Empirical Economics (2021) 61:2827–2864
https://doi.org/10.1007/s00181-020-01963-8

Were they a shock or an opportunity?: The heterogeneous impacts of the 9/11 attacks on refugees as job seekers—a nonlinear multi-level approach

Seonho Shin[1,2]

Received: 27 March 2019 / Accepted: 15 October 2020 / Published online: 21 January 2021

Abstract

This study investigates whether the September 11 terrorist attacks had any impacts on the labor market outcomes of refugees resettled in the United States, who should be distinguished from economic migrants or usual nonnatives. Furthermore, this paper sheds unprecedented light on whether those impacts were heterogeneous depending on a refugee's ethnicity or religion. In terms of econometric methods, this research attempts to allow for the violation of the conventional condition of independently and identically distributed (*i.i.d.*) observations and control for cluster-specific unobservables by using nonlinear multi-level models, considering that refugees form unique networks in their resettlement regions and actively interact with one another within their clusters. Due to the binary dependent variable of this study, the incidental parameters problem is also taken into account. The multi-level estimates of this paper suggest that the September 11 attacks did not uniformly shock all sub-populations of refugees: rather, they presented a unique, substantial opportunity for Asian refugees and a serious threat to African and Arab refugees. One unanticipated finding is that the employment probability of European refugees remained stable, whereas that of Asian refugees markedly increased after the attacks. However, in terms of employment quality, measured by real wages, European refugees were the only ones who benefited from the attacks. Possible explanations for such heterogeneous impacts and different patterns of benefits are discussed, including positive versus negative selection into employment.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00181-020-01963-8.

✉ Seonho Shin
sshin@kli.re.kr

1 Chair of Econometrics, Chair of Macroeconomics and Development, University of Frankfurt, RuW 3.261, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, Germany

2 Korea Labor Institute, Department of Employment Policy Research, Sejong National Research Complex, Economics Building C, 30147 Sejong, South Korea

Keywords Refugee labor market · Clustered observations · Nonlinear multi-level models · Chamberlain–Mundlak's correlated random effects probit model · Conditional logit fixed effects model

JEL Classification J61 · J64 · J71

1 Introduction

In recent years, the number of people forcibly displaced has rapidly increased, ushering in a new age of diaspora. According to the United Nations refugee agency, as of 2017, there are 68.5 million forcibly displaced people worldwide, and 25.4 million of them are officially registered as refugees by the United Nations High Commissioner for Refugees (UNHCR) and the United Nations Relief and Works Agency (UNRWA). Especially, Europe is at the center of giant waves. In 2015 alone, more than one million migrants and refugees crossed into Europe, precipitating the European refugee crisis. While some opponents intransigently raise their voices to close the borders and worry about the myriad social and economic problems expected to occur with their exodus to Europe, others firmly stand against such pessimistic voices and accentuate the positive aspects, the main point of which is the expectation that the influx of refugees will invigorate the rapidly aging continent with an active, younger working population.

Whether the set of advantages will far outweigh that of disadvantages and fulfill the sanguine expectations of pro-refugee proponents depends on how the nations that are in charge help those refugees settle down in the initial stages and enable them to become economically self-sufficient as soon as possible without burdening their host populations that much. The specified goals of the United States Refugee Admission Program (USRAP) do well in reflecting the importance of the successful labor market integration of refugees into their host communities. Furthermore, some scholars argue that the successful labor market integration extends well beyond economic issues because bad labor market outcomes for refugees may contribute to their political radicalization (Dagnelie et al. 2019). Therefore, refugees' smooth labor market integration is of critical importance. For this reason, it is crucial to analyze and understand the uniqueness of the labor market conditions that refugees encounter, and in doing so, contextual heterogeneity should also be taken into careful consideration.

This paper opens with the indisputable fact that refugees are in a highly unfavorable position relative to their host populations, which should be construed as the key underlying feature of refugees as job seekers. Hence, it is obvious that they are much more vulnerable to unexpected external shocks affecting labor market conditions. Additionally, what causes further complications stems from the fact that the magnitude of such vulnerability can be heterogeneous depending on certain factors, especially the cause and agent of a shock. Such heterogeneity is the starting point of the investigation of this paper.

This study questions whether the series of four coordinated terrorist attacks planned and perpetrated by Al-Qaeda against the U.S. on September 11, 2001 triggered any significant changes in the labor market outcomes of refugees resettled in the U.S..[1] While many studies have investigated the impacts of the September 11 attacks on labor markets both in and outside of the U.S., possible consequences on the labor market outcomes of *refugees*, who are distinguished from economic migrants or usual nonnatives, have been far too neglected, with not a single individual study investigating them, to the best of the author's knowledge. Hence, this paper provides the first empirical analysis of the impacts of the attacks on refugees as job seekers. More importantly, this research also sheds new light onto whether the impacts stemming from the attacks were heterogeneous depending on a refugee's ethnicity or religion, considering that the attacks were planned by the Islamic extremist group and perpetrated by its African and Arab hijackers. In light of the fact that refugees usually compete with one another as job seekers in their resettlement regions, if African and Arab refugees, due to the cause and agent of the September 11th attacks, were more affected by the negative impacts stemming from the attacks, it is a sensible hypothesis that the less affected refugees may have rather *benefited* from the attacks, which is to be addressed in this study.

In terms of econometric methods, this paper criticizes common, yet naïve, *single-level* approaches, which ignore the irrefutable fact that refugees form unique networks in their resettlement regions, influence one another, and interact with enclaves composed of their previously resettled compatriots. The important effect of such clustering on nonnatives' labor market outcomes is already corroborated by Edin et al. (2003) and Damm (2009b). Especially, their networks are important in that they provide information on labor market conditions and opportunities for recent refugees (Munshi 2003; Beaman 2012). Hence, it is a natural inference that those networks make refugees not independent but heavily *clustered*, and the use of a single-level approach, which naïvely considers all observations to be independent, is improper in this context.

Such clustering of refugees makes the conventional assumption of independently and identically distributed (hereafter, *i.i.d.*) observations implausible, which is essential for the consistency of many econometric methods, such as maximum likelihood estimation. Especially, in analyzing binary responses, naïve negligence on the violation of *i.i.d.* often leads to severely biased estimates. In contrast to many previous single-level studies that overlooked this important aspect, the present study relies on a *multi-level* approach, which i) allows for the plausible violation of the *i.i.d.* assumption, ii) defines what to cluster over, and iii) controls for cluster-specific heterogeneity in a nonlinear context.

While *nonlinear multi-level* econometric methods, as claimed by Cameron and Trivedi (2005), have not been well understood, it should be noted that colossally different estimates may be obtained depending on the choice of econometric models when binary multi-level analyses are required (Guo and Zhao 2000; Rodriguez and Goldman 2001; Browne and Draper 2006). Therefore, considering the importance, usefulness, and prevalent unfamiliarity of nonlinear multi-level models, the author

[1] The September 11 attacks were a series of airline hijackings and suicide attacks committed by 19 militants associated with Al-Qaeda against targets in the U.S., the deadliest terrorist attacks on American soil. Some 2750 people were killed in New York, 184 at the Pentagon, and 40 in Pennsylvania. Al-Qaeda is a militant Sunni Islamist multi-national organization and operates as a network of Islamic extremists and Salafist jihadists.

hopes that this study provides reliable methodological guidelines for further refugee-related labor market analyses.

The remainder of this paper is structured as follows. Section 2 provides theoretical considerations with a literature review on the September 11 attacks and their socio-economic impacts from a labor market perspective. Details on the data used for this study and explanations on empirical frameworks are given in Sect. 3, and Sect. 4 discusses econometric methods for nonlinear multi-level analyses.[2] The estimation results of this investigation are presented in Sect. 5, while Sect. 6 explains the robustness checks that have been undertaken.[3] Section 7 concludes.

2 Theoretical considerations

Compared to their host populations, refugees are far disadvantaged in the labor markets of their destinations, and they are highly vulnerable to negative labor market shocks due to many reasons. A lack of local language skills, cultural differences, career interruptions, undervalued previous work experiences, and unappreciated educational attainments are just a few examples. In addition, one of the biggest hindrances that refugees may face while seeking work is direct and indirect discrimination. They are likely targets of discrimination and widely thought to heavily suffer from prejudice (Burawoy 1976; Piore 1980; Evans and Kelley 1991). What is more, such discrimination can be intensified by negative incidents, which, in severe cases, can even lead to extreme xenophobia.

The four coordinated airplane hijackings and suicide attacks on the U.S. by Al-Qaeda on September 11, 2001 were one such example of an extreme incident. The September 11 attacks arguably exacerbated discrimination toward refugees resettled in the U.S.. In other words, the attacks must have aggravated refugees' vulnerability as job seekers. Furthermore, it is predictable that the September 11 attacks may have triggered intensified antipathy against persons perceived to be Africans, Arabs, or Muslims due to their ethnic or religious similarity to the terrorists that perpetrated the attacks.[4] According to Singh (2002), polls conducted by various advocacy groups found that between 20 and 60% of Arab Americans and Muslim Americans said they personally experienced discrimination after the September 11 attacks. They also reported increased incidences of discrimination at work (Kaushal et al. 2007).

From a broader perspective, it is plausible that the September 11 attacks induced a substantial increase in anti-immigrant and anti-foreigner rhetoric as well as anti-Arab and anti-Muslim sentiments. Schüller (2016) found that the attacks caused an immediate shift to more negative attitudes toward immigrants and resulted in a considerable decrease in concerns over xenophobic hostility. Considering this more extensive aspect in conjunction with the aforementioned disadvantages that refugees face, I assert that confining our attention only to Arabs and Muslims and viewing them as the only

[2] Full details on *nonlinear multi-level* econometric methods are discussed in Appendix A.1.

[3] For the sake of brevity, additional robustness checks are presented in Online Supplement B.3.

[4] The perpetrators of the September 11 attacks were 19 men from African or Arab countries who were affiliated with Al-Qaeda.

adversely affected individuals is partial and fragmentary. Rather, analyses should be broadened to include all nonnatives (i.e., all refugees in the case of this study) regardless of their relative similarity to the terrorists who perpetrated the September 11 attacks. This is one of the distinguished contributions of this study compared to other previous papers that only investigated the effects of the attacks on Arabs and Muslims such as Dávila and Mora (2005), Sheridan (2006), Kaushal et al. (2007), Braakmann (2009), and Rabby and Rodgers (2011).

The increased animosity toward refugees, or foreigners in general, engendered by the September 11 attacks may have adversely affected their labor market outcomes in the U.S.. On the labor demand side, such negative impacts may have been manifested in a variety of ways as follows. Most importantly, employers' preferences may have been influenced: in addition, consumers' prejudices against refugees may have become aggravated by the attacks. Even inter-employee discrimination among coworkers is possible. In a word, increased hostility after the attacks may have reduced the demand for refugee labor, which accordingly may have decreased refugees' employment probability. On the labor supply side, the fact that the individuals investigated by this study are refugees obviates the need to consider the elasticity of the supply of labor. Different from economic migrants, refugees have no other options such as returning to home countries than settling down and sustaining themselves by finding employment in their resettlement regions.[5]

Meanwhile, although refugees generally face disadvantages in labor markets and this situation may have been exacerbated by the September 11 attacks, the level of such disadvantages specific to each refugee can be heterogeneous. Considering the cause and agent of the September 11 attacks, it is a sensible guess that refugees with a particular ethnicity, despite all the other important factors held constant, were more (or less) affected than refugees with another ethnicity. In other words, the degree of post-attack prejudice and animosity, which could be causative of labor market discrimination against refugees, may have significantly differed across respective ethnic groups. To put it differently, post-attack anti-refugee sentiment may have been selectively *moderated* for some ethnic groups and discriminately *amplified* for others.

In the case of developed economies, it is often argued that foreign-born workers' primary labor market competitors are not their host populations but other foreign-born job seekers. Since the work of Friedberg and Hunt (1995), a number of studies have corroborated that foreign-born workers do not have a considerable adverse impact on the wages and employment opportunities of the native-born population of the receiving country (Peri 2007; D'Amuri et al. 2010; Peri 2011; Manacorda et al. 2012; Ottaviano and Peri 2012).[6] This is because the jobs they compete for are most closely substitutes for the jobs held by other foreign-born workers, which tend to be more *complementary* to the jobs held by natives.[7] The main reason for this labor market separation resides in the fact that foreign-born workers and native workers have different skills and

[5] Moreover, according to Kaushal et al. (2007), previous empirical studies suggest that the elasticity of the supply of labor is relatively small (Killingsworth 1983; Mroz 1987; Blundell and MaCurdy 1998).

