Clicking on Heaven's Door: The Effect of Immigrant Legalization on Crime*

Paolo Pinotti
Università Bocconi and BAFFI Centre

March 2014

Abstract

We estimate the effect of legal status on the number of crimes committed by immigrants in the destination country, exploiting the Italian policy framework as an ideal regression discontinuity design: fixed quotas of residence permits are available each year, applications must be submitted electronically on specific "Click Days" and are processed on a first-come, first-served basis until the available quotas are exhausted. By matching data on applicants – including the timing of the application in milliseconds – with restricted-use police files, we find that obtaining legal status reduces the number of serious crimes committed in the following year from 2.9 to 1.2 per 100 applicants. The effect is greater on economic crimes, in regions offering better economic opportunities to legal immigrants, and under weak enforcement of migration restrictions. These findings are consistent with the predictions of a simple model of crime.

keywords: legal status, crime, regression discontinuity design
JEL codes: J61, K37, K42

*Contacts: Paolo Pinotti, Università Bocconi and BAFFI Centre, Via Roentgen 1, 20136 Milan, Italy. E-mail: paolo.pinotti@unibocconi.it. I would like to thank Alberto Abadie, Jerome Adda, Josh Angrist, Tito Boeri, Christian Dustmann, Andrea Ichino, Steve Machin, Magne Mogstad, Nicola Pavoni, Nicola Persico, Luigi Pistaferri, Imran Rasul, and the seminar participants at the IFS Workshop on Advances in Policy Evaluation, UCL-NHH Crime Conference, European University Institute, and HEC Lausanne for their useful comments. The "Dipartimento Libertà civili e Immigrazione" (Immigration Department) and the "Direzione Centrale della Polizia Criminale" (Criminal Police Department) of the Italian Ministry of Interior kindly provided the data on residence permits and crime. Financial assistance from the Fondazione Rodolfo Debenedetti is gratefully acknowledged.

1 Introduction

The most recent estimates place the number of undocumented immigrants in the United States at 11.5 million, up from 8.5 million in 2000 and close to exceeding the number of legal permanent residents (Hoefer et al., 2012a, Hoefer et al., 2012b). The share of illegals is also high in several European countries, such as Italy, Greece, Portugal, and Spain.[1]

Both in the US and in Europe, illegal immigrants cannot be employed or start a new economic activity, at least officially. Such immigrants can work only in the informal economy and thus receive considerably lower earnings than legal immigrants (Kossoudji and Cobb-Clark, 2002, Kaushal, 2006, Amuedo-Dorantes et al., 2007, Accetturo and Infante, 2010, and Lozano and Sorensen, 2011). According to the Becker-Ehrlich model of crime, a lower income from legitimate activities means a lower opportunity cost of – and thus a higher propensity to participate in – illicit activities.

Therefore, the presence of large illegal populations raises crime concerns in destination countries. According to an annual survey conducted in North American and European countries, approximately two-thirds of the people interviewed are concerned that illegal immigrants increase crime, whereas this share is reduced by half for legal immigrants (Transatlantic Trends, 2009). Moving from perceptions to criminal statistics, illegals constitute between 20% and 30% of all immigrants in Italy, but they represent 80% of those arrested for serious crimes (Italian Ministry of Interior, 2007).

However, these figures may reflect the different composition of the two groups, as opposed to the (causal) effect of legal status. In particular, illegal immigrants in all countries are typically young, single males and are less educated than legal immigrants (Cohn and Passel, 2009, Mastrobuoni and Pinotti, 2014, and Caponi and Plesca, 2013). More generally, the two groups could differ along other (possibly unobserved) dimensions that are relevant to criminal behavior. For instance, individuals who are less risk averse or those who have a higher propensity to violate laws would be more likely to reside illegally in the country and to commit crimes. It is thus difficult to identify the causal effect of legal status on the number of crimes committed by immigrants in the destination country.

The institutional framework in Italy provides an ideal Regression Discontinuity

[1] The Italian case is discussed in detail in the following sections. As for the other European countries, upper and lower bound estimates for the size of the illegal immigrant population are provided by the Clandestino Project (the documentation and data are publicly available online at http://research.icmpd.org/1244.html).

(RD) design that allows this effect to be estimated separately from selection into legal status. The primary method of acquiring legal status in Italy is through a work-related residence permit: fixed quotas of permits are available each year to different categories of immigrants; applications in each category must be submitted online by (prospective) employers of foreign workers starting on a given "Click Day"; and applications are processed on a first-come, first-served basis until the available quotas are exhausted. The rationing of quotas and the frequency of applications during the first hours of Click Days are such that several thousand applicants are denied legal status every year simply because their employer applied a few minutes (or seconds) after the cutoff time.[2]

In this paper, we match data on the applicants for 2007 – including the timing of the application in milliseconds – with restricted-use data on all foreigners prosecuted for having committed a (serious) crime in Italy in the subsequent year, and we compare the number of crimes per applicant between those who applied shortly before and those who applied shortly after the cutoff. For most categories of immigrants, this cutoff occurred fewer than 30 minutes after the start of the Click Day. Most important, the exact timing of the cutoff for each category was unknown ex ante, as it depended on the timing of all applications as well as on how many were rejected for being inaccurate, false, or incomplete. These complexities provide a compelling argument for the fundamental identification assumption that immigrants applying within a narrow bandwidth of the cutoff were unable to *precisely* determine assignment to either side of it. Indeed, density and balance tests cannot reject the hypothesis that residence permits are as good as randomly assigned across this subset of applicants.

The results of the nonparametric kernel regression (Hahn et al., 2001, Porter, 2003) suggest that immigrants who applied shortly before and shortly after the cutoff commit 1.2 and 2.0 crimes per 100 applicants, respectively, during the following year, and the difference in the probability of obtaining legal status between these two groups is 46 percentage points. Given the ratio of the reduced form and the first-stage effect of applying on time, the average effect among those who actually obtained legal status is -1.7 crimes per 100 applicants.

The results are extremely stable when we use alternative bandwidth selection criteria – the theory-based criteria by Imbens and Kalyanaraman (2012) and Cattaneo et al. (2014) and a battery of heuristic bandwidths between 1 and 30

[2]Although applications are materially sent by employers, in the following analysis, for convenience, we use the term "applicant" to denote the immigrant sponsored in the application and the term "sponsor" to denote the employer.

minutes – as well as when we use various specifications of global polynomial regressions. When we distinguish between different types of crimes and different categories of applicants, the effect is greater for economically motivated crimes, applicants experiencing better economic opportunities after obtaining legal status, and applicants subject to a higher risk of expulsion as illegal immigrants.

These findings contribute to the growing body of evidence regarding the effect of immigration on crime.[3] Earlier work by Butcher and Piehl (1998) shows no evidence that immigration affected crime across US counties during the 1980s, whereas Spenkuch (2014) reaches a different conclusion for subsequent periods. Borjas et al. (2010) also find that immigration increases crime, although only indirectly (by raising the crime rate of native black males).

As for other countries, Bianchi et al. (2012) show that the causal effect across Italian provinces is not significantly different from zero, while Alonso-Borrego et al. (2012) obtain an opposite result across Spanish provinces. Finally, Bell et al. (2013) distinguish between the effect of two large immigrant waves in the UK, namely, asylum seekers and the post-2004 inflow from EU accession countries. Interestingly, only the former group, which was characterized by limited economic opportunities in the official labor markets, caused a significant increase in (property) crimes. This last result suggests that average estimates across geographical areas may mask a significant degree of heterogeneity in the effect of different groups of immigrants. Indeed, the propensity to commit crimes should differ across immigrants depending on individual characteristics, such as age, gender, and education, in the same manner as for natives. However, immigrants differ among themselves with respect to another important factor: legal status.

Previous evidence on the effects of legal status is primarily based on the 1986 Immigration Reform and Control Act (IRCA), which granted legal resident status to long-time illegal immigrants in the US. Using distance from the ports of entry and 1960 immigrants' enclaves as instruments for the presence of IRCA applicants, Baker (2013) shows that higher legalization rates caused a reduction in crime across US counties. At the same time, the IRCA enforced stronger control over the hiring of illegal immigrants, creating obstacles to the employment of those who were not legalized. Freedman et al. (2013) analyze the implications of these provisions and document an increase in felony charges filed against Hispanic residents of San Antonio, Texas, after the expiration of the IRCA amnesty deadline. With regard to other countries, in a previous paper with Giovanni Mastrobuoni (2014), we exploit variation in legal status among pardoned prison inmates in Italy

[3] Bell and Machin (2013) provide a survey of this literature.

after the EU expansion of January 2007. We find that after the EU accession, recidivism declined markedly – from 5.8% to 2.3% over a six-month period – among inmates from the new EU member countries, whereas no change occurred in a control group of inmates from EU-candidate member countries.

In this paper, we contribute to this strand of literature by employing a clean quasi-experimental design. Under the assumption of no manipulation of the assignment variable, the RD approach ensures greater internal consistency than other quasi-experimental methods, comparable (at least locally) to that of randomized controlled trials (Lee, 2008). With regard to the external validity of our estimates, the present paper examines the effect of changes in legal status that are routinely induced by migration policy, as opposed to exceptional events such as one-time amnesties or the EU expansion. Indeed, the institutional framework considered here is not specific to the Italian context, as immigration policy in many destination countries (e.g., Austria, Canada, and Spain) is based on analogous quota systems.

This paper is organized as follows. The next section describes the functioning of Click Days. Section 3 provides a simple theoretical framework to guide the empirical analysis. Section 4 discusses the empirical strategy and the data. Section 5 presents the main results. Finally, Section 6 concludes the paper with some policy implications for the international debate on immigration reform.

2 Click Days

This section introduces the main features of Italian migration policy, with a particular focus on Click Days as sources of exogenous variation in legal status.[4]

2.1 The quota system

Italian migration policy is based on a quota system that limits the number of permits available for working purposes by country of origin, province of destination, and employment sector. The total quotas for permits available at the national level are established in the previous year by the government through the Flows Decree ("Decreto Flussi"). Specific quotas are reserved for immigrants coming from 14 "privileged" countries that subscribe to bilateral agreements to control illegal migration. The remaining residence permits are awarded to immigrants

[4]See Del Boca and Venturini, 2005, for a comprehensive analysis of Italian policy.

from other countries with no cap by nationality; rather, fixed quotas are given according to employment sector (domestic work, construction sector, and all other sectors). Table 1 shows the quotas established by the Flows Decree 2007 and the ratio of quotas to applications: slightly more than one-fourth of the applicants eventually obtained legal status.[5]

Table 1: Total quotas fixed by the Flows Decree 2007 and applications for residence permits received, by category of applicant

CATEGORY OF APPLICANTS	QUOTAS	APPLICATIONS	RATIO
Privileged nationalities, total	44,600	352,987	0.13
Albania	4,500	28,564	0.16
Algeria	1,000	1,904	0.53
Bangladesh	3,000	55,070	0.05
Egypt	8,000	18,833	0.42
Ghana	1,000	12,057	0.08
Morocco	4,500	97,079	0.05
Moldova	6,500	31,286	0.21
Nigeria	1,500	5,889	0.25
Pakistan	1,000	27,530	0.04
Philippines	5,000	21,805	0.23
Senegal	1,000	14,835	0.07
Somalia	100	159	0.63
Sri Lanka	3,500	21,966	0.16
Tunisia	4,000	16,010	0.25
Non-privileged, domestic work	65,000	136,576	0.48
Non-privileged, non-domestic work	60,400	120,676	0.50
Total	170,000	616,239	0.27

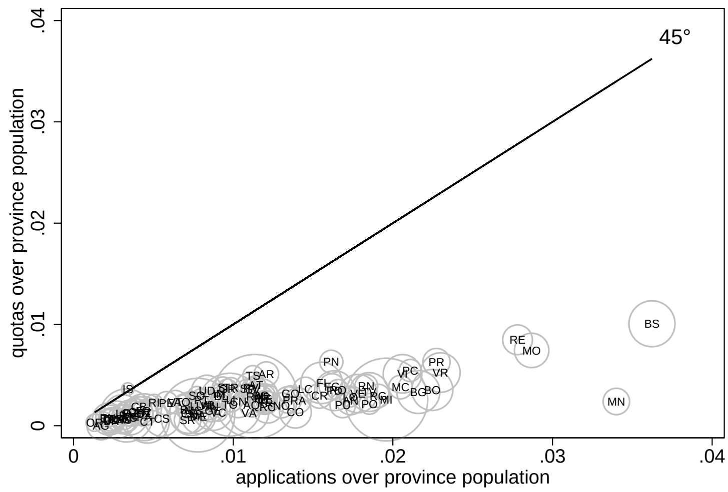
Note: This table reports, for each category of applicant, the quotas fixed by the Flows Decree 2007, the number of applications received, and the ratio of quotas to applications.

The total quota for each category is then partitioned across the 103 Italian provinces based on the expected labor demand for foreign workers communicated in each previous year by the Union of Italian Chambers of Commerce, Industry, and Agriculture ("Unioncamere"). Figure 1 plots the number of permits available in each province compared with the number of applications received. The markers for all but one small province lie below the 45-degree line, indicating that quotas

[5] Note that, contrary to the intentions of policy-makers, the rationing is stronger for the so-called "privileged" nationalities, for whom the average ratio of quotas to applications is as low as 13% (as opposed to 50% for non-privileged nationalities; see the last column of the table). Therefore, individuals who, in principle, should have been favored by the provision of reserved quotas (e.g., Albanians or Moroccans) confronted a much lower probability of obtaining a residence permit compared with immigrants of other nationalities (e.g., Chinese or Latin Americans). Beginning with the Flows Decrees 2009, however, the quotas for the latter group were always set at a low (or zero) level.

fall short of demand for permits almost everywhere, often by a large number. In Milan, for instance, the number of applicants was 10 times greater than the number of available quotas.

