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

Paolo Pinotti

January 2015

Abstract

We estimate the effect of immigrant legalization on the crime rate of immigrants in Italy by exploiting 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. Matching data on applications with individual-level criminal records, we show that legalization reduces the crime rate of legalized immigrants by 0.6 percentage points on average, on a baseline crime rate of 1.1 percent. The effect persists for (at least) two years after legalization.

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

*Contacts: Paolo Pinotti, Bocconi University 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, Ray Fisman, Andrea Ichino, Eliana La Ferrara, Steve Machin, Magne Mogstad, Nicola Pavoni, Nicola Persico, Luigi Pistaferri, Imran Rasul, and the seminar participants at the American Economic Association 2015, NBER Summer Institute 2014 (joint Labor Studies and Crime session), UCLA, UC-Davis, University of Stockholm, Gothenburg University, America Latina Crime and Policy Network (AL CAPONE) in Sao Paulo, HEC Lausanne, European University Institute, UCL-NHH Crime Conference, IFS Workshop on Advances in Policy Evaluation 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. The financial assistance from Fondazione Rodolfo Debenedetti is gratefully acknowledged.

1 Introduction

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 immigrants (Hoefer et al., 2012a, Hoefer et al., 2012b). The share of undocumented is also high in several European countries, such as Italy, Greece, Portugal, and Spain.[1]

Both in the US and in Europe, undocumented immigrants cannot officially work or start a new economic activity. They can only work in the informal economy, and hence face inferior earnings opportunities relative to their legal counterparts (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.

Indeed, the presence of a large undocumented population raises concerns about crime in destination countries. According to an annual survey conducted in North American and European countries, approximately two-thirds of the people interviewed were concerned that irregular immigrants increase crime, whereas only one-third held similar concerns about regular immigrants (Transatlantic Trends, 2009). Moving from perceptions to criminal statistics, the undocumented constitute between 20% and 30% of all immigrants in Italy, but 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, the undocumented are typically young, single males, and are less educated than legal immigrants (Cohn and Passel, 2009, Caponi and Plesca, 2013, and Mastrobuoni and Pinotti, 2015). 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 unofficially 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 Discontinu-

[1]The Italian case is discussed in detail in the rest of the paper, estimates for the other European countries are provided by the EU-funded project "Clandestino" (www.irregular-migration.net).

ity (RD) design that allows us to identify the causal effect of legal status on crime. The primary method of acquiring legal status in Italy is through work-related residence permits sponsored by the immigrant's employer. Typically immigrants enter Italy irregularly, start working unofficially for an employer, and hope that the latter will subsequently sponsor them for a residence permit. Fixed quotas of permits are available each year for different groups of applicants defined by the type of employer and occupation, nationality, and Italian province, for a total of 1751 groups. Applications must be submitted online by employers starting at 8:00 a.m. on given "Click Days" of the year and are processed on a first-come, first-served basis until the exhaustion of the available quota of permits. The rationing of permits 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 employers applied a few minutes (or seconds) after the cutoff time.[2]

In this paper, we match all male immigrants sponsored on December 2007 (for receiving a permit in 2008) with restricted-use data on all male foreigners reported by the police for having committed a serious crime in Italy during the period 2007-2009. Taking advantage of detailed information on the timing of application – available at the millisecond – we compare the probability of being reported for having committed a serious crime in Italy, before and after legalization, across applicants who applied shortly before and shortly after the cutoff. For most groups of applicants, this cutoff occurred fewer than 30 minutes after the start of the Click Day. Most importantly, the exact timing of the cutoff for each group was unknown ex-ante, as it depended on the timing of all applications as well as on how many applications were rejected for being inaccurate, false, or incomplete. These complexities provide a compelling argument for the fundamental identification assumption that applicants within an arbitrarily narrow bandwidth of the cutoff were unable to *precisely* determine their 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 immigrants who applied shortly before or shortly after the cutoff. In particular, the crime rate in 2007 – the year before Click Days – is equal to 1.1 percent in both groups (i.e., 11 in 1000 individuals reported for serious crimes).

In 2008 – the year immediately after Click Days – the crime rate declines to

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

0.8 percent for immigrants who applied before the cutoff, while it remains at 1.1 for those who applied after the cutoff. The difference is statistically significant and it is unaffected when controlling parametrically for a smooth polynomial in the timing of application. Since the difference in the probability of obtaining legal status between these two groups is about 50 percentage points, the two-stage-least-squares (2SLS) estimated effect of legal status is a 0.6 percentage-point reduction in the crime rate (the magnitude is larger when restricting non-parametrically to a narrow bandwidth of the cutoff time).

The decline in crime rate seems persistent, as the crime rate of early applicants remains at 0.8 percent in 2009. In that year, most of the unsuccessful applicants in 2007 were also legalized as a consequence of a one-off procedure that gave priority to previous applicants for obtaining a residence permit in 2009. Upon legalization, the crime rate of this group also declined to 0.8 – the same level observed for immigrants legalized the year before. Overall, these results suggest that the reduced form effect observed for early applicants, equal to -0.3 percentage points, persists for (at least) two years after Click Days. This allows us to exclude the possibility that the drop in crime is explained by the bureaucratic procedures imposed on successful applicants – particularly, the fact that undocumented immigrants awarded a residence permit must go back to the country of origin to collect an entry visa and re-enter Italy.

To gain some insights into the mechanisms behind the relationship between legal status and crime, we parse the data by the type of applicant, offense, and sponsor. The average effect across all applicants is driven by a reduction in the number of economically-motivated crimes committed by applicants sponsored by other immigrants as domestic workers (e.g., housekeepers, elder care givers, and so on) whereas the effect is zero for firm-sponsored employees. In the next section, we preliminarily show that the former category of applicants likely includes a significant number of individuals that would not qualify for a residence permit – male immigrants are typically *not* employed as domestic workers in Italy – but are nevertheless sponsored through fictitious job offers. Such individuals exhibit a higher crime rate before the Click Days (almost 2 percent, twice the average across all applicants), which however declines significantly in the event of obtaining legal status (-1.3 percentage points). Instead, the crime rate for firm-sponsored employees, whose applications are generally backed by real job offers, is very low both before and after Click Days (0.5-0.7 percent).

Based on these findings, we conclude that immigrants in the latter group experience a higher opportunity cost of crime due to the fact they are already (unofficially) employed in the sponsor firm before Click Days. In contrast, in most cases, male immigrants sponsored by other immigrants as domestic workers have neither a job nor a real job offer, so their opportunity cost of committing a crime as undocumented workers is very low. At the same time, this group seems more responsive to legalization, possibly because legal status partly addresses detachment from the labor market. This can be particularly important for individuals at the margin between pursuing licit or illicit activities, who are probably over-represented among applicants with worse labor market opportunities.

The empirical results are extremely stable under a variety of specifications of both parametric and non-parametric regressions. In particular, the coefficient of interest is largely unaffected when varying the degree of the polynomial in the timing of application between 0 and 6, when varying the bandwidth around the cutoff time between 1 and 30 minutes, and when including a full set of fixed effects for applicants competing for the same quota of permits. This last result suggests that the estimated coefficient of legal status is not biased by heterogeneity in the baseline crime rate across groups. Based on a placebo exercise in which we permute the cutoff times across groups, we can also exclude that such an estimate is driven to a significant extent by small sample bias within groups. Finally, additional evidence allows us to exclude that the negative coefficient of legal status is explained by differences in the under-reporting of crimes committed by regular and irregular immigrants. If anything, the strategic misreporting of identity, which should be easier for those not holding official documents, should entail an opposite bias.

This paper contributes to the growing body of evidence regarding the relationship between immigration and crime.[3] Earlier work by Butcher and Piehl (1998) shows no evidence that immigration increased crime across US counties during the 1980s, whereas Spenkuch (2014) reaches an opposite 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) find that immigration increased crime 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

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

accession countries. Interestingly, only the former group, which was characterized by limited access to the official labor market, caused a significant increase in (property) crime.

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 unofficial immigrants in the US. Using the 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 undocumented immigrants, creating obstacles to the employment of those who were not legalized. Freedman et al. (2013) focus on the implications of these additional IRCA provisions and document an increase in the felony charges filed against Hispanic residents of San Antonio, Texas, after the expiration of the amnesty deadline.

With regard to other countries, in a previous paper with Giovanni Mastrobuoni (2015), we exploit variation in legal status across pardoned prison inmates in Italy 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.

The present analysis estimates the causal effect of legal status taking advantage of 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. At the same time, the institutional details of Italian migration policy allow us to estimate the policy effect separately for groups of immigrants characterized by different employment opportunities. This heterogeneity allows for a better understanding of the channels through which legal status affects criminal behavior.

This paper is organized as follows. The next section describes Italian migration policy, the characteristics of different groups of applicants, and the functioning of Click Days. Section 3 discusses the empirical strategy and the data. Section 4 presents the main results. Finally, Section 5 concludes the paper with some policy implications as well as some caveats to our analysis.

2 Institutional framework

In this section we provide an overview of Italian migration policy, focusing on the features that are most relevant for our research design. We emphasize throughout the divergence between the formal rules of admission and residence in the country, on the one hand, and the lax enforcement of migration restrictions on the other.

The regulatory framework designed by Laws 40/1998 and 189/2002 foresaw a rigid demand-driven approach to the management of migration inflows. In particular, the central government establishes at the end of each year (usually in November) the quotas of residence permits available the following year for different categories of workers. On given Click Days, employers in Italy – either natives or legal foreign residents – would apply on behalf of perspective immigrants *before* they entered Italy. Immigrants obtaining a permit would start working for the sponsor employer and they could also search for other jobs, start a new business, and so on. If they remain unemployed, however, they would have six months to find another occupation, after which they should leave Italy.

In practice, however, the system is rarely employed as designed. Most employers in Italy are reluctant to hire complete strangers from abroad – particularly families hiring domestic and care workers (*colf* and *badanti*, in Italian) and the small enterprises that constitute the bulk of Italian industry.[4] As a consequence, most immigrants enter in Italy unofficially, start working in the unofficial economy, and hope to be sponsored at one of the following Click Days. Therefore, migration quotas are "used as crypto-regularizations rather than as an instrument for regulating legal entries" (Cuttitta, 2008).[5]

[4]Law 40/1998 allowed third country nationals to access and reside in Italy for up to 12 months while searching for a job, however this possibility was dismissed with the reform of 2002.

[5]Although irregular immigrants are in principle subject to removal by immigration authorities, this risk is relatively low. Due to the cost of enforcing deportations, most immigrants apprehended by the police for being unofficially present in Italy receive just an injunction to leave the country. Combining estimates of the undocumented population, based on applications for amnesty episodes, and information on removals enforced by the Italian police during the period 2004-2007, Dustmann et al. (2010) estimate the risk of deportation (conditional on being unofficially present in Italy) to be approximately 5 percent.

As for employers, their main incentive in sponsoring applicants on Click Days is avoiding future sanctions for employing irregular workers. To encourage exit from informality, Click Day applications were never used to track undocumented immigrants or their sponsor employers, despite recurrent anomalies such as the excessive number of male applicants allegedly sponsored as domestic and care workers. We next describe the details of the quota system as well as the anomaly just mentioned.

2.1 The quota system

Immigrants can be sponsored for two main types of permits: type-A permits for domestic and care workers employed by individuals and families; and type-B permits for firm employees, further distinguished into B1 and B2 permits for construction or non-construction workers, respectively. In addition, specific quotas without distinction by type of permit are reserved for immigrants coming from 14 "privileged" countries that subscribe to bilateral agreements to control irregular migration.

