Labour Economics 61 (2019) 101757

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

The wage penalty to undocumented immigration ☆

George J. Borjas[a,*], Hugh Cassidy[b]

[a] Harvard University, United States
[b] Kansas State University, United States

A B S T R A C T

This paper examines the determinants of the wage penalty experienced by undocumented workers, defined as the wage gap between observationally equivalent legal and undocumented immigrants. Using recently developed methods that impute undocumented status for foreign-born persons sampled in microdata surveys, the study documents a number of empirical findings. Although the unadjusted gap in the log hourly wage between the average undocumented and legal immigrant is very large (over 35%), almost all of this gap disappears once the calculation adjusts for differences in observable socioeconomic characteristics. The wage penalty to undocumented immigration for men was only about 4% in 2016. Nevertheless, there is sizable variation in the wage penalty over the life cycle, across demographic groups, across different legal environments, and across labor markets. The flat age-earnings profiles of undocumented immigrants, created partly by slower occupational mobility, implies a sizable increase in the wage penalty over the life cycle; the wage penalty falls when legal restrictions on the employment of undocumented immigrants are relaxed (as with DACA) and rises when restrictions are tightened (as with E-Verify); and the wage penalty responds to increases in the number of undocumented workers in the labor market, with the wage penalty being higher in those states with larger undocumented populations.

1. Introduction

The Department of Homeland Security (DHS) estimates that 12.1 million undocumented persons resided in the United States in January 2014 (Baker, 2017). In the past decade, Congress considered (but failed to enact) a number of proposals that would regularize the status of the undocumented population and provide a "path to citizenship." Given the large size of this population, any future change in its immigration status is bound to have significant effects on the labor market and the broader economy. But any evaluation that attempts to predict the economic impact of regularization immediately runs into a major roadblock: We know little about the economic status of the 12 million undocumented persons already living in the United States.

The study of the socioeconomic status of the undocumented is obviously hampered by the fact that no widely available microdata survey reports whether a particular foreign-born person is undocumented or not. In recent years, however, there has been progress in developing methods that attempt to impute the undocumented status of foreign-born persons at the individual level, such as the imputation algorithm for the Current Population Surveys (CPS) developed at the Pew Research Center or Warren's (2014) analogous exercise using the American Community Survey (ACS). These attempts build on the framework first proposed by Warren and Passel (1987) that attempts to estimate the size of the undocumented population. The Passel-Warren methodology, in fact, underlies the "official" estimates of this population as reported by DHS.

The Pew researchers essentially built an algorithm that considers various aspects of a person's demographic background to add a variable to a CPS microdata file indicating if a foreign-born person is "likely authorized" or "likely unauthorized" (Passel and Cohn, 2014). After being granted access to some of the Pew data files, Borjas (2017) used a variant of this algorithm to create an undocumented status identifier in all the post-1994 Current Population Surveys, and used these data to analyze the differences in labor supply behavior among undocumented immigrants, legal immigrants, and natives. The differences in work propensities were striking. Undocumented men had much larger labor force participation and employment rates than other groups in the population; the gap widened substantially over the past two decades; and the labor supply elasticity of undocumented men was close to zero, suggesting that their labor supply is very inelastic. In contrast, undocumented women had much lower participation and employment rates than other groups in the population.

This paper applies the algorithm to the American Community Survey (ACS) to measure the size and examine the determinants of the wage penalty to undocumented immigration.[1] Undocumented immigrants are likely to earn less than equally qualified legal immigrants simply because the undocumented have many fewer options in the labor market.

[1] The ACS data was downloaded from the Integrated Public Use Microdata Series (IPUMS) website. See Ruggles et al. (2018).

☆ Parts of this paper subsume research that appeared in separate unpublished working papers by Borjas (2016) and Cassidy (2017). We are particularly grateful to Mark Lopez and Jeffrey Passel of the Pew Research Center for their generosity in sharing data files. We have also benefited from the comments and reactions of Joan Llull, Joakim Ruist, and two referees.

* Corresponding author.
E-mail address: gborjas@harvard.edu (G.J. Borjas).

https://doi.org/10.1016/j.labeco.2019.101757
Received 10 September 2018; Received in revised form 11 August 2019; Accepted 12 August 2019
Available online 14 August 2019
0927-5371/

Not all jobs are available to undocumented immigrants, and the possibility of detection (and eventual deportation) may lead to exploitation of the undocumented by unscrupulous employers. Our analysis of the ACS data yields a number of potentially important findings:

(1) Although the unadjusted gap in the log hourly wage between undocumented workers and legal immigrants is large (around 35% for both men and women), much of the gap disappears after adjusting for differences in observable socioeconomic characteristics between the two groups. Two variables, educational attainment and English language proficiency, account for nearly half of the observed wage gap between the groups.
(2) The wage penalty to undocumented immigration declined between 2008 and 2016. In 2008, the wage penalty stood between 4 and 6% for both men and women. By 2016, the wage penalty had declined for both groups. Although it is difficult to ascertain why the *average* wage penalty in the national labor market has shrunk, the decline in the wage penalty coincides with the timing of actions by the Obama administration which led to a less restrictive approach to undocumented immigration. In fact, our evidence indicates that the wage penalty to specific groups of immigrants, such as those targeted by the Deferred Action for Childhood Arrivals (DACA) executive action, declined significantly after the relaxation of restrictions.
(3) The finding that the *average* wage penalty is relatively low hides a lot of variation in the penalty among different types of undocumented workers, and among undocumented workers employed in different labor markets. Not surprisingly, the (cross-section) age-earnings profile of undocumented workers lies far below that of legal immigrants (and, of course, of native workers). More strikingly, the age-earnings profile of undocumented workers is almost perfectly flat during much of the prime working years, As a result, the wage penalty for undocumented workers rises significantly over the life cycle.
(4) The evidence indicates that observationally equivalent undocumented workers and legal immigrants are not perfect substitutes. As a result, the wage penalty responds to increases in the relative size of the undocumented population. In particular, the wage penalty is larger in states with relatively larger undocumented populations: A 1 percentage point increase in the fraction of the state's workforce that is undocumented increases the wage penalty for men by about 1 percent. In addition, the wage penalty responds to the enactment of state-level legislation that restricts the employment of undocumented workers, with tighter restrictions leading to significantly larger wage penalties.

This diverse set of findings provides a foundation upon which any eventual analysis of the impact of alternative regularization proposals can be based. It is important to acknowledge at the outset, however, that the robustness of the evidence presented below depends on the validity of the procedure used to impute undocumented status at the micro level.

2. Imputing undocumented status in microdata files

Warren and Passel (1987) introduced the "residual" methodology used by the DHS to calculate the size of the undocumented population. The first step involves estimating how many legal immigrants should reside in the United States at a point in time. Over the years, immigration officials have tracked the number of legal immigrants admitted to the country (i.e., the number of "green cards" granted each year). Other immigration records allow us to determine how many foreign-born persons live in the United States temporarily (e.g., foreign students, business visitors, diplomats, etc.). These data enable us to apply mortality tables to the cumulative count of green cards and predict how many foreign-born persons should be legally residing in the United States at a point in time.

At the same time, many government surveys, such as the decadal census, enumerate the U.S. population and specifically ask where each person was born. These surveys provide estimates of how many foreign-born people are actually living in the country. In rough terms, the difference between the number of foreign-born persons who are actually living in the United States and the number of legal immigrants who should be living in the United States is the Warren–Passel (and now "official" DHS) estimate of the number of undocumented persons.[2]

Jeffrey Passel has continued to work on the enumeration and identification of undocumented immigrants over the past two decades. As a result of these efforts, Passel (and colleagues at the Pew Research Center) developed a comparable methodology that attempts to identify the undocumented immigrants at the *individual level* in survey data. This important extension of the Warren–Passel methodology relies on the same residual approach that was initially used to calculate the size of the undocumented population.

Passel and Cohn (2014) describe the methodology used to add an undocumented status identifier to the Annual Social and Economic Supplement (ASEC) files of the CPS. In rough terms, the algorithm identifies the foreign-born persons in the sample who are likely to be legal, and then classifies the residual group as likely to be undocumented. In closely related work, Warren (2014) used logical edits and other adjustments to impute the legal status of foreign-born persons in the ACS. After being granted access to the 2012–2013 CPS files created by Passel and Cohn (2014), Borjas (2017) "reverse-engineered" the approach and applied the algorithm to all available CPS files to examine the labor supply of undocumented immigrants. The residual method classifies a foreign-born person as a legal immigrant if any of the following conditions hold:

(a) that person arrived before 1980;
(b) that person is a citizen;
(c) that person receives Social Security benefits, SSI, Medicaid, Medicare, or Military Insurance;
(d) that person is a veteran, or is currently in the Armed Forces;
(e) that person works in the government sector;
(f) that person resides in public housing or receives rental subsidies, or that person is a spouse of someone who resides in public housing or receives rental subsidies;
(g) that person was born in Cuba (as practically all Cuban immigrants were granted refugee status);
(h) that person's occupation requires some form of licensing (such as physicians, registered nurses, air traffic controllers, and lawyers);
(i) that person's spouse is a legal immigrant or citizen.[3]

We use this algorithm to create a comparable undocumented status identifier in the American Community Survey (ACS) data beginning in 2008.[4] The only difference in the algorithms applied to the CPS and ACS data arises because the ACS does not identify whether a particular household is living in public housing or receiving subsidized rents, and thus we omit condition *f* from the imputation procedure for the ACS. As Fig. 1 shows, the predicted fraction of undocumented immigrants in the population at any particular age is roughly the same regardless of whether we use the Pew files in our possession (the 2012–2013 cross-sections) or the comparable ACS files, although the ACS tends to slightly overpredict the relative number of undocumented persons at younger

[2] Note that government surveys, including the decadal census, miss many people. Some of the people missed are undocumented immigrants who wish to avoid detection. To calculate an estimate of the size of the undocumented population, the Warren–Passel methodology requires an assumption about the undercount rate. The DHS assumes that the undercount for undocumented persons is 10% (Baker, 2017, p. 7).

[3] Condition *i* implies that a person who does not satisfy any condition between *a* and *h*, but whose spouse satisfies at least one of these conditions, would be considered legal by virtue of their spouse's legal status.

[4] Prior to 2008, the ACS also does not report information on Medicare or Medicaid receipt, so that the classification of undocumented status in the pre-2008 ACS requires further assumptions.

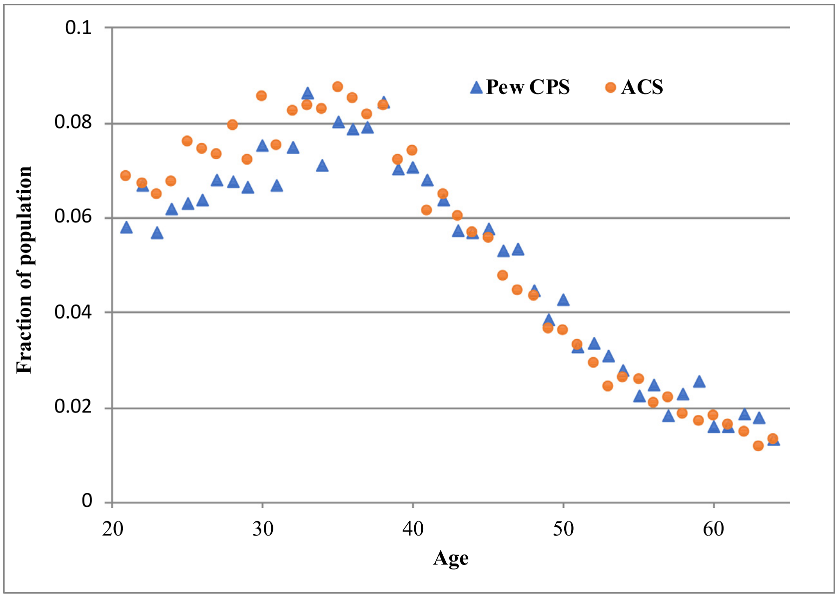


Fig. 1. Percent of population that is undocumented, by age. (Pooled 2012–2013 CPS-ASEC files, pooled 2011–2012 ACS). Notes: The figure calculates the percent of the population (at a particular age) that is foreign-born and is classified as undocumented using either the "likely unauthorized" status indicator created by Jeffrey Passel and colleagues at the Pew Research Center or my reconstruction of the undocumented status indicator in the ACS (see text for details).