Figure 1: Quotas fixed by the Flows Decree 2007 and applications for residence permits received across Italian provinces



Note: This graph plots the quotas for residence permits allocated to each province according to the Flows Decree 2007 compared with the number of applications received, both divided by province population. The size of markers is proportional to the total province population; the 45-degree line is also included in the graph.

The quota system described above partitions applicants into a number of distinct groups: 1,442 groups of immigrants from privileged countries (103 provinces $\times$ 14 nationalities) and 309 groups of immigrants from other countries (103 provinces $\times$ 3 types of occupations), for a total of 1,751 groups. All applicants within each group compete for permits within the same quota. To ensure brevity and to emphasize the primary source of quasi-random variation at the heart of our identification strategy, we subsequently refer to such groups as "lotteries".

2.2 The application procedure

The applications for residence permits must be sponsored and materially sent by employers – Italian native citizens or legal foreign residents – who are willing to hire immigrant workers *before* they enter Italy. However, the matching of vacancies with foreign workers living abroad is extremely difficult (if not impossible). For this reason, most immigrants enter Italy illegally, begin working unofficially, and wait for their employer to sponsor them upon notice of the following Flows Decree.

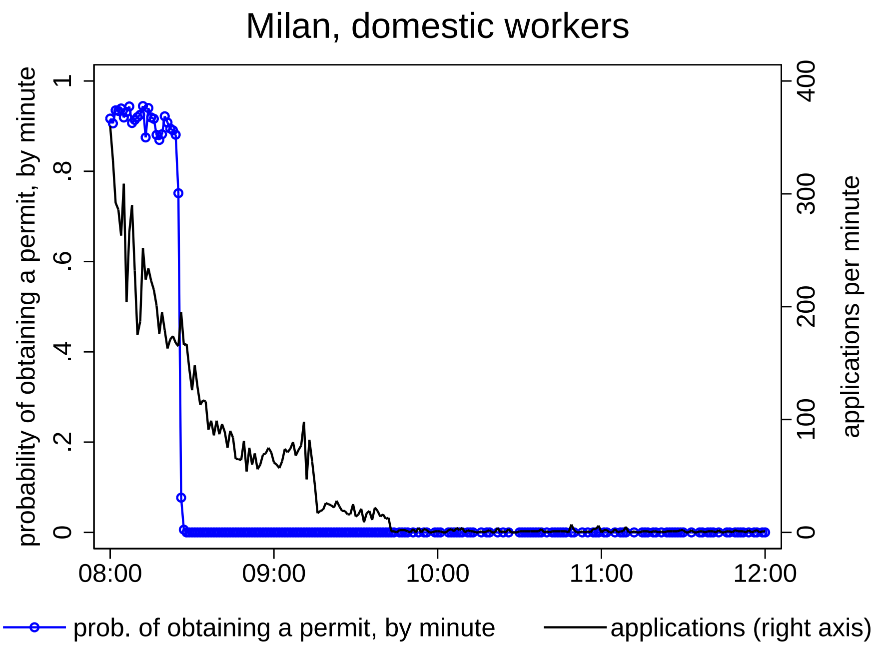
Although this system has been in place since 1998, the application procedure was completely digitized in 2007. Beginning in that year, employers have been required to apply through the internet during specific Click Days each year, and applications are processed on a first-come, first-served basis (within each lottery) based on the exact order in which they are received by the electronic system. The processing of each application by immigration officers involves cross-checking with police records and other administrative archives: if the application is accurate and complete and if the applicant has no criminal record, then (s)he receives the *nulla osta* for a residence permit; however, if some of the information is missing, inaccurate, or false or if the applicant has a criminal record, then the application is rejected. The process continues until the quota of permits available for that lottery is exhausted. In the presence of rationing, this type of mechanism generates a discontinuity in the probability of obtaining a residence permit as of the cutoff time at which the last successful application was sent.

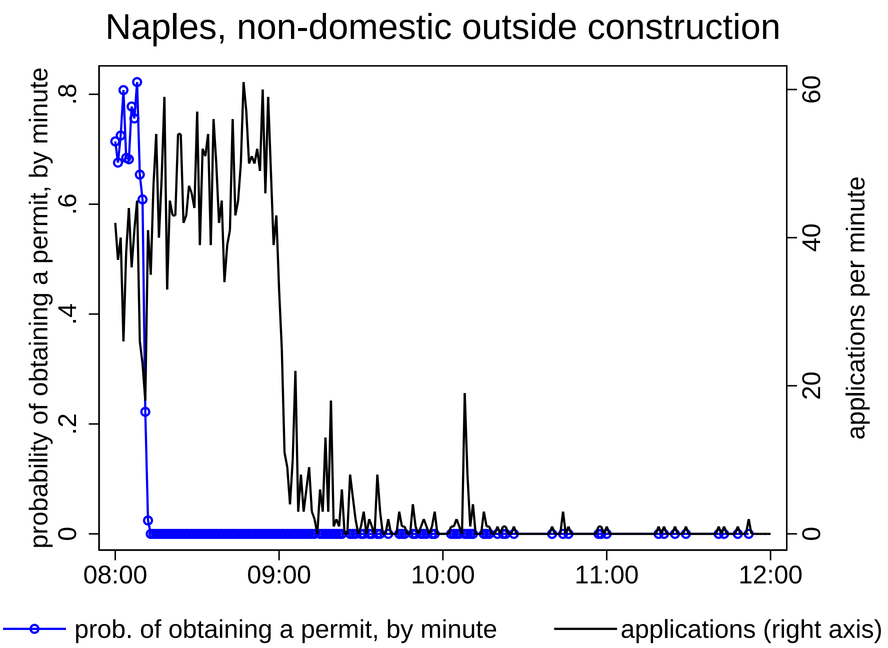
In 2007, the sponsors of immigrants from privileged nationalities could begin applying at 8:00 am on December 15, whereas Click Days for non-privileged nationalities were on December 18 and December 21 for those competing for domestic or non-domestic jobs, respectively. Most applications were received in the first few minutes of the day; by 10:00 am, nearly all quotas were already filled.

Figure 2 shows two examples. The top graph refers to the lottery for domestic workers of non-privileged nationalities in the province of Milan, the largest city in northern Italy, with a high immigrant presence (approximately 13% of the resident population in 2011). The black line shows the total number of applications received at each minute in time. Immediately after 8:00 am, the system received hundreds of applications per minute, and by 9:40 am, the flow had already slowed to only a few applications per minute. The (unconditional) probability of acceptance, also shown in the graph, is initially high (approximately 90%), as the only rejected applications were those deemed false or incomplete by immigration officers or those sponsoring immigrants with a criminal record in Italy. However, the probability of obtaining a residence permit suddenly decreases to zero when the entire quota assigned to this group is exhausted less than half an hour after the start of the lottery – to be precise, the last application accepted into the quota was sent at 8:27:04. The bottom graph shows a similar picture for the lottery for non-domestic workers (outside of construction) in the province of Naples, the most important labor market for immigrants in southern Italy. In this case, the cutoff was even earlier, at 8:10:56.

Applications received after the cutoff time are automatically rejected, and these

Figure 2: Timing of applications and probability of obtaining a residence permit for two lotteries in Milan and Naples





Note: These graphs show the total number of applications received (black line, right axis) and the fraction of those that were eventually accepted (blue line, left axis) at each minute in time between 8:00 and 12:00 for the case of two lotteries: non-privileged, domestic workers in Milan and non-privileged, non-domestic workers outside of the construction sector in Naples.

applicants have no right to reside and work in Italy. As already noted, however, the great majority of these individuals are already present in the country, and they remain there (illegally) even after being refused an official permit. Indeed, the rise in the number of official foreign residents in Italy over the last two decades (from 600,000 at the beginning of the 1990s to 4.5 million in 2011) parallels that of unofficial immigrants. Although no official estimates have been produced, the number of applications for amnesty presented during the frequent legalization episodes provides a lower bound for the size of the illegal immigrant population. The number of applications grew from 100,000 to 250,000 between 1986 and the 1990s, reaching a peak of 700,000 in 2002.

Unauthorized immigrants encounter lower employment prospects than legal immigrants as well as the risk of being expelled from the country. The following section provides a theoretical perspective on how these conditions could affect the number of crimes committed in the destination country.

3 Theoretical framework

This section presents a simple model of the relationship between legal status and crime. The primary purpose of the model is to clarify the policy effect of interest, the main threats to identifying this effect, and the quasi-experiment provided by Click Days.

3.1 Illegal immigrants

Consider a unit mass population of infinitely lived, risk-neutral immigrants who are *illegally* present in the destination country. With probability p, these immigrants are apprehended and expelled by the police at the beginning of each (discrete) period. Those who are not expelled work in the unofficial sector earning $w_0 x$, where x are heterogeneous labor skills distributed according to the cumulative density $G(.)$ and w_0 is the price of skills in the unofficial sector. Henceforth, the subscript 0 denotes (potential) outcomes when illegal.

In addition to working, immigrants may also engage in crime. During each period, immigrants receive a crime opportunity worth z, distributed according to the cumulative density $F(z)$. If the immigrants take this opportunity, then they immediately earn z but confront a probability q of being arrested. Assuming a discount factor $\beta < 1$ and normalizing to zero the utility of being arrested, the expected payoffs from seizing a crime opportunity are

$$K_0(x,z) = z + (1-q)\,\beta \int V_0(x,z')dF(z'), \tag{1}$$

where $V_0(x,z)$ denotes the individual utility conditional on having labor skills equal to x and receiving a crime opportunity z, such that $\int V_0(x,z')dF(z')$ is the expected utility in the case of not being arrested. In turn, the individual utility equals

$$V_0(x,z) = pV^H + (1-p)[w_0 x + \max\{K_0(x,z); \beta \int V_0(x,z')dF(z')\}] \tag{2}$$

where V^H is the utility associated with being expelled and returning to the home country.[6] If the immigrant is not expelled, (s)he earns $w_0 x$ by working in the unofficial economy and must decide whether to commit a crime (conditional on the realization of z). This decision involves a comparison of the return and expected (opportunity) cost of seizing a crime opportunity, respectively, $K_0(x,z)$ and $\beta \int V_0(x,z')dF(z')$ in equation (2). Because only the former depends (positively) on z, there exists a reservation value $\bar{z}_0(x)$ such that, conditional on x, one engages in crime only for $z > \bar{z}_0(x)$. Imposing $K_0(x,\bar{z}_0(x)) = \beta \int V_0(x,z')dF(z')$ in equation (2) and substituting into (1), one obtains the following:

$$\bar{z}_0(x) = \beta q \int V_0(x,z')dF(z'). \tag{3}$$

Intuitively, the payoff required for engaging in crime increases with both the risk and opportunity cost of being arrested. In particular, because $V_0(x,z)$ is increasing in x, $\bar{z}_0(x)$ is also increasing. Additionally, based on a simple revealed preference argument, $V^H < \beta \int V_0(x,z')dF(z')$ for all immigrants in the destination country, which, in turn, implies that $V_0(x,z)$ and $\bar{z}_0(x)$ are decreasing in p (see equation 2).

The reservation value in equation (3) completely characterizes criminal activity. In fact, the individual probability of committing a crime in the destination country equals the probability of not being expelled times the probability of drawing a crime opportunity above $\bar{z}_0(x)$,

$$C_0(x) = (1-p)\left[1 - F(\bar{z}_0(x))\right]. \tag{4}$$

[6]For simplicity, we assume that p and V^H are independent of x; however, all the identification results hold when we relax this assumption.

Intuitively, $C_0(x)$ decreases with working skills x, while the probability of expulsion p has two opposite effects: on the one hand, it reduces the pool of (illegal) immigrants at risk of committing crimes, as captured by the first term on the right-hand side of the equation; on the other, it lowers the opportunity cost of crime for those who are not deported, as captured by $\bar{z}_0(x)$.

3.2 Legalization

There is a continuum of employers interested in hiring immigrants in the official sector. The value of matching an employer with an immigrant worker equals the worker's labor market skills, x, and the shares accruing to the employer and the worker are π and $(1-\pi)$, respectively. If the application is successful, then the immigrant's wage thus changes from w_0 to $w_1 = 1-\pi$, with $w_1 \geq w_0$.[7] Additionally, successful applicants acquire the right to reside in the country, so the probability of expulsion changes from p to 0.

Thus, the individual utility when legal is

$$V_1(x,z) = w_1 x + \max\{K_1(x,z); \beta \int V_1(x,z')dF(z')\}, \tag{5}$$

where $K_1(x,z)$ is defined analogously to (1), and the probability of committing a crime in the destination country is

$$C_1(x) = 1 - F(\bar{z}_1(x)). \tag{6}$$

Note that, in contrast to $C_0(x)$, the probability of committing a crime when legal must not be conditioned on not being expelled from the country, as only the illegal immigrants can be expelled.

3.3 The legalization policy effect

In Rubin's (1974) terminology, $C_1(x)$ and $C_0(x)$ are the *potential outcomes* when legal and illegal, and their difference

$$\Delta C(x) \equiv C_1(x) - C_0(x) \tag{7}$$

is the causal effect of legal status on the probability of committing a crime. Because the individual utility increases with the wage rate and decreases with the

[7]The splitting rule and the restriction that w_1 is not lower than w_0 may be justified as the outcome of Nash bargaining between the employer and the employee.

probability of expulsion, $V_0(x,z) < V_1(x,z)$, and thus, $\bar{z}_0(x) < \bar{z}_1(x)\ \forall x$. Therefore, the probability of committing a crime *conditional on not being expelled* is always greater when illegal, $\tilde{C}_0(x) \equiv 1 - F(\bar{z}_0(x)) > C_1(x)$. On the other hand, illegals who are expelled no longer commit crimes in the destination country. Therefore, the sign of $\Delta C(x)$ depends on the relative strength of these opposite effects. Formally,

$$\Delta C(x) \approx p - [F(\bar{z}_1(x)) - F(\bar{z}_0(x))]\,, \tag{8}$$

where the first term on the right-hand side captures the increase in crime implied by the halting of expulsions, whereas the term in brackets captures the reduction in crime (resulting from the increased opportunity cost) conditional on not being deported.