Column (1) of Table 1 shows the quotas fixed for year 2008. Of 170,000 permits, slightly less than three quarters were awarded to immigrants of non-privileged nationalities, divided almost equally between type-A and type-B permits. The remaining 44,600 permits were reserved for privileged nationalities – which accounted for more than half of all applications. Contrary to the intentions of the policy-maker, such nationalities thus faced a tighter rationing of residence permits relative to non-privileged nationalities – the ratio of quotas over applications being 13 and 49 percent for the former and the latter group, respectively (see the last column of Table 1).[6]

National quotas are 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 provinces lie below the 45-degree line, indicating that quotas 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.

[6]Starting in 2009, however, the quotas for the non-privileged nationalities were always set at a very low level.

Table 1: Quotas of residence permits established for year 2008 and applications received

	(1) **quota**	(2) **applications**	(3) **ratio**
Domestic workers (Type-A permits)	65,000	136,576	0.48
Firm employees (Type-B permits)	60,400	120,676	0.50
Privileged nationalities (Type-A + Type-B)	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
Total	**170,000**	**610,239**	**0.28**

Note: This table reports the quotas established at the end of 2007 for 2008, the number of applications received, and the ratio of quotas to applications.

The quota system described above partitions applicants into 1,751 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 permits: A, B1, and B2). 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".[7]

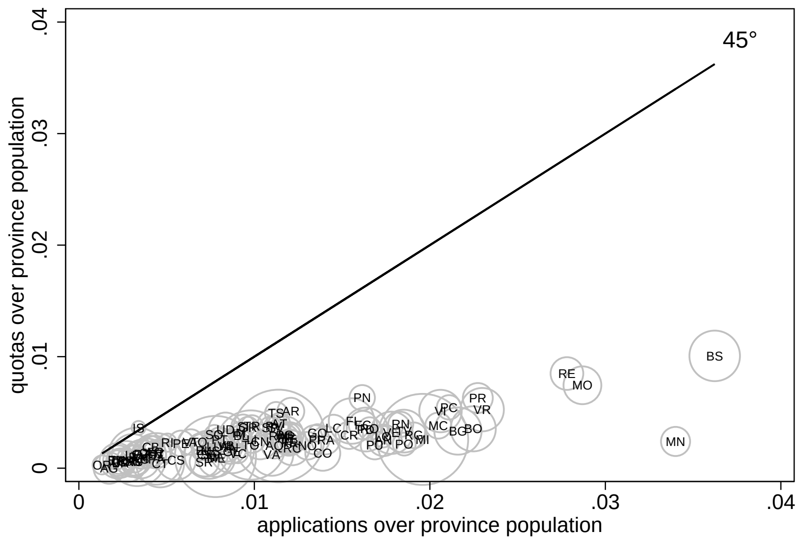
2.2 Type A and type B applicants

In principle, type-A permits should be used by families and individuals to sponsor domestic workers. However, they may also be fraudolently used to legalize immigrants that would not qualify for a residence permit.[8] Immigrants that are unofficially present in Italy but did not receive a job offer may find individuals or

[7]The term *lotteria* (Italian for lottery) has been repeatedly used in the public discourse to describe the Click Days (see, e.g., Italian Senate, 2011)

[8]The existence of a black market for residence permits has long been recognized in Italy and in many other countries (see Sciortino, 1991, and Wasem, 2008, for an exhaustive account of the Italian and the US cases, respectively).

Figure 1: Quotas of residence permits established for year 2008 and applications received, by province



Note: This table reports the provincial quotas established at the end of 2007 for 2008, the number of applications received, and the ratio of quotas to applications, 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.

organizations willing to sponsor them for the payment of a few thousand euros. In such cases, the sponsor would pretend to hire the immigrant as a domestic worker in order to avoid a number of administrative procedures and controls that are imposed on firm employees (e.g., labor and health inspections). In fact, the only prerequisite for sponsoring a domestic worker is having declared an income above 800 euros in the last month before the Click Day.

There is extensive anecdotal evidence of this type of fraud. In 2010, the Corriere della Sera – a leading Italian newspaper – noticed "The Strange Case of Chinese Housekeepers. Where do they work, who hired them, who ever saw them in Italy? Yet, the final data on Click Days uncover 33,000 domestic workers from the People's Republic (...) An anomalous figure indeed: twice as many Ukrainians, who usually work in this occupation (...) A contract as housekeeper is the only way to enter in Italy, it is easier to obtain through family and friends" (March 11, 2011). As for the sponsors, "One Out of Three Chinese People Wants a Housekeeper. According to the data, 1-in-3 Chinese people – including the under-age! – applied to hire (and, thus, to legalize) a housekeeper" (February 16, 2011). Similar anomalies are routinely reported by the Italian press after every Click Day.[9]

To go beyond anecdotal evidence, we compare the incidence and characteristics of type-A applicants in the Click Day files with the results of the ISMU survey of immigrants in the Italian region of Lombardy. The ISMU survey has been conducted annually since 2001 by the NGO "Iniziative e Studi sulla Multietnicitá" on cross-sections of about 9,000 individuals per year, including both regular and irregular immigrants. The sampling of irregular immigrants is based on social networks around a number of "aggregation centers" – train stations, shops, and telephone centers. Sampling weights are provided in order to make the survey representative of the whole foreign population in Lombardy (Blangiardo, 2008).[10]

According to ISMU, in 2007 the share of domestic workers among the employed, male immigrants in the region of Lombardy was only 2.4 percent, increasing to 4.1 percent when expanding this figure to include irregular immigrants; see the first two rows of Table 2. In comparison, 40.9 of male applicants applied for a type-A permit – a clear anomaly; see the third row of the table. The observation is not explained by regional employment differences between Lombardy and Italy as a whole, since the discrepancy increases when we only consider Click Day applications from the region alone (the last row of the table).[11]

Figure 2 provides a more detailed comparison by nationality. For the number of type-A applications to be consistent with the observed patterns of employment in domestic services, the plots should be close to the 45-degree line. However, this is clearly not the case. The anomalies in Click Day applications are particularly evident for males, whose employment in domestic services is traditionally limited to immigrants from Philippines and Sri Lanka, whereas males of all nationalities are sponsored as domestic workers on Click Days (see the left graph in the figure).

Additional anomalies emerge from restricted-use data available to the Italian Ministry of Interior. In particular, sponsors and applicants for type A permits share the same the same nationality in 38 percent of applications, the same city of birth in 32 percent of applications, and the same surname in 21 percent of

[9] Baker (2013) discusses similar evidence for IRCA applications by agricultural workers in the United States.

[10] Lombardy is located in the northwest of Italy, and it is the largest of 20 regions in terms of both population and GDP – approximately 16% and 21% of the total, respectively. The region hosts more than 1 million (regular) immigrants, or one-fourth of all immigrants present in Italy. More information on the ISMU survey is available from the website of the Foundation, www.ismu.org.

[11] The incidence of type-A applications among female applicants at Click Days is also abnormal, although to a lesser extent; see the last column of Table 2.

Table 2: Fraction of domestic workers among employed immigrants in the ISMU survey and type A applicants on Click Days, by gender

EMPLOYED AS A DOMESTIC WORKER:	ALL	MALES	FEMALES
ISMU	0.181	0.025	0.431
ISMU, only undocumented	0.234	0.041	0.522
Click Day	0.562	0.409	0.829
Click Day, only Lombardy	0.589	0.461	0.844

Note: This table compares the incidence of domestic workers in the ISMU survey and among Click Day applicants in year 2007. In particular, the first row of the table shows the fraction of domestic workers among the employed individuals in the ISMU survey, whereas the second row refers to the sub-sample of undocumented immigrants in the survey. The third and fourth rows report the fraction of type-A applicants among all applicants at Click Days and those among the applicants in the region of Lombardy, respectively.

the applications (Italian Ministry of Interior, 2009). Overall, both anecdotal evidence and statistical evidence suggest that type-A applicants may include a large number of unemployed (undocumented) immigrants disguised as domestic workers. This fact will be important for interpreting heterogeneity in the effect of legal status across different lotteries.

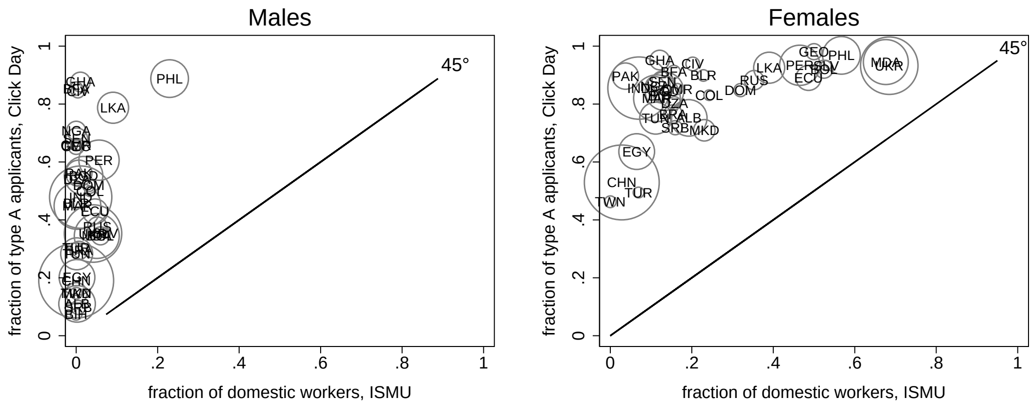
2.3 The application procedure

Until 2006, applications for residence permits were submitted on paper and processed on a rolling basis. This system, described at length in Cuttitta (2008), imposed tremendous delays in the processing of applications, because the manual data entry of paper forms required several months. To remedy this situation, in 2007 the application procedure was completely digitized, and applications are now submitted by employers through the Internet during specific Click Days each year. For instance, in 2007 the Click Day for privileged nationalities was December 15, whereas applications for type-A and type-B permits for other (non-privileged) immigrants had to be sent on December 18 and December 21, respectively.

Starting at 8:00 a.m. on Click Days, employers can log in to the website of the Ministry of the Interior and sponsor one or more immigrants for a residence permit.[12] Applications are then processed on a first-come, first-served basis within each lottery according to the order in which they are received by the electronic system. The processing of each application involves cross-checking

[12] Figure A1 in the Web Appendix shows a screenshot of the dedicated page on the website of the Ministry of Interior.

Figure 2: Fraction of domestic workers among employed immigrants in the ISMU survey and type-A applicants on Click Days by gender and nationality



Note: The two graphs compare the fraction of domestic workers among the employed individuals in the ISMU survey (on the horizontal axis) with the incidence of type-A applicants among all applicants at Click Days (on the vertical axis) distinguished by gender and nationality. The size of the markers is proportional to the size of the national-gender groups among Click Day applicants; the 45-degree line is also included in the graph.

the criminal record of the sponsor and, only for type-B permits, the electronic registries of firms and self-employees maintained by UNIONCAMERE. If there are severe legal impediments – e.g., if the sponsor has been previously convicted for human smuggling – or if part of the information provided in the application is missing, inaccurate, or manifestly fraudulent, the application is rejected. If there are no impediments, instead, the sponsor is notified that the application was successful (*nihil obstat*) and an entry visa allowing the applicant to enter in Italy is sent to the Italian consulate in his/her country of origin. This is because applicants should in principle enter Italy *after* having obtained a residence permit. As discussed before, however, most of them are already (unofficially) in Italy at the time of Click Days. Nevertheless, to comply with formal rules they need first to go back to the country of origin and then re-enter Italy, pretending they have never been there before.[13]

When the number of accepted applications fills the quota established for a given lottery, no further applications are accepted. This mechanism generates a discontinuity in the probability of obtaining a residence permit between those who applied just before and just after the moment in which the quota is exhausted.

