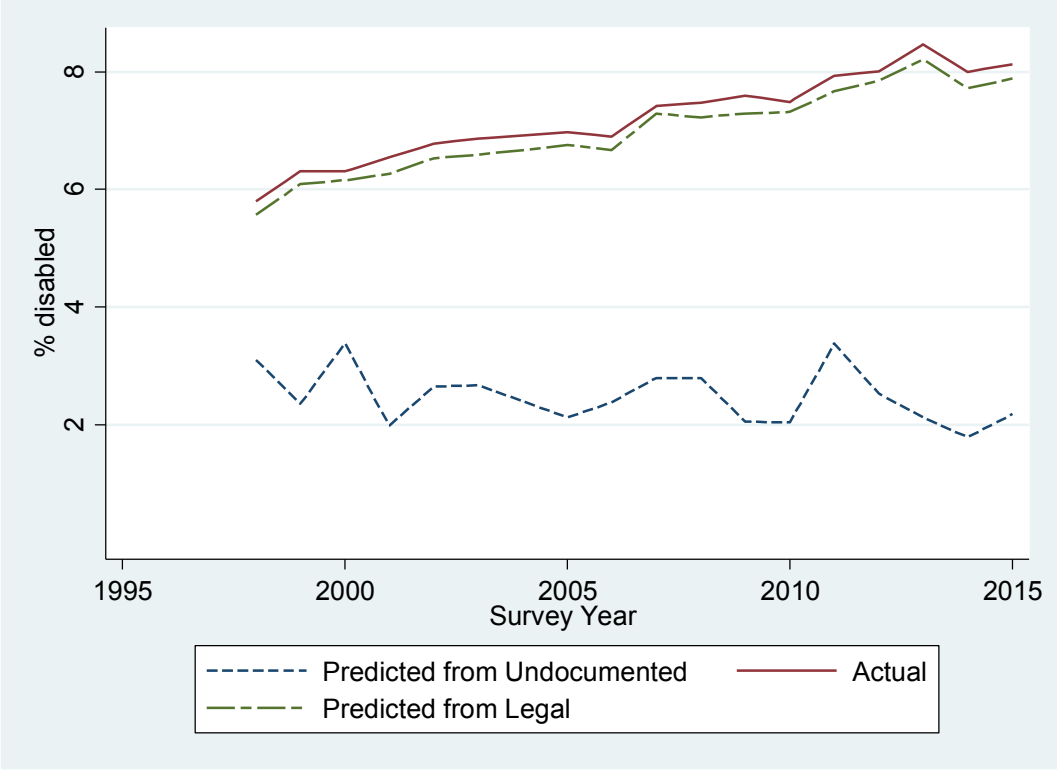
Notes: NHIS Sample Adult 18-65. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

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 trend in the disability rate of the legal immigrant 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 legal immigrants?

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

This counterfactual analysis helps us address two crucial questions: 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? This exercise allows us to estimate both extensive margins: the impact of the removal of benefits and the impact of the introduction of benefits. This is contrast to previous studies (Mueller, Rothstein, and von Wachter 2016; Gelber, Moore, and Stand 2017; Milligan and Schirle 2017) have mostly focused on the intensive margin (i.e., changes in the generosity of benefit) or the interaction with other programs (e.g., unemployment insurance).

Figure 4: Predicted trend in Disability Rates for Legal Immigrants and Native Born



Notes: NHIS Sample Adult 18-65. Weighted. Uses Logit model from above.

Figure 4 shows the actual and predicted disability rates for the legal immigrant and native born population. 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 legal

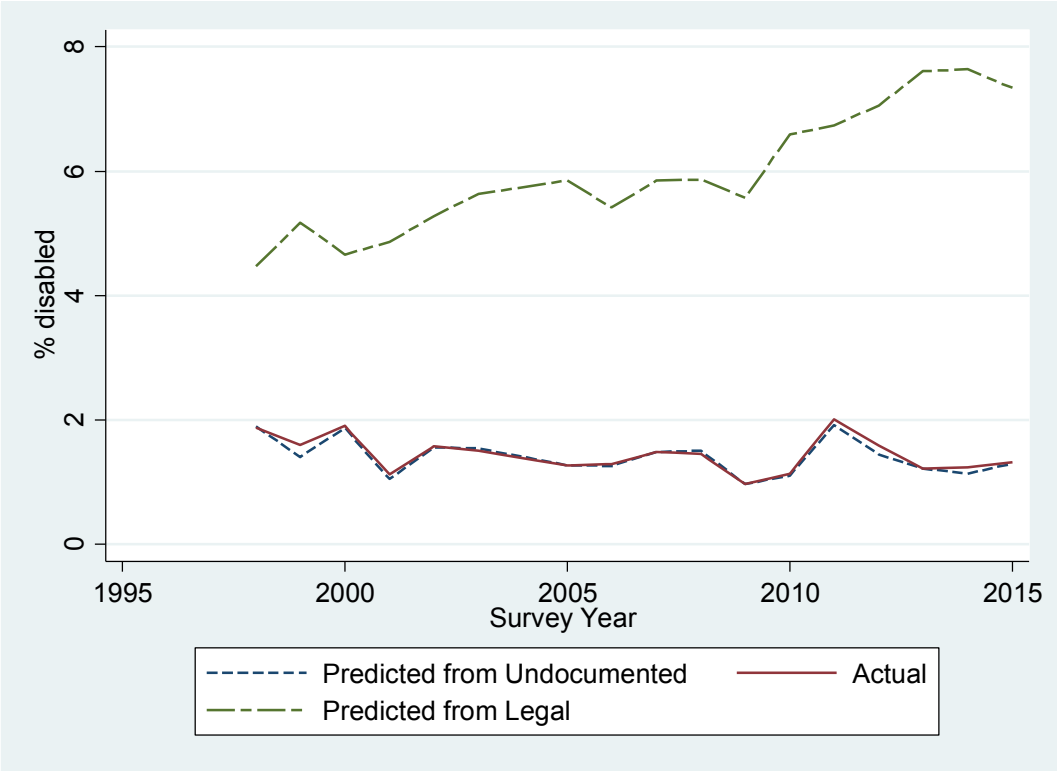
immigrant population. Second, the disability rate as predicted by the model fitted using the sample of undocumented persons.

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 shown 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 legal immigrants 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 pooled native born and legal immigrant 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 legal immigrants (partly because they become 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.

Figure 5: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: NHIS Sample Adult 18-65. Weighted. Uses Logit model from above.

V. 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 two alternative data sets, the California Health Insurance Survey (CHIS) and the CPS-ASEC. These sensitivity tests show that our evidence is 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.

We first replicate the Oaxaca-Blinder using many more measures of health status beyond those used by the Social Security Administration (2017b) in determining disability status. Table A-1 shows the results. The first column replicates the evidence from our earlier analysis. The

second column adds the following 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 "legal" population, whether legal immigrants or legal immigrants and natives, is explained mostly by differences in the coefficients that determine disability status. 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 illustrate the robustness of the key evidence. In particular, we first compare two groups who should not have any difference in the ability to claim disability benefits: native-born and legal immigrants who came into the country as children.[6] 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 A1 and A2 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

[6] The categories for NHIS's variable for years in the US 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. Therefore, Figures A3 and A4 repeat figures A1 and A2 for only those 18-32. This also results in us dropping the age bucket dummy variables from this entire analysis, with the exception of the one for 18-27 (vs. 28+).

undocumented immigrants. As Figures A5 and A6 show, the trend in disability rates in these two groups is 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 regression models using CHIS (2017). It is much more difficult to apply the 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 dropped) and there are only extremely broad occupation and industry codes (limiting dropping those 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 conditions: 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 it only converges for a probit specification. Nevertheless, in Figure A12 we show that our age/health profile result holds (where the legal and native born population is sicker at every age). In Table A5 we shown that our Oaxaca-Blinder decomposition result (that the difference is coming from the coefficients) also holds. In Figures A13 and A14 we show that as above predicting for the legal population using the undocumented model reduced the level and removes the trend, and vice versa increases the level and introduces a trend.

Finally, we also check our results using the CPS ASEC data from IPUMS (Flood et al.

2017), as used in Borjas (2017a; 2017b). First, given the large sample, with reasonable confidence we can check the sums of survey weights for undocumented immigrants for California against the sums of survey weights for the CHIS data. This result is in Figure A15. 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. Nevertheless, we repeated the logistic regressions from above of disability status (defined by not working for reasons of illness or disability) on these variables. We then used the coefficients from those two regressions to predict the two counterfactuals described above. Figures A16 and A17 show the results of this analysis, which are largely consistent with Figures 4 and 5 above.