[6] The empirical study of Ceritoglu et al. (2017) argues that Syrian refugees have negatively affected the employment outcomes of natives in Turkey. However, the labor market of Turkey is distinct from that of the U.S. primarily due to the level of its economic development and the prevalence of informal employment.

[7] Such an argument is called *nonnative-native complementarities* in production (Ruist 2013).

characteristics, and nonnatives, especially in developed economies, may take jobs that natives do not want (Peri 2014; Foged and Peri 2016). Applied to the case of refugees, this argument implies that refugees compete with one another as job seekers. Beaman (2012) provides empirical evidence corroborating this aspect that refugees' *mutual competition* decreases their probability of obtaining employment. Dagnelie et al. (2019) also underlines that a refugee's employment is negatively affected by the number of refugees already employed, which substantiates the presence of such mutual competition among refugees in their labor markets. Hence, if refugees with a particular ethnicity were less affected by the negative impacts stemming from the September 11 attacks compared to those with another ethnicity, the less affected presumably *benefited* from the attacks, which selectively increased animosity against the more affected ethnic groups. This is what the present paper seeks to address by examining and quantifying differences in the estimated impacts of the September 11 attacks on labor market outcomes across different ethnic subgroups of refugees resettled in the U.S. Religious subgroups are taken into consideration as well.

3 Data and empirical frameworks

3.1 Data

The repeated cross-sectional data used for this paper describe refugees resettled in the U.S. between 2001 and 2004. Those refugees were distributed to destination cities by a refugee resettlement agency, and this exogenous distribution precludes individual refugees from any systematic sorting.[8] The data were collected by the International Rescue Committee (IRC), a large resettlement agency that helps refugees. In total, the data set contains 1478 adult refugees randomly chosen from a full administrative data set, 65% of whom became employed within 90 days after their arrival in the U.S.. This data set is unique in terms of several key aspects as follows.

First, all individuals in the data set are refugees, not economic migrants. A refugee refers to persons unable to stay in the country of their nationality due to serious persecution or a well-founded fear of persecution. In order to be granted refugee status for immigration purposes, the foreign national must meet clearly defined requirements, the main point of which is that the refugee has a credible and reasonable fear of severe persecution. Since refugees are *forcibly displaced*, their decision function of leaving their home countries is generally less sensitive than that of economic immigrants. Hence, the typical *selection issue* (i.e., selection into defection) is not considered to be serious in the case of this study.[9] Furthermore, this feature also plays an important role as an identification condition for estimating the impacts of the September 11 attacks as such a rigid decision function guarantees that refugees who arrived *before*

[8] This context is similar to the case of Damm (2009a), Damm (2014) and Dustman et al. (2019), which exploit a dispersal policy that assigned refugees to municipalities on a quasi-random basis.

[9] The negative selection hypothesis argued by Borjas (1987) and corroborated by Rooth and Saarela (2007) and Moraga (2011), which claims that it is the lowest-wage, less skilled men who exhibit a stronger tendency to *migrate*, holds only in the case economic migrants (i.e., selection into migration) whose purpose of migration is wealth maximization.

and *after* the attacks are similar in terms of both observables and unobservables. More explanations on identification conditions are provided in Sect. 3.3.[10]

Second, sample respondents did not have any family members who had already settled down in the U.S. and who could thus assist in their resettlement. In other words, they had no consanguinity-based ties in the U.S. by the time they arrived. This feature makes the effect of individual-level covariates, including the regressor of the September 11 attacks, much clearer, since they have no family members to rely on for their initial resettlement. In many empirical studies on refugees' labor market outcomes, having family members established in the host country before a refugee's arrival contaminates estimation results in the form of omitted variable bias unless relevant information is recorded: this data set is completely free from such an issue. Moreover, this aspect also makes it even more important to allow for cluster-specific heterogeneity as their within-cluster (i.e., nationality-based local networks in their resettlement cities) reciprocity is expected to be much stronger compared to refugees who have family members already living in the U.S. that can help them to settle down.[11]

In terms of the dimensions of covariates, various individual characteristics of refugees were recorded at the time of placement, enabling this research to allow for many controls and avoid omitted variable bias as much as possible, such as age, household size, ethnicity, religion, date of birth, initial English language level, previous work experience, education received before coming to the U.S., and so forth. Summary statistics on these extensive covariates are presented in Table 14 in Online Supplement B.4.

Observations on refugees' job acquisition status, the primary dependent variable of this study, were collected at 90 days after each refugee's arrival in the U.S..[12] Such short-term records can be seen as one of the shortcomings of this data set; however, as previously mentioned, since I accentuate the importance of refugees' early integration into host communities' labor markets in the initial phase upon arrival, job acquisition measurement at 90 days after arrival is not improper. Also, it is often said in labor economics that the period of 90 days is not too short for refugees to find jobs (Beaman 2012; Dagnelie et al. 2019). In addition, such a short-term nature makes it even clearer whether the September 11 attacks triggered any changes in the early employment of refugees with significant heterogeneity across different ethnic subgroups, which is the primary question of this study.

Another shortcoming of this data set resides in the fact that it is only composed of male refugees: however, I think this matter is not much problematic in the context of

[10] Also, as previously mentioned in Sect. 2, the fact that all individuals in the data set are refugees obviates the need to consider the elasticity of the labor supply, making the analysis less complicated.

[11] For more details on what to cluster over, see Sect. 4.2.

[12] One caveat that needs to be mentioned is that the variable of the September 11 attacks is dichotomous and coded one for those whose 90-day job search periods were affected by the attacks regardless of the number of affected days. However, this matter is not considered serious in the context of this study because the variables of primary interest are the interaction terms between the attacks variable and the ethnicity variables. Debatable observations are those whose 90-day job search periods started between June 14, 2001 (i.e., one day was affected by the attacks) and September 10, 2001 (i.e., 89 days were affected by the attacks), and it is unlikely that there is a systematic difference in terms of the number of such debatable observations depending on a refugee's ethnicity.

this study due to the reasons as follows. First, the key question of this paper is whether or not the September 11 attacks caused *ethnicity-specific* labor market impacts. Hence, if all ethnic subgroups to be compared were in the same circumstance in terms of their sample composition (i.e., only males), the findings would be still meaningful. Second, while gender may play an important role in affecting labor supply decisions, males usually show comparable patterns in obtaining employment across nations: hence, the labor market outcomes of male refugees can be objectively compared.[13] Third, this study delves into increased prejudice and discrimination as the main source of negative labor market impacts stemming from the September 11. As the paper of Kaushal et al. (2007) argues, there may be less prejudice toward women because women are usually not associated with terrorism. Thus, in studying the labor market impacts of the September 11 driven by employers' prejudice and discrimination, the shortcoming of using the sample only composed of male refugees would not be that much critical.

3.2 Empirical frameworks

The key objective of this study is to shed new light onto whether the impacts stemming from the September 11 attacks were heterogeneous depending on a refugee's ethnicity, an important aspect in consideration of the fact that the attacks were planned by Al-Qaeda and perpetrated by its African and Arab hijackers. For the reason that refugees, as discussed in Sect. 2, usually compete with one another as job seekers in their resettlement regions, if African and Arab refugees were more affected by the negative impacts of the attacks due to the cause and agent of the attacks, it is a sensible hypothesis that refugees of other ethnicities may have benefited from the attacks. This paper seeks to address such a facet by examining and quantifying differences in the estimated impacts of the September 11 attacks on refugees' labor market outcomes across different ethnic subgroups as follows.

$$y_i^m = \theta_0 + \theta_1 \text{Post911}_i + \sum_{k=2}^{5} \pi_k D_{k,i} + \sum_{k=2}^{5} \phi_k (\text{Post911}_i \cdot D_{k,i}) + \mathbf{q}_i' \boldsymbol{\varphi} + \tau_m + u_i \tag{1}$$

In (1), the dependent variable y_i^m is the labor market outcome of refugee i resettled in city m. The regressor Post911_i is a dummy variable and coded one if a refugee's job search periods are affected by the September 11 attacks: zero otherwise. $D_{k,i}$ is also a dummy variable coded one if refugee i's ethnicity belongs to a certain ethnic subgroup k (i.e., $i \in k$), while there are five categories: Africans, Arabs, Asians, Europeans, and South Americans, among which Africans are the baseline reference group in (1). $\mathbf{q}_i$ is a vector of individual characteristics that include age, household size, religion, initial English language level, previous work experience, education received before

[13] Females' labor supply decisions are more difficult to compare due to their varying patterns of selection into work. For details, see Neal (2004), Blau et al. (1990s), Mulligan and Rubinstein (2008), and Olivetti and Petrongolo (2008). Especially, Neal (2004) underlines that relationships between labor force participation and family structure differ notably by ethnicity in the case of females, which can lead to non-comparable selection-into-employment patterns.

coming to the U.S., issuance of a social security number, and so forth. In addition, τ_m is resettlement city fixed effects, and u_i represents unmeasured factors affecting refugee i's labor market outcome. Most importantly, this study delves into the heterogeneity of ϕ_k, and the marginal effect of the attacks on y_i^m is separately estimated for each ethnic subgroup on the basis of θ_1 and ϕ_k.

On the other hand, it is plausible, as mentioned earlier in Sect. 1, that refugees form unique clusters in their resettlement regions, interact with their cluster members, and share some cluster-specific unobservable factors, which can affect a refugee's labor market outcome y_i^m. This facet means that u_i in (1) consists of two components: one caused by clustering and the other purely stemming from idiosyncratic elements. If the former is correlated with regressors, u_i in (1) can cause coefficient estimates to be biased. Therefore, such non-independence among observations caused by clusters should be modeled by *decomposing* u_i into cluster-specific unobservables α_c and idiosyncratic elements ε_i as follows:

$$
\begin{aligned}
y_i^{c,m} = \theta_0 + \theta_1 \text{Post911}_i + \sum_{k=2}^{5} \pi_k D_{k,i} \\
+ \sum_{k=2}^{5} \phi_k (\text{Post911}_i \cdot D_{k,i}) + \mathbf{q}_i' \boldsymbol{\varphi} + \tau_m + \alpha_c + \varepsilon_i,
\end{aligned} \tag{2}
$$

where y_i^m in (1) is changed to $y_i^{c,m}$ in an attempt to highlight the fact that refugee i belongs to cluster c. With regard to α_c, one important caveat should be made that we, in general, cannot simply put cluster-specific dummies in a bid to control for α_c when $y_i^{c,m}$ is binary because of the incidental parameters problem (Neyman and Scott 1948; Hall and Severini 1998).[14] This is why nonlinear multi-level models should be used for the present study, which is discussed in full detail in Sect. 4 and Appendix A.1.

3.3 Identification conditions

Three conditions are required to properly identify the heterogeneous impacts of the September 11 attacks on refugees' labor market outcomes. First, it should be tenable that overall labor market conditions in the U.S. remained stable during the sample period of four years. If fulfilled, this condition makes it more persuasive that no other serious shocks affecting refugees' labor market outcomes occurred, aside from the September 11 attacks. The descriptive statistics from the U.S. Department of Labor depict that the unemployment rate slightly increased starting at the beginning of 2001 due to the recession in the U.S.: however, during the sample period, the difference between before and after the attacks was less than one percentage point. Therefore, the first condition is not considered to be seriously violated.

[14] One exception is the linear probability model.

Second, the primary aim of this study is to address whether the impacts of the September 11 attacks on refugees' labor market outcomes were heterogeneous depending on a refugee's ethnicity. Therefore, the time trends for each ethnic subgroup before the September 11 attacks should not be significantly different, similar to the common (pre-treatment) trend requirement for the difference-in-differences approach. If validated, this common trend, as a necessary condition, makes it more plausible that the heterogeneous impacts were absent prior to the attacks, and hence they, if found significant after the attacks, can be convincingly attributed to the attacks.[15]

Third, as previously mentioned in Sect. 3.1, the fact that the individuals investigated by this study are not economic migrants but refugees plays an important role as an identification condition. Since refugees are forcibly displaced, their decision function of leaving their home countries is generally less sensitive than that of economic immigrants. Such a rigid decision function makes it probable that refugees that arrived before and after the attacks are not much different in terms of both observables and unobservables. Hence, the typical selection issue is not expected to be much serious in the context of this study.