Figure 3 shows two examples. The left graph refers to the lottery for type-A

[13]We discuss the implications of this provision for our empirical strategy in Section 4.

applicants of non-privileged nationalities in the province of Milan, the largest city in northern Italy, in which immigrants represent almost 15% of the resident population. The black line shows the total number of applications received at each minute in time. Immediately after 8:00 a.m., the system was flooded with hundreds of applications per minute, and by 9:40 a.m. the flow had already decreased to nearly zero. It is evident from this pattern that, although the system was adopted in 2007 for the first time, applicants were well-aware of the "first-come, first-served" rule.

Most importantly, the ratio of accepted applications (also shown in the graph) is initially very high, as only fraudulent or incomplete applications were rejected, and it decreases to zero when the entire quota assigned to this group is exhausted. This occurred half an hour after the start of the lottery – to be precise, the last accepted application was received at 8:27:04.

As another example, the right graph in Figure 3 shows the lottery for type-B1 applicants (i.e., non-domestic workers outside 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. Additionally, the ratio of applications sent on time that were rejected by immigration officers is larger for type-B applicants in Naples than for type-A applicants in Milan. This difference could reflect a higher incidence of fraudulent applications, more stringent scrutiny by local immigration authorities, or both.[14]

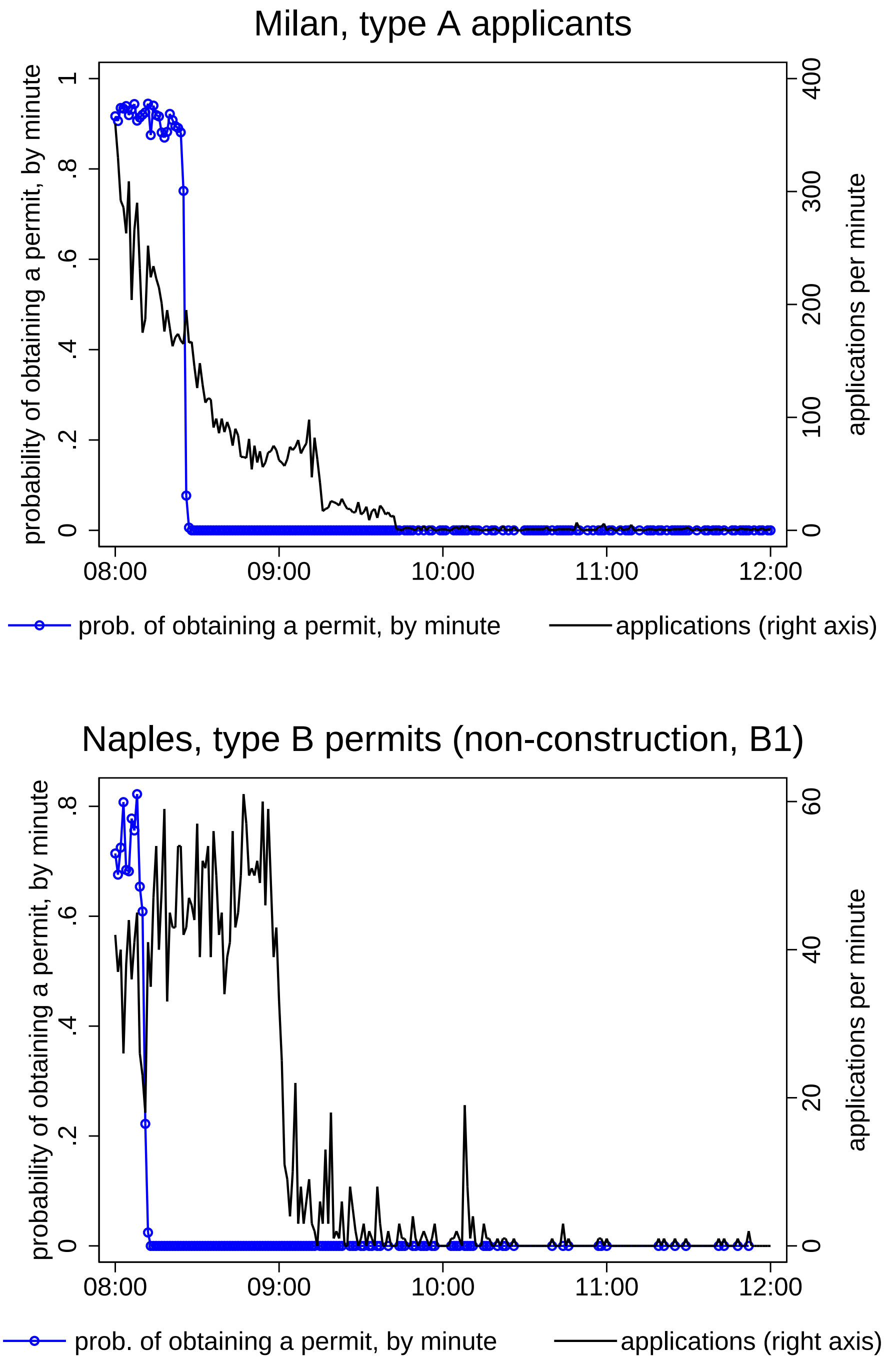
Applications sent after the cutoff time were automatically rejected, and these applicants had no right to reside and work in Italy. As already noted, however, the great majority of these individuals were already present in the country, and they typically remained there (unofficially) even after being refused an official permit.

Our empirical analysis is based on comparing the criminal records of immigrants whose applications were received just before and just after the cutoff. These individuals should be similar in all respects except legal status, so any difference in criminal activity in the following year (2008) can be attributed to the causal effect of legalization.

Finally, applicants that did not obtain a residence permit at the Click Days 2007 were exceptionally given priority one year later. To clear the backlog accumulated in the previous year, no new applications were accepted in 2008. Rather,

[14] In general, rejections of applications sent before the cutoff are more frequent for type-B applicants, possibly because information on sponsor firms is checked against administrative registries that are not available for individuals and families sponsoring type-A applications.

Figure 3: 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: type-A applicants in Milan and type-B applicants (non-construction, B1) in Naples.

the first 150,000 applicants excluded in 2007 were considered for a residence permit in 2009 conditional upon renewing the application between December 15, 2008, and January 3, 2009. Therefore, late applicants also had the possibility of obtaining legal status in 2009, so legal status should differ on average across applicants on opposite sides of the cutoff only in 2008.

3 Empirical strategy and data

This section presents the methods and data employed to estimate the regression discontinuity design described in the previous section. It also discusses potential sources of measurement error and their implications for estimating bias.

3.1 The policy effect

Let L be an indicator variable equal to 1 for applicants that obtained legal status at the Click Days 2007 ($L = 0$ otherwise) and $C_t = 1$ for those who commit a crime in the destination country during year $t = 2007, 2008, 2009$ ($C_t = 0$ otherwise). Following Rubin (1974), the observed difference in crime rates between applicants who obtain and do not obtain legal status can be expressed in terms of "potential outcomes" when regular and irregular, C_t^1 and C_t^0 respectively:

$$E(C_t|L=1)-E(C_t|L=0) = \underbrace{E(C_t^1 - C_t^0|L=1)}_{\tau}+\underbrace{\left[E(C_t^0|L=1) - E(C_t^0|L=0)\right]}_{\varepsilon}. \tag{1}$$

The first term on the right-hand side, $\tau \equiv E(C_t^1 - C_t^0|L = 1)$, is the causal effect of legal status on the crime rate in the destination country across legalized immigrants. The parameter τ depends both on the change in criminal behavior conditional on not being expelled and the change in the probability of expulsion. In the absence of information on expulsions (or, for that matter, mobility across the border), it is impossible to separately identify these two components.

Notice that the same limitation applies generally to empirical studies of crime, 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 this problem is admittedly more relevant for immigrants, we should stress that obtaining a residence permit in Italy does *not* allow for free mobility in the rest of the European Union – or, for that matter, in any other country – so there is no reason to expect greater mobility across the border by legalized immigrants.[15]

Most importantly, the overall effect of legalization on crime 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. The latter parameter would, in fact, over-estimate the reduction in crime achieved through legalization. In the end, when weighing the advantages and disadvantages of legalization policies, we wish to know the change in the number of crimes caused by the concession of legal status to the potential pool of applicants; our estimates specifically identify this parameter.

The main threat to empirically identifying τ is the selection bias ε, which summarizes the effect of other determinants of criminal behavior that are potentially correlated with legal status – for instance, labor market ability and risk aversion. However, differences along these other dimensions should tend to zero when we restrict to applicants within a narrow bandwidth of the cutoff time.

3.2 The RD design

Let X be the timing of applications, with $X = 0$ at the cutoff. If all and only immigrants applying before the cutoff obtained legal status (i.e., $L = 1 \Leftrightarrow X \leq 0$), then

$$\lim_{x \to 0^-} E(C_t|X = x) - \lim_{x \to 0^+} E(C_t|X = x) = E(C_t^1 - C_t^0|X = 0), \tag{2}$$

i.e., the difference in crime rates between those applying just before and just after the cutoff equals the causal effect of legal status on applicants at the cutoff.[16]

However, applicants before the cutoff may still have the application rejected (see Figure 3), so legal status does not change sharply at the cutoff. In this case, commonly referred to as "fuzzy" RD design, the average causal effect on applicants actually obtaining legal status equals the ratio of the "reduced form" effect in (2) over the "first stage" change in the probability of obtaining legal

[15]If anything, the opposite might be true, as some immigrants could decide to leave the country upon having the application rejected.

[16]Equation (2) implicitly assumes continuity of τ and ϵ in X at $X = 0$.

status

$$\frac{\lim\limits_{x\to 0^-} E(C_t|X=x) - \lim\limits_{x\to 0^+} E(C_t|X=x)}{\lim\limits_{x\to 0^-} E(L|X=x) - \lim\limits_{x\to 0^+} E(L|X=x)}, \tag{3}$$

see Hahn et al. (2001).

The ratio (3) can be estimated using both parametric and non-parametric Wald-type estimators. In the former case, we estimate the effect of L on C_t by two-stage least squares (2SLS) using $Z \equiv 1\{X \leq 0\}$ as an instrument for L, and including on the right-hand side a J-th order polynomial in X and its interaction with Z (Lee and Lemieux, 2010). Specifically, the reduced-form and first-stage estimating equations are

$$\begin{aligned} C_t &= \theta_0 + \beta Z + \sum_{j=1}^{J} \theta_j X^j + \sum_{j=1}^{J} \vartheta_j X^j \cdot Z + \epsilon && (4) \\ L &= \pi_0 + \alpha Z + \sum_{j=1}^{J} \pi_j X^j + \sum_{j=1}^{J} \varpi_j X^j \cdot Z + \nu, && (5) \end{aligned}$$

where ϵ and ν are error terms summarizing the effect of other omitted factors. Controlling for smooth trends in X, the coefficients β and α capture the extent of any discontinuity in C_t and L at the cutoff, and their ratio provides an estimate for (3). Following Gelman and Imbens (2014), we focus on a quadratic polynomial in X in the main specification, although we explore the sensitivity of the results to using any polynomial degree between 0 and 6.

