The Crime Effect of Refugees*

Mevlude Akbulut-Yuksel†
Dalhousie University and IZA

Naci Mocan‡
Louisiana State University and NBER

Semih Tumen§
TED University and IZA

Belgi Turan¶
TOBB-ETU

October 16, 2022

Abstract

We analyze the impact on crime of 3.7 million refugees who entered and stayed in Turkey as a result of the civil war in Syria. Using a novel administrative data source on the flow of offense records to prosecutors' offices in 81 provinces of the country each year, and utilizing the staggered movement of refugees across provinces over time, we estimate instrumental variables models that address potential endogeneity of the number of refugees and their location, and find that an increase in the number of refugees leads to more crime. We estimate that the influx of refugees between 2012 and 2016 generated additional 75,000 to 150,000 crimes per year, although it is not possible to identify the distribution of these crimes between refugees and natives. Additional analyses reveal that low-educated native population has a separate, but smaller, effect on crime. We also highlight the pitfalls of employing incorrect empirical procedures and using poor proxies of criminal activity, which produce the wrong inference about the refugee-crime relationship. Our results underline the need to quickly strengthen the social safety systems, to take actions to dampen the impact on the labor market, and to provide support to the criminal justice system for mitigating the repercussions of massive influx of individuals into a country, and to counter the social and political backlash that typically emerges in the wake of such large-scale population movements.

*We thank Philip Marx, Ayca Akarcay, Murat Kirdar, Katrin Sommerfeld, and Colette Salemi for useful comments. The usual disclaimer holds.

†Dalhousie University, Department of Economics, Canada. Email: mevlude@dal.ca
‡Louisiana State University, Department of Economics, US. Email: mocan@lsu.edu
§TED University, Department of Economics, Turkey. Email: semih.tumen@tedu.edu.tr
¶TOBB University of Economics and Technology, Department of Economics, Turkey. Email: belgituran@etu.edu.tr

1 Introduction

Because of potential economic and social impacts of international population movements, there has been a long-standing literature in economics analyzing labor market repercussions of immigration. A substantial body of work has investigated the extent to which immigration has an impact on wages and employment prospects of native workers, along with other related questions such as whether undocumented immigrants are negatively or positively selected, the elasticity of substitution between immigrant and native labor, and productivity gains and economic spillovers related to immigration.[1]

To the extent that criminal activity is determined by expected costs and benefits of participation in legal and illicit markets as specified by Becker (1968) and Ehrlich (1973), a natural area of inquiry is the investigation of whether immigration has an impact on crime. Given the substantial economic costs of crime and the political narrative surrounding the immigration-crime nexus, this is an important question to investigate, both scientifically and for public policy.[2] Consequently, a burgeoning literature has analyzed the impact of immigration on criminal activity.

Earlier work on immigration-crime relationship has focused on the U.S. context and reported no significant association between immigration inflows and crime rates of cities, although foreign-born youth were reported to be less likely to engage in crime (Butcher and Piehl, 1998a,b). Recent research had access to data sets better suited to investigate the relationship between immigration and criminal activity, and they took advantage of plausibly exogenous changes in the policy environment in a variety of countries to identify the impact of immigrants on crime.[3]

As international migration continues to be one of the issues at the forefront of political debate in developed countries, a more salient phenomenon over the last decade has been the rise in the magnitude of displaced populations. UNHCR, the Refugee Agency of the UN, reports that there were 82.4 million people who were forcibly displaced worldwide by the end of 2020, 24.6 million of whom were refugees not including Palestinians, and another 4.1 million were asylum seekers (UNHCR, 2021). This large-scale movement of refugees has been a significant issue with its associated political, economic and environmental concerns, especially for low- and middle-

[1] Examples include Battisti et al. (2018), Dustmann et al. (2016), Dustmann et al. (2013), Dustmann et al. (2012), Bratsberg and Raaum (2012), Borjas (2003), Card (2001), and LaLonde and Topel (1991).

[2] Ferraz and Soares (2022) calculate that the expenditures on the criminal justice system and the cost of victimization add up to $450 billion in the U.S., and Anderson (2021) estimates the cost of crime in the U.S. as $3-4 trillion without considering the transfers from the victims to criminals.

[3] Bianchi et al. (2012) reported an increase in robberies in Italy caused by exogenous changes in the Italian immigration population. Piopiunik and Ruhose (2017) took advantage of the exogenous relocation of East Germans after the unification of Germany and showed that the arrival of ethnic German immigrants from the East significantly increased crime. Bell et al. (2013) found that the inflow of asylum seekers in the late 1990s/early 2000s led to a modest rise in property crime in the U.K., but that immigration from the EU accession countries had a small negative impact. Using a panel of U.S. counties, Spenkuch (2014) reported a positive impact of immigrants on crime. Researchers have also shown that better labor market opportunities for undocumented immigrants (e.g., through legalization of their status) had a negative effect on crime both in the U.S. (Freedman et al., 2018; Baker, 2015) and in Italy (Pinotti, 2017; Mastrobuoni and Pinotti, 2015).

income destination countries which hosted 86 percent of all refugees by the end of 2020.[4] Recent influx of more than 5 million Ukrainians into various European countries in the wake of the Russian attack on Ukraine, and the assessment of the United Nations that climate change might lead to displacement of around 20 million people annually, underline that large-scale population movements will likely remain a major policy issue for years to come.

The analysis of the impact of refugees on receiving countries is conducted within the same conceptual framework, although there are key differences between immigration and refugee flows. First, while much of the immigration literature is concerned, implicitly or explicitly, with the flow of undocumented immigrants who have limited opportunities in the labor market because of their illegal status, refugees, for the most part, are provided with legal rights of residency in receiving countries along with some other benefits such as (limited) access to social welfare services. Second, while the nature of self-selection of immigrants from the country of origin to the destination country is debatable, refugees represent large groups of individuals who leave their home countries not by choice, but by necessity, due to traumatic events such as civil war. We discuss the specific theoretical implications of this phenomenon in Section 2. Third, unlike immigration flows, refugee movements generally take place in very large numbers in relatively short periods of time, generating significant shocks to and bottlenecks in a variety of sectors in receiving countries, ranging from labor and housing markets to the provision of key government services. Thus, the impact of refugee movements on socio-economic outcomes, including crime, could be significant.

In this paper, we analyze the impact on crime in Turkey of the influx of Syrian refugees, who were displaced by the war and civil conflict in Syria. We utilize data on refugee inflows into 81 provinces of the country, and on the number of crimes committed each year in each province (as reported to the offices of the prosecutors). Potential endogeneity of refugee location is addressed by estimating instrumental variables models, the results of which reveal a positive and significant effect of the number of refugees on local criminal activity. There were no Syrian refugees in Turkey prior to 2012, while the number reached 3.7 million in 2016, which translated into a 4.5 percent increase in the population of the country. We calculate that this influx of refugees led to 75,000 to 150,000 additional crimes per year. As we describe in the paper, these additional offenses could have been committed by both the refugees and the natives as a result of the interplay between the labor market reactions to the refugee influx and the personal attributes of the refugees as well as natives. Additional analyses reveal that an increase in the number of low-skilled natives (those with an elementary school or middle school education) also has a positive impact on crime, although an increase in refugee population by a given magnitude generates

[4] According to the definition provided by the UNHCR, a refugee is a person who has crossed an international border due to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion.

a larger increase in total crime in comparison to an increase in unskilled native population by the same magnitude. The results are robust to extensive sensitivity analyses ranging from the use of alternative instruments to variations in sample composition, from alternative model specifications to placebo exercises. We also highlight the pitfalls of employing incorrect empirical procedures and using poor proxies of criminal activity, which produce the wrong inference about the refugee-crime relationship.

The rest of the paper proceeds as follows. The next section describes the conceptual framework and the related literature. Section 3 provides the background on the Syrian refugee inflow into Turkey. Section 4 describes data sources and the proper measurement of crime. Section 5 presents the descriptive statistics. Section 6 lays out empirical framework, and illustrates the consequences of incorrect model specification which plagued some recent work in this area. Section 7 presents the main results. The sensitivity analyses and robustness checks are explained in Section 8, and Section 9 presents the results of the translog regressions analyzing the extent to which unskilled natives impact crime. Section 10 concludes the paper.

2 Conceptual framework and the existing literature

Following the seminal work of Becker (1968) and Ehrlich (1973), and their extensions (Lochner, 2004; Mocan et al., 2005), participation in criminal activity is determined by individuals' expected returns from the labor market, captured by such factors as the propensity for employment and relevant market wages for individuals, expected returns to illegal activity, as well as deterrence variables such as the probability of apprehension and the severity of punishment. Also important is the risk aversion and time discount of the decision-maker. Within this framework, the mechanism through which refugees can impact crime has two primary channels. First, the influx of refugees increases the supply of labor in the relevant labor market. If the average human capital of refugees is lower than that of natives, and if jobs with higher skill contents require the aptitude of speaking the language of the host country, refugee labor is expected to influence the market for unskilled labor, increasing the unemployment rate and decreasing unskilled wages. This effect is expected to raise the propensity for crime not only for refugees, but also for native unskilled workers who now face diminished labor market opportunities. Evidence for this phenomenon has been provided by Borjas et al. (2010), who show that black workers in the U.S. increased their criminal activity in reaction to their worsening labor market conditions due to immigration.[5]

[5] If unskilled and skilled workers are complements, depending on the elasticity of substitution, a mitigating factor could be an increase in the demand for skilled workers and the ensuing increase in production and income (assuming the expansion of capital). In other words, if the influx of refugees and its depressing effect on the market for unskilled workers trigger an enhanced demand for unskilled workers, which then generates an increase in production, this could have a dampening effect on the expected rise in crime. A full discussion of the wage effects within a general equilibrium framework can be found in Borjas (2013).

The second channel through which refugees can influence crime is one which is provided by research on the impact of violence exposure. Recent work has identified the impact of negative life shocks and trauma on individuals' risk preference, time discount, and subsequent behavior. For example, Eckel et al. (2009) show that exposure to a severe hurricane and having been forced to evacuate to a different city decreases people's risk aversion. Hanaoka et al. (2018) report that individuals who experienced stronger intensity of the Great East Japan Earthquake of 2011 became more tolerant of risk one year after the earthquake. Page et al. (2014) provide the same result based on the analysis of behavioral changes following a major flood.[6] Voors et al. (2012) report the results of field experiments in Burundi and show that exposure to violence impacts behavior, possibly due to changes in preferences.[7] In a related domain, Moya and Carter (2019) find that, among the internally displaced people in Colombia, those who are exposed to more severe violence are more likely to predict facing extreme poverty in the following year, and they are more likely to believe in diminished prospects of economic mobility. Similarly, Rohner et al. (2013) show that those who are exposed to more intense fighting during the ethnic conflict in Uganda between 2002 and 2005 have lower generalized trust and heightened ethnic identity.

There is also evidence indicating that people's exposure to conflict and violence makes them more prone to being violent themselves. Miguel et al. (2011) show that all else the same, soccer players in professional leagues of Europe, who are from countries which experienced civil conflict, behave more violently on the soccer field. Noe and Rieckmann (2013) find that exposure to violent civil conflict in Colombia increases women's risk of victimization by domestic violence by partners who presumably were exposed to the same violence. The same impact on domestic violence stemming from exposure to civil war atrocities has been reported in Rwanda (La Mattina, 2017) and Peru (Gutierrez and Gallegos, 2016). More directly related to our work, Couttenier et al. (2019) analyze a unique data set of all crimes reported in Switzerland, which includes information on the nationalities of the perpetrators and the victims. Focusing on asylum seekers, the authors find that cohorts which are exposed to civil conflicts or mass killings are 35 percent more likely to commit violent crime in Switzerland.

In summary, an increase in refugee population, especially a sudden inflow, is expected to have a non-negative impact on the crime *rate* of the host country, unless the baseline proclivity for criminal activity is lower for refugees in comparison to natives, and the inflow of refugees has no impact on the labor markets. The *average* crime rate of refugees in the host country could be lower than that of the natives if expected sanctions are stiffer for refugees. One such example is the policy of deportation which sends refugees to their country of origin if refugees are convicted

[6]It should be mentioned that there also exists research which reports the opposite result. For example, Cassar et al. (2017) find that the 2004 tsunami led to an increase in risk aversion in rural Thailand.

[7]While the authors found that people who experienced violence or who live in communities that have been violently attacked displayed more altruistic behavior, they also reported that such individuals were more risk-seeking, and they had higher discount rates.

of a crime. On the other hand, the average refugee crime rate could be higher than the crime rate of natives if the former group faces worse labor market opportunities (due to lower skills, discrimination, lack of language skills or some other reason). Refugees' marginal propensity for crime could be higher, especially for violent crimes, if their exposure to violence in the country of origin had an impact on their risk aversion, time preference; or if violence exposure has altered their behavioral patterns in some way, as discussed above. Thus, it is unclear *a priori* whether and how the baseline crime rates would be different between natives and refugees. Importantly, and relevant for our research, no matter what the benchmark initial crime rate is for either group, there are theoretical reasons to expect an increase in crime rates of both natives and refugees in response to an inflow of refugees. This is because an increase in the number of refugees, especially by the magnitude analyzed in this paper, is expected to have detrimental effects in the labor market for unskilled labor, which will likely influence the marginal native criminal as well.

Refugees and the labor market. The impact on labor markets of refugee inflows has been analyzed using the arrival of Cubans to Miami (Anastasopoulos et al., 2021; Peri and Yasenov, 2019; Borjas, 2017; Card, 1990), the entry of Albanians to France in 1962 (Hunt, 1992), the movement of Soviet Jews to Israel in the early 1990s (Friedberg, 2001; Borjas and Monras, 2017), and the population movements after the Balkan War in the 1990s (Angrist and Kugler, 2003; Borjas and Monras, 2017).[8]

Recent work has demonstrated the impact of Syrian refugees on the Turkish labor market, which is the context of this paper. Syrian refugees in Turkey are less educated than natives. For example, while 33 percent of natives have at least a high school degree, the rate is 5.5 percent among Syrian refugees [Tumen (2018), Table 1]. Syrian refugees enter the Turkish labor market through informal manual jobs and displace natives who are employed in those jobs (Del Carpio and Wagner, 2015; Tumen, 2016; Bagir, 2018; Ceritoglu et al., 2017; Aksu et al., 2022; Altindag et al., 2020).[9] It has also been shown that informally employed refugee workers provide labor cost advantages to firms, and that they suppress wage growth in the lower segment of the labor market (Balkan and Tumen, 2016). Informal refugee workers, employed in manual tasks, are complementary to formal native workers employed in more complex tasks (Akgunduz et al., 2018; Akgunduz and Torun, 2020), and formal wages have moderately increased in response to refugee inflows (Tumen, 2016). These findings suggest that competition between refugee and native workers for low-skill jobs imposes a downward pressure on employment probabilities and

[8] See Clemens and Hunt (2019) for a summary and re-analysis of these papers.

[9] In contrast to previous findings in the literature, Cengiz and Tekguc (2022) depict a more optimistic picture about the impact of refugees on labor market outcomes of natives. The authors write that "...the native lower-skilled workers in Turkey experienced small wage and employment losses after the Syrian migration, whereas the higher-skilled workers have seen gains." Different from most papers in the literature, the authors use aggregated (region-level) data, which may mask the underlying/true causal effects. Time-varying shocks in their specification come from differentiated trends across industries, which is also substantially different from the convention in the literature that relies on a more general structure of time-region interaction terms.

potential wages in the lower segment of the labor market. Conversely, the increased availability of formal jobs with higher skill requirements has also been reported. It is important to note, however, that the negative impact of refugees on low-skilled natives is stronger than their positive impact on more skilled workers, suggesting an overall negative labor market impact (Ceritoglu et al., 2017). Thus, the refugee effect can be labeled as a low-skilled labor supply shock, rather than a productivity and growth enhancing human capital inflow. These labor market effects, would in turn, impact the propensity to commit crime for the marginal criminal refugee both on the extensive and intensive margins (the probability to commit a crime, and the number of offenses). Furthermore, the increase in the supply of refugees and the resultant labor market adjustments are expected to have a spillover effect on the marginal native criminal in the same fashion.

Literature on refugees and crime. There is limited work on the impact of refugees on crime. Analyzing the impact of the drop in the number of refugees admitted to the U.S. following the Executive Order of President Trump, Masterson and Yasenov (2021) do not find a change in the crime rates of counties that reduced their refugee admissions. The extent to which this finding has external validity for other countries is questionable for two reasons. First, the U.S. admits a small number of refugees each year.[10] Second, refugees entering the U.S. undergo an investigation process conducted by the Department of State and the Department of Homeland Security. Hence, this is a different "treatment" than those experienced by host countries elsewhere in the world, which admitted millions of refugees with urgent humanitarian needs in very short periods of time. For example, Megalokonomou and Vasilakis (2020) analyze data on refugee inflows to Greek islands, and find that an increase in the refugee share of the island population generated a significant increase in property crimes, knife attacks and rapes in comparison to neighboring unexposed islands.

Researchers who focused on the German experience in the wake of the recent refugee inflows using different data sets, time periods, and identification strategies revealed a positive impact of refugees on crime (Gehrsitz and Ungerer, 2022; Lange and Sommerfeld, 2021; Dehos, 2021). Couttenier et al. (2019) used an exhaustive dataset on all crimes in Switzerland over a seven-year period, which includes information on the perpetrators. The authors reported that asylum seekers who were exposed to civil conflict were substantially more prone to violent crime in comparison to the average cohort of asylum seekers. Taking advantage of the data on both victims and perpetrators, the authors also found that past conflict exposure made asylum seekers more crime-prone towards Swiss natives.

[10] For example, in the U.S. the number of refugees was 53,716 between October 1, 2016 and September 30, 2017. The corresponding number was 30,000 in 2019 and 11,814 in 2020. Even before the presidency of Donald Trump, the largest number of refugees over the last two decades was 84,994 in 2016. This amounts to 2 refugees per 10,000 residents, socio-economic impact of which is expected to be negligible.

In contrast to the research summarized above, two recent papers reported *crime-reducing* effects of Syrian refugees in Turkey (Kayaoglu, 2022; Kirdar et al., 2022). These papers use a similar identification strategy to the one we use here. They, however, have major drawbacks, including using poor proxies of crime, and incorrect empirical implementation which lead to the wrong inference.[11] We provide a summary assessment of these papers in Appendix B.

3 The Syrian refugee inflow

Following the internal conflict in Syria, which started in March 2011, millions of Syrians were caught under fire, and they were involved in, or otherwise exposed to atrocities of a violent civil war. Masses migrated from Northern Syria to the Southeastern regions of Turkey. Most of the refugees were originated from Syrian provinces of intense conflict, close to the Syrian-Turkish border. During the initial phase of this crisis, (late 2012, and 2013), most refugees fled to Turkey from the following six Syrian cities which are very close to the Turkish border: Aleppo (residents of which constituted 36% of Syrian refugees in Turkey), Idlib (21%), Raqqa (11%), Latakia (9%), Al-Hasakah (5.4%), and Hama (7.5%). The remaining 10% of refugees arrived from other cities relatively far from the border. Subsequent waves resembled a similar composition.

The Syrian refugee inflows started in early 2012 and accelerated over time (see Figure 1). There were no Syrian refugees in Turkey prior to the civil war, but their numbers have reached 3.7 million by 2021. Initially, refugees were mostly located in the accommodation centers or camps constructed and operated by the Turkish government near the Syrian border. But, over time, the refugee population residing outside of the camps increased sharply, and the camps became almost idle.

Until 2014, refugees were mostly located close to the Turkey-Syria border for two reasons. First, in the early days of the entry, there were hopes of going back to Syria once the crisis was resolved. Second, the Turkish government built large refugee camps along the border regions to provide basic services—such as health, security, food, education, etc. After the mid-2014, following the involvement of Russia in the Syrian civil war, it became clear that the conflict would not end soon. Consequently, refugees started to actively seek permanent homes. Some of them preferred to stay in and around the border regions close to Syria, while others chose to move out of the

[11] First and most important, both papers use the same empirical specification which imposes a mechanical negative association between refugees and the crime rate. Although both papers employ an instrumental variable strategy, the manner in which their empirical design is implemented invalidates the exclusion restriction. These issues are discussed in detail in Section 6. Third, the authors use inaccurate proxies of crime. Kirdar et al. (2022) employ *prison intake* (the number of convicted offenders entering prison each year) as a proxy for crime. As the vast literature in economics of crime reveals, prison intake is not a valid proxy for criminal activity for a number of important reasons (Kuziemko, 2013; Buonanno and Raphael, 2013; Drago et al., 2009; Levitt, 1996). We explain in Appendix B the pitfalls of using prison intake as an indicator of crime, but suffice it to say to say that doing so leads the authors to drastically underreport the crime rate of the country. More specifically, they report the crime rate in Turkey as **196** per 100,000 residents [Kirdar et al. (2022), Table 1, p. 573], when the correct crime rate of the country is 20 times higher. Similarly, Kayaoglu (2022) employs an inaccurate proxy of crime. These measurement problems are summarized in Appendix B.

southeastern regions toward the western regions of the country. Figure 2 displays these location choice patterns. The majority of refugees who lived outside of the camps reported that they left Syria for security reasons and they chose Turkey as their destination because of the close proximity of the border, and ease of transportation to it.

Syrian refugees in Turkey are less educated than natives. About 80 percent of refugees have no high school degree, and about 55 percent have no middle school education (Tumen, 2018). This aggregate human capital profile of refugees is worse than the aggregate Turkish profile where over the period between 2012 and 2016, 52-53 percent of the Turkish population over the age of 15 has primary or junior school degree, 21-22 percent has a high school degree and 13-16 percent of population has university degree or higher, although southeastern Turkey, which received the bulk of the refugee flow, is associated with lower levels of human capital.