4 Econometric methods

4.1 Non-independence caused by cluster-specific heterogeneity

As previously mentioned in Sect. 3, it is highly plausible that refugees composing the sample are *not independent* for the reason that they form unique clusters in their resettlement regions, interact with their cluster members, and share some cluster-specific unobservable factors. If so, naïve single-level approaches, merely assuming independence among observations and not allowing for cluster-specific heterogeneity, can lead to biased estimates. For such a reason, a multi-level approach is essential in this study.

For multi-level analyses, suppose the ith observation in the overall sample is the jth observation (*i.e., level-one*) in the cth cluster (*i.e., level-two*), while there are C clusters in total, respectively, having N_c observations.[16] Using the binary employment outcome at 90 days after arrival as its main dependent variable, this study uses the following two-level specification:

$$\Pr(y_{jc} = 1 \mid \mathbf{x}_{jc}, \alpha_c) = \Pr(y_{jc} = 1 \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = E(y_{jc} \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c), \quad (3)$$

where observable covariates $\mathbf{x}_{jc}$ are partitioned into $\mathbf{v}_{jc}$ and $\mathbf{i}_c$ for reasons to be elaborated in a later section. $\mathbf{v}_{jc}$ is a vector of level-one covariates varying within clusters, while $\mathbf{i}_c$ is that of level-two covariates invariant within clusters. Most importantly, α_c is cluster-specific heterogeneity in the form of random intercepts. Unlike $\mathbf{v}_{jc}$ and $\mathbf{i}_c$ that

[15] Whether this requirement is satisfied or not was tested by exploiting that the data set segregates the periods before the September 11 attacks into three quarters, with which the time trends for each ethnic subgroup before the attacks can be compared. The null hypothesis of the same time trends before the attacks could not be rejected.

[16] It is assumed throughout this paper that N_c is exogenous.

are observed, α_c, cluster-specific fixed effects commonly shared by all the members of a certain cluster c, is not directly observed in the data *per se*, and it substantially differs depending on what to cluster over. Before discussing what to cluster over, it is worthwhile pinpointing how clustering affects estimations.

Non-independence among observations caused by clustering affects two distinct aspects of estimations. First, if unobserved cluster-specific heterogeneity α_c is correlated with covariates $\mathbf{x}_{jc}$, which implies $E\left(\alpha_c|\mathbf{x}_{jc}\right) \neq 0$, its existence concerns the estimation of coefficients. Hence, in such a case, ignoring the presence of α_c leads to omitted variable bias and causes inconsistent coefficient estimates. This issue can be circumvented by directly modeling clusters and including their fixed effects α_c. However, in the case of nonlinear response models, another complication arises in the sense that we, in general, should not simply put cluster dummies in a bid to control for α_c because of the incidental parameters problem (Neyman and Scott 1948). This difficulty is further discussed in Appendix A.1.

The second aspect that needs attention is the existence of within-cluster error correlation, which affects variance estimates. If observations are clustered, which can cause errors to be non-independent, the usual variance formula underestimates the true variance matrix, assuming positive within-cluster error correlation (Cameron and Trivedi 2005). Pertaining to this matter, the solution is simpler: the variances of the regression parameters need to be adjusted in consideration of clustered errors.

The previous two paragraphs can be summarized as follows. If α_c is a pure random effect uncorrelated with covariates $\mathbf{x}_{jc}$, then only the variances of the regression parameters need to be adjusted, and a single-level approach is not much problematic. However, if α_c is correlated with $\mathbf{x}_{jc}$, then the regression parameters based on the omission of α_c are inconsistently estimated, and a multi-level approach, which allows for the presence of α_c in a true model and considers $E\left(\alpha_c|\mathbf{x}_{jc}\right) \neq 0$, is necessary. On the positive side, whether α_c is correlated with $\mathbf{x}_{jc}$ or not is testable: the test results are discussed in Online Supplement B.2.

4.2 What to cluster over

As described in the previous section, non-independence among observations caused by clustering necessitates considering two distinct fronts—coefficient estimates (affected by α_c) and variance estimates (affected by within-cluster error correlation). The former aspect should be dealt with by directly modeling clusters (i.e., see (1) and (2)), and the latter one should be taken care of by using cluster-robust variance matrices. While both econometric measures require a clear-cut decision on what to cluster over, there is no broadly accepted guidance encompassing both. Thus, it is often not clear what to cluster over and how to define clusters. Moreover, there is not even a formal test of the level at which to cluster.

Therefore, empirical thoughts should play a role in determining what to cluster over. To begin with, what can cause the existence of α_c should be taken into account. As mentioned in Sect. 3, since the refugees investigated by this study do not have any family members who had already settled down in the U.S. and who could thus assist in their resettlement, the importance of considering compatriot-based networks and enclaves should be more distinguished, and the unobserved features of such networks that can affect y_{jc} are expected to lead to the existence of α_c. Therefore, concerning how to model clusters and their fixed effects α_c, which is one of two factors broaching the subject of what to cluster over, this study assumes that refugees i) from the same home country and ii) being assigned to the same resettlement city form their unique networks and interact with existing enclaves of their compatriots in that city. This is in accordance with how clusters are defined in Beaman (2012) and can be expressed as follows:

$$\Pr(y_{imh} = 1 \mid \mathbf{v}_{imh}, \mathbf{i}_{mh}, \alpha_{mh}) = \Pr(y_{jc} = 1 \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c). \tag{4}$$

In (4), h denotes a refugee's home country, and m represents a refugee's resettlement city in the U.S.. c is a refugee's cluster defined by the unique combination of h and m (i.e., clustering at the *city-nationality* level). This decision on what to cluster over is also supported by the argument of Topa (2001) and Damm (2014) that there are residence-based job information networks stratified by where a nonnative comes from.

The second aspect to be considered concerns within-cluster error correlation, the issue of which sometimes cannot be completely resolved even by modeling clusters and including their fixed effects α_c.[17] In this study, the same logic described above in (4) is applied to how to adjust the variances of the regression parameters as well. On the other hand, we, however, cannot rule out the possibility that refugees form more aggregated networks, for example, based on whether they come from the same continent given that they reside in the same resettlement city. This leads to the likelihood that within-cluster error correlation is made within bigger, more aggregated clusters (e.g., clustering at the *city-continent* level) although α_c itself is only caused by subordinate lower clusters. Hence, this plausible error correlation within more aggregated clusters is taken into consideration in Online Supplement B.3.2 for a robustness check. It is motivated by the fact that different assumptions (i.e., different clustering levels) can be made about each of α_c and within-cluster error correlation: it is further encouraged by the suggestion of Pepper (2002) and Cameron and Miller (2015) that the consensus, when it comes to the use of cluster-robust variance matrices, is to be conservative by using bigger and more aggregated clusters.

4.3 Nonlinear multi-level approaches for clustered observations

Once what to cluster over has been determined, the next step is to decide which multi-level model to use, and in doing so, the nonlinear feature of the primary dependent variable of this study should be taken into consideration.[18] As previously discussed, non-independence caused by clustering affects two aspects of estimations—coefficient estimates through α_c and variance estimates through within-cluster error correlation. However, economists have been mainly focused on the latter without paying enough attention to the former.

[17] For more details on this facet, see Cameron and Miller (2015).

[18] See, among others, Pendergast et al. (1996) for an overview of the extensive literature on nonlinear multi-level models in biostatistics.

When the ith observation in the overall sample is the jth observation in the cth cluster, the most general specification for a binary multi-level approach is

$$\begin{aligned} y^*_{jc} &= \mathbf{x}'_{jc}\boldsymbol{\beta} + u_{jc} \\ &= \mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + u_{jc} \\ &= \mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c + \varepsilon_{jc}, \end{aligned} \tag{5}$$

where y^*_{jc} is a latent continuous variable—all the other notations are the same as defined above. Whether y^*_{jc} crosses a threshold or not decides the observed binary outcome y_{jc} in the following way:

$$y_{jc} = 1[\mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c + \varepsilon_{jc} > 0], \tag{6}$$

where $1[\cdot]$ refers to the indicator function. Multi-level approaches are distinguished from single-level approaches in the sense that they decompose u_{jc} into α_c and ε_{jc} and do not set $\alpha_c = \alpha$ for those belonging to different c.

Generalized linear models (GLM), which can be applied to an entire class of models with dependent variables following a distribution from the exponential family, can be extended to analyze multi-level clustered data as shown below:

$$\Pr(y_{jc} = 1 \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = E(y_{jc} \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = g(\mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c), \tag{7}$$

where g refers to link functions (Neuhaus and Kalbfleisch 1998; Neuhaus and McCulloch 2006; Brumback et al. 2010). The model (7) above, which is often called a generalized linear mixed model (GLMM), stands on the assumption that the effects of level-one covariates $\mathbf{v}_{jc}$ are the same across all clusters; hence, α_c solely captures cluster-specific heterogeneous features.[19]

Concerning the choice of g in (7), the cumulative distribution function of ε_{jc} in (6) should be considered. If probit is used as a link function based on the assumption

$$\varepsilon_{jc} \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c \sim N(0, \sigma^2_{\varepsilon}), \tag{8}$$

(7) leads to

$$\Pr(y_{jc} = 1 \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = \Phi(\mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c), \tag{9}$$

where $\Phi[\cdot]$ is the standard normal cumulative distribution function. If logit is to be used as an alternative link function in place of probit, we can replace $\Phi[\cdot]$ with $\Lambda[\cdot]$, which refers to the logistic cumulative distribution function. By the same token, if the linear probability model (LPM), which does not use a cumulative distribution function, is to be used, we can simply substitute the identity link for $\Phi[\cdot]$.[20]

[19] If this assumption is expected to be violated, the model should become more flexible by additionally including cluster-specific random slopes in addition to random intercepts, which can cause a greater computational burden.

[20] For details on the linear probability model, see Appendix A.1.3.

Another key decision to be made concerns which model to use between two dichotomized multi-level models—*the fixed effects* (FE) *model* and *the random effects* (RE) *model*. As mentioned above, multi-level approaches are distinguished from single-level ones in that they consider $u_{jc} = \alpha_c + \varepsilon_{jc}$ and do not set $\alpha_c = \alpha$, and different assumptions on the correlation between α_c and covariates lead to those two different models.

The random effects model, which regards α_c as purely random intercepts whose distribution does not depend on any covariates, uses both within- and between-cluster variations and accordingly enables estimating the effects of both individual-specific level-one variables $\mathbf{v}_{jc}$ and cluster-invariant level-two variables $\mathbf{i}_c$. However, the required assumption that cluster-specific unobservables are not correlated with regressors is very restrictive and often implausible, and the violation of this assumption leads to inconsistent estimates. On the other hand, the fixed effects model, which also regards α_c as random unobservables but allows for the possible correlation between α_c and covariates, can provide unbiased estimates of level-one variables $\mathbf{v}_{jc}$. However, the fixed effects model is less efficient than the random effects model and cannot provide estimates for cluster-invariant level-two variables $\mathbf{i}_c$ because it only uses within-cluster variations. In a nutshell, each of these two models takes an opposite stand concerning the common econometric trade-off between efficiency versus consistency. Considering each model's advantages and disadvantages, this study uses both models. Full details on *nonlinear multi-level* econometric methods are discussed in Appendix A.1.

5 Estimation results

5.1 The impacts of the September 11 attacks on refugees' employment

Before starting a full discussion on main estimation results, I present two diagnostic tests and their results that adduce convincing evidence on why the multi-level approach is indispensable for this study. For the sake of brevity, their details are discussed in Online Supplement B.2.

Table 1 compares estimates from different model specifications based on the Chamberlain–Mundlak's correlated random effects probit model, while Table 2 makes comparisons of those based on the conditional logit fixed effects model. Note that the estimated coefficients in Tables 1 and 2 are different from marginal effects due to the nonlinear nature of both models. The signs of coefficient estimates, however, show whether a certain regressor has a positive or negative marginal effect on the response probability (Wooldridge 2003). Therefore, a discussion based on the signs and statistical significance of coefficient estimates is presented first, and it is followed by explanations on marginal effects and odds ratios in later sections.