Importantly, the sign of $\Delta C(x)$ can differ across the skill distribution, depending on the shape of $C_1(x)$ and $C_0(x)$. This variation is shown in the first diagram of Figure 3. While $C_1(x) < \tilde{C}_0(x)\ \forall x$, $\Delta C(x) < 0$ only for the high-skilled ($x = x''$), for whom access to the official labor market brings a significant improvement in earning opportunities. Instead, this type of increase is limited for low-skilled workers ($x = x'$), in which case the probability of committing a crime increases, $\Delta C(x') > 0$.

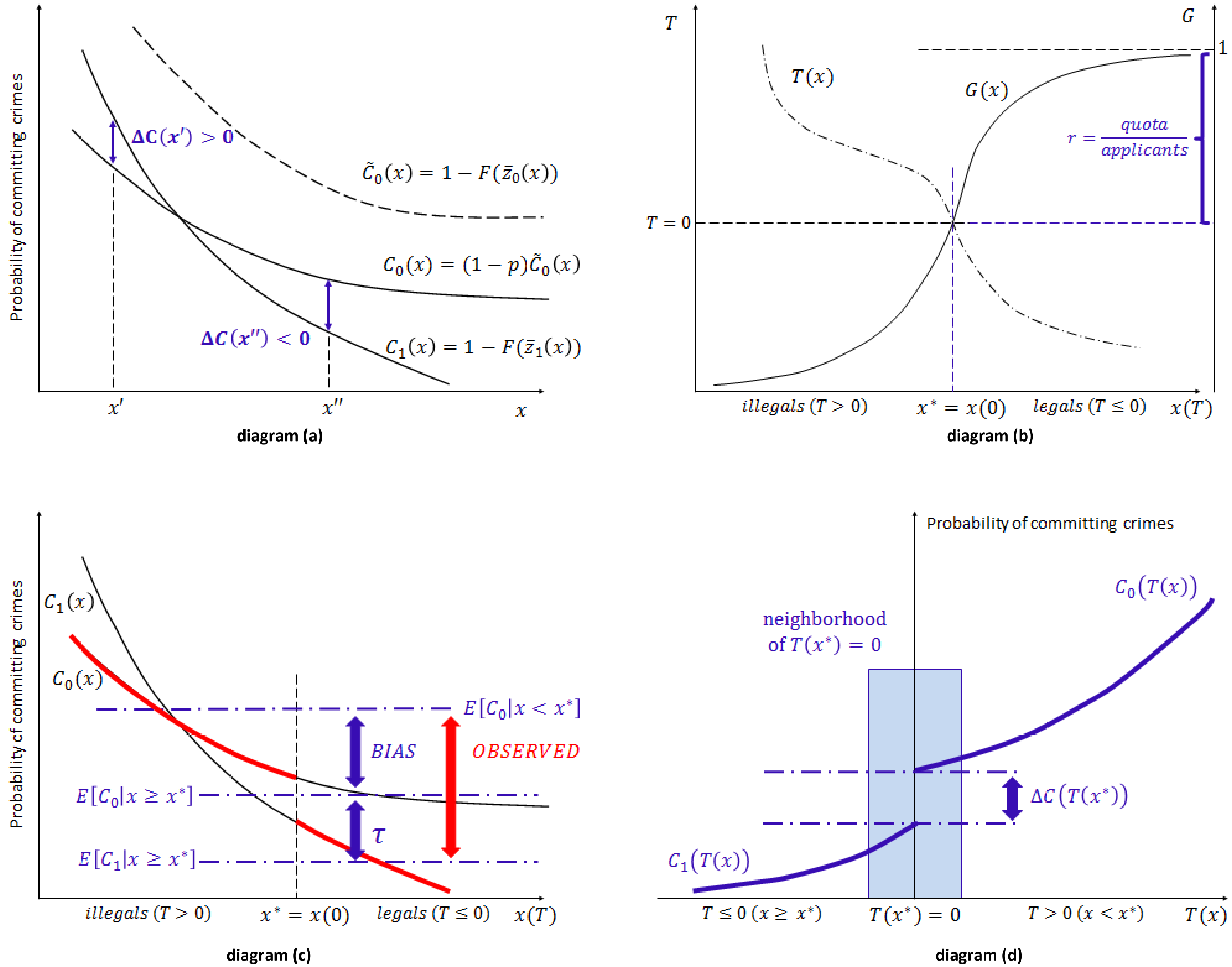
The effect of the legalization policy on the number of crimes per applicant committed in the destination country is

$$\tau \equiv E\left[\Delta C\left(x\right)|L=1\right], \tag{9}$$

where $L = 1$ for applicants who obtain legal status ($L = 0$ otherwise). The parameter τ depends on both effects in equation (8) – changes in criminal behavior conditional on not being expelled and changes in the probability of expulsion – but in the absence of individual-level data on expulsions, it is impossible to identify them separately.[8] From a policy perspective, however, the overall effect τ is of primary interest, as it reveals the sign and magnitude of the effect of legalization on the number of crimes per applicant committed in the destination country. Next, we discuss how to obtain consistent estimates of this parameter.

[8] Even if data on expulsions were available, immigrants' mobility across the border – in particular, differential mobility between legal and illegal immigrants – and other sources of attrition would prevent us from distinguishing between the two components of the effect.

Figure 3: Theoretical model and identification strategy



3.4 Selection bias

The primary threat to empirically identifying τ in (9) is that one never observes both potential outcomes for the same group of individuals but rather observes C_1 for $L = 1$ and C_0 for $L = 0$. The difference in average outcomes between the two groups differs from τ because of selection bias:

$$E(C_1|L=1)-E(C_0|L=0) = \underbrace{E(\Delta C|L=1)}_{\tau}+\underbrace{E(C_0|L=1)-E(C_0|L=0)}_{selection\ bias}, \quad (10)$$

where the bias is generally different from zero because of selection into legal status. For instance, individuals with better earning opportunities in the labor market (higher x) are likely to have a better likelihood of obtaining a job offer and thus obtaining legal status. As shown in Figure 3, such individuals would also have a lower probability of committing crimes, in which case the difference in crimes per applicant between legal and illegal immigrants would overestimate the policy effect of legal status.[9]

Click Days generate quasi-random variation in L across a particular subsample of applicants, thus eliminating selection bias. We illustrate this argument by modeling the allocation of residence permits.

3.5 The RD design

To hire immigrant workers, employers can sponsor them for a residence permit through the procedure described in Section 2: they submit applications beginning at a given time, and these applications are processed on a first-come, first-served basis until the available permits are exhausted.

Applications are costly (e.g., in terms of effort, organization, use of ICT equipment), and the cost increases with the speed of the application. Formally, this cost equals $\gamma(d)$, where $d \geq 0$ is the delay in submitting the application relative to the start of the Click Day and $\gamma(.)$ is a continuous, decreasing, convex function. As is observed in reality, there is uncertainty regarding the cutoff delay δ at which quotas are exhausted. Such uncertainty takes the form of a uniform distribution of common priors, $U(.)$, defined over some interval $[\underline{\delta}, \bar{\delta}]$.

[9]Relaxing the assumption of a common probability of expulsion and utility in the home country (i.e., $p(x) = p$ and $V_H(x) = V_H\ \forall x$), the bias could point in any direction depending on the shape of $C_1(x)$ and $C_0(x)$. However, all the identification results that follow carry over to this more general setting.

Based on these assumptions, the problem of the employer is

$$\max_{d} \left[1 - U(d)\right] \pi x - \gamma(d). \tag{11}$$

The solution of (11) implies that the delay of each application is a continuous function of the worker's skills, $d = d(x)$, with $d'(x) < 0$ (employers apply earlier for the most productive workers).

When applications are processed, the cutoff delay δ is revealed, and all applications sent before (after) the cutoff are accepted (rejected). With $T(x) = d(x) - \delta$ representing the timing of each application relative to the cutoff, $E(L|T \leq 0) = 1 - E(L|T > 0) = 1$.[10]

Because the function $T(.)$ decreases monotonically with x, it is invertible; thus, we may express an applicant's skills as a function of the timing of the application, $x = x(T)$. Then, there exists a cutoff level of skills, $x^* = x(0)$, such that all and only workers with skills above x^* obtain legal status (i.e., $L = 1$ if and only if $x > x^*$).[11]

The cutoffs δ and x^* ultimately depend on the rationing of permits. Letting r be the ratio of permits available to the total number of applications received from illegal immigrants, $1 - G(x^*) = r$, $d(x^*) = \delta$, and $T(x^*) = 0$; see diagram (b) in Figure 3. Then, diagram (c) shows how differences in observed outcomes between accepted and rejected applicants (i.e., applicants with skills above and below x^*, respectively) can be decomposed into the causal effect of legal status and selection bias.

Finally, the last diagram in Figure 3 plots the (expected) potential outcomes as a function of the timing of application, T (rather than of skills x, as in the other diagrams). Since C_0, C_1 and T are all inversely related to x, C_0 and C_1 will increase with T. The diagram shows that, as we restrict to an arbitrary small neighborhood of $T(x^*) = 0$, differences in outcomes between accepted and rejected applicants approximate the causal effect of legal status for the marginal applicant

[10] For the sake of simplicity, we abstract here from the rejections of inaccurate, false, or incomplete applications, i.e., we assume that the RD is "sharp". However, in the empirical analysis, we address the "fuzzy" nature of our RD design.

[11] Uncertainty surrounding the cutoff time guarantees a smooth distribution of the running variable. In fact, without uncertainty, the number of applications would change discontinuously to zero at the cutoff. The same result holds if the delay with which applications are received involves a stochastic component, which could reverse the order in which the applications were sent by the sponsors. Such elements are indeed present in reality, for instance, as a result of random delays in transmission through the internet, and these instances make an even stronger case for quasi-random assignment near the cutoff.

(i.e., the applicant with skills $x = x^*$); formally,

$$\lim_{t \to 0^-} E(C|T=t) - \lim_{t \to 0^+} E(C|T=t) = E\left[\Delta C(x)|T=0\right] = E\left[\Delta C(x)|x=x^*\right]. \quad (12)$$

The next section presents the methods and data available to empirically estimate equation (12).

4 Empirical strategy

If all and only the immigrants applying before the cutoff obtained legal status – as was the case in the simple model in Section 3 – then the difference in the limits in equation (12) would identify the causal effect of legal status for the marginal applicant. However, applications that are sent on time may still be rejected by immigration officers for being false, inaccurate, or incomplete (see Figure 2). In this case, commonly referred to as *fuzzy* RD design, the average effect of legal status on "compliers" equals the ratio of the difference in outcomes to the difference in the probability of obtaining legal status between early and late applicants,

$$\frac{\lim_{t \to 0^-} E(C|T=t) - \lim_{t \to 0^+} E(C|T=t)}{\lim_{t \to 0^-} E(L|T=t) - \lim_{t \to 0^+} E(L|T=t)}. \quad (13)$$

4.1 Estimation

The ratio (13) can be estimated using both parametric and nonparametric Wald-type estimators. In the former case, one can estimate the effect of L on C by two-stage least squares (2SLS) using $Z \equiv 1\{T \leq 0\}$ as an instrument for L, and including on the right-hand side a J-th order polynomial in T and its interaction with Z (Lee and Lemieux, 2010). Specifically, the reduced-form and first-stage estimating equations are

$$C = \theta_0 + \beta Z + \sum_{j=1}^{J} \theta_j T^j + \sum_{j=1}^{J} \vartheta_j T^j \cdot Z + \varepsilon \quad (14)$$

$$L = \pi_0 + \alpha Z + \sum_{j=1}^{J} \pi_j T^j + \sum_{j=1}^{J} \varpi_j T^j \cdot Z + \nu, \quad (15)$$

where ε and ν are error terms summarizing the effect of other omitted factors. After controlling through the polynomial in T for any smooth trend of C to the left and to the right of the cutoff – respectively, the curves C_1 and C_0 in diagram

(d) of Figure 3 – the coefficient β estimates the difference in potential outcomes at the cutoff. We will explore the sensitivity of the results to alternative specifications of the polynomial regression.

Alternatively, nonparametric estimators restrict the sample to applicants within an arbitrarily small bandwidth of the cutoff $T(x^*) = 0$, applying kernel weights and conditioning on the running variable to eliminate residual differences between accepted and rejected applicants. Following Hahn et al. (2001), the reduced-form and first-stage local kernel regressions are

$$\min_{\beta,\gamma'} \quad \textstyle\sum_i K\left(\frac{T_i}{h}\right) \quad \left(Y_i - \beta Z - \gamma_0 - \gamma_1 T_i - \gamma_2 T_i \cdot Z_i\right)^2 \tag{16}$$

$$\min_{\alpha,\delta'} \quad \textstyle\sum_i K\left(\frac{T_i}{h}\right) \quad \left(L_i - \alpha Z - \delta_0 - \delta_1 T_i - \delta_2 T_i \cdot Z_i\right)^2, \tag{17}$$

where $i = 1, 2, ..., N$ denote the individuals in the sample and $K\left(\frac{T_i}{h}\right)$ is a triangular kernel attaching positive weights only to observations within a bandwidth h around the cutoff.