It should be noted that, at the early stages of the entry, Syrian refugees were not officially allowed to work in Turkey as registered workers—mostly due to a lack of legal framework recognizing them as refugees or visitors under temporary protection. Hence, refugees initially entered Turkish labor markets through informal jobs. The Turkish labor market, especially the labor market in the Southeastern Turkey and some major metropolitan provinces, offers extensive informal employment opportunities. After 2016, a new policy has allowed issuance of formal work permits for Syrian refugees. However, the number of formally employed Syrians has remained very small due to three main reasons. First, work permits are mostly issued upon the request of firms if they claim that they could not fill a vacancy with native workers, or the refugee worker has a unique skill/talent that the firms need. Refugees are on average much less skilled/educated than natives, so the number of such requests is minuscule. Second, there are some additional bureaucratic barriers that make it difficult for refugees to obtain a work permit, and because of political reasons the government is reluctant to ease out those barriers. Finally, Syrian refugees themselves, regardless of work permit issues, are more willing to be employed informally rather than formally as most of them benefit from cash transfer programs (mainly, the CCTE and ESSN programs).[12] These transfer payments stop once the refugee pursues formal employment. Furthermore, regardless of employment status, refugees have access to health benefits, which are allowed by their temporary protection status in Turkey. Consequently, the number of formally employed Syrians remained ignorable in our period of analysis.

4 Data sources and measurement

Refugee data and control variables. Data on Syrian refugees in each of the 81 provinces of the country are obtained from Turkish Disaster and Emergency Management Presidency

[12] For more information about the CCTE (Conditional Cash Transfer for Education) and ESSN (Emergency Social Safety Net) programs, see: https://data2.unhcr.org/en/documents/download/62207.

(AFAD) for 2013; Erdogan (2014) for 2014; and Ministry of Interior Directorate General of Migration Management for 2015 and 2016. Figure 1 presents the aggregate time-series pattern of the number of refugees. As discussed in the previous section, the number of Syrian refugees was zero prior to 2013, and it rose dramatically since then, reaching a total of almost 3.6 million in 2018.[13] Figure 2 displays the regional and temporal variation of refugee intensity in Turkey, and reveals that refugees were clustered around the Syrian border until 2014 and later they spread beyond the Southern provinces of Turkey.

Province attributes such as native population, native population by education (which is utilized later in the paper), and the number of hospitals per 100k population are obtained from Turkish Statistical Institute (TurkStat). Information on the availability of natural gas is gathered from the Turkish Natural Gas Journal. We also use data on the province-level public expenditures, provided by the Presidency of Strategy and Budget, the Office of the Presidency of the Republic of Turkey.[14] As detailed in Section 6, we do not include to the model variables that gauge the economic conditions of provinces, because such variables are likely endogenous. Following Rose and Shem-Tov (2021), Corman and Mocan (2005), and Levitt (1996), we also control for the lagged value of inmates entering the prison system as a deterrence variable, the source of which is the annual *Prison Statistics* reports of the Turkish Ministry of Justice.[15]

The measurement of crime. Comprehensive crime and judicial data are obtained from publicly available official annual reports of the Ministry of Justice.[16] These reports provide detailed information on the number of crimes and suspects handled by state prosecutors in each province of the country, as well as information on the adjudication process, including the number of offenses and defendants in case files in courts. This detail allows us to cross-validate the profile of criminal activity in the country in various stages of the judicial process, and it also enables us to compare various pieces of information to data provided by international agencies.

Because proper measurement of crime is central to the analysis, we provide a thorough description of its definition and construction. The standard measure of criminal activity is the number of offenses reported to the police, although this measure underrepresents the true incident of crime in any society because not all offenses are reported to the law enforcement agencies.[17] Information on crimes reported to the police are not available in Turkey. However, in most countries, including Turkey, case files of offenses handled by the police are transferred to the

[13] Refugees started entering Turkey in January 2012, but they were retained in refugee camps during 2012.

[14] https://www.sbb.gov.tr/yatirimlarin-illere-gore-dagilimi/.

[15] There are no publicly-available police data at the province level.

[16] Data used in the analysis are extracted from Judicial Statistics between 2006-2018; see, https://adlisicil.adalet.gov.tr/Home/SayfaDetay/adalet-istatistikleri-yayin-arsivi.

[17] Measurement error in reported crime could also be systematic if the propensity to report is correlated with the perception of police efficiency and the clearance rate of particular crimes. An alternative measure of criminal activity can be created using surveys of crime victimization. These surveys provide information about the perpetrator attributes, but victimization surveys are also prone to reporting bias, and they are not systematically available in many countries, including Turkey.

office of the prosecutor, and this information is available from the Ministry of Justice. Figure 3 displays the summary information for Turkey for the year 2013, provided by the Turkish Ministry of Justice annual reports.[18] As mentioned earlier, no information is available from the police; hence Box (I) is empty. Box (II) reveals, however, that offices of the prosecutors received new cases related to 3,396,695 offenses in 2013, which include cases with unknown suspects. This *flow* of offenses into the offices of the prosecutors in a given year is the correct measure of the extent of criminal activity, regardless of whether the perpetrator is known or arrested. Consequently, we employ the new cases received by prosecutors each year (Box II) as our main crime measure.

Prosecutors dispose of their cases by either forwarding them to courts, or by dropping them because of lack of evidence or lack of a suspect. Also, some cases are not disposed of in a given year by the prosecutors' offices due to overcrowding, high workload, or because the investigation has not been completed in that same year. These cases are listed in annual reports as pending (staying in the prosecutors' offices until the following year). The arrow from Box (II) to Box (III) signifies the transfer of files from the prosecutors to the courts in 2013, and shows that 1,315,457 criminal case files were forwarded to courts. This number includes criminal cases with maximum sentence lengths of 10 years, as well as felonies with longer sentences. It also includes cases in criminal courts of peace where cases are primarily related to disputes between the police and the citizens, prosecutorial decisions, search and seizure warrants, and so on. Juvenile cases are also part of this aggregate, although they constitute only about five percent of all cases, and nine percent of cases excluding criminal courts of peace.

The number of adjudicated cases in Box III in Figure 3 is not a sound indicator of criminal activity because the number of cases in courts is always smaller than the number of offenses as some of the offenses in prosecutors' offices have no suspects and they therefore cannot be referred to courts. Furthermore, the offense-to-case ratio can change over time.[19] Box III also depicts the number of charges in these court cases. Although the number of charges in court cases is also smaller than the number of offenses handled by the prosecutors' offices, it is a better indicator of the extent of criminal activity than the number of cases.[20] As shown in Box III of Figure 3, in 2013 there were 3,388,613 charges in courts nationwide.[21] To investigate the sensitivity of our results, we use the number of charges in courts as our alternative (albeit noisy) measure of

[18] We chose the year 2013 because it is the midpoint of our analysis period. Other years provide the same picture as the one portrayed in Figure 3 regarding both the number of cases and how they filter through the criminal justice system.

[19] A simple example is the following. If the use of firearms becomes more prevalent over time, various crimes may be more likely to involve a firearm and this would add a weapons charge to these crimes, increasing the offense-to-case ratio. Or, if crime goes up on the intensive margin (marginal criminals committing multiple offenses), this will also increase the offense-to-case ratio.

[20] In some cases, some charges are being dropped while others being retained for a given defendant. For example, the file in the prosecutor's office may include charges of assault, drug use, and robbery for the same suspect, while some of these charges may be dropped before the case is forwarded to the court.

[21] However, 562,000 of these pertain to charges related to bankruptcy, loan and contract disputes and other cases which are arguably less impacted by the refugee inflow. While this break-down is available at the national level, the way the data are reported by the Turkish Ministry of Justice does not allow us to deduct these non-felony charges in court cases from the total number of charges at the province level.

criminal activity.

As revealed by Box (IV), about 445,000 defendants are convicted, which indicates a conviction rate of 37 percent. As in the case in every country, not all convicted defendants are incarcerated. Convictions could lead to outcomes other than prison terms, such as fines, suspended sentences, probations and other sanctions. As a result, only 161,711 individuals entered the prison system in 2013 (Box V in Figure 3).

In summary, as the number of reported crimes flows through the criminal justice system (as we move from left to right in Figure 3), it becomes less and less appropriate to consider them as correct indicators of criminal activity. And it is particularly problematic to use the intake into the prison system, which is the final stage of the judicial process, as an indicator of crime (as was done by Kirdar et al. (2022)). Additional important details of the pitfalls of using prison entry as an indicator of criminal activity (e.g., the mismatch between the year in which the crime is committed and the year the perpetrator is incarcerated) are discussed in Appendix B.

The magnitude of the measurement error, which arises if entry into the prison system is used as a measure of criminal activity, can be seen by comparing the inaccurate crime rate based on the number of inmates entering prison with the correct crime rate based on the new offenses reported to the prosecutors. The incorrect measure (prison intake) implies a crime rate of 196 per 100,000 residents [Kirdar et al. (2022), Table 1], while the correct crime rate (calculated using the number of offenses received by prosecutors used in our paper), reveals that the crime rate is 4,500 per 100,000 residents.[22]

Figure 4 displays the time-series pattern of the number of new crimes received by the prosecutors' offices. This is our primary measure of criminal activity (Box II in Figure 3), which exhibits a significant jump in 2017 and 2018. This jump is due to the repercussions of a failed coup d'état on July 15, 2016. In the wake of this unsuccessful coup attempt, the government declared a state of emergency. A number of executive orders, issued by the government, granted special authority to law enforcement agencies in order to find, arrest, and prosecute individuals who were suspected collaborators or supporters of the coup attempt. Consequently, tens of thousands of alleged collaborators were arrested and prosecuted, which led to the unusual jump in the number of crimes handled by the prosecutors' offices in 2017 (see Altindag and Kaushal (2021) for additional details). Thus, to avoid confounding due to this event, we exclude the years 2017 and 2018 from the analyses.[23] To demonstrate that the jump in the number of offenses received

[22]To put this mismeasurement into perspective, crime rates (crimes per 100k residents) are 1,500, 1,800, and 2,200 respectively, in Bulgaria, Greece, and Spain. The crime rate is about 3,200 in Romania, 3,500 in Portugal, and 3,800 in Malta. The rate is 4,500 in Italy, 5,000 in Scotland, 7,000 in England and Wales, 7,500 in Germany (Aebi et al., 2014). The crime rate in the EU was around 7,000 in 2010 (Buonanno et al., 2018). The U.S. crime rate was 2,500 in 2019, although it was more than 4,000 in 2,000 and 3,500 in 2010 (FBI, Uniform Crime Reports).

[23]The data include terrorism charges handled by the prosecutors' offices. There were about 15,000 terrorism-related cases each year nation-wide handled by the prosecutors' offices until 2016 (about 0.4 percent of all crimes).

by the prosecutors' offices after 2016 is not concentrated in one region of the country, we present in Appendix A Figure 9 the behavior of the same variable in five regions of the country (West, East, North, South, and Central).

Figure 5 presents the national time-series pattern of our alternative measure of crime: the number of offenses charged in cases handled by courts. As mentioned above, this alternative variable is measured with error, because unlike our main crime indicator (depicted in Figure 4), it does not include cases with unknown suspects, or cases that were not forwarded to courts by prosecutors for other reasons (e.g., insufficient evidence). On the other hand, it includes charges which did not involve the offices of the prosecutors, such as contract disputes and bankruptcy resolutions. To the extent that these charges are not related to refugees, this alternative crime measure contains additional noise, although it can be argued that bankruptcies, disputes between tenants and landlords, and other related cases can also be impacted by the repercussions in the labor and housing markets following the inflow of refugees.

The pattern of this variable in Figure 5 calls for a couple of comments. First, there is a significant drop in 2016, the year of the coup attempt. Recall that the coup attempt took place on July 15, 2016. It has been quickly determined that a large number of collaborators of the leader of the coup were employed as judges at the Ministry of Justice. There was a swift and large-scale purge of these judges, who were subsequently replaced by new hires. Between July 15, 2016 and December 31, 2017, 4,385 judges and public prosecutors were officially investigated under the suspicion of being a member of FETÖ/PDY terrorist organization, and following these investigations 3,945 judges and public prosecutors, which corresponds to 26 percent of judges and prosecutors on duty as of 26 July, 2016, were dismissed.[24] The drop in the number of charges in courts in 2016 (Figure 5) is likely the reflection of this phenomenon. Nevertheless, to be consistent between the use of the two crime measures, we keep the year 2016 in the analysis sample when we employ this alternative, albeit noisy, crime indicator. Second, note that the two measures of crime do not display the same time-series pattern. While the number of crimes received by the offices of the prosecutors exhibits a secular positive trend (Figure 4), the number of offenses in court cases is quite noisy, and exhibits no clear systematic time-series pattern (Figure 5). As the econometric analyses reveal, however, both measures of crime lead to the same conclusion regarding the impact of refugees on crime, with similar magnitudes of the estimated effect.

[24]Council of Judges and Prosecutors, 2017 Annual Activity Report; see, https://www.hsk.gov.tr/Eklentiler/files/HSK%202017%20YILI%20FAAL%C4%B0YET%20RAPORU.pdf.

5 Descriptive statistics

Table 1 presents the descriptive statistics of key variables for the entire sample and for the years before the beginning of the Syrian refugee inflow (2006-2012) as well as the years afterwards (2013-2016). The provision of public services has expanded during the time period analyzed. For example, while natural gas was available as a means for energy in 23 percent of the provinces in 2006, the rate went up to 65 percent in 2009; and 79 percent of provinces had access to natural gas in 2012. Ninety-four percent of the 81 provinces had natural gas in 2016. Public expenditures per 100k population were 27,772 TL (in nominal terms) between 2006 and 2012, which went up to 51,758 TL (in nominal terms) between 2013 and 2016. The number of hospitals per 100k population also rose over the time period analyzed, from an average of 2.38 in 2010 to 2.43 in 2016.

To the extent that improvements in economic conditions have crime-reducing effects, these changes point to a decrease, rather than increase, in crime. Yet, there was an increase in the number of crimes handled by the offices of the prosecutors from an average of about 38,000 offenses per province per year (about 3 million offenses for the country per year) to about 43,400 offenses (3.65 million offenses country-wide). As a result, the number of convicted felons entering the prison system, which is a function of lagged criminal activity, doubled from an average of 1,128 inmates per year per province in the pre-refugee period to 2,126 inmates in the post-refugee period.

As mentioned in the previous section, we will also employ the number of offenses in criminal courts as an alternative measure of criminal activity. The average number of offenses filed in criminal court cases went down from about 41,724 to 39,253 per province-year from the pre-refugee to post-refugee period. It should, however, be noted that the data for this variable are available only starting in 2010—i.e., the pre-refugee average is based on two years of data from each province. More importantly, the decline in this variable during the post-period is primarily because of the drop in 2016 due to purge of judges after the coup attempt, explained above (see Figure 5). If the year 2016 is excluded, the average number of charges in courts is 40,704 (between 2013 and 2015.)

In our empirical design, the effect of refugees on crime will be identified by exploring province-by-time variation in the number of refugees. In this design, it is important to investigate whether there is a systematic relationship between refugee existence in a province and observable province attributes. If refugees are distributed between regions and over time "almost randomly," there should not be a relationship between refugee intensity and observable attributes of the provinces in which they are located. The results of this balance test are provided in Table 2. In this exercise, we regress the logarithm of Syrian refugees in a given province-and-year on other time-varying province characteristics to investigate whether time-varying province indicators systematically predict the Syrian refugee share.[25]

The result in column (1) of Table 2 shows that only the size of native population of the province is significantly related to the number of Syrian refugees, but the coefficient becomes smaller and statistically insignificant when the model includes time trends and region fixed effects in column (2). As displayed by the F-tests, province characteristics fail to jointly predict the size of the Syrian refugee population after we control for province and region effects. Column (3) shows that the results are not sensitive to the exclusion of the mega city Istanbul.

6 Empirical framework

Following Ehrlich (1973) (p.534) and the empirical literature that follows, it can be postulated that the aggregate supply of offenses can be described by a production function at the province level (ignoring the subscripts) as follows:

$$CR = f(A, N, R), \tag{1}$$

where CR is the number of offenses committed in the province, N represents the number native residents, and R stands for the size of the refugee population. A is a vector that includes other determinants of crime, including socio-economic factors, deterrence measures, and unobserved cultural attributes and other characteristics of the province. This production function can further be specified as

$$CR = AR^{\beta} N^{\gamma} e^{\varepsilon}, \tag{2}$$

the empirical counterpart of which is:

$$\ln(CR) = \alpha + \beta \ln(R) + \gamma \ln(N) + \varepsilon, \tag{3}$$

where α contains observable exogenous characteristics of the province, as well as various fixed effects that soak up unobserved province and regional attributes. Equation 3 and its alternative versions will be employed in our empirical analyses as explained below.

Equation 3 can be converted into different forms. For example, adding $[-\beta \ln(N+R) - \ln(N+R)]$ to both sides and rearranging terms yields

$$\ln\left(\frac{CR}{N+R}\right) = \alpha + \beta \ln\left(\frac{R}{N+R}\right) + \gamma \ln(N) + [(\beta - 1)\ln(N+R)] + \varepsilon, \tag{4}$$

[25] +1 is added to the refugee population.

which can also be written as

$$\ln\left(\frac{CR}{N+R}\right) = \alpha + \beta\ln\left(\frac{R}{N+R}\right) + \gamma\ln(N) + \nu, \tag{5}$$

where $\nu = [(\beta - 1)\ln(N+R)] + \varepsilon$.

The left-hand-side of Equation 5 is (log of) the crime rate, and the key variable on the right-hand side is the (log of) share of refugees in total population. Although Equation 5 is a rearrangement of Equation 3, it is not appropriate to use Equation 5 in an effort to estimate the impact of refugees on crime. This is because of the following reasons:

(*i*) Suppose that there is no true relationship between the refugee share ($R/(N+R)$) and the crime rate; that is, $\beta = 0$ in Equation 5. The error term of Equation 5 reveals, however, that an increase in the number of refugees (R) will nevertheless induce a negative relationship between refugees and the crime rate. This mechanical negative relationship, imposed by the transformation of Equation 3 to Equation 5 persists as long as the elasticity of the crime rate with respect to refugee share (β) is less than one. Put differently, fitting Equation 5 to data underestimates β.

(*ii*) Ignoring the issue highlighted in point (*i*), another problem in using Equation 5 is that the variable of interest, R, is both in the numerator of the key explanatory variable, and in the denominator of the dependent variable. This property of Equation 5 also imposes a mechanical negative relationship between refugees and the crime rate by construction.

(*iii*) Because the size of the refugee population (R) in a province is likely an endogenous variable, estimation of Equation 5 would benefit from an instrument for R. However, related to points (*i*) and (*ii*) above, any instrument that is correlated with R is invalid in Equation 5, because the exclusion restriction is violated and the estimated β is biased. More specifically, consider Equation 5 again. The probability limit of the instrumental variables estimate of β is: $\text{plim}\,\widehat{\beta} = \beta + \frac{Cov(\ln(Z),\nu)}{Cov(\ln(R/(N+R)),\ln(Z))}$, where Z is the instrument. Any instrument Z, which would generate a movement in R, is also correlated with the error term (ν) of Equation 5 as the error term contains R. More specifically, if $Cov(Z,R) > 0$, this would imply that $Cov(Z,\nu) < 0$ if $\beta < 1$, and $Cov(Z,\nu) > 0$ if $\beta > 1$. The instrument is uncorrelated with the error term only if $\beta = 1$, but even in this special case the instrument is invalid, because the endogenous variable (R) also appears in the denominator of the dependent variable.

(*iv*) Finally, even if none of these vital issues existed, a basic problem would have been the use of the same divisor both in the dependent and the independent variable. More specifically, using the crime rate as the dependent variable, and then using population (which is the denominator of the crime rate) as the deflator of the key explanatory variable creates bias.[26]

[26] Consider the model $[CR/(N+R)] = \alpha + \beta(R/N) + \varepsilon$, where R is the refugee population, N is the native population, and

Because of these reasons, we use the formulation in Equation 3. More, specifically, we estimate the following specification:

$$\ln(CR_{prt}) = \alpha + \beta \ln(R_{pt}) + \gamma \ln(N_{pt}) + \delta_p + \Psi_t + \boldsymbol{\Omega}'\boldsymbol{X}_{prt} + \tau_{rt} + \varepsilon_{prt}, \tag{6}$$

where CR_{prt} represents the number of crimes committed in province p, region r, and year t. It is measured by the new cases incoming to the offices of the prosecutors each year. As discussed in Section 4, CR includes felony violent and property crimes which were handled by prosecutors. R is the refugee population and N stands for the native population. Because the sample includes observations with zero refugees, +1 is added to all refugee populations before taking logs. Fixed effects to control for unobserved differences between provinces are represented by δ_p; and Ψ_t stands for year fixed effects. Eighty-one provinces of the country are divided into 12 standard statistical regions, and these regional differences (which also vary by time) are filtered out by the inclusion of region-by-year dummies τ_{rt}. The 12 regions of the country are further classified into five upper-regions (West, East, North, South and Central). The model also includes a set of upper-region time trends for the five broad geographic regions of the country. Standard errors are clustered at the province level.

Aside from time and region shocks that can be captured by year and province fixed effects, there may exist region-level time-varying developments which can mistakenly be attributed to refugee inflows if not appropriately controlled for. For example, economic activity, political tendencies, socio-demographic factors, and even government policy can change over time and across regions. These time-varying regional shocks would ideally be captured by the inclusion of province-year interaction dummies. However, the treatment (i.e., regional distribution of refugees) is specified at province-year level in Equation 6, and province-year fixed effects would be collinear with the treatment variable. Thus, we follow the convention in the literature (see, e.g., Stephens and Yang (2014)) and include region-year interaction terms (τ_{rt}) by using a higher-level (i.e., more aggregated, NUTS1) regional classification than province-level. We also use region-level time trends to control for the remaining time-varying unobserved factors.

