

Ortega, Francesc

Working Paper

Occupational Barriers and the Productivity Penalty from Lack of Legal Status

CReAM Discussion Paper Series, No. 18/21

Provided in Cooperation with:

Rockwool Foundation Berlin (RF Berlin)

Suggested Citation: Ortega, Francesc (2021) : Occupational Barriers and the Productivity Penalty from Lack of Legal Status, CReAM Discussion Paper Series, No. 18/21, Centre for Research & Analysis of Migration (CReAM), Department of Economics, University College London, London

This Version is available at:

<https://hdl.handle.net/10419/295651>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



CREAM

**Centre for Research &
Analysis of Migration**

Discussion Paper Series

CDP 18/21

- ▶ **Occupational Barriers and the Productivity Penalty from Lack of Legal Status**
- ▶ **Francesc Ortega and Amy Hsin**

Centre for Research and Analysis of Migration
Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

www.cream-migration.org

Occupational Barriers and the Productivity Penalty from Lack of Legal Status*

Francesc Ortega[†]
CUNY, Queens College

Amy Hsin[‡]
CUNY, Queens College

May 28, 2021

Abstract

Wage gaps between documented (including natives) and undocumented workers reflect employer exploitation, endogenous occupational sorting and productivity losses associated with lack of legal status. Identification of the productivity penalty is crucial to estimate the net economic gains from legalization. Our paper presents a model-based strategy to identify the productivity penalty associated with lack of legal status. In the model, heterogeneous workers choose occupations and undocumented workers are subject to employer discrimination and experience productivity loss in occupations characterized by tasks that require legal status. The theoretical analysis provides guidance on how to identify occupational barriers and on how to compute a lower bound for the undocumented productivity penalty. Applying this strategy to individual-level data that imputes undocumented status, we estimate that the productivity penalty associated with lack of legal status in the United States is at least 5%. This implies that legalization of undocumented workers not only improves their wages, but also increases GDP.

JEL Classifications: J3, J7

Keywords: Immigration; Undocumented; Legalization; Discrimination

*We thank the William T. Grant Foundation and PSC/CUNY for generous financial support. We also thank Catalina Amuedo-Dorantes, Marianne Bertrand, Janet Calvo, Michael Clemens, Albrecht Glitz, Laura Giuliano, Jennifer Hunt, Fernando Lozano, Joan Monras, Giovanni Peri, Amy Ellen Schwartz, Kevin Shih, Todd Sorensen, Chad Sparber, Suleyman Taspinar and Marcos Vera-Hernandez for helpful comments.

[†]Francesc Ortega is the Dina A. Perry Professor of Economics, Queens College, CUNY. E-mail: francesc.ortega@qc.cuny.edu. Homepage: (<http://qcpages.qc.cuny.edu/~fortega>).

[‡]Associate Professor of Sociology, Queens College, CUNY. amy.hsin@qc.cuny.edu. Homepage: (<https://sites.google.com/view/amyhsin>)

1 Introduction

Many countries are home to large numbers of unauthorized immigrants.¹ Despite lacking the right to reside or work legally, unauthorized immigrants contribute in significant ways to the economies of the host countries. Roughly 11 million unauthorized immigrants live in the United States, comprising 5% of the labor force and contributing over 3% of GDP (Edwards and Ortega (2017)). Legalization of undocumented workers is widely debated by policy-makers and social scientists. While questions of human rights and ethics are foundational to these debates, so are questions related to the economic consequences of legalization for host countries.

A large body of literature shows that the wages and working conditions of undocumented immigrants increase when they gain legal status. In the context of the United States, many studies have supported this claim based on the 1986 IRCA legalization (Rivera-Batiz (1999), Kossoudji and Cobb-Clark (2002), Amuedo-Dorantes et al. (2007), Lozano and Sorensen (2011), Pan (2012)) and, more recently, on the 2012 DACA program providing temporary work permits to undocumented youth (Pope (2016), Amuedo-Dorantes and Antman (2017)).

However, the previous evidence is insufficient to answer some key concerns in the debate about the economic effects legalization. Namely, what would be the aggregate effects on the GDP and fiscal balance of the host country? Providing an answer to these questions requires distinguishing how much of the wage increase stemming from gaining legal status can be attributed to a gain in productivity versus other factors, such as the loss of employers' ability to exploit undocumented workers. While the latter mainly entails income redistribution from employers to formerly undocumented workers, productivity increases generate a net increase in income for the host country.

Identifying the productivity gains associated with gaining legal status is a challenging task. While hard to quantify with precision, several studies have shown that unauthorized immigrants suffer wage exploitation (Gleeson and Gonzales (2012), Brown et al. (2013), Bartolucci (2014)). At the same time, there is clear evidence that occupational barriers have created talent misallocation and penalized the productivity of women and blacks in the labor market (Weeden (2002), Hsieh et al. (2019)). The labor market opportunities of undocumented workers are almost certainly diminished by occupational barriers in similar ways (Abrego (2011), Amuedo-Dorantes and Antman (2017)). These

¹The United Nations Development Programme estimated the worldwide unauthorized immigrant population to be over 50 million people in 2009.

barriers vary importantly across occupations, reflecting regulatory constraints, such as legal residence requirements associated with licenses, as well as the nature of the specific tasks involved in each occupation. For instance, the need to hold face-to-face interactions with customers or government agencies, or to travel exposes undocumented workers to apprehension and deportation.

Our paper presents a new strategy to identify the productivity penalty associated with lack of legal status. We lay out a theoretical model where heterogeneous workers choose occupations (as in the Roy model). Some occupations entail tasks that require legal status. As a result, undocumented workers in these occupations suffer a productivity loss that entails lower wages and acts as an entry barrier into those occupations. In addition, employers may exploit undocumented workers in all occupations, paying them wages below productivity. The theoretical analysis suggests an empirical strategy to identify which occupations have entry barriers for undocumented workers, clarifies the factors that determine the productivity and wage gaps between documented (which includes natives) and undocumented workers, and shows how to estimate a lower bound for the undocumented productivity penalty. Additionally, the model also illustrates the labor market effects of legalization in terms of occupational switching, wage growth and aggregate economic gains.

Our main contribution is methodological, developing a model-based strategy to identify the productivity loss associated with lack of legal status in a data environment constrained by the absence of large datasets with reliable information on legal status. Accurate estimates of the undocumented productivity penalty are crucial in structural analyses aimed at estimating the economic contribution of undocumented workers (e.g. [Edwards and Ortega \(2017\)](#)) and to conduct simulations of the effects of legalization policies on the economy (e.g. [Machado \(2017\)](#) and [Ortega et al. \(2019\)](#)).

We apply our approach to individual-level data from the American Community Survey that also includes a sophisticated imputation to identify likely undocumented individuals. Our empirical analysis has two main findings. First, we identify the occupations with the largest entry barriers to undocumented workers. These occupations either require legal status (e.g. teachers, nurses or law enforcement) or entail tasks that involve driving, long-distance travel or face-to-face interaction with the public and government officials (e.g. managers, secretaries or salespersons), which entail a large risk of apprehension. Secondly, we estimate substantial wage penalties associated with lack of legal status. The penalties are substantially higher in occupations featuring high entry barriers (10 log points) relative to occupations without barriers (5 log points). Based on

the theoretical analysis, these estimates imply that the productivity penalty associated with lack of legal status is at least 5 log points.

Our analysis is not only relevant in the United States. Unauthorized immigration is pervasive in high-income countries that are in geographical proximity to countries with demographic, economic or political pressures (Orrenius and Zavodny (2016)). Several studies have used European data to analyze the economic effects of legalization. For instance, Monras et al. (2017) empirically analyze a large legalization process in Spain. Among other findings, they show that legal status increased the labor market opportunities of immigrants. Along similar lines, Devillanova et al. (2018) study the effect of the *prospect* of legal status on the employment of undocumented immigrants in Italy, finding a positive effect. Inevitably, an important factor in the discussions on whether to provide legal status to undocumented workers in receiving countries is the consequences of such a policy for GDP and the public coffers. As argued above, these effects rely crucially on whether legal status increases the productivity of undocumented workers or simply redistributes income from employers to employees.

The structure of the paper is the following. Section 2 summarizes the relevant literature. Section 3 presents our theoretical analysis. Section 4 presents the data and descriptive statistics. Section 5 estimates the size of the occupational barriers faced by undocumented workers, Section 6 estimates the documented-undocumented wage gaps, and Section 7 concludes.

2 Related Literature

Our paper draws on the broad literature analyzing wage gaps by race and gender and applies a similar approach to estimating wage gaps by legal status. The most relevant study for our paper is Hsieh et al. (2019) which identifies the role of occupational barriers in generating misallocation of talent by race and gender in the United States. The authors use a generalized Roy model (Roy (1951)) and consider four groups of individuals, defined by race and gender. In their model, individuals first choose education and later enter the labor market by choosing occupations. These groups face barriers to human capital accumulation and occupational choice to different degrees. These barriers are modeled as *wedges* that increase the cost of acquiring education and entering certain occupations. Using Census and ACS data they back out the evolution of these frictions over time and conduct simulations to evaluate their effects. Their results suggest that about one quarter of US economic growth in the last fifty years can be explained by the

reduction in frictions and the resulting improvement in talent allocation. In our paper, occupational barriers also play a central role as a determinant of the occupational choices of the minority group (i.e. undocumented workers). However, our theoretical model is simpler because we do not consider human capital accumulation, which will understate the overall long-run effects of undocumented status on individual and aggregate income.

Our paper is also connected to the literature on the labor market outcomes of undocumented workers. This literature demonstrates a large wage differential between documented and undocumented workers with similar skills. Using the Survey of Income and Program Participation, [Hall et al. \(2010\)](#) estimated a 17 percent wage gap between documented and undocumented male Mexicans. A number of studies have examined the wage effects of the 1986 IRCA amnesty, estimating undocumented wage penalties ranging between 5 and 20 percent ([Rivera-Batiz \(1999\)](#), [Kossoudji and Cobb-Clark \(2002\)](#), [Amuedo-Dorantes et al. \(2007\)](#), [Lozano and Sorensen \(2011\)](#) and [Pan \(2012\)](#)). [Orrenius and Zavodny \(2015\)](#) provide additional evidence of the wage penalty associated with undocumented status by showing that the introduction of E-Verify, a program that allows employers to verify the legal status of employees, led to a reduction on the wages of undocumented workers.