Alternatively, non-parametric estimators restrict the sample to applicants within an arbitrarily small bandwidth of the cutoff $X = 0$, 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

$$\begin{aligned} &\min_{\beta,\gamma'} \quad \textstyle\sum_i K\left(\frac{X_i}{h}\right) \quad \left(C_{t,i} - \beta Z - \gamma_0 - \gamma_1 X_i - \gamma_2 X_i \cdot Z_i\right)^2 && (6) \\ &\min_{\alpha,\delta'} \quad \textstyle\sum_i K\left(\frac{X_i}{h}\right) \quad \left(L_i - \alpha Z - \delta_0 - \delta_1 T_i - \delta_2 T_i \cdot Z_i\right)^2, && (7) \end{aligned}$$

where $i = 1, 2, ..., N$ denote the individuals in the sample, and $K\left(\frac{X_i}{h}\right)$ 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. We will explore the sensitivity of the results to alternative criteria for choosing the bandwidth, namely, the theory-based rules of Imbens and Kalyanaraman (2012) and Cattaneo et al. (2014) as well as a battery of heuristic bandwidths.

3.3 Data and measurement issues

To implement the empirical strategy described above, we obtained from the Italian Ministry of Interior the anonymized data on all applications presented at the Click Days on December 2007. Each record includes the country of origin, age, and gender of the applicant, the type of permit sponsored in the application, the province in which it was presented, the timing in milliseconds, and the outcome (i.e., whether the applicant eventually obtained a residence permit). 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 – limiting ourselves to males, the final sample includes 110,337 applicants.[17]

The Immigration Department and the Criminal Police Department of the Ministry of Interior agreed to match the data on applications with the individual 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, and/or rape – in any year between 2007 and 2009. The two archives were matched by name, surname, nationality, and year of birth, so 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. Also, the final data are subject to under-reporting, as reported crimes always underestimate the true number of committed crimes (MacDonald, 2002).

To explore the implications of measurement error from these two sources, let $\tilde{C}_t = 1$ for immigrants reported by the police ($\tilde{C}_t = 0$ otherwise). If under-reporting is symmetric between regular and irregular immigrants, then $E(\tilde{C}_t|C_t = 1, L = 1) = E(\tilde{C}_t|C_t = 1, L = 0) = \mu \leq 1$. In this case, differences in reported crimes would have the same sign of differences in the actual crime rates

[17] As it is typically the case, the crime rate of females is very close to zero (at least as far as serious crimes are concerned).

between the two groups, and bound their magnitude from below:

$$E(\tilde{C}_t|L=1) - E(\tilde{C}_t|L=0) = \mu\left[E(C_t|L=1) - E(C_t|L=0)\right].$$

This is likely the case in our data, as 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 alternative measures to incarceration, such as home detention). Moreover, violations of migration law do not constitute a serious crime, so differences in the observed crime rates between regular and irregular immigrants do not depend on the fact that the latter can be reported for being unofficially 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, undocumented immigrants have a better probability of hiding their true identities, as they do not carry official documents. Indeed, this phenomenon is well known in Italy. For instance, Barbagli (2008) reports the case of an undocumented immigrant arrested for having committed a crime who subsequently confessed to have used 15 aliases in previous years (Table A1 in the Web Appendix).

For all of these reasons, the matched data set would overestimate the crime rate of legal immigrants relative to that of undocumented immigrants, $E(\tilde{C}_t|C_t = 1, L=1) > E(\tilde{C}_t|C_t = 1, L=0)$, thus biasing the estimates toward finding a positive effect of legal status on crime. Therefore, 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.

In summary, 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.

3.4 Implementation

The information reported on each application (specifically, the country of origin, the province of destination, and the type of permit) 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.[18]

The largest lotteries were those for type-A applicants 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 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 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 effect of legal status (the slope parameter).[19] We will also perform a permutation exercise to exclude the possibility that our estimates average a small sample bias across lotteries.

The running variable is the delay of each application relative to its lottery's cutoff. In contrast to most RD designs in which treatment assignment depends on a predetermined threshold rule, the cutoff in this case is unknown ex-ante, depending on the realized timing and outcome of all applications submitted in the same lottery. On the one hand, this feature reinforces the fundamental identifying assumption that applications are as good as randomly assigned near the cutoff, on the other it raises the issue of its definition.

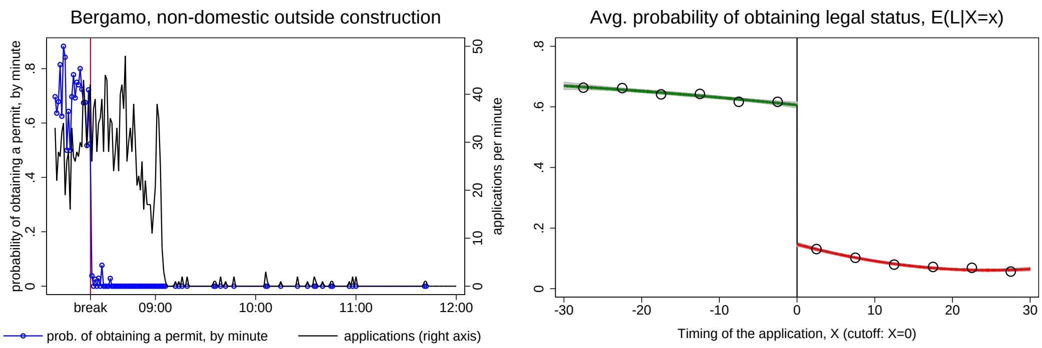
When the probability of obtaining a residence permit immediately drops to zero, as in Figure 3, the cutoff would coincide with the timing of the last accepted application. This situation actually occurs for most lotteries. In a few cases, however, the acceptance rate falls to a low but still positive level for several minutes. The lottery for B1 permits in the province of Bergamo (a city near Milan) provides an example. The probability of acceptance declines from 70% to 5% at approximately 8:22 am, but it reaches zero only a few minutes later; see the left graph in 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

[18] Specifically, the system blocked most applications from Sri Lankan immigrants due to the unusual length of their names. The issue was brought up also in Sri Lankan news, see http://www.lankanewspapers.com/news/2007/12/22962.html. Notice that the article also makes reference to the black market for residence permits, discussed in Section 2.2 above.

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

in which the next applicants were re-contacted sometimes subverted the initial ordering.

Figure 4: Non-compliance with treatment assignment



Note: The left 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 type-B applicants (non-construction, B1) in the province of Bergamo. The vertical line shows the timing of the structural breakpoint according to the Andrews (1993) test. The right graph shows the average probability of obtaining legal status conditional on the timing of application X across all lotteries, with $X = 0$ being the estimated cutoff. 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.

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 subversion – for instance, because of voluntarily manipulation by the immigration officials – can be easily accommodated in the fuzzy RD design, as the Wald-type estimators in (4)-(5) and (6)-(7) exploit the variation in legal status that depends only on the initial ordering of applications near the cutoff (as opposed to later subversions of such ordering).

However, in the presence of fuzziness on both sides the timing of the last accepted application is no longer an accurate measure of the cutoff, since the probability of obtaining legal status could have dropped significantly several minutes earlier (as in the case of Bergamo, Figure 4). Confronted with the 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 treatment assignment on a 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 an unknown breakpoint. The test identifies the "most likely break point" and allows us to test for the significance of the structural break.

In the left graph 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 8:39:06 and the majority of quotas were exhausted before 9:00 (the distribution of cutoff points across applicants is reported in Figure A2 of the Web Appendix).[20]

The right graph in Figure 4 plots the average probability of obtaining legal status conditional on the timing of the application X across all lotteries, with $X = 0$ being the estimated cutoff. The circles are average probabilities across five-minutes bins of X to the left and right of the cutoff, whereas the solid line and shaded area are the predicted outcomes and associated confidence intervals, respectively, based on a quadratic polynomial regression (i.e., equation 5 with $J = 2$): this will be the first stage of the fuzzy RD. Based on this regression, the probability of obtaining legal status decreases on average by about 50 percentage points at the lottery cutoff $X = 0$.

4 Results

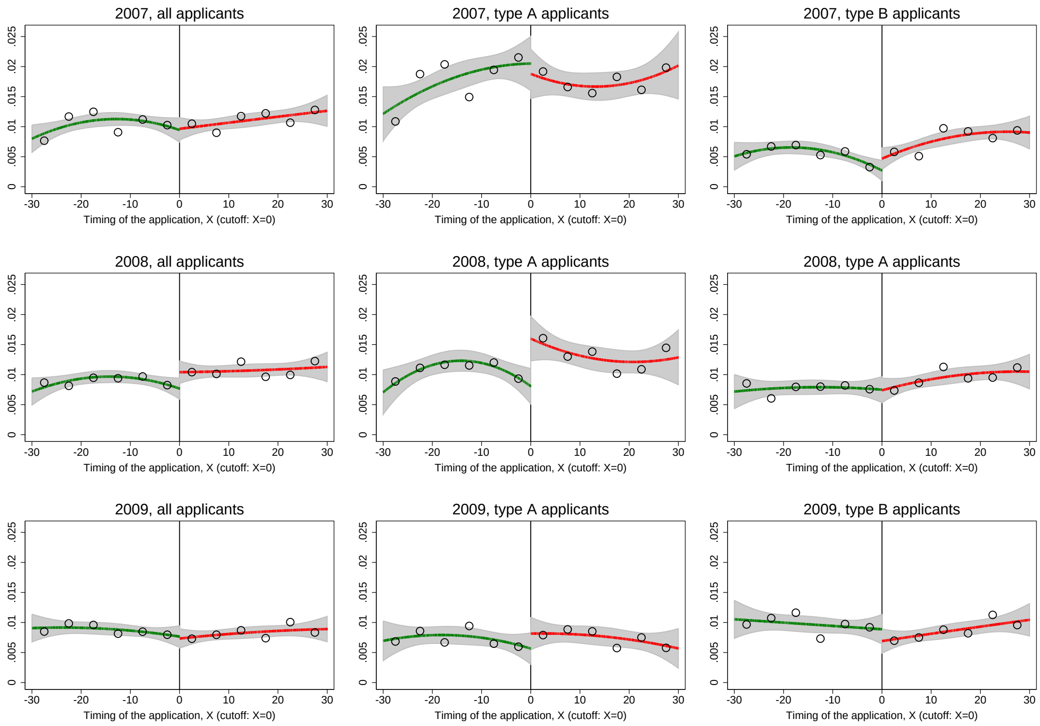
The RD approach allows for a transparent graphical representation of the effect of interest. We thus begin with this analysis and then move to parametric and non-parametric regression methods.

4.1 Baseline estimates

Figure 5 plots the crime rate of Click Day applicants in the years before and after Click Days, conditional on the timing of the application X, for different types of applicants. The circles are averages across five-minutes bins of X to the left and right of the cutoff, whereas the solid line and shaded area are the predicted outcomes and associated confidence intervals, respectively, based on a quadratic

[20] Card et al. (2008) argue that search bias when estimating the unknown cutoff leads to over-reject the null hypothesis that the discontinuity is not different from zero. This is an issue when the discontinuity in the running variable is the outcome of main interest, however this is not the case of our analysis. The main implication of search bias, in the present context, is that we would in the sample some lotteries with no discontinuity in the probability of obtaining legal status. This would bias the estimated effect of legal status toward zero.