We also control for time-varying province-level attributes, which are not collinear with the treatment variable due to their continuous nature. It is possible to obtain data on such province attributes as per capita income, the number of physicians, the number of hospital beds, teacher-pupil ratio, the number of motor vehicles, and so on. These attributes, however, would respond to the inflow of refugees. For example, if the refugees affect economic conditions (positively or negatively), indicators of economic activity, ranging from income to unemployment will be

(R/N) represents the refugee share. Because $(N + R) \approx N$, this regression would produce a spurious relationship between the crime rate and refugee share because both the dependent variable and the explanatory variable have (almost) the same denominator. As explained by Kronmal (1993) and Bazzi and Clemens (2013), and as highlighted with examples by Clemens and Hunt (2019), the denominators that are the same or very similar will generate spurious correlation between the two variables when the true β is zero. See Kronmal (1993) for theoretical and empirical examples, and proposed solutions.

correlated with the size of the refugee population. Similarly, variables measuring various aspects of government services, such as the number of physicians, the number of hospital beds, and the number of teachers, are endogenous to refugee population because in Turkey the overwhelming share of these services are provided by the central government, and the government can adjust the staff size in the health and education sectors in response to increased demand due to refugee inflows. Consequently, in this particular analysis, where the key variable of interest is the refugee population, it is inappropriate to include variables to the model that are likely to be impacted by the refugee population. However, there exist a handful of variables that may be considered as pre-determined, or reasonably insensitive to the change in refugee population. One such variable is the existence of natural gas in the province as an energy source for heating. Since the early 1990s, Turkey underwent a major effort to replace coal with natural gas. The installation of the natural gas lines took years and its entry to different provinces took place in different years. Given that the existence of natural gas is exogenous to the refugee population, we use it as a control variable. A similar variable is per capita government expenditures in the province. Another control variable is the number of hospitals per 100k capita. Although the number of physicians and to some extent the number of hospital beds can be adjusted to the rise in demand in services, increasing the number of hospitals takes significantly longer. Thus, we use the number of hospitals per 100k capita as an indicator for the infrastructure of the health care delivery system in the province. Along the same lines, we use per capita government expenditures as a control variable. We also control for the twice-lagged value of the number of inmates entering prisons in a province in a given year.[27] Prison entry is lagged to minimize concerns of reverse causality (see, Corman and Mocan (2005), etc.), although using its once-lagged, or contemporaneous value, or omitting it from the model has no impact on the results.

The instrument. Even though observable characteristics of provinces of the country are not correlated with the size of the refugee population in these provinces as revealed by Table 2, it may be the case that the intensity of the refugee inflow to a destination province is related to some unobserved province attributes. To address the potential endogeneity of refugee location, the literature almost always relies on distance-based IV strategies (Del Carpio and Wagner, 2015; Tumen, 2018, 2021), although there are slight differences between the construction of particular instruments. In our main specifications, we employ the most commonly-used instrument, which uses the average distance from each Syrian governorate to Turkish provinces weighted by aggregate annual refugee entry into the country. To demonstrate the insensitivity of the results to the choice of instrument, we also employ two other, slightly different instruments used by other researchers (Aksu et al., 2022; Altindag et al., 2020).

[27] Total prison population, rather than prison entry, would have been a better measure of deterrence, but the number of prisoners is not available at the province level. Similarly, the size of the police force is available only for the year 2015 at the province level.

Our instrument, proposed by Del Carpio and Wagner (2015), explores the variation in the distance between the origin (the source governorates in Syria) and the destination provinces in Turkey. It is motivated by the unanticipated nature of the civil conflict in Syria, and Syrians' immediate need of fleeing the war environment to the closest provinces in neighboring Turkey. Based on the international community's expectations that the civil conflict would be short-lived, the Turkish government built temporary refugee camps in Turkish provinces near the Syrian border to help accommodate the large number of displaced Syrians. Similarly, the Syrian refugees affirmed that their displacement would be temporary, and they had no legal status in Turkey which posed significant hindrance to their labor market prospects. Consequently, Syrian refugees stayed close to the Turkish-Syrian border with the expectation to return to Syria when the armed conflict ended. This led to rapid overpopulation of Turkish provinces located at the border after the onset of the Syrian civil conflict. As hypothesized in the gravity models of immigration, it is plausible to expect a higher share of displaced Syrians to originate from more populous Syrian governorates in response to the armed conflict. Therefore, the instrument also incorporates the fraction of population in each Syrian governorate before the civil war erupted to allow for the potential heterogeneity across Syrian governorates.

More specifically, our instrument for the Syrian refugees in year t in a given province p is defined as follows:

$$Z_{pt} = RI_t \sum_g \pi_g \left(\frac{1}{\omega_{pg}} \right), \tag{7}$$

where ω_{pg} represents the shortest travel distance between each Syrian governorate g and the Turkish province p. We obtained the shortest distance between each pair of 81 Turkish provinces and 14 Syrian governorates using Google Maps. This pairing yields 1,134 distinct potential routes for Syrian refugees. The fraction of Syrian population in each Syrian governorate g in the pre-conflict period (2010 being the pre-conflict year) is represented by π_g; thereby, alluding to the possibility that the more populous Syrian governorates likely experienced more significant displacements as a result of the civil war. The total number of Syrian refugees in Turkey in any given year is represented by RI_t. The identifying assumption of the instrument is that the shortest travel distance across pairs of Syrian and Turkish localities affects the changes in the local crime rates only through increasing the number of refugees, our treatment variable. Another assumption, implicitly invoked, is that the total annual inflow of refugees does not lead to movement of native population. Balkan and Tumen (2016) provide evidence that the Syrian refugee inflows have not generated significant changes in internal migration patterns of the native population. Moreover, our empirical specifications include time-region interaction terms and region-specific trends, which control for all time-varying regional shocks—including any potential internal migration effects.

7 Results

Table 3 presents our benchmark estimates. The analysis sample covers 81 provinces of the country over the years 2006 to 2016. The dependent variable is the logarithm of the felony crime cases handled by prosecutors' offices in each province. In the interest of space, we only report the estimated coefficients of $\ln(R)$ and $\ln(N)$. The specification reported in column (1) includes only year and province fixed effects. In columns (2) to (4), we sequentially add time-varying province attributes, region-specific time trends for the five regions of the country (North, South, East, Central, and West) and area-year fixed effects pertaining to 12 NUTS1 regions. The results of the most extensive specification, reported in column (4), are obtained from a specification which also includes province trends interacted with the pre-refugee level of crime (in year 2011) of each province.

Panel A presents the OLS results, which indicate a positive effect of refugees on crime. Panel B of Table 3 displays the results of instrumental variables regressions. The first-stage is strong with an F-value of greater than 40 for the instrument. The results indicate that the refugee influx has led to a statistically significant increase in the number of new offenses handled by the prosecutors' offices. The assumption that the instrument has an impact on crime only through its influence on the refugee population is ultimately non-testable. Therefore, we also report the result of the estimated reduced form in column (5), which reveals a significant impact of the instrument on criminal activity.

The average of the refugee population is about 9,000 in this sample because the refugee population is zero between 2006 and 2012 in all provinces. The native population average is 922,000 and the average annual number of crimes is about 40,000 per province. Thus, the instrumental variable results indicate that an increase in the number of refugees by 2,000 would generate an additional 480 crimes received by prosecutors' offices in a typical province.

Our alternative crime measure is the number of charges in cases filed at the criminal courts in that year. This measure represents a noisy proxy for criminal activity in comparison to the number of new cases handled by the prosecutors' offices. This is because the number of offenses handled by the prosecutors' offices include crimes that do not end up on the court's docket for a variety of reasons, ranging from lack of sufficient evidence to lack of a suspect (see the flow chart presented in Figure 3). In addition, as discussed earlier, the number of charges in court cases includes non-felony cases that may not be meaningfully impacted by refugee inflows.

Table 4 presents both the OLS and IV estimates where the extent of criminal activity is measured by the number of charges in courts. Because these data are available starting in 2010, there are 567 observations in the regressions. The OLS coefficient of the refugee population is positive, small and not different from zero. The IV estimates reported in Panel B, on the other hand, are highly statistically significant, regardless of the specification, despite the small sample size. This result is important because the national aggregate time series patterns of the two crime indicators are *negatively* correlated (see Figures 4 and 5), and the province level correlation between the two is only weakly positive.[28] Put differently, the two measures of crime exhibit different patterns over time. Nevertheless, they both yield the same inference that an increase in the refugee population leads to a rise in crime. The IV estimate in column (4) of Panel B in Table 4 implies that if the province's refugee population went up by 2,000, this would generate an additional 485 charges brought to courts. Thus, even though the crime measures are different in regressions reported in Tables 3 and 4, the estimated effects are nearly identical.

It should be noted that the identified rise in the incidence of crime due to the increase in refugee population does not signify the number of refugees who commit crimes. This is because of two reasons. First, the dependent variable is the number of crimes, and it therefore measures the intensive margin. Given that a typical offender commits multiple crimes, the number of offenders involved in these crimes is smaller than the number of crimes. Second, although the impact mentioned in the exercise above is propagated by an increase in the number of refugees, the number of crimes committed as a result of this increase is not entirely attributable to refugees. Rather, the rise in criminal activity is generated by both the refugees and natives, because the dependent variable is the total number of crimes handled by the criminal justice system, and not the number of crimes committed by refugees. As discussed earlier, under the assumption that a large increase in the refugee population has labor market repercussions, the unskilled segment of the native population will be impacted by this surge, which will in return contribute to the rise in criminal activity. Later in the paper we provide additional insights into particular sub-populations that may be contributing to the rise in refugee-induced crime.

It is also interesting to note that the coefficients reported in Tables 3 and 4 indicate no impact of native population on crime (with the exception of the specification reported in column (4) of Table 3). This is likely because only about 33 percent of the total population of the country is in the most crime-prone age interval of 15 to 35 (TurkStat), and cohorts who were in this age group around the time period of 2010-2016 should have received at least a middle school education due to a compulsory education reform which was implemented in 1997.[29] This means that it is difficult to detect an impact on crime of an overall increase in the size of the general population without focusing on the most crime-prone segment of that population. To address this issue, we used the size of the native population between the ages of 15 and 44, and between

[28] A regression at the province-year level (no of observations=567) of $\ln(CR1)$ on $\ln(CR2)$, where $CR1$ is the number of new offenses handled by prosecutors' offices and $CR2$ is the number of offenses in court cases yields of a coefficient of 0.33 (p=0.00) in a model with province fixed effects and year fixed effects. When the model also includes region-by-year dummies, the coefficient is 0.26 (p=0.00).

[29] The reform increased the compulsory years of education from 5 years to 8, starting in 1997(Akyol and Mocan, 2020; Cesur and Mocan, 2018; Torun, 2018).

15 and 54. The results did not change; i.e., the impact of refugees on crime remained the same in magnitude and significance, and the coefficients of population 15-44 or 15-54 were not different from zero. However, to tease out the impact of native population on the incidence of crime further, we analyze various sub-groups of the native population, categorized by education. We perform this analysis in Section 9 below.

8 Sensitivity analyses and placebo tests

In this section, we report the results of a number of sensitivity analyses, ranging from the use of alternative instruments to focusing on various sub-samples based on geography and time-period. We also provide the results of event study and placebo exercises.

Alternative instruments. We examine the robustness of the results to the use of two alternative instruments. The first one is similar to our instrument as it is based on the shortest travel distance between Syrian governorates and Turkey, but also considers distances to other neighboring countries (Kirdar et al., 2022). The remaining set-up of the instruments mimics our instrument. The second alternative instrument makes use of the fact that Syrian refugees are Arabs, and assumes that they would be attracted to provinces which have a larger share of Arabic-speaking population. Hence, the interaction of the fraction of the Arabic-speaking population in a given province in 1965 and the number of internationally displaced Syrians in a given point in time serves as an instrument for the Syrian refugees in a given province over time. This instrument is motivated by the historical settlement of the Arabic-speaking population in various Turkish provinces and assumes that the location of residents in 1965 with Arabic heritage is exogenous to the current outcomes of interest (Altindag et al., 2020). Although this set-up invokes the relatively strong assumption that the geographic distribution of the Arabic-speaking minority has not changed appreciably over the last half century, we nevertheless employ it as the second alternative instrument.

Appendix A Tables 7 and 8 present the results obtained by using these alternative instruments for our main crime indicator (the number of crimes handled by the offices of prosecutors), and the number of criminal charges in courts, respectively. In both tables, Panel A displays the results pertaining to the first alternative instrument (modified Shift IV) and the results in Panel B are based on the second instrument (Arabic IV). The results are consistent with those reported in our main models (Tables 3 and 4). The sample size is almost halved in Table 8, which reduces the precision, but not the magnitude of the estimates.

Refugee ratio as a measure of refugee intensity. There is an extensive literature on the wage gap between skilled and unskilled workers, and how this gap is influenced by skill-biased technological change. Models estimated in this literature employ the ratio of skilled-to unskilled labor[30] as the key explanatory variable (Autor et al., 2008). Motivated by these specifications, some analysts investigated the impact of refugee-to-native ratio on labor market outcomes (Altindag et al., 2020; Aksu et al., 2022; Ceritoglu et al., 2017; Del Carpio and Wagner, 2015). To evaluate the robustness of our results, we employ the refugee ratio (the ratio of refugees to the native population) as an alternative measure of refugee presence in a given province in a given year. More specifically, we estimate the model depicted by Equation 6, by replacing Refugees (R) with Refugee Ratio (R/N). While this measure has been widely used in prior research on Syrian refugees, our model theoretically suggests that the logarithm of refugees in a given province should be the main variable of interest in the reduced form estimations. However, using the logarithm of the refugee ratio, by construction, provides the same coefficients as those displayed in Tables 3 and 4.[31] Therefore, we test the robustness of our results by using the refugee ratio (without logs) as a measure of the refugee population in a given province over time.

Results with refugee ratio are shown in Appendix A Tables 9 and 10 for the number of crimes handled by the Prosecutor's Office and the number of criminal charges in courts, respectively. In these tables, we first present the OLS estimates followed by the main IV, modified-shift-IV, and Arabic-IV results. As indicated in these tables, our results remain robust and similar in magnitude across all specifications and instruments, supporting our baseline findings.

Different time periods and different regions. We further investigate the sensitivity of the results to time periods. As discussed earlier, we restrict the end of the analysis sample to 2016 because of the coup attempt in Turkey in July 2016, which led to dismissal of a sizable number of judges, and other related upheavals of the criminal justice system in an effort to purge and prosecute alleged collaborators of the coup d'état. Recall that the refugee influx started in 2013 (see Figures 1 and 2). Put differently, the identifying variation is obtained from four years of data (2013 to 2016) for each province. Although shortening the sample period would reduce identifying variation, we nevertheless estimated the models using the samples that end in 2015, and in 2014. The results, reported in Appendix A Table 11 reveal no change in inference despite the decline in sample sizes.

A number of recent studies have suggested that two-way fixed effects models may not provide an estimate with a causal interpretation when effects are heterogeneous (De Chaisemartin and d'Haultfoeuille, 2020).[32] Our IV specification would significantly mitigate such a concern because

[30]Note that this literature does not employ the share of skilled labor-to all labor as an explanatory variable. Rather it uses the ratio of skilled-to-unskilled labor. More specifically, the skill-biased technological change literature regresses the ratio of high-to-low skill wages on the ratio of high-to-low skill employment. This formulation can be derived using the theory of production, profit maximization and competitive markets (Autor et al., 2008; Card and DiNardo, 2002).

[31]Note that $\ln(CR) = \alpha + \beta\ln(R/N) + \gamma\ln(N) + \varepsilon$, reduces to $\ln(CR) = \alpha + \beta\ln(R) + \delta\ln(N) + \varepsilon$, where $\delta = (\gamma - \beta)$. This indicates that the estimated impact of R should be the same in both versions.

[32]Our analysis reveals that the estimated weights are not systematically related to the factors that could increase the effects of the Syrian refugees, specifically year of arrival, and the logarithm of refugees itself.

the construction of the instrument is structured around the predetermined potential migration pathways between Syria and Turkey instead of the gradual inflow of the Syrian refugees across Turkish provinces over time. Appendix A Table 12 provides additional evidence in support of the robustness of our results and the lack of heterogeneity across different subsamples. In the first five columns of Appendix A Table 12, we exclude each region from the analysis sample to test whether our results are driven by a particular region. The last two columns of the table focus exclusively on the regions in the Southern and Eastern Turkey which are closer to Syria and where the Syrian refugees primarily settled when the civil war erupted. The estimates are stable across the samples, indicating homogeneity of the estimated effect across regions of the country.

In addition to the previous exercise, we also test the robustness of the results by excluding the largest and most populous cities such as Istanbul, Ankara, and Izmir, as well as cities with the highest refugee to native ratios. The results are displayed in Appendix A Table 13. In panel A, the first column reports the main estimates from Table 3. Column (1) of Panel B presents the estimates from Table 4. Column 2 presents the results that are obtained when Istanbul (which has a population of about 15 million) is dropped. Column (3) repeats the same exercise by dropping the three biggest cities of the county: Istanbul, Ankara, and Izmir. Finally, the last column drops the cities of Kilis and Sirnak, which had the highest refugee ratio over time. In these subsamples, the results mirror those reported in previous tables, in which we find that a larger influx of Syrian refugees leads to more local crime, as measured by the number of cases handled by the Prosecutor's Office, or by criminal charges in courts. Taken together, these results indicate the results are not driven by a particular region of the country.

We also investigate whether the estimated effects vary across Turkish provinces with varying education and poverty levels prior to the arrival of the Syrian refugees. It is possible that provinces with better pre-war institutional capacity and human capital would potentially better navigate the population shock generated by the refugee inflows. To formally test this conjecture, we divided Turkish provinces into 3 categories by the percentage of the population with high school degrees in 2011 (in Panel A of Appendix A Table 14) and the percentage of households that belong to the lowest wealth category in the 2013 Turkish DHS (in Panel B of Appendix A Table 14). Although the results in column (1) of Panel A suggest that the impact on crime of a rise in the refugee population is larger in provinces with lower levels of initial education, we cannot reject the hypothesis of the equality of the estimated coefficients. Taken as a whole, the results of Appendix A Table 14 reveal no significant heterogeneity of the refugee impact on crime by province-level pre-refugee education or wealth.

Finally, we estimated the model by aggregating the data to the 26 NUTS2 regions of the country. Doing so reduces the number of observations to 286. Our instrument remains strong with a first-stage F-value of 16.21. In this specification, the estimated IV coefficient is 0.05 (std error=0.01), yielding the same inference. The two other alternative instruments, discussed earlier, also generate statistically significant effects of the refugees on crime with coefficient estimates of 0.063 (std error=0.017), and 0.104 (std error=0.045), although these instruments are not powerful in these aggregate regressions with first-stage F-values of 5.66 and 3.71.

Event study and placebo analyses. We performed event-study style analyses that aim to explore the differences in crime trends before the arrival of refugees. To test the parallel trends assumption between low and high exposure provinces we divide them into two groups. The first group consists of provinces with the refugee-to population ratio ($\times 100$) lower than 0.1 in 2016, and those with a ratio ($\times 100$) of 0.1 or higher. This partition divides the sample into two groups of provinces, where the latter group consists of 65 "treated" provinces (80 percent of the provinces), while the former contains 16 "control" provinces (20 percent of the provinces). While the cutoff is admittedly arbitrary, it divides the sample between more heavily and less heavily treated provinces. Using an indicator variable that determines the heavily treated group and interacting this indicator dummy with year dummies allows us to analyze whether the difference in criminal activity between the groups was trending before the onset of the refugee influx in 2013. The assumption we invoke here is that the provinces in the first group do not differ from those in the second group in ways that directly impact their criminal activity. With that proviso, we present the results in Figure 6. The left-out category is the year 2011, and as revealed by the graph, there is no statistically significant difference between control and treatment provinces until 2014. On the other hand, the impact of having experienced a heavier refugee influx leads to a higher rise in crime, as evidenced by positive and statistically significant coefficients after 2013. To demonstrate the robustness of this exercise and to shed some light into potential selection into groups, we also partitioned provinces into treatment and control groups using the threshold of the refugee-to-native population ratio ($\times 100$) at 0.067, which corresponds to the 15th percentile of the ratio distribution in 2016. Similarly, we used the 25th percentile as the cutoff (refugee-to-native ratio $\times$ 100 = 0.15). These classifications generated the same inference, as displayed in Appendix A Figures 10 and 11. These results indicate no differential pre-trends between more-heavily and less-heavily treated provinces.

As a final exercise, we implemented placebo analyses in which we randomly re-assigned the values of refugees within a province, and re-estimate the models. Each province was exposed to refugee inflows from 2013 to 2016. We take the actual refugee population of a particular province for each year between 2013 and 2016, and randomly re-assign these refugee values to those years. We do this for each province and estimate our models. We repeat this exercise 1,000 times, each time randomly re-distributing the number of refugees of each province. If random assignment of refugee population produces coefficients similar to the ones estimated in the paper, this would raise doubts about the validity of our design and estimates.

The distribution of the estimated placebo coefficients of the instrumental variable regression for our first crime variable (cases handled by the offices of the prosecutors) are displayed in Figure 7, along with the coefficient estimated in our model in column (4) of Table 3, which is depicted by the vertical line. The placebo estimates are very different from our estimates, and the probability of our estimated impact coming from this distribution is zero.[33] We repeat the same analysis for the reduced form model. Here we randomize the values of the instrument in each province between the years 2013-2016. The distribution of 1,000 placebo coefficients is displayed in Figure 8, along with our estimate (0.017) from Table 3. Our estimated reduced form for impact has zero probability to have come from this placebo distribution. Thus, Figures 7 and 8 refute the hypothesis that our results could have been produced by any distribution of refugees within each province.

9 Digging deeper

The production of offenses depicted by Equation 1 can be expanded to the following form:

$$CR = f(N_1, N_2, N_3, R, \boldsymbol{X}), \tag{8}$$

where CR is the number of offenses committed. N_1, N_2, and N_3 represent the number of natives over age 15, categorized by skill. More specifically, N_1 stands for the number of natives who are 15 years of age or older who are illiterate, N_2 is the number of individuals in the same age group who have an elementary school or middle school diploma (education $\leq$ 8 years), and N_3 represents the number of people with at least a high school diploma who are 15 and older. R stands for the size of the refugee population, and $\boldsymbol{X}$ is a vector that includes other determinants of crime, including deterrence measures.