VI. Discussion and Implications

We can use the estimates from our analysis of the NHIS data 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?

Table 6 shows the each element of our calculation:

Table 6: Cost savings of disability reduction for native born and legal immigrants, 18-65

Total population	187 million (sum of survey weights)
Disabled population	15.2 million (sum of survey weights)
Disability rate	0.0812 (= 15.2 million / 187 million)
Counterfactual disability rate	0.0218 (using counterfactual prediction)
Counterfactual disabled population	4.08 million (= 187 million * 0.0218)
Change in population disabled	-11.0 million (=4.08 million – 15.2 million)
Share of disabled legal and native born receiving SSDI	0.383 (using survey response)
Population no longer receiving SSDI	-4.21 million (= -11.0 million * 0.383)
Average monthly benefits for SSDI	$1,171.25 (from Social Security)
Monthly savings from SSDI	-$4.93 billion (=-4.21 million * $1,171.25)
Share of disabled legal and native born receiving SSI	0.261 (using survey response)
Population no longer receiving SSI	-2.87 million (= -11.0 million * 0.261)
Average monthly benefits for SSI	$542.5 (from Social Security)
Monthly savings from SSI	-$1.56 billion (=-2.87 million * $542.5)
Total monthly savings	=-$6.49 billion (=-$4.21 billion – $1.56 billion)
Total annual savings	=$77 billion ($6.49 billion * 12)

In 2015 (the last year of NHIS data used in our analysis), the sum of the survey weights corresponds to 187 million native born and legal immigrants ages 18-65. Figure 4 shows the disability dropping from the measured 8.1% (~15.2 million individuals) to only 2.3% (~4.1 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 18-65, in 2015 38.3% receive SSDI and 26.1% receive SSI.[7] 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 $6.5 billion per month ($77 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). So this represents a 43% decline in payouts. That said, there are

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

likely individuals in an intermediate space where they are not definitely disabled (the 2.3%) nor in the tail of the 8.1% and claiming fraudulently who would still claim disability status and benefits in a harsher regime. In other words, while there are some savings to be had, the estimate of $77 billion is almost certainly an upper bound.

Another way to look at this is to see that there is no trend in the level of undocumented for those with legal 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% - is likely nonessential benefits and can be explained by changing coefficients, and not by a population that is getting older and sicker.

The second exercise, and quite 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 7 shows the each element of our calculation:

Table 7: Cost of disability benefits for undocumented immigrants

Total population	11.8 million (sum of survey weights)
Disabled population	156 K (sum of survey weights)
Disability rate	0.0132 (= 156 K / 11.8 million)
Counterfactual disability rate	0.0734 (using counterfactual prediction)
Counterfactual disabled population	866 K (= 11.8 million * 0.0734)
Share of disabled legal and native born receiving SSDI	0.383 (using survey response)
Population now receiving SSDI	+332 K (= +866 K * 0.383)
Average monthly benefits for SSDI	$1,171.25 (from Social Security)
Monthly cost from SSDI	+$388 million (=+332 K * $1,171.25)
Share of disabled legal and native born receiving SSI	0.261 (using survey response)
Population now receiving SSI	+226 K (= +866 K * 0.261)
Average monthly benefits for SSI	$542.5 (from Social Security)
Monthly cost from SSI	+$123 million (=+226 K * $542.5)
Total monthly cost	=+$511 million (=$388 million + $123 million)
Total annual cost	=+$6.1 billion (+$511 million * 12)

The most recent DHS estimate is that there are 11.4 million undocumented immigrants (Department of Homeland Security 2017), which closely matches the sum of survey weights from our analysis (11.8 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). The predicted increase in the share of undocumented immigrants who are disabled if they were "treated like" legal immigrants was from 1.3% to 7.2%. Allowing all of the 7.2% of these individuals to claim benefits (as even the ones who previously reported disability can now claim) would lead to an increase in federal liabilities of $6.1 billion per year, which represents an increase of 3.4%. Note, however, that many undocumented immigrants may already be paying taxes to the disability system but have not qualified for benefits (Goss et al. 2013; Social Security 2015; Gee, Gardner, and Wiehe 2016).

Finally, we can use our empirical results to answer the opening question of this paper: 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 18-65 population (of any immigration status) was 5.6% in 1998 and 7.7% 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%. So the aging of the population can only explain 29% of the increase. The rest must be due to changes in other factors such as the impact of medical conditions increase the probability that a person did not work in the reference week due to health reasons.

VII. 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 both of persons legally in the country (both legal immigrants and natives), and allows us to calculate counterfactuals that help us understand how disability status 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 a 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 status 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 level of the disability rate for the legal population drops substantially and there is no longer the upward sloping time trend in disability over the past two decades. In the second, the level of the disability rate substantially increases and an upward sloping time trend appears.

These results suggest two policy applications. The first is that there is likely both substantial fraud in the current disability benefits system and numerous situations where an individual with improved economic circumstances 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 second is that legalizing the undocumented population might 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? NBER Working Paper No. 21987.

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.

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.

Blinder, A.S.: 1973. Wage discrimination: reduced form and structural estimates. J. Hum. Resour. 8, 436–455.

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

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.

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.

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

Department of Homeland Security. 2017. Estimates of the Unauthorized Immigrant Population Residing in the United States. https://www.dhs.gov/immigration-statistics/population-estimates/unauthorized-resident

Sarah Flood, 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

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, and Meg Wiehe. 2016. Undocumented Immigrants' State & Local Tax Contributions. The Institute on Taxation & Economic Policy.

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.

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.

Giuntella, Osea and Luca Stella. 2016. The Acceleration of Immigrant Unhealthy Assimilation IZA Discussion Paper No. 9664.

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.

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

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

Milligan, Kevin and Tammy Schirle. 2017. Push and Pull: Disability Insurance, Regional Labor Markets, and Benefit Generosity in Canada and the United States. NBER Working Paper No. 23405.

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, R. 1973. Male-female wage differentials in urban labor markets. Int. Econ. Rev. 14, 693–709

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.

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

Sommers, Ben. 2010. Targeting in Medicaid: The costs and enrollment effects of Medicaid's citizenship documentation requirement. Journal of Public Economics 94: 174-182.

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.

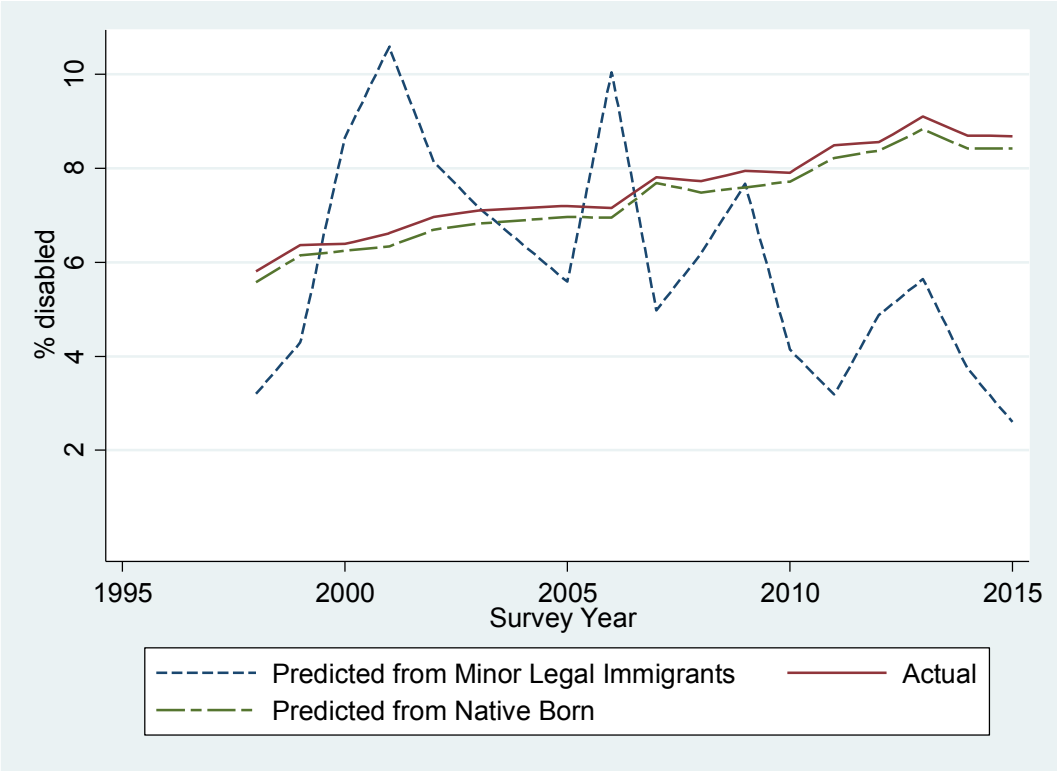
Warren, Robert E. and Jeffrey S. Passel. 1987. "A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 United States Census." *Demography* 24(3): 375–393.