Table 1
Comparison of summary statistics for male workers, 2012–2013.

	Natives	Legal No correction.	H1B correction	Undocumented No correction.	H1B correction.
A. Pew					
Percent of pop.	80.9	12.4	12.6	6.6	6.5
Average age	41.9	42.7	42.6	37.4	37.5
Education:					
High school dropouts	5.4	20.1	19.9	44.7	45.6
High school graduates	31.2	23.5	23.3	29.4	30.0
Some college	29.3	17.7	17.5	10.1	10.3
College graduates	23.4	21.3	21.5	9.1	8.6
Postcollege	10.8	17.3	17.8	6.6	5.5
State of residence:					
California	9.1	26.1	26.1	22.5	22.5
New York	5.4	11.1	11.0	6.7	6.8
Texas	8.1	10.0	9.9	14.8	14.9
Log wage gap	0.000	−0.070	−0.062	−0.438	−0.460
Sample size	66,632	15,794	15,936	7016	6874
B. ACS					
Percent of pop.	81.5	11.3	11.6	7.2	6.9
Average age	42.0	43.6	43.5	36.8	37.0
Education:					
High school dropouts	5.6	19.2	19.0	42.6	44.5
High school graduates	31.8	25.7	25.4	28.9	30.1
Some college	31.7	20.4	20.3	10.7	11.2
College graduates	21.0	19.2	19.3	9.4	7.9
Postcollege	10.0	15.5	16.0	8.4	6.3
Speaks English	–	58.4	58.5	29.5	27.4
State of residence:					
California	8.9	25.8	25.8	23.6	23.7
New York	5.4	11.0	11.0	8.2	8.4
Texas	7.7	9.9	9.9	13.5	13.7
Log wage gap	0.000	−0.040	−0.025	−0.398	−0.439
Sample size	980,270	121,699	124,433	60,889	58,155

Notes: The statistics are calculated in the sample of men aged 21–64 who are not enrolled in school, are not self-employed, and report positive wage and salary income, weeks worked, and usual hours worked weekly.

ages. The figure also shows that the life cycle trend in the fraction of persons who are imputed to be undocumented in the ACS closely tracks the fraction predicted in the original Pew CPS files.

To further document that our application of the algorithm to the ACS leads to very similar results as those implied by the Pew CPS files, Table 1 reports summary statistics for the samples of male working natives, legal immigrants, and undocumented persons in both the Pew CPS and the ACS 2012–2013 cross-sections. The corresponding results for women are reported in Appendix Table A1. The sample is restricted to persons aged 21–64 who are not enrolled in school, and who report positive wage and salary income in the previous calendar year, positive weeks worked, and positive usual hours worked weekly.

As illustrated in Table 1, the Pew residual method suggests that a strikingly large number of undocumented workers have high levels of

educational attainment. For example, 17.8% of the undocumented male population in the ACS have at least a college degree. Although this surprising result has not been explored in any of the previous studies that impute an undocumented status indicator in micro data, we suspect that the typical imputation algorithm misclassifies many highly educated immigrant workers. Specifically, the algorithms do not "filter out" the large sample of high-skill immigrants who are in the United States temporarily under the auspices of the "high tech" H-1B program. In fact, Albert (2019) reports that the algorithm (and the Pew methodology it is based on), while very accurate for low-skilled immigrants, mistakenly classified around 25% of college educated immigrants as undocumented. This inaccuracy suggests that accounting for the legal status of H-1B immigrants may be appropriate.

In this paper, we refine the Pew algorithm by adding an additional filter to the list above, further classifying a person as a legal immigrant if he or she is likely to be an H-1B visa-holder. Specifically, we assume that a foreign-born person is likely to be in the country with an H-1B visa if: (1) the immigrant works in an occupation that commonly employs H-1B visa holders (such as computer programmer)[5]; (2) the immigrant has resided in the United States for six years or fewer (i.e., the maximum length of time an H-1B visa is valid); and (3) the immigrant is at least a college graduate. As Table 1 shows, the application of the H-1B filter reduces the fraction of undocumented immigrant men with at least a college degree from 17.8 to 14.2%. Note, however, that both the original Pew files and our imputation in the ACS still produce a relatively large number of undocumented workers with high levels of educational attainment. We use the H-1B filter throughout the empirical analysis reported in this paper.[6]

There is a lot of similarity in the socioeconomic characteristics of the three demographic groups across the two data extracts. Among men, for example, the fraction of the population that is undocumented is 6.5% in the Pew CPS and 6.9% in the ACS. The average age of undocumented immigrants is the same in the two files (about 37 years). And 45.2 of undocumented men in the Pew files are high school dropouts, as compared to 44.0% in the ACS files.

We also calculated the hourly wage rate for each worker in the sample (defined as wage and salary income divided by the product of weeks worked in the past year and usual hours worked weekly). Table 1 also shows that the log wage gap between undocumented workers and natives is similar across the data sets. The wage disadvantage of undocumented men is −0.398 log points in the Pew CPS and −0.414 log points in the ACS data (equivalent to about a 33% wage gap between the two groups). The comparable statistics reported in Table A1 for undocumented women imply that an equally large wage disadvantage (of about −0.385 log points in the ACS).

The validity of the evidence presented below hinges on the accuracy of the undocumented status indicator in the original Pew algorithm. In the absence of administrative data on the characteristics of the undocumented population, it is not possible to quantify the direction and magnitude of any potential bias. We can, however, compare key socioeconomic characteristics in our sample with comparable data in samples of undocumented immigrants created by other researchers using different methods. For instance, the Center for Migration Studies (CMS) has also developed an analogous method of imputing legal status in the ACS (Warren, 2014).[7] The CMS method uses individual characteristics (including birthplace, occupation, or the receipt of public benefits) to classify some immigrants as likely legal. The CMS also makes further adjustments by country of origin and incorporates a correction for the undercount of undocumented persons (our methodology does *not* perform any reweighting).[8] Table 2, adapted from Warren (2014, Table 2), compares the total predicted size of the undocumented population as well as its geographic distribution using alternative methods, and adds results from our own imputation in the ACS. It is evident that the geographic distribution of undocumented immigrants in our imputed ACS data is broadly consistent with the distribution predicted by the four alternative methodologies summarized in the Warren study. It seems, therefore, that our approach closely duplicates the undocumented population examined in other studies.

3. Estimating the wage penalty

We calculate the wage penalty to undocumented status by estimating the following Mincerian wage regression in the sample of working immigrants:

$$\log w_i = \beta h_i + \beta_L L_i + \varepsilon_i, \tag{1}$$

where w_i is the hourly wage rate of worker *i; h_i* is a vector of socioeconomic characteristics that affect earnings; and L_i is a dummy variable that equals one if the worker is a legal immigrant. The coefficient β_L measures the wage penalty, with a positive value indicating the earnings advantage enjoyed by legal immigrants over observationally equivalent undocumented workers.

It is also possible to calculate the wage penalty by instead performing an Oaxaca–Blinder decomposition that yields the predicted wage disadvantage of the average undocumented immigrant arising from differential treatment in the labor market (i.e., allows the coefficient vector β, the returns to socioeconomic charactersitics, to vary by legal status). One alternative definition of the wage penalty would calculate by how much the earnings of the average undocumented immigrant increased if he or she were "treated" just like an observationally equivalent legal immigrant in the labor market.

It is unlikely, however, that observationally equivalent legal immigrants and undocumented immigrants are perfect substitutes in production (and the empirical evidence reported below indeed shows that they are not). Even putting aside the possibility the two groups might have different unobservable skill sets, legal restrictions prevent employers from viewing one type of immigrant as a clone of the other type.[9] As a result, the relative number of undocumented immigrants in a particular labor market actually affects the structure of wages (and hence the

[5] The list of occupations assumed to commonly employ H-1B visa holders are computer and information system managers; computer and mathematical occupations; architecture and engineering occupations; and postsecondary teachers. These occupations account for over 80% of all H-1B petitions filed in 2017 (U.S. Citizenship and Immigration Services, 2018a).

[6] Our H-1B filter identifies 598,000 foreign-born persons as H-1B visa holders, which is in the ballpark of what one would expect to be the steady-state number of that population (i.e., the visa is capped at 85,000 visas per year, and the visa lasts 6 years). It may be that H-1B visa holders stay in the country beyond the sixth year while waiting to adjust their status because of country-specific quotas on the number of green cards available. An alternative filter might define H-1B status only by education and occupation. However, the predicted number of H-1B visa holders if one ignores the 6-year limitation is 2.1 million, which seems far too large to be consistent with the number that is expected to reside in the country.

[7] See also Van Hook et al. (2015), which evaluates various methods of imputing the legal status of immigrants using Monte Carlo simulations.

[8] Specifically, the CMS algorithm assigns each immigrant a likely legal status based on individual characteristics, in a manner similar to our approach. Two additional steps are then performed: (1) likely undocumented are randomly sampled at a rate that varies by country of origin; and (2) undercounting of undocumented immigrants is accounted for by re-weighting the microdata, depending on year of arrival.

[9] Cotton (1988, p.238) makes a related point in the context of measuring racial wage discrimination. In his discussion of whether to use the black or the white regression coefficients to measure discrimination he writes: "…each assumption abstracts from the central reality of wage and other forms of economic discrimination: not only is the group discriminated against undervalued, but the preferred group is overvalued, and the undervaluation of the one subsidizes the overvaluation of the other. Thus, the white and black wage structures are both functions of discrimination and we would not expect either to prevail in the absence of discrimination."

Table 2
Geographic distribution of the undocumented population in 2010, using alternative methodologies.

	Borjas–Cassidy	CMS	Warren and Warren	DHS	Pew Research Center
US Total (thousands)	12,256	11,725	11,725	11,570	11,400
Distribution by state (%):					
California	23.6	24.9	25.0	25.2	21.9
Texas	13.7	14.7	13.7	15.4	14.5
New York	7.9	7.8	6.0	6.0	7.0
Florida	7.3	6.7	8.5	6.3	7.9
Illinois	4.8	5.1	5.0	4.8	4.4
New Jersey	4.3	4.1	3.5	3.8	4.4
Georgia	3.4	3.4	3.4	3.7	n/a
North Carolina	2.8	2.9	3.2	3.4	n/a
Arizona	2.4	2.6	2.9	3.0	n/a
Washington	2.0	2.0	2.2	2.2	n/a
Other states and DC	27.7	25.9	26.6	26.3	n/a

Notes: The first column shows the geographic distribution of the undocumented by state in 2010 applying our methodology across the whole population. The remaining columns show the distribution using alternative methodologies, with data derived from Table 2 in Warren (2014).

coefficient vector β) for both groups.[10] Any large-scale legalization initiative would then influence the wage-setting decisions by employers and change the β vectors for both legal and undocumented workers. Running a Mincerian wage regression on the pooled sample of legal and undocumented immigrants, where the vector β gives the returns to socioeconomic characteristics for the average worker, bypasses this problem.[11]

To document that our application of the Pew residual method to the ACS does not alter the nature of the empirical evidence, we initially focus on the 2012–2013 period. As noted earlier, we have access to the pooled 2012–2013 CPS files created by the Pew Research Center, which allows us to compare the estimates of the wage penalty in those years to those obtained in the ACS. After we establish the similarity between the estimates, we can then expand the analysis to other periods and other samples in the much larger ACS data files.[12] To simplify the presentation, we pool the two cross-sections and treat them as a single data set.