[33]The coefficients that are not distributed around zero are attributable to the fact that we randomize four values of refugee population (between 2013 and 2016) for each city and the refugee numbers are strongly correlated over time in a city. Therefore, even if these four values are randomized (e.g., the number of refugees in a city in 2016 is moved to 2014, etc.), they are still likely to explain crime, albeit with smaller magnitudes.

The empirical counterpart to Equation 8 is the translog production function below.[34]

$$\ln(CR_{pt}) = \delta_0 + \sum_{i=1}^{4} \psi_i \ln(N_{ipt}) + \frac{1}{2}\sum_{i=1}^{4}\sum_{j=1}^{4} \lambda_{ij} \ln(N_{jpt}) + \sum_{k=1}^{m} \Omega_k X_{kpt} + \mu_{pt}, \tag{10}$$

where, for notational simplicity, N_4 represents the size of the refugee population (R), μ is the error term, and $\lambda_{ij} = \lambda_{ji}$. The vector $\boldsymbol{X}$ contains the same explanatory variables included in Equation 6.

Translog production functions (as well as the corresponding cost functions) do not impose restrictions on the underlying technology, regarding homotheticity, homogeneity, elasticities of substitution, economies of scale, and so on.[35] Because they represent a local second-order approximation to an arbitrary production function, the estimated technology parameters are more robust at the point of approximation, which is the sample mean, and translog and other flexible functional forms may perform poorly for data away from the approximation point (Caves and Christensen, 1980; Wales, 1977). Thus, we normalize the explanatory variables in Equation 9 by dividing them by their sample mean before taking logarithms.[36] Doing so gives an additional advantage that the elasticities are the first order parameters (ψ_i), when the estimated model is evaluated at the means.

Because this flexible specification involves the linear and quadratic terms of three native skill groups and the refugees, as well as their interactions, estimating it with instrumental variables is not feasible. However, it should be noted that all results obtained in the paper using OLS tend to underestimate the causal effect identified by the IV. Thus, estimating Equation 10 by OLS is not expected to provide an upwardly-biased estimate. With that proviso, we report the results in Table 5. The dependent variable is the number of new crimes handled by the prosecutors' offices in a year. Because data on population by education are not available at the province level prior to 2008, the sample period is 2008-2016. The four columns of the table display the result obtained from the same four specifications employed in Table 3. Because the explanatory variables are mean-scaled, the coefficients of the first-order terms are elasticities,

[34] As an alternative formulation, consider $CR = f(N_1^*, N_2^*, N_3^*, R^*, \boldsymbol{X})$, where N_1^*, N_2^*, N_3^*, and R^* represent the number of "efficiency-adjusted" individuals in each skill group skill. Consider w_1, w_2, w_3, and w_R as inverse-efficiency augmenting parameters, where

$$N_1 = (w_1)^{-1} N_1^*, \;\; N_2 = (w_2)^{-1} N_2^*, \;\; N_3 = (w_3)^{-1} N_3^*, \;\text{ and }\; R = (w_R)^{-1} R^*. \tag{9}$$

An increase (decrease) in w_1, w_2, w_3, and w_R would make individuals in the associated groups less (more) crime prone. An example of (w) would be the relevant labor market wages, variations in which would impact the number of crimes committed by the group without changing the size of the group (thus, the terminology of efficiency-adjusted). Using the elements of Equation 9 in Equation 8, and considering that the crime production function 8 can be described by a general translog specification yields Equation 10. The logarithmic transformation of the key input variables N_1, N_2, N_3, and R implies that the efficiency-augmenting terms w_1, w_2, w_3, and w_R are absorbed by the error term of Equation 10, and subsumed in the set of province, year, and region-by-year fixed effects which are also part of $\boldsymbol{X}$.

[35] Translog production functions have enjoyed widespread applications in a number of domains, ranging from production of health in Indonesia (Rosenzweig and Schultz, 1983), manufacturing in Japan (Nakamura, 1985), manufacturing in the U.S. (Kim, 1992; Berndt and Christensen, 1973), to fisheries (Kirkley et al., 1998), to agriculture (Chen and Gong, 2021).

[36] See Mocan (1997) and Vita (1990).

and higher-order terms drop out when evaluated at the means.

Table 5 reveals that the number of crimes is not influenced by an increase in the number of natives who are illiterate or who have no diploma. While this finding may be surprising at first sight, it should be noted that this uneducated group consists of older individuals, and those who are more likely to reside in rural areas.[37] Figure 12 in the Appendix A shows that the share of this group in the population has been declining, and Figure 13 reveals that the total number of people in this group is decreasing over time because older people with no education are dying, and younger generations have at least some education. Thus, although these older individuals constitute the lowest skill group, they are not the most crime-prone.

Table 5 shows that an increase in the size of native population with an elementary school or middle school degree, on the other hand, exerts a positive impact on crime. Specifically, if the number of natives in this low-skill group goes up by 2,000, this generates an increase in total offenses by 172. The number of refugees is highly significantly related to crime as well. An increase in refugee population by 2,000 triggers an increase in crime by 197 offenses received by the offices of the prosecutors.

The results displayed in Table 6 are obtained from translog regressions where the dependent variable is the number of charges in court cases. The magnitude of the refuge impact reported in this table is similar to the one obtained from Table 5. More specifically, using the information displayed in column (4) of Table 6, we calculate that if the refugee population goes up by 2,000 in an average province (which would constitute a 14 percent increase) 190 new charges would be generated in court cases. The impact on crime of native unskilled population is small and not different from zero in this specification.[38]

Note that the coefficient estimates of the refugee population from the translog specification reported in Tables 5 and 6 are smaller than those reported in Tables 3 and 4. This is likely because translog models are estimated using OLS, and OLS understates the coefficients in this case (compare the OLS and IV coefficients in Tables 3 and 4). But, regardless of the magnitude, an increase in refugee population leads to an increase in the incidence of crime.

Figure 14 in the Appendix A shows a linear trend in the logarithm of the annual stock of refugees. There was a total of 560,000 refugees in Turkey in 2013, and the number went up to 3.6 million in 2016. This indicates that a constant growth rate of 0.86 explains this pattern of growth. The average number of crimes was 40,000 offenses per province-year during this period. Using the

[37]The mandatory years of education in Turkey is eight years since the enactment of a law in 1997. The first fully impacted cohort was 1987—see Cesur and Mocan (2018). Compliance with the law is more than 90 percent, which means that individuals who were in their late teens-early twenties around 2015-2016 (a few years after the start of the refugee inflow) would have an average of more than 8 years of education.

[38]The statistically insignificant point estimate implies that an increase in the native population with an elementary or middle school degree by 2,000 would bring about only 25 additional crimes.

estimates of Tables 3 and 5, we find that the average annual rise in refugee population generated an additional 920 to1,858 crimes per province, or 74,500 to 150,500 additional crimes nationwide, where the former is likely an underestimate as it is obtained from translog models estimated by OLS. This corresponds to an increase in total incidence of crime in the range of approximately 2 percent to 4.5 percent annually.

10 Summary and discussion

Consequences of labor movements across nations have received increasing attention from social scientists. Economists analyzed the impact of international migration on a number of outcomes, ranging from the influence on the destination country labor markets to the effect on remittances, and welfare impacts in the countries of origin, to the extent of brain drain. Several analysts argue that eliminating barriers to immigration and allowing people to freely cross borders would generate substantial welfare gains (Clemens, 2011; Kennan, 2013; Moses and Letnes, 2004; Hamilton and Whalley, 1984), although a different perspective is provided by Borjas (2015). An optimal immigration policy, however, is difficult to determine and complicated to implement because of the distributional effects of immigration in destination countries, and also because of the resultant social and political repercussions (Dustmann and Preston, 2019).[39]

While international immigration is a voluntary phenomenon, refugee movements are involuntary, and they take place rapidly and in large magnitudes as exemplified by the recent influx of about four million Ukrainian refugees who fled their country in March and April of 2022, following the attack of Russia. The Refugee Agency of the United Nations (UNHCR) reports that more than 82 million people were forcibly displaced worldwide at the end of 2021, and that about 25 million of them were refugees, not including the Palestinians and the recent Ukrainian refugees. In addition to the refugee movements that are generally triggered by war and civil conflict, climate change is expected to force 20 million people to leave their homes and move to other areas each year, some of whom will be moving to neighboring countries (UNHCR).

The rapid and continued increase in refugee movements are expected to have socio-economic and political repercussions in destination countries with associated shocks in a variety of sectors ranging from labor and housing markets to the education and criminal justice systems. It is therefore important to investigate the economic and social repercussions of these large and

[39] A similar dilemma exists regarding free trade policy. While global free trade is welfare enhancing, resistance to free trade has been mounting along with its political consequences both because of the distributional effects of trade, and also because the impact of free trade has been observed in a variety of social outcomes. For example, analyzing the impact of China's rapid opening to international trade since the early 2000s following its entrance into the World Trade Organization, economists have shown that this China shock (Autor et al., 2013) lead to a decrease in employment and earnings of young adult men relative to women, which led to increases in the fraction of mothers who are unmarried, in the fraction of kids in poverty, and in premature mortality (Autor et al., 2019; Pierce and Schott, 2020). The exposure to this Chinese trade shock is also shown to impact schooling (Greenland and Lopresti, 2016), and crime (Che et al., 2018) in the U.S. It further led to political polarization with consequences on election outcomes (Autor et al., 2020).

sudden population movements because these analyses can help determine policies that would mitigate potential negative economic or social effects of these shocks to destination countries. This is all the more important for developing nations because the destination countries of refugee inflows, by and large, tend to be low-to-middle income countries with limited resources.

In this paper we analyze the impact of the Syrian refugee influx on criminal activity in Turkey. During the civil war in Syria, millions of Syrian refugees entered Turkey. While some of these refugees eventually went to various European countries, 3.7 million Syrians remained in Turkey, which generated an increase in the population of the county by 4.5 percent. We employ province-level data between 2006 and 2016 to investigate the extent to which variations in the refugee population within provinces generate an increase in crime. Considering the potential endogeneity of the geographic distribution of the refugee population, we use a distance-to-the border-based instrument (Tumen, 2021; Del Carpio and Wagner, 2015). We use the flow of new crimes handled by the prosecutors' offices each year as our main crime measure, although we also use a noisy crime indicator (offenses charged in courts) as an alternative outcome. The results show that an increase in the refugee population has a significant positive impact on crime. This result is robust to a wide range of sensitivity analyses including the use of alternative instruments, altering the model specification, changing the analysis sample by time period, or by region, and so on. Placebo exercises further indicate the credibility of the estimated parameters.

The finding that the refugee population has an impact on total crime committed does not imply that the entire increase in crime is attributable to the increase in refugee population. This is because, part (or all) of the increase in crime may have been generated by native population in response to the changes in labor market conditions, triggered by the refugee inflow. To investigate this point further, we estimate translog crime production functions, which consider three groups of natives, classified by education, and refugees as distinct inputs. These models confirm that the refugee population is a significant determinant of crime. They, however, also reveal that an increase in low-skilled native population generates an increase in total criminal activity as well. These results are consistent with theoretical models of crime which postulate that the propensity to engage in criminal activity is driven by both labor market variables and individual attributes of marginal criminals. We briefly discuss, in Appendix C, the way in which the interplay between refugee labor and unskilled native labor (and the degree of substitutability between them) relates to the relative difference between the impacts of refugees and natives on crime. We also demonstrate in Appendix B how incorrect inference regarding the impact of refugees on crime can be obtained by using the wrong model specification as was done by a couple of recent papers.

Our results show that the increase in refugee population in Turkey between 2013 and 2016 led to an increase in the incidence of crime by 2-4.75 percent per year, which corresponds to about 75,000 to 150,000 additional crimes per year. Crime is associated with substantial costs to society (Anderson, 2021). Because developing nations are bigger targets of refugee inflows from their low-income neighbors, and because they possess modest levels of human capital, a rise in criminal activity arguably poses more significant problems for developing countries. This is because, it has been shown that switching away from the legal labor market to crime lowers legal human capital, increases criminal human capital, and creates path-dependence in criminal activity (Mocan and Bali, 2010; Mocan et al., 2005). An increase in crime has also intergenerational implications (Damm and Dustmann, 2014; Hjalmarsson and Lindquist, 2012). Thus, a rise in crime may generate potential development bottlenecks for low-income countries. An increase in crime has also implications for refugees, including increased animosity towards them which could impact the well-being of refugees directly (e.g., discrimination in the labor and housing markets), and indirectly through domestic politics. Thus, our results highlight the need to quickly strengthen the social safety systems, take actions to counter the impact on the labor market, and provide support to the criminal justice system to mitigate the repercussions of massive influx of individuals into a country.[40] Other methods could include procedures that would enable integration of refugees into natives' social networks (Bailey et al., 2022), and increasing the level of local civic engagement (Barreto et al., 2022).

[40] For example, Angrist and Kugler (2003) show that the negative impact of immigrants on labor markets (e.g., native job losses) in Europe are larger in countries with more rigid labor and product markets.

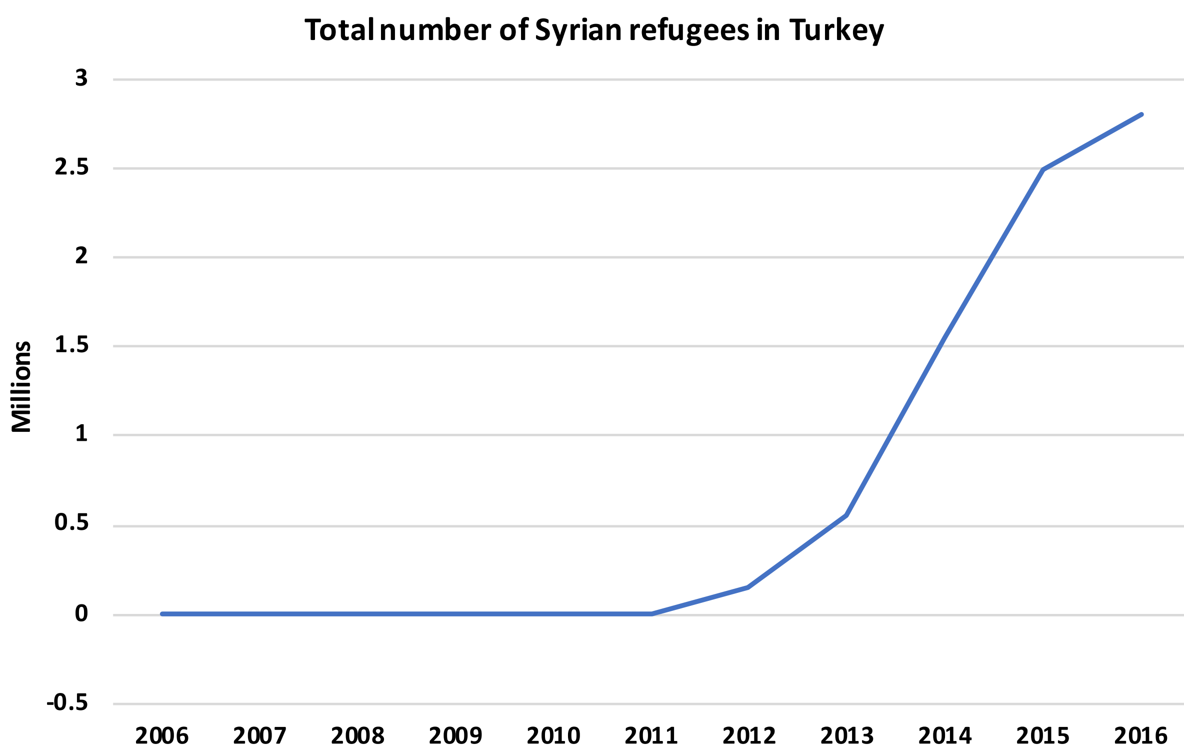

Figure 1: Source: Aggregate data on Syrian refugees are obtained from the UNHCR; see, https://data.unhcr.org/en/situations/syria/location/113.

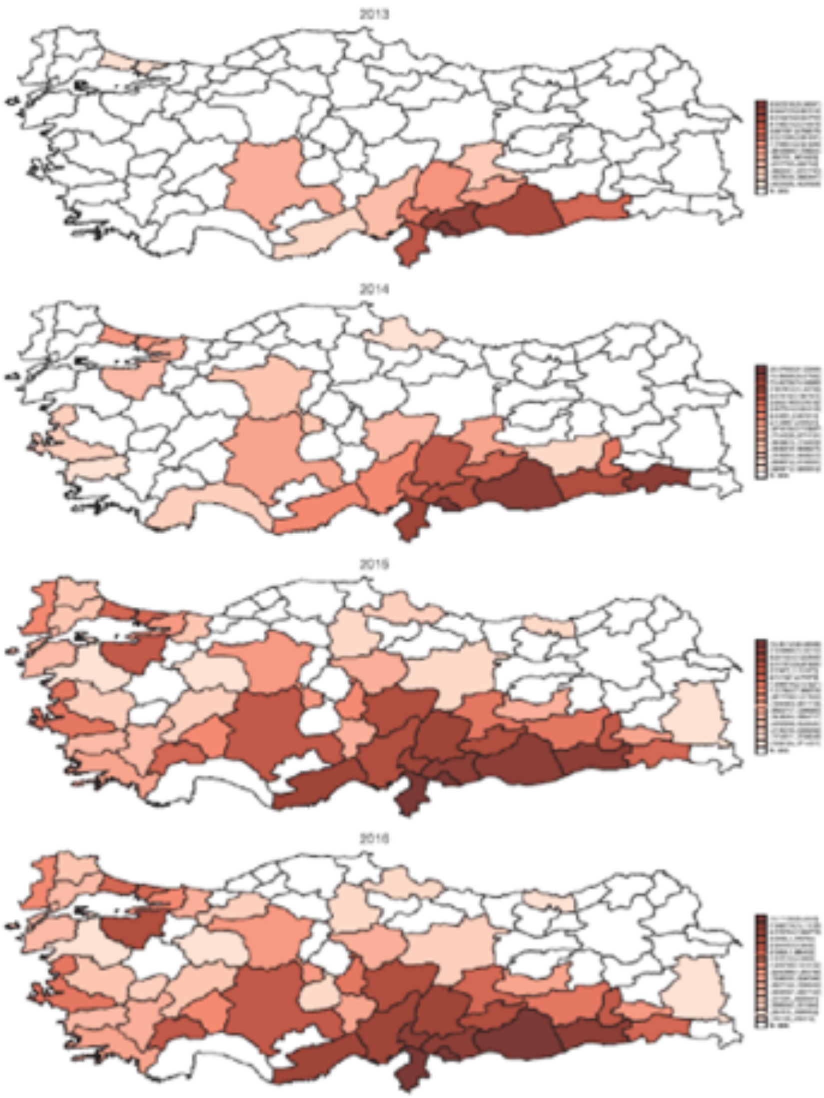

Figure 2: Province-level refugee concentration by year. Data on Syrian refugees are obtained from Turkish Disaster and Emergency Management Presidency (AFAD) for 2013; Erdogan (2014) for 2014; and Ministry of Interior Directorate General of Migration Management for 2015 and 2016.

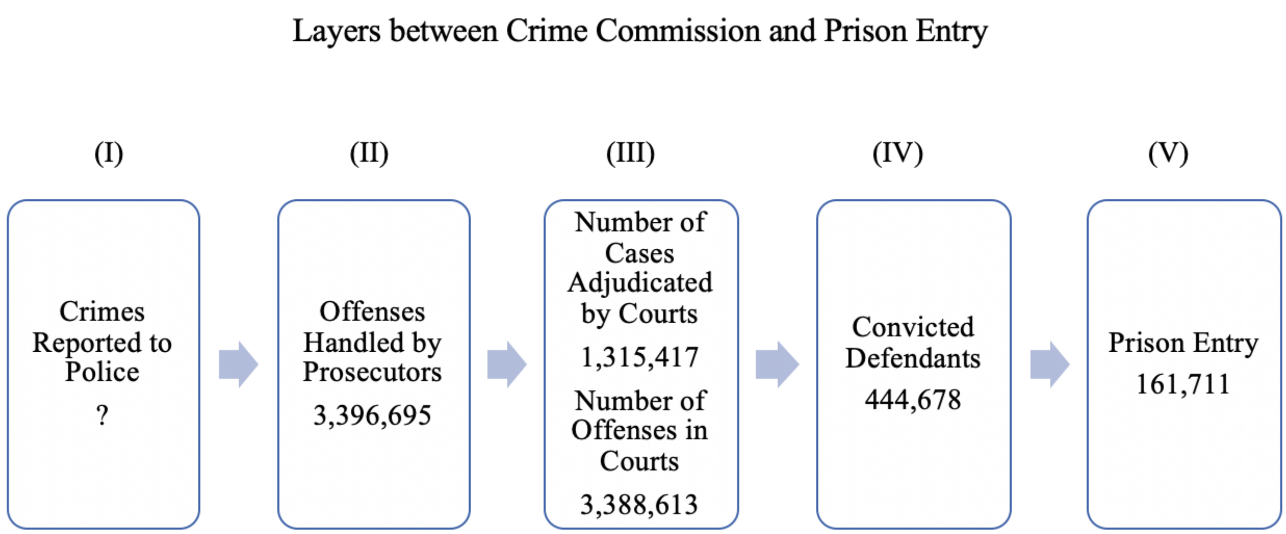

Figure 3: Layers between crime commission and prison entry.