Above all, Tables 1 and 2 show apparently that the coefficient estimates of the September 11 attacks variable (i.e., Post 9/11) are significantly negative, meaning that the attacks appear to have lowered refugees' employment probability significantly. Moreover, the magnitudes, in both Tables 1 and 2, are far more substantial than those of the English level 3 variable and of the summer variable, two other important factors arguably affecting refugees' labor market outcomes.

Table 1 CRE—Coefficient estimates on employment

DV=Employed (1) or not (0) Estimator Link function	[1] CRE Probit	[2] CRE Probit	[3] CRE Probit	[4] CRE Probit
Ethnicity (Ref.: African)	Yes	Yes	Yes	Yes
Religion (Ref.: Christian)	Yes	Yes	Yes	Yes
Local language level (Ref.: English level 2)				
English level 1	−0.159	−0.160	−0.172	−0.168
	[0.116]	[0.117]	[0.117]	[0.118]
English level 3	0.293**	0.304**	0.291**	0.301**
	[0.120]	[0.123]	[0.120]	[0.124]
External shock				
Post 9/11	-0.484***	-0.811***	-0.688***	-0.922***
	[0.149]	[0.177]	[0.220]	[0.209]
Seasonality				
Summer	−0.210*	−0.194*	−0.204*	-0.193*
	[0.116]	[0.114]	[0.115]	[0.114]
Interaction terms (Post 9/11 & Ethnicity)				
Post 9/11 × South American		0.120		0.361
		[0.544]		[0.686]
Post 9/11 Asian		1.439***		1.480***
		[0.478]		[0.545]
Post 9/11 × Middle Eastern		0.161		0.133
		[0.340]		[0.319]
Post 9/11 × European		0.989***		0.884**
		[0.338]		[0.361]
Interaction terms (Post 9/11 & Religion)				
Post 9/11 × Muslim			0.243	0.189
			[0.267]	[0.262]
Mundlak test				
p-value	< 0.001	< 0.001	< 0.001	< 0.001
AIC	1665.08	1660.23	1668.42	1668.05
No. of observations	1478	1478	1478	1478

(i) Cluster-robust standard errors are shown in brackets. (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively. (iii) Other controls: Age, age squared, household size, household size squared, English level missing category, education level, education level missing category, religion missing category, no job market experiences, professional jobs in home countries, exemption from employment by IRC, matching grant program participation, delayed issuance of a social security number, and the number of refugees and migrants resettled in the same city from years t through $t-4$ with the same nationality as the surveyed individuals (i.e., collected by Beaman (2012))

Column 1 of Tables 1 and 2 already contains a wide range of control variables such as basic demographic information, education levels, English levels, and job market

Table 2 CML FE—Coefficient estimates on employment

DV=Employed (1) or not (0) Estimator Link function	[1] CML FE Logit	[2] CML FE Logit	[3] CML FE Logit	[4] CML FE Logit
Ethnicity (Ref.: African)	Yes	Yes	Yes	Yes
Religion (Ref.: Christian)	Yes	Yes	Yes	Yes
Local language level (Ref.: English level 2)				
English level 1	−0.278	−0.270	−0.291	−0.276
	(0.202)	(0.205)	(0.203)	(0.206)
English level 3	0.538**	0.574**	0.529**	0.566**
	(0.250)	(0.253)	(0.251)	(0.254)
External shock				
Post 9/11	−0.841***	−1.466***	−1.235***	-1.689***
	(0.204)	(0.356)	(0.334)	(0.424)
Seasonality				
Summer	−0.313**	-0.259	-0.297*	-0.258
	(0.157)	(0.160)	(0.158)	(0.161)
Interaction terms (Post 9/11 & Ethnicity)				
Post 9/11 × South American		0.322		0.728
		(0.749)		(0.810)
Post 9/11 × Asian		2.489***		2.552***
		(0.902)		(0.972)
Post 9/11 × Middle Eastern		0.300		0.263
		(0.482)		(0.502)
Post 9/11 × European		1.787***		1.604**
		(0.619)		(0.639)
Interaction terms (Post 9/11 & Religion)				
Post 9/11 × Muslim		0.465		0.372
		(0.427)		(0.453)
AIC	1117.01	1111.38	1119.09	1115.71
No. of observations	1295	1295	1295	1295

(i) Conventional standard errors are shown in parentheses. Cluster-robust standard errors are not used here since the conditional logit fixed effects model relies on the conditional independence assumption, completely controlling for cluster-specific heterogeneity. (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively. (iii) Other controls: Age, age squared, household size, household size squared, English level missing category, education level, education level missing category, religion missing category, no job market experiences, professional jobs in home countries, exempted from employment by IRC, matching grant program participation, delayed issuance of a social security number, and the number of refugees and migrants resettled in the same city from years t through $t-4$ with the same nationality as the surveyed individuals (iv) The number of observations is smaller since some observations are dropped if their within-cluster variation in $N_c \overline{y}_c$ is zero

experiences before coming to the U.S.. However, the interaction terms between the September 11 attacks and the ethnicity variables are omitted.

Column 2 of Tables 1 and 2 additionally includes the interaction terms between the September 11 attacks and the ethnicity variables, the baseline reference group of which is African refugees. The inclusion of these interaction terms considerably increases the estimated magnitudes of the September 11 attacks by 67% in Table 1 and 74% in Table 2, respectively. This is because the inclusion of such interaction terms makes the coefficient estimates reflect the impact of the September 11 attacks specifically to African refugees. Moreover, although all interaction terms show positive signs, their magnitudes and statistical significance remarkably vary, implying highly heterogeneous moderation effects depending on a refugee's ethnicity.

Columns 3 and 4 of Tables 1 and 2 intend to check whether the impacts of the attacks differ depending on a refugee's religion. The attacks-religion interaction terms, not fully shown for the sake of brevity, are both marginally and jointly insignificant when included without (Column 3) and with (Column 4) the attacks-ethnicity interaction terms.

This insignificance of the attacks-religion interaction terms is understandable, since job seekers often are not necessarily required to disclose their religion while searching for a job. On the other hand, ethnicity is, of course, easily exposed through many channels, the most explicit one of which may be visual identification and names, as shown by the experiment of Widner and Chicoine (2011). Moreover, race and gender information is officially required to be collected under the relevant law in the U.S., while religion information is not mandated.[21] Therefore, it can be concluded that a refugee's religion neither intensifies nor alleviates the impacts of the September 11 attacks. Instead, the observed heterogeneous moderation effects can be solely attributed to a refugee's ethnicity.

On the other hand, the summer variable, a seasonal shock to refugees' labor market, also seems to have negatively affected their employment. This is unambiguous, as many employment decision makers take vacation during the summer, which delays the recruiting process. Interestingly, the interaction terms between the summer and ethnicity variables, not shown for the purpose of brevity, turn out to be both marginally and jointly insignificant when tested with and without the attacks-ethnicity interaction terms. Taken together, these results suggest that summer was an equally negative shock, whereas the September 11 attacks, due to their contextual reasons, were not equally adverse for all refugees.

Table 3 compares coefficient estimates from different models based on the final specification.[22] Note that Columns 2, 3, and 4 of Table 3 cannot be directly compared, since the link function of each model is different. However, both of Columns 1 and 2 are based on probit and thus are directly comparable. The key difference resides in the fact that Column 1 is from a naïve single-level approach, whereas Column 2 is from a multi-level approach. As the substantial differences between Columns 1 and 2 in Table 3 and the results of the Mundlak test in Table 1 show, it is apparent that the naïve single-level approach results in biased estimates, primarily underestimating

[21] For example, the Equal Employment Opportunity (EEO) laws collect such information in order to prohibit specific types of job discrimination in certain workplaces.

[22] Year-specific dummies turn out to be both marginally and jointly insignificant , and the inclusion of year dummies does not substantially affect the estimation results. For a comparison, see Table 15 in Online Supplement B.4.

Table 3 RE, CRE, CML FE, and LPM—Coefficient estimates on employment

DV=Employed (1) or not (0)	[1]	[2]	[3]	[4]
Level	Single-level	Multi-level	Multi-level	Multi-level
Estimator	Pooled	CRE	CML FE	LPM FE
Link function	Probit	Probit	Logit	Identity
Cluster-specific heterogeneity	No	Yes$^{(iv)}$	Yes	Yes
Ethnicity	Yes	Yes	Yes	Yes
Religion	Yes	Yes	Yes	Yes
Local language level (Ref.: English level 2)				
English level 1	-0.083	−0.160	-0.270	-0.043
	[0.113]	[0.117]	(0.205)	[0.034]
English level 3	0.301***	0.304**	0.574**	0.084**
	[0.114]	[0.123]	(0.253)	[0.035]
External shock				
Post 9/11	−0.650***	-0.811***	-1.466***	-0.221***
	[0.149]	[0.177]	(0.356)	[0.046]
Seasonality				
Summer	−0.214**	−0.194*	−0.259	−0.053*
	[0.106]	[0.114]	(0.160)	[0.032]
Interaction terms (Post 9/11 & Ethnicity)				
Post 9/11 × South American	0.191	0.120	0.322	0.044
	[0.446]	[0.544]	(0.749)	[0.123]
Post 9/11 × Asian	0.981**	1.439***	2.489***	0.355***
	[0.400]	[0.478]	(0.902)	[0.130]
Post 9/11 × Middle Eastern	0.134	0.161	0.300	0.021
	[0.277]	[0.340]	(0.482)	[0.102]
Post 9/11 × European	0.604**	0.989***	1.787***	0.262***
	[0.294]	[0.338]	(0.619)	[0.087]
No. of observations	1478	1478	1295$^{(v)}$	1478

(i) Cluster-robust standard errors are in brackets; conventional standard errors are shown in parentheses. (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively. (iii) Other controls: Age, age squared, household size, household size squared, English level missing category, education level, education level missing category, religion missing category, no job market experiences, professional jobs in home countries, exempted from employment by IRC, matching grant program participation, delayed issuance of a social security number, and the number of refugees and migrants resettled in the same city from years t through $t-4$ with the same nationality as the surveyed individuals (iv) Through the linear projection of cluster-means of covariates (v) The number of observations is smaller since some observations are dropped if their within-cluster variation in $N_c\overline{y}_c$ is zero

the coefficients of the attacks-ethnicity interaction terms and failing to show clear heterogeneity across ethnic subgroups. On the other hand, all the other multi-level estimates, despite different link functions, consistently show evident heterogeneity depending on a refugee's ethnicity, with similar statistical significance.

The estimates from all multi-level models in Columns 2, 3, and 4 of Table 3 suggest that Asian and European refugees were significantly less affected by the negative

Table 4 Marginal effects and odds ratios of the September 11 attacks on employment

DV=Employed (1) or not (0) Interpretation	[1] AME	[2] AME	[3] Odds ratio
Estimator Link function	CRE Probit	LPM FE Identity	CML FE Logit
African refugees	−0.1743***	-0.2209***	0.2308***
	[0.0337]	[0.0456]	95% CI: [0.1149, 0.4636]
South American refugees	−0.1763	−0.1770	0.3184*
	[0.1262]	[0.1168]	95% CI: [0.0841, 1.2056]
Asian refugees	0.1148*	0.1344	2.7811
	[0.0695]	[0.1213]	95% CI: [0.5488, 14.0941]
Middle Eastern refugees	-0.1584**	-0.1996**	0.3116**
	[0.0688]	[0.0939]	95% CI: [0.1535, 0.6329]
European refugees	0.0124	0.0412	1.3781
	[0.0187]	[0.0772]	95% CI: [0.5100, 3.7239]
No. of observations	1478	1478	$1295^{(iii)}$

(i) Cluster robust standard errors are in brackets. (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively. (iii) The number of observations is smaller since some observations are dropped if their within-cluster variation in $N_c\overline{y}_c$ is zero. (iv) Standard errors of odds ratios are less meaningful: thus, for odds ratios, 95% confidence intervals are shown instead

impacts of the September 11 attacks compared to African refugees. However, such significant differences are not observed in the case of Middle Eastern and South American refugees. In sum, the multi-level estimates imply that only Asian and European refugees were significantly *cushioned* by their ethnic backgrounds from the post-attack anti-refugee sentiments.

The heterogeneity of these interaction terms necessitates separately examining the marginal effects of the September 11 attacks for each ethnic subgroup. Columns 1 and 2 of Table 4 present the average marginal effects (AME) of the September 11 attacks on a refugee's employment probability. The marginal effects in Column 1 are estimated by the Chamberlain–Mundlak's correlated random effects probit model, and those in Column 2 are estimated by the linear probability model with cluster fixed effects.