Table 3 reports the wage penalty results. For the Pew data, we only report a single specification where the vector h includes age, state of residence, years since migration, educational attainment, and country of birth.[13] For the ACS regression, we add a vector of fixed effects that characterizes the worker's English language proficiency, a variable that is not available in the CPS but which is likely an important component of an immigrant's human capital stock. In fact, there are sizable differences between the English language skills of undocumented and legal immigrants, with legal immigrants being far more proficient. The ACS data indicate that 16.3% of undocumented immigrants reported not speaking English at all, as compared to only 4.1% of legal immigrants. Similarly, 59.4% of legal immigrants reported they spoke either only English or English "very well," as compared to only 28.9% of undocumented immigrants.

The top three rows of Table 3 show the overall "raw" difference in log wages between legal and undocumented immigrants, the wage gap that is explained by the control variables, and the unexplained portion, which is our estimate of the wage penalty.

There are several interesting findings in the table. First, the Pew CPS and ACS data generate very similar estimates of the raw wage gap between legal and undocumented immigrants, as well as of the adjusted wage penalty. Among men, for example, the raw wage gap is approximately 39.8% in the Pew CPS and 41.3% in the ACS. Adjusting for age, state of residence, years since migration, educational attainment, and country of birth implies an estimated wage penalty of 6.0% in the Pew CPS and of 8.6% in the ACS. Among women, the estimated wage penalty is 4.6% in the CPS and 6.3% in the ACS. In short, our application of the residual methodology to the ACS data yields similar estimates of the wage penalty as those obtained in the Pew CPS files.

It turns out, however, that these estimates of the wage penalty are "too big," as adding English language proficiency fixed effects to the regression model further reduces the wage penalty in the ACS, from 8.6 to 6.1% for men and from 6.3 to 4.2% for women. In short, after controlling for an extensive set of observable individual characteristics, we find there is a positive and significant wage penalty to undocumented immigration, but it is numerically small—on the order of 4–6%.[14]

This striking finding raises a number of interesting questions. For example, *which* differences in observable characteristics play a larger role in generating the observed wage gap between legal immigrants and undocumented workers? In other words, while introducing the full set of characteristics dramatically lowers the estimate of the wage penalty β_L, how much does each set of covariates contribute to the reduction?

Gelbach (2016) presents a methodology that allows us to decompose the contribution of each set of covariates (e.g., education) to the change in the estimated wage penalty. The advantage of this approach over the more common procedure of sequentially adding each set of covariates

10 Ortega, Edwards, and Hsin (2018) simulate the impact of DACA on the wage structure of both legal and undocumented workers who do not change status. Because of the small number of DACA recipients relative to the immigrant population, the effects are minimal.

11 The estimate of the wage penalty given by the regression in Eq. (1) is numerically identical to that implied by the Oaxaca-Blinder decomposition method if the reference coefficients for the socioeconomic characteristics are estimated on the pooled sample of legal and undocumented immigrants.

12 Because the CPS reports earnings in the previous calendar year, the analysis uses the comparable 2011 and 2012 cross-sections of the ACS.

13 Age is included as a vector of fixed effects indicating a worker's age 5-year bands (20–24, 25-29, and so on); state of residence is included as a vector of 51 fixed effects; years since migration is included as a fourth-order polynomial; educational attainment is included as a vector of fixed effects indicating if the worker has less than 12 years of schooling, 12 years, 13–15 years, 16 years, or more than 16 years; and country of birth is included as a vector of fixed effects using all the information in the CPS or ACS data. The vector of fixed effects indicating English proficiency uses all the information contained in the English language variable in the ACS.

14 Ortega and Hsin (2018) use the ACS data from 2010–2012 which contains legal status based on the CMS methodology. The authors find that, due to occupational barriers, lacking legal status reduces undocumented immigrants' productivity by 12%. They also find wage gaps (see Table 4 of their paper) between legal and undocumented immigrants that are larger than those reported in our paper, though their imputation methodology does not correct for potential H-1B immigrants. Hotchkiss and Quispe-Agnoli (2013), who identify undocumented workers using state administrative data, also find that the large difference in wages between legal and undocumented immigrants is mostly attributable to differences in observed characteristics.

Table 3
Wage penalty to undocumented status in the 2012–2013 cross-section.

	Men			Women		
	Pew	ACS (1)	(2)	Pew	ACS (1)	(2)
Difference	0.398	0.413	0.413	0.358	0.385	0.385
	(0.009)	(0.003)	(0.003)	(0.011)	(0.004)	(0.004)
Explained	0.338	0.327	0.352	0.311	0.322	0.343
	(0.005)	(0.003)	(0.003)	(0.009)	(0.003)	(0.003)
Unexplained	0.060	0.086	0.061	0.046	0.063	0.042
	(0.009)	(0.003)	(0.003)	(0.011)	(0.004)	(0.004)
Fraction explained by:						
Age	0.011	0.021	0.035	−0.002	0.005	0.016
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
State of residence	0.002	0.003	0.004	0.007	0.009	0.009
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
YSM	0.053	0.080	0.057	0.061	0.090	0.066
	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)
Education	0.195	0.168	0.144	0.171	0.160	0.136
	(0.005)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)
Birthplace	0.078	0.056	0.042	0.075	0.059	0.048
	(0.005)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)
English	–	–	0.071	–	–	0.067
	–	–	(0.001)	–	–	(0.002)

Notes: Standard errors are reported in parentheses. The dependent variable gives a worker's log hourly wage rate. The statistics reported in the table are the results from a Mincerian wage regression that includes controls for survey year, age, educational attainment, state of residence, years-since-migration, and birthplace, while ACS columns (2) add English language proficiency. The rows labeled "Difference", "Explained", and "Unexplained" indicate the raw wage gap between legal and undocumented immigrants, the amount of that gap that is explained by the covariates, and the amount that remains unexplained, respectively. Each covariate row under "Fraction explained by:" indicate the fraction of the explained portion of the wage gap explained by that set of covariates (Gelbach, 2016). The years-since-migration variable is introduced as a fourth order polynomial; the age, education, state of residence, birthplace, and English language proficiency variables are introduced as vectors of fixed effects.

and simply documenting the change in the coefficient is that the Gelbach methodology accounts for the correlations among sets of covariates. In the presence of such correlations, the order in which each set of covariates is added impacts the interpretation of the results, whereas the Gelbach decomposition is independent of any sequential introduction of sets of covariates.[15]

The bottom panel of Table 3 reports the part of the wage gap "explained" by each of the covariate groups in our regression model. For example, differences between the two groups in the values of the covariate group "age" (which stands for a vector of nine age fixed effects) leads to a 3.5 percentage point wage gap for men, while the covariate group "state of residence" generates only a 0.4 percentage point wage gap. It is evident that the covariate groups that "matter," in terms of explaining a large part of the observed wage gap, are years since migration (with undocumented immigrants having been in the United States for a shorter period), educational attainment, and English proficiency. For men, these three sets of variables together generate a 27.2% wage gap, about two-thirds of what is actually observed; and differences in educational attainment alone generate a 14.4% wage gap, about a third of what is actually observed. Similar results are obtained for women.[16]

Having established the similarity between the Pew CPS and the ACS results, we can now extend the analysis to other ACS cross sections and subgroups of the population. We first explore how the wage penalty evolved over the past decade. Specifically, we conduct our decomposition exercise separately in each of the ACS cross-sections between 2008 and 2016, using the full model specification that includes English language proficiency. The top panel of Fig. 2 illustrates the trend in the wage penalty for the entire male workforce, as well as for low-skill (i.e., at most a high school education) and high-skill (i.e., at least some college) workers.[17] The bottom panel of the figure duplicates the analysis for the female workforce.[18]

It turns out that the wage penalty for undocumented men was relatively stable at about 5–6% through 2013, at which time it began a noticeable, and statistically significant, decline. In 2013, for example, the wage penalty for the average male worker was 6.7 percentage points (with a standard error of 0.6), but it declined to 4.1 percentage points by 2016 (with a standard error of 0.6).

The figure also illustrates the analogous trends in the wage penalty for low- and high-skill workers. Both groups exhibit the post-2013 decline in the wage penalty, with the decline being steeper for high skill undocumented workers. The wage penalty for low-skill workers stood at 8.3% in 2013, before beginning its decline and ending up at 6.6% in 2016. In contrast, the wage penalty for high-skill workers was 6.1% 2013, but by 2016 had declined to 2.7%. As the descriptive statistics reported in Table 1 show, there are a surprisingly large number of high-skill workers in the undocumented population. Both the Pew CPS and the ACS suggest that about 14% of undocumented men have at least a college diploma (even after applying the filter for H-1B status), and that an additional about 11% have some college education. The debate over undocumented immigration in the United States has focused on its

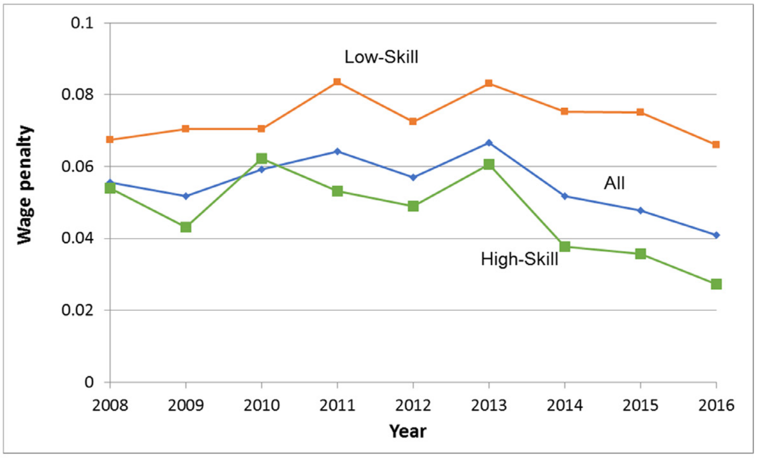
[15] We use the Stata package "b1x2" to perform the decomposition.

[16] Adding occupation controls to the decomposition further lowers the wage penalty to 2.7% for men and to near zero for women in the ACS, and occupation explains 15.3 and 17.7 percentage points of the wage gap between legal and undocumented immigrants for men and women, respectively.

[17] The standard error of the wage penalty in any given year is about 0.006 for men and 0.008 for women.

[18] The wage penalty for low- or high-skill workers is calculated by estimating the regression model separately in the samples of low- or high-skill workers.

A. Men



B. Women

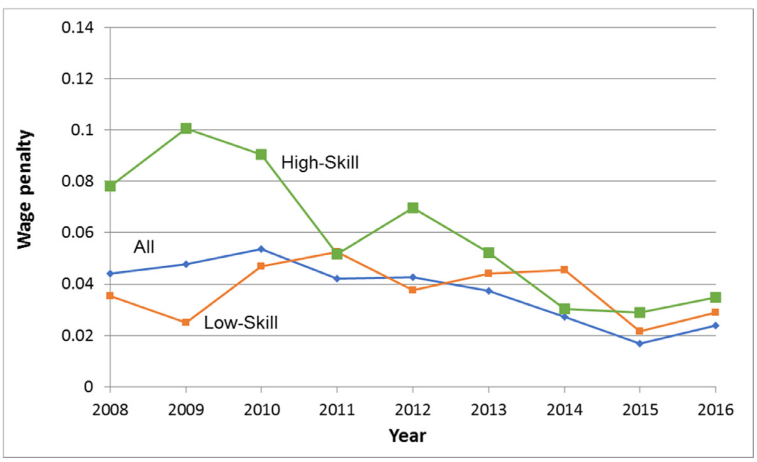


Fig. 2. Trend in the wage penalty for undocumented workers. Notes: Figures show the log hourly wage penalty between legal and undocumented immigrants calculated with Mincerian wage regressions estimated separately in each cross-section that include controls for age, educational attainment, state of residence, years-since-migration, birthplace, and English language proficiency. The wage penalty values shown are the coefficients on legal status. "Low-skill" and "high-skill" include workers who are high school graduates or less and workers with more than a high school degree, respectively. All results calculated from the ACS.

impact on the low-skill labor market, and the presence and labor market impact of high-skill undocumented immigrants has been ignored.