More recently, [Albert \(2021\)](#) documents differences in wages and job finding rates by legal status. He finds that, conditional on observable characteristics, undocumented immigrants earn 8% less and have a 7 percentage-point higher probability of finding a job than natives. His analysis is based on data from the Current Population Survey. Using these data, the author imputes legal status and also presents a model where firms prefer workers with lower bargaining power (because they can extract more surplus) and are able to discriminate between hiring native and immigrant workers (extending [Chassambouli and Peri \(2015\)](#)). More recently, [Borjas and Cassidy \(2019\)](#) produce estimates of the wage gaps between observationally equivalent immigrants differing in documentation status. They impute documentation status in the 2008-2016 waves of the American Community Survey with an approach similar to the one used in our data. Similar to our own findings, they find that (in 2012-2013) the wages of undocumented workers were roughly 6 log points lower than for documented workers with comparable education and demographic information. The longer time span in their data allows them to describe the evolution over time in the documentation wage gap, producing interesting new findings, such as the reduction in the wage gap after the implementation of the DACA program, which provided temporary legal status to undocumented youth who arrived to the United States as children (also known as Dreamers).

In the last few years, several studies have focused on the effects of DACA on the labor market and educational outcomes of Dreamers. [Pope \(2016\)](#) and [Amuedo-Dorantes and Antman \(2017\)](#) use data from the ACS and CPS, respectively. Lacking information on immigrants' legal status, these authors were forced to assume that non-citizens in a given age range are undocumented. Both studies find positive effects of DACA on employment, but disagree on the effects on schooling. [Hsin and Ortega \(2018\)](#) use administrative data that allows for a precise identification of students' legal status. They find that DACA led to a large increase in dropout rates among undocumented college students enrolled at 4-year colleges (though not among those attending community college). In a recent study, [Kuka et al. \(2020\)](#) provide evidence that DACA incentivized human capital investments among teenagers. In comparison, our study uses data for the period immediately prior to DACA and focuses on the quantification of documented-undocumented wage gaps and identification of the factors generating those gaps.

As noted earlier, the positive effect of legalization on the wages of undocumented workers does not necessarily imply an increase in their productivity and, consequently, on overall GDP. It might simply reflect the strengthening in these workers' bargaining power and the enhanced ability to enforce their rights as workers. Undocumented immigrants reside in the country without work authorization and are a deportable population which makes them especially vulnerable to employer exploitation. Through qualitative analysis, [Gleeson and Gonzales \(2012\)](#) find that undocumented workers are commonly subjected to workplace violation of labor laws and are deterred from filing complaints because they fear employers will retaliate by reporting them to immigration authorities. [Brown et al. \(2013\)](#) analyze administrative data from Georgia state and identify which firms employ undocumented workers on the basis of erroneous social security numbers. The results suggest that firms with undocumented workers experience a competitive advantage, which translates into a higher rate of survival. These studies show that in many industries undocumented workers are desirable source of labor because they are pliable and exploitable. Undocumented workers are also more likely to work in jobs that are physically strenuous and hazardous and receive no compensating differentials for working in dangerous work environments [Hall and Greenman \(2015\)](#), which implicitly amounts to reduced wages relative to documented workers with similar educational attainment.

Nonetheless, a number of studies have documented that illegality negatively affects worker productivity in multiple ways. For instance, the threat of deportation and heavily restricted labor market opportunities increases the risk of depression and anxiety among

undocumented youth (Abrego (2011), Gonzales (2011), Hainmueller et al. (2017), Patler and Pirtle (2018)). Furthermore, undocumented workers also face large occupational barriers. They are prevented from working in occupations that require legal status (e.g. teachers, nurses, law enforcement) and professions that make them vulnerable to immigration enforcement, such as professions that require long-distance travel and significant face-to-face interactions with the public or government officials (e.g. managers, secretaries and salespersons). In this sense, our work also relates to the literature on occupational licensing. Kleiner and Krueger (2013) documented that licensing is associated with higher wages, and Kleiner and Vortnikov (2017) argue that relaxing licensing constraints amounts to a reduction in occupational barriers that leads to lower prices and higher consumer welfare. In a recent study, Blair and Chung (2017) have argued that occupational licensing can be a powerful tool to reduce the wage gaps of women and blacks (relative to white men) by reducing information asymmetries regarding worker productivity.

3 Theoretical Framework

Consider an economy with two occupations, indexed by $o = 1, 2$. Workers are heterogeneous in their idiosyncratic productivity vector $\varepsilon = (\varepsilon_1, \varepsilon_2)$, drawn from joint distribution $f(\varepsilon_1, \varepsilon_2)$ with domain \mathbf{R}^2 . Each worker chooses the occupation that maximizes her income.² There are also two types of workers: documented (which includes natives) and undocumented ($d = D, U$). The measure of documented workers is normalized to 1 and the measure of undocumented workers is $u \leq 1$.

3.1 The occupational choices of documented workers

For documented workers, wages are a function of productivity and each worker chooses the wage-maximizing occupation. As in Borjas (1987) version of the Roy (1951) model, log wages are given by

$$\omega_{io} = \mu_o + \varepsilon_{io}, \tag{1}$$

where μ_o is the occupation-specific mean and ε_{io} the productivity of worker i in occupation o . We will assume that occupation 2 has higher average wages: $\mu_2 \geq \mu_1$.

²We note that we do not require normality or any other distributional assumption, as is often done in most versions of the Roy model.

Each individual i faces a vector of *potential* wages $(\omega_{i1}, \omega_{i2})$. But her *actual* wage depends on the chosen occupation. Thus, individuals' **optimal choice** is summarized by the rule: choose occupation $o = 2$ if and only if

$$\begin{aligned}\omega_{i2} &\geq \omega_{i1} \\ \mu_2 + \varepsilon_{i2} &\geq \mu_1 + \varepsilon_{i1} \\ \varepsilon_{i2} - \varepsilon_{i1} &\geq \mu_1 - \mu_2.\end{aligned}$$

Hence, individuals self-select into the occupation that gets them the highest earnings.

The **optimal allocation** of workers to occupations is as follows. Let D_o denote the set of types (for documented workers) that choose occupation o .³ Then

$$\begin{aligned}D_1 &= \{(\varepsilon_1, \varepsilon_2) : \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2)\} \\ D_2 &= \{(\varepsilon_1, \varepsilon_2) : \varepsilon_2 \geq \varepsilon_1 + (\mu_1 - \mu_2)\}.\end{aligned}$$

As a result, wages vary at the individual level within and across occupations:

$$\omega_i = \begin{cases} \mu_1 + \varepsilon_{i1} & \text{if } (\varepsilon_{i1}, \varepsilon_{i2}) \in D_1 \\ \mu_2 + \varepsilon_{i2} & \text{if } (\varepsilon_{i1}, \varepsilon_{i2}) \in D_2, \end{cases}$$

where set D_1 (D_2) contains the documented workers that choose occupation 1 (occupation 2).

Employment levels can be computed integrating over the appropriate support of the density of types for documented workers.

$$Emp_1^D = Pr(D_1) = \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \int_{\varepsilon_2=-\infty}^{\varepsilon_2=\varepsilon_1+(\mu_1-\mu_2)} f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 d\varepsilon_1$$

and

$$Emp_2^D = Pr(D_2) = \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \int_{\varepsilon_2=\varepsilon_1+(\mu_1-\mu_2)}^{\varepsilon_2=\infty} f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 d\varepsilon_1,$$

³It is helpful to consider a graphical representation in the $(\varepsilon_1, \varepsilon_2)$ -space. Above the indifference line, all workers choose occupation 2 (D_2). Below this line, all workers choose occupation 1 (D_1).

where it is clear that $Emp_1^D + Emp_2^D = 1$.

Marginal distributions given occupational choices. It is helpful to define the following marginal distributions for actual productivity that can be used to compute aggregate employment and average wages. These distributions are a function of the actual occupational choices of the individuals in the group:

$$\begin{aligned} f_1^D(\varepsilon_1) &= \frac{1}{Pr(D_1)} \int_{\varepsilon_2=-\infty}^{\varepsilon_2=\varepsilon_1+(\mu_1-\mu_2)} f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 \\ f_2^D(\varepsilon_2) &= \frac{1}{Pr(D_2)} \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\varepsilon_2-(\mu_1-\mu_2)} f(\varepsilon_1, \varepsilon_2) d\varepsilon_1, \end{aligned}$$

where $Emp_o^D = Pr(D_o)$, for $o=1,2$.

We make the following **symmetry assumption** on joint density f :

$$E(\varepsilon_1|D_1) = E(\varepsilon_2|D_2) = 0$$

so that the average documented worker in occupation $o = 1, 2$ earns log wage $\omega_o = \mu_o$.

Average wages of documented workers in each of the two occupations:

$$\begin{aligned} E(w_1|D_1) &= \mu_1 + \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \varepsilon_1 f_1^D(\varepsilon_1) d\varepsilon_1 \\ &= \mu_1 + E(\varepsilon_1|D_1) = \mu_1, \end{aligned}$$

and

$$\begin{aligned} E(w_2|D_2) &= \mu_2 + \int_{\varepsilon_2=-\infty}^{\varepsilon_2=\infty} \varepsilon_2 f_2^D(\varepsilon_2) d\varepsilon_2 \\ &= \mu_2 + E(\varepsilon_2|D_2) = \mu_2. \end{aligned}$$

3.2 The occupational choices of undocumented workers

The idiosyncratic productivity distribution for undocumented is assumed to be identical to the one for documented workers: $f(\varepsilon_1, \varepsilon_2)$. However, undocumented workers are subject to **exploitation** by their employers, captured by parameter $\tau \geq 0$. As a result, undocumented workers' wages (ω_{io}^U) are lower than their corresponding productivity by a factor τ , assumed to be the same in both occupations.

In addition, undocumented workers also face **entry barriers** that distort their choice of occupation. Specifically, $\phi \geq 0$ will denote the entry barrier into occupation 2 and we assume that occupation 1 has no entry barriers. We view the entry barrier as stemming from the nature of the *tasks* required to perform occupation 2 to its full extent. More specifically, we adopt the view that undocumented workers cannot carry out the whole bundle of tasks in occupation 2 due to certain limitations, such as the inability to drive legally, travel by plane, obtain an occupational license, or freely interact with customers and government officials without the risk of apprehension. These limitations entail a *productivity penalty*. In contrast to a setup with purely psychic costs, here the employer pays a lower wage to an undocumented worker (in occupation 2) because of her diminished productivity relative to a documented worker with the same potential productivity.