A huge external shock to the labor market of refugees resettled in the U.S., the September 11 terrorist attacks significantly lowered African refugees' employment probability by 17 to 22% points. Furthermore, the attacks reduced Arab refugees' employment probability by 15 to 19% points. In the case of European refugees, the null hypothesis of no impact cannot be rejected with the magnitudes very close to zero due to the strong ethnicity-based moderation effect.

The more surprising finding resides in the fact that the September 11 attacks, against common expectations, substantially increased Asian refugees' employment probability by 11–13% points. Despite such substantial magnitudes, the lack of statistical significance is not unusual and can be attributed to the fact that the average marginal effects of the attacks are determined by the sum of two estimates with opposite signs. Moreover, according to Greene (2007), the inference on heterogeneity should be made

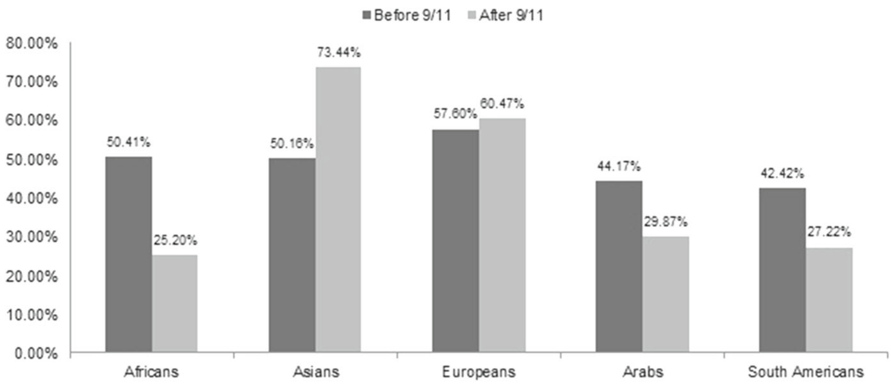

Fig. 1 Changes in predicted employment probabilities

not with marginal effects but with coefficients.[23] Therefore, it should be noted that the heterogeneous impacts stemming from the September 11 attacks are already corroborated in Table 3.

Estimates from the conditional logit fixed effects model are shown in Column 3 of Table 4 in the form of odds ratios. The odds ratio of the September 11 attacks is 0.23 and 0.31 for refugees from Africa and the Middle East, respectively. This means that the odds of being employed significantly decreased by a factor of 0.23 and 0.31 in multiplicative terms. For the other ethnic subgroups, the results are statistically insignificant. However, the odds ratio for Asian refugees, albeit insignificant, suggests that the September 11 attacks, when all other important covariates are controlled for, made Asian refugees' odds of being employed 2.78 times higher. The odds ratio for European refugees is close to one, which means almost no impact in multiplicative terms.

Based on these estimation results, it can be concluded that the September 11 attacks had varying impacts on each ethnic subgroup of refugees seeking employment in the U.S. Due to the context of the attacks, their negative impacts on employment were significant only in the case of African and Arab refugees. A plausible explanation for such heterogeneous impacts is that the September 11 attacks triggered increased animosity selectively against African and Arab refugees due to their ethnic similarity to the terrorists. Also, employers may have perceived hiring African or Arab refugees as risky and costly, due to either security concerns or uncertainty over the permanency of their stay in the U.S.. Therefore, the shock of the attacks was not equally negative for all refugees, and the degrees of post-attack prejudice and discrimination must have differed across ethnic subgroups. Figure 1 sums up these findings in the form of changes in predicted employment probabilities for each ethnic subgroup.[24]

[23] This is because a hypothesis tested, in the case of marginal effects, is about a function of all coefficients (Greene 2002).

[24] Figure 1 is based on those who had some job experiences in their home countries, completed secondary education and could speak English at an intermediate level when entering the U.S.. These features form one of the most common types that we can think of for job-seeking refugees.

Above all, the most interesting finding lies in the fact that the September 11 attacks gave Asian refugees better career opportunities, as the attacks appear to have substantially increased their employment probability. Nevertheless, it is unlikely that Asian refugees experienced an advantage over American citizens; instead, they arguably held an advantage over other refugees as their direct job market competitors. As discussed in Sect. 2, Friedberg and Hunt (1995), Peri (2007), Ottaviano and Peri (2012), and Foged and Peri (2016) argue that foreign-born workers' primary labor market competitors, in developed economies, are not their host populations but other nonnatives. While refugees compete with one another as job seekers, any selectively increased antipathy against African and Arab refugees triggered by the September 11 attacks may have boosted American employers' preference for hiring Asian refugees for jobs that all refugees had equally competed for prior to the attacks.

Meanwhile, Table 4 leaves one question open concerning why Arab refugees were slightly less affected than African refugees, although they were likely perceived as being more similar to the September 11 hijackers, the group of which was composed of Arabs that outnumbered and only a few Africans. One possible explanation is that the strong social networks and solidarity of Arab Muslims in the U.S. may have exclusively helped their members find jobs, while they were not required to disclose their religion, presumably unfavorable information, to employers when searching for a job. To check the plausibility of this elucidation, the marginal effects of the September 11 attacks on employment were separately estimated for each ethnicity-religion subgroup. As shown in Table 5, it seems that the negative impacts of the attacks affected Arab Muslim refugees far less than Arab non-Muslim refugees. In the case of Arab non-Muslim refugees, the negative marginal effects are even greater than those of African refugees from a magnitude standpoint. In other words, being an Arab Muslim was not a *double disadvantage*; rather, it may have functioned as a *buffer* against the negative impacts stemming from the September 11 attacks. Interestingly, such Muslim solidarity does not seem to have worked among African refugees. Further studies regarding this aspect would be worthwhile.

5.2 Interstate regional heterogeneity

Following the September 11 attacks, political researchers found increases in both individuals' trust in government and conservative political attitudes (Ford et al. 2001; Bonanno and Jost 2006; Hersh 2013). Moreover, such increases were found to be linked to authoritarian political views (Jost et al. 2003; Huddy et al. 2005; Huddy and Feldman 2011). For this reason, it is likely that the Americans most mentally and emotionally affected by the attacks became more Republicans relative to those less affected, and the gubernatorial elections held between 2001 (after the attacks) and 2004 may have reflected this facet. Therefore, in a bid to examine interstate regional heterogeneity in terms of the labor market impacts stemming from the September 11 attacks, the 16 resettlement cities were classified into two types, as shown in Table 16 in Online Supplement B.4. Regions where Democratic candidates won were categorized as liberal areas, while regions where Republican candidates won were categorized

Table 5 Marginal effects of the September 11 attacks on employment

DV=Employed (1) or not (0)	
Estimator	LPM FE
Link function	Identity
	AME
African refugees	
African Christian refugees	−0.2134***
	[0.0518]
African Muslim refugees	−0.2716***
	[0.0804]
African non-believer refugees	No obs.
	-
African refugees with other religions	No obs.
	-
Arab refugees	
Arab Christian refugees	-0.3109**
	[0.1410]
Arab Muslim refugees	-0.1393**
	[0.0684]
Arab non-believer refugees	-0.4241***
	[0.1469]
Arab refugees with other religions	−0.3697
	[0.3846]
No. of observations	1478

(i) Cluster robust standard errors are in brackets. (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively

as conservative areas. Following this, three-way interaction terms were included for estimating marginal effects in each category.

Table 6 presents the results in light of regional heterogeneity and reveals some noteworthy empirical findings as follows. African refugees were uniformly affected by the negative impacts of the attacks, yet Arab refugees were far more adversely affected in conservative regions than in liberal regions. In other words, Arab refugees were the hardest hit group in conservative regions. For European refugees, the impact magnitudes were universally very close to zero. Finally, the comparative benefit that Asian refugees enjoyed in employment was far greater in conservative regions. Taken together, these results mean that there was remarkable interstate regional heterogeneity when it comes to the impacts stemming from the terrorist attacks.

In a bid to confirm such regional heterogeneity, each region's friendliness to non-natives was additionally used as another classification criterion. For this purpose, the General Social Survey conducted in 2000 was used, which contained a relevant question.[25] Based on this, the 16 resettlement cities were categorized into two types (i.e., nonnative-friendly regions versus nonnative-unfriendly regions), and three-way inter-

[25] The question asked whether a respondent thinks the number of immigrants from foreign countries who are permitted to come to the U.S. to live should be decreased.

Table 6 Marginal effects of the September 11 attacks on employment—Regional heterogeneity based on gubernatorial election results

DV=Employed (1) or not (0)	
Estimator	LPM FE
Link function	Identity
Estimand	AME
African refugees	
In liberal regions	−0.2897***
	[0.0806]
In conservative regions	−0.2067***
	[0.0554]
South American refugees	
In liberal regions	−0.2529**
	[0.1269]
In conservative regions	−0.0087
	[0.0376]
Asian refugees	
In liberal regions	0.0920
	[0.1897]
In conservative regions	0.3208**
	[0.1493]
Middle Eastern refugees	
In liberal regions	0.0201
	[0.1222]
In conservative regions	−0.3271***
	[0.1088]
European refugees	
In liberal regions	0.0124
	[0.1367]
In conservative regions	0.0734
	[0.0736]
No. of observations	1478

(i) Cluster robust standard errors are in brackets. (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively

action terms were included for estimating marginal effects in each type.[26] The results, presented in Table 17 in Online Supplement B.4, are in line with what is shown in Table 6.

5.3 The impacts of the September 11 attacks on refugees' wage levels

Notably, the September 11 attacks had a substantially positive impact on the employment probability of Asian refugees. As refugees compete with one another in looking

[26] This paper considers a region to be non-friendly toward immigration if more than 40% of respondents say that they think the number of immigrants to the U.S. should be limited.

for employment, it could be argued that Asian refugees benefited from the attacks probably due to employers' selectively increased antipathy against African and Arab refugees. If that is the case, however, why did European refugees not benefit from the attacks? On the other hand, it is possible that European refugees conceivably benefited from the attacks in terms of employment quality (i.e., higher earnings), not employment probability. To answer this question, whether the September 11 attacks affected refugees' earnings is investigated.

Before proceeding to the wage analysis to determine whether the September 11 attacks improved (or worsened) the quality of jobs held by refugees, a sample selection problem must be thoroughly checked because refugees with recorded wage observations are *selected samples* based on values taken by the dependent variable of the selection (i.e., employment) equation, which could lead to the dependence of the employment equation and the wage equation in their unobservables (i.e., error terms) and hence cause wage estimates to be inconsistent if overlooked.

Based on the well-known bivariate maximum likelihood (i.e., often abbreviated as ML) sample selection model and the Heckman two-step estimator from Heckman (1979), the existence of the sample selection problem was checked. The results, not shown for the sake of brevity, are comforting in that both models indicate the absence of the selection issue.[27] Accordingly, I now proceed to the question of whether the September 11 attacks improved or worsened the quality of jobs held by refugees, which is measured by real hourly wages.

As can be seen in Table 7, only European refugees saw a statistically significant rise in their real hourly wages at the 10% significance level, in addition to the greatest magnitude. Despite its weak statistical significance, which requires further investigations with more observations, the estimate of European refugees, from a magnitude standpoint, is economically substantial in that their 0.491 dollars per hour increase in wages is 37% of a standard deviation of the wage distribution of all employed refugees during the same period. This result indicates that the quality of jobs held by European refugees substantially improved after the September 11 attacks. For Asian refugees, whose employment probability markedly increased after the attacks, their wage level increase experienced the second-highest magnitude but without statistical significance even at the 10% level. Meanwhile, African, Middle Eastern, and South American refugees saw neither a statistically significant nor a substantial increase in their wages. Taken together, these results suggest that the quality of jobs improved only for European refugees after the attacks: such a rise in wages may have offset the otherwise possible increase in the employment probability of European refugees. Asian refugees, on the other hand, did not reap the additional benefit of getting better quality jobs.

Then, one necessary question worth following up is what caused such different patterns of benefits between Asian and European refugees. A likely cause for the divide is that Asian and European refugees had different preferences in the labor market of refugees—easier and quicker job acquisition for Asian refugees and better quality jobs for European refugees. These heterogeneous preferences can be attributed to the

[27] The selection parameter ρ was estimated to be very close to zero from a magnitude standpoint along with its statistical insignificance.