Table A1: Oaxaca-Blinder Decomposition for Logit with Severity

	(1) SSA Controls	(2) SSA Controls + Severity	(3) SSA Controls + Severity + Cancer Type
Means:			
Legal	0.0700***	0.0700***	0.0700***
	(4.48E-06)	(4.41E-06)	(4.41E-06)
Undocumented	0.0139***	0.0139***	0.0139***
	(8.58E-06)	(8.54E-06)	(8.51E-06)
Difference in means	0.0561***	0.0561***	0.0561***
	(9.67E-06)	(9.61E-06)	(9.58E-06)
Share due to:			
Endowments	0.0108***	0.0115***	0.0109***
	(2.36E-05)	(2.50E-05)	(2.33E-05)
Coefficients	0.0467***	0.0452***	0.0452***
	(1.16E-05)	(1.15E-05)	(1.15E-05)
Interaction	-0.00132***	-0.000613***	4.25e-05*
	(2.45E-05)	(2.59E-05)	(2.42E-05)
Observations	410,123	410,123	410,123

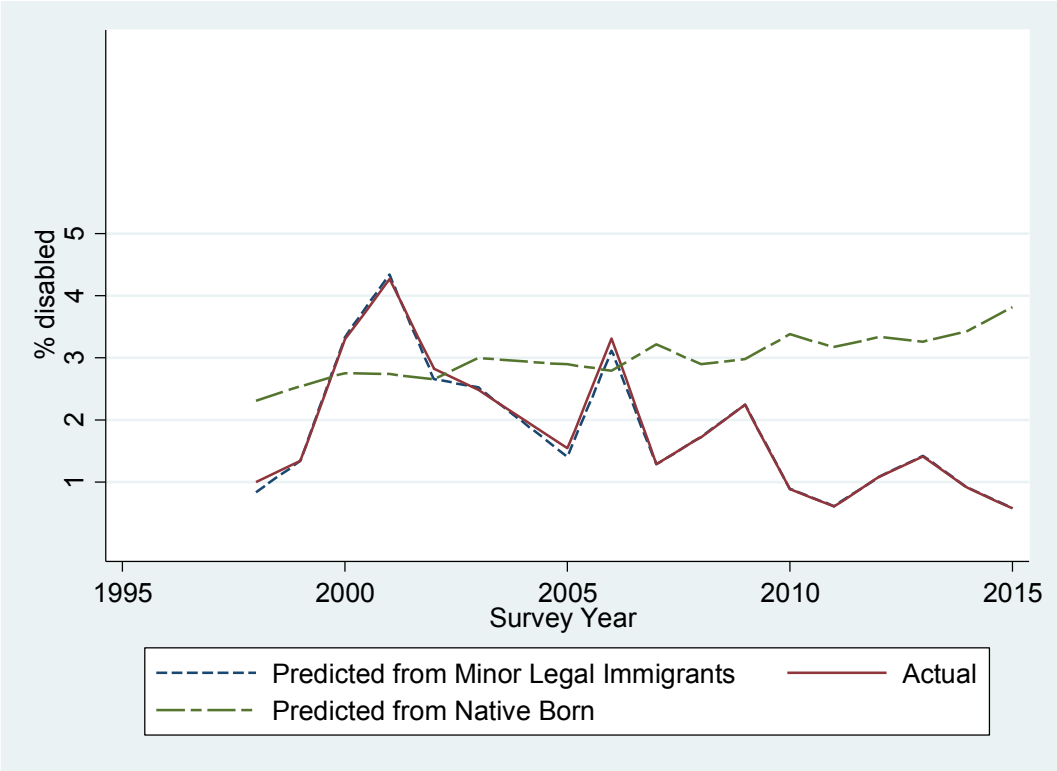
Notes: NHIS Sample Adult 18-65. Weighted. Robust standard errors in parentheses. Column (2) contains dummy variables for heart attack, angina, other heart disease, stroke, kidney disease. Column (3) additionally contains dummy variables for common types of cancer: breast, cervical, colon, kidney, leukemia, lung, lymphoma, thyroid, uterine. *** p<0.01, ** p<0.05, * p<0.1

Figure A1: Predicted trend in Disability Rates for Native Born



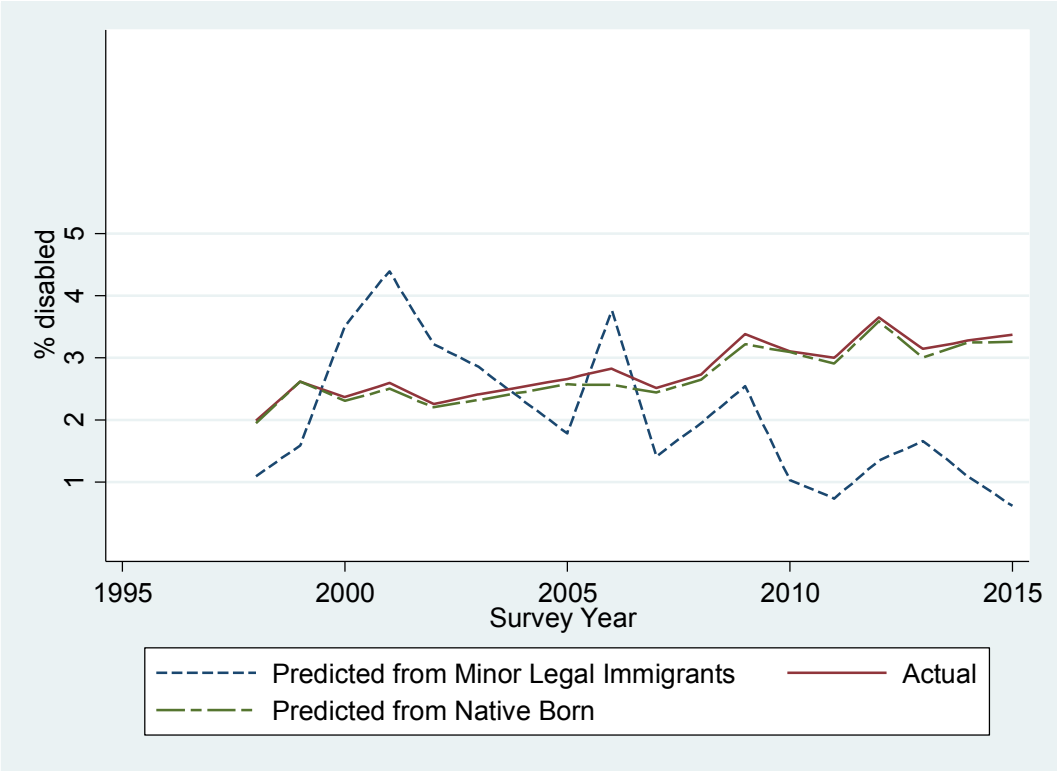
Notes: NHIS Sample Adult 18-65. Weighted. Uses Logit model.