In sum, depending on their choice of occupation, the wages of undocumented workers will be given by:

$$\omega_i^U = \begin{cases} (\mu_1 + \varepsilon_{i1}) - \tau & \text{if } (\varepsilon_{i1}, \varepsilon_{i2}) \in U_1 \\ (\mu_2 + \varepsilon_{i2} - \phi) - \tau & \text{if } (\varepsilon_{i1}, \varepsilon_{i2}) \in U_2, \end{cases}$$

where set U_1 (U_2) contains the undocumented workers that choose occupation 1 (occupation 2). The occupation-specific wages above make clear that the wage gap between equally skilled documented and undocumented workers in occupation 1 will stem exclusively from the exploitation term. In contrast, the wage gap in occupation 2 will also reflect the *productivity penalty* associated with lack of legal status.

Because of the productivity loss associated with lack of legal status, the occupational choices of undocumented immigrants will be distorted, leading to under-representation in some occupations and over-representation in others on the basis of the task bundle of each occupation (relate to [Peri and Sparber \(2009\)](#)). Namely, the **optimal occupation choice** for undocumented workers is to choose occupation $o = 2$ if and only if:

$$\begin{aligned} \omega_{i2}^U &\geq \omega_{i1}^U \\ \mu_2 + \varepsilon_{i2} - \phi - \tau &\geq \mu_1 + \varepsilon_{i1} - \tau \\ (\varepsilon_{i2} - \varepsilon_{i1}) &\geq (\mu_1 - \mu_2) + \phi, \end{aligned}$$

which does not depend on the degree of employer exploitation τ .⁴

⁴In the $(\varepsilon_1, \varepsilon_2)$ -space, the indifference line for undocumented workers $\varepsilon_2 = \varepsilon_1 + (\mu_1 - \mu_2) + \phi$ lies

The resulting **allocation** of undocumented workers to occupations is as follows. Let U_o denote the set of undocumented types that choose occupation $o = 1, 2$. Then

$$\begin{aligned} U_1 &= \{(\varepsilon_1, \varepsilon_2) : \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) + \phi\} \\ U_2 &= \{(\varepsilon_1, \varepsilon_2) : \varepsilon_2 \geq \varepsilon_1 + (\mu_1 - \mu_2) + \phi\}. \end{aligned}$$

3.3 Gaps in employment shares

Mismatch region. The entry barrier into occupation 2 induces occupational mismatch among a subset of undocumented workers: some undocumented workers that would find it more beneficial to choose occupation 2 in the absence of the entry barrier, inefficiently choose occupation 1. These worker types are only slightly better at occupation 2 and do not find it worthwhile to “pay” the cost to enter that occupation. Put otherwise, identical individuals make different occupational choices purely on the basis of documentation status. Accordingly, the **mismatch region** is given by

$$MM(\phi) = \{(\varepsilon_1, \varepsilon_2) : \varepsilon_1 + (\mu_1 - \mu_2) < \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) + \phi\}.$$

We note that function MM maps values of ϕ into subsets of the type space. Similarly, $Pr(MM(\phi))$ is a function mapping values of ϕ into the unit interval $[0, 1]$ with the following properties: $Pr(MM(0)) = 0$ and $Pr(MM(\phi))$ is increasing in ϕ . Hence, the size of the mismatch region uniquely identifies the value of ϕ .

Importantly, the occupation allocations of the two types of workers are related as follows:

$$\begin{aligned} U_1 &= D_1 \cup MM \\ D_2 &= U_2 \cup MM. \end{aligned}$$

Thus the set of undocumented workers in occupation 1 equals the set of documented workers in that same occupation together with the mismatch set (i.e. undocumented workers that inefficiently choose occupation 1). Similarly, the set of documented workers in occupation 2 equals the union of the sets of undocumented workers in that same occupation and the set of missing undocumented workers who were mismatched into occupation 1. These relationships will be useful below when we compare the wage gaps between documented and undocumented workers.

higher up than the corresponding line for documented workers ($\varepsilon_2 = \varepsilon_1 + (\mu_1 - \mu_2)$).

Employment levels can be computed integrating over the appropriate support of the density of types. For undocumented workers, we have

$$Emp_1^U = uPr(U_1) = u \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \int_{\varepsilon_2=-\infty}^{\varepsilon_2=\varepsilon_1+(\mu_1-\mu_2)+\phi} f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 d\varepsilon_1$$

and

$$Emp_2^U = uPr(U_2) = u \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \int_{\varepsilon_2=\varepsilon_1-(\mu_1-\mu_2)+\phi}^{\varepsilon_2=\infty} f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 d\varepsilon_1,$$

where it is clear that $Emp^U = Emp_1^U + Emp_2^U = u$.

Similarly, the employment levels for documented workers are given by

$$\begin{aligned} Emp_1^D &= Pr(D_1) \\ Emp_2^D &= Pr(D_2) \end{aligned}$$

and $Emp^D = Emp_1^D + Emp_2^D = 1$.

It is easy to compute overall employment in each occupation using

$$\begin{aligned} Emp_1 &= Pr(D_1) + uPr(U_1) = (1 + u)Pr(D_1) + uPr(MM) \\ Emp_2 &= Pr(D_2) + uPr(U_2) = (1 + u)Pr(D_2) - uPr(MM). \end{aligned}$$

Occupation shares. One of the goals of the empirical analysis will be the comparison of the employment distributions across occupations for documented and undocumented workers. Specifically, we can characterize the documented-undocumented gap in occupational shares as follows:

$$egap_1 = \frac{Emp_1^D}{Emp^D} - \frac{Emp_1^U}{Emp^U} = -Pr(MM(\phi)) \quad (2)$$

$$egap_2 = \frac{Emp_2^D}{Emp^D} - \frac{Emp_2^U}{Emp^U} = Pr(MM(\phi)), \quad (3)$$

where we used that $Pr(D_1) = Pr(U_1) - Pr(MM)$ and $Pr(D_2) = Pr(U_2) + Pr(MM)$. Clearly, we note that the gaps in employment shares are monotonic functions of ϕ and the gaps vanish when entry barriers are removed ($\phi = 0$). We collect these observations in the following proposition.

Proposition 1. *When all workers choose their utility maximizing occupation,*

1. *The employment share in occupation 2 is lower for undocumented workers than for documented workers and the gap is increasing in entry barrier ϕ . In particular, the documented-undocumented gap is given by*

$$egap_2 = Pr(MM(\phi)) \geq 0.$$

2. *Conversely, the employment share in occupation 1 is higher for undocumented workers than for documented workers and the gap (in absolute value) is increasing in ϕ and given by*

$$egap_1 = -Pr(MM(\phi)) \leq 0.$$

This finding will play an important role in the empirical analysis of the paper. Specifically, it shows that the comparison of the occupation shares of documented and undocumented workers identifies which occupations have entry barriers that discourage undocumented workers. Furthermore, the size of the gaps is informative regarding the size of the entry barrier.

3.4 Documented-undocumented gaps in wages

As a result of their occupational choices, the wage schedule for undocumented workers is given by:

$$\omega_i^U = \begin{cases} \mu_1 + \varepsilon_{i1} - \tau & \text{if } \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) \\ \mu_1 + \varepsilon_{i1} - \tau & \text{if } \varepsilon_1 + (\mu_1 - \mu_2) < \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) + \phi \\ \mu_2 + \varepsilon_{i2} - \tau - \phi & \text{if } \varepsilon_2 > \varepsilon_1 + (\mu_1 - \mu_2) + \phi. \end{cases} \quad (4)$$

It is helpful to describe the wage schedule for documented workers using the same partition of the type space, namely,

$$\omega_i^D = \begin{cases} \mu_1 + \varepsilon_{i1} & \text{if } \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) \\ \mu_2 + \varepsilon_{i2} & \text{if } \varepsilon_1 + (\mu_1 - \mu_2) < \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) + \phi \\ \mu_2 + \varepsilon_{i2} & \text{if } \varepsilon_2 > \varepsilon_1 + (\mu_1 - \mu_2) + \phi. \end{cases} \quad (5)$$

As we did earlier for documented workers, it is helpful to construct the marginal type distributions for undocumented workers given their occupational choices. Importantly, these distributions are a function of entry barrier ϕ :

$$\begin{aligned} f_1^U(\varepsilon_1) &= \frac{1}{Pr(U_1)} \int_{\varepsilon_2=-\infty}^{\varepsilon_2=\varepsilon_1+(\mu_1-\mu_2)+\phi} f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 \\ f_2^U(\varepsilon_2) &= \frac{1}{Pr(U_2)} \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\varepsilon_2-(\mu_1-\mu_2)-\phi} f(\varepsilon_1, \varepsilon_2) d\varepsilon_1, \end{aligned}$$

where $Emp_o^U = Pr(U_o)$, for $o=1,2$.

Self-selection and productivity gaps. The above expressions facilitate the comparison of the average *productivity* levels of documented and undocumented workers employed in the same occupation. Clearly, in occupation 1, the average realized productivity of undocumented workers $E(\mu_1 + \varepsilon_{i1}|U_1)$ will be *lower* than for the average documented worker in the same occupation $E(\mu_1 + \varepsilon_{i1}|D_1)$ because the distorted occupational choices of undocumented workers induce *negative selection* into occupation. Namely, $E(\varepsilon_{i1}|U_1) < E(\varepsilon_{i1}|D_1)$.

Let us now turn to occupation 2. Because of the entry barrier into this occupation, undocumented workers will be *positively selected* into the occupation, in the sense of $E(\varepsilon_{i2}|U_2) > E(\varepsilon_{i2}|D_2)$. However, their average productivity *need not* be higher than for the average documented worker in this occupation because of the productivity penalty faced by undocumented workers. Specifically, the average productivity of undocumented workers in occupation 2 will be higher than for the average documented worker in the occupation only if

$$\phi < E(\varepsilon_{i2}|U_2) - E(\varepsilon_{i2}|D_2). \quad (6)$$

Intuitively, the productivity penalty needs to be smaller than the degree of self-selection. The latter finding has an important implication for our empirical analysis: even in the absence of pay discrimination ($\tau = 0$), undocumented workers may be paid less than documented workers in *all* occupations due to the productivity loss associated with lack of legal status.