Figure 5: Number of crimes per applicant before and after the Click Days, conditional on the timing of application



Note: The graphs show the average number of crimes per applicant committed during the years before and after Click Days, conditional on the timing of application. The scatterplots are averages within five-minute bins, and the solid line and shaded areas are the predicted outcomes and associated confidence intervals, respectively, based on quadratic polynomial regressions. The graphs in the top, middle, and bottom row show the crime rates in the year before Click Days (2007), the year immediately after (2008), and the second year after Click Days (2009), respectively. The left, middle, and right column show the results for all applicants, the subsample of type A applicants, and the subsample of type B applicants, respectively.

polynomial regression – i.e., equation (4) with $J = 2$. The discontinuity in the predicted crime rate at the cutoff equals the reduced form coefficient β in equation (4), which is reported in Table 3 together with the baseline crime rate for those applying after the cutoff, $E(C_t|Z = 0)$.

Figure 5 and Table 3 convey several interesting results. First, in the year before Click Days (2007) the crime rate is identical – 1.1 every 100 applicants reported for having committed a serious crime – for immigrants applying before and after the cutoff, respectively. This provides indirect evidence in favor of the fundamental identification assumption that, after controlling for a smooth polynomial in the timing of application, applicants are as good as randomly assigned to either side of the cutoff.

Second, in the year after Click Days (2008) the crime rate declines to 0.8

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

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline estimates			**Lottery FE, clustered s.e.**		
	all	**type A**	**type B**	**all**	**type A**	**type B**
Year 2007:						
Reduced form	-0.000	0.002	-0.002	-0.000	0.002	-0.002
	(0.002)	(0.004)	(0.001)	(0.002)	(0.004)	(0.002)
$E(C\|Z=0)$	[0.011]	[0.018]	[0.008]	[0.011]	[0.018]	[0.008]
Year 2008:						
Reduced form	-0.003*	-0.008***	0.000	-0.003	-0.008**	0.000
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
2SLS effect of legal status	-0.006*	-0.013***	0.000	-0.006	-0.013**	0.001
	(0.003)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
$E(C\|Z=0)$	[0.011]	[0.013]	[0.009]	[0.011]	[0.013]	[0.009]
First stage coeff. of Z	0.459***	0.610***	0.372***	0.463***	0.621***	0.374***
	(0.007)	(0.010)	(0.009)	(0.029)	(0.032)	(0.041)
first stage F-stat. for Z	4304.08	3492.26	1645.23	248.73	369.39	83.15
Year 2009:						
Reduced form	0.000	-0.002	0.002	0.000	-0.003	0.002
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
$E(C\|Z=0)$	[0.008]	[0.008]	[0.009]	[0.008]	[0.008]	[0.009]
Observations	110,337	40,451	69,886	110,337	40,451	69,886

Note: This table reports reduced form and 2SLS estimates of the effect of legal status on the crime rate of Click Day applicants in the years before and after legalization. The dependent variable is a dummy $C = 1$ for applicants that committed at least one serious offense in a given year (2007, 2008, or 2009). The explanatory variable of interest is a dummy $L = 1$ for applicants obtaining legal status in year 2008 (most of the other applicants obtained legal status in 2009). The first stage instrument is a dummy $Z = 1$ for having applied before the cutoff time at the Click Day 2007. The first stage coefficient of Z in the 2SLS regression for year 2008 and the F-statistic for the excluded instrument adjusted for heteroskedastic and clustered standard errors (i.e., the Kleibergen-Paap statistic) are also reported. 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 (4)-(6) also include lottery-fixed effects and a quadratic polynomial in age. Robust standard errors are reported in parentheses, and they are clustered by lottery in columns (4)-(6). *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively. The mean of the dependent variable for $Z = 0$ in each year is reported in square brackets.

for immigrants applying on time, while it remains at 1.1 for other applicants. Therefore, the reduced form effect (-0.3 percentage points, significant at the 90 percent confidence level) reflects a decrease in the crime rate of applicants to the left of the cutoff, rather than an increase in the crime rate of late applicants. The 2SLS estimated effect of legal status, obtained as the ratio between the reduced form and first stage coefficients, amounts to a 0.6 percentage point reduction in the crime rate.

Third, the crime rate of immigrants applying after the cutoff also decreases in the following year (2009), when they also obtained legal status, equaling that of applicants that were legalized the year before. This evidence suggests that the effect observed in the year immediately after Click Day is not transitory, nor is it explained by the fact that applicants had to go back to their country of origin to collect the entry visa. If this was the case, in fact, their crime rate would have increased again the following year, possibly above that of applicants that were legalized in year 2009 (and had themselves to collect the visa).[21]

Fourth, the fall in crime rate is entirely driven by type A applicants. These applicants start from a higher crime rate before Click Days (1.7 percent) but exhibit the greatest decrease when legalized (-1.3 percentage points), see the middle column of Figure 5 and column (2) of Table 3. By contrast, type-B applicants are characterized by a lower crime rate before the Click Days (0.7 percent), which remains largely unaffected during the following years, see the right column of Figure 5 and column (3) of Table 3.[22]

In columns (4)-(6) we include lottery fixed effects, a quadratic polynomial in age (the only additional individual characteristic included in our data), and we cluster standard errors by lottery to account for interactions in crime among groups of immigrants who apply for the same lottery. All coefficients are unaffected by the inclusion of fixed effects, suggesting that results are driven neither by aggregation bias across lotteries (discussed in Section 3.4) nor by compositional effects.[23] Figure 6 confirms indeed that applicants close to the cutoff are

[21]Our data include information on residence permits awarded under the 2008 quotas, but not on those awarded under the 2009 quotas. For this reason, we report only the reduced form effect for year 2009.

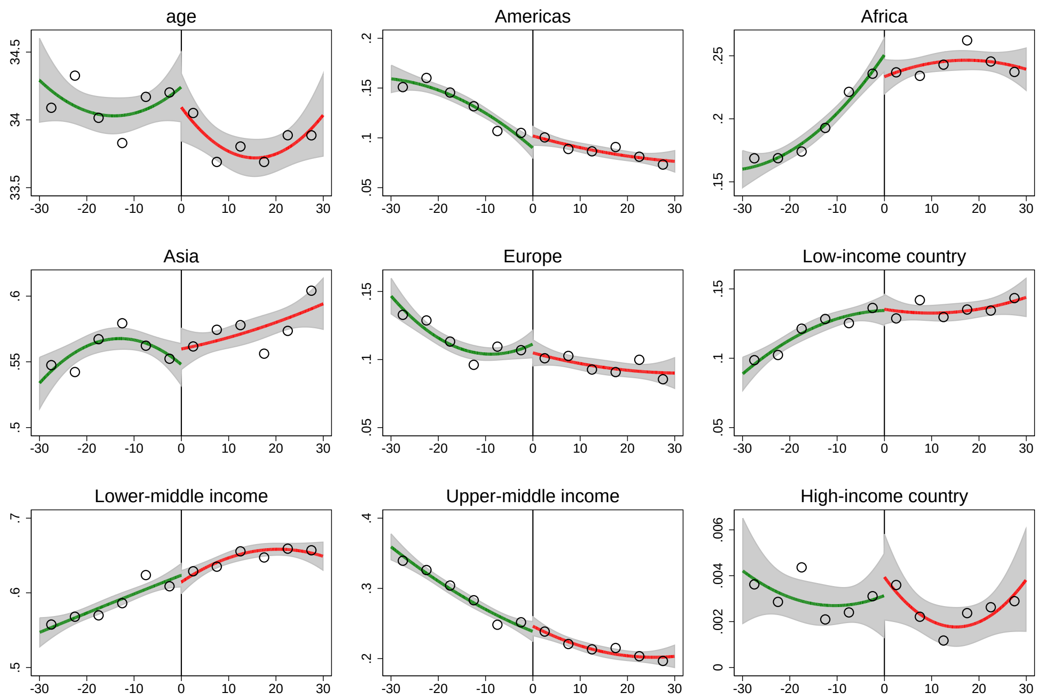
[22]The difference in coefficients between the two groups is statistically significant, see the interaction coefficients reported in the Appendix Table A2.

[23]Lottery fixed effects implicitly absorb province fixed effects interacted with indicators for the 14 privileged nationalities, and with a dummy for the type of permit for applicants of non-privileged nationalities. Table A3 in the Web Appendix shows that results are unaffected when we use either province fixed effects or province fixed effects interacted with a full set of nationality fixed effects, and we cluster standard errors at the corresponding level of aggrega-

on average identical and the empirical density of the running variable in Figure 7 provides additional support for the assumption that legal status is as good as randomly assigned near the cutoff. Consistent with the visual evidence, the McCrary (2008) test does not reject the null hypothesis of no discontinuity in the density of application times.

This evidence suggests that neither manipulation of the running variable nor differences in average characteristics can explain the discontinuity in the probability of committing crimes at the cutoff. This is consistent with the fact that the timing of the cutoff was unknown to applicants during the submission process.

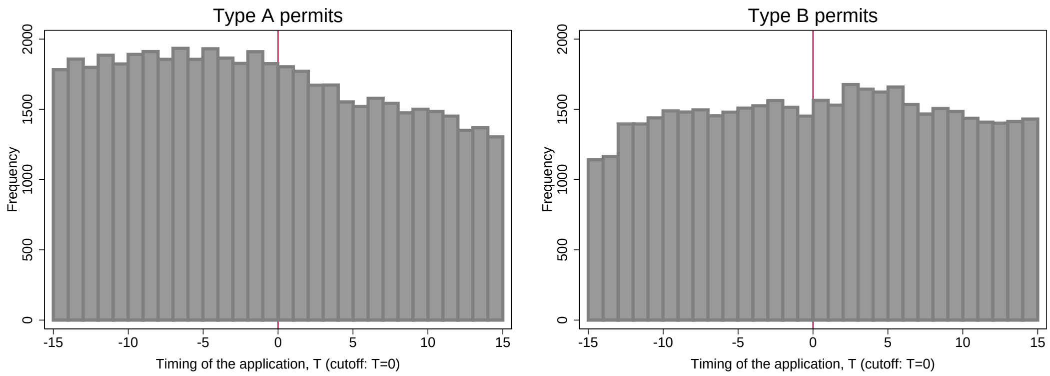
Figure 6: Individual characteristics, conditional on the timing of application



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

tion.

Figure 7: Frequency of applications near the cutoff



Note: The graph shows the density of applications received by type-A and type-B applicants at each minute in time before and after the cutoff.

4.2 Robustness

We examine the sensitivity of the baseline estimates in Table 3 to using non-parametric methods as well as alternative specifications of both parametric and non-parametric regressions. We perform a permutation test to detect the presence of any systematic bias in our estimates, and we further discuss the issue of measurement error.

Starting with the first issue, Table 4 shows the results of 2SLS kernel local linear regressions on equations (6) and (7). To choose the bandwidth, we follow the theory-based criteria proposed by Imbens and Kalyanaraman (2012, IK2012 henceforth) and Cattaneo et al. (2014, CCT2014 henceforth). The criterion used in each specification, the size of the bandwidth, and the number of observations included within the bandwidth are reported in the table. The size of the bandwidth remains very similar when applying these two criteria. The 2SLS estimated effect in year 2008 is around -1.0 percentage point, larger in magnitude than the parametric estimate and statistically significant at the 95% confidence level (columns 1 and 4).

When distinguishing between different groups of applicants, the crime rate declines significantly for type-A applicants obtaining legal status (columns 2 and 5), and the magnitude of the effect is virtually identical to the estimate obtained using parametric methods (Table 3). The absence of any effect on the crime rate of type-B applicants is also confirmed (columns 3 and 6).