In this case, the properties of the estimator depend crucially on the choice of h. Larger bandwidths increase asymptotic bias as a result of extrapolation away from the cutoff, whereas smaller bandwidths reduce bias at the cost of greater asymptotic variance. Along this bias-variance tradeoff, Imbens and Kalyanaraman (2012, IK2012 henceforth) choose the bandwidth that minimizes the expected squared error loss. Cattaneo et al. (2014, CCT2014 henceforth) argue that this criterion typically yields bandwidths that are "too large" for the distributional approximations invoked and that the RD estimate would be biased as a consequence. They therefore amend the IK2012 criterion in two ways: re-centering the estimated coefficient and standard errors to correct for the asymptotic bias and choosing a smaller bandwidth. We will examine the sensitivity of the results using both the IK2012 and CCT2014 criteria as well as a battery of different (heuristic) bandwidths.

4.2 Data

To implement the empirical strategy described above, we merge the applications presented on Click Days in 2007 with the criminal records of all offenders reported by the police in the following year.

Specifically, the Italian Ministry of Interior provided us with the administrative records of all applications processed by immigration officers. We obtained records on 403,741 applications, of a total of approximately 700,000 (because the quotas of available permits totaled 170,000, the last 300,000 applications were not even considered). To avoid extrapolation bias in the global polynomial regressions, we restrict the sample to applications received within a one-hour symmetric time window around the cutoff, which leaves us with a total of 186,608 observations.[12]

Each record includes the nationality of the applicant and the province in which it was presented, the type of employment for which the applicant was sponsored (domestic work, construction, all other sectors), the timing of the application by the millisecond, and the outcome (i.e., whether the applicant eventually obtained a residence permit).

Columns (1)-(2) of Table 2 report the sample averages for the individual characteristics available in our data, namely, age and country of origin, among all male and female applicants. There is a prevalence of males over females, and the average age for the two groups is 34 and 39 years, respectively. The majority of applicants come from Asian countries, followed by European countries outside of the EU (because of free mobility within the EU, citizens of EU member countries do not need a permit to reside and work in Italy). Consistent with the finding that immigration to Italy is largely an economic phenomenon, slightly fewer than two-thirds of applicants reside in Northern Italy, which is characterized by better economic opportunities for foreign workers.

The data on applications were matched, at the individual level, with the criminal records of all offenders reported by the police for having committed (at least) one serious crime: robbery, theft, drug trafficking, smuggling, extortion, kidnapping, murder, violent assault, and rape. We limit ourselves to crimes committed in 2008 because the applicants that did not fall within the quotas for 2007 were reconsidered one year later by the Flows Decree 2008.[13]

The final data are potentially subject to measurement error from two primary sources. First, reported crimes always underestimate the true number of committed crimes; see MacDonald (2002) for a thorough discussion. Second, errors can arise in the matching between Click Day applications and criminal records. In

[12]The results for the entire sample (i.e., also including applications far away from the cutoff) are reported in the Web Appendix and discussed briefly in the next section. The estimated coefficients are qualitatively similar to the baseline estimates, and the standard errors are much lower because of the increased sample size.

[13]To clear the backlog accumulated in the previous year, no new applications were allowed in 2008. Rather, the first 150,000 excluded in 2007 were re-processed upon renewal of the applications (to be sent between December 15, 2008, and January 3, 2009). Therefore, applicants to the right of the cutoff, who constitute our control group, also had the possibility of obtaining legal status in 2009.

Table 2: Descriptive statistics

	(1)	(2)	(3)	(4)
	ALL LOTTERIES		SIGNIFICANT BREAK	
	males	*females*	*males*	*females*
age	33.860	39.178	33.866	39.203
	(0.023)	(0.040)	(0.024)	(0.040)
Africa	0.172	0.062	0.166	0.059
	(0.001)	(0.001)	(0.001)	(0.001)
Americas	0.085	0.188	0.087	0.189
	(0.001)	(0.001)	(0.001)	(0.001)
Asia	0.541	0.339	0.545	0.340
	(0.001)	(0.002)	(0.001)	(0.002)
Europe (outside of the EU)	0.201	0.411	0.202	0.412
	(0.001)	(0.002)	(0.001)	(0.002)
Northern Italy	0.643	0.615	0.648	0.618
	(0.001)	(0.002)	(0.001)	(0.002)
Central Italy	0.212	0.208	0.210	0.206
	(0.001)	(0.001)	(0.001)	(0.001)
Southern Italy	0.145	0.177	0.142	0.176
	(0.001)	(0.001)	(0.001)	(0.001)
REPORTED CRIMES IN 2008, PER 100 APPLICANTS				
all types of crime	1.418	0.080	1.427	0.081
	(0.048)	(0.017)	(0.049)	(0.017)
theft	0.253	0.035	0.255	0.036
	(0.017)	(0.008)	(0.018)	(0.008)
robbery	0.208	0.003	0.210	0.003
	(0.016)	(0.002)	(0.016)	(0.002)
drug trafficking	0.162	0.022	0.161	0.022
	(0.015)	(0.014)	(0.015)	(0.014)
smuggling	0.039	0.007	0.040	0.007
	(0.009)	(0.003)	(0.009)	(0.003)
kidnapping	0.028	0.000	0.026	0.000
	(0.005)	.	(0.005)	.
extortion	0.060	0.009	0.062	0.010
	(0.007)	(0.004)	(0.008)	(0.004)
illegal carrying of firearms	0.161	0.001	0.162	0.001
	(0.013)	(0.001)	(0.013)	(0.001)
murder	0.058	0.000	0.058	0.000
	(0.009)	.	(0.009)	.
violent assault	0.290	0.003	0.293	0.003
	(0.016)	(0.002)	(0.016)	(0.002)
rape	0.160	0.000	0.160	0.000
	(0.017)	.	(0.017)	.
observations	112,772	73,836	110,338	73,097

Note: This table shows the average characteristics (top panel), the number of reported crimes per 100 applicants (bottom panel), and the total number of observations (last row) for the applicants to all lotteries (columns 1-2) and to lotteries exhibiting a significant structural break in the probability of treatment assignment at the cutoff (columns 3-4). Separate figures by gender are presented, and standard errors are reported in parentheses.

particular, because the two archives were matched by name, surname, nationality, and year of birth, any mistake in the coding of these variables in either of the two archives (or both) could generate Type I errors, namely, whenever an applicant has been reported for having committed a crime but the algorithm fails to detect the match, or Type II errors, namely, whenever an applicant was never reported but the algorithm still finds a match.

Let $Y = 1$ for immigrants reported by the police ($Y = 0$ otherwise). If the probability of being reported conditional on having committed a crime is the same for legal and illegal immigrants, $E(Y|C = 1) = \mu \leq 1$, then differences in reported crimes would have the same sign and would provide a lower bound for the magnitude of differences in actual crime rates between the two groups:

$$E(Y|L = 1) - E(Y|L = 0) = \mu \left[E(C|L = 1) - E(C|L = 0)\right].$$

This is the case when the extent of under-reporting is symmetric between legal and illegal immigrants. The use of police charges (as opposed to incarceration rates) eliminates distortions resulting from the differing treatment of suspected offenders by the judicial system (e.g., only legal immigrants can ask for measures alternative to incarceration, such as home detention). Moreover, violations of migration law do not constitute a serious crime; thus, differences in observed crime rates between legal and illegal immigrants do not depend on the fact that the latter can be reported for being illegally present in the country.

Turning to the match between applications and criminal records, errors that occur at random would also bias the estimated effect of legal status toward zero. However, individuals who are arrested by the police may *intentionally* misreport their identity to maintain a clean criminal record for their true identity or to avoid the application of aggravating circumstances (if they are recidivists). Clearly, illegal immigrants have a better probability of hiding their true identity, as they do not carry official documents. Indeed, this phenomenon is well known in Italy. For instance, Barbagli (2008) reports the case of an illegal immigrant who was arrested in Bologna and who subsequently confessed to have used 15 aliases in previous years (see Table A1 in the Web Appendix).

For all these reasons, the matched data set would overestimate the crime rate of legal immigrants relative to that of illegal immigrants, $E(Y|C = 1, L = 1) > E(Y|C = 1, L = 0)$, thus biasing the estimates toward finding a positive effect of legal status on crime:

$$\begin{aligned} E(Y|L=1) - E(Y|L=0) &= E(Y|C, L=1)E(C|L=1) - E(Y|C, L=0)E(C|L=0) \\ &> E(C|L=1) - E(C|L=0) \end{aligned}$$

In summary, we can interpret a negative estimated effect of legal status as a lower bound (in terms of magnitude) of the true effect, whereas a positive estimate could reflect either measurement error or a positive effect of legal status or both. Therefore, our empirical strategy only allows us to reject the null hypothesis of no effect of legal status against the alternative of a negative effect, but not of a positive effect.

The bottom panel of Table 2 shows the number of (reported) crimes per 100 applicants for all individuals in our sample. As is typically the case, men commit more crimes than women (the number of recorded crimes for females are actually close to zero). With regard to the type of crimes, immigrants were primarily reported for property crimes (thefts and robberies) and violent assaults.

4.3 Implementation

The information reported on each application (specifically, the country of origin, the province of destination, and the type of job) allows us to identify the lottery in which each applicant participated. In particular, Click Days in 2007 generated 1,580 out of 1,751 potential lotteries (see Section 2.1) because 68 lotteries received no applications and because the 103 provincial lotteries for Sri Lankan immigrants were corrupted by a computer bug.

The largest lotteries were those for domestic workers of non-privileged nationalities in Rome and Milan, with 9,974 and 12,207 (processed) applications, respectively; meanwhile, dozens of lotteries for privileged nationalities in smaller provinces received only one or a few applications. The mean and median number of applications processed across all lotteries are 246 and 57, respectively, and there were 311 lotteries receiving fewer than 10 applications. Following common practice in RD studies (see, for instance, Fredriksson et al., 2013 and Dahl et al., 2014), we pool the observations from all lotteries. We then use the delay of each application relative to the cutoff of its lottery as a common running variable and include lottery fixed effects to prevent heterogeneity in the baseline crime rate across lotteries (the intercept of the regression) from biasing the estimate of the differential effect of legal status (the slope parameter).[14] We will also perform a permutation exercise to exclude the possibility that our estimates average a small sample bias across lotteries.

To compute the running variable, we must define the cutoff point $t = 0$ in each lottery. In contrast to most RD designs, in which treatment assignment depends on a predetermined threshold rule, the cutoff time in this case depends on the realized (ex-post) timing of all applications and the realized acceptance rate. This feature reinforces the fundamental identifying assumption that applications are as good as randomly assigned near the cutoff; however, it raises the issue of how to define this cutoff.

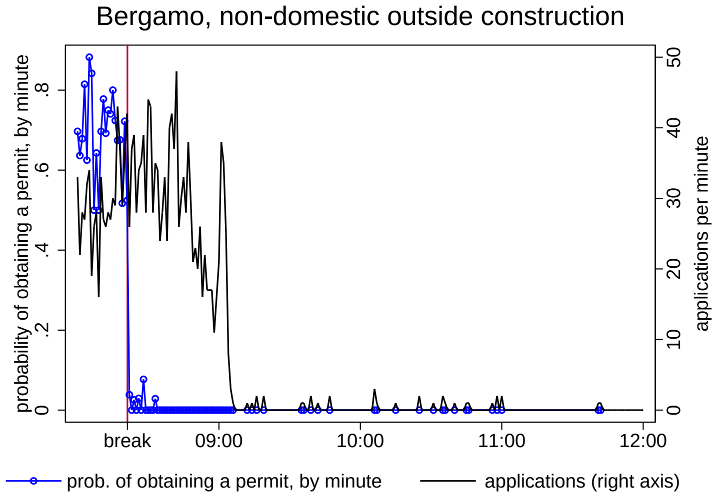
When the probability of obtaining a residence permit immediately drops to zero, as in Figure 2, the cutoff would coincide with the timing of the last accepted application. This situation actually occurs for most lotteries. In a few cases, however, the acceptance rate falls to a low, but still positive, level for several minutes. The lottery for non-construction workers in the province of Bergamo (a city near Milan) provides an example. The probability of acceptance declines markedly, from 70% to 5%, at approximately 8:22 am, but it reaches zero only a few minutes later; see Figure 4. This situation occurs because when an application has been rejected because of missing, inaccurate, or false information or when the applicant did not collect the permit that had been authorized, the ordering in which the next applicants were re-contacted sometimes subverted the initial ordering.

As long as the subversion results from idiosyncratic implementation errors, it would not cause any systematic bias in the comparison of immigrants obtaining or not obtaining a residence permit (in fact, it could make an even stronger case for random treatment assignment). Moreover, even non-random subversions – for instance, because of voluntarily manipulation by the immigration officials – can be easily accommodated in the fuzzy RD design. Indeed, the Wald-type estimators in (14)-(15) and (16)-(17) exploit variation in legal status that depends only on the initial ordering of applications near the cutoff (as opposed to later subversions of such ordering).

The only complication that arises when the discontinuity is fuzzy on both sides is that the timing of the last accepted application may no longer be an adequate measure of the cutoff, as the probability of obtaining legal status could have already changed several minutes earlier (as in Figure 4). Confronted with the

[14] This particular form of omitted variable bias is sometimes called the "Yule-Simpson Paradox," see, e.g., Chen et al. (2009).