Wage gaps. It is helpful to define the documented-undocumented (log) wage gap

within each occupation ($o = 1, 2$) by

$$wgap_o = E(\omega_o|D_o) - E(\omega_o|U_o). \quad (7)$$

We can now state the following proposition, which clarifies the relationships between the idiosyncratic skill, productivity and wage gaps of documented and undocumented workers:

Proposition 2: Productivity and wage gaps in occupation 2. *Let $\phi > 0$. When all workers choose their utility maximizing occupation, we observe:*

1. *Positive selection of undocumented workers in occupation 2:*

$$s_2(\phi) = E(\varepsilon_2|U_2) - E(\varepsilon_2|D_2) > 0 \quad (8)$$

2. *The degree of positive selection in occupation 2 intensifies in the size of ϕ . Let $\phi_1 > \phi_0 > 0$. Then $s_2(\phi_1) - s_2(\phi_0) > 0$.*
3. *The average productivity of undocumented workers in occupation 2 will be higher than the corresponding value for documented workers when $\phi < s_2(\phi)$.*
4. *The documented-undocumented log wage gap in occupation 2 is given by*

$$wgap_2(\phi) = [\mu_2 + E(\varepsilon_2|D_2)] - [\mu_2 - \tau - \phi + E(\varepsilon_2|U_2)] \quad (9)$$

$$= \tau + \phi - s_2(\phi) < \tau + \phi. \quad (10)$$

Proof Proposition 2. See Appendix.

It is worth noting that $wgap_2(\phi)$ will generally *not* be monotonic because an increase in entry barrier ϕ will *negatively* affect the productivity of undocumented workers in occupation 2 but, at the same time, it will induce more intense *positive* selection in this occupation.

Analogous arguments prove that undocumented workers employed in the no-barrier occupation will be negatively selected and will exhibit wage gaps (relative to documented workers in the same occupation) that exceed the exploitation wedge. Namely,

Proposition 3: Productivity and wage gaps in occupation 1. *Let $\phi > 0$. When all workers choose their utility maximizing occupation, it is the case that:*

1. *Negative selection of undocumented workers in occupation 1:*

$$s_1(\phi) = E(\varepsilon_1|D_1) - E(\varepsilon_1|U_1) > 0 \quad (11)$$

2. *The degree of negative selection in occupation 1 intensifies in the size of ϕ . Let $\phi_1 > \phi_0 > 0$. Then $s_1(\phi_1) - s_1(\phi_0) > 0$.*
3. *The average productivity of undocumented workers in occupation 1 will be lower than the corresponding value for documented workers.*
4. *The documented-undocumented log wage gap in occupation 1 is given by*

$$wgap_1 = [\mu_1 + E(\varepsilon_1|D_1)] - [\mu_1 - \tau + E(\varepsilon_1|U_1)] \quad (12)$$

$$= \tau + s_1(\phi) > \tau. \quad (13)$$

Proof Proposition 3. See Appendix.

Empirical prediction. Next, we show how the previous results can be used to compute a lower bound for the productivity penalty emerging from lack of legal status. By virtue of point 4 in Proposition 3, the exploitation wedge cannot be larger than the documented-undocumented wage gap in occupation 1, that is, $\tau < wgap_1$. Similarly, point 4 in Proposition 2 implies that $\phi > wgap_2 - \tau$. Combining both conditions yields the following result:

Proposition 4: Lower bound for productivity penalty. *The previous Propositions imply that $\phi > wgap_2 - wgap_1$.*

What is important to note is that, provided $wgap_2 > wgap_1$, the lower bound is informative. Let us provide an example to help fix ideas. Suppose that $wgap_2 = 0.10$ and $wgap_1 = 0.05$. Then $\phi > 0.05$, that is, lack of legal status entails a productivity loss equal to 5 log points or larger. Implicit in *Proposition 4*, the model assumes that the degree of exploitation is similar in occupation 1 and 2, namely, employers in occupations with task-based entry barriers do not extract a larger surplus from undocumented workers than

employers in other occupations. This is a plausible assumption if we view employers' ability to exploit undocumented workers as stemming from the lack of legal protection experienced (or perceived) by workers fearing deportation.

3.5 The Effects of Legalization

Granting legal status to undocumented workers eliminates the frictions ($\tau = \phi = 0$). The following proposition gathers all the effects of such a measure.

Proposition 5. *As a result of legalization ($\phi = \tau = 0$), we expect:*

1. *Occupational switch: the previously mismatched undocumented workers switch from occupation 1 to occupation 2, and the gaps in employment shares vanish: $egap_1 = egap_2 = 0$.*
2. *The productivity and wages of workers with identical idiosyncratic productivity types converge, that is, the wage schedule for all types is given by [Equation \(5\)](#).*
3. *Wages increase for all undocumented workers according to the following schedule:*

$$\Delta\omega_i^U = \begin{cases} \tau & \text{if } \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) \\ \tau + (\mu_2 - \mu_1) + (\varepsilon_{i2} - \varepsilon_{i1}) & \text{if } \varepsilon_1 + (\mu_1 - \mu_2) < \varepsilon_2 < \varepsilon_1 + (\mu_1 - \mu_2) + \phi \\ \tau + \phi & \text{if } \varepsilon_2 > \varepsilon_1 + (\mu_1 - \mu_2) + \phi. \end{cases}$$

The statements in the Proposition follow easily from the earlier results. The elimination of the documented-undocumented gaps in employment shares across occupations follows from the vanishing of the mismatch area (i.e. $Pr(MM(0)) = 0$). Next, setting $\tau = \phi = 0$ implies that the wage schedule for undocumented workers collapses to the schedule for documented workers ([Equation \(5\)](#)). Then the last point simply follows from recognizing that the post-legalization wage schedule for previously undocumented workers is the same as the wage schedule for documented workers so that $\Delta\omega_i^U = \omega_i^D - \omega_i^U$.

The predictions in the previous proposition are largely consistent with the empirical findings of legalization processes, such as the 1986 IRCA in the United States ([Rivera-Batiz \(1999\)](#), [Kossoudji and Cobb-Clark \(2002\)](#), [Amuedo-Dorantes et al. \(2007\)](#), [Lozano and Sorensen \(2011\)](#) and [Pan \(2012\)](#)). In addition, the proposition has two additional corollaries. First, the largest wage increases are experienced by previously undocumented

workers switching from occupation 1 to occupation 2, and by those already employed in occupation 2.⁵

Second, aggregate output (defined as the integral of the individual productivity terms) in the economy rises. This happens for two reasons: previously undocumented workers are now matched to their optimal frictionless occupations, and the productivity penalty disappears so that previously undocumented workers employed in occupation 2 effectively enjoy a boost in productivity.⁶ It is worth noting that in the absence of an undocumented productivity penalty (i.e. with $\phi = 0$), the wages of undocumented workers would still be lower than the wages of comparable documented workers due to pay discrimination (arising from τ). However, legalization would not change the productivity of undocumented workers and, as a result, it would not entail a net increase in GDP. Rather, it would simply redistribute income from employers toward undocumented workers. As a result, assessing whether an undocumented productivity penalty exists (and its magnitude) is crucial for the aggregate economic implications of legalization.

The next section presents the data we will use to compute the documented-undocumented gaps in occupational shares and wages, which will then be used to estimate a lower bound for the productivity penalty due to lack of legal status.

4 Data and Summary Statistics

We use a special extract of the American Community Survey provided by the [Center for Migration Studies \(2014\)](#). Besides the usual information on employment, skills and wages, this confidential dataset contains a sophisticated imputation for documentation status developed by [Warren \(2014\)](#). These data have been used to estimate, by means of calibration and simulation methods, the economic contribution of undocumented workers ([Edwards and Ortega, 2017](#)) and the consequences of providing legal status to Dreamers ([Ortega et al., 2019](#)). In recent studies, [Albert \(2021\)](#) and [Borjas and Cassidy \(2019\)](#) have developed their own imputations and applied it to the CPS and ACS, respectively.

The unauthorized status imputation was first proposed in the 1990's and many authors have contributed to their development over the last few decades ([Passel and Clark \(1998\)](#), [Baker and Rytina \(2013\)](#), [Warren and Warren \(2013\)](#), [Passel and Cohn \(2015\)](#),

⁵If $(\mu_2 - \mu_1) + (\varepsilon_{i2} - \varepsilon_{i1}) > \phi$ then the occupation switchers will experience the largest wage increase. Else, it will be the workers that were already employed in occupation 2 prior to legalization.

⁶Quantification of the actual increase in GDP will also depend on the price of the output produced in each occupation.

and [Warren \(2014\)](#), among several others). The procedure is a 2-step process: (1) applying ‘logical edits’ to identify legal residents on the basis of the information in the ACS; and (2) re-weighting individual observations to match official unauthorized population estimates by country of origin. The main logical edits rely on information on year of arrival (because of the 1986 IRCA amnesty), country of origin, occupation, industry, and receipt of government benefits.⁷ Strictly speaking, we should refer to likely unauthorized individuals but, for simplicity, we will often simply refer to unauthorized (or undocumented) individuals.

We use data for years 2010-2012. The reason is that President Obama’s *Deferred Action on Childhood Arrivals* was rolled out starting at the very end of 2012. This program provided beneficiaries with reprieve from deportation and two-year renewable work permits, which has been shown to have improved substantially the labor market outcomes of its recipients (e.g. [Pope \(2016\)](#) and [Amuedo-Dorantes and Antman \(2017\)](#)). Since we cannot identify DACA recipients in the data, it is preferable to restrict the analysis to the pre-DACA period.

Our data (for year 2012) show that most undocumented have been present in the United States for 16 years or more, and some have resided in the country for three decades ([Figure 1](#)). As a result, undocumented immigrants are deeply rooted in their local communities and make up for 4.5% of the U.S. population. Furthermore, about a third of the undocumented (amounting to approximately 3 million individuals) were brought to the country as children (Dreamers). We restrict the analysis to adult full-time employees. In our data, we estimate that slightly over 5 million workers are likely undocumented, accounting for about 5 percent of full-time employment. In addition, about 1 in 4 undocumented arrived in the United States at age 17 or younger, and are often referred to as Dreamers.

Let us now present descriptive statistics for the variables we will use in the estimation ([Table 1](#)) (using the unweighted sample). The data contain 2.7 million observations and about 4% correspond to likely undocumented workers. The rest are documented workers, which contains both US-born individuals and foreign-born with legal status. The mean log hourly wage is 2.9 across all workers (pooling both documented and undocumented). About 7% of the sample did not graduate from high-school and 35% obtained a 4-year

⁷[Warren \(2014\)](#) argues that the imputation accounts for 89 percent of unauthorized residents, which increases to 93 percent if we add individuals that were unauthorized at some point in the past. Other studies assessing the validity of this methodology are [Pastor and Scoggins \(2016\)](#) and [Van Hook et al. \(2015\)](#).

college degree.⁸ Last, the table also reports the continent of origin of each individual, the degree of English fluency and an indicator for having arrived in the country before the age of 10. Specifically, 7 percent of the individuals in the sample were born in South or Central America (and 5% in Asia), 92% are fluent in English and 87% arrived in the United States by age 10.⁹

5 Estimation of Occupational Barriers

Occupational barriers can be the result of licensing requirements, which often include legal residence. However, these barriers may also stem from the tasks involved in a given occupation. Some occupations require wide exposure to the public, the ability to drive legally, or extensive travel, all of which increase the risk of apprehension faced by undocumented workers.