Figure A2: Predicted trend in Disability Rates for Minor Legal Immigrants



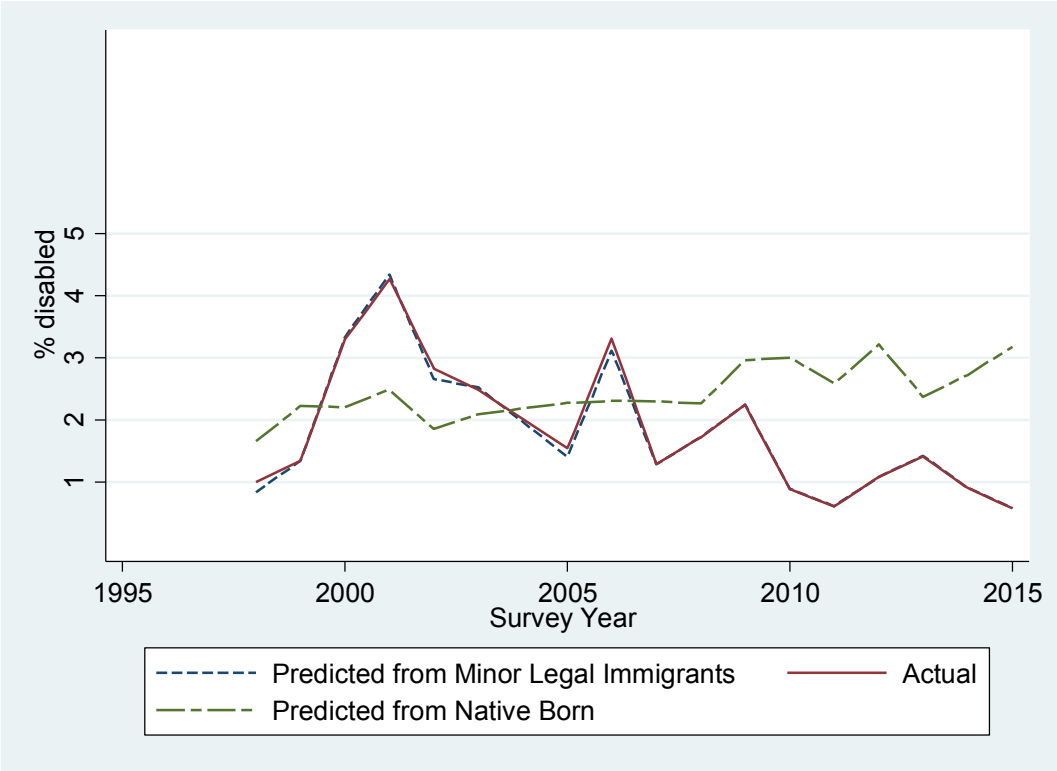
Notes: NHIS Sample Adult 18-65. Weighted. Uses Logit model.

Figure A3: Predicted trend in Disability Rates for Native Born (Ages 18-32)



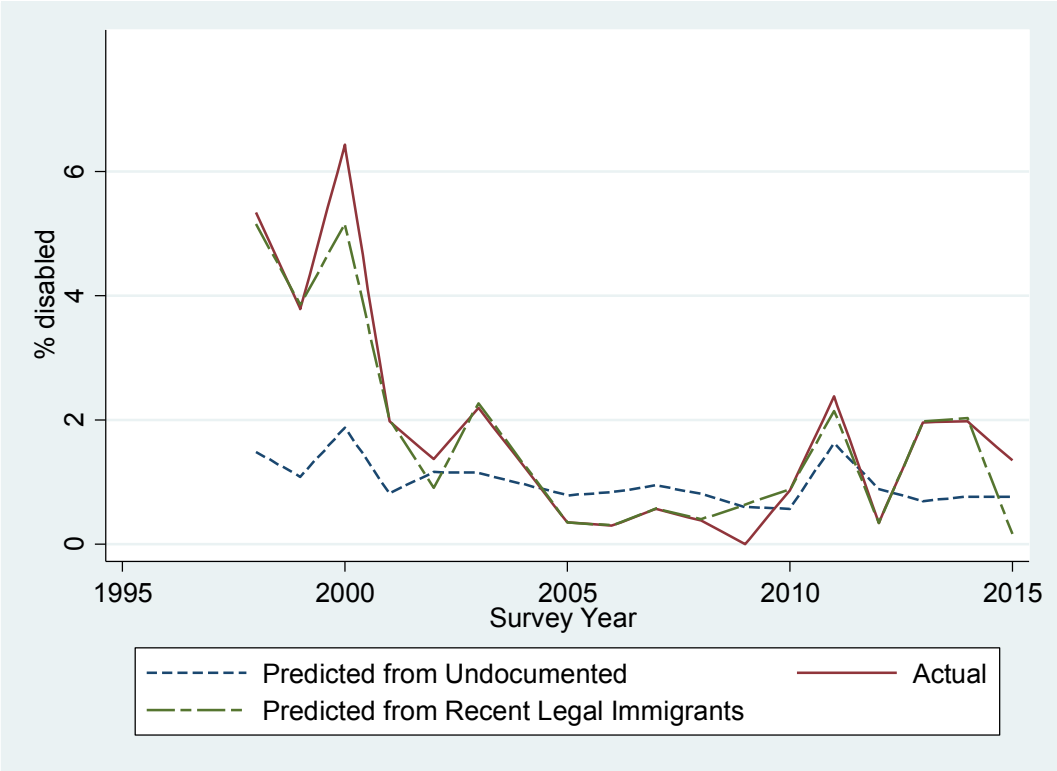
Notes: NHIS Sample Adult 18-32. Weighted. Uses Logit model.

Figure A4: Predicted trend in disability Rates for Minor Legal Immigrant (Ages 18-32)



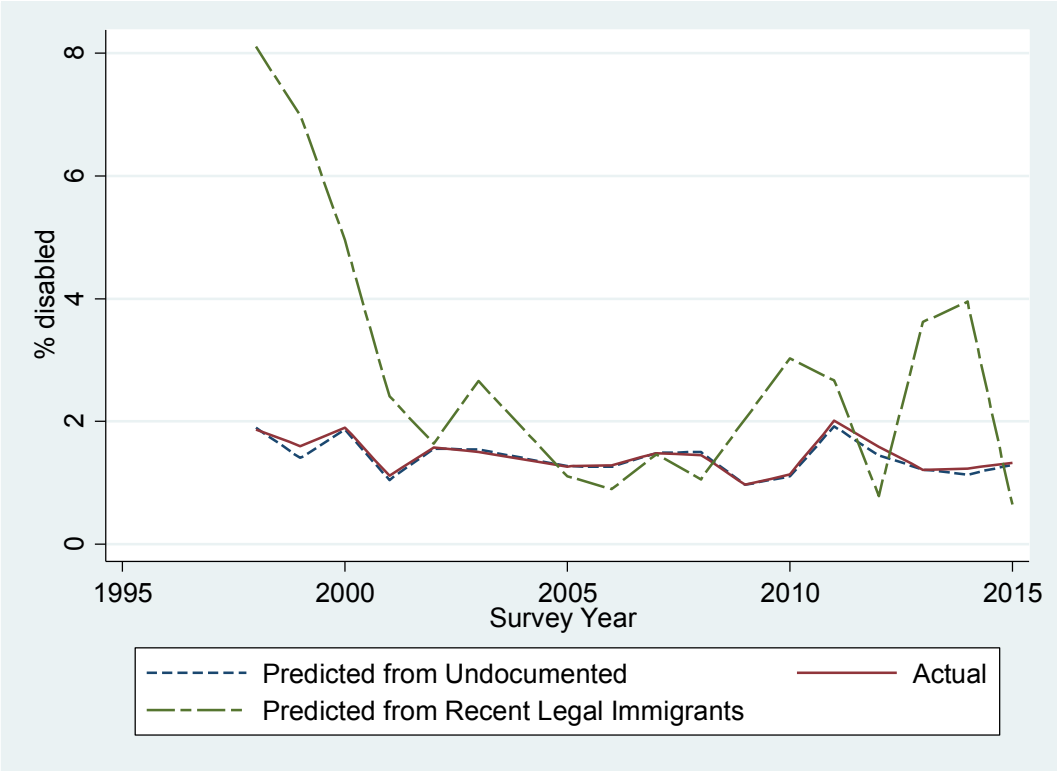
Notes: NHIS Sample Adult 18-32. Weighted. Uses Logit model.