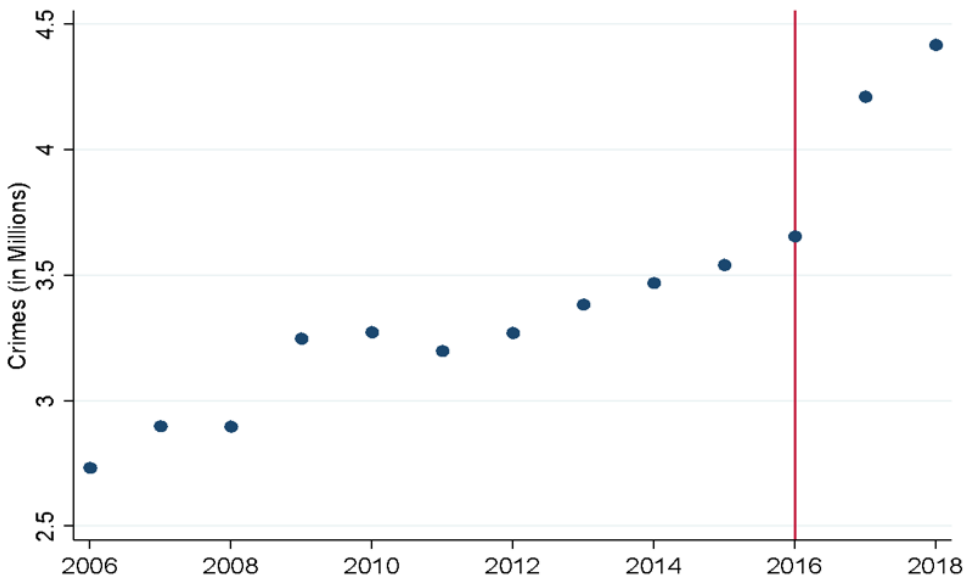

Figure 4: Crimes handled by prosecutors' offices. Source: TurkStat.

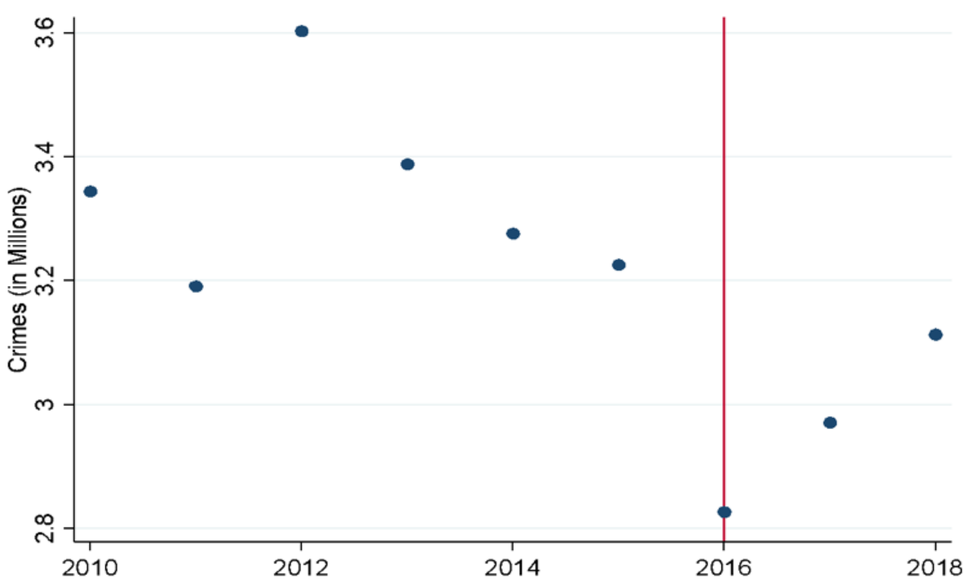

Figure 5: Number of charges in courts. Source: TurkStat.

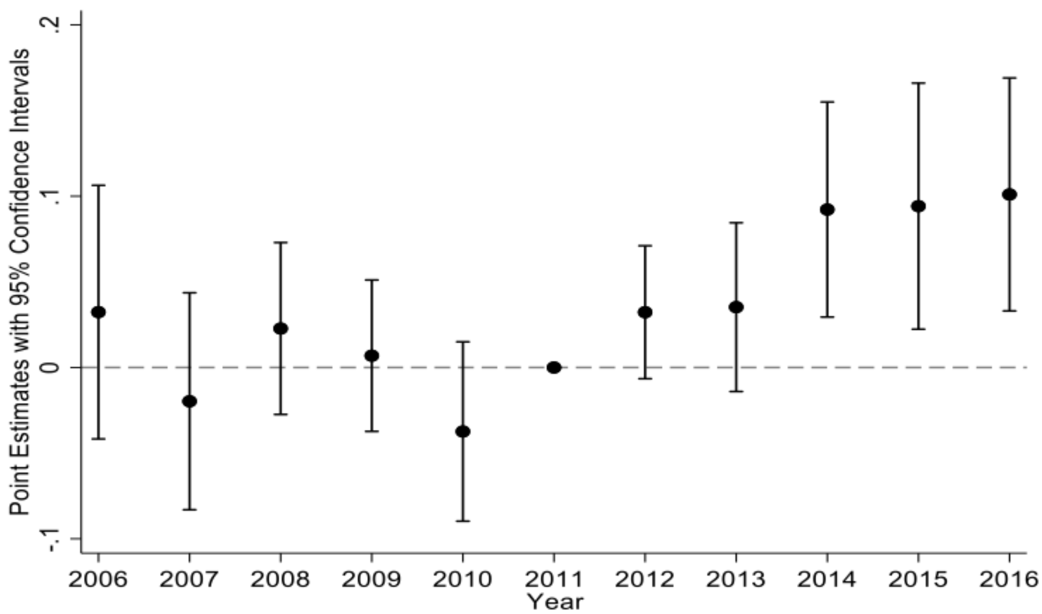

Figure 6: Event study: The difference in refugee impact between high vs. low refugee exposure provinces.

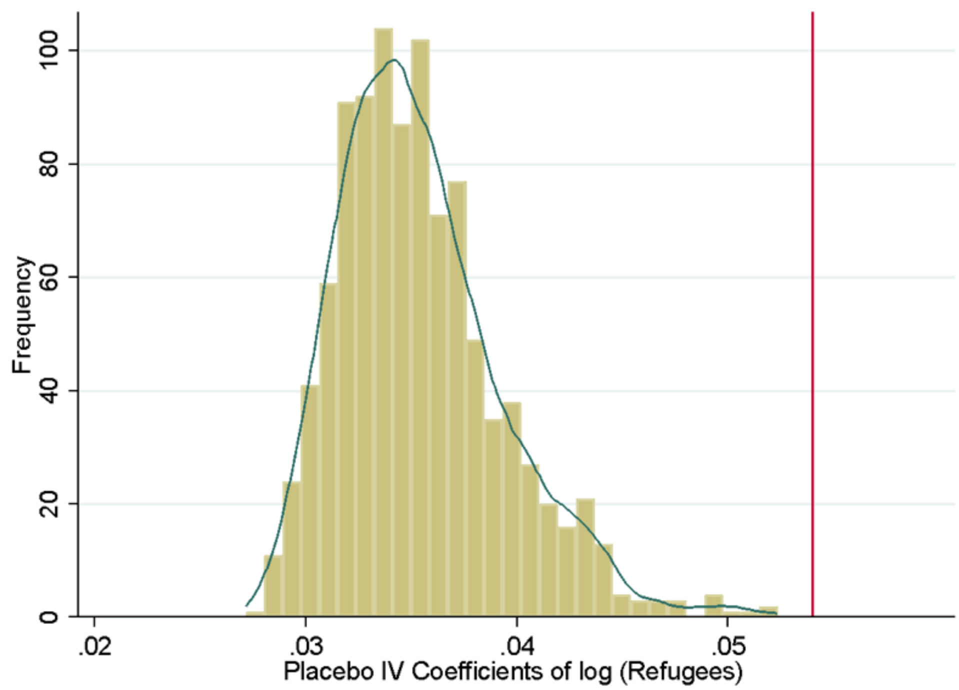

Figure 7: The distribution of placebo IV coefficients vs. the estimated IV coefficient.

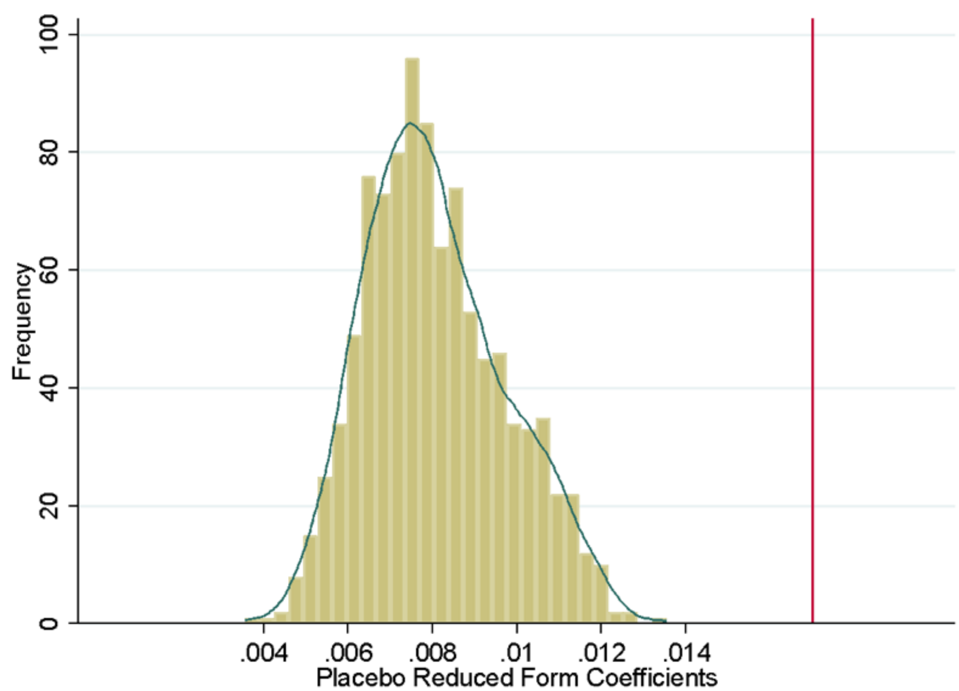

Figure 8: The distribution of the placebo reduced-form coefficients vs. the estimated reduced form coefficient.

Descriptive Statistics

	All	Before Syrian refugees	After Syrian refugees
		(2006-2012)	(2013-2016)
	(1)	(2)	(3)
Number of refugees	8,892	0	24,453
	(44,323)	(0.000)	(70,932)
Number of new crimes handled by prosecutors' offices	40,039	38,110	43,414
	(80,720)	(81,236)	(79,823)
Number of new charges in crime courts	40,312	41,724	39,253
	(77,612)	(85,151)	(71,573)
Population	922,180	897,253	965,802
	(1,627,010)	(1,560,297)	(1,739,242)
Prison intake	1,491	1,128	2,126
	(2,547)	(1,919)	(3,285)
Hospitals per-100k population	2.354	2.289	2.468
	(0.992)	(0.986)	(0.993)
Natural gas	0.681	0.564	0.886
	(0.466)	(0.496)	(0.319)
Public budget per-100k population	36,494	27,772	51,758
	(39,081)	(38,707)	(34,870)
Number of observations	891	567	324

Table 1: Data are obtained from TurkStat. The entries are the means; standard deviations are reported in parentheses. The unit of observation is province-year.

The relationship between refugee flows and province attributes
Dependent variable: Log of the number of refugees in each province

	(1)	(2)	(3)
Log of population	11.545***	3.174	3.174
	(3.136)	(2.467)	(2.451)
Log of hospitals per-100k population	0.231	-0.041	-0.041
	(0.561)	(0.368)	(0.366)
Any natural gas	0.070	-0.231	-0.231
	(0.309)	(0.215)	(0.214)
Log of public expenditure per-100k population	0.111	0.122	0.122
	(0.249)	(0.212)	(0.210)
Prison intake	-0.269	-0.272	-0.272
	(0.345)	(0.236)	(0.234)
Adjusted R^2	0.835	0.920	0.917
Joint F-test	3.04	0.92	0.92
p-value of joint F-test	0.015	0.471	0.464
Year FE	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
5-region linear and quadratic trends	No	Yes	No
12-region-year FE	No	Yes	Yes
Drop Istanbul	No	No	Yes
Number of observations	891	891	880

Table 2: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01).

The impact of refugees on crime (2006-2016)

Dependent variable: Number of new crimes handled by prosecutors' offices

	(1)	(2)	(3)	(4)	(5)
	Panel A: OLS				
$\ln(R)$	0.010**	0.010**	0.004	0.014**	
	(0.005)	(0.005)	(0.005)	(0.005)	
$\ln(N)$	0.539***	0.532**	0.287	0.440**	
	(0.198)	(0.217)	(0.207)	(0.189)	
	Panel B: IV				
$\ln(R)$	0.052***	0.049***	0.055*	0.054**	
	(0.018)	(0.016)	(0.031)	(0.021)	
$\ln(N)$	0.068	0.084	0.125	0.424**	
	(0.301)	(0.303)	(0.264)	(0.201)	
					Reduced form
Instrument/1000					0.017**
					(0.007)
$\ln(N)$					0.290
					(0.194)
F-test for the first stage	51.81	50.14	40.92	48.65	
Mean of the dependent variable	40,039	40,039	40,039	40,039	40,039
Mean of refugees	8,892	8,892	8,892	8,892	8,892
Mean of native population	922,180	922,180	922,180	922,180	922,180
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes	Yes
12-region-year FE	No	No	Yes	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes	No
Number of observations	891	891	891	891	891

Table 3: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per-100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per-100k population, and the lag of prison intake.

The impact of refugees on crime (2010-2016)

Dependent variable: Number of charges in court cases

	(1)	(2)	(3)	(4)	(5)
	Panel A: OLS				
$\ln(R)$	0.004	0.005	0.001	0.007*	
	(0.003)	(0.003)	(0.004)	(0.004)	
$\ln(N)$	0.703**	0.683**	0.616*	0.672**	
	(0.282)	(0.294)	(0.318)	(0.337)	
	Panel B: IV				
$\ln(R)$	0.045**	0.040***	0.079*	0.084*	
	(0.020)	(0.015)	(0.046)	(0.044)	
$\ln(N)$	0.075	0.118	0.417	0.642	
	(0.431)	(0.375)	(0.458)	(0.473)	
					Reduced form
Instrument/1000					0.018***
					(0.007)
$\ln(N)$					0.569*
					(0.320)
F-test for the first stage	31.35	31.08	30.89	34.32	
Mean of the dependent variable	40,312	40,312	40,312	40,312	40,312
Mean of refugees	13,973	13,973	13,973	13,973	13,973
Mean of native population	947,080	947,080	947,080	947,080	947,080
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes	Yes
12-region-year FE	No	No	Yes	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes	No
Number of observations	567	567	567	567	567

Table 4: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per-100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per-100k population, and the lag of prison intake.

The impact of refugees on crime (2008-2016) – Translog models
Dependent variable: Number of new crimes handled by prosecutors' offices

	(1)	(2)	(3)	(4)
$\ln(N_1)$	-0.004	-0.039	0.018	-0.049
	(0.153)	(0.156)	(0.187)	(0.212)
$\ln(N_2)$	0.867***	0.903***	0.649**	0.774**
	(0.297)	(0.297)	(0.305)	(0.322)
$\ln(N_3)$	-0.139	-0.172	-0.141	-0.068
	(0.300)	(0.297)	(0.268)	(0.360)
$\ln(R)$	0.032***	0.032***	0.026***	0.026**
	(0.008)	(0.009)	(0.010)	(0.010)
Mean of the dependent variable	41,183	41,183	41,183	41,183
Mean of N_1	86,851	86,851	86,851	86,851
Mean of N_2	371,460	371,460	371,460	371,460
Mean of N_3	235,858	235,858	235,858	235,858
Mean of refugees	10,868	10,868	10,868	10,868
Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes
12-region-year FE	No	No	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes
Number of observations	729	729	729	729

Table 5: N_1, N_2, and N_3 refer to natives with no education/illiterate, elementary/middle school education, and high school education or above, respectively. Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per-100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per-100k population, and the lag of prison intake. The model also includes the interactions between the four population categories listed in the table, as well as their quadratic terms.

The impact of refugees on crime (2010-2016) – Translog models
Dependent variable: Number of charges in court cases

	(1)	(2)	(3)	(4)
$\ln(N_1)$	0.210	0.264	0.035	0.177
	(0.333)	(0.337)	(0.493)	(0.529)
$\ln(N_2)$	0.109	0.156	0.227	-0.005
	(0.399)	(0.393)	(0.531)	(0.592)
$\ln(N_3)$	0.043	-0.053	-0.126	0.439
	(0.289)	(0.287)	(0.344)	(0.428)
$\ln(R)$	0.029***	0.029***	0.027*	0.030*
	(0.009)	(0.008)	(0.015)	(0.017)
Mean of the dependent variable	40,312	40,312	40,312	40,312
Mean of N_1	80,916	80,916	80,916	80,916
Mean of N_2	375,500	375,500	375,500	375,500
Mean of N_3	248,807	248,807	248,807	248,807
Mean of refugees	13,973	13,973	13,973	13,973
Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes
12-region-year FE	No	No	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes
Number of observations	567	567	567	567

Table 6: N_1, N_2, and N_3 refer to natives with no education/illiterate, elementary/middle school education, and high school education or above, respectively. Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per-100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per-100k population, and the lag of prison intake. The model also includes the interactions between the four population categories listed in the table, as well as their quadratic terms.

References

Aebi, M. F., G. Akdeniz, G. Barclay, C. Campistol, S. Caneppele, B. Gruszczynska, S. Harrendorf, M. Heiskanen, V. Hysi, J. M. Jehle, and A. Jokinen (2014): *European sourcebook of crime and criminal justice statistics*, Helsinki European United Nations Institute.

Akdeniz, G. (2019): "Homicide trends in Turkey: A review of publicly available data," *Journal of Penal Law and Criminology*, 7, 1–25.

Akgunduz, Y. E. and H. Torun (2020): "Two and a half million Syrian refugees, tasks and capital intensity," *Journal of Development Economics*, 145, 102470.

Akgunduz, Y. E., M. van den Berg, and W. Hassink (2018): "The impact of the Syrian refugee crisis on firm entry and performance in Turkey," *World Bank Economic Review*, 32, 19–40.

Aksu, E., R. Erzan, and M. G. Kirdar (2022): "The impact of mass migration of Syrians on the Turkish labor market," *Labour Economics*, 76, 102183.

Akyol, S. P. and H. N. Mocan (2020): "Education and consanguineous marriage," NBER Working Paper #28212.

Altindag, O., O. Bakis, and S. V. Rozo (2020): "Blessing or burden? Impacts of refugees on businesses and the informal economy," *Journal of Development Economics*, 146, 102490.

Altindag, O. and N. Kaushal (2021): "Do refugees impact voting behavior in the host country? Evidence from Syrian refugee inflows to Turkey," *Public Choice*, 186, 149–178.

Anastasopoulos, L. J., G. J. Borjas, G. G. Cook, and M. Lachanski (2021): "Job vacancies and immigration: Evidence from the Mariel supply shock," *Journal of Human Capital*, 15, 1–33.

Anderson, D. A. (2021): "The aggregate cost of crime in the United States," *Journal of Law and Economics*, 64, 857–885.

Angrist, J. D. and A. D. Kugler (2003): "Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives," *Economic Journal*, 113, F302–F331.

Autor, D., D. Dorn, and G. Hanson (2013): "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 103, 2121–2168.

——— (2019): "When work disappears: Manufacturing decline and the falling marriage market value of young men," *American Economic Review: Insights*, 1, 161–178.

Autor, D., D. Dorn, G. Hanson, and K. Majlesi (2020): "Importing political polarization? The electoral consequences of rising trade exposure," *American Economic Review*, 110, 3139–3183.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008): "Trends in US wage inequality: Revising the revisionists," *Review of Economics and Statistics*, 90, 300–323.

Bagir, Y. K. (2018): "Impact of the Syrian refugee influx on Turkish native workers: An ethnic enclave approach," *Central Bank Review*, 18, 129–147.

Bailey, M., D. M. Johnston, M. Koenen, T. Kuchler, D. Russel, and J. Stroebel (2022): "The social integration of international migrants: Evidence from the networks of Syrians in Germany," NBER Working Paper #29925.

Baker, S. R. (2015): "Effects of immigrant legalization on crime," *American Economic Review*, 105, 210–213.

Balkan, B. and S. Tumen (2016): "Immigration and prices: Quasi-experimental evidence from Syrian refugees in Turkey," *Journal of Population Economics*, 29, 657–686.

Barbarino, A. and G. Mastrobuoni (2014): "The incapacitation effect of incarceration: Evidence from several Italian collective pardons," *American Economic Journal: Economic Policy*, 6, 1–37.

Barreto, C., P. Berbee, K. G. Torres, M. Lange, and K. Sommerfeld (2022): "The civic engagement and social integration of refugees in Germany," *Nonprofit Policy Forum*, 13, 161–174.

Battisti, M., G. Felbermayr, G. Peri, and P. Poutvaara (2018): "Immigration, search and redistribution: A quantitative assessment of native welfare," *Journal of the European Economic Association*, 16, 1137–1188.

Bazzi, S. and M. A. Clemens (2013): "Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth," *American Economic Journal: Macroeconomics*, 5, 152–186.

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

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

Berndt, E. R. and L. R. Christensen (1973): "The translog function and the substitution of equipment, structures, and labor in US manufacturing 1929-68," *Journal of Econometrics*, 1, 81–113.

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

Borjas, G. J. (2003): "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market," *Quarterly Journal of Economics*, 118, 1335–1374.

——— (2013): "The analytics of the wage effect of immigration," *IZA Journal of Migration*, 2, 1–25.

——— (2015): "Immigration and globalization: A review essay," *Journal of Economic Literature*, 70, 1077–1110.

——— (2017): "The wage impact of the Marielitos: A reappraisal," *Industrial and Labor Relations Review*, 70, 1077–1110.

Borjas, G. J., J. Grogger, and G. H. Hanson (2010): "Immigration and the economic status of African-American men," *Economica*, 77, 255–282.

Borjas, G. J. and J. Monras (2017): "The labour market consequences of refugee supply shocks," *Economic Policy*, 32, 361–413.

Boylan, R. T. and H. N. Mocan (2014): "Intended and unintended consequences of prison reform," *Journal of Law, Economics, and Organization*, 30, 558–586.