Proposition 1 suggested a strategy to identify empirically which occupations exhibit entry barriers to undocumented workers. Specifically, in these occupations undocumented workers are under-represented, relative to documented workers with similar characteristics (similar to [Hsieh et al. \(2019\)](#) in their analysis of gender and race discrimination). The characteristics of the occupations exhibiting the largest barriers to the entry of undocumented workers will be informative regarding the nature of those barriers.

5.1 Unconditional gaps in occupational shares

We begin with the simplest implementation of the strategy. Namely, we compute the occupational shares of documented and undocumented workers, disregarding their individual characteristics, and simply compute the documented-undocumented gap for each occupation as in [Equation \(2\)](#).

Unconditional occupational shares can be computed easily as the proportion of individuals employed in each occupation relative to the total number of full-time employed individuals in the group. Accordingly, for each occupation o and group $g = D, U$, we

⁸The ACS data provide 10 categories for the educational attainment of individual respondents. The lowest level is for individuals with completed education up to 4th grade, followed by individuals that completed up to 8th grade. The top two educational categories are a 4-year college degree, and having completed 5 or more years of college (including graduate studies). We aggregate these categories into high-school dropouts (HSD), high-school graduates (HSG) and College graduates (CoGrad).

⁹We consider that US-born individuals arrived in the United States before age 10.

compute $p_o^g = \frac{Emp_o^g}{Emp^g}$. Then the *unconditional* documented-undocumented gap in occupational shares is simply $egap_o = p_o^U - p_o^D$, computed using the appropriate sampling weights. We note that the size of the gap reflects both the differences in the occupational distribution of the two groups and the overall size of the occupation.

Table 2 presents the 20 occupations with the largest gaps. The top 5 occupations in the list are (23) Teachers and Instructors, (32) Nurses and Therapists, (04) Managers and administrators, (57) Secretaries and (08) Accountants, Auditors and Financial specialists. Columns 1 and 2 report the estimated employment levels in each occupation for documented and undocumented workers, respectively. For instance, we estimate that 12.1 million documented workers were employed in occupation 23 (Teachers up to secondary education). In comparison, only 60,000 undocumented workers were employed in this occupation. As shown in columns 3 and 4, the corresponding occupational shares also vary by group. This occupation accounts for 4.4 percent of employment for documented workers, but only for 0.4 percent for undocumented workers, which results in a 4 percentage-point gap.¹⁰ The second row shows a 2.6 percentage-point gap in occupation 32 (Nurses and Therapists), which partly reflects the smaller (employment) size of this occupation relative to Teachers.

More generally, we note that many of the occupations in the Table require occupational licenses (e.g. teachers and healthcare professionals). However, we also find occupations that do not require licensing but entail face-to-face interactions with customers or government officials (e.g. secretaries, retail sales and clerks) or often require driving (e.g. salesmen or mail carriers).

Last, we note that some of the occupations in the Table entail college degrees, such as teachers, registered nurses or lawyers. Thus, the differences in the occupational shares between documented and undocumented workers might also reflect differences in educational attainment. Next, we extend our analysis to take into account individual characteristics in age, gender, education and state of residence. This will allow us to produce a cleaner estimate of occupational entry barriers.

¹⁰Until 2014 licensing requirements for teachers in all U.S. states required legal residence. Thus the estimated 60,000 likely undocumented workers employed in this occupation are a combination of foreign teachers on temporary visas (who are misclassified as undocumented by the imputation) and truly undocumented individuals working as instructors (in community organizations or after-school programs).

5.2 Conditional gaps in occupational shares

Clearly, sociodemographic differences between documented and undocumented workers will also shape their occupations of employment. More specifically, our data show that the average age among documented workers is 40 but it is 4.6 years lower for undocumented workers. Similarly, the share of females among documented workers is 45%, or 17 percentage-points higher than among undocumented. The data also reveal important differences in educational attainment: the share of workers with 4 years of college education is 35% among documented workers, that is, 21 percentage-points higher than for undocumented workers.

To account for the large differences in individual characteristics and characterize *conditional* gaps in occupational shares, we estimate a series of occupation-specific binomial Probit models. Specifically, let d_{io} denote an indicator function taking a value of one if individual i is employed in occupation o , and zero otherwise. Then we postulate that

$$Prob(d_{io} = 1|X_i) = \Phi(\alpha_o + \beta_o Undoc_i + \gamma_o X_i), \quad (14)$$

where Φ is the CDF of the standard normal distribution. A coefficient $\beta_o < 0$ indicates that there exists a barrier to occupation o affecting the entry of undocumented workers. On the basis of the (maximum likelihood) estimates of the coefficients above, we compute *conditional* average effects of undocumented status on occupational shares:

$$E^U(d_{io}|X_i) - E^D(d_{io}|X_i) = \frac{1}{N^U} \sum_{i \in U} \Phi(\hat{\alpha}_o + \hat{\beta}_o + \hat{\gamma}_o X_i) - \frac{1}{N^D} \sum_{i \in D} \Phi(\hat{\alpha}_o + \hat{\gamma}_o X_i),$$

where E^g indicates that the expectation integrates over the subset of individuals belonging to group $g = D, U$, and N^g denotes the corresponding sample size. Clearly, this expression maps easily into the documented-undocumented wage gap in [Equation \(7\)](#): $egap_o = -(E^U(d_{io}|X_i) - E^D(d_{io}|X_i))$.

Columns 1-6 in [Table 3](#) compare the mean age, share of females and share of college graduates among documented and undocumented workers in each occupation. Typically, undocumented workers are younger and less likely to be female. In contrast, in some occupations undocumented workers are more likely to be college graduates than documented workers (e.g. Teachers & Instructors or Salespersons), while the converse is true in other occupations (e.g. Nurses & Therapists or Secretaries).

Columns 7 and 8 report the predicted occupational share in each occupation for doc-

umented and undocumented workers, respectively, evaluated at the mean values of the individual characteristics for the corresponding group. Combining these estimates, column 9 presents the *conditional* documented-undocumented gaps in occupational shares. The estimates show that the set of occupations where undocumented status entails a larger reduction in the probability of employment are practically identical to those based on unconditional gaps (Table 2).

Summing up, our analysis has shown large differences in the occupational distributions of workers on the basis of their legal status that are not accounted for by differences in individual characteristics, such as age, gender, education or state of residence. Thus, our estimates suggest the existence of high barriers of entry into a number of occupations for undocumented individuals. While in some cases these barriers are based on regulations, in others the barriers reflect the tasks required in those occupations.

In light of the findings in Hsieh et al. (2019), the large disparities in the employment distributions across occupations of documented and undocumented workers suggest that occupational barriers associated with lack of legal status lead to misallocation of talent and negatively affect economic growth. In addition, occupational barriers entail a wasteful use of resources, particularly in the case of Dreamers, many of whom have attended public schools and even graduated from college.¹¹

6 Wage gaps and the undocumented productivity penalty

Next, we turn to estimate the wage gaps between observationally equivalent documented and undocumented workers (employed in the same occupations), and to use these estimates to learn about the gap in productivity between the two types of workers.

To bridge the gap between wages and productivity, we need to address two challenges. First, we need to adjust for differences in observable characteristics between documented and undocumented workers. Besides the basic sociodemographic characteristics (age, gender, education), labor economists have long recognized (Chiswick (1991), Chiswick et al. (2005)) that immigrants with an imperfect command of English will suffer a productivity and wage loss until they acquire a higher command of the language. The richness of the ACS allows us to build detailed controls to mitigate this problem.

¹¹Since 2014, several states (such as California and New York) have adopted changes in licensure requirements to allow DACA recipients access to these occupations (Calvo (2017)).

Specifically, we add continent of origin dummies, an indicator for English fluency (taking a value of one for native speakers and for individuals with high fluency), and an indicator for having arrived in the country before the age of 10 (as in [Bleakley and Chin \(2010\)](#)).

The second challenge is more specific to the peculiar labor market status of undocumented workers. As discussed earlier, workers lacking legal status may be subject to exploitation, that is, may receive wages below their productivity. Additionally, their productivity may be diminished because they are unable to carry out some tasks, such as driving, long-distance traveling or interactions with government officials. Moreover, the latter productivity loss may induce self-selection across occupations, as illustrated in the theory presented earlier. As we showed in *Proposition 4*, the comparison between the documented-undocumented wage gaps in occupations with and without entry barriers can be used to produce a *lower bound* for the undocumented productivity penalty.¹² In essence, some tasks inherent to some occupations entail a productivity loss for undocumented workers, which also lowers their wage and acts as an entry barrier. The self-selection induced by these barriers tends to *reduce* the documented-undocumented wage gap in occupations with entry barriers ($wgap_2 = \tau + \phi - s_2(\phi)$) but to *increase* the wage gap in occupations without entry barriers ($wgap_1 = \tau - s_1(\phi)$). As a result, the *difference* between the wage gaps in occupations with entry barriers and occupations without will be higher or equal to the productivity penalty ϕ .

[Table 4](#) collects the estimated wage gaps needed to compute the lower bound for the productivity penalty. To set the stage, the top panel presents estimates of log wage gaps using all occupations.¹³ As seen in column 1, the raw hourly wage for undocumented workers is about 55 log points. Controlling for state of residence, age and gender hardly affects this gap. However, controlling for educational attainment lowers the gap to 25 log points. If we also account for continent of origin, English fluency and arriving in the country as a child, the gap shrinks down to 8 log points. Columns 5-8 shows that the wage gaps are also present within (2-digit) occupations (5 log points) and across different levels of education (5 to 8 log points). All in all, these estimates show that undocumented workers earn lower hourly wages than observationally similar documented workers.¹⁴ However, as our theory made clear, these wage gaps reflect several factors, including

¹²See point 4 in *Propositions 2 and 3*.

¹³The Table presents the estimate corresponding to the undocumented status indicator. Thus, it should be interpreted as the (conditional) undocumented-documented wage gap.