Figure A5: Predicted trend in Disability Rates for Recent Legal Immigrants



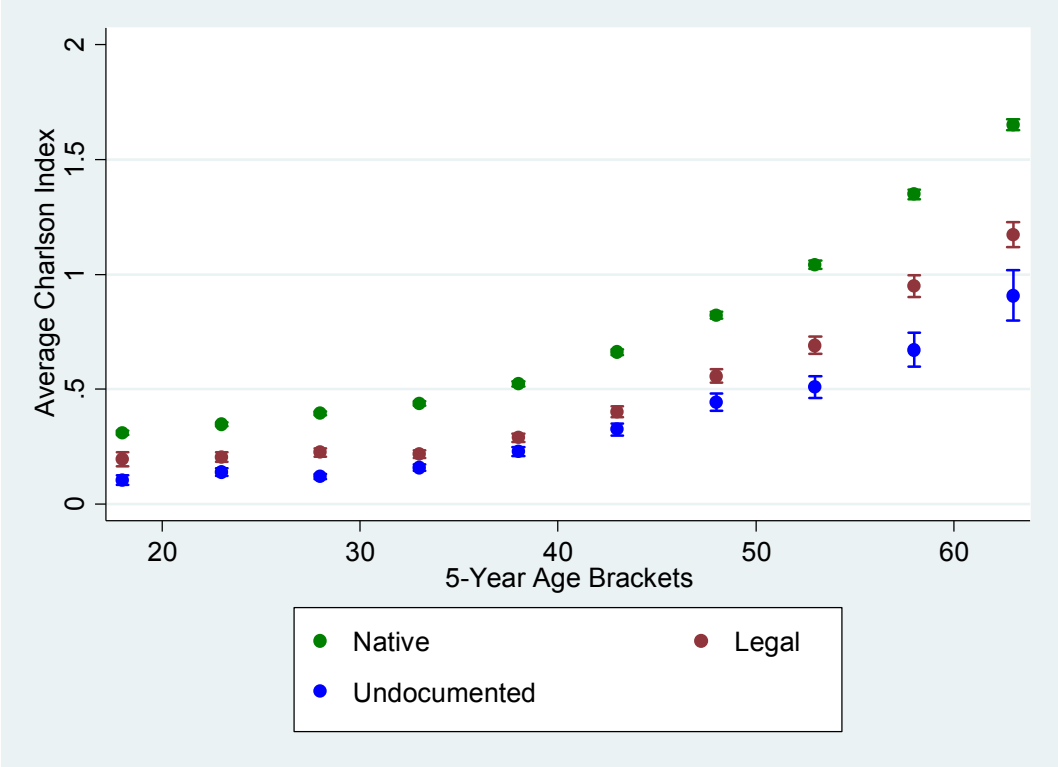
Notes: NHIS Sample Adult 18-65. Weighted. Uses Logit model.

Figure A6: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: NHIS Sample Adult 18-65. Weighted. Uses Logit model.

Figure A7: Charlson Index by Age for Undocumented, Legal, and Native Born



Notes: NHIS Adult Sample, 18-65. Weighted.

Table A2: Predicting Disability Status with Medical Conditions

	(1) Native & Legal	(2) Native & Legal	(3) Native	(4) Legal	(5) Undocumented
Heart Disease	0.0505***	0.0501***	0.0525***	0.0293***	0.0100***
	(1.76e-05)	(1.76e-05)	(1.89e-05)	(4.64e-05)	(4.73e-05)
Cancer	0.0296***	0.0286***	0.0285***	0.0349***	0.0136***
	(1.47e-05)	(1.47e-05)	(1.57e-05)	(4.21e-05)	(5.04e-05)
Diabetes	0.0397***	0.0403***	0.0422***	0.0252***	0.00839***
	(1.23e-05)	(1.23e-05)	(1.34e-05)	(2.91e-05)	(2.67e-05)
Hypertension	0.0280***	0.0274***	0.0274***	0.0255***	0.00716***
	(9.63e-06)	(9.61e-06)	(1.03e-05)	(2.46e-05)	(2.22e-05)
Asthma	0.0252***	0.0242***	0.0252***	0.0119***	0.00325***
	(1.16e-05)	(1.16e-05)	(1.23e-05)	(3.64e-05)	(3.84e-05)
Emphysema	0.0553***	0.0540***	0.0562***	0.0309***	0.00876***
	(2.43e-05)	(2.43e-05)	(2.56e-05)	(9.36e-05)	(8.15e-05)
Liver disease	0.0686***	0.0691***	0.0733***	0.0343***	0.0124***
	(2.22e-05)	(2.22e-05)	(2.41e-05)	(5.27e-05)	(4.60e-05)
Joint pain	0.0464***	0.0451***	0.0455***	0.0404***	0.00959***
	(9.12e-06)	(9.11e-06)	(9.80e-06)	(2.34e-05)	(2.08e-05)
Ulcers	0.0295***	0.0290***	0.0306***	0.0173***	0.00171***
	(1.27e-05)	(1.27e-05)	(1.36e-05)	(3.56e-05)	(3.71e-05)
Bronchitis	0.0284***	0.0276***	0.0282***	0.0217***	0.00809***
	(1.58e-05)	(1.57e-05)	(1.67e-05)	(5.14e-05)	(5.14e-05)
Legal immigrant		-0.0285***			
		(1.67e-05)			
Observations	380,663	380,663	333,938	46,725	30,106

Notes: NHIS Sample Adult 18-65. Weighted. 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 A3: Oaxaca Blinder for Native Born vs. Undocumented Immigrants

	(1) OLS	(2) Probit	(3) Logit
Means:			
Native Born	0.0736***	0.0734***	0.0736***
	(5.13e-06)	(4.85e-06)	(4.84e-06)
Undocumented	0.0139***	0.0139***	0.0139***
	(8.68e-06)	(8.61e-06)	(8.58e-06)
Difference in means	0.0597***	0.0595***	0.0597***
	(1.01e-05)	(9.88e-06)	(9.85e-06)
Share due to:			
Endowments	0.0112***	0.0113***	0.0115***
	(1.15e-05)	(2.36e-05)	(2.53e-05)
Coefficients	0.0626***	0.0550***	0.0540***
	(1.29e-05)	(1.30e-05)	(1.27e-05)
Interaction	-0.0142***	-0.00687***	-0.00584***
	(1.42e-05)	(2.51e-05)	(2.66e-05)
Observations	364,044	364,044	364,044

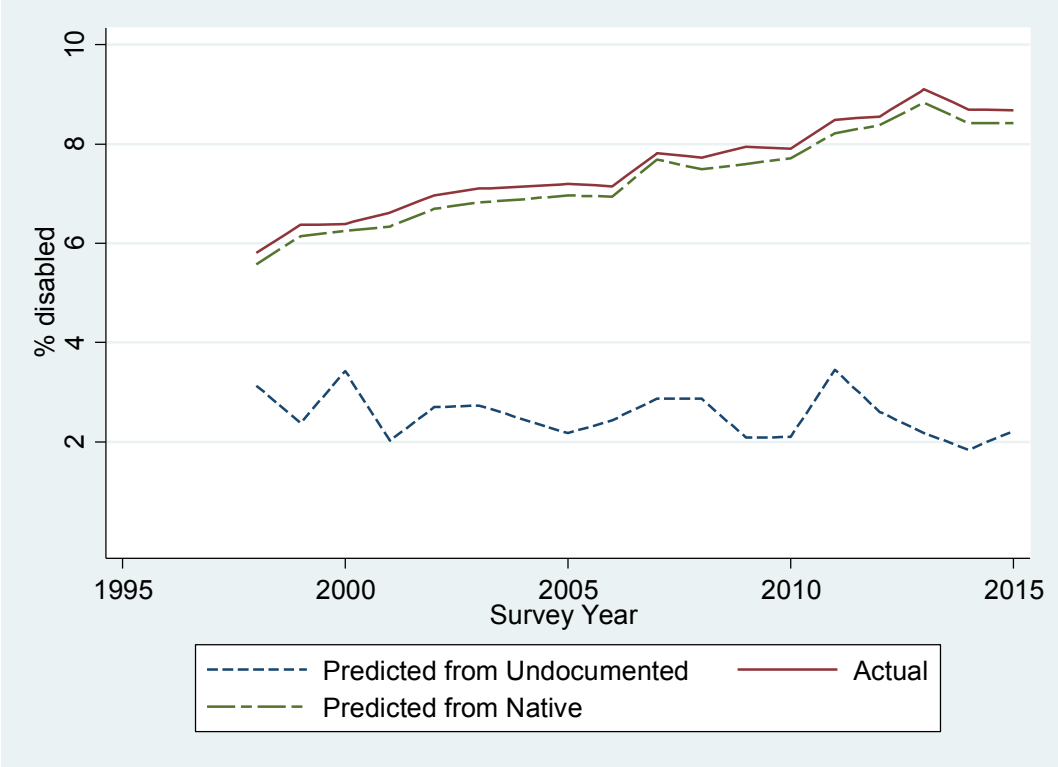
Notes: NHIS Sample Adult 18-65. Native Born and Undocumented only. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A4: Oaxaca Blinder for Legal vs. Undocumented Immigrants

	(1) OLS	(2) Probit	(3) Logit
Means:			
Legal Immigrants	0.0448***	0.0446***	0.0448***
	(1.16e-05)	(1.11e-05)	(1.11e-05)
Undocumented	0.0139***	0.0139***	0.0139***
	(8.68e-06)	(8.61e-06)	(8.58e-06)
Difference in means	0.0308***	0.0307***	0.0308***
	(1.45e-05)	(1.40e-05)	(1.40e-05)
Share due to:			
Endowments	0.00643***	0.00590***	0.00591***
	(7.80e-06)	(1.16e-05)	(1.19e-05)
Coefficients	0.0219***	0.0195***	0.0193***
	(1.70e-05)	(1.46e-05)	(1.44e-05)
Interaction	0.00254***	0.00535***	0.00565***
	(1.25e-05)	(1.44e-05)	(1.45e-05)
Observations	76,831	76,831	76,831