Table 7 Coefficient estimates of the September 11 attacks on wage levels

DV=Real hourly wage	
Level	Multi-level
Estimator	Heckman selection
African refugees	−0.0796
	[0.2187]
South American refugees	0.0958
	[0.2990]
Asian refugees	0.2666
	[0.3897]
Middle Eastern refugees	−0.1709
	[0.2569]
European refugees	0.4913*
	[0.2645]

(i) Bootstrap standard errors are shown in brackets. (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively. (iii) The sample selection problem is confirmed to be absent

possibility that their expectations for future labor market outcomes differed during the initial resettlement period, in which they received financial aid and diverse care from the US government as legally accepted refugees and thus did not have impending concerns about their livelihoods yet.[28]

If most European refugees tend to have greater expectations of finding better jobs after the initial period of support due to some reasons, such as cultural and racial similarities with the mainstream American society, the level of their reservation wages is likely higher during the initial resettlement period than that of most Asian refugees, who have lower expectations for their future labor market outcomes. Accordingly, some European refugees with sanguine expectations may have *intentionally delayed* employment if proposed market wages fell short of their high reservations wages. Despite the fact that this *wait-and-see* strategy takes time, they, with a lack of urgency stemming from the financial aid and diverse care from their host government, may have still preferred waiting for good quality jobs rather than hastily accepting mediocre employment offers that would only make them no longer qualify for the aid and care.

If this explanation is reasonable, the selection into employment of European refugees should be negatively made because such *wait-and-see* behaviors and intentional delay in obtaining employment of those with higher reservation wages can cause the wage level of employed refugees to be lower than that of non-employed ones.[29] Furthermore, those with higher reservation wages usually have a higher rate of unobserved capabilities and skills, which are unobservable components in both the employment equation and the wage equation. Hence, their *wait-and-see* strategy for targeting high paying jobs, which *increases* wage outcomes but *decreases* employment probability, can cause the employment equation and the wage equation to be negatively correlated in their unobservables (i.e., error terms), leading to the negative selection

[28] For details on the initial resettlement support from the US government, see Online Supplement B.1.

[29] In this context, *selection* refers to *selection on unobservables*.

Table 8 Selection parameter estimates

Sample selection check Bivariate selection ML	
ρ_1 of European refugees	-0.214
	[0.202]
ρ_2 of non-European refugees	0.928***
	[0.026]
No. of observations	1478

(i) Robust standard errors in brackets (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively. (iii) The dummy variable of delayed issuance of a social security number is used as an exclusion restriction between the selection equation and the wage equation

into employment. Thus, in order to check the plausibility of this inference, the selection parameter of European refugees (i.e., ρ_1) and that of non-European refugees (i.e., ρ_2) were separately estimated based on a single likelihood function. See Appendix A.2 for estimation-related details.

The result of $\hat{\rho}_1 < 0$ and $\hat{\rho}_2 > 0$, shown in Table 8, corroborates that European refugees and other refugees are distinguishable in terms of their contrasting selection patterns into employment. The finding that ρ_1 of European refugees is estimated to be negative, albeit statistically insignificant, bolsters the inference mentioned above and explains why European refugees, in contrast to Asian refugees, benefited from the attacks in terms of employment quality, not employment probability. The statistical insignificance of $\hat{\rho}_1$, which may be attributed to a limited number of European refugees in the sample, motivates future research with more observations. On the other hand, $\hat{\rho}_2 > 0$ with strong statistical significance is expectable in that positive self-selection into employment is considered common and natural in usual labor markets, which means that those with high earnings potential and a higher rate of unobserved capabilities and skills are more likely to take up employment (Smith and Ward 1989; Blau et al. 1990s; Blundell et al. 2007; Olivetti and Petrongolo 2008).[30]

6 Robustness checks

6.1 Unobserved time-varying factors

As was suggested by Kaushal et al. (2007), in order to isolate the effects of the September 11 attacks from seasonal, cyclical factors and time-varying unobservables as much as possible, a group-comparison approach with regression adjustments was additionally used as a strategy for checking the robustness of the estimates of this study. This approach can be regarded as a difference-in-differences (i.e., hereafter abbreviated as DD) procedure with repeated cross-sections. As highlighted in previous sections, the

30 The selection parameter of Asian refugees cannot be solely estimated due to the non-convergence of the maximum likelihood function.

sample observations are not independent but heavily clustered: hence, cluster-specific heterogeneity should still be taken into consideration in using the DD approach.

The empirical strategy required for the DD approach is that the employment probability should become affected for some groups (i.e., treatment group) by the September 11 attacks, while it is expected to be unaffected for other groups (i.e., control group) (Åslundslund and Rooth 2005). Thus, properly defining treatment and control groups is of paramount importance for the DD approach. According to the estimation results shown in Sect. 5, it seems to be a tenable argument that European refugees' employment probability was neither significantly nor substantially affected by the attacks. Therefore, European refugees can be regarded as controlled observations in the context of this study; on the other hand, African, Middle Eastern, and South American refugees can be thought of as (negatively) treated observations.

However, for the reason that we cannot rule out the possibility that *all* refugees, even including European refugees, were exposed to the impacts of the attacks after they had been perpetrated, it should be carefully noted that the DD estimate here between the treated (i.e., African, Middle Eastern, and South American refugees) and the controlled (i.e., European refugees) should be construed as the effect stemming from the combination of *two factors*: the September 11 attacks *and* the ethnic disadvantage of being Africans, Arabs, or South Americans. Accordingly, in the strict sense, European refugees are not free from the first factor but only free from the second factor, and the treatment should be regarded as the concurrent combination of those two factors. Many previous studies with the DD approach interpreted their DD estimates as the effects of the September 11 attacks: however, unless it can be guaranteed that their controlled observations were completely free from any impacts of the attacks, interpreting their DD estimates as the impacts of the attacks as one sole factor is erroneous.

The DD approach requires two underlying assumptions for identification. First, the effects of time should be common across treated and controlled observations (Cameron and Trivedi 2005). This assumption is testable and should be satisfied when either panel or cross-section data are investigated. Thus, the common time trend assumption was thoroughly checked, and the null hypothesis of the parallel time trend before the attacks could not be rejected. Second, this section uses the DD approach with repeated cross-sections: therefore, the composition of the treatment group and that of the control group should be stable before and after the attacks so that unobserved individual features should play no role. As previously mentioned in Sect. 3, this condition is not expected to be severely violated since the observations of this study are not economic migrants but refugees, whose decision function of leaving their home countries is expected to be far less sensitive than that of economic migrants. Based on these two assumptions, the DD model with a time trend ω_t is specified as shown below, where *post* refers to the periods *after* the September 11 attacks.

$$\begin{aligned}&\Pr(y_{jc,t} = 1 \mid \text{Treatment}, \text{Post}, \mathbf{v}_{jc}, \mathbf{i}_c, \omega_t, \alpha_c)\\&= g(\beta_1 \text{Treatment} + \beta_2 \text{Post} + \beta_{\text{DD}}(\text{Treatment} \times \text{Post}) + \mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \omega_t + \alpha_c)\end{aligned} \tag{10}$$

Table 9 DD Estimates of the treatment on employment—Treatment group: African, Middle Eastern, and South American refugees

DV=Employed (1) or not (0) Estimator	DD
β_{DD}	−0.1703*
	[0.0961]
Common trend assumption	*Satisfied*
No. of observations	1379

(i) Cluster-robust standard errors in brackets (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively

Table 10 DD Estimates of the treatment on employment—Treatment group: Asian refugees

DV=Employed (1) or not (0) Estimator	DD
β_{DD}	0.1356
	[0.1299]
Common trend assumption	*Satisfied*
No. of observations	289

(i) Cluster-robust standard errors in brackets (ii) *** implies significance at the 1% significance level, ** at 5%, * at 10%, respectively

While all the other notations are the same as defined earlier, g refers to a certain binary link function, such as $\Lambda[\cdot]$ (i.e., the logistic cumulative distribution function) in the case of the conditional logit fixed effects model and $1[\cdot]$ (i.e., the identity link function) for the linear probability model. While both models support the robustness of the main findings discussed in Sect. 5, this section presents the results based on the latter for the reason that the only direct interpretation of the former is using odds ratios in multiplicative terms, which is less intuitive than additive marginal effects (Kleinman and Norton 2009; Karaca-Mandic et al. 2012).

Table 9 shows the DD estimates between European refugees as controlled observations and African, Middle Eastern, and South American refugees as treated observations. The negative β_{DD} estimate implies the negative effect of the combination of two factors—the September 11 attacks and the ethnic disadvantage of being Africans, Arabs, or South Americans. It suggests that the treatment lowered the employment probability of the treated by 17 % points in additive terms.

On the other hand, Table 10 shows the DD estimates between European refugees as controlled observations and Asian refugees as (positively) treated observations. Again, the DD estimates capture the effect of the combination of two factors— the September 11 attacks and Asian ethnicity. The loss of statistical significance may be attributed to the greatly reduced number of observations. The positive β_{DD} estimate corroborates that the treatment increased the employment probability of the treated by 13% points in additive terms.

Taken together, the DD estimates in Tables 9 and 10 support the fact that the estimates of this study presented in Sect. 5 are robust when seasonal, cyclical factors and time-varying unobservables are taken into account.

6.2 Additional robustness checks

In addition to unobserved time-varying factors tackled in Sect. 6.1, further issues for robustness are a nonlinear projection for cluster heterogeneity in the case of the Chamberlain–Mundlak's correlated random effects probit model (i.e., discussed in Appendix A.1.2) and more aggregated clusters with regard to what to cluster over (i.e., discussed in Sect. 4.2). This study establishes the robustness of its main findings in these dimensions in Online Supplement. See Online Supplement B.3.1 for a nonlinear projection and B.3.2 for more aggregated clusters.

7 Conclusion

This paper investigates the impacts of the September 11 terrorist attacks on refugees as job seekers in the U.S.. To the best of the author's knowledge, this paper is the first study of the attacks' impacts on refugees that qualify for official refugee status and hence are distinguished from economic migrants or general nonnatives. Moreover, this investigation sheds unprecedented light on the heterogeneity of the impacts by ethnic subgroups, remembering that the attacks were planned by the Islamic extremist group and perpetrated by its African and Arab hijackers. Quantifying the heterogeneity of the impacts is a significant addition to previous studies, which simply presumed that the attacks had homogeneously negative impacts.

In terms of econometric methods, this study, the observations of which are not independent but clustered, shows the enormous importance of considering cluster-specific heterogeneity by using nonlinear multi-level models. The multi-level estimates of this study suggest that the September 11 attacks did not uniformly shock all sub-populations of refugees: rather, they presented a unique, substantial opportunity for Asian refugees and a serious threat to African and Arab refugees. One unexpected finding, on the other hand, is that the employment probability of European refugees remained stable, whereas that of Asian refugees markedly increased after the attacks. However, in terms of employment quality, measured by real wages, European refugees were the only ones who benefited from the attacks. Such different patterns of benefits between Asian and European refugees seem to have been caused by their different reservation wages and heterogeneous expectations for future labor market outcomes, as evidenced by their contrasting patterns of selection into employment during the initial resettlement period.

This research provides the first comprehensive assessment of the heterogeneous impacts of the September 11 terrorist attacks on refugees' labor market outcomes and establishes an econometric framework for controlling for cluster-specific unobservables by disentangling several econometric complications that are intertwined, such as non-independent and clustered observations, the binary dependent variable, and the incidental parameters problem. In addition, various robustness checks show the insensitivity of the findings of this research. Hence, this study makes meaningful strides toward enhancing our understanding of the heterogeneous labor market impacts of the September 11 terrorist attacks on refugees as job seekers.

Acknowledgements I thank the editor, the associate editor, and two anonymous referees for their careful reading and helpful suggestions despite the long-lasting pandemic situation. Earlier versions of this article benefited from comments by Horst Entorf, Nicola Fuchs-Schündeln, Cornelia Storz (U. Frankfurt), Jin-Young Choi (Xiamen U.), Jo-Yup Ahn, Yoon-Gyu Yoon, Jae-Min Seong (KLI), Le Van Cuong (U. Paris I), Marie Lalanne (Università degli Studi di Torino), and numerous participants at conferences including the Econometric Society Meetings in 2019. I thank Birgit Herrmann for excellent administrative assistance and gratefully acknowledge various support from the Federal Government of Germany, Frankfurter Wirtschaftswissenschaftliche Gesellschaft, Development and Policies Research Center, and the French National Center for Scientific Research (CNRS).