Figure 4: Timing of applications and probability of obtaining a residence permit for a lottery in the city of Bergamo



Note: This graph shows the total number of applications received (black line, right axis) and the fraction of those that were eventually accepted (blue line, left axis) at each minute between 8:00 and 12:00 for the lottery of non-privileged, non-construction workers in the city of Bergamo. The vertical line shows the timing of the structural breakpoint, as estimated by the Andrews (1993) test.

same problem (i.e., estimating an unknown cutoff point in a fuzzy RD design), Chay et al. (2005) and Bertrand et al. (2010) run a battery of regressions of a dummy for treatment assignment on another dummy that equals 1 after each possible cutoff point and choose the one that maximizes the R^2 of the regression. Following the same idea, we conduct a formal Andrews (1993) test for the existence (in each lottery) of a structural break with unknown breakpoint on the probability of success at each second in time. The test identifies the "most likely break point" and allows us to test for the significance of the structural break. Such a test is important in our context because in some lotteries the permits were not rationed and the only immigrants who were eventually excluded were those whose applications were rejected.

In the example in Figure 4, the estimated structural break point is represented by the vertical line. In general, for the great majority of lotteries, the test rejects the null hypothesis that there is no structural break. Indeed, these lotteries constitute 98% of the applicants in our sample; the median cutoff time is 08:39:06, and the majority of quotas were exhausted well before 9:00. Columns (3)-(4) in Table 2 show that the average individual characteristics across applicants in this subsample, which will be used for estimation, are similar to those for the original sample.[15]

5 Results

The RD approach allows for a graphical representation of treatment effects that is both intuitive and transparent. We thus begin with this analysis and then move to parametric and nonparametric regression methods.

5.1 Graphical analysis and parametric estimates

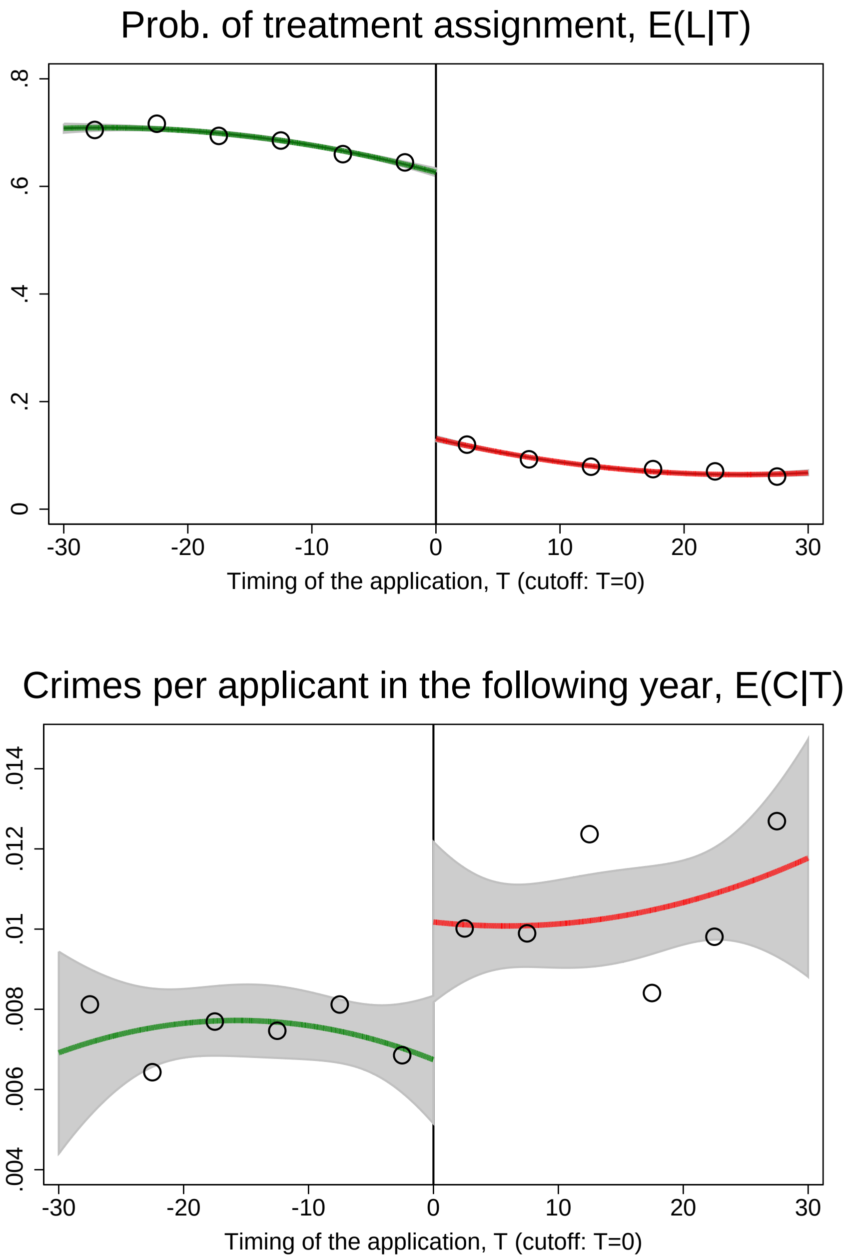
Figure 5 plots the average probability of obtaining legal status and the average number of crimes per applicant committed in Italy over the following year, conditional on the timing of the application T. The circles are average probabilities across 5-minutes bins of T to the left and right of the cutoff, whereas the solid lines and shaded areas are predicted outcomes and associated confidence intervals, respectively, based on quadratic polynomial regressions, i.e., regressions (14) and (15), with $J = 2$. This specification was validated against more flexible alternatives ($J > 2$) based on the Akaike information criterion (Lee and Lemieux, 2010).

The left graph in Figure 5 shows that the average probability of obtaining legal status decreases from 63% to 13% at $T = 0$. The right graph is the empirical counterpart of diagram (d) in Figure 3, and it shows that the drop in the probability of obtaining legal status coincides with a significant increase in the number of crimes per applicant committed in the following year.

To quantify the magnitude of such effects, the first and second rows of Table 3 report the estimated coefficients β and α in equations (14) and (15), respectively. In the baseline specification in column (1), the reduced-form coefficient equals -0.3 crimes per 100 applicants, while the first stage amounts to a 50 percentage point increase in the probability of obtaining legal status. Taking the ratio of the two coefficients (the third row of the table), the estimated causal effect of legal status is -0.7 crimes per 100 applicants, which is significant at the 95% confidence level. Distinguishing applicants by gender, the effect doubles for males, while it is not different from zero for females (columns 2 and 3, respectively), which is consistent with the finding that females exhibit crime rates that are close to zero.

[15] Table A2 in the Web Appendix provides the same comparison for the full sample (i.e., also including applicants more distant from the cutoff).

Figure 5: Probability of obtaining legal status and number of crimes per applicant, conditional on the timing of application



Note: The figure shows the probability of obtaining legal status (top graph) and the average number of crimes per applicant committed over the following year (bottom graph) conditional on the timing of application. The scatterplots are averages within 5-minute bins, and the solid lines and shaded areas are the predicted outcomes and associated confidence intervals, respectively, based on a quadratic polynomial regression.

Table 3: The effect of legal status on the number of crimes per applicant, global polynomial regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	BASELINE SPECIFICATION			DEMOGRAPHIC CONTROLS			LOTTERY FEs, CLUSTERED S.E.		
	all	*males*	*females*	*all*	*males*	*females*	*all*	*males*	*females*
Reduced form	-0.003**	-0.006**	0.001	-0.003**	-0.006**	0.001	-0.003*	-0.006**	0.001
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.001)
First stage	0.495***	0.459***	0.557***	0.495***	0.459***	0.556***	0.498***	0.463***	0.561***
	(0.005)	(0.007)	(0.008)	(0.005)	(0.007)	(0.008)	(0.028)	(0.029)	(0.037)
2SLS estimate	-0.007**	-0.013**	0.001	-0.007**	-0.013**	0.002	-0.007*	-0.013**	0.002
	(0.003)	(0.005)	(0.001)	(0.003)	(0.005)	(0.001)	(0.004)	(0.007)	(0.001)
Observations	183,432	110,337	73,095	183,431	110,337	73,094	183,431	110,337	73,094
First stage F-stat	8738.1	4303.3	4820.7	8722.9	4304.1	4820.9	306.2	248.73	232.77

Note: This table reports 2SLS estimates for the effect of legal status on the number of crimes committed in Italy by Click Day applicants. The dependent variable is a dummy $C = 1$ for individuals committing at least 1 serious offense in 2008, the explanatory variable of interest is a dummy $L = 1$ for those obtaining a residence permit, and the first-stage instrument is a dummy $Z = 1$ for having applied before the cutoff time. All regressions control for a quadratic polynomial in the time elapsed since the cutoff (by the millisecond) and its interaction with Z, and the specifications in columns (7)-(9) also include lottery fixed effects. The first row reports the reduced-form regression of C on Z, the second row reports the first stage regression of L and Z, and the third row reports the 2SLS-estimated effect of L on C. The last row of the table reports the first-stage F statistic for the excluded instrument, Z. Robust standard errors are reported in parentheses, and they are clustered by lottery in columns (7)-(9). *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

In columns (4) to (6), we control for the individual characteristics reported in our data, namely, gender and age (both linear and squared), whereas in columns (7) to (9), we include lottery fixed effects and cluster standard errors by lottery to account for interactions in crime among groups of immigrants who apply for the same lottery. Controlling for lottery fixed effects is particularly important to exclude the possibility that our results are driven by aggregation bias across lotteries (see Section 4.3). This possibility does not appear to be a concern, as the coefficient of legal status remains identical to the baseline estimate in column (1).

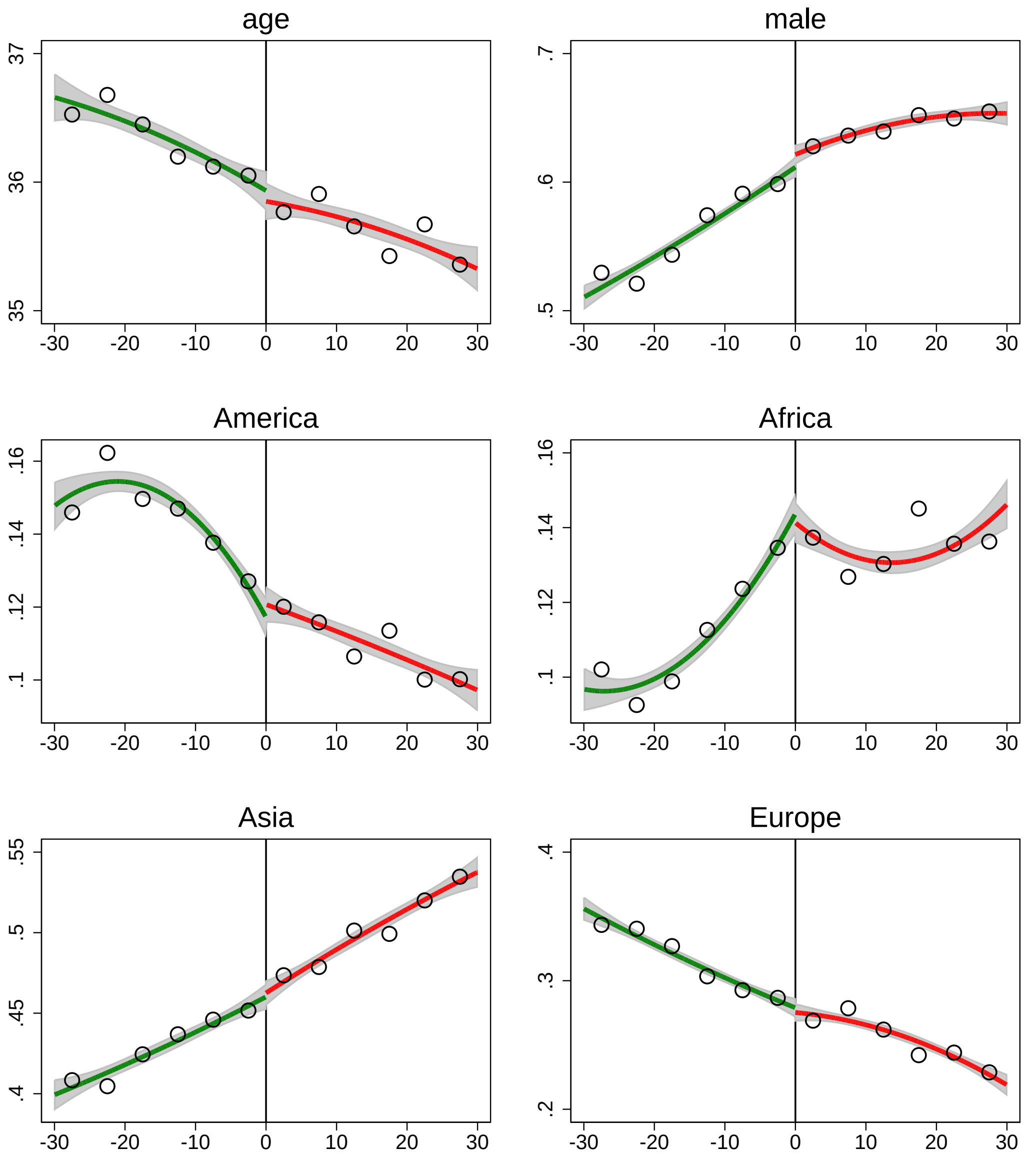
More generally, the stability of the coefficients across all different specifications in Table 3 appear to exclude the possibility that differences in criminal behavior between legal and illegal immigrants depend on compositional effects. This finding is further confirmed in Figure 6, which plots the relationship between application times and average individual characteristics. Although there are clear trends in T (for instance, females and younger workers applied first), there is no significant discontinuity at the cutoff.

Of course, early and late applicants could differ along other dimensions. Unfortunately, the administrative data on applications do not report any information on important determinants of criminal activity (e.g., income and educational levels), and in any case, many other relevant factors would be difficult to measure (e.g., attitudes toward illegal activities and risk aversion). Nevertheless, the balance in observable characteristics is consistent with the assumption that legal status is as good as randomly assigned near the cutoff.