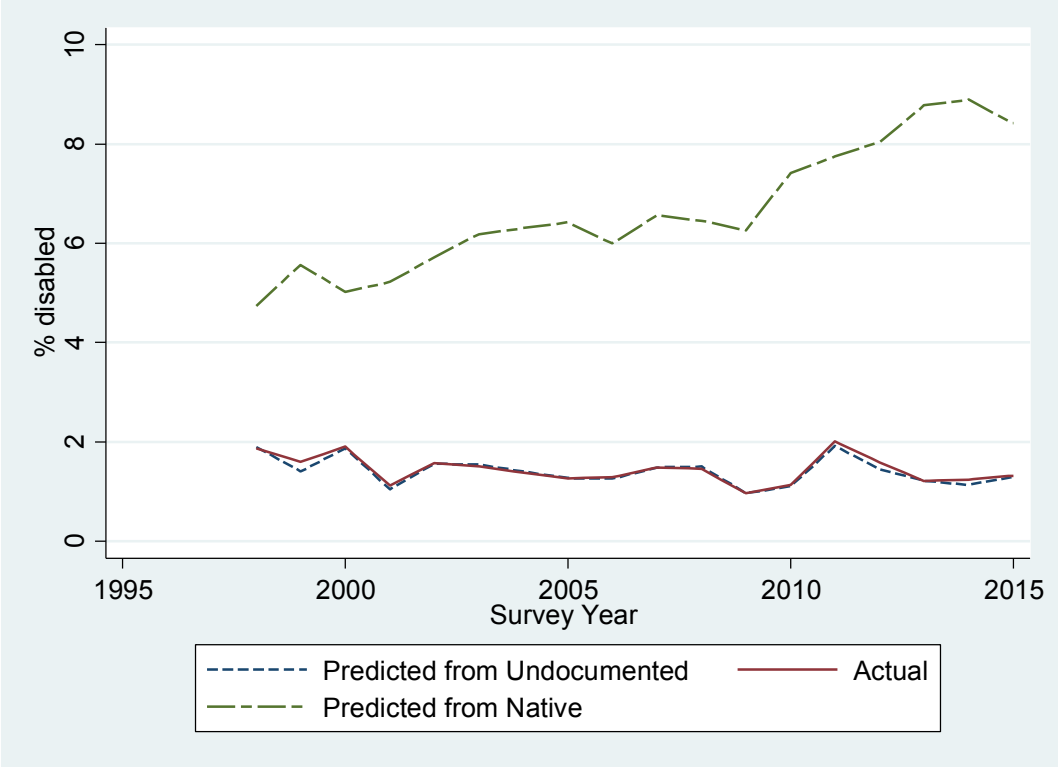
Notes: NHIS Sample Adult 18-65. Weighted. Legal immigrants and Undocumented only. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure A8: Predicted trend in Disability Rates for Native Born



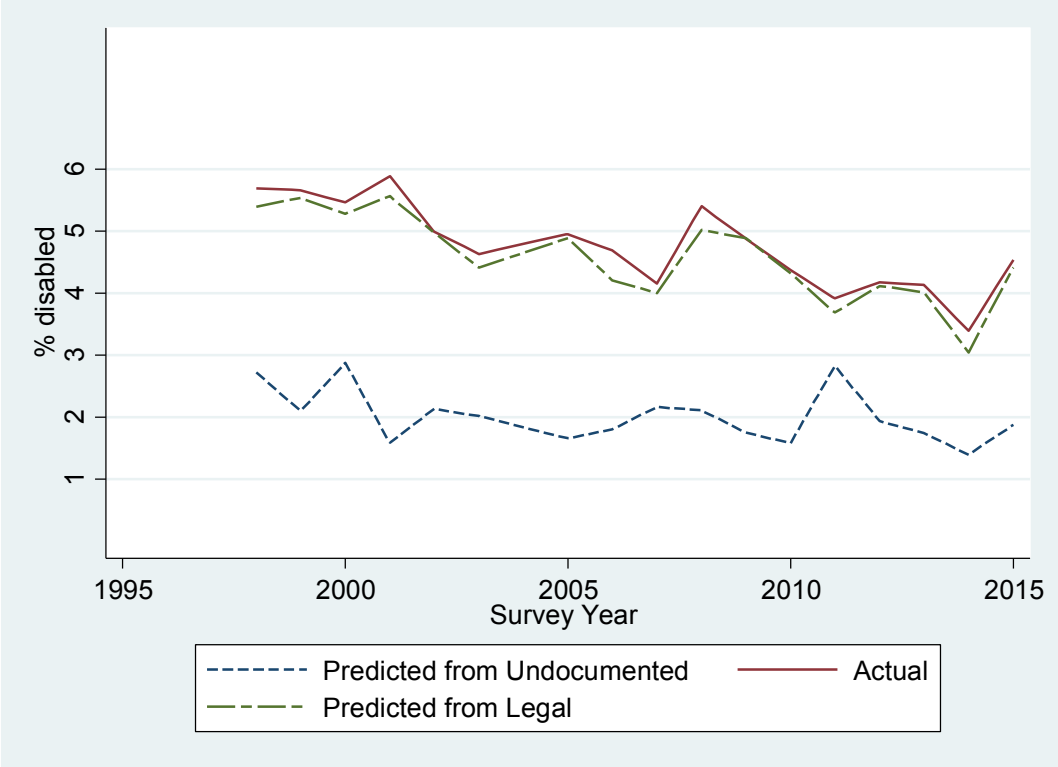
Notes: NHIS Sample Adult 18-65. Native Born and Undocumented only. Weighted. Uses Logit model.

Figure A9: Predicted trend in Disability Rates for Undocumented Immigrants



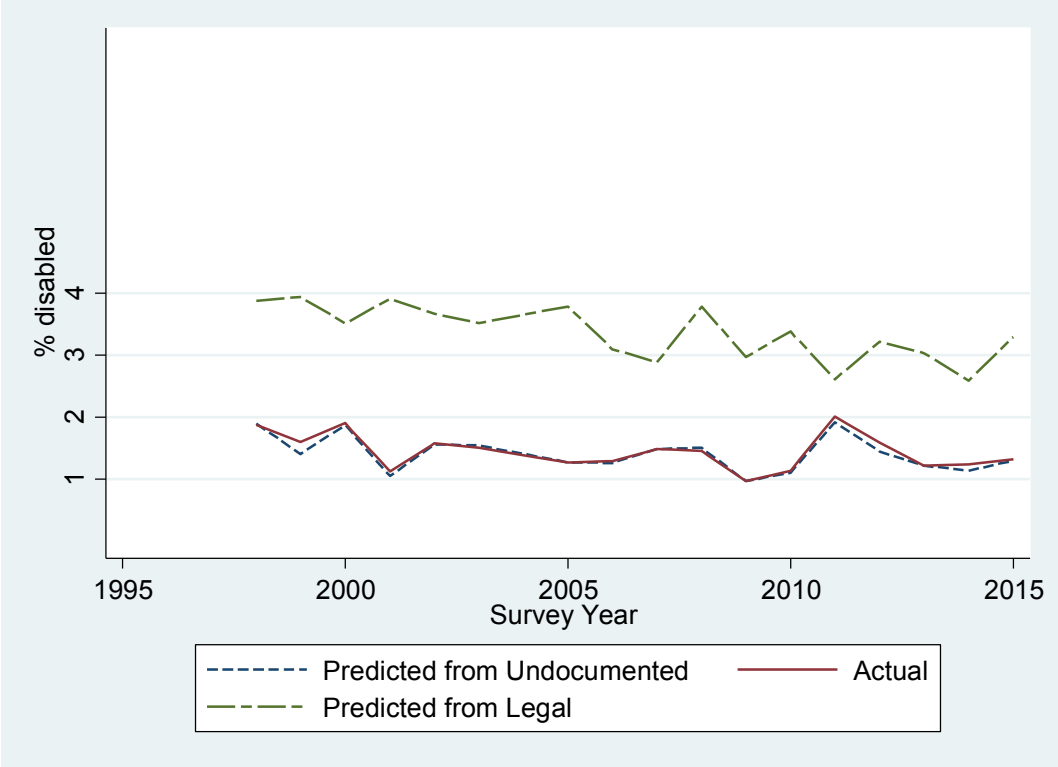
Notes: NHIS Sample Adult 18-65. Native Born and Undocumented only. Weighted. Uses Logit model.