¹⁴Our estimate of the undocumented wage penalty is similar to those reported by [Albert \(2021\)](#), using the 1994-2016 CPS (8 log points), and by [Borjas and Cassidy \(2019\)](#), using the ACS (6 log points for years 2010-2012).

pay discrimination, productivity loss due to lack of legal status and self-selection across occupations.

Guided by the theoretical predictions of our model, we now estimate documented-undocumented wage gaps separately for occupations with and without entry barriers. The middle panel of [Table 4](#) restricts the estimation sample to occupations with entry barriers, that is, occupations where undocumented workers are under-represented relative to documented workers with similar characteristics (age, gender, state of residence and education).¹⁵ As in the previous panel, the unconditional wage gaps are very large (44 log points) but shrink down as we gradually control for individual differences in characteristics. For our purposes, the main specification is reported in column 5. The estimated coefficient for the undocumented indicator shows that undocumented workers employed in occupations with entry barriers have hourly wages that are 10 log-points lower than observationally similar documented workers in those same occupations.

The bottom panel estimates the undocumented wage penalty for occupations without entry barriers, defined as those where undocumented workers are equally or over-represented relative to documented workers with the same characteristics. The estimate in column 5 shows that undocumented workers earn hourly wages that are 5 log-points lower than similar documented workers. Hence, by virtue of *Proposition 4*, the lower bound for the undocumented productivity penalty is $\phi \geq 0.10 - 0.05 = 0.05$ log points (displayed at the bottom of the Table).

Columns 6-8 provide estimates for sub-samples that differ by educational attainment. The estimates show the lower bound for ϕ is zero for workers without a high-school degree, but rises to 6 and 7 log points, respectively, for workers with a high-school degree and with 4 years of college education.

Last, columns 9 and 10 examine the robustness of our estimated lower bound to changes in the set of occupations with entry barriers. The estimates in column 9 are based on the 10 occupations with the highest entry barriers and those in column 10 on the 30 occupations with the highest entry barriers. In both cases the estimated lower bound for ϕ is 0.04, only slightly lower than in column 5.

In sum, the analysis in this section implies that undocumented workers suffer a substantial productivity loss because of their lack of legal status. This loss is found to be, at a minimum, approximately 5% and is likely to be larger for undocumented

¹⁵We focus on the top 20 occupations by documented-undocumented gap in occupational shares. We will also examine the sensitivity of the estimates to changes in the number of restricted-entry occupations used in the analysis.

workers with medium to high educational attainment. We also note that our approach likely underestimates the overall productivity loss because we are taking educational attainment as given. To the extent that the uncertain labor market prospects due to lack of legal status discourage educational investments, the overall productivity penalty is likely to be substantially larger.

7 Conclusions

In policy discussions around legalization of undocumented workers, one of the most prominent and controversial issues is whether legalization entails an increase in GDP and, if so, of what magnitude. A large number of studies quantify this effect using calibrated general equilibrium models. In these analyses, the crucial parameter is the size of the labor productivity increase accompanying legalization. This parameter is typically calibrated on the basis of empirical estimates of the wage gaps between documented and undocumented workers, or the within-person change in the wages of undocumented workers following legalization. Implicitly, this practice equates wages and productivity (either in levels or in changes). As several authors have pointed out, undocumented workers' wages are likely to reflect employer exploitation, which probably changes discontinuously when gaining legal status.

This paper has developed a based strategy to obtain a lower bound for the productivity penalty associated with lack of legal status in the context of a model where undocumented workers may be subject to pay discrimination and self-select across occupations. Our analysis has found that lack of legal status lowers the productivity of undocumented workers by at least 5 percent. This finding implies that legalizing undocumented workers entails a net gain in GDP and provides support for the existing estimates based on calibrated general equilibrium models.

It is important to realize that the overall productivity loss due to lack of legal status may be substantially higher if undocumented youth under-invest in human capital because of the uncertain returns to educational investments, as suggested by [Kuka et al. \(2020\)](#) and [Liscow and Woolston \(2018\)](#), or if lack of legal status reduces productivity through other channels, such as increases in stress and anxiety ([Hainmueller et al. \(2017\)](#), [Patler and Pirtle \(2018\)](#)).

References

- Abrego, Leisy J, “Legal consciousness of undocumented Latinos: Fear and stigma as barriers to claims-making for first-and 1.5-generation immigrants,” *Law & Society Review*, 2011, 45 (2), 337–370.
- Albert, Christoph, “The Labor Market Impact of Immigration: Job Creation versus Job Competition,” *American Economic Journal: Macroeconomics*, January 2021, 13 (1), 35–78.
- Amuedo-Dorantes, Catalina and Francisca Antman, “Schooling and labor market effects of temporary authorization: evidence from DACA,” *Journal of Population Economics*, January 2017, 30 (1), 339–373.
- , Cynthia Bansak, and Steven Raphael, “Gender Differences in the Labor Market: Impact of IRCA,” *American Economic Review*, May 2007, 97 (2), 412–416.
- Baker, Bryan C. and Nancy Rytina, “Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2012,” 2013. Washington, DC: Office of Immigration Statistics, Policy Directorate, US Department of Homeland Security. http://www.dhs.gov/sites/default/files/publications/ois_ill_pe.2012_2.pdf.
- Bartolucci, Cristian, “Understanding the Native-Immigrant Wage Gap Using Matched Employer-Employee Data,” *ILR Review*, October 2014, 67 (4), 1166–1202.
- Blair, Peter and Bobby Chung, “Occupational Licensing Reduces Racial and Gender Wage Gaps: Evidence from the Survey of Income and Program Participation,” Technical Report 2017-50, HCEO working paper 2017.
- Bleakley, Hoyt and Aimee Chin, “Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants,” *American Economic Journal: Applied Economics*, January 2010, 2 (1), 165–192.
- Borjas, George J, “Self-Selection and the Earnings of Immigrants,” *American Economic Review*, September 1987, 77 (4), 531–53.
- Borjas, George J. and Hugh Cassidy, “The wage penalty to undocumented immigration,” *Labour Economics*, 2019, 61, 101757.
- Brown, J. David, Julie L. Hotchkiss, and Myriam Quispe-Agnoli, “Does employing undocumented workers give firms a competitive advantage?,” *Journal of Regional Science*, 2013, 53 (1), 158–170.
- Calvo, Janet, “Professional Licensing and Teacher Certification for Non-Citizens: Federalism, Equal Protection And A States Socio-Economic Interests,” *Columbia Journal of Race and Law*, 2017, 8 (1).
- Center for Migration Studies, “Estimates of the Unauthorized Population,” 2014. Dataset based on the augmented American Community Survey datafiles hosted by IPUMS (Integrated Public Use Microdata Series), 2010 to 2014.
- Chassambouli, Andri and Giovanni Peri, “The Labor Market Effects of Reducing the Number of Illegal Immigrants,” *Review of Economic Dynamics*, October 2015, 18 (4), 792–821.

- Chiswick, Barry R, "Speaking, Reading, and Earnings among Low-Skilled Immigrants," *Journal of Labor Economics*, April 1991, 9 (2), 149–170.
- Chiswick, Barry, Yew Lee, and Paul Miller, "Family matters: the role of the family in immigrants' destination language acquisition," *Journal of Population Economics*, November 2005, 18 (4), 631–647.
- Cho, Esther Yoona, "Revisiting Ethnic Niches: A Comparative Analysis of the Labor Market Experiences of Asian and Latino Undocumented Young Adults," *RSF*, 2017.
- Devillanova, Carlo, Francesco Fasani, and Tommaso Frattini, "Employment of Undocumented Immigrants and the Prospect of Legal Status: Evidence from an Amnesty Program?," *ILR Review*, August 2018, 71 (4), 853–881.
- Edwards, Ryan and Francesc Ortega, "The Economic Contribution of Unauthorized Workers: An Industry Analysis," *Regional Science and Urban Economics*, 2017, 67, 119–134.
- Gleeson, Shannon and Roberto G. Gonzales, "When Do Papers Matter? An Institutional Analysis of Undocumented Life in the United States," *International Migration*, 2012, 50 (4), 1–19.
- Gonzales, Roberto G., "Learning to Be Illegal: Undocumented Youth and Shifting Legal Contexts in the Transition to Adulthood," *American Sociological Review*, 2011, 76 (4), 602–619.
- Hainmueller, Jens, Duncan Lawrence, Linna Martén, Bernard Black, Lucila Figueroa, Michael Hotard, Tomás R Jiménez, Fernando Mendoza, Maria I Rodriguez, Jonas J Swartz et al., "Protecting unauthorized immigrant mothers improves their childrens mental health," *Science*, 2017, 357 (6355), 1041–1044.
- Hall, Matthew and Emily Greenman, "The Occupational Cost of Being Illegal in the United States: Legal Status, Job Hazards, and Compensating Differentials," *International Migration Review*, 2015, 49 (2), 406–442.
- , —, and George Farkas, "Legal Status and Wage Disparities for Mexican Immigrants," *Social Forces*, 2010, 89 (2), 491–513.
- Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow, "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 2019, 87 (5), 1439–1474.
- Hsin, Amy and Francesc Ortega, "The Effects of Deferred Action for Childhood Arrivals on the Educational Outcomes of Undocumented Students," *Demography*, Jun 2018.
- Kleiner, Morris M. and Alan B. Krueger, "Analyzing the Extent and Influence of Occupational Licensing on the Labor Market," *Journal of Labor Economics*, 2013, 31 (S1), 173–202.
- and Evgeny Vorotnikov, "Analyzing occupational licensing among the states," *Journal of Regulatory Economics*, October 2017, 52 (2), 132–158.
- Kossoudji, Sherrie A. and Deborah A. Cobb-Clark, "Coming out of the Shadows: Learning about Legal Status and Wages from the Legalized Population," *Journal of Labor Economics*, July 2002, 20 (3), 598–628.