Bratsberg, B. and O. Raaum (2012): "Immigration and wages: Evidence from construction," *Economic Journal*, 122, 1177–1205.

Buonanno, P., F. Drago, R. Galbiati, and P. Vertova (2018): "How much should we trust crime statistics? A comparison between EU and US," *European Journal of Law and Economics*, 46, 343–357.

Buonanno, P. and S. Raphael (2013): "Incarceration and incapacitation: Evidence from the 2006 Italian collective pardon," *American Economic Review*, 103, 2437–2465.

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

——— (1998b): "Recent immigrants: Unexpected implications for crime and incarceration," *Industrial and Labor Relations Review*, 51, 654–679.

Card, D. (1990): "The impact of the Mariel boatlift on the Miami labor market," *Industrial and Labor Relations Review*, 43, 245–257.

——— (2001): “Immigrant inflows, native outflows, and the local labor market impacts of higher immigration,” *Journal of Labor Economics*, 19, 22–64.

Card, D. and J. E. DiNardo (2002): “Skill-biased technological change and rising wage inequality: Some problems and puzzles,” *Journal of Labor Economics*, 20, 733–783.

Cassar, A., A. Healy, and C. Von Kessler (2017): “Trust, risk, and time preferences after a natural disaster: experimental evidence from Thailand,” *World Development*, 94, 90–105.

Caves, D. W. and L. R. Christensen (1980): “Global properties of flexible functional forms,” *American Economic Review*, 70, 422–432.

Cengiz, D. and H. Tekguc (2022): “Is it merely a labor supply shock? Impacts of Syrian migrants on local economies in Turkey,” *Industrial and Labor Relations Review*, 75, 741–768.

Ceritoglu, E., H. B. Gurcihan-Yunculer, H. Torun, and S. Tumen (2017): “The impact of Syrian refugees on natives’ labor market outcomes in Turkey: Evidence from a quasi-experimental design,” *IZA Journal of Labor Policy*, 6, 1–28.

Cesur, R. and N. Mocan (2018): “Education, religion, and voter preference in a Muslim country,” *Journal of Population Economics*, 31, 1–44.

Che, Y., X. Xu, and Y. Zhang (2018): “Chinese import competition, crime, and government transfers in US,” *Journal of Comparative Economics*, 46, 544–567.

Chen, S. and B. Gong (2021): “Response and adaptation of agriculture to climate change: Evidence from China,” *Journal of Development Economics*, 148, 102557.

Clemens, M. A. (2011): “Economics and emigration: Trillion-dollar bills and the sidewalk?” *Journal of Economic Perspectives*, 25, 83–106.

Clemens, M. A. and J. Hunt (2019): “The labor market effects of refugee waves: Reconciling conflicting results,” *Industrial and Labor Relations Review*, 72, 818–857.

Corman, H. and H. N. Mocan (2005): “Carrots, sticks, and broken windows,” *Journal of Law and Economics*, 48, 235–266.

Couttenier, M., V. Petrencu, D. Rohner, and M. Thoenig (2019): “The violent legacy of conflict: Evidence on asylum seekers, crime, and public policy in Switzerland,” *American Economic Review*, 109, 4378–4425.

Damm, A. P. and C. Dustmann (2014): “Does growing up in a high crime neighborhood affect youth criminal behavior?” *American Economic Review*, 104, 1806–1832.

De Chaisemartin, C. and X. d’Haultfoeuille (2020): “Two-way fixed effects estimators with heterogeneous treatment effects,” *American Economic Review*, 110, 2964–2996.

Dehos, F. T. (2021): "The refugee wave to Germany and its impact on crime," *Regional Science and Urban Economics*, 88, 103640.

Del Carpio, X. V. and M. C. Wagner (2015): "The impact of Syrian refugees on the Turkish labor market," World Bank Policy Research Working Paper #7402.

Drago, F., R. Galbiati, and P. Vertova (2009): "The deterrent effects of prison: Evidence from a natural experiment," *Journal of Political Economy*, 117, 257–280.

Dustmann, C., T. Frattini, and I. P. Preston (2012): "Rethinking the effect of immigration on wages," *Journal of the European Economic Association*, 10, 152–197.

——— (2013): "The effect of immigration along the distribution of wages," *Review of Economic Studies*, 80, 145–173.

Dustmann, C. and I. P. Preston (2019): "Free movement, open borders, and the global gains from labor mobility," *Annual Review of Economics*, 11, 783–803.

Dustmann, C., U. Schonberg, and J. Stuhler (2016): "The impact of immigration: Why do studies reach such different results?" *Journal of Economic Perspectives*, 30, 31–56.

Eckel, C. C., M. A. El-Gamal, and R. K. Wilson (2009): "Risk loving after the storm: A bayesian-network study of Hurricane Katrina evacuees," *Journal of Economic Behavior & Organization*, 69, 110–124.

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

Erdogan, M. M. (2014): "Perceptions of Syrians in Turkey," *Insight Turkey*, 16, 65–75.

Ferraz, E. and R. R. Soares (2022): "Socially optimal crime and punishment," IZA Discussion Paper #15053.

Freedman, M., E. Owens, and S. Bohn (2018): "Immigration, employment opportunities, and criminal behavior," *American Economic Journal: Economic Policy*, 10, 117–151.

Friedberg, R. M. (2001): "The impact of mass migration on the Israeli labor market," *Quarterly Journal of Economics*, 116, 1373–1408.

Gehrsitz, M. and M. Ungerer (2022): "Jobs, crime and votes: A short-run evaluation of the refugee crisis in Germany," *Economica*, 89, 592–626.

Greenland, A. and J. Lopresti (2016): "Import exposure and human capital adjustment: Evidence from the US," *Journal of International Economics*, 100, 50–60.

Gutierrez, I. A. and J. V. Gallegos (2016): "The effect of civil conflict on domestic violence: The case of Peru," RAND Working Paper WR-1168, RAND Corporation.

Hamilton, B. and J. Whalley (1984): "Efficiency and distributional implications of global restrictions on labour mobility: Calculations and policy implications," *Journal of Development Economics*, 14, 61–75.

Hanaoka, C., H. Shigeoka, and Y. Watanabe (2018): "Do risk preferences change? Evidence from the great East Japan earthquake," *American Economic Journal: Applied Economics*, 10, 298–330.

Hjalmarsson, R. and M. J. Lindquist (2012): "Like godfather, like son: Exploring the intergenerational nature of crime," *Journal of Human Resources*, 47, 550–582.

Hunt, J. (1992): "The impact of the 1962 repatriates from Algeria on the French labor market," *Industrial and Labor Relations Review*, 45, 556–572.

Johnson, R. and S. Raphael (2012): "How much crime reduction does the marginal prisoner buy?" *American Economic Journal: Economic Policy*, 55, 275–310.

Kayaoglu, A. (2022): "Do refugees cause crime?" *World Development*, 154, 105858.

Kennan, J. (2013): "Open borders," *Review of Economic Dynamics*, 16, L1–L13.

Kim, H. Y. (1992): "The translog production function and variable returns to scale," *Review of Economics and Statistics*, 74, 546–552.

Kirdar, M. G., I. L. Cruz, and B. Turkum (2022): "The effect of 3.6 million refugees on crime," *Journal of Economic Behavior & Organization*, 194, 568–582.

Kirkley, J., D. Squires, and I. E. Strand (1998): "Characterizing managerial skill and technical efficiency in a fishery," *Journal of Productivity Analysis*, 9, 145–160.

Kronmal, R. A. (1993): "Spurious correlation and the fallacy of the ratio standard revisited," *Journal of the Royal Statistical Society: Series A*, 156, 379–392.

Kuziemko, I. (2013): "How should inmates be released from prison? An assessment of parole versus fixed-sentence regimes," *Quarterly Journal of Economics*, 128, 371–424.

La Mattina, G. (2017): "Civil conflict, domestic violence, and intra-household bargaining in post-genocide Rwanda," *Journal of Development Economics*, 124, 168–198.

LaLonde, R. J. and R. H. Topel (1991): "Immigrants in the American labor market: Quality, assimilation, and distributional effects," *American Economic Review*, 81, 297–302.

Lange, M. and K. Sommerfeld (2021): "Do refugees impact crime? Causal evidence from large-scale refugee immigration," ZEW Working Paper.

Levitt, S. D. (1996): "The effect of prison population size on crime rates: Evidence from prison overcrowding litigation," *Quarterly Journal of Economics*, 111, 319–351.

Lochner, L. (2004): "Education, work, and crime: A human capital approach," *International Economic Review*, 45, 811–843.

Masterson, D. and V. Yasenov (2021): "Does halting refugee resettlement reduce crime? Evidence from the US refugee ban," *American Political Science Review*, 115, 1066–1073.

Mastrobuoni, G. and P. Pinotti (2015): "Legal status and the criminal activity of immigrants," *American Economic Journal: Applied Economics*, 7, 175–206.

Megalokonomou, R. and C. Vasilakis (2020): "The effect of the exposure to refugees on crime activity: Evidence from the Greek islands," SSRN Working Paper #3566564.

Miguel, E., S. M. Saiegh, and S. Satyanath (2011): "Civil war exposure and violence," *Economics & Politics*, 23, 59–73.

Mocan, H. N. (1997): "Cost functions, efficiency, and quality in day care centers," *Journal of Human Resources*, 32, 861–891.

Mocan, H. N., S. C. Billups, and J. Overland (2005): "A dynamic model of differential human capital and criminal activity," *Economica*, 72, 655–681.

Mocan, N. H. and T. G. Bali (2010): "Asymmetric crime cycles," *Review of Economics and Statistics*, 92, 899–911.

Moses, J. W. and B. Letnes (2004): "The economic costs to international labor restrictions: Revisiting the empirical discussion," *World Development*, 32, 1609–1626.

Moya, A. and M. R. Carter (2019): "Violence and the formation of hopelessness: Evidence from internally displaced persons in Colombia," *World Development*, 113, 100–115.

Nakamura, R. (1985): "Agglomeration economies in urban manufacturing industries: A case of Japanese cities," *Journal of Urban Economics*, 17, 108–124.

Noe, D. and J. Rieckmann (2013): "Violent behaviour: The effect of civil conflict on domestic violence in Colombia," Discussion Paper #136, University of Göttingen, Courant Research Centre.

Page, L., D. A. Savage, and B. Torgler (2014): "Variation in risk seeking behaviour following large losses: A natural experiment," *European Economic Review*, 71, 121–131.

Peri, G. and V. Yasenov (2019): "The labor market effects of a refugee wave synthetic control method meets the Mariel boatlift," *Journal of Human Resources*, 54, 267–309.

Pierce, J. R. and P. K. Schott (2020): "Trade liberalization and mortality: Evidence from US counties," *American Economic Review: Insights*, 2, 47–64.

PINOTTI, P. (2017): "Clicking on heaven's door: The effect of immigrant legalization on crime," *American Economic Review*, 107, 138–168.

PIOPIUNIK, M. AND J. RUHOSE (2017): "Immigration, regional conditions, and crime: Evidence from an allocation policy in Germany," *European Economic Review*, 92, 258–282.

ROHNER, D., M. THOENIG, AND F. ZILIBOTTI (2013): "Seeds of distrust: Conflict in Uganda," *Journal of Economic Growth*, 18, 217–252.

ROSE, E. K. AND Y. SHEM-TOV (2021): "How does incarceration affect reoffending? estimating the dose-response function," *Journal of Political Economy*, 129, 3302–3356.

ROSENZWEIG, M. R. AND T. P. SCHULTZ (1983): "Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight," *Journal of Political Economy*, 91, 723–746.

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

STEPHENS, J. M. AND D. Y. YANG (2014): "Compulsory education and the benefits of schooling," *American Economic Review*, 104, 1777–1792.

TORUN, H. (2018): "Compulsory schooling and early labor market outcomes in a middle-income country," *Journal of Labor Research*, 39, 277–305.

TUMEN, S. (2016): "The economic impact of Syrian refugees on host countries: Quasi-experimental evidence from Turkey," *American Economic Review*, 106, 456–460.

——— (2018): "The impact of low-skill refugees on youth education," IZA Discussion Paper #11869.

——— (2021): "The effect of refugees on native adolescents' test scores: Quasi-experimental evidence from PISA," *Journal of Development Economics*, 150, 102633.

UNHCR (2021): "Global Trends in Forced Displacement – 2020," Tech. rep., The UN Refugee Agency.

VITA, M. G. (1990): "Exploring hospital production relationships with flexible functional forms," *Journal of Health Economics*, 9, 1–21.

VOORS, M. J., E. E. NILLESEN, P. VERWIMP, E. H. BULTE, R. LENSINK, AND D. P. VAN SOEST (2012): "Violent conflict and behavior: A field experiment in Burundi," *American Economic Review*, 102, 941–964.

WALES, T. J. (1977): "On the flexibility of flexible functional forms: An empirical approach," *Journal of Econometrics*, 5, 183–193.

A Appendix: Additional analyses and robustness checks

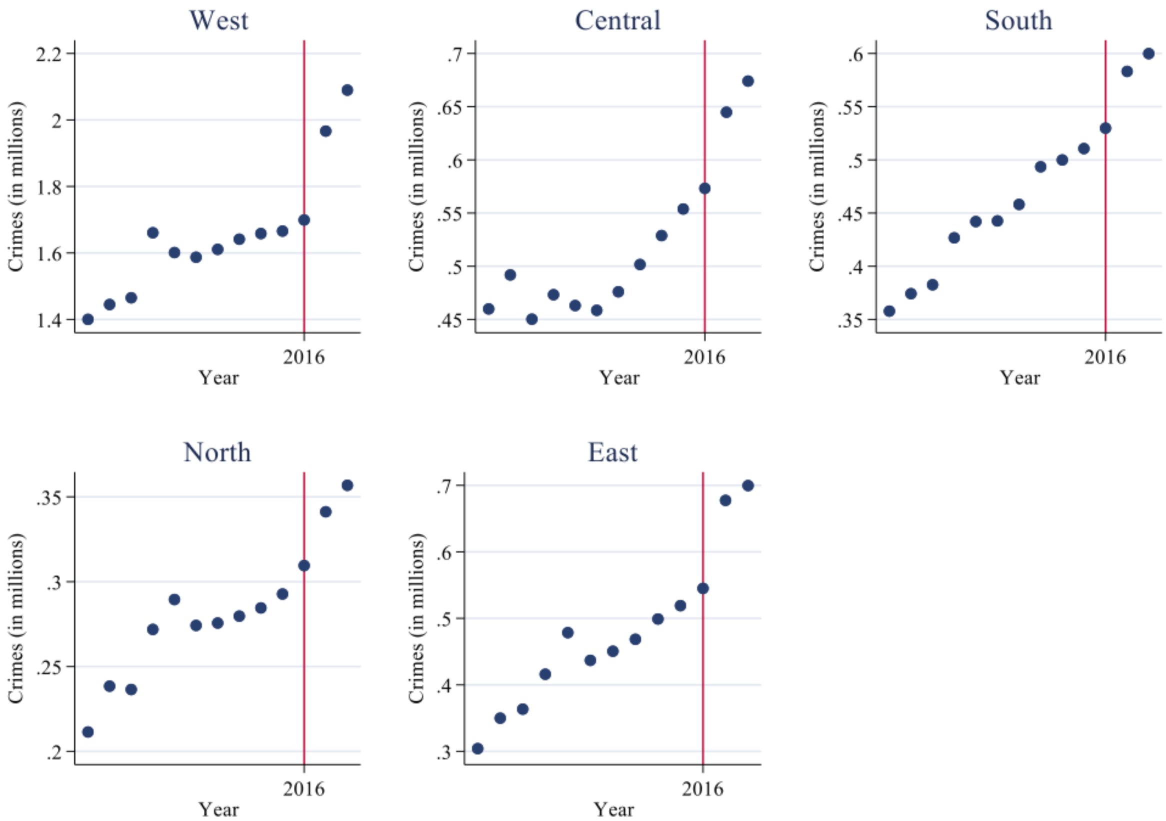

Figure 9: Crimes at prosecutors' offices per region. Source: TurkStat.

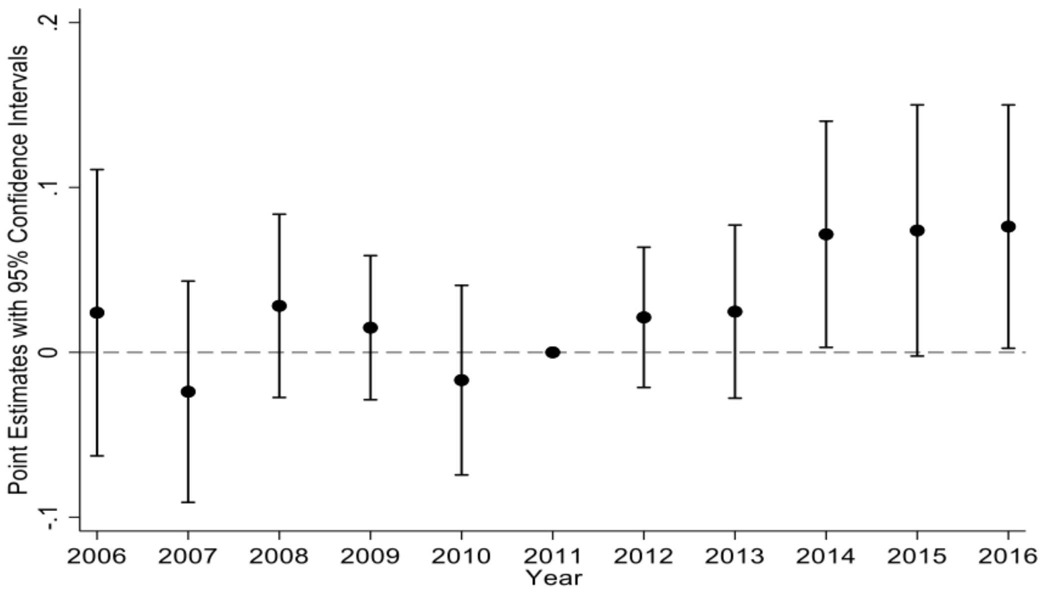

Figure 10: Event Study: The difference in refugee impact between high vs. low refugee exposure provinces—cutoff: 15th percentile.

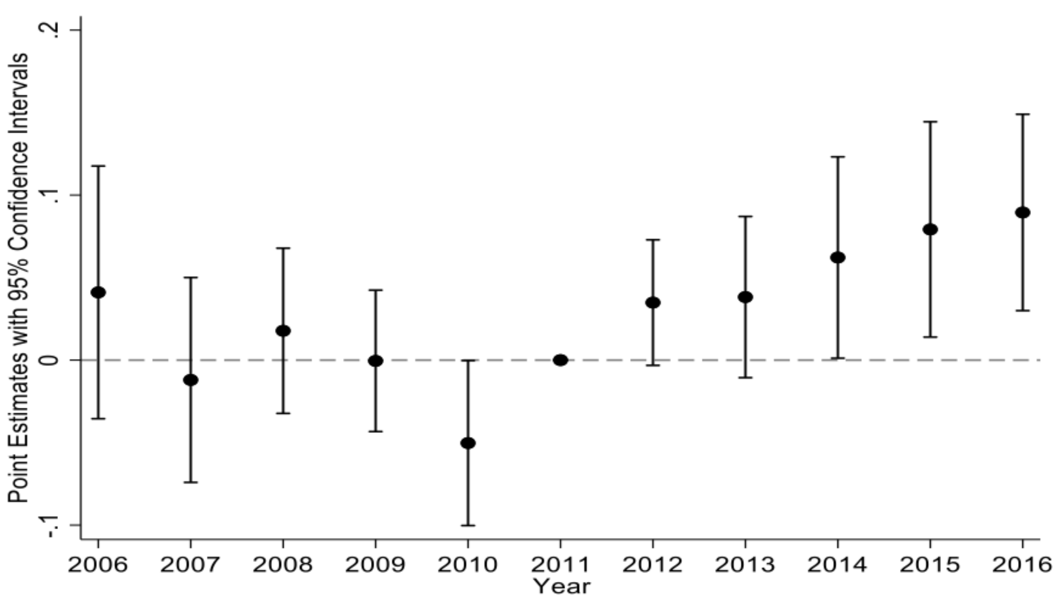

Figure 11: Event Study: The difference in refugee impact between high vs. low refugee exposure provinces—cutoff: 25th percentile.

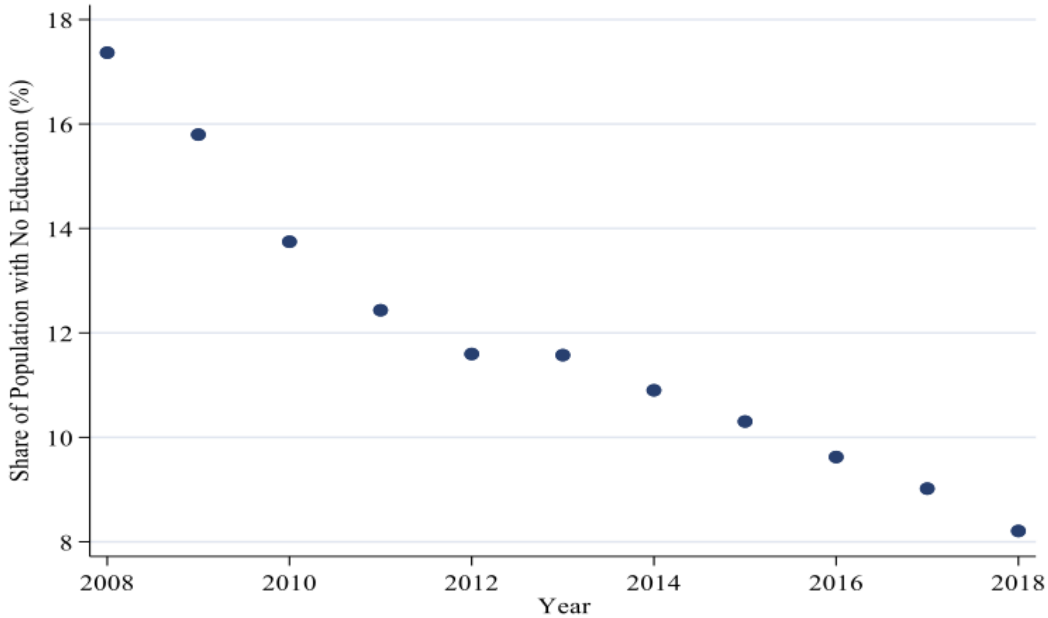

Figure 12: Share of Turkish population with no education: The graph plots the share of population in Turkey who are illiterate or without any diploma. Source: TurkStat.