Figure A10: Predicted trend in Disability Rates for Legal Immigrants



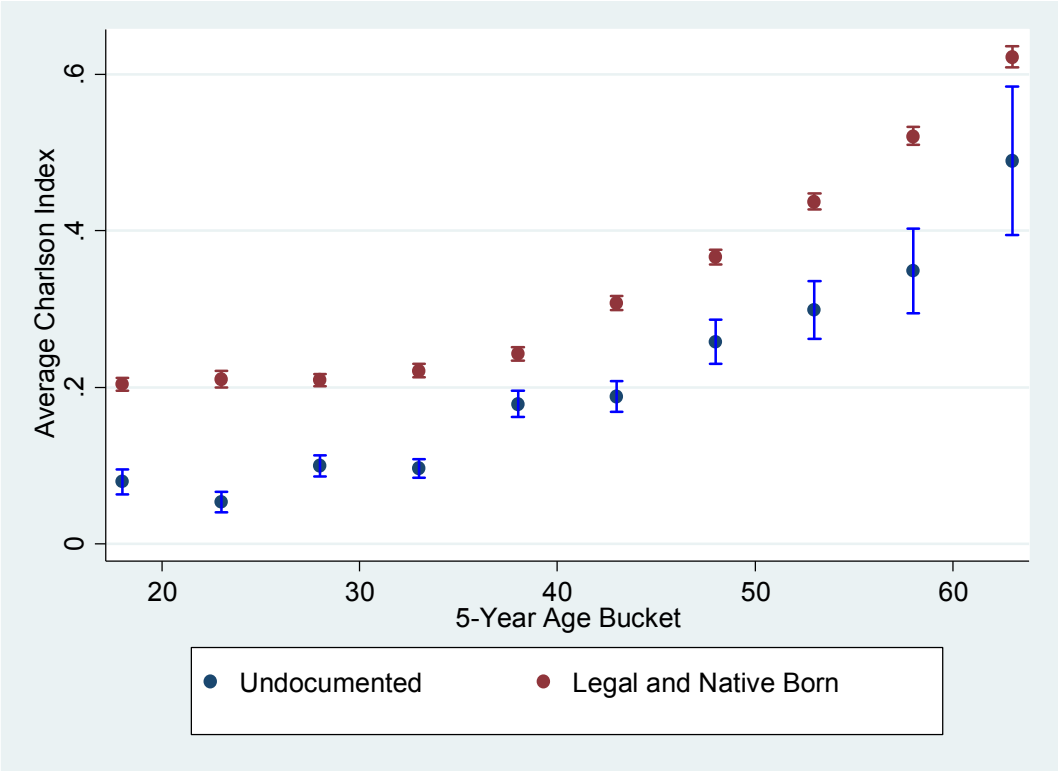
Notes: NHIS Sample Adult 18-65. Legal immigrants and Undocumented only. Weighted. Uses Logit model.

Figure A11: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: NHIS Sample Adult 18-65. Legal immigrants and Undocumented only. Weighted. Uses Logit model.

Figure A12: Charlson Index by Age for Undocumented and Legal Populations



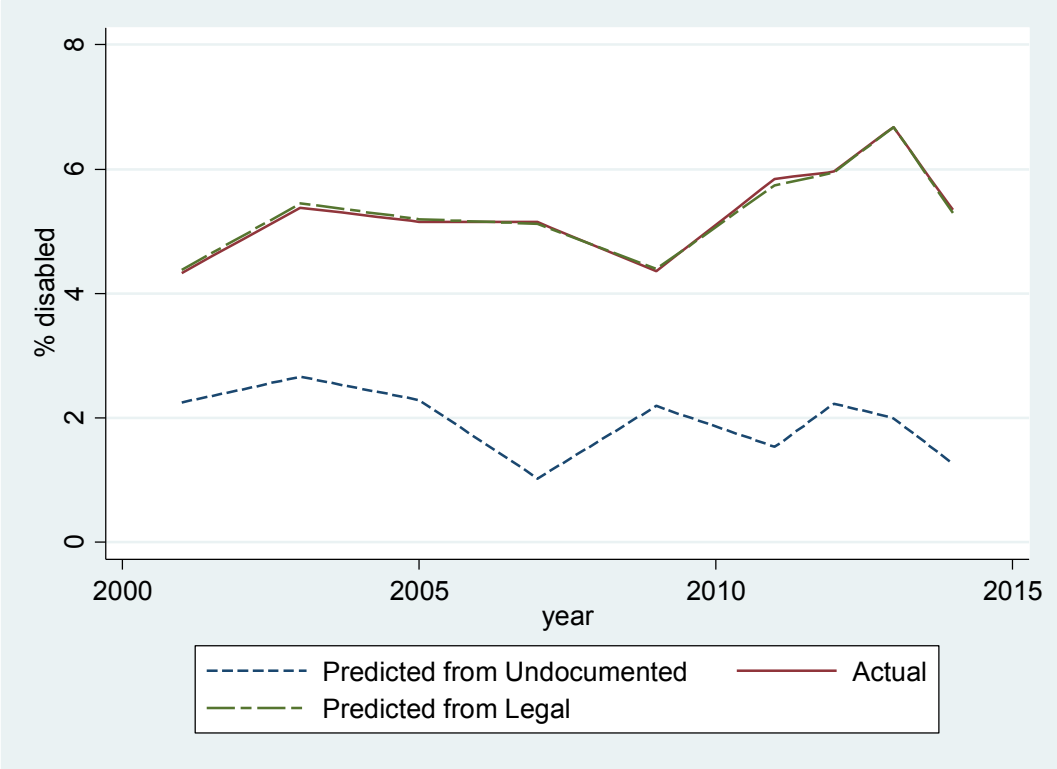
Notes: CHIS Adult Sample, 18-65. Weighted.

Table A5: Oaxaca Blinder

	(1) OLS	(2) Probit	(3) Logit
Means:			
Legal Immigrants	0.0540***	0.0540***	0.0540***
& Native Born	(1.66e-05)	(1.63e-05)	(1.63e-05)
Undocumented	0.0153***	0.0153***	0.0153***
	(2.37e-05)	(2.34e-05)	(2.33e-05)
Difference in	0.0387***	0.0386***	0.0387***
means	(2.89e-05)	(2.85e-05)	(2.85e-05)
Share due to:			
Endowments	0.00676***	0.00532***	0.00545***
	(2.53e-05)	(4.31e-05)	(4.53e-05)
Coefficients	0.0435***	0.0408***	0.0407***
	(3.40e-05)	(3.56e-05)	(3.53e-05)
Interaction	-0.0116***	-0.00750***	-0.00745***
	(3.15e-05)	(4.80e-05)	(4.98e-05)
Observations	245,422	245,422	245,422

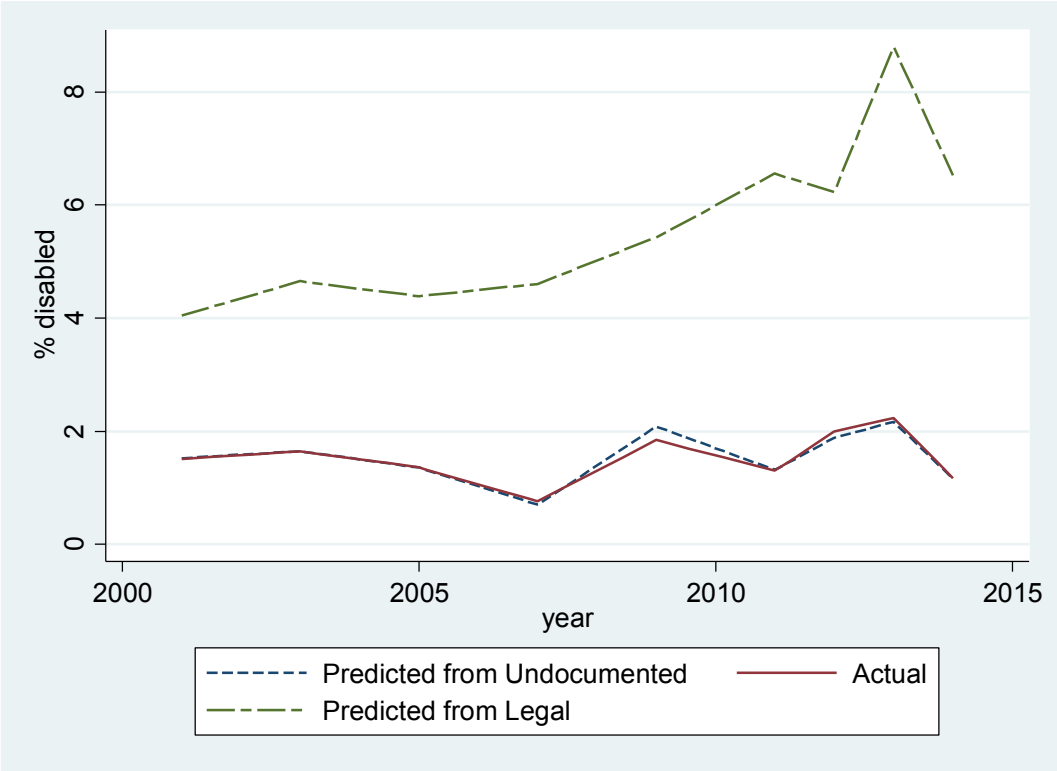
Notes: CHIS Sample Adult 18-65. Weighted. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure A13: Predictedtrend in Disability Rates for Legal Immigrants and Native Born



Notes: CHIS Adult Sample 18-65. Weighted. Uses Probit model.