The bottom panel of Fig. 2 illustrates the analogous trends in the wage penalty estimated in the sample of women. As with men, the key finding is that there has been a long-term decline in the average wage penalty to undocumented women, with the decline beginning a bit earlier (around 2010). In 2010, the wage penalty for women stood at over 5%. By 2016, it had fallen to about 2%. The decline in the female wage penalty was also steeper for high-skill women. As noted earlier, however, undocumented women have very low employment rates, so that it is difficult to disentangle the impact of self-selection biases in the labor supply decision from secular trends in the wage penalty.[19]

It is of interest to compare our estimate of the wage penalty obtained from adding an undocumented identifier to the ACS to existing estimates of how much legalization raises the wage of undocumented workers. Almost all existing estimates of this wage penalty come from studies that examine what happened to the earnings of the persons who received amnesty in 1986 as part of the Immigration Reform and Control Act (IRCA). Nearly 3 million undocumented immigrants received amnesty at the time, and contemporaneous surveys tracked those immigrants as they received their legal working papers (Kossoudji and Cobb-Clark, 2002; and Kaushal. 2006). Their wage rose by at most 6% between 1989 and 1992. The estimates of the wage penalty implied by the ACS around 2008 (the earliest year available where the ACS provides the requisite information required to identify undocumented status), are very similar (around 4–6%). In short, the existing estimates of the wage penalty (based on measuring the wage impact of the IRCA amnesty) closely resemble the penalty implied by the wage data in the early years of our ACS cross-sections.[20]

It is difficult to identify precisely which factor drove the decline in the wage penalty in the national labor market after 2013.[21] A number

[19] We also estimated the wage penalty and its trend using the alternative approach of holding constant the demographic composition of the immigrant population, and then using those fixed characteristics to compute the average wage for legal and undocumented immigrants in each ACS cross-section. Specifically, we calculated (by gender) the distribution of immigrants across demographic cells using the pooled 2008-2016 ACS (where the cells are defined in terms of education, English language proficiency, age, years-since-migration, and state of residence). We then use those shares to get a weighted average of the log wage for legal and undocumented immigrants each year. This approach also reveals a decline of 3–+5 percentage points in the wage penalty starting around 2012 or 2013.

[20] Rivera-Batiz (1999, p. 106) looks specifically at Mexican undocumented immigrants using the 1990 Census. His results are similar to those reported in this paper, and he concludes that: "The most important characteristics in explaining the wage gap are: schooling, English proficiency, and recency of immigration."

[21] One stumbling block is that the composition of the undocumented population has changed in unknown ways during this period. The estimated number of undocumented immigrants (as reported by the DHS) rose between 2000 and 2006, and held relatively steady through 2016. The constant number of un-

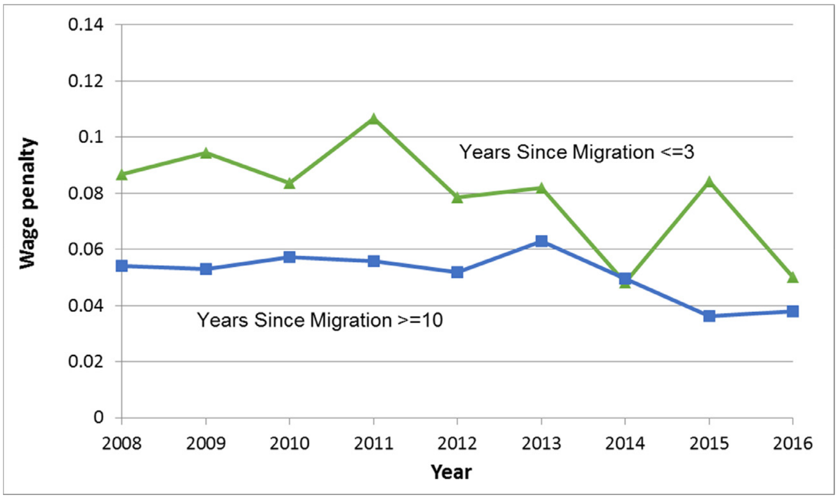


Fig. 3. Trend in the wage penalty for undocumented workers in specific cohorts.
Notes: Figures show the log hourly wage penalty between legal and undocumented immigrants calculated with Mincerian wage regressions estimated separately in each cross-section that include controls for age, educational attainment, state of residence, years-since-migration, birthplace, and English language proficiency. The wage penalty values shown are the coefficients on legal status. "Low-skill" and "high-skill" include workers who are high school graduates or less and workers with more than a high school degree, respectively. All results calculated from the ACS.

of sensitivity exercises can be conducted, however, that help to further identify the groups that experienced a substantial decline in the wage penalty and that may suggest a potential source for the decline. For example, we can examine what happened to the entry wage disadvantage of *new* undocumented immigrants over the past decade. We define a recent immigrant as someone who arrived in the 3-year period prior to the ACS cross-section, and we define an "older" immigrant as someone who has been in the United States more than 10 years. Because of the small number of "new" immigrants (only about 5% of legal immigrants and 12% of undocumented immigrants are recent arrivals), we pool the sample of male and female workers to calculate the wage penalty.[22] Fig. 3 illustrate the wage trends for the new and the earlier immigrants.

It is evident that the wage penalty associated with undocumented status for the newly arrived immigrants shrank substantially in the post-2011 period. The wage penalty to new immigrants fell from 10.7% in 2011 to 5.0% (with a standard error of 1.5) by 2016. In contrast, the trend in the wage penalty accruing to undocumented immigrants who have been in the United States more than 10 years was more stable, with the wage penalty declining by only about 2 percentage points (from about 6% in 2013 to 4% in 2016).

One plausible explanation for the decline in the wage penalty for the newly arrived immigrants is that there was a favorable shift in the legal environment regarding undocumented immigration during the years of the Obama administration. It seems plausible to argue that the shift would particularly benefit newly arrived immigrants, as they better represent the "marginal" worker in the labor market that will most quickly be affected by the implied changes in the legal environment. Unfortunately, the time-series giving the trend in the national wage penalty do not provide sufficient information that would help identify the impact of such *economy-wide* changes in the labor market for undocumented workers. There is evidence, however, suggesting that changes in the legal environment at the federal level *do* affect the national wage penalty (and we will show below that corresponding changes in the local legal environment also influence the wage penalty in the local labor market).

On June 15, 2012, President Barack Obama issued an executive action that grants undocumented immigrants who entered the United States as children a temporary reprieve from the threat of deportation as long as some eligibility requirements were met. The undocumented persons who qualify for the Deferred Action for Childhood Arrivals (DACA) program are immigrants who entered the United States under the age of 16, were at most 31 years old at the time the executive action was taken, and had at least a high school (or equivalent) education. The executive action permits these immigrants to work as if they were legal immigrants. In other words, the DACA program potentially represents a substantial change in labor market opportunities for the eligible undocumented workers in the national labor market, and it would be important to determine if it led to a reduction in the wage penalty for the affected workers.

We can use the ACS data to determine if the wage penalty for the DACA-eligible population fell towards the end of our sample period.[23] Because of the relatively small sample of undocumented immigrants who can potentially benefit from DACA, we use a simpler strategy to estimate how the wage penalty responded to the executive action. In particular, we pool the sample of all immigrants (legal and undocumented) who satisfy the *demographic* requirements for DACA eligibility: the immigrant must have migrated to the United States before the age of 16, be at most 31 years old in 2012, and have at least a high school education. In the 2012 ACS, 30.1% of the workers in this sample were undocumented and would qualify for the benefits provided by DACA.

We estimate a regression in this sample of persons relating the worker's log hourly wage rate on a variable indicating if the worker was a legal immigrant, holding constant the set of demographic characteristics used throughout this section (i.e., age, sex, educational attainment, English language proficiency, state of residence, and country of birth). The coefficient of the legal status indicator, of course, measures the wage penalty. To isolate the impact of the DACA executive action on the wage penalty just before and after the 2012 announcement, we restrict the analysis to the 2010–2016 ACS cross-sections. We then interact the legal status indicator with variables indicating if the observation

documented persons does *not* imply that the flow of undocumented immigrants stopped altogether in 2006. Some of the undocumented persons present in the United States in 2006 may have left the country and many may have been able to adjust their immigration status and obtain a green card. These "exits" were then replaced by a similarly sized flow of new undocumented immigrants. We lack the requisite information to precisely measure how much of the decline in the wage penalty can be accounted for by changes in the sample composition of the relevant populations over the past decade.

[22] Note that although the pooling of male and female workers helps alleviate the small sample issue, it also introduces a problem. Nearly half of the undocumented women do not work so that wage trends in this sample are likely influenced by sample selection.

[23] Pope (2016) also uses the ACS to test the impact of DACA and finds that it increased the labor force participation and reduced the unemployment of eligible unauthorized immigrants, though only raised income for unauthorized immigrants in the bottom of the income distribution. Amuedo-Dorantes and Antman (2017) find that DACA reduced the probability of school attendance, consistent with a lack of legal work status leading to a substitution away from work and towards schooling.

Table 4
The impact of DACA on the wage penalty.

	Schooling > 12		Schooling = 12, not enrolled	
	Excludes enrolled (1)	Includes enrolled (2)	DACA eligible (3)	Not DACA eligible, but age <31 as of 2012 (4)
Legal status indicator	0.064	0.069	0.068	0.039
	(0.008)	(0.007)	(0.011)	(0.012)
Legal status indicator interacted with:				
2010–2011	−0.013	−0.011	−0.004	0.030
	(0.011)	(0.010)	(0.016)	(0.017)
2012–2013	–	–	–	–
2014	−0.023	−0.022	−0.016	0.022
	(0.012)	(0.011)	(0.017)	(0.018)
2015	−0.017	−0.023	−0.023	0.024
	(0.013)	(0.011)	(0.017)	(0.018)
2016	−0.038	−0.045	−0.045	0.028
	(0.012)	(0.011)	(0.017)	(0.017)
Includes school enrollment indicator	No	Yes	No	No
Number of observations	89,759	119,231	34,433	32,982

Notes: Standard errors are reported in parentheses. The sample in columns (1) and (2) consists of working immigrants who meet the demographic qualifications for DACA: aged 31 or less in 2012, have at least a high school education, and who migrated to the United States when they were 16 years old or younger. The sample in column (3) adds the further restriction that the immigrants have exactly 12 years of schooling. The sample in column (4) consists of workers who do not meet the demographic qualifications for DACA, but were 31 years old or younger in 2012. The regression includes vectors of fixed effects for age, gender, educational attainment, English language proficiency, state of residence, and birthplace.

is drawn from a particular cross-section, allowing us to document the trend in the wage penalty. Table 4 presents the relevant coefficients.

Before proceeding to discuss the coefficients, it is worth noting that it took a while for the DACA program to go into effect. Only 1687 applications had been approved by the end of the 2012 calendar year, and many more (472,378) were approved during the 2013 calendar year (U.S. Citizenship and Immigration Services, 2014). Much of the initial implementation of the program, therefore, took place over the 2012–2013 period, and we use this period as the baseline for our analysis.

The first column of Table 4 shows that the wage penalty in the demographic sample potentially affected by DACA stood at 6.4% during this baseline period. However, note that the wage penalty began to decline after 2014. By 2017, it had dropped by 3.8 percentage points.

The DACA executive action obviously encourages further education for the affected undocumented immigrants (as one needs at least a high school diploma to qualify for the benefits that DACA imparts).[24] Our empirical study of the wage penalty has been restricted to workers not enrolled in school. In the DACA context, however, this restriction might generate results that miss some of the potential impact of the executive action. The second column of the table replicates the analysis using the larger sample of DACA-eligible immigrants, which includes those who are enrolled in school (but report earnings). The regression suggests that the measured decline in the wage penalty in the post-DACA period is slightly larger, about 4.5 percentage points.