The empirical density of T, presented in Figure 7, provides additional evidence in this respect. If immigrants near the cutoff were able to select on either side of it and they did so in a monotonic way (i.e., all individuals manipulate the running variable in the same direction), then the density of T would exhibit a discontinuity at $T = 0$. McCrary (2008) provides a formal test based on the (log) height of the distribution at the cutoff. This test does not reject the null hypothesis of no discontinuity in the distribution of T (the log-difference in the height of the distribution at the cutoff equals 0.022, with a standard error of 0.020).

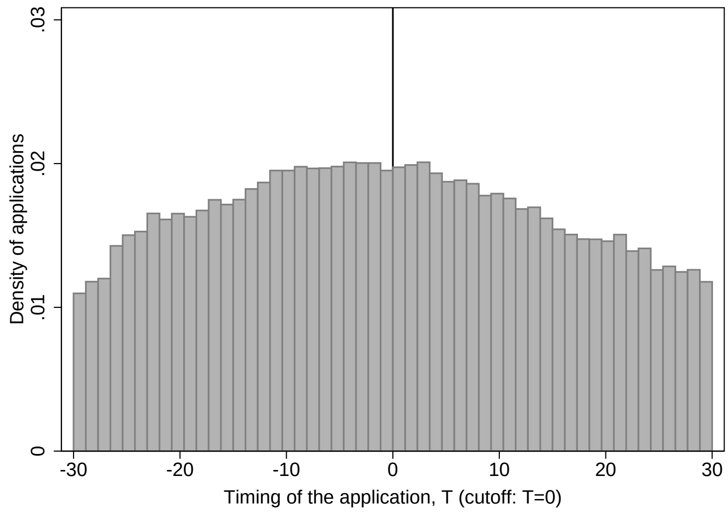
This evidence suggests that neither manipulation of the running variable nor differences between the average characteristics of individuals to the left and right of the cutoff can explain the discontinuity in the probability of committing crimes at $T = 0$.

Figure 6: Individual characteristics, conditional on the timing of application



Note: These graphs show the average individual characteristics of Click Day applicants, conditional on the timing of application. The scatterplots are averages within 5-minute bins, and the solid lines and shaded areas are the predicted values and associated confidence intervals, respectively, based on a quadratic polynomial regression.

Figure 7: Frequency of applications near the cutoff



Note: The graph shows the density of applications received at each minute in time, before and after the cutoff.

5.2 Nonparametric estimates

Table 4 shows the main results for the kernel local linear regressions on equations (16) and (17).[16] The criteria used to select the bandwidth and its size are reported at the top of each column.

The IK2012 criterion selects a bandwidth of 20:44 minutes, which includes approximately 83 thousand observations. The estimated reduced-form effect is -0.45 crimes per 100 applicants, which is statistically significant at the 95% confidence level. Dividing this estimate by the first-stage coefficient (-0.46) yields a 2SLS estimate of -1.0 crimes per 100 applicants (column 1).

As discussed in Section 4.1, the optimal bandwidth under a squared error loss function may be "too large," inducing a non-negligible bias in the estimated coefficients. The specification in column (2) employs the bias-correction procedure by CCT2014, re-centering the coefficient estimate and adjusting the standard errors to account for bias estimation. When this procedure is used, both the reduced form and the 2SLS coefficient increase threefold, which may be a symptom of extrapolation bias away from the cutoff.

In column (3), we thus select the bandwidth according to the CCT2014 criterion. The new bandwidth is much smaller, only 6:29 minutes, and comprises slightly fewer than 30 thousand observations. Within this narrow neighborhood

[16]In view of the previous results in Table 3, we focus on the subsample of males.

Table 4: The effect of legal status on the number of crimes per applicant, kernel local linear regression

	(1)	(2)	(3)	(4)
BANDWIDTH SELECTION:	IK2012 [20:44]		CCT2014 [06:29]	
BIAS CORRECTION:	NO	YES	NO	YES
Reduced form	-0.004**	-0.014**	-0.007**	-0.008**
	(0.002)	(0.006)	(0.003)	(0.004)
First stage	0.463***	0.456***	0.443***	0.441***
	(0.006)	(0.020)	(0.011)	(0.012)
2SLS estimate	-0.010**	-0.030**	-0.015**	-0.017**
	(0.005)	(0.013)	(0.007)	(0.008)
Total observations	110,338		110,338	
obs. inside the BW	83,394		29,219	

Note: This table reports nonparametric estimates of the effect of legal status on the number of crimes committed in Italy by Click Day applicants. The dependent variable is a dummy $C = 1$ for individuals committing at least 1 serious offense in 2008, the explanatory variable of interest is a dummy $L = 1$ for those obtaining a residence permit, and the first-stage instrument is a dummy $Z = 1$ for having applied before the cutoff time. All coefficients are estimated using a kernel local linear regression in a symmetric bandwidth around the cutoff. The table reports the bandwidth selection rule, IK2012 or CCT2014, at the top of each column; the size of the bandwidth in minutes and seconds [mm:ss]; and whether the re-centering procedure by CCT2014 is employed. The first row reports the reduced-form regression of C on Z, the second row reports the first-stage regression of L on Z, and the third row reports the estimated effect of L on C obtained as the ratio between the reduced form and the first stage. Robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

Table 5: Standardized difference in observable characteristics between applicants to the left and right of the cutoff (inside the bandwidth)

	IK2012 [20:44]			CCT2014 [06:29]		
	$E(X\|T \leq 0)$	$E(X\|T > 0)$	SMD	$E(X\|T \leq 0)$	$E(X\|T > 0)$	SMD
age	33.898	33.859	0.005	33.924	33.963	-0.005
	(0.039)	(0.038)		(0.067)	(0.065)	
America	0.096	0.079	0.058	0.090	0.083	0.023
	(0.001)	(0.001)		(0.002)	(0.002)	
Africa	0.162	0.171	-0.023	0.175	0.175	-0.001
	(0.002)	(0.002)		(0.003)	(0.003)	
Asia	0.530	0.556	-0.052	0.530	0.543	-0.027
	(0.002)	(0.002)		(0.004)	(0.004)	
Europe	0.212	0.193	0.045	0.205	0.198	0.018
	(0.002)	(0.002)		(0.003)	(0.003)	

Note: This table compares the average characteristics of applicants within a given bandwidth (indicated at the top of the table) to the left and the right of the cutoff (columns $E(X|T \leq 0)$ and $E(X|T > 0)$, respectively). The table also shows the standardized mean difference (SMD) between the two groups. The SMD is considered to be "large" when it exceeds 0.20 (Rosenbaum and Rubin, 1983).

of the cutoff, the 2SLS estimated coefficient of legal status is -1.5 crimes per 100 applicants, and it remains stable when we employ the bias-correction procedure (column 4).

Importantly, within all subsamples and bandwidths, applicants to the left and right of the cutoff are balanced in terms of observable characteristics, as the standardized difference between the two groups is always small (well below the critical value of 0.20 suggested by Rosenbaum and Rubin, 1983); see Table 5.

5.3 Robustness

We next examine the sensitivity of the baseline estimates to alternative specifications of the parametric and nonparametric regressions, perform a falsification test to detect the presence of any systematic bias in our estimates, and further discuss the issue of measurement error.

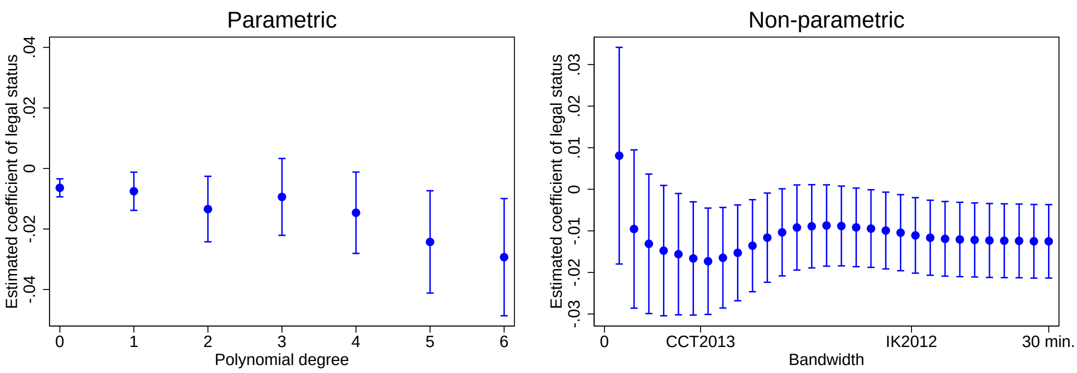
Beginning with the first issue, we note that the consistency of the RD estimates relies on the assumption that controlling parametrically for a flexible polynomial in T or restricting the sample nonparametrically to observations in a neighborhood of the cutoff removes the effect of other factors that vary (smoothly) around the cutoff. Both methods involve a trade-off between bias and efficiency: in general, higher-order polynomials and smaller bandwidths reduce the bias at the cost of greater asymptotic variance.

To examine the sensitivity of our results to different choices among these trade-offs, the graphs in Figure 8 plot the estimated coefficient of legal status and the associated confidence intervals for different specifications of parametric and nonparametric regressions. In particular, the left graph plots the results of the global polynomial regression when varying the order of the polynomial, J, between 0 and 6. The coefficient of legal status is always negative and is generally statistically significant (it is marginally non-significant at conventional confidence levels for $J = 3$).

As a further robustness check, we also run parametric regressions on the entire sample of Click Day applicants, including also those who applied outside of the one-hour symmetric time window around the cutoff. These estimates are potentially subject to a severe extrapolation bias, and simply controlling for the polynomial in T may not be sufficient to remove any heterogeneity between applicants who are far away from the cutoff on either side. Nevertheless, the estimated coefficients remain qualitatively similar to those in Table 3 and stronger in terms of statistical significance.[17]

Turning to the nonparametric estimates, the right graph in Figure 8 shows the results of the local kernel regression for each possible bandwidth between 1 and 30 minutes (the IK2012 and CCT2014 bandwidths are also indicated on the horizontal axis). With the exception of the smallest bandwidth (1 minute), for which the variance is decidedly too large, the estimates remain remarkably stable and similar to those obtained using parametric methods.

Figure 8: Sensitivity analysis



Note: The left graph shows the 2SLS estimated coefficient of legal status for different specifications of the parametric polynomial regression in Table 3. In particular, the scatterplots are the point estimates obtained for different degrees of the polynomial in the running variable (on the horizontal axis), and the range plots are the associated confidence intervals. The right graph shows the point estimates and confidence intervals estimated by nonparametric local kernel regression within all possible (integer) bandwidths between 1 and 30 minutes. The bandwidths selected according to the IK2012 and CCT2014 criteria are also reported on the horizontal axis.

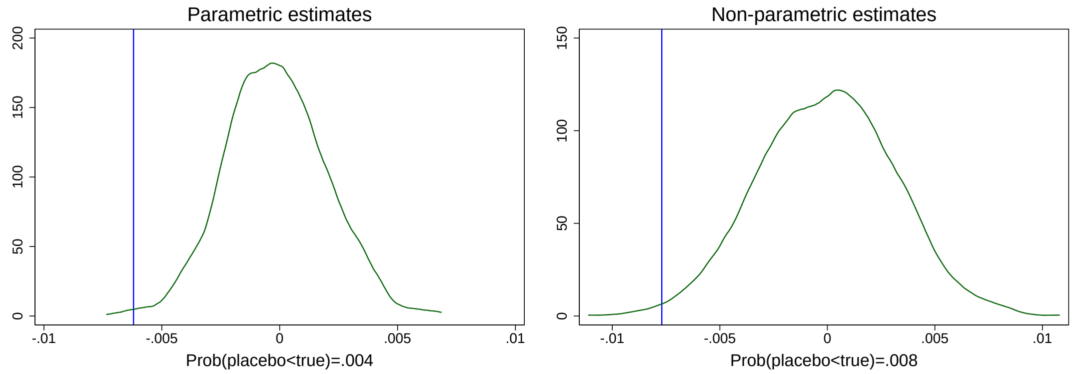
Figure 9 compares these estimates with a distribution of 1,000 placebos. Each placebo estimate is obtained by permuting the cutoff points at random across lotteries, computing a "fake" running variable $\tilde{T}$ as the difference between the timing of each application and the placebo cutoff, and estimating the discontinuity in the probability of committing crimes at $\tilde{T} = 0$. Because the placebo cutoffs have, by construction, no predictive power for the probability of obtaining a residence permit – the first stage of the regression – we consider only the reduced-form coefficients.

The distributions of parametric and nonparametric placebos are centered at zero, and the probability of obtaining estimates as extreme as those at the true cutoffs are 0.004 and 0.008, respectively. These p-values can be interpreted as the probability that, under the null hypothesis of no effect of legal status, the estimating bias is sufficiently large to account for the magnitude of the estimated

[17]The complete results are reported in the Web Appendix (Table A3).

coefficient. In particular, these results appear to exclude the possibility that the baseline estimates only average a small sample bias across lotteries.

Figure 9: Reduced-form effect of legal status on the number of crimes per applicant, placebo estimates



Note: These graphs show the distribution of reduced-form estimates obtained by parametric and nonparametric methods (left and right graphs, respectively) for 1,000 random permutations of the cutoff points across lotteries. The vertical line represents the estimate obtained for the true cutoff point (from Tables 3 and 4, respectively). The fraction of placebo estimates lower than the actual estimate is reported at the bottom of each graph.

Finally, we consider a further source of bias, namely, errors in the match between permit applications and criminal records. As discussed in Section 4.2, random matching errors and intentional misrepresentation of identity by illegal immigrants would bias the estimated coefficient toward zero. The possibility of an opposite bias arises if illegal status increases the probability of being reported for serious offenses – conditional on the same level of criminal activity. For instance, immigrants stopped by the police and found without documents could undergo a closer inspection, which could reveal evidence of additional offenses. If this situation occurred, then we would expect a non-significant coefficient when restricting to immigrants who were *not* reported for illegal status.