The consistency of parametric and nonparametric estimates requires, respec-

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

	(1)	(2)	(3)	(4)	(5)	(6)
	Bandwidth: IK2012			**Bandwidth: CCT2014**		
	all	**type A**	**type B**	**all**	**type A**	**type B**
Year 2007:						
Reduced form	-0.001	-0.003	0.002	-0.001	-0.004	0.001
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)
size of bandwidth	10:20	25:43	11:15	7:27	9:51	7:31
obs. inside bandwidth	45,472	36,362	31,160	33,310	15,905	21,371
Year 2008:						
Reduced form	0.004**	0.008**	0.001	0.005**	0.009**	0.002
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
2SLS effect of legal status	-0.008**	-0.013**	-0.003	-0.010**	-0.014**	-0.005
	(0.004)	(0.005)	(0.006)	(0.005)	(0.006)	(0.007)
size of bandwidth	9:26	10:52	9:55	6:37	8:17	7:57
obs. inside bandwidth	41,743	17,378	27,801	29,805	13,454	22,532
Year 2009:						
Reduced form	-0.000	0.000	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
size of bandwidth	27:54	11:11	11:32	10:15	8:28	7:34
obs. inside bandwidth	105,033	17,917	31,903	45,150	13,747	21,530
Observations	110,337	40,451	69,886	110,337	40,451	69,886

Note: This table reports reduced form and 2SLS non-parametric estimates of the effect of legal status on the crime rate of Click Day applicants in the years before and after Click Days. The dependent variable is a dummy $C = 1$ for applicants that committed at least one serious offense in a given year (2007, 2008, or 2009). The explanatory variable of interest is a dummy $L = 1$ for applicants obtaining legal status in year 2008 (most of the other applicants obtained legal status in 2009). The first stage instrument is a dummy $Z = 1$ for having applied before the cutoff time at the Click Day 2007. 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, the size of the bandwidth (expressed in [minutes:seconds]), and the number of observations included in the bandwidth. Robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

tively, that controlling for a flexible polynomial in X or restricting the sample 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 regarding these trade-offs, the graphs in Figure 8 plot the estimated 2SLS coefficient of legal status and the associated confidence intervals for different specifications of parametric and non-parametric regressions. In particular, the first two graphs plot the results for type-A and type-B applicants, respectively, of the global polynomial regression when varying the order of the polynomial between 0 and 6; the last two graphs plot the results of local kernel regressions for each possible bandwidth between 1 and 30 minutes (the IK2012 and CCT2014 bandwidths are also indicated on the horizontal axis). The estimates are remarkably similar when adopting different specifications of parametric and non-parametric methods.[24]

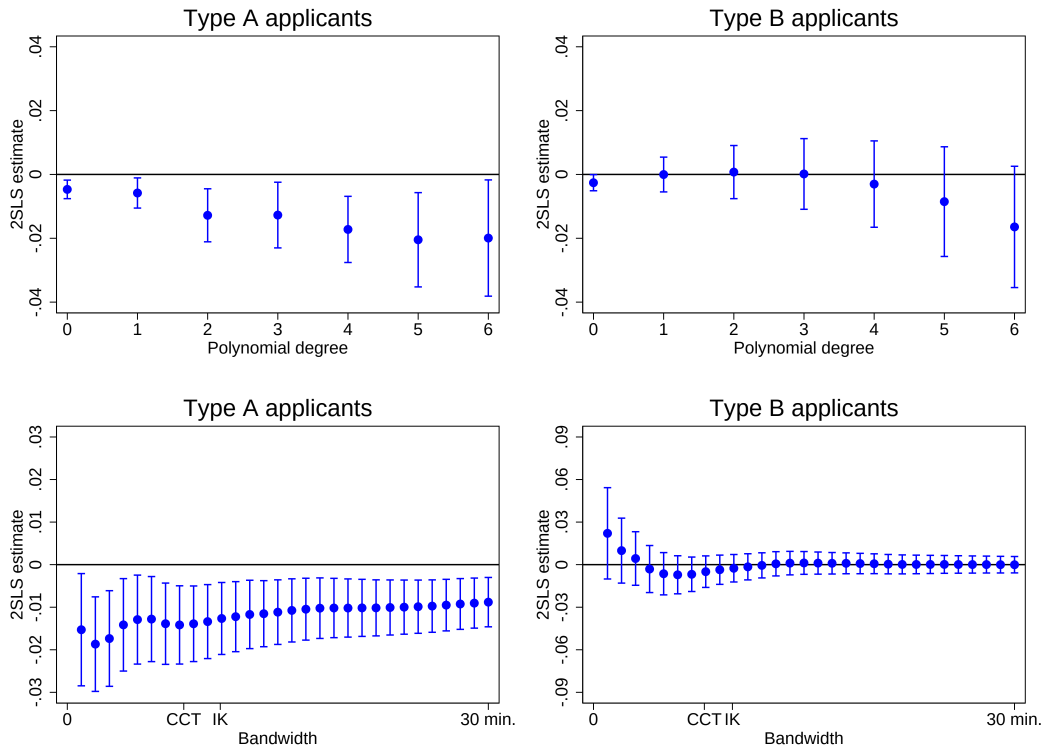
Figure 9 compares such estimates with a distribution of 1,000 placebos. Each placebo estimate is obtained by permuting the cutoff points at random across lotteries, computing a "fake" running variable $\tilde{X}$ as the difference between the timing of each application and the placebo cutoff, and estimating the discontinuity in the probability of committing crimes at $\tilde{X} = 0$.[25] The distributions of the parametric and non-parametric placebos are centered at zero, and the probability of obtaining values below the estimates at the true cutoff for type-A applicants (i.e., the estimates reported in Table 3) is as small as 0.003. 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 coefficient. In particular, these results appear to exclude the possibility that the baseline estimates only average a small sample bias across lotteries.

Finally, we consider a further source of bias, namely, errors in the match between permit applications and criminal records. As discussed in Section 3.3, random matching errors and the intentional misrepresentation of identity by undocumented immigrants imply that the (negative) estimated effect of legal status should be biased toward zero. The possibility of an opposite bias arises

[24] The same is true when looking at the reduced form effect for all years (2007, 2008, and 2009), see Figures A3 and A4 of the Web Appendix.

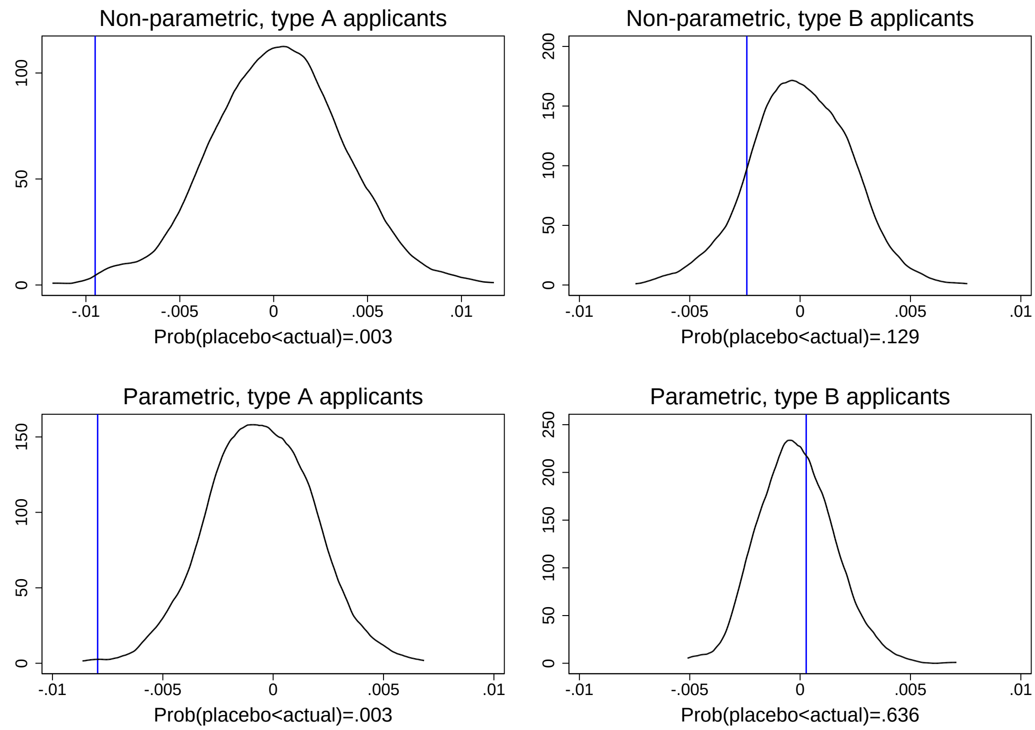
[25] We consider only the reduced-form coefficients because the placebo cutoffs have, by construction, no predictive power for the probability of obtaining a residence permit, so the first coefficients of the placebo regressions would be zero.

Figure 8: Sensitivity analysis



Note: The graphs show the 2SLS estimated coefficients of legal status for different specifications of parametric and non-parametric regressions, distinguishing between type-A and type-B applicants. In particular, the two graphs on the top show the point estimates and associated confidence intervals when varying the degree of the parametric polynomial regression between 0 and 6. The two graphs on the bottom show the point estimates and confidence intervals when varying the bandwidth of the non-parametric regressions between 1 and 30 minutes; the bandwidths selected according to the IK2012 and CCT2014 criteria are also reported on the horizontal axis.

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



Note: The graphs show the distribution of the reduced-form estimates obtained for 1,000 random permutations of the cutoff points across lotteries using parametric and non-parametric methods and distinguishing between type-A and type-B applicants. The vertical lines represent the estimates obtained at the true cutoff point (i.e., the estimates reported in Tables 3 and 4). The fraction of placebo estimates lower than the actual estimate is also reported at the bottom of each graph.

if irregular 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 the analysis to immigrants who were *not* reported for irregular status. On the contrary, the estimated coefficient of legal status remains statistically significant and very similar in magnitude after we exclude immigrants who were (also) reported for violating the migration law.

4.3 Channels

Two main conclusions can be drawn from the results obtained so far. First, legal status decreased the crime rate of immigrants between 2007 and 2008. Second, the average coefficient hides an extreme heterogeneity by the type of applicant. In particular, the effect is high in magnitude and very precisely estimated for type-A applicants, whereas there is no effect on the crime rate of type-B applicants. All estimates are extremely stable when using different methods and specifications. As discussed in Section 2.2, the former type of applicants are characterized by worse labor market opportunities, with a large number of them being unemployed, whereas the latter are likely employed, although unofficially, by sponsor firms before the Click Days.

A potential explanation for the absence of any effect on the crime rate of type-B applicants is thus that they have a very high opportunity cost of crime also before legalization. This cost would include the salary paid by the firm as well as the possibility of re-applying in the future in case the application in 2007 was not successful. By contrast, type-A applicants confront meager employment opportunities before the Click Days. At the same time, legal status triggers a significant change in the behavior of type-A individuals, possibly because improved opportunities following the concession of legal status induce immigrants that were previously detached from the labor market to search for a job. Put differently, the group of type-A applicants may include a large number of individuals at the margin between pursuing a criminal career and entering the labor market.