Note that the regression analysis reported in Table 3 is, in an important sense, "tracking" a particular cohort of immigrants (those who satisfy the demographic restrictions in DACA, whether legal or not) across ACS cross-sections. For example, the average age of a worker in our sample is 25.2 in 2010 and 28.5 in 2016. As a result, there may be life cycle effects on the wage penalty that contaminate the secular trend, and we might be mistakenly attributing any life cycle effects to DACA.

A simple way of showing that DACA does indeed seem to have an impact is to further refine the sample to workers not enrolled in school who have exactly 12 years of schooling, leading to a much more focused "tracking" of a particular set of workers.[25] In 2012, 64.5% of the sample of DACA-eligible undocumented workers had exactly 12 years of schooling. Column (3) of the table re-estimates the regression in this subsample of the DACA-eligible population and shows that the wage penalty in the baseline period 2012–2013 was 6.8% and had declined by 4.5 percentage points by 2016.

We can document that this decline in the wage penalty is not reflecting a life cycle effect by simply showing what happened to the trend in a comparable population that is *not* DACA-eligible. In particular, column 4 estimates the regression using the sample of immigrants who are not DACA-eligible, but were high school graduates and were at most 31 years old in 2012.[26] It is evident that the wage penalty in this comparable, but non-eligible, sample did not decline over time. If anything, the wage penalty was *rising* somewhat over the life cycle in this "counterfactual" sample (a trend consistent with the life cycle effects documented in the next section). In sum, the evidence in Table 4 suggests that the DACA executive action significantly improved the labor market conditions facing the affected undocumented workers and reduced the wage penalty by at least 4 or 5 percentage points.[27]

4. The wage penalty over the life cycle

The last section documented the differences in the wage penalty across different groups of undocumented workers, and the differential trends in the penalty experienced by the different groups. It turns out that the wage penalty will also vary for a given worker along the life cycle.

We begin our analysis of the life cycle variation in the wage penalty by illustrating the differences in the (cross-sectional) age-earnings profiles of natives, legal immigrants, and undocumented immigrants, shown in Fig. 4.[28] The age-earnings profiles of undocumented workers lie far below those of the other two groups *and* are relatively flat. At the age of

[24] Hsin and Ortega (2018) find that DACA, which is effectively a work permit program, serves to incentivize work over schooling, and the effect of DACA on university and community college attendance depends on how accommodating schools are of working students.

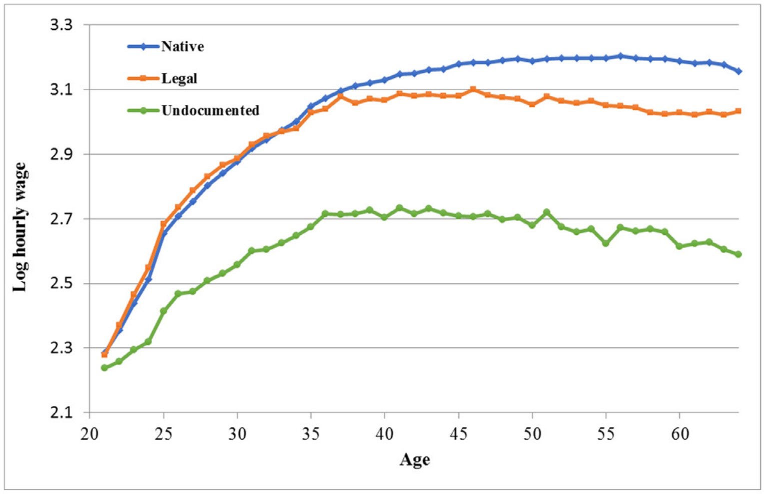
[25] The sample restriction avoids the sample composition problem created by the fact that some of the workers who do not appear in the early years of the sample because they enrolled in school eventually show up in the labor force in the later cross-sections as college graduates.

[26] By construction, the only difference between the two samples is that workers in the DACA-eligible sample migrated before age 16, while non-eligible undocumented workers migrated after age 16.

[27] Ortega et al. (2018) report that DACA recipients experienced a wage increase of around 12%, although they find no evidence that undocumented immigrants with a college degree experienced a wage increase.

[28] The analysis reported in this section pools the 2008–2016 cross-sections of the ACS.

A. Men



B. Women

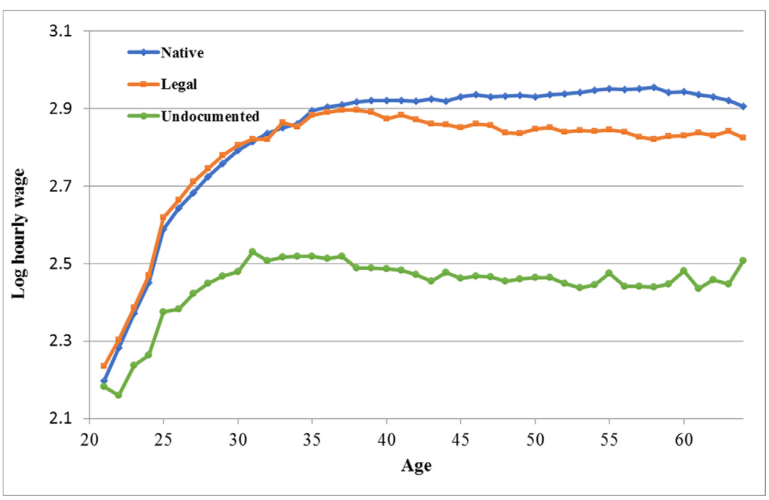


Fig. 4. Age-earnings profiles of workers.
Notes: The age-earnings profiles report the average log hourly wage of workers in each of the nativity groups at each age.

25, for example, the hourly wage of undocumented men in the ACS is 0.24 log points below that of natives and 0.27 log points below that of legal immigrants. By age 45, the wage gap between natives and undocumented immigrants rose to 0.47 log points, while the wage gap between legal and undocumented immigrants rose to 0.37 log points. The bottom panel of Fig. 3 shows similar life cycle effects for women.

It is important to emphasize that it is difficult to interpret the cross-section age-earnings profiles of both legal, and particularly, undocumented workers as measuring some type of wage evolution over the life cycle. It is well known (Borjas, 1985) that cross-section age-earnings profiles of immigrants are affected by both assimilation effects, the wage growth that occurs as a particular immigrant gets older, and by cohort effects, the differences in earnings potential across waves of immigrants that entered the United States at different times. The wage evolution of the undocumented sample is also affected by the fact that some of the undocumented will be able to “filter themselves” out and obtain green cards as they age, joining the legal sample, and by the fact that changes in the legal infrastructure regulating undocumented immigration (such as non-enforcement of existing laws or enactment of new penalties) might affect the flow of undocumented workers in and out of the country over time.

An important factor in understanding the evolution of earnings over the life cycle, particularly for undocumented versus legal immigrants, may be occupational attainment. A lack of legal immigration status likely acts as a barrier in the occupational mobility of undocumented immigrants as some occupations may be more difficult (or nearly impossible to attain) in the absence of legal status. To understand the importance of occupations in explaining the life cycle pattern of wages, we use a task-based approach to occupational attainment. Each occupation is assigned a vector of task requirements that summarize what is required to perform that job. The task requirements for each occupation are derived from the U.S. Department of Labor’s O*NET, with details of the procedure used to assign the task requirements discussed in Appendix A.

To simplify the presentation, we focus on only two tasks that efficiently summarize the difference in the types of jobs held by legal and undocumented immigrants: cognitive and non-cognitive tasks. An occupation that has a high level of cognitive task requirement might involve, for example, high levels of mathematical and deductive reasoning. In our data, the occupations with the highest cognitive task requirements are actuaries and physicists and astronomers. In contrast, occupations that require high levels of non-cognitive tasks typically involved physical strength and stamina, and the two occupations with the highest levels of non-cognitive task requirements are millwrights and dancers.

Fig. 5 shows the age-task requirement profiles (analogous to the age-earnings profiles in Fig. 4) for our cognitive and non-cognitive occupational task requirement measures. These figures mirror the age-earnings profiles. The cognitive task, which is strongly and positively associated

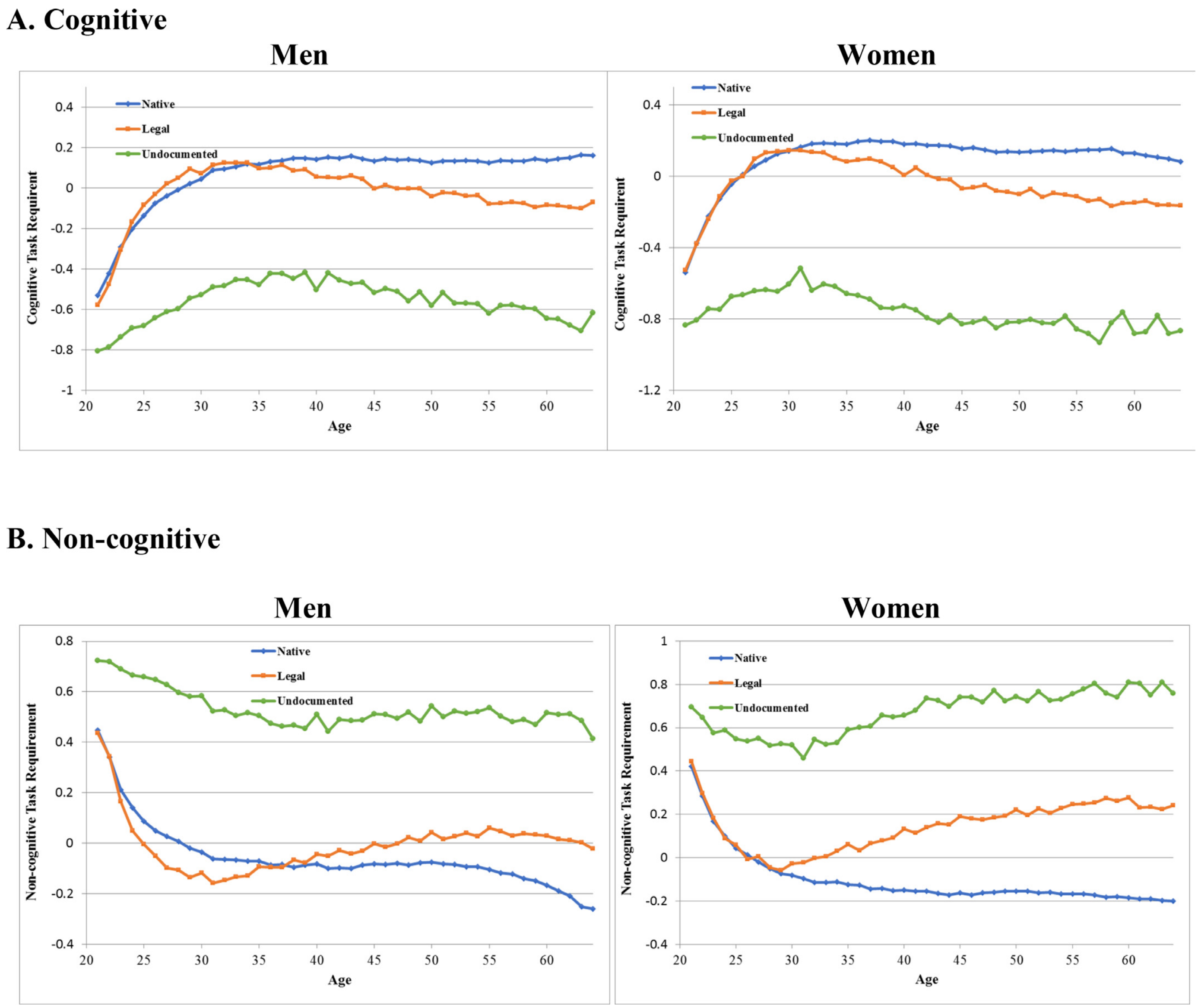


Fig. 5. Age-task requirement profiles of workers.
Notes: The age-task requirement profiles report the average cognitive and non-cognitive task requirements of workers in each of the nativity groups at each age.