Instead, after we exclude immigrants who were reported (also) for violating the migration law, the estimated coefficient of legal status remains negative and statistically significant, although lower in magnitude. In our preferred specification (the kernel local linear regression using the CCT2014 method), the coefficient changes to -0.013 from -0.017 in the baseline specification of Table 4.[18] Note that the estimate for the reduced sample may be biased toward zero by the exclusion of illegal immigrants who were reported for illegal status *after* having been arrested for another (serious) offense.

[18]The full results are presented in the Web Appendix (Table A4).

5.4 Local crime rates

The results presented thus far suggest that the concession of legal status lowers the number of crimes committed per applicant during the period after Click Days. However, the effect of legalization estimated at the aggregate level could differ from that estimated at the individual level owing to interactions in crime between applicants who obtained legal status and those who did not or between immigrants and natives. In particular, the effect of legalization on local crime rates would be higher (lower) if there were complementarity (substitutability) between the criminal enterprises of such groups. As an extreme case, if the number of crime opportunities in a given area is fixed, the concession of legal status will change the identity of offenders but not their total number.

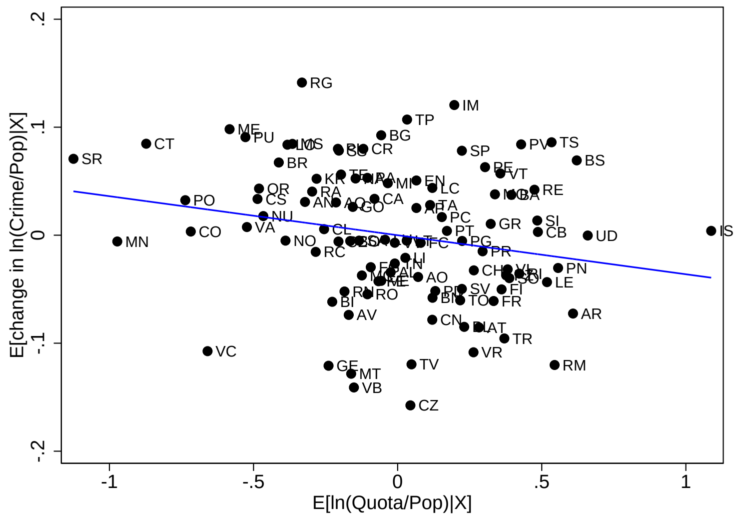
We thus examine the relationship between crime trends across provinces, before and after Click Days, as well as the quotas of permits allocated to each province. Figure 10 shows that, keeping other province characteristics constant – the log crime rate in 2007, the log of the province population, and the log of the number of applications over the province population – the log-change in the number of crimes between 2007 and 2008 is negatively related to the log of the number of permits over the province population.[19]

The slope of the partial regression is reported in column (1) of Table 6; this slope is only slightly affected when we also control for the employment rate in the province and the (log) GDP per worker in the region (column 2).[20] In the remaining columns of Table 6, we exploit the variation over a wider time window (2004-2009). Column (3) presents a baseline difference-in-differences specification in which we interact the quota allocated to each province with a dummy for the period after Click Days; interestingly, higher quotas are associated with a decline in crime only after such event. The estimated interaction coefficient is robust to the inclusion of province and year fixed effects (columns 4 and 5, respectively), the province employment rate and regional GDP per worker (column 6), and province-specific trends, thus controlling for differences in the dynamics (and not only the level) of the crime rate over the previous period.

[19] The univariate relationship between the two variables (i.e., without controlling for additional covariates) is even stronger.

[20] GDP per worker is not available at the provincial level.

Figure 10: Quotas of permits allocated across Italian provinces and change in crime rates



Note: The figure plots the partial regression of the log-change in crime rates between 2007 and 2008 on the log of the number of permits over the province population across Italian provinces, controlling for the log crime rate in 2007, the log of the province population, and the log of the number of applications over the province population.

Table 6: The effect of legalization on crime rates across Italian provinces

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	change 2007-08		difference-in-differences, 2004-09				
ln(Quota/Pop)	-0.036**	-0.030*	0.013				
	(0.015)	(0.016)	(0.008)				
Post			0.010	0.036			
			(0.052)	(0.045)			
Post×ln(Quota/Pop)			-0.024**	-0.027***	-0.020**	-0.021**	-0.058***
			(0.009)	(0.009)	(0.009)	(0.009)	(0.015)
Observations	101	101	505	505	505	505	505
Additional controls	NO	YES	NO	NO	NO	YES	YES
Province FE	NO	NO	NO	YES	YES	YES	YES
Year FE	NO	NO	NO	NO	YES	YES	YES
Province FE×trends	NO	NO	NO	NO	NO	NO	YES
R-squared	0.970	0.971	0.423	0.496	0.581	0.581	0.864

Note: This table reports OLS estimates for the effect of legalization of Click Day applicants on the number of crimes committed across Italian provinces. The dependent variable is the yearly log-change in crime rates, and the main explanatory variable is the quota of residence permits over the total population in each province. The specifications in columns (1) and (2) consider only the change between 2007 and 2008, while the specifications in columns (3)-(7) include all years in the 2004-2009 period and interact the explanatory variable with a dummy, Post, that equals one for the period after Click Days. The log of the crime rate in the previous year, the log of the province population, and the log of the number of applicants over the province population are included in all specifications. The additional control variables that are included in columns (2), (6), and (7) are the employment rate and the log of regional GDP per worker. Robust standard errors are reported in parentheses, and they are clustered by province in columns (3)-(7). *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

5.5 Heterogeneity

Overall, individual-level RD estimates and cross-province OLS regressions provide the same result – namely, that legalization reduces the number of crimes committed in the destination country – although the identification assumption is much stronger in the former type of analysis. For this reason, in the last part of this section, we focus on the RD results to distinguish the estimated effect for different types of crimes and different categories of applicants.[21]

The first row in Table 7 reports the estimated effect of legal status, $\hat{\tau}$, obtained using the CCT2014 "robust" approach (column 4 of Table 4); the average number of crimes per 100 applicants committed by those who obtained legal status, $E(C_1|L=1)$; and the counterfactual crime rate in the event of being denied legal status, $E(C_0|L=1) = E(C_1|L=1) - \hat{\tau}$. According to these estimates, the concession of legal status implies a 60% reduction in the number of crimes committed in Italy, from 2.9 to 1.2 per 100 applicants.

Table 7: Effect of legal status and actual and counterfactual number of crimes per applicant among legalized immigrants, by group of applicants

	obs.	$\hat{\tau}$	$E(C_1\|L=1)$	$E(C_0\|L=1)$
Average effect	110,338	-0.017**	0.012	0.029
		(0.008)		
Economic crimes	110,338	-0.014**	0.008	0.022
		(0.006)		
Violent crimes	110,338	-0.005	0.006	0.011
		(0.005)		
Center-North regions	71,478	-0.023**	0.012	0.035
		(0.010)		
Southern regions	38,860	-0.005	0.011	0.016
		(0.016)		
Bilateral enforcement	30,335	-0.015*	0.003	0.018
		(0.009)		
No bilateral enforcement	80,003	-0.021**	0.015	0.036
		(0.012)		

Note: This table reports the number of observations in each subsample, the 2SLS effect of legal status on the number of crimes per applicant, the actual crime rate among legalized immigrants in the following year, and the counterfactual crime rate if they had not been legalized. *, **, and *** denote statistical significance of the estimated coefficient at the 90%, 95%, and 99% confidence levels, respectively.

The second and third rows show that the average effect is driven entirely by a reduction of property and other economically motivated crimes (e.g., drug trafficking), whereas the coefficient on violent crimes is not significantly different from

[21] The complete results are reported in the Web Appendix (Table A5).

zero. This result is consistent with the idea that legal status affects criminal behavior by changing the relative returns on legitimate activities and, moreover, violent crimes depend to a lesser extent on economic considerations (Machin and Meghir, 2004). This result is also consistent with previous evidence provided by Baker (2013), Freedman et al. (2013), and Mastrobuoni and Pinotti (2014).

To further investigate this channel, we exploit variation in the relative employment opportunities of legal and illegal immigrants across geographical areas. Although information on the labor market outcomes of illegal immigrants is not systematically available across Italian regions, it is well known that northern and central regions exhibit better income opportunities in the official sector, whereas a large share of the labor force in southern regions is employed in the shadow economy.[22] Interestingly, the counterfactual crime rate in the absence of legalization is higher in the north than in the south, but the two areas converge to the same number of crimes after legalization (fourth and fifth rows of Table 7). An explanation consistent with our theoretical framework is that the larger size of the unofficial economy in the south offers better employment opportunities for illegal immigrants. This finding is also consistent with previous results in Mastrobuoni and Pinotti (2014) using a different sample of immigrants and a different source of variation in legal status.

Finally, the model in Section 3 has ambiguous implications for the effect of expulsions on the number of crimes committed by illegal immigrants: those who are expelled no longer commit crimes in the destination country, but a higher probability of expulsion reduces the future expected utility – and, thus, the opportunity cost of crime – of those who are not expelled.

To explore this dimension of the model, we estimate the effects separately for privileged and non-privileged nationalities. As discussed in Section 2.1, immigrants in the former group come from countries that signed bilateral agreements with Italy for the enforcement of migration restrictions, particularly allowing for the faster repatriation of illegals apprehended by the Italian police. In the absence of such agreements, opposition by border authorities in origin countries typically represents a severe obstacle to the effective expulsion of illegal immigrants, who would then receive only a written injunction to leave the Italian territory. Given this situation, the probability of an illegal immigrant actually being expelled should be higher for immigrants of privileged nationalities. The results

[22]According to the official statistics, at the end of 2012, the GDP per capita was 75% higher in the center-north than in the south. At the same time, the (estimated) size of the unofficial economy, as a share of total GDP, was 50% lower in the former region than in the latter region.

reveal that the reduction in crime is lower for such immigrants (last two rows of Table 7). Therefore, the actual enforcement of expulsions – as opposed to the simple injunction to leave the country – appears to be effective in reducing the number of crimes committed by illegal immigrants. However, in the absence of data on the apprehensions and expulsions of Click Day applicants and in consideration of the large differences between the two groups in terms of both potential outcomes, this last result must be interpreted with caution.

6 Conclusions

Illegality imposes a heavy toll on foreign immigrants in terms of poorer employment opportunities, lower incomes, and lesser access to social services, all of which imply a lower opportunity cost for engaging in crime. Indeed, the results of the present paper show that legal status significantly reduces the number of serious crimes committed by immigrants in Italy – from 2.9 to 1.2 per 100 applicants in our preferred specification.

Our theoretical framework highlights two primary channels through which the effect of legal status could operate: first, a decrease in the probability of committing crimes conditional on not being deported as a result of better opportunities in the official labor market; second, a potential increase in the number of crimes committed by legalized immigrants who are no longer being expelled. The estimated reduction in crime suggests that the first effect may prevail, although we can *not* distinguish the effect of legal status on the individual probability of committing crimes (conditional on not being deported) from the effect of differences in the probability of expulsion between legal and illegal immigrants (zero for the former and positive for the latter). This limitation is inherent to our data, as we cannot observe the mobility of immigrants across the border, either because of expulsion or for other reasons.

Notice that the same limitation applies to empirical studies of crime in general, as the crime rate is typically measured by the number of crimes committed in a given period (e.g., one year) over the beginning-of-period stock of potential offenders (e.g., the population in a given area). Insofar as there is population mobility across areas during the period of interest, the estimated effect of any crime determinant captures the combined effect on the stock of potential offenders and on the crime probability of those who did not change location (although the problem is admittedly more relevant for immigrants).

Most important, the effect of legalization on the number of crimes committed in the destination country is arguably the relevant parameter from a policy perspective, as opposed to the (unobserved) change in criminal behavior conditional on not being expelled, which would overestimate the reduction in crime achieved through legalization. In the end, when weighing the advantages and disadvantages of legalization policies, we want to know the change in the number of crimes that is caused by the concession of legal status to one additional applicant: our estimates specifically identify this parameter.

Indeed, we believe that our results can inform the current debate on immigration policy. In the context of increasing pressures at the border as a result of economic and political turmoil in several areas of the world, immigration reform is near the top of the agenda in the United States as well as in many other countries. Some countries are reacting by raising further barriers to the entry of new immigrants. In Switzerland, for instance, a referendum has been passed as recently as February 2014 to reduce immigration quotas and to limit the free mobility of European Union citizens. The results of this paper suggest that the effect of such restrictions will crucially depend on the effective degree of enforcement. If enforcement is low, then the ultimate effect of the policy could be to exclude a number of immigrants from the official labor market and thus raise the risk that they become involved in criminal activity.

References

Accetturo, A. and L. Infante (2010). Immigrant earnings in the italian labour market. *Giornale degli Economisti 69*(1), 1–28.

Alonso-Borrego, C., N. Garoupa, and P. Vázquez (2012). Does immigration cause crime? evidence from spain. *American law and economics review 14*(1), 165–191.

Amuedo-Dorantes, C., C. Bansak, and S. Raphael (2007). Gender differences in the labor market: Impact of irca. *American Economic Review 97*(2), 412–416.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica 61*(4), 821–856.

Baker, S. (2013). Effects of immigrant legalization on crime: The 1986 immigration reform and control act. Unpublished manuscript, Stanford University.

Barbagli, M. (2008). *Immigrazione e sicurezza in Italia*. Il mulino.

Becker, G. S. (1968). Crime and punishment: An economic approach. *The Journal of Political Economy 76*(2), pp. 169–217.