Figure A14: Predicted trend in Disability Rates for Undocumented Immigrants



Notes: CHIS Adult Sample 18-65. Weighted. Uses Probit model.

Figure A15: California undocumented population from CPS and CHIS

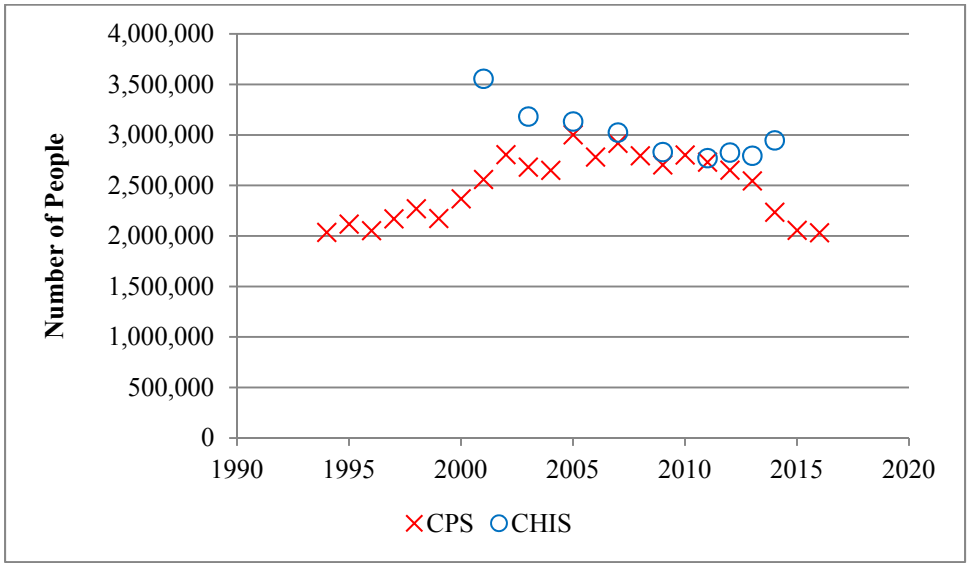
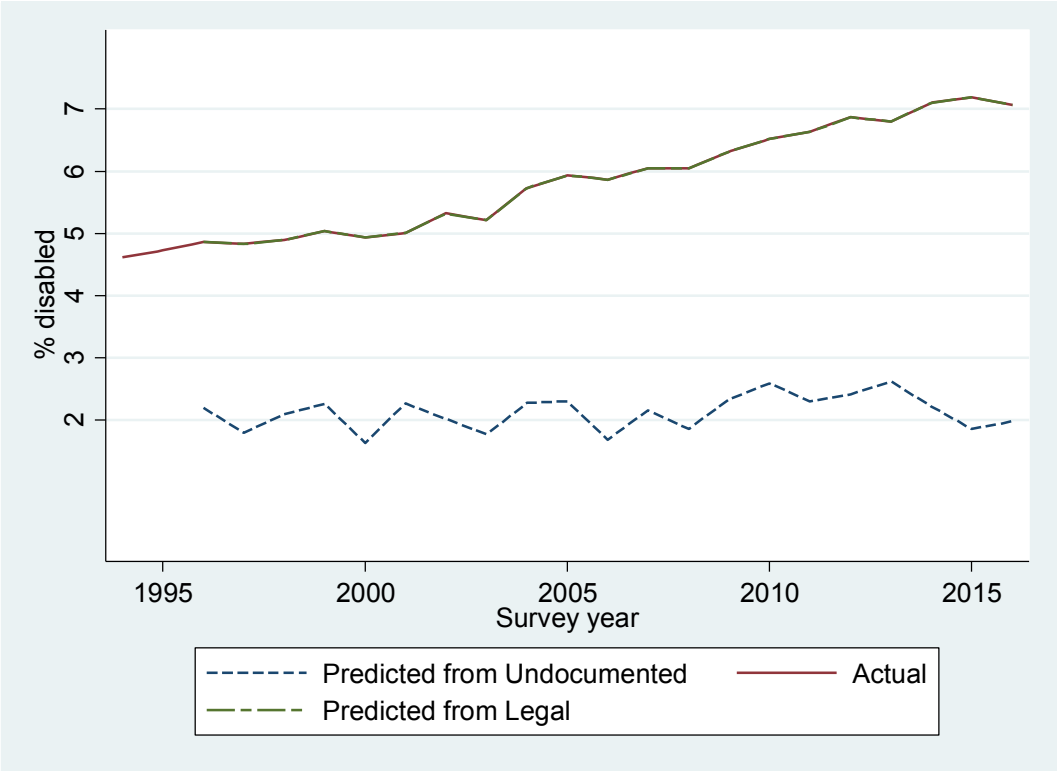
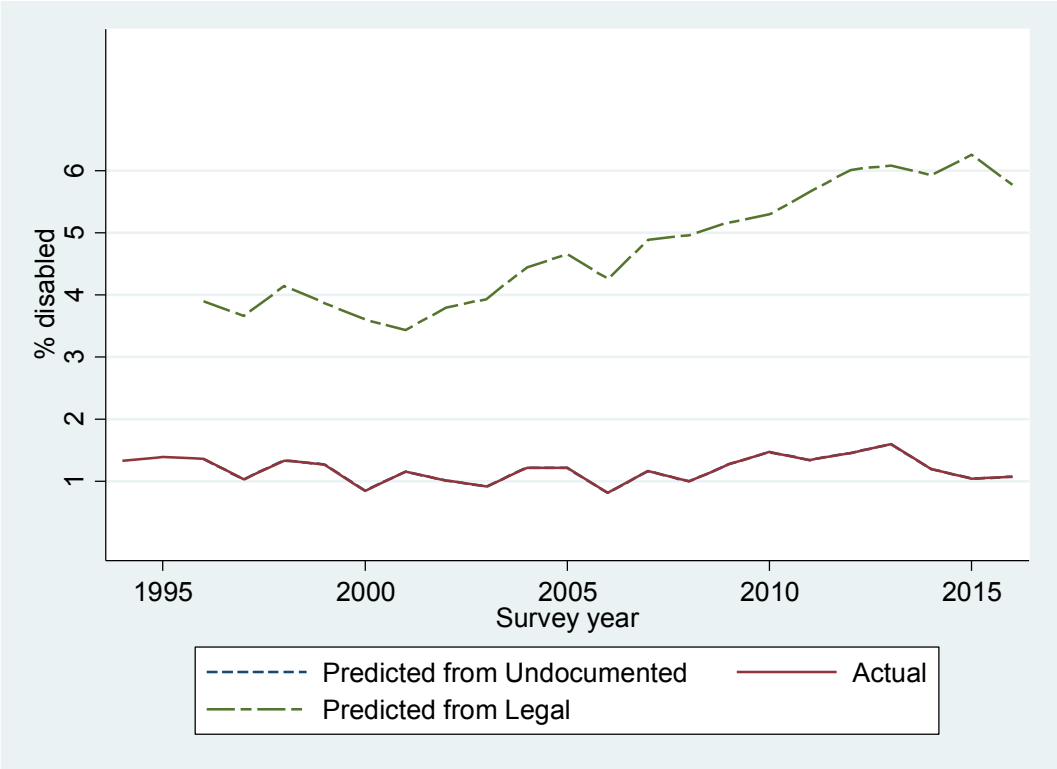


Figure A16: Predicted Trend in Disability Rates for Legal Immigrants and Native Born



Notes: CPS 18-65. Weighted. Uses Logit model.

Figure A17: Predicted Trend in Disability Rates for Undocumented Immigrants



Notes: CPS 18-65. Weighted. Uses Logit model.