Funding None

Compliance with ethical standards

Conflict of interest I, as the author of this manuscript, certify that I have no conflict of interest to declare at the time of submission.

Ethical approval This study does not contain any studies with human participants or animals performed by the author.

Data availability The data that support the findings of this study are not publicly available but are available from the International Rescue Committee upon request. Restrictions apply to the availability of these data according to the relevant regulations of the US government.

Code availability The program codes that generate the final results of this study are available from the author upon request.

A Appendix

A.1 Details on multi-level econometric models

A.1.1 Conditional maximum likelihood model

In general, the fixed effects model is considered more conservative in that it allows for the possible correlation between cluster-specific fixed effects α_c and covariates. However, nonlinear maximum likelihood variants of the cluster-specific fixed effects model entail further complications—especially in the case of data with a small number of cluster members N_c. This is because data with small N_c have the problem of too many *incidental parameters* (i.e., $\alpha_1, \ldots, \alpha_C$), while the number of observations for estimating each α_c is not enough (Neyman and Scott 1948). Unlike linear cases, it is generally not possible to eliminate such nuisance parameters in nonlinear cases (Hall and Severini 1998). This is the reason why the present study, the dependent variable of which is binary, cannot simply use the dummy variable fixed effects model. Cluster-specific dummies, in nonlinear cases with the finite number of cluster members N_c, fail to properly pick up α_c and render maximum likelihood estimates inconsistent (Neyman and Scott 1948; Chamberlain 1980; Lancaster 2000; Greene 2002).

One alternative method for obtaining consistent estimates that eliminates unwanted α_c is using *the conditional maximum likelihood model* (CML). It is based on a log density for the jth individual (in the cth cluster) that conditions on $\sum_{j=1}^{N_c} y_{jc}$, which refers to the total number of outcomes equal to one (i.e., employed) for a given cluster (Chamberlain 1980). When a binary logit model with cluster fixed effects α_c specifies

$$\Pr(y_{jc} = 1 \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = \frac{\exp(\mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c)}{1 + \exp(\mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c)}, \tag{11}$$

where all notations are the same as defined in Section 4, the joint conditional probability for the cth cluster is calculated as follows:

$$\begin{aligned} &\Pr{}_c(y_{1c}, y_{2c}, y_{3c}, \ldots, y_{N_c c} \mid N_c\overline{y}_c) \\ &= \frac{\exp(\sum_{j=1}^{N_c} y_{jc}(\mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c))}{\sum_{d \in \widetilde{B}_c} \exp(\sum_{j=1}^{N_c} d_{jc}(\mathbf{v}'_{jc}\boldsymbol{\delta} + \mathbf{i}'_c\boldsymbol{\eta} + \alpha_c))} \\ &= \frac{\exp(\sum_{j=1}^{N_c} y_{jc}\mathbf{v}'_{jc}\boldsymbol{\delta})}{\sum_{d \in \widetilde{B}_c} \exp(\sum_{j=1}^{N_c} d_{jc}\mathbf{v}'_{jc}\boldsymbol{\delta})}, \end{aligned} \tag{12}$$

where $\widetilde{B}_c = \{(d_{1c}, d_{2c}, d_{3c}, \ldots, d_{N_c c}) \mid d_{jc} = 0$ or 1, and $\sum_{j=1}^{N_c} d_{jc} = \sum_{j=1}^{N_c} y_{jc} = N_c\overline{y}_c\}$ (Hamerle and Ronning 1995; Cameron and Trivedi 2009; Hosmer et al. 2013).[31] This approach is uniquely possible only with logit by virtue of its feature that $\exp(\cdot)$ appears both in the numerator and in the denominator, which enables two sets of common factors — $\exp(\sum_{j=1}^{N_c} y_{jc}\mathbf{i}'_c\boldsymbol{\eta}) = \exp(\sum_{j=1}^{N_c} d_{jc}\mathbf{i}'_c\boldsymbol{\eta})$ and $\exp(\sum_{j=1}^{N_c} y_{jc}\alpha_c) = \exp(\sum_{j=1}^{N_c} d_{jc}\alpha_c)$ conditioning on $N_c\overline{y}_c$ — to be canceled out in (12).[32] Finally, the conditional likelihood is $\prod_{c=1}^{C} \Pr_c(y_{1c}, y_{2c}, y_{3c}, \ldots, y_{N_c c} \mid N_c\overline{y}_c)$, the product of (12) over all C clusters, where C refers to the total number of clusters.

Under the between-cluster independence assumption and the conditional independence assumption (CIA) that within-cluster observations $\{y_{1c}, y_{2c}, y_{3c}, \ldots, y_{N_c c}\}$ are independent conditioning on $\mathbf{v}_{jc}$, $\mathbf{i}_c$, and α_c, the conditional maximum likelihood model eliminates α_c and yields consistent coefficient estimates of level-one covariates $\mathbf{v}_{jc}$ (Chamberlain 1980). The primary advantage of the conditional maximum likelihood model resides in the fact that it does not rely on any assumptions concerning the distribution of α_c. This approach is referred to as *the Chamberlain's fixed effects logit model* in econometrics and applicable even when cluster sizes vary.

While eliminating α_c in (12) is uniquely feasible without the incidental parameters problem, this approach, however, still entails some substantial problems, the most critical point of which lies in the fact that it leads to the loss of observations if y_{jc} is either 0 for all j or 1 for all j. On top of this, all clusters with $N_c = 1$ are excluded, further impairing efficiency (Cameron and Trivedi 2005). Moreover, while this method is useful for obtaining consistent coefficient estimates, it is generally not possible to estimate additive marginal effects because they depend on eliminated α_c, unlike linear

[31] In plain terms, $\widetilde{B}_c$ indicates all possible combinations of 0 and 1 given $N_c\overline{y}_c$.

[32] Thus, it is often also referred to as *the conditional logit fixed effects model*. On the other hand, in the case of the multi-level probit model, there is no solution to the incidental parameters problem, which makes it impossible to include fixed effects (Lancaster 2000; Cameron and Trivedi 2005).

fixed effects models (Beck 2020). Some researchers report marginal effects evaluated at a certain value $\alpha_c = q$: however, that is not much meaningful as where to evaluate α_c is a completely arbitrary decision. As an alternative, a common way of interpreting coefficients is using an odds ratio, which measures the probability of $y = 1$ relative to the probability of $y = 0$. However, such multiplicative interpretation is usually less intuitive than additive marginal effects. This is the point at which the more flexible random effects model should also be considered.

A.1.2 Chamberlain–Mundlak's correlated random effects probit model

When the simplest model $y_{jc} = \mathbf{x}'_{jc}\boldsymbol{\beta} + u_{jc}$ is supposed with the error term decomposition $u_{jc} = \alpha_c + \varepsilon_{jc}$, the conventional random effects model assumes

$$\alpha_c \mid \mathbf{x}_{jc} \sim N(0, \sigma_\alpha^2), \tag{13}$$

and this is undeniably a very strong assumption. It should be carefully noted that (13) requires two conditions to be satisfied concurrently: first, α_c and $\mathbf{x}_{jc}$ should not be correlated, and second, α_c should be normally distributed with homoskedasticity (i.e., constant variance σ_α^2).

Due to the implausibility of the assumption (13) in the case of the data under investigation, this study does not use the conventional random effects model; instead, the approach of Chamberlain (1980) is leveraged to clustered observations, which had been originally devised in the context of panel data. Based on the method of Chamberlain (1980), the correlation between α_c and $\mathbf{x}_{jc}$ can be tolerated by replacing α_c with its *linear projection* onto the cluster-specific means of covariates (i.e., $\overline{\mathbf{x}}_c$) including a projection error (Wooldridge 2010). By allowing α_c to be determined by $\overline{\mathbf{x}}_c$, α_c can be expressed as

$$\alpha_c = \psi + \overline{\mathbf{x}}'_c\,\boldsymbol{\xi} + w_c, \tag{14}$$

where w_c denotes the projection error.[33] Then, the Mundlak (1978) version of Chamberlain's assumption can be written as

$$\alpha_c \mid \mathbf{x}_{jc} \sim N(\psi + \overline{\mathbf{x}}'_c\boldsymbol{\xi}, \sigma_w^2). \tag{15}$$

This signifies that α_c can be correlated with regressors through $\overline{\mathbf{x}}_c$: therefore, the inclusion of $\overline{\mathbf{x}}_c$ is expected to control for the correlation between cluster-specific heterogeneous features and covariates. However, this assumption (15), albeit more flexible than (13), is still restrictive in that it specifies the conditional distribution of α_c. In plain terms, (15) means that α_c given $\mathbf{x}_{jc}$ should be normally distributed with mean $\psi + \overline{\mathbf{x}}'_c\boldsymbol{\xi}$ and variance σ_w^2.

While the fundamental logic of Mundlak (1978) is to let $\overline{\mathbf{x}}_c$ in (14) include all regressors, within-cluster invariant regressors do not provide any information for this projection. Hence, regressors are categorized into two types: within-cluster variant

[33] In (14), it is assumed that the projection error w_c has zero mean and is uncorrelated with $\overline{\mathbf{x}}_c$.

regressors $\mathbf{v}_{jc}$ and within-cluster invariant regressors $\mathbf{i}_c$, as previously mentioned with (3) in Sect. 4. Considering this separation, cluster-specific α_c can be rewritten as

$$\alpha_c = \psi + \overline{\mathbf{v}}_c' \boldsymbol{\tau} + w_c, \tag{16}$$

where $\overline{\mathbf{v}}_c$ refers to the averages of level-one covariates $\mathbf{v}_{jc}$ within each cluster. Naturally, cluster-invariant level-two covariates $\mathbf{i}_c$ are still included as exploratory variables. As a result, the generalized linear model along with probit as a binary link function can be defined as

$$\Pr(y_{jc} = 1 \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = E(y_{jc} \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = \Phi(\psi + \mathbf{v}_{jc}' \boldsymbol{\zeta} + \overline{\mathbf{v}}_c' \boldsymbol{\tau} + \mathbf{i}_c' \boldsymbol{\varphi}), \tag{17}$$

where Φ denotes the standard normal cumulative distribution function. This method is called *the Chamberlain–Mundlak's correlated random effects probit model* (CRE). The test of $\boldsymbol{\tau} = 0$ in (17), which is known as the Mundlak test, can be easily implemented in a bid to figure out whether the assumption of the conventional random effects model (13) is valid (i.e., if $\boldsymbol{\tau} = 0$) or not (i.e., if $\boldsymbol{\tau} \neq 0$) (Mundlak 1978). The results of the test applied to this study are discussed in Online Supplement B.2.

The Chamberlain–Mundlak's correlated random effects probit model yields unbiased estimates of level-one covariates (Mundlak 1978; Neuhaus and Kalbfleisch 1998; Snijders and Berkhof 2008). Hence, this study also uses the Chamberlain–Mundlak's correlated random effects probit model along with the conditional maximum likelihood model.

A.1.3 Linear probability model with cluster fixed effects

As previously mentioned in Sect. 4.3, if we use the identity link for g in (7), the simple linear probability model with cluster fixed effects can be defined as shown below.

$$\Pr(y_{jc} = 1 \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = E(y_{jc} \mid \mathbf{v}_{jc}, \mathbf{i}_c, \alpha_c) = \mathbf{v}_{jc}' \boldsymbol{\delta} + \mathbf{i}_c' \boldsymbol{\eta} + \alpha_c \tag{18}$$

Despite its simplicity and computational convenience, the simple linear probability model with cluster fixed effects has some shortcomings as follows. First, the linear probability model, which uses neither a cumulative distribution function nor a latent variable model, has no structural room for error terms. Second, some predicted probabilities based on the linear probability model may have nonsensical values that are less than zero or greater than one (i.e., outside the unit interval). Third, the linear probability model can even lead to negative variances (Greene 2002).

However, on a more positive note, the linear probability model requires no distributional assumptions of disturbances, the violation of which can make maximum likelihood estimates inconsistent (Bera et al. 1984). Hence, this study also uses the linear probability model for comparisons. It is also used when a maximum likelihood function fails to converge because of the large number of parameters to be estimated.