Bell, B. and S. Machin (2013). Crime and immigration: What do we know? In P. Cook, S. Machin, O. Marie, and G. Mastrobuoni (Eds.), *Lessons from the Economics of Crime: What Reduces Offending?* MIT Press.

Bell, B., S. Machin, and F. Fasani (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics 95*(4), 1278–1290.

Bertrand, M., R. Hanna, and S. Mullainathan (2010). Affirmative action in education: Evidence from engineering college admissions in india. *Journal of Public Economics 94*(1), 16–29.

Bianchi, M., P. Buonanno, and P. Pinotti (2012). Do immigrants cause crime? *Journal of the European Economic Association 10*(6), 1318–1347.

Borjas, G. J., J. Grogger, and G. H. Hanson (2010). Immigration and the economic status of african-american men. *Economica 77*(306), 255–282.

Butcher, K. F. and A. M. Piehl (1998). Cross-city evidence on the relationship between immigration and crime. *Journal of Policy Analysis and Management 17*(3), 457–493.

Caponi, V. and M. Plesca (2013). Empirical Characteristics of Legal and Illegal Immigrants in the U.S. IZA Discussion Papers 7304, Institute for the Study of Labor (IZA).

Cattaneo, M., S. Calonico, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. Unpublished manuscript, University of Michigan.

Chay, K. Y., P. J. McEwan, and M. Urquiola (2005). The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review 95*(4), 1237–1258.

Chen, A., T. Bengtsson, and T. K. Ho (2009). A regression paradox for linear models: Sufficient conditions and relation to simpsons paradox. *The American Statistician 63*(3).

Cohn, J. and J. Passel (2009). A portrait of unauthorized immigrants in the united states. Pew hispanic center.

Dahl, G. B., K. V. Lken, and M. Mogstad (2014). Peer Effects in Program Participation. *American Economic Review (forthcoming).*

Del Boca, D. and A. Venturini (2005). Italian migration. In K. Zimmermann (Ed.), *European Migration : What Do We Know?* Oxford University Press.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *The Journal of Political Economy 81*(3), pp. 521–565.

Fredriksson, P., B. ckert, and H. Oosterbeek (2013). Long-Term Effects of Class Size. *The Quarterly Journal of Economics 128*(1), 249–285.

Freedman, M., E. Owens, and S. Bohn (2013). Immigration, employment opportunities, and criminal behavior. Unpublished manuscript, Cornell University.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Hoefer, M., N. Rytina, and B. C. Baker (2012a). Estimates of the legal permanent resident population in 2011. Population estimates, Office of Immigration Statistics, Department of Homeland Security.

Hoefer, M., N. Rytina, and B. C. Baker (2012b). Estimates of the unauthorized immigrant population residing in the united states: January 2011. Population estimates, Office of Immigration Statistics, Department of Homeland Security.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies 79*(3), 933–959.

Italian Ministry of Interior (2007). Rapporto sulla criminalità in italia. analisi, prevenzione, contrasto. Technical report.

Kaushal, N. (2006). Amnesty programs and the labor market outcomes of undocumented workers. *Journal of Human Resources 41*(3), 631–647.

Kossoudji, S. A. and D. A. Cobb-Clark (2002). Coming out of the shadows: Learning about legal status and wages from the legalized population. *Journal of Labor Economics 20*(3), 598–628.

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics 142*(2), 675–697.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *The Journal of Economic Literature 48*(2), 281–355.

Lozano, F. A. and T. A. Sorensen (2011). The Labor Market Value to Legal Status. IZA Discussion Papers 5492, Institute for the Study of Labor (IZA).

MacDonald, Z. (2002). Official crime statistics: Their use and interpretation. *The Economic Journal 112*(477), F85–F106.

Machin, S. and C. Meghir (2004). Crime and economic incentives. *Journal of Human Resources 39*(4), 958–979.

Mastrobuoni, G. and P. Pinotti (2014). Legal status and the criminal activity of immigrants. Unpublished manuscript, Bocconi University.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Porter, J. (2003). Estimation in the regression discontinuity model. Unpublished manuscript, University of Wisconsin at Madison.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66*(5), 688 – 701.

Spenkuch, J. L. (2014). Understanding the impact of immigration on crime. *American Law and Economics Review 16*(1), 177–219.

Transatlantic Trends (2009). Immigration survey. German Marshall Fund, Compagnia di San Paolo, Barrow Cadbury Trust, and Fundacin BBVA. available at http://trends.gmfus.org/immigration.

Web Appendix

Table A1: Example of misreporting of identity by an illegal immigrant

NAME	SURNAME	NATIONALITY	BIRTHDATE
Samri	Jamli	Algeria	23.7.1965
Abdi	Aldhi	Algeria	23.7.1970
Abdila	Aldhi	Algeria	23.7.1970
Abdihdi	Abdila	Algeria	23.7.1970
Samir	Karim	Marocco	23.7.1970
Ali	Nasire	Marocco	24.3.1974
Neighebouti	Razki	Algeria	23.7.1967
Abdlhah	Abdlhdi	Marocco	23.7.1970
Sadaci	Mohamed	Marocco	23.8.1975
Galesi	Manim	Per	23.7.1970
Saadi	Karimi	Marocco	23.7.1973
Sissmr	Yamlih	Albania	23.7.1970
Sarim	Karim	Albania	23.7.1970
Chebouti	Akzki	Marocco	23.7.1970
Samri	Yamlih	Algeria	23.7.1966
Sadeik	Sakkipei	Sarajevo	23.7.1970

Note: The table presents the list of alias used, over the years, by an illegal immigrant arrested in the city of Bologna. This judicial case is described at length in Barbagli (2008).

Table A2: Descriptive statistics, full sample

	(1)	(2)	(3)	(4)
	ALL LOTTERIES		SIGNIFICANT BREAK	
	males	*females*	*males*	*females*
age	33.924	39.089	33.856	39.137
	(0.015)	(0.028)	(0.017)	(0.030)
Africa	0.269	0.086	0.260	0.079
	(0.001)	(0.001)	(0.001)	(0.001)
Americas	0.066	0.174	0.071	0.178
	(0.000)	(0.001)	(0.001)	(0.001)
Asia	0.487	0.323	0.479	0.318
	(0.001)	(0.001)	(0.001)	(0.001)
Europe (outside the EU)	0.177	0.417	0.190	0.424
	(0.001)	(0.001)	(0.001)	(0.001)
Northern Italy	0.626	0.605	0.648	0.616
	(0.001)	(0.001)	(0.001)	(0.001)
Centre Italy	0.205	0.228	0.205	0.226
	(0.001)	(0.001)	(0.001)	(0.001)
Southern Italy	0.169	0.168	0.147	0.158
	(0.001)	(0.001)	(0.001)	(0.001)
REPORTED CRIMES IN 2008, PER 100 APPLICANTS				
all types of crime	1.220	0.079	1.263	0.083
	(0.030)	(0.010)	(0.033)	(0.012)
theft	0.231	0.038	0.236	0.040
	(0.011)	(0.006)	(0.012)	(0.006)
robberie	0.171	0.004	0.181	0.005
	(0.009)	(0.002)	(0.011)	(0.002)
drug-trafficking	0.182	0.014	0.181	0.015
	(0.011)	(0.007)	(0.012)	(0.008)
smuggling	0.024	0.004	0.026	0.005
	(0.004)	(0.002)	(0.005)	(0.002)
kidnapping	0.022	0.000	0.023	0.000
	(0.003)	.	(0.003)	.
extortion	0.044	0.010	0.046	0.010
	(0.004)	(0.003)	(0.005)	(0.003)
illegal carrying of firearms	0.131	0.001	0.132	0.002
	(0.008)	(0.001)	(0.009)	(0.001)
murder	0.043	0.002	0.046	0.002
	(0.005)	(0.001)	(0.006)	(0.001)
violent assault	0.241	0.006	0.256	0.005
	(0.010)	(0.002)	(0.011)	(0.002)
rape	0.132	0.000	0.136	0.000
	(0.010)	.	(0.011)	.
observations	256,703	147,034	212,039	128,411

Note: This table shows the average characteristics (top panel), the number of reported crimes per 100 applicants (bottom panel), and the total number of observations (last row) for the applicants to all lotteries (columns 1-2) and to lotteries exhibiting a significant structural break in the probability of treatment assignment at the cutoff (columns 3-4), including also those whose application was received outside the symmetric one-hour bandwidth around the cutoff. Separate figures by gender are presented, and standard errors are reported in parentheses.

Table A3: The effect of legal status on the number of crimes per applicant, global polynomial regression including all applicants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	BASELINE SPECIFICATION			DEMOGRAPHIC CONTROLS			LOTTERY FE, CLUSTERED S.E.		
	all	*males*	*females*	*all*	*males*	*females*	*all*	*males*	*females*
Reduced form	-0.004***	-0.004***	0.000	-0.002***	-0.004***	0.000	-0.003***	-0.004***	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
First stage	0.601***	0.561***	0.648***	0.596***	0.561***	0.647***	0.586***	0.544***	0.646***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.026)	(0.021)	(0.035)
2SLS estimate	-0.006***	-0.007***	0.000	-0.004***	-0.007***	0.000	-0.004***	-0.007***	0.000
	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.000)
Observations	340,447	212,038	128,409	340,444	212,038	128,406	340,444	212,037	128,347
First stage F-stat	79418.7	39832.3	39951.0	78039.8	39846.9	39927.9	499.43	651.93	338.2

Note: This table reports 2SLS estimates for the effect of legal status on the number of crimes committed in Italy by Click Day applicants, including those whose application was received outside the one-hour symmetric time window around the cutoff. The dependent variable is a dummy $C = 1$ for individuals committing at least 1 serious offense in 2008, the explanatory variable of interest is a dummy $L = 1$ for those obtaining a residence permit, and the first-stage instrument is a dummy $Z = 1$ for having applied before the cutoff time. All regressions control for a quadratic polynomial in the time elapsed since the cutoff (by the millisecond) and its interaction with Z, and the specifications in columns (7)-(9) also include lottery fixed effects. The first row reports the reduced-form regression of C on Z, the second row reports the first stage regression of L and Z, and the third row reports the 2SLS-estimated effect of L on C. The last row of the table reports the first-stage F statistic for the excluded instrument, Z. Robust standard errors are reported in parentheses, and they are clustered by lottery in columns (7)-(9). *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

Table A4: The effect of legal status on the number of crimes per applicant, kernel local linear regression (excluding immigrants reported *also* for illegal status)

	(1)	(2)	(3)	(4)
BANDWIDTH SELECTION:	IK2012	[21:54]	CCT2014	[06:55]
BIAS CORRECTION:	NO	YES	NO	YES
Reduced form	-0.003*	-0.011*	-0.005*	-0.006*
	(0.002)	(0.006)	(0.003)	(0.003)
First stage	0.467***	0.458***	0.445***	0.442***
	(0.006)	(0.022)	(0.011)	(0.012)
2SLS estimate	-0.007*	-0.023*	-0.011*	-0.013*
	(0.004)	(0.013)	(0.006)	(0.007)
Total observations	106,472		106,472	
obs. inside the BW	84,232		30,019	

Note: This table reports nonparametric estimates of the effect of legal status on the number of crimes committed in Italy by Click Day applicants, excluding those there were subsequently reported for illegal status (possibly in addition to other offenses). The dependent variable is a dummy $C = 1$ for individuals committing at least 1 serious offense in 2008, the explanatory variable of interest is a dummy $L = 1$ for those obtaining a residence permit, and the first-stage instrument is a dummy $Z = 1$ for having applied before the cutoff time. All coefficients are estimated using a kernel local linear regression in a symmetric bandwidth around the cutoff. The table reports the bandwidth selection rule, IK2012 or CCT2014, at the top of each column; the size of the bandwidth in minutes and seconds [mm:ss]; and whether the re-centering procedure by CCT2014 is employed. The first row reports the reduced-form regression of C on Z, the second row reports the first-stage regression of L on Z, and the third row reports the estimated effect of L on C obtained as the ratio between the reduced form and the first stage. Robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

Table A5: The effect of legal status on the number of crimes per applicant, by type of crime and category of applicant

	(1)	(2)	(3)	(4)	(5)	(6)
	TYPE OF CRIME		GEOGRAPHICAL AREA		BILATERAL ENFORC.	
	ECON	VIOL	NORTH	SOUTH	YES	NO
Reduced form	-0.006**	-0.002	-0.012**	-0.002	-0.008*	-0.008**
	(0.003)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)
First stage	0.450***	0.641***	0.511***	0.317***	0.561***	0.396***
	(0.011)	(0.019)	(0.016)	(0.022)	(0.020)	(0.014)
2SLS estimate	-0.014**	-0.003	-0.023**	-0.005	-0.015*	-0.021**
	(0.006)	(0.004)	(0.010)	(0.016)	(0.009)	(0.012)
Total observations	110,338	110,338	71,478	38,860	30,335	80,003
inside the BW	38,544	26,060	17,540	11,159	9,279	24,080

Note: This table reports nonparametric estimates of the effect of legal status on the number of crimes committed in Italy by Click Day applicants, distinguishing between different types of crime and different categories of applicants (reported on top of each column). The dependent variable is a dummy $C = 1$ for individuals committing at least 1 serious offense in 2008, the explanatory variable of interest is a dummy $L = 1$ for those obtaining a residence permit, and the first-stage instrument is a dummy $Z = 1$ for having applied before the cutoff time. All coefficients are estimated by kernel local linear regression using the CCT2014 bandwidth and applying the bias-correction procedure. The first row reports the reduced-form regression of C on Z, the second row reports the first-stage regression of L on Z, and the third row reports the estimated effect of L on C obtained as the ratio between the reduced form and the first stage. Robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.