with wages, starts lower for undocumented immigrants than for natives or legal immigrants, rises more slowly with age, and actually begins to decline quite early in the life cycle. In contrast, the non-cognitive task shows the opposite pattern, falling more slowly for undocumented immigrants than the other groups, flattening out for men after about age 35, and actually starting to rise early in the life cycle for women. Note that the divergence between legal and undocumented immigrants in the occupational task requirements occurs between the ages of 21 and 35. More generally, Fig. 5 demonstrates the striking difference in the jobs the two groups perform, how this difference widens prior to age 35, and how the substantial gap then persists over the lifecycle.[29]

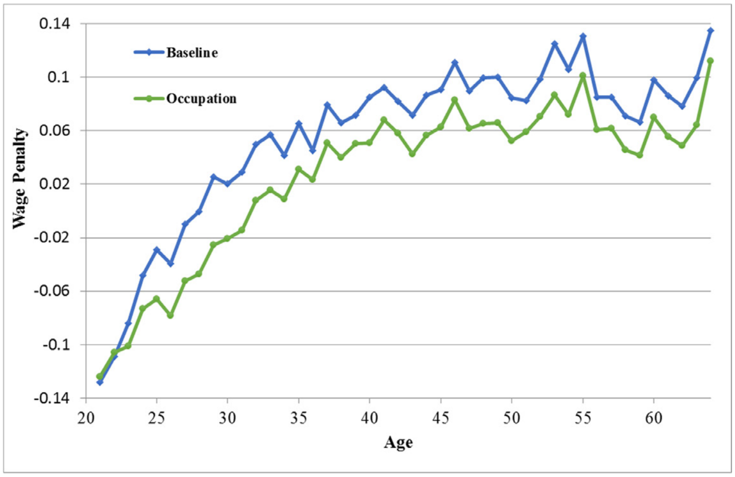
The "raw" age-earnings profiles illustrated in Fig. 4 do not adjust for differences in other worker characteristics such as educational attainment and English language proficiency, but they do suggest that the wage penalty to undocumented immigration is not constant over the life cycle, while Fig. 5 suggests that differences in occupational mobility (particularly at younger ages) may be an important factor in explaining both the overall wage penalty as well as in understanding the evolution of the wage penalty over the life cycle. To study the variation in the wage penalty over the life cycle, we estimate a Mincerian log wage regression that allows us to measure the difference in the slope of the age-earnings profile between legal and undocumented immigrants. In particular, consider the following regression model estimated in the pooled 2006–2016 ACS sample (separately for men and women):

$$\log\ w_{it} = \beta h_i + \theta_t + A_i + \pi_A\left(L_i \times A_i\right) + \varepsilon_{it}, \tag{3}$$

where w_{it} gives the wage of worker i in year t; h_i is a vector of the worker's socioeconomic characteristics (i.e., education, years since migration, English proficiency, state of residence, and country of birth); θt is a vector of calendar year fixed effects; At is a vector of age fixed effects, with each value of age having its own fixed effect; and these age fixed effects are interacted with L_i, a variable indicating if worker i is a legal immigrant. The coefficient vector πA measures the wage penalty at a particular age.

[29] The most common occupation among low-skilled men is truck driver for legal immigrants but construction laborer for undocumented immigrants, which is consistent with truck drivers often requiring an occupational license and these licenses being more difficult to undocumented immigrants to acquire. See Cassidy and Dacass (2019) for a more thorough discussion of occupational licensing and immigrants.

A. Men



B. Women

Fig. 6. Wage penalty for undocumented workers over the life-cycle.
Notes: Figures show the wage penalty between legal and undocumented immigrants in log hourly wage at different points in the life cycle calculated with a Mincerian wage regression that includes controls for survey year, educational attainment, state of residence, years-since-migration, birthplace, and English language proficiency. The wage penalty values shown are the coefficients on legal status interacted with age. The line "occupation" adds occupation controls to the baseline specification. All results calculated from the ACS.

Fig. 6 illustrates the "baseline" life cycle trend in the measured wage penalty. Consistent with Figs. 4 and 5, we find that the wage penalty increases steadily over the life cycle until approximately ages 45–50, when it plateaus for men and begins to decline slightly for women. Note that the measured wage penalty is negative for the youngest undocumented workers. Given the unadjusted age-earnings profiles illustrated in Fig. 4, the finding of a negative wage penalty at younger ages is not surprising. After all, the average wage of undocumented immigrants is basically equal to that of legal immigrants for workers in their 20s. At the same time, however, the undocumented population has far less education and is much less English proficient, generating a negative wage penalty. The relatively superior economic performance of young undocumented workers seems like an empirical finding that deserves much further study.

We explore the role of occupational mobility in generating the life cycle trend in the wage penalty by adding a vector of occupation fixed effects to the log wage regression in Eq. (3). Fig. 6 shows that the introduction of the occupation fixed effects noticeably reduces the growth in the wage penalty between ages 21 and 35, particularly for men. For example, the baseline wage penalty for men grows by 19.4 percentage points (from −12.9–+6.5%) through age 35, while the wage penalty that adjusts for the widening gap in occupational attainment grows by only 15.5 percentage points (from −12.4–+3.1). The evidence, therefore, suggests that occupational mobility (or, more specifically, the lack thereof) is a determinant of the life cycle trend in the wage penalty.

5. The wage penalty across states

The analysis reported in the previous sections suggests that the *average* wage penalty to undocumented immigration was quantitatively small for both undocumented men and women by 2016. As we have seen, however, this conclusion does not necessarily imply that there was little wage penalty throughout the U.S. labor market. We have documented important differences in the wage penalty as a worker ages, between new immigrants and older immigrants, and over time as relaxed restrictions on undocumented immigration affected some groups of workers. This section continues the analysis of the dispersion in the wage penalty by exploiting the fact that the relative number of undocumented immigrants varies substantially across states. According to the official DHS statistics (Baker, 2017), 56% of undocumented immigrants in January 2104 lived in only 5 states (California with 24%; Texas with 16%; Florida with 6%; and New York and Illinois, each with 5%).

Further, the labor market environment facing undocumented immigrants in the past decade changed differently across states, due perhaps to geographic differences in the impact of the Great Recession (and subsequent recovery) or to state-specific legislation that made it more difficult for undocumented immigrants to work in particular regions (more on this below). These differences may account for some of the observed changes in the relative number of undocumented immigrants choosing to settle in some states over time. For example, the official DHS statistics (Baker, 2017) reveal a sizable decline between 2007 and 2014 in the number of undocumented immigrants in Arizona (from 530,000 to

Table 5
Interstate variation in undocumented immigration and the wage penalty for men.

	Number of undocumented workers (1000s)		Undocumented share of workforce (%)		Wage penalty	
State	2008	2016	2008	2016	2008	2016
Arizona	251.0	140.1	8.6	4.8	−0.039	0.011
California	1940.1	1491.9	11.3	8.3	0.080	0.061
Colorado	138.6	124.6	5.4	4.5	0.120	−0.012
Florida	558.3	519.2	6.7	5.8	0.035	0.039
Georgia	280.5	260.1	6.1	5.6	0.072	0.053
Illinois	400.8	332.3	6.5	5.5	0.045	0.049
Maryland	150.0	165.9	5.3	5.6	0.045	0.062
Massachusetts	140.2	103.2	4.3	3.1	−0.036	−0.020
Nevada	139.6	119.7	11.0	8.9	−0.032	0.010
New Jersey	339.5	324.5	8.0	7.7	0.079	0.086
New York	676.9	562.0	7.3	6.1	0.051	0.058
North Carolina	215.0	202.9	4.9	4.5	0.066	0.048
Pennsylvania	102.7	103.7	1.7	1.8	−0.002	−0.001
Texas	997.3	1084.1	8.8	8.6	0.015	0.065
Virginia	173.3	182.5	4.4	4.5	0.086	0.066
Washington	149.0	164.9	4.6	4.8	−0.022	0.070

Notes: The 16 states listed in this table had the largest number of undocumented workers (at least 100,000) in 2008. The undocumented share is the fraction of undocumented workers in the state's workforce. The wage penalty values are for male immigrants.

270,000), a stable undocumented population in New York (at 640,000), and a slight increase in the number of undocumented persons in North Carolina (from 380,000 to 400,000).

The potential interstate differences in the labor market conditions facing undocumented immigrants suggest a novel use of the ACS data. Specifically, we can estimate the wage penalty to undocumented immigration in each state/year cell and then determine whether this variation responds to factors that describe the local labor market, including the relative number of undocumented immigrants, aggregate economic conditions in the state, and state-specific legislative changes that made it more difficult for employers to hire undocumented workers.[30]

To calculate the wage penalty for each state-year cell, we again estimate a Mincerian earnings function using the pooled ACS data over the entire sample period 2008–2016:

$$\log\ w_{ist} = \beta h_i + \theta_{rt} + \pi_{rt}\left(L_i \times \theta_{rt}\right) + \varepsilon_{irt}, \tag{4}$$

where w_{ist} gives the wage of worker i residing in state r in year t; h_i is a vector of the worker's socioeconomic characteristics (which now also includes a vector of age fixed effects in 5-year bands); θrt is a vector of state-year interaction fixed effects; and these fixed effects are interacted with L_i, the variable indicating if worker i is a legal immigrant. The coefficient vector πrt measures the wage penalty in state r at time t.

We first illustrate the sizable interstate variation in both the number of undocumented workers and in the measured wage penalty for male immigrants in the 2008 and 2016 ACS cross-sections. Table 5 reports the number of undocumented workers (aged 21–64) in the 16 states that employed at least 100,000 undocumented workers in 2008. The table also reports the share of undocumented workers as a fraction of the state's total workforce.

It is evident that the number of undocumented workers fell significantly in some states, while rising in others. For example, the number of undocumented workers fell by over 40% in Arizona (from 251.0 thousand to 140.1 thousand), while rising by nearly 10% in Texas (from 997.3 thousand to 1084.1 thousand).

The table also shows sizable differences in both the level and the trends in the undocumented share. The fraction of the state's workforce composed of undocumented workers fell from 8.6 to 4.8% in Arizona and from 11.3 to 8.3% in California. In contrast, it declined slightly in Texas from 8.8 to 8.6 % and rose slightly from 4.6 to 4.8 in Washington. Of the 16 states listed, only Washington, Pennsylvania, and Maryland experienced an increase in the *share* of their workforce that is undocumented, and those increases were modest.

Although the average wage penalty in the national labor market hovered between 4 and 6% throughout much of the period, there was much greater interstate variation in the penalty. Table 5 also reports the estimated wage penalty for men for each of the states in 2008 and 2016. The wage penalty rose by 5 percentage points in Arizona (from −3.9% to 1.1%), by 0.7 percentage point in New Jersey (from 7.9 to 8.6%), and fell by 1.9 percentage points in California (from 8.0 to 6.1%). Fig. 7 illustrates the dispersion in the size of the wage penalty for men across the 16 states with the largest number of undocumented workers. Note that most of the penalties estimated for each state-year cell are positive, i.e., legal immigrants have higher wages that otherwise similar undocumented immigrants.