To further investigate the importance of labor market opportunities, we exploit additional heterogeneity in the type of offense and sponsor.[26] In particular,

[26]The estimates reported in this section are obtained using the parametric specification in

Table 5: heterogeneous effects of legal status by type of offense and nationality of the sponsor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	type A applicants				**type B applicants**			
	type of crime		**sponsor**		**type of crime**		**sponsor**	
	econ	**viol**	**foreign**	**native**	**econ**	**viol**	**foreign**	**native**
2SLS effect of legal status	-0.009**	-0.002	-0.015**	-0.008	-0.000	-0.001	-0.003	0.002
	(0.004)	(0.002)	(0.007)	(0.007)	(0.004)	(0.002)	(0.006)	(0.008)
Observations	40,451	40,386	23,390	16,859	69,886	69,840	30,441	39,308

Note: This table reports parametric 2SLS estimates of the effect of legal status on the crime rate of Click Day applicants in the year after Click Days by type of applicant, offense, and nationality of the sponsor. The dependent variable is a dummy $C = 1$ for individuals committing at least one serious offense in 2008. The explanatory variable of interest is a dummy $L = 1$ for those obtaining legal status in 2008, 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, lottery-fixed effects, and a quadratic polynomial in age. Robust standard errors clustered by lottery are reported in parentheses. *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

if 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, we should expect that the average effect is driven by economically motivated crimes (Machin and Meghir, 2004). This is exactly the picture that emerges from Table 5. Distinguishing between economically motivated crimes (thefts, robberies, drug-trafficking, extortion, and smuggling) and purely violent crimes (murders and sex offenses), legal status only affects the former type of offenses.[27] This last result is consistent with previous evidence in Baker (2013), Freedman et al. (2013), and Mastrobuoni and Pinotti (2015).

In Table 5 we also distinguish between applicants sponsored by foreign and native employers. As discussed in Section 2.2, fictitious job offers aimed at legalizing unemployed immigrants who otherwise would not qualify for a residence permit typically come through type-A applications sponsored by other immigrants. Indeed, the negative coefficient estimated across type-A applicants is entirely driven by this sub-sample of individuals – approximately 23.5 thousand of a total of 40.5 thousand type-A applicants.

column (6) of Table 3 (i.e., second-order polynomial regression controlling for lottery fixed effects and clustering standard errors by lottery). As it was the case for the baseline estimates, the results obtained by non-parametric methods are very similar and are available upon request.

[27] We excluded two types of crime – kidnappings and illegal carrying of arms – that cannot be exclusively classified as either economically motivated or violent.

5 Conclusions

Irregularity imposes a heavy toll on foreign immigrants in terms of poorer employment opportunities, lower incomes, and lower access to social services. These conditions lower the opportunity cost of crime, potentially increasing the propensity of undocumented immigrants to engage in crime. The present paper documents this effect for immigrants applying for a residence permit in Italy.

We believe that our results can inform the current debate on immigration policy. Immigration amnesty is near the top of the agenda in the United States as well as in many other countries due to the large presence of undocumented immigrants. The present paper highlights a potentially important payoff from amnesty programs, namely, a reduction in the number of crimes committed by legalized immigrants in a destination country. The heterogeneity in the effect across different types of applicants also sheds light on which groups would be most responsive to legalization – those with worse employment opportunities when irregular.

Of course, these conclusions are subject to a number of caveats. First of all, the results obtained for Italy cannot be immediately generalized to countries characterized by a very different institutional context. In particular, undocumented immigrants in the United States confront a higher risk of being expelled from the country conditional upon committing an offense, which may lower their propensity to engage in crime. Indeed, Butcher and Piehl (2007) argue that selective deportations are the main reason why immigrants exhibit *lower* incarceration rates than natives in the United States, contrary to what happens in most other countries. In particular, in all European countries immigrants are generally characterized by a high crime rate relative to natives, limited access to employment opportunities, and a very low risk of expulsion.

Second, even within Italy, the results of our analysis cannot be generalized to the whole population of immigrants, as some of them may in principle decide not to apply for a residence permit. In practice, however, the number of applications received on Click Days is very close to the estimated number of immigrants unofficially present on the Italian territory. For instance, the "high-quality" estimate of the undocumented population provided by the Clandestino Project for 2006 – just before the Click Days – reached 650 thousand, just above the 610 thousand applications for residence permits received the following year (see Clandestino Project, 2009, and Table 1). Moreover, those actually applying for legal status may be the most interesting population from a policy perspective, as legalization programs typically presume the active involvement of to-be-legalized immigrants as well as their employers.

Third, in light of previous evidence on the effects of legal status, we emphasize the role of improvements in employment opportunities and wages as a likely driver of changes in criminal behavior by newly legalized immigrants. In principle, legal status could affect the opportunity cost of crime through other channels, notably access to social welfare. In practice, however, undocumented immigrants in Italy have access (with the same conditions as legal residents) to essential services such as basic healthcare and schooling; at the same time, no universal income support scheme exists in Italy (not even for legal residents). For both reasons, legal status should have rather limited effects on this additional dimension.

Finally, each legalization episode potentially generates expectations of further amnesties in the future and thus increases migration pressures at the border. Although previous studies have not found clear evidence of such effects following amnesty programs in the United States (Donato et al., 1992; Orrenius and Zavodny, 2003), a complete evaluation of legalization policies would need to take this possibility into account; we leave this issue to further research.

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.

Butcher, K. F. and A. M. Piehl (2007). Why are immigrants' incarceration rates so low? evidence on selective immigration, deterrence, and deportation. NBER Working Papers 13229, National Bureau of Economic Research, Inc.

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).

Card, D., A. Mas, and J. Rothstein (2008). Tipping and the dynamics of segregation. *The Quarterly Journal of Economics 123*(1), 177–218.

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).

Clandestino Project (2009). Stocks of irregular migrants: Estimates for italy. Technical report, CLANDESTINO Technical Report.

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

Cuttitta, P. (2008). Yearly quotas and country-reserved shares in italian immigration policy. *Migration Letters 5*(1), 41–51.

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

Donato, K. M., J. Durand, and D. S. Massey (1992). Stemming the tide? assessing the deterrent effects of the immigration reform and control act. *Demography 29*(2), 139–157.

Dustmann, C., F. Fasani, and B. Speciale (2010). Illegal migration and consumption behavior of immigrant households. Technical report, Centre for Research and Analysis of Migration (CREAM).

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.

Gelman, A. and G. Imbens (2014). Why high-order polynomials should not be used in regression discontinuity designs. Technical report, National Bureau of Economic Research.

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.

Italian Ministry of Interior (2009). Secondo rapporto sull'attività dei consigli territoriali per limmigrazione. Technical report.

Italian Senate (2011). Informativa del ministro dell'interno sulla regolarizzazione di lavoratori extracomunitari e sui flussi di ingresso. 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 (2015). Legal status and the criminal activity of immigrants. *American Economic Journal: Applied Economics, forthcoming*.

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

Orrenius, P. M. and M. Zavodny (2003). Do amnesty programs reduce undocumented immigration? evidence from irca. *Demography 40*(3), 437–450.

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

Sciortino, G. (1991). Immigration into europe and public policy: do stops really work? *Journal of Ethnic and Migration Studies 18*(1), 89–99.

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.

Wasem, R. E. (2008). Immigration fraud: Policies, investigations, and issues. *Congressional Research Service, Library of Congress*.

Web Appendix – For Online Publication

Figure A1: Screenshot of an application sent through the website of the Ministry of the Interior during a Click Day

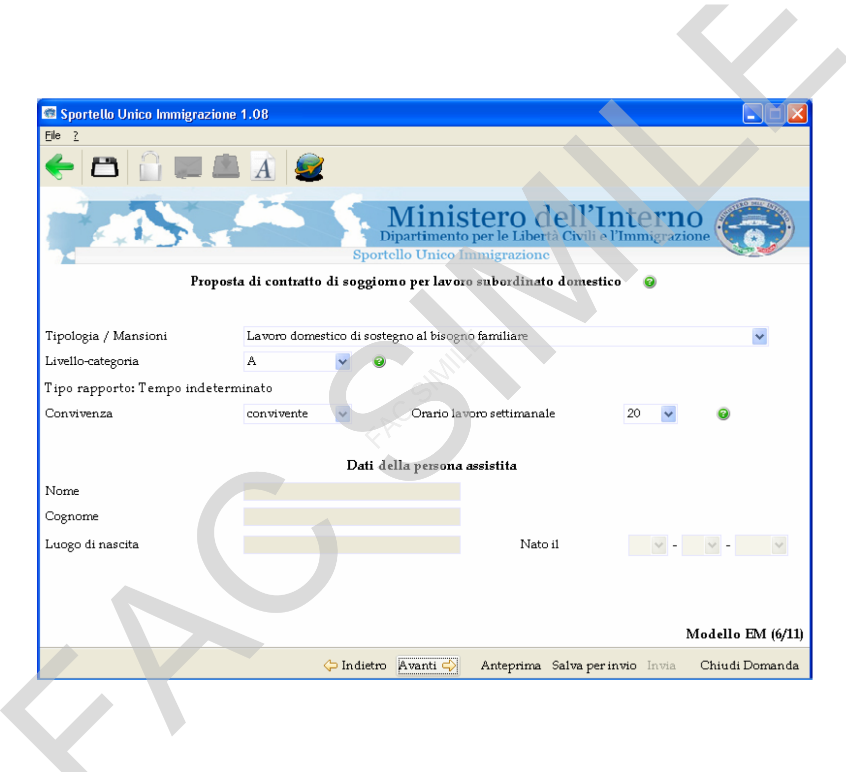


Table A1: Example of the misreporting of identity by an undocumented immigrant

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 several years by an undocumented immigrant arrested in the city of Bologna. This judicial case is described at length by Barbagli (2008).

Figure A2: Density of cutoff points across applicants

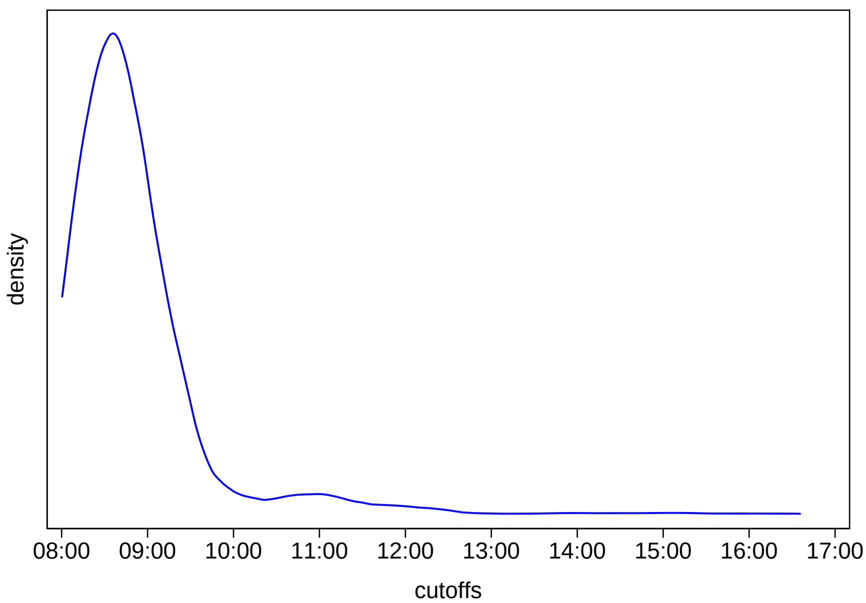


Table A2: Differential effect of legal status on type A and type B applicants, 2SLS estimates