- Kuka, Elira, Na'ama Shenhav, and Kevin Shih, "Do Human Capital Decisions Respond to the Returns to Education? Evidence from DACA," *American Economic Journal: Economic Policy*, February 2020, 12 (1), 293–324.
- Liscow, Zachary and William Gui Woolston, "Does Legal Status Matter for Educational Choices? Evidence from Immigrant Teenagers," *American Law and Economics Review*, 09 2018, 20 (2), 318–381.
- Lozano, Fernando A. and Todd A. Sorensen, "The Labor Market Value to Legal Status," IZA Discussion Papers 5492, Institute for the Study of Labor (IZA) February 2011.
- Machado, Jol, "Dealing with Undocumented Immigrants: The Welfare Effects of Amnesties and Deportations," *JODE - Journal of Demographic Economics*, December 2017, 83 (4), 445–492.
- Monras, Joan, Javier Vzquez-Grenno, and Ferran Elias Moreno, "Understanding the Effects of Legalizing Undocumented Immigrants," IZA Discussion Papers 10687, Institute of Labor Economics (IZA) April 2017.
- Orrenius, Pia M. and Madeline Zavodny, "The impact of E-Verify mandates on labor market outcomes," *Southern Economic Journal*, 2015, 81 (4), 947–959.
- and —, "Irregular immigration in the European Union," Working Papers 1603, Federal Reserve Bank of Dallas January 2016.
- Ortega, Francesc, Ryan Edwards, and Amy Hsin, "The Economic Effects of Providing Legal Status to DREAMers," *IZA Journal of Labor Policy*, 2019, 9:2.
- Pan, Ying, "The Impact of Legal Status on Immigrants Earnings and Human Capital: Evidence from the IRCA 1986," *Journal of Labor Research*, 2012, 33 (2), 119–142.
- Passel, Jeffrey S. and D'Vera Cohn, "Share of Unauthorized Immigrant Workers in Production, Construction Jobs Falls Since 2007: In States, Hospitality, Manufacturing and Construction are Top Industries," March 26 2015. Washington, D.C.: Pew Research Center.
- and Rebecca L. Clark, "Immigrants in New York: Their Legal Status, Incomes and Taxes," 1998. Washington, DC: Urban Institute.
- Pastor, Manuel and Justin Scoggins, "Estimating the Eligible-to-Naturalize Population," March 8 2016. Center for the Study of Immigrant Integration.
- Patler, Caitlin and Whitney Laster Pirtle, "From undocumented to lawfully present: Do changes to legal status impact psychological wellbeing among latino immigrant young adults?," *Social Science Medicine*, 2018, 199, 39 – 48. The role of Racism in Health Inequalities: Integrating Approaches from Across Disciplines.
- Peri, Giovanni and Chad Sparber, "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, July 2009, 1 (3), 135–169.
- Pope, Nolan G., "The Effects of DACAmentation: The Impact of Deferred Action for Childhood Arrivals on Unauthorized Immigrants," *Journal of Public Economics*, 2016, 143 (C), 98–114.
- Rivera-Batiz, Francisco L., "Undocumented workers in the labor market: An analysis of the earnings of legal and illegal Mexican immigrants in the United States," *Journal of Population Economics*, 1999, 12 (1), 91–116.

- Roy, A. D., "Some Thoughts On The Distribution Of Earnings," *Oxford Economic Papers*, 1951, 3 (2), 135–146.
- Van Hook, Jennifer, James D. Bachmeier, Donna L. Coffman, and Ofer Harel, "Can We Spin Straw Into Gold? An Evaluation of Immigrant Legal Status Imputation Approaches," *Demography*, February 2015, 52 (1), 329–354.
- Warren, Robert, "Democratizing Data about Unauthorized Residents in the United States: Estimates and Public-Use Data, 2010 to 2013," *Journal on Migration and Human Security*, December 2014, 2 (4), 305–328.
- and John Robert Warren, "Unauthorized Immigration to the United States: Annual Estimates and Components of Change, by State, 1990 to 2010," *International Migration Review*, Summer 2013, 47 (2), 296–329.
- Weeden, Kim A., "Why Do Some Occupations Pay More than Others? Social Closure and Earnings Inequality in the United States," *American Journal of Sociology*, 2002, 108 (1).

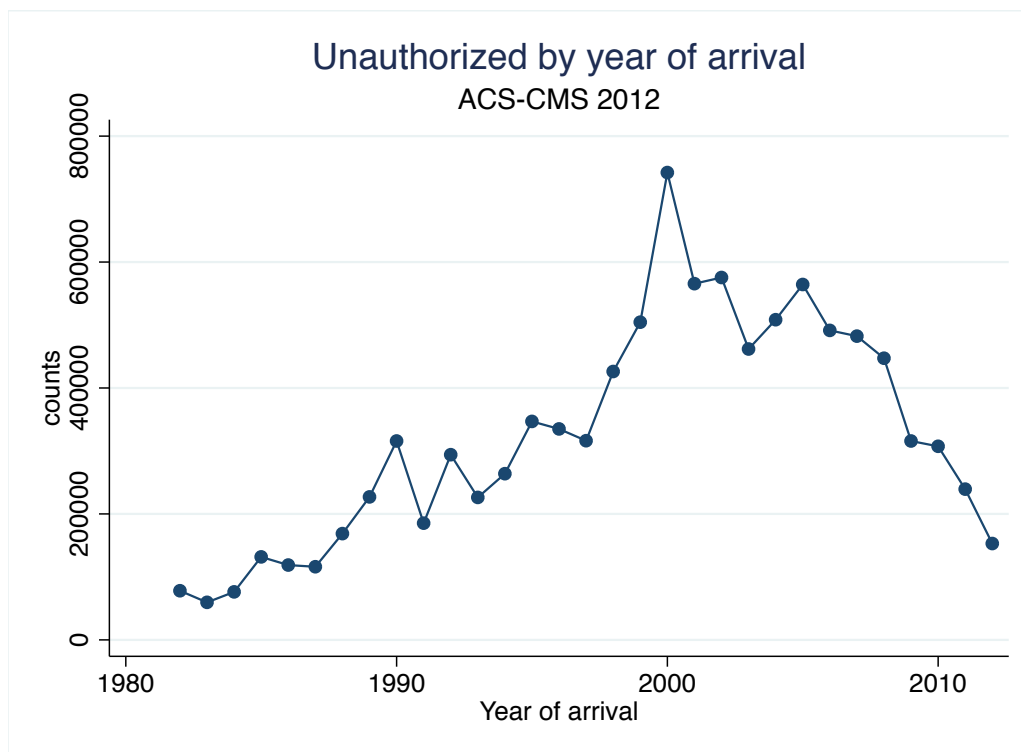


Table 1: Summary statistics ACS 2010-2012

Variable	Obs	Mean	Std. Dev.	Min	Max
Variable	Obs	Mean	Std. Dev.	Min	Max
Year	2,763,538	2011.00	0.82	2010	2012
Undocumented	2,763,538	0.04	0.18	0	1
Dreamer	2,763,538	0.01	0.09	0	1
Age	2,763,538	44.10	12.38	18	77
Female	2,763,538	0.45	0.50	0	1
HSD	2,763,538	0.07	0.25	0	1
HSG	2,763,538	0.23	0.42	0	1
CoGrad	2,763,538	0.35	0.48	0	1
Log hourly wage	2,763,252	2.87	0.77	-7.16	5.85
Origin South or Central America	2,763,538	0.07	0.25	0	1
Origin Europe	2,763,538	0.02	0.15	0	1
Origin Asia or Oceania	2,763,538	0.05	0.21	0	1
Origin Africa	2,763,538	0.01	0.07	0	1
Fluent English	2,763,538	0.92	0.27	0	1
Arrival by age 10	2,763,538	0.87	0.34	0	1

Notes: Pooled data for the CMS-ACS for period 2010-2012. Unweighted statistics. The sample restricts to full-time employed individuals (with over 30 weekly work hours), older than 18 years old. HSD is an indicator for high-school dropouts, HSG is an indicator for high-school graduation (but no more education) and CoGrad is an indicator for having completed 4 years of college (or more). All US-born individuals are assumed to be fluent in English and considered to have arrived in the country before age 10.

Table 2: Gaps in occupational shares (unconditional). Top 20 occupations by size of the gap

Occ2digit	Description	(1)		(2)		(3)		(4)		(5)	
		Doc Emp (th.)	Undoc Emp (th.)	Doc Emp (th.)	Undoc Emp (th.)	Doc Share (%)	Undoc Share (%)	Doc Share (%)	Undoc Share (%)	Doc-Undoc Gap Share (%)	Doc-Undoc Gap Share (%)
23	Teachers (up to secondary) and Instructors	12,100	60	4.4	0.4					4.0	
32	Nurses and Therapists	7,441	10	2.7	0.1					2.6	
04	Managers and administrators	8,698	136	3.2	0.9					2.3	
57	Secretaries	7,506	76	2.7	0.5					2.2	
08	Accountants, Auditors and Financial specialists	6,184	18	2.3	0.1					2.1	
00	CEO, Managers and administrators	6,846	82	2.5	0.5					2.0	
48	Salespersons, Advertising	6,786	106	2.5	0.7					1.8	
01	Financial Managers, Human Resources	7,149	127	2.6	0.8					1.8	
20	Social Workers, Counselors, Clergy	5,156	50	1.9	0.3					1.5	
47	Retail Sales Clerks, Cashiers	17,400	733	6.3	4.8					1.5	
51	Bookkeepers, billing clerks, bank tellers	5,749	90	2.1	0.6					1.5	
38	Police, Law Enforcement	3,470	0	1.3	0.0					1.3	
58	Secretaries, insurance adjusters and examiners	4,956	80	1.8	0.5					1.3	
52	Customer service reps, file clerks	5,783	128	2.1	0.8					1.3	
21	Lawyers, legal assistants	3,334	10	1.2	0.1					1.1	
50	Office supervisors	3,607	45	1.3	0.3					1.0	
30	Physicians, Pharmacists, Dentists, Nutritionists	2,737	3	1.0	0.0					1.0	
02	Managers in education, Farmers	3,308	47	1.2	0.3					0.9	
05	Purchasing managers, Insurance adjusters	2,445	24	0.9	0.2					0.7	
55	Mail carriers postal service, Dispatchers	2,585	35	0.9	0.2					0.7	

Notes: Pooled data for the CMS-ACS for period 2010-2012, full-time employed individuals, older than 18 years old. Two-digit occupations based on 2000-2017 ACS occupational codes (recoded by IPUMS in variable *OC*). Columns 1 and 2 report estimated employment in thousands. Columns 3 and 4 report each occupation's share in overall employment for the corresponding group. Column 5 is computed as column 3 minus column 4, that is, the difference in the proportion of undocumented workers in an occupation and the proportion of documented workers in that same occupation, computed using sampling weights.