We exploit this variation to determine if there are systematic factors that explain the differences in the wage penalty that undocumented immigrants face in different geographic labor markets at different times. Specifically, we estimate second-stage regressions that relate the wage penalty in a state-year cell to variables that describe local labor market conditions facing the undocumented.[31] We consider three specific variables that might determine the size of the wage penalty: (1) the relative number of undocumented immigrants in the local labor market; (2) the presence of state-level legislation that restricts the employment of undocumented immigrants; and (3) the impact of the Great Recession on local labor market conditions. This second-stage regression is estimated using the entire sample of 459 state-year observations (9 annual observations for each of the 51 "states," which includes the District of Columbia). The regression also includes vectors of state fixed effects and year fixed effects. The regression is weighted by the number of observations used to calculate the dependent variable (i.e., the wage penalty in the state-year cell), and the standard errors are clustered at the state level.[32]

The top panel of Table 6 presents the relevant OLS coefficients of the regression model. Consider initially the regressions estimated using the sample of male workers. As column 1 shows, there is a positive

[30] Related work by Massey and Gentsch (2014), using Mexican Migration Project data and state-year undocumented population estimates from Warren and Warren (2013), find that the percent of a state in a given year is negatively related to the wage of undocumented Mexican immigrants.

[31] Note that we are using the state as the geographic definition of the local labor market. It might be preferable to look at smaller geographic units, such as metropolitan areas or commuting zones, but the sample size of undocumented immigrants would fall substantially in many of these smaller geographic units.

[32] More precisely, the weight is given by $(nL \times nU)/(nL + nU)$, where nL and nU give the number of observations in the state-year cell for legal immigrants and undocumented workers, respectively.

Fig. 7. Distribution of wage penalty for men in states with largest number of undocumented workers, 2008–2016.
Notes: The wage penalty is calculated at the state-year cell and is regression coefficient of state-year interacted with legal immigrant status in a Mincerian wage regression that includes controls for survey year, age, educational attainment, state of residence, years-since-migration, birthplace, and English language proficiency.

Table 6
Determinants of variation in wage penalty across states, 2008–2016.

	All men (1)	All men (2)	Low-skill men (3)	High-skill Men (4)	Women (5)
OLS estimates					
Undocumented share	0.009	0.011	0.012	0.017	0.015
	(0.003)	(0.004)	(0.005)	(0.006)	(0.004)
E-Verify	–	0.045	0.038	0.048	−0.015
		(0.014)	(0.027)	(0.042)	(0.011)
Unemployment rate	–	−0.003	−0.002	−0.001	−0.002
		(0.005)	(0.005)	(0.007)	(0.005)
IV estimates					
Undocumented share	0.009	0.011	0.013	0.019	−0.002
	(0.003)	(0.006)	(0.006)	(0.009)	(0.011)
E-Verify	–	0.045	0.037	0.047	−0.003
		(0.013)	(0.025)	(0.039)	(0.014)
Unemployment rate	–	−0.003	−0.003	−0.002	0.004
		(0.005)	(0.005)	(0.008)	(0.007)

Notes: Standard errors in parentheses clustered at the state level. The undocumented share gives the percent of the state's total workforce that is composed of undocumented workers. The E-Verify variable is set to unity if employers in a given state/year cell were required to use E-Verify in their new hiring. The "low-skill" regressions are estimated using the sample of immigrants who have at most a high school education; and the "high skill" regressions are estimated in the sample of immigrants who have more than a high school education. Columns (1), (2), and (4), have 459 observations, column (3) has 456 observations, and column (5) has 458 observations.

and statistically significant relationship between the wage penalty and the fraction of the state's workforce that is undocumented. Moreover, the impact is quantitatively sizable: An increase of 1 percentage point in the undocumented share raises the wages penalty by about 0.9 percentage points.[33]

It is worth noting that the positive relationship between the wage penalty and the relative size of the undocumented workforce suggests that legal and undocumented immigrants are not perfect substitutes in production. If the two groups were perfect substitutes, the relative wage of undocumented immigrants would not depend on their relative number. In fact, if we assume that the wage data was generated in labor markets where profit-maximizing competitive firms faced the technology implied by a nested CES aggregate production function, we can use our data to estimate the elasticity of substitution between the two groups. It is well known that the elasticity of substitution is given by the reciprocal of the coefficient of a regression of the log wage ratio between any two labor inputs on the log quantity ratio of those two inputs. The estimated regression for men is:

$$\pi_{rt} = \theta_r + \theta_t - 0.059 \log \frac{E_{Lrt}}{E_{Urt}}, \tag{5}$$

where θr and θt denote vectors of state and time fixed effects; and *Ejrt* gives the number of workers of type *j* in state *r* at time *t*. The implied elasticity of substitution between the two groups is 17.0. For women, the coefficient on the log ratio of legal to undocumented immigrants is −0.087 (with a standard error of 0.029), which yields an implied elasticity of 11.5. Note that the coefficient is statistically significant for both men and women, so that we can reject the hypothesis that legal and undocumented immigrants are perfect substitutes. In short, there are factors—perhaps due to unobservable differences in the skill set of the two groups, or because the two groups face different labor market constraints—which imply that the labor market does not view the two groups as interchangeable. As a result, an important insight from this type of empirical analysis is that the relative size of the undocumented

[33] Massey and Gentsch (2014) also find that a one percentage point increase in the share of a state that is undocumented lowers the wages of undocumented Mexican immigrants by one percentage point.

population is itself a major determinant of the wage penalty. A larger undocumented population generates a substantially larger wage penalty.

Of course, the OLS correlation between the wage penalty and the undocumented share is contaminated by the potentially endogenous settlement of both undocumented and legal immigrants in some states and not in others. We address the endogeneity problem by using a variant of the generic shift-share instrument that has become popular in the immigration literature (although see Jaeger et al., 2018, for a critical appraisal). Specifically, we use the pooled Current Population Surveys between 1995 and 2000 to obtain the baseline interstate distribution of both legal and undocumented immigrants from a particular country of origin, where the CPS surveys use the residual method described in section II of the paper to impute an undocumented immigration status variable for each observation.[34]

Let *scr* be the share of the undocumented population from country *c* living in state *r* in the pooled CPS. Suppose that the ACS survey in cross-section year *t* reveals the presence of $Uc(t)$ undocumented immigrant workers from that country. We then predict the number of undocumented immigrants from country *c* residing in state *r* in year *t* to be the product $(scr \times Uc(t))$. We conduct a similar calculation for legal immigrants. By adding up this predicted number across all countries of birth, we can then obtain the number of undocumented and legal immigrants that would be predicted to be in state *r* in cross-section *t*, or $\hat{U}_r(t)$ and $\hat{L}_r(t)$. The instrument for the undocumented immigrant share variable is then given by:

$$S = \frac{\hat{U}_r(t)}{\hat{U}_r(t) + \hat{L}_r(t) + N_r(t)}, \tag{4}$$

where $Nr(t)$ gives the size of the native workforce in state *r* in year *t*.

The first-stage regression shows that there is indeed there is a significant positive correlation between the actual undocumented immigrant share in the state and the predicted share. The relevant coefficient of the first stage is 0.916, with a standard error of 0.188. The bottom panel of Table 6 shows that the IV coefficient of the share is again positive and statistically significant. In fact, the magnitude of the coefficient is nearly identical to that obtained using OLS. A one percentage point increase in the undocumented share again increases the size of the wage penalty by about 0.9 percentage points.

The second column of the table adds two additional regressors to the model that attempt to explain the interstate variation in the wage penalty—a variable that measures state-specific legislative measures to restrict undocumented employment and the state's unemployment rate (as reported by the BLS).

During the period under analysis, the federal inaction on resolving the status of undocumented workers led some states to take state-specific actions that made it more difficult for employers to employ the undocumented. The best known of these attempts was the 2010 legislation in Arizona that, among many things, "[required] law enforcement officers to determine immigration status during any lawful stop; [created] state crimes and penalties for failure to carry federally-issued alien registration documents; [made] it unlawful for an unauthorized alien to knowingly apply for or perform work in Arizona; and [permitted] an officer to make a warrantless arrest if the officer has probable cause to believe the person has committed any public offense that makes the person removable from the United States" (National Conference of State Legislatures, 2012a).[35]

One common provision in the restrictive state-level statutes was the requirement that employers use E-Verify to authenticate the legal status of new hires. As the Department of Homeland Security describes it: "E-Verify is a web-based system that allows enrolled employers to confirm the eligibility of their employees to work in the United States. E-Verify employers verify the identity and employment eligibility of newly hired employees by electronically matching information provided by employees on the Form I-9, Employment Eligibility Verification, against records available to the Social Security Administration (SSA) and the Department of Homeland Security (DHS)."[36] During the period under analysis, four states enacted legislation mandating that *all* private employers in those states use the E-Verify system to confirm the employment eligibility of new hires: Arizona beginning in 2008, Alabama in 2012, Mississippi in 2011, and South Carolina in 2010.[37]

We introduced a variable into our regression model indicating if an E-Verify provision was in effect in a particular state at time *t*. Table 6 shows that the legislation mandating the use of E-Verify had a significant positive impact on the wage penalty to undocumented immigration, raising the wage penalty by 4.5 percentage points (in both the OLS and IV regressions). Not surprisingly, legal restrictions that make it harder for employers to hire undocumented immigrants (and make legal and undocumented immigrants less substitutable) increases the wage penalty. It is important to note, however, that our evidence exploits the enactment of the E-Verify system in only a very small number of states, so that it would not be prudent to generalize from this exercise to a prediction of what would happen to the wage penalty if the system were adopted nationwide.

Our results showing an impact of the E-Verify program on the wage penalty are generally consistent with related evidence reported in other studies, although these studies typically examine the link between E-Verify and the economic outcomes of Hispanic (mainly Mexican) immigrants (Bohn et al., 2015; Orrenius and Zavodny, 2015, 2016; and Orrenius et al., 2018). In rough terms, these studies find that the adoption of the E-Verify system reduces the hourly wage of undocumented Mexican men; leads to an outflow of undocumented immigrants from the states that adopt the system; and reduces the employment propensities of Hispanic workers.[38]

The second column of the table also includes the BLS unemployment rate in the state-year cell. Surprisingly, the impact of the local unemployment rate on the wage penalty is near zero, both statistically and numerically. It seems that changes in local labor market conditions arising from the business cycle affect aggregate wages, and likely affect immigrant wages, but do not seem to affect the relative wage of undocumented workers.

Column 5 of the table replicates the analysis using the sample of female workers. The results are far less stable, although they still show a significant positive relation between the wage penalty and the relative size of the undocumented population (but only in the OLS specification). As noted earlier, the analysis of wage trends in the female undocumented sample is problematic, as undocumented women have a very low labor force participation rate, so that selection biases likely play a significant role in determining both the interstate variation in the wage penalty and the trend in that penalty within a particular state. Borjas (2017, Table 1) reports that the employment rate in the 2012–2013 CPS

[34] Ideally, the instrument would employ the geographic settlement of legal and undocumented immigrants many years prior to the sample period of 2008–2016. The longer lag would reduce the probability that serial correlation in economic conditions at the state level invalidate the instrument. Unfortunately, there are no large-scale micro surveys that can be used to impute the legal status of a foreign-born person prior to the 1994 CPS.

[35] Some of the provisions in this legislation were later ruled to be unconstitutional by the U.S. Supreme Court.

[36] U.S. Citizenship and Immigration Services (2018b).

[37] NumbersUSA (2016); see also National Council of State Legislatures (2012b). A few other states passed versions of the E-Verify requirement, but these states typically exempted many employers (such as small firms) and had a phase-in period before the requirement was fully operational.

[38] We defined the states that use an E-Verify system using a strict definition of the regulation, specifically isolating the states and time periods where E-Verify applied to all workers. The Orrenius–Zavodny–Gutierrez studies use a less strict definition, classifying some states as having an E-Verify system even though many employers are exempt from the legislation or there is a phasing-in period when the regulation was not enforced.

is 84.7% for legal immigrant men and 88.1% for undocumented men. In contrast, the employment rate is 64.4% for legal immigrant women and only 56.7% for undocumented women. The selection biases in wage regressions are likely to be substantial when nearly half the sample self-selects out of the workforce.

Finally, Table 6 also reports regressions estimated separately for low-skill (defined as persons with at most a high school education) and high-skill (those with more than a high school education) male immigrants. These regressions tend to reinforce the finding that the wage penalty is larger in states that have a relatively larger undocumented population, and that restrictions on the employment of undocumented immigrants tend to raise the wage penalty.