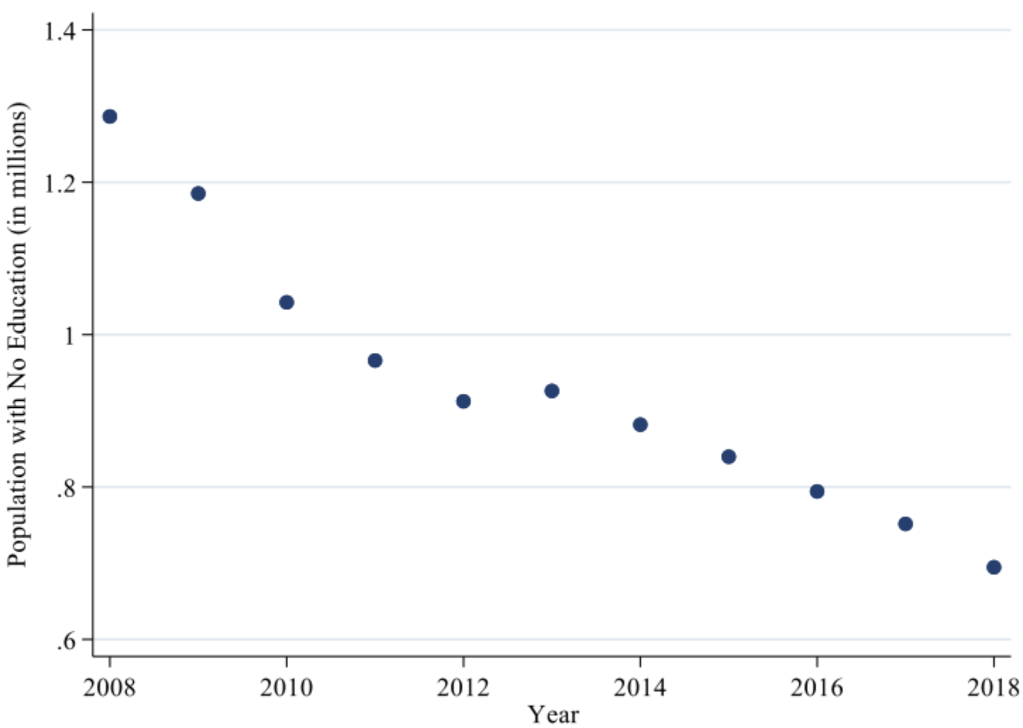

Figure 13: Turkish population with no education: The graph plots the size of population who are illiterate or without any diploma. Source: TurkStat.

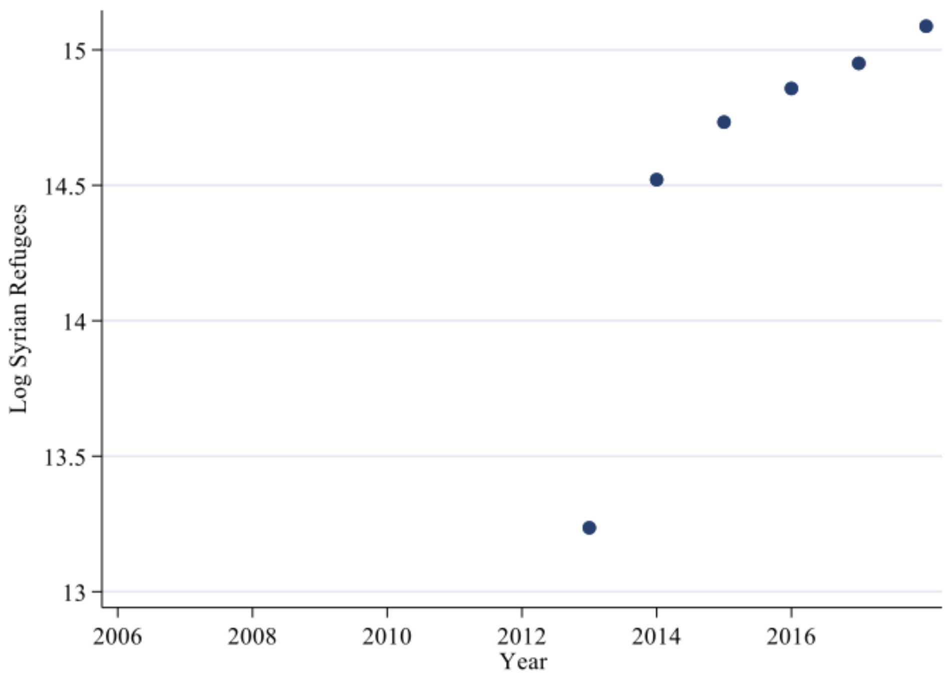

Figure 14: Logarithm of the number of Syrian refugees. Source: UNHCR.

The impact of refugees on crime—using alternative instruments
Dependent variable: Number of new crimes handled by prosecutors' offices

	(1)	(2)	(3)	(4)	(5)
	Panel A: Modified shift IV				Reduced form
ln(R)	0.059***	0.056***	0.064*	0.061***	
	(0.024)	(0.022)	(0.040)	(0.022)	
ln(N)	-0.027	-0.011	0.093	0.440**	
	(0.384)	(0.379)	(0.312)	(0.217)	
F-test for the first stage	47.86	46.04	39.06	49.00	
Instrument/1000					0.036***
					(0.014)
ln(N)					0.305
					(0.196)
	Panel B: Arabic IV				Reduced form
ln(R)	0.055***	0.050***	0.085**	0.070**	
	(0.009)	(0.007)	(0.036)	(0.028)	
ln(N)	-0.250	-0.108	-0.070	0.371	
	(0.299)	(0.250)	(0.337)	(0.234)	
F-test for the first stage	48.79	46.69	35.71	42.54	
Instrument/1000					0.003***
					(0.001)
ln(N)					0.328*
					(0.178)
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes	Yes
12-region-year FE	No	No	Yes	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes	No
Number of observations (panel A)	810	810	810	810	810
Number of observations (panel B)	737	737	737	737	737

Table 7: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

The impact of refugees on crime—using alternative instruments
Dependent variable: Number of charges in court cases

	(1)	(2)	(3)	(4)	(5)
	Panel A: Modified shift IV				Reduced form
ln(R)	0.053***	0.047**	0.090*	0.090**	
	(0.022)	(0.016)	(0.049)	(0.042)	
ln(N)	0.026	0.084	0.378	0.702	
	(0.475)	(0.407)	(0.530)	(0.502)	
F-test for the first stage	26.99	26.51	27.62	33.12	
Instrument/1000					0.035***
					(0.012)
ln(N)					0.595*
					(0.354)
	Panel B: Arabic IV				Reduced form
ln(R)	0.057**	0.044***	0.100	0.091	
	(0.027)	(0.016)	(0.067)	(0.055)	
ln(N)	-0.095	0.076	0.075	0.838	
	(0.602)	(0.441)	(0.799)	(0.540)	
F-test for the first stage	31.03	30.36	28.43	33.33	
Instrument/1000					0.003**
					(0.001)
ln(N)					0.700**
					(0.275)
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes	Yes
12-region-year FE	No	No	Yes	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes	No
Number of observations (panel A)	486	486	486	486	486
Number of observations (panel B)	469	469	469	469	469

Table 8: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

The impact of refugees on crime—using refugee-to-native ration

Dependent variable: Number of new cases handled by prosecutors' offices

	(1)	(2)	(3)	(4)	(5)
	Panel A: OLS				Reduced form
Refugees/Natives	0.675***	0.670***	0.469***	0.426***	
	(0.048)	(0.048)	(0.097)	(0.101)	
$\ln(N)$	0.544***	0.560***	0.311	0.443**	
	(0.202)	(0.212)	(0.188)	(0.185)	
Number of observations	891	891	891	891	
	Panel B: IV				
Refugees/Natives	0.893***	0.888***	0.551***	0.545**	0.017**
	(0.209)	(0.202)	(0.197)	(0.214)	(0.007)
$\ln(N)$	0.512**	0.531**	0.313*	0.442**	0.290
	(0.203)	(0.210)	(0.186)	(0.183)	(0.194)
Number of observations	891	891	891	891	891
	Panel C: Modified shift IV				
Refugees/Natives	0.934***	0.932***	0.613***	0.603**	0.036***
	(0.198)	(0.190)	(0.193)	(0.216)	(0.014)
$\ln(N)$	0.518**	0.547**	0.340*	0.467**	0.305
	(0.204)	(0.211)	(0.184)	(0.184)	(0.196)
Number of observations	810	810	810	810	810
	Panel D: Arabic IV				
Refugees/Natives	1.299***	1.256***	1.239***	1.191***	0.003***
	(0.301)	(0.293)	(0.299)	(0.291)	(0.001)
$\ln(N)$	0.242	0.363**	0.219	0.298*	0.328*
	(0.171)	(0.165)	(0.172)	(0.178)	(0.178)
Number of observations	737	737	737	737	737
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes	Yes
12-region-year FE	No	No	Yes	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes	No

Table 9: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

The impact of refugees on crime—using refugee-to-native ration
Dependent variable: Number of charges in court cases

	(1)	(2)	(3)	(4)	(5)
	Panel A: OLS				Reduced form
Refugees/Natives	0.505***	0.524***	0.537***	0.520***	
	(0.072)	(0.081)	(0.156)	(0.159)	
$\ln(N)$	0.646**	0.610**	0.579*	0.627*	
	(0.293)	(0.290)	(0.319)	(0.331)	
Number of observations	567	567	567	567	
	Panel B: IV				
Refugees/Natives	0.598***	0.636***	0.607***	0.616**	0.018***
	(0.164)	(0.173)	(0.222)	(0.239)	(0.007)
$\ln(N)$	0.623**	0.579**	0.574*	0.618*	0.569*
	(0.292)	(0.287)	(0.320)	(0.331)	(0.320)
Number of observations	567	567	567	567	567
	Panel C: Modified shift IV				
Refugees/Natives	0.582***	0.622***	0.596***	0.594***	0.035***
	(0.162)	(0.167)	(0.198)	(0.208)	(0.012)
$\ln(N)$	0.680**	0.653**	0.617*	0.685*	0.595*
	(0.317)	(0.313)	(0.353)	(0.360)	(0.354)
Number of observations	486	486	486	486	486
	Panel D: Arabic IV				
Refugees/Natives	0.965*	0.974*	1.339**	1.307**	0.003**
	(0.490)	(0.369)	(0.655)	(0.587)	(0.001)
$\ln(N)$	0.600**	0.595**	0.485	0.708**	0.700**
	(0.251)	(0.252)	(0.329)	(0.339)	(0.275)
Number of observations	469	469	469	469	469
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes	Yes
12-region-year FE	No	No	Yes	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes	No

Table 10: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

Heterogeneity in estimated effects of refugees on crime: IV results
Models estimated for different time periods

	(2006-2016)	(2006-2015)	(2006-2014)
	(1)	(2)	(3)
Panel A: Number of crimes handled by prosecutors' offices			
ln(R)	0.054**	0.063**	0.051**
	(0.021)	(0.024)	(0.020)
ln(N)	0.424**	0.338	0.283
	(0.201)	(0.210)	(0.188)
Number of observations	891	810	729
Panel B: Number of criminal charges in courts			
ln(R)	0.084*	0.077*	0.034
	(0.044)	(0.039)	(0.025)
ln(N)	0.642	0.610	0.518
	(0.472)	(0.503)	(0.393)
Number of observations	567	486	405
Year FE	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Time-varying province controls	Yes	Yes	Yes
5-region linear and quadratic trends	Yes	Yes	Yes
12-region-year FE	Yes	Yes	Yes
Trends by pre-Syrian province crime	Yes	Yes	Yes

Table 11: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

Heterogeneity in estimated effects of refugees on crime: IV results
Models estimated for different locations of the country

	Excluding					Only	
	West	Central	South	North	East	South	East
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Number of crimes handled by prosecutors' offices							
ln(R)	0.052**	0.052**	0.061**	0.055**	0.051**	0.050**	0.055**
	(0.020)	(0.022)	(0.026)	(0.021)	(0.025)	(0.025)	(0.026)
ln(N)	0.399*	0.546**	0.369*	0.489**	0.279	0.510	0.630*
	(0.232)	(0.225)	(0.200)	(0.226)	(0.244)	(1.510)	(0.359)
Number of observations	649	770	803	715	627	88	264
Panel B: Number of criminal charges in courts							
ln(R)	0.067*	0.087*	0.122*	0.087*	0.060	0.054*	0.098
	(0.034)	(0.044)	(0.073)	(0.048)	(0.039)	(0.027)	(0.068)
ln(N)	1.016**	0.605	0.409	0.652	0.539	1.167	0.999
	(0.457)	(0.529)	(0.480)	(0.625)	(0.490)	(2.329)	(0.897)
Number of observations	413	490	511	455	399	56	168
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	No	No	No	No	No
12-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trends by pre-Syrian province crime	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 12: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

Sensitivity of estimated effects of refugees on crime: IV results
Models estimated by excluding various provinces

		Excluding		
	Baseline	Istanbul	3 major provinces	Kilis and Sirnak
	(1)	(2)	(3)	(4)
Panel A: Number of crimes handled by prosecutors' offices				
$\ln(R)$	0.054**	0.054**	0.053**	0.042*
	(0.021)	(0.021)	(0.021)	(0.024)
$\ln(N)$	0.424**	0.424**	0.440**	0.468**
	(0.201)	(0.200)	(0.200)	(0.197)
Number of observations	891	880	858	869
Panel B: Number of criminal charges in courts				
$\ln(R)$	0.084*	0.084*	0.083*	0.059
	(0.044)	(0.043)	(0.042)	(0.049)
$\ln(N)$	0.642	0.642	0.618	0.635
	(0.472)	(0.469)	(0.473)	(0.434)
Number of observations	567	560	546	553
Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Time-varying province controls	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	Yes	Yes	Yes	Yes
12-region-year FE	Yes	Yes	Yes	Yes
Trends by pre-Syrian province crime	Yes	Yes	Yes	Yes

Table 13: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

Sensitivity of estimated effects of refugees on crime: IV results
Models estimated by controlling initial province education and wealth

	Crimes in the prosecutors' office	Criminal charges in courts
	(1)	(2)
Panel A: By initial education		
$\ln(R)$ × Least educated	0.054**	0.089**
	(0.022)	(0.044)
$\ln(R)$ × Mid educated	0.052**	0.084*
	(0.021)	(0.044)
$\ln(R)$ × Most educated	0.049**	0.092**
	(0.021)	(0.046)
Panel B: By initial wealth		
$\ln(R)$ × Least wealth	0.048**	0.082*
	(0.021)	(0.044)
$\ln(R)$ × Mid wealth	0.054***	0.086**
	(0.020)	(0.042)
$\ln(R)$ × Most wealth	0.046**	0.084*
	(0.021)	(0.043)
Year FE	Yes	Yes
Province FE	Yes	Yes
Time-varying province controls	Yes	Yes
5-region linear and quadratic trends	Yes	Yes
12-region-year FE	Yes	Yes
Trends by pre-Syrian province crime	Yes	Yes

Table 14: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

B Appendix: Comparison to related papers

In this appendix, we synthesize two recent papers which addressed the same research question we ask in our paper, but reached the conclusion that an increase in refugees lowered crime in Turkey. Specifically, Kirdar et al. (2022) and Kayaoglu (2022) both employ the same refugee data we use in this paper, and they implement an instrumental variables analysis using the same or very similar instruments to ours. They, however, approximate the criminal activity by different measures, as opposed to using the number of offenses reported to the prosecutors' offices as was done in our paper. Kayaoglu (2022) uses a noisy indicator of crime which leads to underestimating criminal activity and misrepresentation of the true crime trend in the country. Kirdar et al. (2022) use an erroneous approximation of crime which is incorrect conceptually, and which woefully underreports its true prevalence. More importantly, both papers estimate a particular empirical specification which produces a downward-biased estimate of the impact of refugees on crime. Put differently, the econometric model used by these papers mechanically produces a negative relationship between refugees and crime, when in fact the true relationship is positive. In this appendix, we describe these pitfalls in some detail although we discussed them throughout the paper.

Problems with crime indicators

Kirdar et al., (2022). Kirdar et al. (2022) use the number of individuals who enter the prison as the measure of criminal activity. The first author of Kirdar et al. (2022) graciously shared their data with us. Thus, we are able to confirm that the authors, in fact, used as their crime indicator *the number of convicted felons who entered the prison system* as reported by the Turkish Statistical Institute (Prison Statistics, Table 2.7: *Convicts Received into Prison*, for all years until 2013. The data for post-2013 can be downloaded from this link).

As the vast literature in economics of crime reveals, however, prison intake is not a valid proxy for criminal activity for a number of important reasons. First, incarceration has an impact on crime itself. This is both because of the incapacitation effect of incarceration (incarcerated individuals being unable to commit crime while in prison), and because incarceration is a deterrent to crime. This means that prison population is a determinant of crime, rather than being an indicator of crime itself (Barbarino and Mastrobuoni, 2014; Johnson and Raphael, 2012; Drago et al., 2009; Corman and Mocan, 2005; Levitt, 1996). Prison population is, of course, influenced by the extent of criminal activity. However, that incarceration is impacted by crime does not imply that the former can be used as a proxy for the latter. This is described in Figure 3 of our paper. In 2013, there were about 3.4 million recorded criminal acts in Turkey, but that same year only 161,711 individuals entered prisons. This is because, as shown in Figure 3, not all crimes end up in courts to be adjudicated (because of unknown suspects, lack of evidence, and so on). Furthermore, if a case goes to trial, not all defendants are convicted; and only some of the convicted criminals are imprisoned—due to other resolutions implemented by the courts such as suspended sentences, fines, probations, etc.

Using the number of convicted felons who enter prison as a proxy for the incidence of crime lead the authors to report the crime rate of Turkey as **196** offenses per 100,000 people [Kirdar et al. (2022), Table 1]. The true crime rate of the country, based on crimes handled by the offices of the prosecutors is 4,500 per 100,000 population. It should be noted that with a few exceptions with questionable crime reporting, there is virtually no country with a crime rate in the range of a few hundred per 100,000 people. The crime rate in the EU was 7,000 in 2010 (Buonanno et al., 2018). The current crime rate in EU countries ranges from 1,500 in Bulgaria, to 3,500 in Portugal, to 4,500 (Italy) to 7,500 in Germany (European Sourcebook of Crime and Criminal Justice Statistics, 2021, 6th edition). The crime rate in the U.S. was 3,500 in 2010, and 2,500 in 2019 (FBI, Uniform Crime Reports).

Relatedly, Kirdar et al. (2022) claim that the crime rate for most felony crimes are in the single digits. For example, the authors argue that the homicide rate is **8** per 100,000 people, and the robbery rate is **6.6**. They also argue that with the exception of assault, there are provinces with **zero crime** for all other crime categories [Kirdar et al. (2022), Table 1]. This peculiar picture emerges because there are in fact very few individuals who enter prison for narrow crime categories in a given province.

Another issue with the attempt to use prison intake in a particular year as a proxy for crime for that year lies in the fact that the timing of prison entry does not match the timing of the commission of the crime. Judicial process is slow, which translates into a mismatch between the year in which a crime is committed and when the offender enters prison. In Turkey, the average time for the office of the prosecutors to process files with known suspects was 91 days in 2013, and it steadily rose over time, reaching 131 days in 2016 (Turkish Justice Statistics Yearbook, 2016). The average duration of cases at the courts was 231 days for cases adjudicated in 2013, rising to 274 days in 2016 (Turkish Justice Statistics Yearbook, 2016). This means that the time span between a crime reported to the office of the prosecutors and its final court resolution was 322 days in 2013, and 405 days in 2016. This, in turn, implies that perpetrators who got arrested for their offense in March or later in a particular year are expected to enter prison (if convicted) during the following year. A perpetrator who committed a crime in December 2014 is expected to hear the decision of the judge in January 2016.

To make matters worse, the manner in which the prison intake is reported in the Correction Statistics of the Turkish Statistical Institute *does not* refer to defendants who are convicted and received a prison sentence in that year. Rather, prison intake refers to the resolution of cases after they have completed the appeal process. More specifically, although the defendant stays in prison while his case is evaluated by the relevant Appellate Court, this individual is not counted as a "prison entry" until after the final decision of the Appellate court.[41] The average duration of the appellate process in criminal cases is over 1,000 days (Akdeniz, 2019). Thus, the interval between the commission of a crime and the time the perpetrator shows up in prison statistics as "prison intake" could be four years or longer. Clearly, the timing of prison intake does not match the timing of the criminal activity.

There are other concerns as well. The delays in the judicial process are not exogenous, nor are they time-invariant. Rather, judicial delays depend on the caseload of the system, which in turn is a function of the extent of criminal activity. Put differently, a rise in crime increases the caseload of the criminal justice system, which leads to further delays in processing the defendants, and widens the time span between arrest and prison entry.[42]

A final complication pertains to prison capacity. It is well-established that prison infrastructure and physical capital cannot be expanded quickly (Boylan and Mocan, 2014; Levitt, 1996). This implies that judicial decisions are expected to be impacted by prison overcrowding, and judicial leniency goes up when prisons are operating at or near full capacity. This, in turn, has a positive impact on crime as it signifies a reduction in deterrence. In summary, prison intake should not be used as a proxy for the extent of criminal activity.