Table 3: Gaps in occupational shares (conditional). Top 20 occupations by size of the gap

occ index	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Age Doc	Age U-Doc	Fem. Doc	Fem. U-Doc	CoG Doc	CoG U-Doc	ShOcc Doc	ShOcc U	ShOcc Doc-U
	All Occup.	40.0	-4.6	45.0	-17.0	35.0	-21.0			
23	Teachers (up to sec.) and Instructors	40.2	-3.1	75.6	5.3	88.0	-26.7	4.5	0.5	4.0
32	Nurses and Therapists	41.4	-5.1	87.9	-14.2	60.6	20.2	2.8	0.1	2.7
04	Managers and administrators	42.6	-5.0	41.1	-7.1	57.8	2.1	3.2	0.9	2.2
08	Accountants, Auditors, Fin. spec.	40.0	-5.7	56.6	-17.5	79.0	2.0	2.3	0.1	2.2
57	Secretaries	42.8	-8.0	95.6	-8.1	19.9	8.1	2.6	0.5	2.0
00	CEO, Managers and administrators	42.5	-3.2	34.4	-2.3	64.3	9.6	2.5	0.6	1.9
01	Financial Managers, Human Resources	42.1	-4.8	42.6	-13.9	58.3	14.6	2.7	0.9	1.8
48	Salespersons, Advertising	40.3	-4.5	35.4	-0.4	51.7	-10.1	2.5	0.7	1.8
47	Retail Sales Clerks, Cashiers	38.5	-4.1	47.1	1.6	23.1	-6.1	6.3	4.7	1.6
20	Social Workers, Counselors, Clergy	40.6	-1.4	64.6	-24.1	74.5	-7.7	1.8	0.3	1.5
51	Bookkeepers, billing clerks, bank tellers	40.8	-6.0	84.7	-14.7	17.8	12.9	2.0	0.6	1.4
38	Police, Law Enforcement	38.8	NA	19.1	NA	28.7	NA	1.3	0.0	1.3
52	Customer service reps, file clerks	38.3	-3.4	69.2	-12.9	24.5	2.6	2.1	0.8	1.3
58	Secretaries, ins. adjustors, examiners	40.6	-5.9	79.6	-8.0	20.8	1.5	1.8	0.5	1.2
21	Lawyers, legal assistants	40.1	-4.9	57.9	12.2	76.6	-0.9	1.3	0.1	1.2
50	Office supervisors	42.1	-5.4	62.3	-15.8	31.1	1.2	1.3	0.3	1.0
30	Physicians, Pharm., Dentists, Nutrit.	40.2	-1.2	46.4	41.8	97.0	-62.1	1.0	0.0	1.0
02	Managers in education, Farmers	42.7	-4.4	40.6	-32.1	62.8	-46.9	1.2	0.4	0.8
05	Purchasing managers, Ins. adjusters	41.2	-5.6	55.6	-14.0	48.1	4.7	0.9	0.2	0.7
06	Personnel, HR, Training	40.1	-3.4	62.6	-3.6	54.5	0.4	0.9	0.1	0.7

Notes: Pooled data for the CMS-ACS for period 2010-2012, full-time employed individuals, older than 18 years old. Two-digit occupations based on 2000-2017 ACS occupational codes (recoded by IPUMS in variable *OC*). Column 1 reports the mean age for documented workers in the corresponding occupation. Column 2 reports the gap in the mean age of undocumented workers relative to documented workers in the same occupation. Similarly, columns 2-6 report the shares of female and college graduates, by occupation, and the corresponding undocumented-documented gaps. Columns 7 and 8 are based on the estimation of bivariate Probit models (for the choice of that particular occupation) with controls for age intervals, gender, education levels and state of residence (dummies), using sampling weights. Column 7 (Column 8) reports the predicted probability (occupation share in employment) evaluated at the mean values for documented (undocumented) workers. Column 9 is the gap in the predicted probabilities. NA values in the Table appear in occupations with zero employment of (likely) undocumented workers.

Table 4: Wage gaps and lower bound productivity penalty

Dep. Var. $\ln w$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
All Occupations										
Undoc	-0.55*** [0.003]	-0.53*** [0.003]	-0.25*** [0.003]	-0.08*** [0.003]	-0.05*** [0.003]	-0.05*** [0.005]	-0.08*** [0.005]	-0.05*** [0.007]		
Observations	2,717,902	2,717,902	2,717,902	2,717,902	2,717,902	178,322	1,580,750	958,830		
R-squared	0.025	0.144	0.302	0.310	0.391	0.160	0.267	0.313		
Occ with Barriers										
Undoc	-0.44*** [0.009]	-0.42*** [0.009]	-0.28*** [0.008]	-0.15*** [0.008]	-0.10*** [0.008]	-0.06*** [0.022]	-0.13*** [0.011]	-0.10*** [0.013]	-0.09*** [0.009]	-0.09*** [0.007]
Observations	1,210,473	1,210,473	1,210,473	1,210,473	1,210,473	20,738	588,254	601,481	840,491	1,459,237
No Barriers										
Undoc	-0.36*** [0.003]	-0.34*** [0.003]	-0.18*** [0.004]	-0.06*** [0.004]	-0.05*** [0.004]	-0.06*** [0.006]	-0.07*** [0.006]	-0.03*** [0.011]		
Observations	788,015	788,015	788,015	788,015	788,015	128,195	550,565	109,255		
Sample	All	All	All	All	All	HSD	HSGrad	CoGrad	All	All
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age and Fem	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Educ FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin Fluent Child	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occ FE	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
ϕ Lower Bound	0.08	0.08	0.1	0.09	0.05	0	0.06	0.07	0.04	0.04

Notes: Pooled data for the CMS-ACS for period 2010-2012. The sample restricts to full-time employed individuals, age 19-67. The dependent variable is the hourly wage in logs. The sample in panel 1 includes all occupations. Panel 2 restricts to occupations with entry barriers (based on conditional gaps in occupational shares). Panel 3 restricts to occupations where the share of employment is higher or equal among undocumented workers. Occupations are defined at the 2-digit level. Heteroskedasticity-robust standard errors reported in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix

A Proofs

Proof Proposition 2. Let us first show that $s_2(\phi) > 0$.

1. In the ability-type space, the set of documented workers who choose occupation 2 includes the set of undocumented workers that make the same choice. Provided $\phi > 0$ there is a non-trivial region where the two sets do not overlap, which we define as $MM(\phi)$. We also note that this region expands as ϕ increases.
2. The average ability of the documented workers who choose occupation 2 can be expressed as a weighted average of the average ability of those workers over the regions U_2 (where undocumented workers also choose occupation 2) and $MM(\phi)$.
3. Therefore

$$s_2(\phi) = E(\varepsilon_2|U_2) - E(\varepsilon_2|D_2) \quad (\text{A.1})$$

$$= E(\varepsilon_2|U_2) - \lambda E(\varepsilon_2|U_2) - (1 + \lambda)E(\varepsilon_2|MM) \quad (\text{A.2})$$

$$= (1 - \lambda) [E(\varepsilon_2|U_2) - E(\varepsilon_2|MM)]. \quad (\text{A.3})$$

4. Thus, proving that $s_2(\phi) > 0$ requires showing that $E(\varepsilon_2|U_2) - E(\varepsilon_2|MM) > 0$. Writing out the integrals, $E(\varepsilon_2|U_2) - E(\varepsilon_2|MM)$ becomes

$$\begin{aligned} &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \left[\int_{\varepsilon_2=\varepsilon_1+\mu_1-\mu_2+\phi}^{\varepsilon_2=\infty} \varepsilon_2 f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 \right] d\varepsilon_1 - \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \left[\int_{\varepsilon_2=\varepsilon_1+\mu_1-\mu_2}^{\varepsilon_2=\varepsilon_1+\mu_1-\mu_2+\phi} \varepsilon_2 f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 \right] d\varepsilon_1 \\ &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \left[\int_{\varepsilon_2=\varepsilon_1+\mu_1-\mu_2+\phi}^{\varepsilon_2=\infty} \varepsilon_2 f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 - \int_{\varepsilon_2=\varepsilon_1+\mu_1-\mu_2}^{\varepsilon_2=\varepsilon_1+\mu_1-\mu_2+\phi} \varepsilon_2 f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 \right] d\varepsilon_1 \\ &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} [I(\varepsilon_1, \phi) - J(\varepsilon_1, \phi)] d\varepsilon_1. \end{aligned}$$

5. Now we can bound this expression above zero as follows. Integral $I(\varepsilon_1, \phi) \geq \varepsilon_1 + \mu_1 - \mu_2 + \phi$, which would result from concentrating the probability on the lower bound for ε_2 . Likewise $J(\varepsilon_1, \phi) \leq \varepsilon_1 + \mu_1 - \mu_2 + \phi$, which would result from concentrating the probability on its upper bound for ε_2 . Hence,

$$I(\varepsilon_1, \phi) - J(\varepsilon_1, \phi) \geq (\varepsilon_1 + \mu_1 - \mu_2 + \phi) - (\varepsilon_1 + \mu_1 - \mu_2 + \phi) = 0.$$

6. Integration over all possible values of ε_1 will therefore also yield a non-negative value.

Let us now turn to show that $s_2(\phi)$ is an increasing function of ϕ . Consider two values $\phi_1 > \phi_0 > 0$. It follows that

$$s_2(\phi_1) - s_2(\phi_0) = E(\varepsilon_2|U_2, \phi_1) - E(\varepsilon_2|U_2, \phi_0),$$

given that $E(\varepsilon_2|D_2)$ does not depend on ϕ . Now let us examine the expression for the average ability in occupation 2 for undocumented workers who choose that occupation in its integral form:

$$E(\varepsilon_2|U_2, \phi) = \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \left[\int_{\varepsilon_2=\varepsilon_1+\mu_1-\mu_2+\phi}^{\varepsilon_2=\infty} \varepsilon_2 f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 \right] d\varepsilon_1.$$

The important observation is that the lower bound of the integral inside the square brackets is increasing in ϕ . Thus, $E(\varepsilon_2|U_2, \phi)$ will increase as ϕ increases.

End of Proof.

Proof Proposition 3. Analogous to the proof of Proposition 1. **End of Proof.**

B Tables

Table 5: Full-time employed by documentation status

year	Undoc	Dreamers	Undoc/All	Dreamers/All
2010	5,126,668	1,178,365	5.4%	1.2%
2011	5,065,635	1,180,535	5.3%	1.2%
2012	5,093,545	1,193,015	5.2%	1.2%

Notes: CMS-ACS 2010-2012. Full-time, employed individuals older than 18. Dreamers defined as likely undocumented individuals arrived in the United States at age 17 or younger. Survey weights used.