6. Summary

The past decade has witnessed a series of attempts to create some type of "path to citizenship" for the over 12 million undocumented immigrants now residing in the United States. This paper uses newly developed algorithms that impute undocumented status for each person in microdata files, including the Current Population Surveys and the American Community Surveys, to examine the determinants of what is perhaps the key indicator of their economic well-being, their earnings in the U.S. labor market.

The analysis yields a number of new insights into the determination of earnings for the large undocumented population:

(1) The age-earnings profiles of undocumented workers lies far below that of legal immigrants and of native workers. Moreover, the age-earnings profile of undocumented workers is almost perfectly flat during the prime working years.

(2) The unadjusted gap in the log hourly wage between legal immigrants and undocumented workers is large (around 35% for both men and women). Much of this gap disappears once the calculation adjusts for differences in observable socioeconomic characteristics, including age, education, state of residence, country of birth, and English language proficiency. As a result, the wage penalty—the wage disadvantage suffered by undocumented workers relative to statistically comparable legal immigrants—hovered around 6% until about 2013 for men and 4% for women, at which point it began a noticeable decline. Between 2013 and 2016, the wage penalty to both male and female undocumented workers had shrunk to about only 2–4%.

(3) There are important differences in the level and trend in the wage penalty over the life cycle and across labor markets. The wage penalty rises over the life cycle partly because undocumented immigrants do not experience the same extent of occupational mobility as legal immigrants; and the wage penalty is larger in states with a larger undocumented population and in states that have enacted state-specific legislation that restricts the employment of undocumented immigrants.

It is important to emphasize that the analysis reported in this paper represents a first step in any evaluation of the proposals being discussed to regularize the status of undocumented workers. Much more information about the economic well-being of the undocumented population needs to be documented and examined before a full evaluation can be made. Similarly, it is crucial to continue to assess the validity of the statistical methods that are used to impute a person's undocumented status in microdata surveys.

Appendix A. Deriving occupational task requirements

In this appendix, we briefly describe the procedure used to assign cognitive and non-cognitive task requirements to each *occ1990* occupation code in the ACS.

Table A1
Comparison of summary statistics for workers, 2012–2013, Women.

		Legal		Undocumented	
	Natives	No correction	H1B correction.	No correction.	H1B correction
A. Pew					
Percent of pop.	84.6	11.4	11.5	4.0	4.0
Average age	42.5	43.5	43.5	39.1	39.2
Education:					
High school dropouts	3.8	14.2	14.2	38.2	38.6
High school graduates	25.4	24.2	24.1	28.0	28.4
Some college	32.3	20.7	20.7	14.4	14.6
College graduates	25.1	26.0	26.1	12.4	12.1
Postcollege	13.3	14.9	15.1	7.0	6.3
State of residence:					
California	8.9	26.2	26.2	21.1	21.0
New York	5.5	11.6	11.6	7.7	7.7
Texas	7.7	8.1	8.1	14.1	14.2
Log wage gap	0.000	−0.048	−0.045	−0.391	−0.403
Sample size	64,173	11,864	11,924	4553	4493
B. ACS					
Percent of pop.	84.6	11.4	11.6	3.9	3.8
Average age	42.7	43.9	43.8	38.3	38.5
Education:					
High school dropouts	3.7	15.1	14.9	36.2	37.2
High school graduates	25.8	24.3	24.1	29.5	30.4
Some college	34.5	23.4	23.2	14.2	14.6
College graduates	23.2	23.2	23.3	12.0	11.1
Postcollege	12.7	14.1	14.5	8.0	6.7
Speaks English	–	60.2	60.3	32.9	31.7
State of residence:					
California	8.7	25.5	25.4	24.0	24.1
New York	5.7	12.5	12.5	8.7	8.8
Texas	7.6	8.5	8.5	12.0	12.1
Log wage gap	0.000	−0.024	−0.019	−0.381	−0.404
Sample size	933,459	114,975	115,854	31,398	30,519

Notes: The statistics are calculated in the sample of women aged 21–64 who are not enrolled in school, are not self-employed, and report positive wage and salary income, weeks worked, and usual hours worked weekly.

Each occupation in the O*NET (version 17.0) contains a vector of job characteristics, e.g., number facility. Using the ACS variable *occsoc*, we merge these O*NET job characteristics with the ACS for years 2010–2016, which are the years in the ACS where the *occsoc* code most closely matches the SOC codes used in the O*NET. Each individual in the ACS with a valid *occsoc* code between ages 21–64 is assigned a vector of O*NET job characteristics. Since we will use the *occ1990* occupation code to merge the task requirements with the other ACS samples, we average across the O*NET job characteristics within each *occ1990* occupation code.

We follow the approach of Yamaguchi (2012) and Imai et al. (2018) and perform an *a priori* grouping of some of these job characteristics into our two groups: cognitive and non-cognitive. The cognitive characteristics follow the analytical category used in Imai et al. (2018), and contains: Deductive Reasoning, Inductive Reasoning, Information Ordering, Category Flexibility, Mathematical Reasoning, Number Facility, Analytical Thinking, Making Decisions and Solving Problems, and Mathematics. The non-cognitive characteristics follow the physical strength category from Imai et al. (2018), and include: Static Strength, Dynamic Strength, Trunk Strength, Stamina, Performing General Physical Activities, and Handling Moving Objects.

To reduce the job characteristics in the cognitive and non-cognitive groups to a single task measure each, we perform principal component analysis separately for each group of characteristics and extract the first component, which yields our cognitive and non-cognitive task requirements. We then rescaled each task requirement to have a mean of zero and a standard deviation of one in the 2010–2016 ACS sample. This entire procedure is performed separately for men and women, whose task measures may differ depending on how the *occsoc* codes map into the *occ1990* codes.

At the end of this procedure, we have, for each *occ1990* code (and separately for men and women), a two-dimensional vector of cognitive and non-cognitive occupational task requirements. We then merge these task requirements for all of our ACS sample years (2008–2016).

References

Amuedo-Dorantes, C., Antman, F., 2017. Schooling and labor market effects of temporary authorization: evidence from daca. J. Popul. Econ. 31 (1), 339–373.
Albert, C., 2019. The Labor Market Impact of Immigration: Job Creation vs. Job Competition. CEMFI Working Paper.
Baker, B., 2017. Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2014. Department of Homeland Security, Office of Immigration Statistics, Washington, DC.
Bohn, S., Lofstrom, M., Raphael, S., 2015. Do E-Verify mandates improve labor market outcomes of low-skilled native and legal immigrant workers? South. Econ. J. 81 (4), 960–979.
Borjas, G.J., 1985. Assimilation, changes in cohort quality, and the earnings of immigrants. J. Labor Econ. 3 (4), 463–489.
Borjas, G.J., 2016. The Earnings of Undocumented Immigrants. NBER Working Paper No. 23236.
Borjas, G.J., 2017. The labour supply of undocumented immigrants. Labour Econ. 46, 1–13.
Cassidy, H., 2017. Kansas State University Working Paper.
Cassidy, H., Dacass, T., 2019. Kansas State University Working Paper.
Cotton, J., 1988. On the decomposition of wage differentials. Rev. Econ. Stat. 70 (2), 236–243.
Gelbach, J.B., 2016. When do covariates matter? And which ones, and how much? J. Labor Econ. 34 (2), 509–543.
Hotchkiss, J.L., Quispe-Agnoli, M., 2013. The expected impact of state immigration legislation on labor market outcomes. J. Policy Anal. Manag. 32 (1), 34–59.
Hsin, A., Ortega, F., 2018. The effects of deferred action for childhood arrivals on the educational outcomes of undocumented students. Demography 55 (4), 1487–1506.
Imai, S., D. Stacey, and C. Warman (2018): “From engineer to taxi driver? Language proficiency and the occupational skills of immigrants,” Working Paper.
Jaeger, D.A., Ruist, J., Stuhler, J., 2018. Shift-Share Instruments and the Impact of Immigration. National Bureau of Economic Research NBER Working Paper No. 24285.
Kossoudji, S.A., Cobb-Clark, D.A., 2002. Coming out of the shadows: learning about legal status and wages from the legalized population. J. Labor Econ. 20 (3), 598–628.
Kaushal, N., 2006. Amnesty programs and the labor market outcomes of undocumented workers. J. Hum. Resour. 16 (3), 631–647.
Massey, D.D., Gentsch, K., 2014. Undocumented migration to the United States and the wages of Mexican immigrants. Int. Migr. Rev. 48 (2), 482–499.
National Conference of State Legislatures. 2012a. “State omnibus immigration legislation and legal challenges,” August 27. Website: http://www.ncsl.org/research/immigration/omnibus-immigration-legislation.aspx.
National Conference of State Legislatures. 2012b. “E-Verify,” August 27. Website: http://www.ncsl.org/research/immigration/everify-faq.aspx.
NumbersUSA. 2016. “Map of states with E-verify laws,” December 18. Website: https://www.numbersusa.com/resource-article/everify-state-map.
Orrenius, P.M., Zavodny, M., 2015. The impact of E-Verify mandates on labor market outcomes. South. Econ. J. 81 (4), 947–959.
Orrenius, P.M., Zavodny, M., 2016. Do state work eligibility verification laws reduce unauthorized immigration? IZA J. Migr. 5 (5), 1–17.
Orrenius, P.M., Zavodny, M., Gutierrez, E., 2018. Do state employment eligibility verification laws affect job turnover? Contemp. Econ. Policy 36 (2), 394–409.
Ortega, F., Hsin, A., 2018. IZA Discussion Paper No. 11680.
Ortega, F., Edwards, R., Hsin, A., 2018. IZA Discussion Paper No. 11281.
Passel, J.S., Cohn, D’V., 2014. Unauthorized Immigrant Totals Rise in 7 States, Fall in 14 States: Decline in Those From Mexico Fuels Most State Decreases. Pew Research Center, Washington, DC.
Pope, N.G., 2016. The effects of DACAmentation: the impact of deferred action for childhood arrivals on unauthorized immigrants. J. Public Econ. 143, 98–114.
Rivera-Batiz, F.L., 1999. Undocumented workers in the labor market: an analysis of the earnings of legal and illegal Mexican immigrants in the United States. J. Popul. Econ. 12 (1), 91–116.
Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., Sobek, M..
U.S. Citizenship and Immigration Services. 2014. “Deferred action for childhood arrivals (DACA) report (Aug. 15, 2012–Dec. 31, 2013),” Washington, DC. https://www.uscis.gov/sites/default/files/USCIS/Resources/ReportsandStudies/ImmigrationFormsData/AllFormTypes/DACA/DACA-06-02-14.pdf.
U.S. Citizenship and Immigration Services. 2018a. “Characteristics of H-1B specialty occupation workers,” Washington, DC. https://www.uscis.gov/sites/default/files/reports-studies/Characteristics-of-Specialty-Occupation-Workers-H-1B-Fiscal-Year-2017.pdf.
U.S. Citizenship and Immigration Services. 2018b. “About E-Verify,” Washington, DC. https://www.uscis.gov/e-verify.
Van Hook, J., James, D, Donna Coffman, B., Harel, O., 2015. Can we spin straw into gold? An evaluation of immigrant legal status imputation approaches. Demography 52 (1), 329–354.
Warren, R., 2014. Democratizing data about unauthorized residents in the United States: estimates and public-use data, 2010 to 2013. J. Migr. Hum. Secur. 2 (4), 305–328.
Warren, R., Warren, J.R., 2013. Unauthorized immigration to the United States: annual estimates and components of change, by state, 1990 to 2010. Int. Migr. Rev. 47 (2), 296–329.
Warren, R., Passel, J.S., 1987. A count of the uncountable: estimates of undocumented aliens counted in the 1980 United States census. Demography 24 (3), 375–393.
Yamaguchi, S., 2012. Tasks and heterogeneous human capital. J. Labor Econ. 30 (1), 1–53.