	(1)	(2)	(3)	(4)
	Baseline OLS		**FEs, clustered s.e.**	
Second stage:				
Legal status	-0.006*	0.000	-0.006	0.001
	(0.003)	(0.005)	(0.004)	(0.005)
Type A	0.005***	0.010***	0.001	0.006*
	(0.002)	(0.003)	(0.002)	(0.003)
Legal Status × Type A		-0.013*		-0.014*
		(0.007)		(0.007)
First stage for Legal status:				
Z	0.460***	0.372***	0.463***	0.373***
	(0.007)	(0.009)	(0.029)	(0.041)
Type A	0.030***	-0.083***	0.062***	-0.056**
	(0.007)	(0.007)	(0.013)	(0.026)
$Z\times$ Type A		0.237***		0.246***
		(0.014)		(0.051)
F-statistic (excluded instruments)	4346.45	1645.26	249.91	83.37
First stage for Legal status X Type A:				
Z		-0.000***		0.001
		(0.000)		(0.001)
Type A		0.093***		0.082***
		(0.005)		(0.018)
$Z\times$ Type A		0.610***		0.619***
		(0.010)		(0.032)
F-statistic (excluded instruments)		3488.91		367.89
Observations	110,337	110,337	110,337	110,337

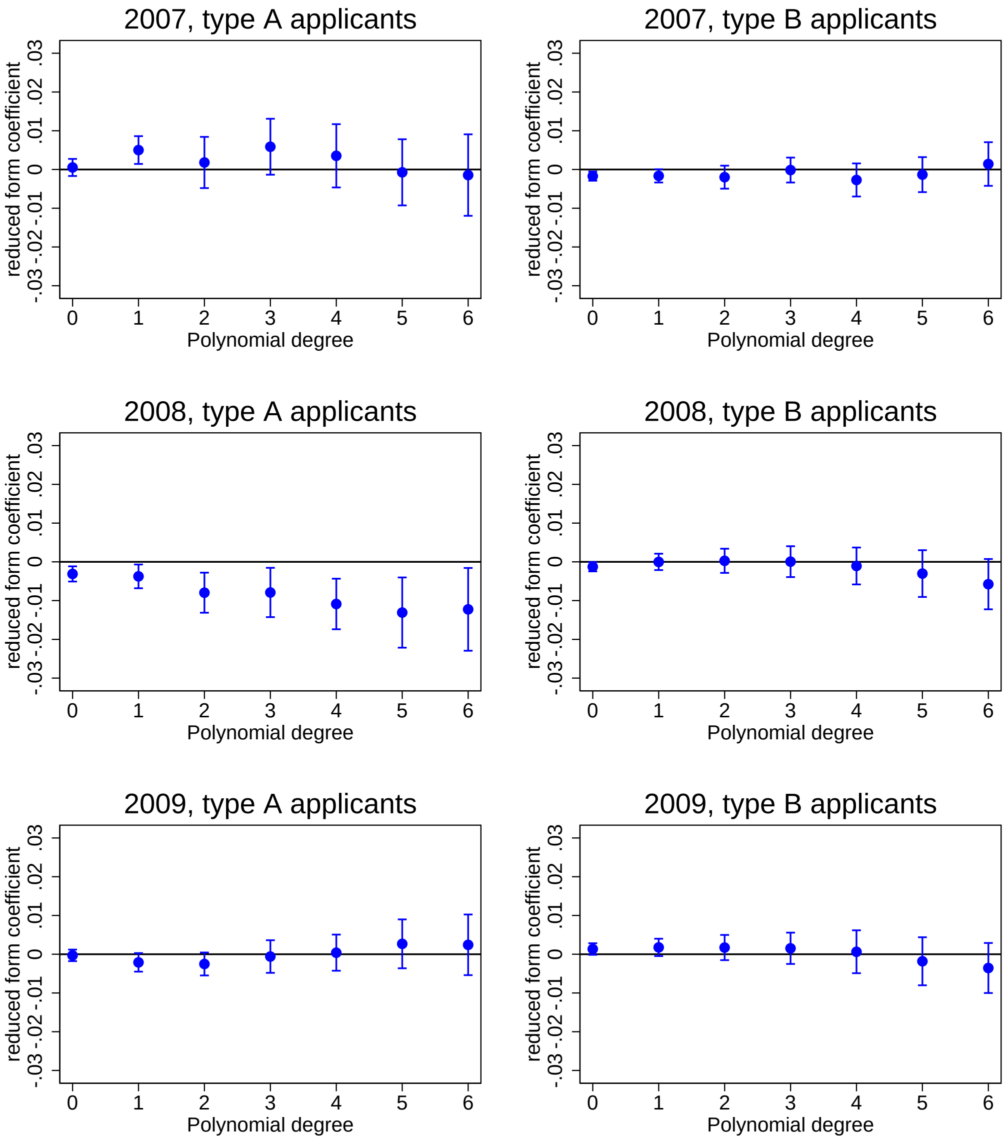
Note: This table reports the 2SLS estimated effect of legal status on the crime rate of Click Day applicants. The dependent variable is a dummy $C = 1$ for individuals that committed at least one serious offense in the year after Click Days (2008); the explanatory variables of interest are a dummy $L = 1$ for applicants obtaining legal status in year 2008 and its interaction with a dummy for Type A applicants; the first stage instruments are a dummy $Z = 1$ for having applied before the cutoff time at the Click Day 2007 and its interaction with the dummy for Type A applicants. The first stage coefficientsand the F-statistic for the excluded instrument adjusted for heteroskedastic and clustered standard errors (i.e., the Kleibergen-Paap statistic) are also reported. All regressions control for a quadratic polynomial in the time elapsed since the cutoff (by the millisecond) and its interaction with Z, the specifications in columns (2) and (4) further interact the dummy for type A applicants with the polynomial and its interaction with Z, and the specifications in columns (3) and (4) also include lottery-fixed effects and a quadratic polynomial in age. Robust standard errors are reported in parentheses, and they are clustered by lottery in columns (3)-(4). *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

Table A3: Reduced form effect on the number of crimes per applicant, global polynomial regression (robustness)

	(1)	(2)	(3)	(4)	(5)	(6)
	province fixed effects			**province X nationality FE**		
	all	**type A**	**type B**	**all**	**type A**	**type B**
Year 2007:						
Reduced form	-0.000	0.002	-0.002	-0.000	0.002	-0.002
	(0.002)	(0.004)	(0.001)	(0.002)	(0.004)	(0.002)
Year 2008:						
Reduced form	-0.003*	-0.008***	0.000	-0.003	-0.008**	0.000
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
2SLS effect of legal status	-0.006*	-0.013***	0.001	-0.006*	-0.013***	0.001
	(0.003)	(0.005)	(0.005)	(0.003)	(0.005)	(0.005)
Year 2009:						
Reduced form	0.000	-0.002	0.002	0.000	-0.003	0.002
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Observations	110,337	40,451	69,886	110,337	40,451	69,886

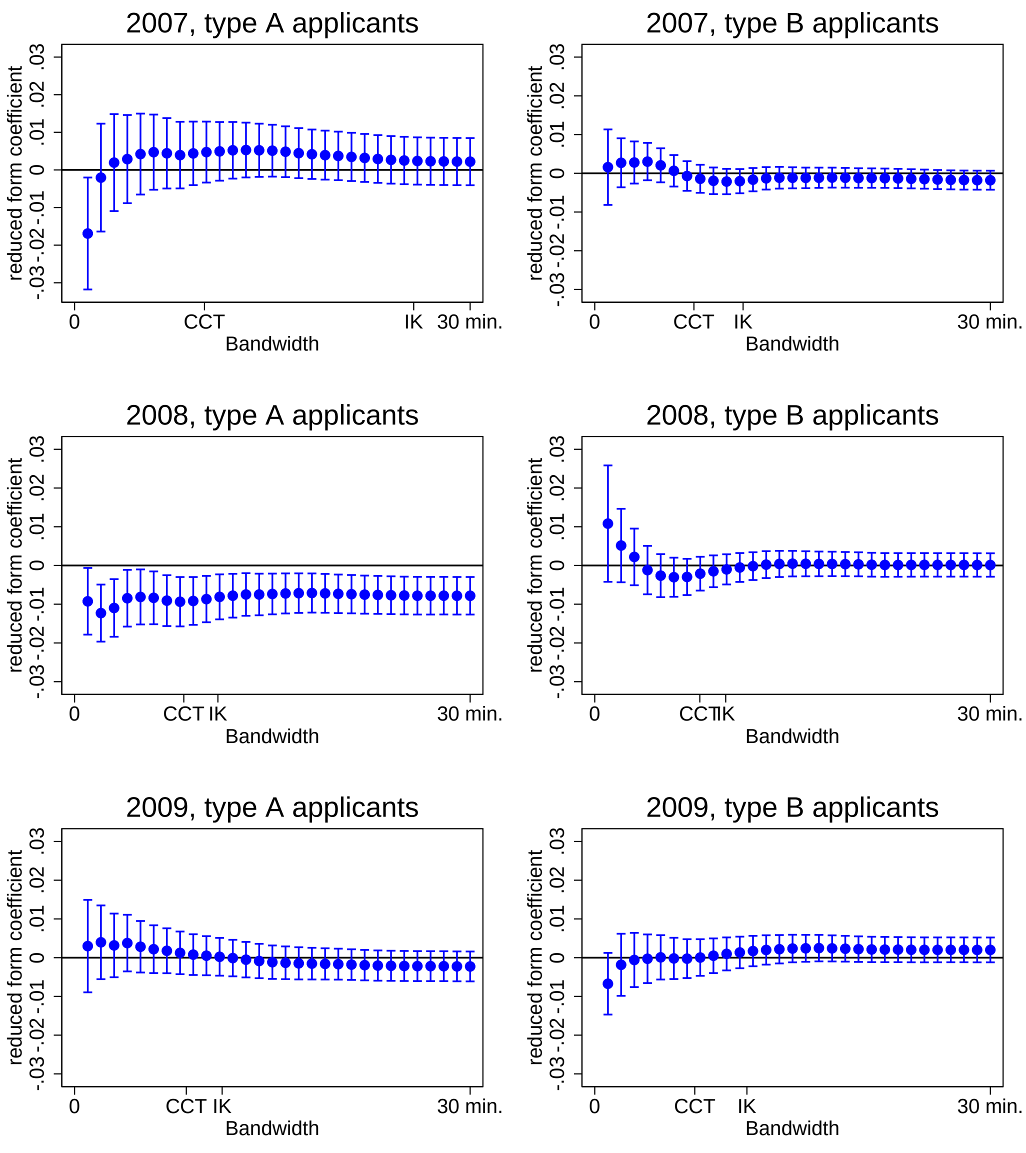
Note: This table reports the estimated reduced form effect of applying on time on the crime rate of Click Day applicants in the years before and after Click Days. The dependent variable is a dummy $C = 1$ for individuals that committed at least one serious offense in a given year (2007, 2008, or 2009), the explanatory variable of interest 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 for a quadratic polynomial in age. The specifications in columns (1)-(3) include province fixed effects and cluster robust standard errors by province, whereas the specifications in columns (4)-(6) include province $\times$ nationality fixed effects and cluster robust standard errors by province $\times$ nationality. *, **, and *** denote statistical significance at the 90%, 95%, and 99% confidence levels, respectively.

Figure A3: Sensitivity analysis, parametric estimates (reduced form)



Note: The graphs show the reduced form coefficients estimated in parametric regressions when varying the degree of the polynomial in the running variable between 0 and 6. The graphs distinguish between the effect in different years and for different types of applicants.

Figure A4: Sensitivity analysis, nonparametric estimates (reduced form)



Note: The graphs show the reduced form coefficients estimated in nonparametric regressions when varying the between 1 and 30 minutes – the bandwidths selected according to the IK2012 and CCT2014 criteria are also reported on the horizontal axis. The graphs distinguish between the effect in different years and for different types of applicants.