Kayaoglu (2022). Kayaoglu (2022) uses the number of cases in criminal courts as her crime indicator. As we discussed in Section 4, the number of cases in courts is not a good proxy for the incidence of crime for a number of reasons. First, the number of cases is, by definition, smaller than the number of offenses because some defendants are charged with multiple offenses. Second, some suspects and arrestees are not pursued further by prosecutors because of insufficient evidence. In these situations, the case files are not forwarded to the courts although there were suspects in these cases. Finally, there are offenses with unknown suspects. This means that, although a crime has been committed and that there is a record in the files of the police and the prosecutors' offices, no perpetrator has been identified. The upshot is that the number of court cases underestimates the true incidence of crime.

Box II of Figure 3 of our paper reveals that there were about 3.4 million new crimes handled by the prosecutors' offices in 2013. In contrast, Box III shows that there were about 1.3 million new cases that came into the dockets of criminal courts in that year. Kayaoglu (2022) uses the sum of cases in basic criminal courts and criminal courts of peace as her main outcome. In 2013, there are about 1.17 million new cases in these courts—The upper line in Figure 2 of Kayaoglu

[41] This is explained on page XV of the of the Prison Statistics 2013, Turkish Statistical Institute.

[42] Each year Turkish criminal courts roll over about 1.7 million cases to the following year (Turkish Justice Statistics Yearbooks, various years).

(2022) presents the behavior of these court cases per 100,000 people.[43]

The middle (dashed) line in Figure 15 below replicates the crime rate in Figure 2 of Kayaoglu (2022), which is the sum of the number of cases in basic criminal courts and criminal courts of peace. Figure 16 presents the same information using the actual number of cases, rather than rates per 100,000 people. According to these figures, there was a steady decline in the number of court cases (Figure 16) between 2012 and 2017, the period during which Syrian refugee influx took place.[44] The solid lines in both figures display the cases in prosecutors' offices with unknown perpetrators. These cases are not forwarded to the courts because there were no defendants identified. There has been a steady increase in the number of cases with unknown perpetrators over time. These are criminal acts reported by the police to the prosecutors' offices, but ultimately no suspects were identified. These unresolved cases may reflect the resource constraint faced by law enforcement agencies during a period of rising crime.

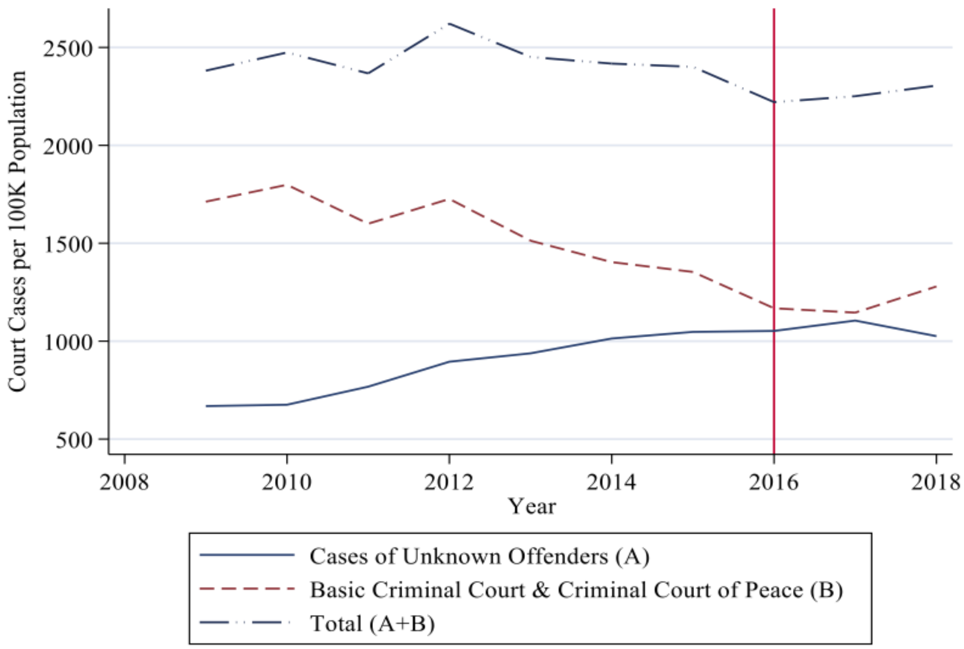

Figure 15: The number of cases in some courts per 100k population, and cases with unknown perpetrators per 100k population.

The true number of criminal cases can be portrayed by the sum of the cases in courts (the middle lines in Figures 15 and 16; as used by Kayaoglu (2022)) and the number of cases with unknown offenders (the bottom, solid line in Figures 15 and 16). This total is shown by the top line (the

[43]In Figure 2 of Kayaoglu (2022), this variable is titled "Basic Criminal Court Cases," although it consists of the sum of cases in basic criminal courts and cases in criminal courts of peace. The bottom line of the same figure represents the number of cases for felonies with associated sentences of 10 years and longer. The author calls these High Criminal Court Cases. There were 69,732 of such cases in 2013.

[44]The decline in the number of court cases in Figure 16 translates into a decline in the crime rate based on these court cases in Figure 15.

broken line) in both figures, which still does not represent the actual criminal activity, but it brings the calculated crime rate and its evolution over time closer to reality.

Nevertheless, we used the same crime rate proxy employed by Kayaoglu (2022), and estimated our model depicted by Equation 6. That is, we used the sum of the number of cases in basic criminal courts and cases in criminal courts of peace (i.e., the middle line in Figure 16), and estimated our instrumental variables specification. The results, reported in Table 15 below reveal that an increase in refugee population has a positive and significant impact on crime even when this particular crime proxy is employed as the outcome. This result confirms the theoretical discussion in Section 6 of the paper, which reveals that the specific empirical model used by both Kayaoglu (2022) and Kirdar et al. (2022) imposed a mechanical negative relationship between refugees and crime. We repeat the explanation of this pitfall below.

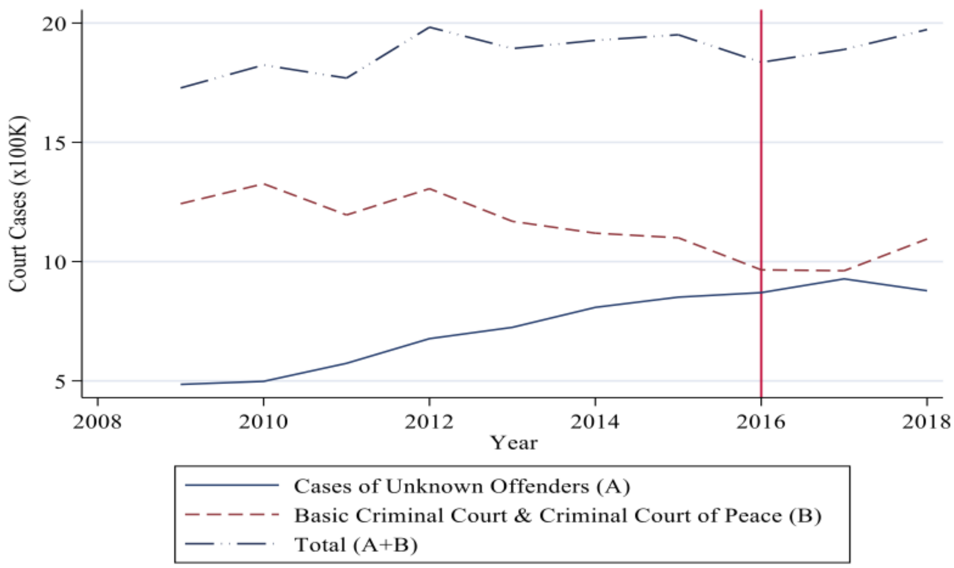

Figure 16: The number of cases in some courts and cases with unknown perpetrators.

Incorrect empirical model that produces a negative bias on the estimated refugee effect. This material is presented in Section 6 of the paper. We summarize it again here for completeness. The production function for crime can be specified as

$$CR = AR^{\beta}N^{\gamma}e^{\varepsilon}, \tag{11}$$

where CR stands for the number of offenses, R represents the number of refugees, and N is the size of the native population. The empirical counterpart of Equation 11 is:

$$\ln(CR) = \alpha + \beta\ln(R) + \gamma\ln(N) + \varepsilon, \tag{12}$$

The impact of refugees on crime (2006-2016)

Dependent variable: Number of cases in criminal courts and criminal courts of peace

	(1)	(2)	(3)	(4)	(5)
	Panel A: OLS				
$\ln(R)$	0.005	0.005	0.002	0.009	
	(0.005)	(0.005)	(0.005)	(0.006)	
$\ln(N)$	0.471*	0.419	0.131	0.215	
	(0.238)	(0.253)	(0.234)	(0.257)	
	Panel B: IV				
$\ln(R)$	0.027	0.027	0.075*	0.074**	
	(0.019)	(0.017)	(0.044)	(0.032)	
$\ln(N)$	0.220	0.163	-0.099	0.188	
	(0.301)	(0.308)	(0.296)	(0.286)	
					Reduced form
Instrument/1000					0.023***
					(0.009)
$\ln(N)$					0.125
					(0.223)
F-test for the first stage	43.12	50.20	16.31	16.15	
Mean of the dependent variable	13,569	13,569	13,569	13,569	13,569
Mean of refugees	8,892	8,892	8,892	8,892	8,892
Mean of native population	922,180	922,180	922,180	922,180	922,180
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Time-varying province controls	No	Yes	Yes	Yes	Yes
5-region linear and quadratic trends	No	No	Yes	Yes	Yes
12-region-year FE	No	No	Yes	Yes	Yes
Trends by pre-Syrian province crime	No	No	No	Yes	No
Number of observations	891	891	891	891	891

Table 15: Standard errors, clustered at province level, are in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). Time-varying province controls include the log of hospitals per 100k population, the presence of natural gas lines in the province in a given year, the log of public expenditures per 100k population, and the lag of prison intake.

where α contains observable exogenous characteristics of the province, as well as various fixed effects that soak up unobserved province and regional attributes. Province and time subscripts are suppressed.

Equation 12 can be converted into different forms. For example, adding $[-\beta \ln(N+R) - \ln(N+R)]$ to both sides and rearranging terms yields

$$\ln\left(\frac{CR}{N+R}\right) = \alpha + \beta \ln\left(\frac{R}{N+R}\right) + \gamma \ln(N) + [(\beta - 1)\ln(N+R)] + \varepsilon, \tag{13}$$

which can also be written as

$$\ln\left(\frac{CR}{N+R}\right) = \alpha + \beta \ln\left(\frac{R}{N+R}\right) + \gamma \ln(N) + \nu, \tag{14}$$

where $\nu = [(\beta - 1)\ln(N+R)] + \varepsilon$.

The left-hand-side of Equation 14 is (log of) the crime rate, and the key variable on the right-hand side is the (log of) share of refugees in total population as used by both Kayaoglu (2022) and Kirdar et al. (2022) to estimate the impact of refugees on crime.

Although Equation 14 is a rearrangement of Equation 12, it is not appropriate to use Equation 14 in an effort to estimate the impact of refugees on crime. This is because of the following reasons:

(i) Suppose that there is no true relationship between the refugee share ($R/(N+R)$) and the crime rate; that is, assume that $\beta = 0$ in Equation 14. The error term of Equation 14 reveals, however, that an increase in the number of refugees (R) will nevertheless produce a negative relationship between refugees and the crime rate. This mechanical negative relationship, imposed by the transformation of Equation 12 to Equation 14 persists as long as the elasticity of the crime rate with respect to refugee share (β) is less than one. Put differently, fitting Equation 14 to data underestimates β.

(ii) Ignoring the issue highlighted in point (i), another problem in using Equation 14, as was done by Kayaoglu (2022) and Kirdar et al. (2022), is that the variable of interest, R, is both in the numerator of the key explanatory variable, and in the denominator of the dependent variable. This property of Equation 14 also imposes a mechanical negative relationship between refugees and the crime rate by construction.

(iii) Related to points (i) and (ii) above, any instrument that is correlated with R is invalid in Equation 14 because the exclusion restriction is violated and the estimated β is biased. More specifically, consider Equation 14 again. The probability limit of the instrumental variables estimate of β is: $\text{plim}\,\widehat{\beta} = \beta + \frac{Cov(\ln(Z),\nu)}{Cov(\ln(R/(N+R)),\ln(Z))}$, where Z is the instrument. Any instrument Z, which would generate a movement in R, is also correlated with the error term (ν) of Equation 14 as the error term contains R. More specifically, if $Cov(Z, R) > 0$, this would imply that $Cov(Z, \nu) < 0$ if $\beta < 1$, and $Cov(Z, \nu) > 0$ if $\beta > 1$. The instrument is uncorrelated with the error term only if $\beta = 1$, but even in this special case the instrument is invalid, because the endogenous variable (R) also appears in the denominator of the dependent variable.

(iv) Finally, even if none of these vital issues existed, a basic problem would have been the use of the same divisor both in the dependent and the independent variable. More specifically, using the crime rate as the dependent variable, and then using population (which is the denominator of the crime rate) as the deflator of the key explanatory variable creates bias.[45]

In summary, estimating the specification shown in Equation 14 produces a downward bias which explains the surprising result reported by Kayaoglu (2022) and Kirdar et al. (2022) that refugee inflows have a crime-reducing effect. To demonstrate this point empirically, we used the same crime indicators employed by these papers (prison intake as used by Kirdar et al. (2022); and the sum of the number of cases in basic criminal courts and cases in criminal courts of peace as used by Kayaoglu (2022)). We first replicated their results, and then demonstrated how they changed under correct model specification. This exercise is summarized in Table 16 below.

Panel A of Table 16 reports the instrumental variables results using the same crime proxy (prison intake) and the same instrument of Kirdar et al. (2022). These authors used data spanning 2008-2019, but they dropped the year 2012 from the analysis sample. We employed the same sample as they did, used the same incorrect empirical specification (Equation 14), and replicated their results as was reported in the first row of their Table 3 [Kirdar et al. (2022), p. 576]. This replication, displayed in the top section of Panel A, reveals a negative impact of the refugee ratio on the crime rate, as reported by the authors. The bottom section of Panel A, uses the same data and the same instrument, but employs the correct empirical model (Equation 12) as discussed above. Doing so reverses the results and reveals that an increase in refugees leads to more crime.

Panel B of Table 16 repeats the same exercise for Kayaoglu (2022). The top section of Panel B replicates the author's results using the same crime proxy, the same instrument, and the same incorrect model specification as employed in that paper.[46] Here, as in Table 5 of Kayaoglu

[45] More specifically, consider the model $[CR/(N + R)] = \alpha + \beta(R/N) + \varepsilon$, where CR stands for crime, R is the refugee population, N is the native population, and (R/N) represents the refugee share. Because $(N + R) \approx N$, this regression would produce a spurious relationship between the crime rate and refugee share because both the dependent variable and the explanatory variable have (almost) the same denominator. As explained by Kronmal (1993) and Bazzi and Clemens (2013), and as highlighted with examples by Clemens and Hunt (2019), the denominators that are the same or very similar will generate spurious correlation between the two variables when the true β is zero. See Kronmal (1993) for theoretical and empirical examples, and proposed solutions.

[46] The author writes in the paper that she used the period 2009-2016, which would have produced 648 observations. However, Table 5 of Kayaoglu (2022) reports 567 observations, which implies a span of 7 years. Thus, we estimated the model between 2010-2016, which produced the results reported in the top section of Panel B. Using the time period of 2009-2016 or 2012-2018 provided the same results.

(2022), there is a negative impact of the refugee share on the crime rate. The bottom section of Panel B reports the results based on the same data and the same instrument, but here we use the correct model (Equation 12), which does not impose a negative association between refugees and crime. As was the case with the Kirdar et al. (2022) in Panel A, doing so flips the sign of refugees from negative to positive.

In summary, as displayed in Table 16, the impact of refugees on crime is positive, and the results reported by these authors are an artifact of the incorrect specification employed by them.

The impact of refugees on crime

The influence of model specification when using poor/wrong proxies of crime: IV specifications

	Panel A: Kirdar et al. (2022) Crime proxy: Convicts entering prison	
Incorrect model: Equation 14	(I)	(II)
Dependent variable: $\ln(CR/(N+R))$	-157.282*	-140.377
Explanatory variable: $\ln(R/(N+R))$	(89.023)	(138.970)
Correct model: Equation 12	(I)	(II)
Dependent variable: $\ln(CR)$	0.117**	0.194*
Explanatory variable: $\ln(R)$	(0.046)	(0.109)

	Panel B: Kayaoglu (2022) Crime proxy: Cases in criminal courts and criminal courts of peace	
Incorrect model: Equation 14	(I)	(II)
Dependent variable: $\ln(CR/(N+R))$	-0.012***	-0.006***
Explanatory variable: $\ln(R/(N+R))$	(0.004)	(0.002)
Correct model: Equation 12	(I)	(II)
Dependent variable: $\ln(CR)$	0.011	0.043*
Explanatory variable: $\ln(R)$	(0.019)	(0.023)

Table 16: Following Kirdar et al. (2022), Panel A models are based on 891 observations spanning 2008-2019, omitting 2012. Column (I) in Panel A corresponds to the first column in Table 3 (page 576) of Kirdar et al. (2022), and controls for province and year fixed effects and province specific controls. Column (II) corresponds to the fifth specification in Table 3 of Kirdar et al. (2022), and controls for province and year fixed effects, 5-Region time trends, NUTS1 time trends, 5-Region-year fixed effects, NUTS1-year fixed effects, and province specific controls. The correct model in Panel A also controls for population.

Following Kayaoglu (2022), the results reported in Panel B are based on 567 observations, spanning the years 2010-2016. The specification in Column (I) in Panel B corresponds to Column 4 in Table 5 (page 15) of Kayaoglu (2022), and Column (II) corresponds to Column 5 in Table 5 (page 15) of Kayaoglu (2022). The correct model in Panel B also controls for population.

Standard errors are clustered at the province level in all regressions. *, **, and *** indicate statistical significance at the 10%, 5% and 1%, respectively.

C Appendix: Connections to the labor market

In the paper, we demonstrated that an increase in the refugee population generates an increase in criminal activity. The same is true for an increase in the unskilled native population, although the impact of the refugee population is bigger. With this result in mind, consider that an increase in the refugee population exerts two effects on total crime: (i) non-labor market effect, which signifies the rise in total number of crimes simply because of additional individuals in the society, and (ii) the labor market effect, which impacts crime through the influence of refugees on wages.

Specifically, consider

$$\frac{\partial CR_{\text{Non-labor market}}}{\partial R} + \left(\frac{\partial CR_R}{\partial w_R}\right)\left(\frac{\partial w_R}{\partial R}\right) + \left(\frac{\partial CR_N}{\partial w_R}\right)\left(\frac{\partial w_R}{\partial R}\right) = C^* + \left(\frac{\partial w_R}{\partial R}\right)\left[\left(\frac{\partial CR_R}{\partial w_R}\right) + \left(\frac{\partial CR_N}{\partial w_R}\right)\right]. \tag{15}$$

The first term of Equation 15 represents the increase in crime due to the non-labor market effect of an increase in the refugee population (R). As described in the introduction, this reflects an increase in crime simply because of the increase in the number of people who have attributes (e.g., risk aversion, time preference, exposure to violence, and so on) which would influence their criminal proclivity one way or the other. The second term captures the change in crimes committed by refugees (CR_R), induced by the change in wages triggered by a rise in the refugee population. This second term summarizes the labor market effect on refugees' crime of a change in refugee wages. The third term depicts how a change in refugee wages, due to an increase in the number of refugees, impacts crime committed by natives (CR_N). Collecting the terms produces the expression on the right-hand-side of the equality sign, where C^* stands for $\partial CR_{\text{Non-labor market}}/\partial R$.

Equation 16 displays the same idea for an increase in the native population (N). The term C^{**} on the right-hand-side of Equation 16 represents the change in crime because of an increase in native population, without altering the labor market conditions.

$$\frac{\partial CR_{\text{Non-labor market}}}{\partial N} + \left(\frac{\partial CR_N}{\partial w_N}\right)\left(\frac{\partial w_N}{\partial N}\right) + \left(\frac{\partial CR_R}{\partial w_N}\right)\left(\frac{\partial w_N}{\partial N}\right) = C^{**} + \left(\frac{\partial w_N}{\partial N}\right)\left[\left(\frac{\partial CR_N}{\partial w_N}\right) + \left(\frac{\partial CR_R}{\partial w_N}\right)\right]. \tag{16}$$

Our empirical analyses show that an increase in the number of refugees has a larger impact on total crime than an increase in the native population. This implies that the right-hand side of Equation 15 is greater than the right-hand side of Equation 16.

If native unskilled labor and refugee labor are perfect substitutes, this would imply the existence of one prevailing wage in the market for both groups ($w_R = w_N$). Under this scenario, the last terms in brackets on the right-hand-side of Equations 15 and 16 would be the same, and it would also be the case that $(\partial w/\partial R) = (\partial w/\partial N))$. Thus, it would follow that $C^* > C^{**}$, which would in turn imply that, even if there were no labor market effect on crime, an increase in refugee population generates a larger increase in crime in comparison to an equivalent increase in native population.

Alternatively, suppose that $C^* = C^{**}$. That is, assume that absent any labor market effect, an injection of refugees or natives in a community by a given magnitude would impact crime equally. Further assume that unskilled native workers and refugees are not perfect substitutes. In this case, our finding that the magnitude produced by Equation 15 being greater than the magnitude generated by Equation 16 implies that

$$\left[\left(\frac{\partial CR_N}{\partial w_N}\right) + \left(\frac{\partial CR_R}{\partial w_N}\right)\right] < \left[\left(\frac{\partial CR_R}{\partial w_R}\right) + \left(\frac{\partial CR_N}{\partial w_R}\right)\right].$$

This inequality depends on elements such as the elasticity of labor demand for refugee labor and for native labor, the responsiveness of refugee crime to refugee wages, and the responsiveness of native crime to native wages. It also depends on the responsiveness of refugee crime (native crime) to native wages (refugee wages) through the elasticity of substitution between refugee labor and native labor.

The upshot is that the results identified in the paper can emerge theoretically under a number of different scenarios involving the structure of the labor markets (which also reflect the production technology).