A.2 Bivariate selection model

Conventional selection models are composed of two sequential equations—one for employment (i.e., also often called *selection* or *participation*) and the other for wage outcomes (Amemiya 1985). An employment equation with binary outcomes can be expressed as shown below, where y_d^* is a latent variable that determines whether to work or not.

$$y_{d_i}^{\text{Employment}} = \begin{cases} 1 & \text{if } y_{d_i}^* > 0 \\ 0 & \text{if } y_{d_i}^* \leq 0 \end{cases} \tag{19}$$

While Cameron and Trivedi (2005) explains that y_d^* is construed as the unobserved desire or propensity to work, Heckman (1974), from a labor supply standpoint, notes that it can be regarded as the difference between a refugee's market wage (i.e., wage offer) and his or her reservation wage. On the other hand, a resultant market wage equation with continuous outcomes can be expressed as follows, where latent y_w^* determines how much to work.

$$y_{w_i}^{\text{Wage}} = \begin{cases} y_{w_i}^* & \text{if } \ y_{d_i}^* > 0 \\ \text{Unobserved} & \text{if } \ y_{d_i}^* \leq 0 \end{cases} \tag{20}$$

The market wage equation (20) denotes that the wage outcome of a refugee is observed if and only if a refugee is employed (i.e., $y_{d_i}^{\text{Employment}} = 1$) with $y_{d_i}^* > 0$ in (19). In other words, whether we can observe a refugee's market wage level or not entirely depends on his or her labor supply decision. The canonical approach for modeling selection specifies linear models with additive error terms in the following manner (Cameron and Trivedi 2005).

$$\begin{cases} y_{d_i}^* = \mathbf{x}_{d_i}' \boldsymbol{\beta}_d + \varepsilon_{d_i} & \text{for employment} \\ y_{w_i}^* = \mathbf{x}_{w_i}' \boldsymbol{\beta}_w + \varepsilon_{w_i} & \text{for market wage levels} \end{cases} \tag{21}$$

The correlation between ε_d and ε_w in (21) is the core of sample selection models, which, if overlooked, can cause problems in estimating $\boldsymbol{\beta}_w$ (Greene 2002). In the case of the bivariate sample selection model, suggested by Tobin (1958), estimation by maximum likelihood is straightforward given the additional assumption of

$$\begin{bmatrix} \varepsilon_d \\ \varepsilon_w \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \text{Var}(\varepsilon_d) = \sigma_d^2 & \text{Cov}(\varepsilon_d, \varepsilon_w) \\ \text{Cov}(\varepsilon_d, \varepsilon_w) & \text{Var}(\varepsilon_w) = \sigma_w^2 \end{bmatrix}\right], \tag{22}$$

which, in plain terms, means that the correlated errors are *joint normally distributed* with homoskedasticity (Cameron and Trivedi 2005).[34] Gronau (1974), Heckman (1979), and Keane et al. (1988) are among previous studies using this bivariate normality assumption: for further details on this assumption, see Amemiya (1985) and Moffitt (1999). Based on the assumption (22), the bivariate sample selection model

[34] In (22), a normalization of $\text{Var}(\varepsilon_d) = \sigma_d^2 = 1$ is commonly used (Greene 2002).

maximizes the likelihood function

$$L = \prod_{i=1}^{N} \{\Pr[y^*_{d_i} \leq 0]\}^{1-y_{d_i}} \{\Pr[y^*_{d_i} > 0] \times f(y_{w_i} \mid y^*_{d_i} > 0)\}^{y_{d_i}}, \tag{23}$$

and the use of probit as a link function leads to

$$\begin{aligned} L = &\prod_{i=1}^{N} \{\Phi(-\mathbf{x}'_{d_i}\boldsymbol{\beta}_d)\}^{1-y_{d_i}} \\ &\times [\sigma_w^{-1}\phi\left(\frac{y_{w_i} - \mathbf{x}'_{w_i}\boldsymbol{\beta}_w}{\sigma_w}\right)\Phi\left\{\frac{\mathbf{x}'_{d_i}\boldsymbol{\beta}_d + (y_{w_i} - \mathbf{x}'_{w_i}\boldsymbol{\beta}_w)\rho/\sigma_w}{\sqrt{1-\rho^2}}\right\}]^{y_{d_i}}, \end{aligned} \tag{24}$$

where ρ refers to the correlation coefficient between ε_d and ε_w. As is customary, Φ denotes the standard normal cumulative distribution function, whereas ϕ is the standard normal probability density function. The log of (24) is the objective function of the bivariate sample selection model. However, in this context, the selection parameter of European refugees and that of non-European refugees should be separately estimated, and if they are estimated in a single likelihood function, they can be directly compared. Therefore, (24) is modified as follows.

$$\begin{aligned} L = \prod_{i=1}^{N} &\left\langle \{\Phi(-\mathbf{x}'_{d_i}\boldsymbol{\beta}_d)\}^{1-y_{d_i}} \left[\sigma_1^{-1}\phi\left(\frac{y_{w_i} - \mathbf{x}'_{w_i}\boldsymbol{\beta}_w}{\sigma_1}\right)\Phi\left\{\frac{\mathbf{x}'_{d_i}\boldsymbol{\beta}_d + (y_{w_i} - \mathbf{x}'_{w_i}\boldsymbol{\beta}_w)\rho_1/\sigma_1}{\sqrt{1-\rho_1^2}}\right\}\right]^{y_{d_i}} \right\rangle^{S_i} \\ &\left\langle \{\Phi(-\mathbf{x}'_{d_i}\boldsymbol{\gamma}_d)\}^{1-y_{d_i}} \left[\sigma_2^{-1}\phi\left(\frac{y_{w_i} - \mathbf{x}'_{w_i}\boldsymbol{\gamma}_w}{\sigma_2}\right)\Phi\left\{\frac{\mathbf{x}'_{d_i}\boldsymbol{\gamma}_d + (y_{w_i} - \mathbf{x}'_{w_i}\boldsymbol{\gamma}_w)\rho_2/\sigma_2}{\sqrt{1-\rho_2^2}}\right\}\right]^{y_{d_i}} \right\rangle^{1-S_i} \end{aligned} \tag{25}$$

While all notations are the same as defined hitherto, $S_i = 1$ means that an individual i is a European refugee (i.e., $i \in S$).[35] Thus, ρ_1 refers to the correlation between ε_d and ε_w of European refugees, and σ_1 denotes the standard deviation of European refugees' ε_w. Likewise, ρ_2 refers to the correlation between ε_d and ε_w of non-European refugees, and σ_2 denotes the standard deviation of non-European refugees' ε_w. The likelihood function (25) is flexible in the sense that the coefficients of $\mathbf{x}_d$ and $\mathbf{x}_w$ are allowed to be different in those two groups (i.e., $\boldsymbol{\beta}_d, \boldsymbol{\beta}_w, \boldsymbol{\gamma}_d, \boldsymbol{\gamma}_w$).[36]

References

Åslundslund O, Rooth DO (2005) Shifts in attitudes and labor market discrimination: Swedish experiences after 9–11. J Popul Econ 18(4):603–629

35 The modification in (25) is based on the key feature of maximum likelihood estimation that an overall (i.e., joint) likelihood function is the product of individual (i.e., marginal) probability density functions (Greene 2002; Baltagi 2011). In this context, individuals are categorized into two groups: $S = 1$ and $S = 0$.

36 $\boldsymbol{\gamma}_d$ and $\boldsymbol{\gamma}_w$ correspond to $\boldsymbol{\beta}_d$ and $\boldsymbol{\beta}_w$, respectively.

Abadie A, Athey S, Imbens GW, Wooldridge J (2017) When should you adjust standard errors for clustering? Working Paper No. 24003, National Bureau of Economic Research

Amemiya T (1985) Advanced econometrics. Harvard University Press, Cambridge

Baltagi BH (2011) Econometrics. Springer, Berlin

Beaman LA (2012) Social networks and the dynamics of labour market outcomes: evidence from refugees resettled in the U.S. The Rev Econ Stud 79(1):128–161

Beck N (2020) Estimating grouped data models with a binary-dependent variable and fixed effects via a logit versus a linear probability model: The impact of dropped units. Polit Anal 28(1):139–145

Bera AK, Jarque CM, Lee LF (1984) Testing the normality assumption in limited dependent variable models. Int Econ Rev 25(3):563–578

Blau FD, Kahn LM (1990s) The U. S. gender pay gap in the 1990s: slowing convergence. Ind Labor Relat Rev 60(1):45–66

Blundell R, MaCurdy T (1998) Labor supply. In: Handbook of labor economics, vol 3, Elsevier, pp 1559–1695

Blundell R, Gosling A, Ichimura H, Meghir C (2007) Changes in the distribution of male and female wages accounting for employment composition using bounds. Econometrica 75(2):323–363

Bonanno GA, Jost JT (2006) Conservative shift among high-exposure survivors of the september 11th terrorist attacks. Basic Appl Soc Psychol 28(4):311–323

Borjas GJ (1987) Self-selection and the earnings of immigrants. Am Econ Rev 77(4):531–553

Braakmann N (2009) The impact of september 11th, 2001 on the employment prospects of Arabs and Muslims in the German labor market. J Econom Stat 229(1):2–21

Browne WJ, Draper D (2006) A comparison of Bayesian and likelihood-based methods for fitting multilevel models. Bayesian Anal 1(3):473–514

Brumback BA, Dailey AB, Brumback LC, Livingston MD, He Z (2010) Adjusting for confounding by cluster using generalized linear mixed models. Stat Probab Lett 80(21):1650–1654

Burawoy M (1976) The functions and reproduction of migrant labor: comparative material from southern Africa and the united states. Am J Soc 81(5):1050–1087

Cameron AC, Miller DL (2015) A practitioner's guide to cluster-robust inference. J Hum Resour 50(2):317–372

Cameron AC, Trivedi PK (2005) Microeconometrics. Cambridge Press, Cambridge

Cameron AC, Trivedi PK (2009) Microeconometrics using Stata. Stata Press, College Station

Ceritoglu E, Yunculer HBG, Torun H, Tumen S (2017) The impact of Syrian refugees on natives' labor market outcomes in turkey: evidence from a quasi-experimental design. IZA J Labor Policy 6(1):5

Chamberlain G (1980) Analysis of covariance with qualitative data. Rev Econ Stud 47(1):225–238

Dagnelie O, Mayda AM, Maystadt JF (2019) The labor market integration of refugees to the united states: do entrepreneurs in the network help? Eur Econ Rev 111:257–272

Damm AP (2009a) Determinants of recent immigrants' location choices: quasi-experimental evidence. J Popul Econ 22(1):145–174

Damm AP (2009b) Ethnic enclaves and immigrant labor market outcomes: quasi-experimental evidence. J Labor Econ 27(2):281–314

Damm AP (2014) Neighborhood quality and labor market outcomes: evidence from quasi-random neighborhood assignment of immigrants. J Urban Econ 79:139–166

D'Amuri F, Ottaviano GI, Peri G (2010) The labor market impact of immigration in Western Germany in the 1990's. Eur Econ Rev 54(4):550–570

Dustman C, Vasiljeva K, Damm AP (2019) Refugee migration and electoral outcomes. Rev Econ Stud 86(5):2035–2091

Dávila A, Mora MT (2005) Changes in the earnings of Arab men in the US between 2000 and 2002. J Popul Econ 18(4):587–601

Edin PA, Fredriksson P, Åslundslund O (2003) Ethnic enclaves and the economic success of immigrants: evidence from a natural experiment. Q J Econ 118(1):329–357

Evans MDR, Kelley J (1991) Prejudice, discrimination, and the labor market: attainments of immigrants in Australia. Am J Sociol 97(3):721–759

Foged M, Peri G (2016) Immigrants' effect on native workers: new analysis on longitudinal data†. Am Econ J Appl Econ 8(2):1–34

Ford CA, Udry J, Gleiter K (2001) Chantala K (2003) Reactions of young adults to september 11. J Adolesc Health 32(2):124–125

Friedberg RM, Hunt J (1995) The impact of immigrants on host country wages, employment and growth. J Econ Perspect 9(2):23–44
Greene WH (2002) Econometric analysis. Pearson Education, London
Greene WH (2007) Discrete choice modeling. In: Palgrave handbook of econometrics, Springer, pp 473–556
Gronau R (1974) Wage comparisons —a selectivity bias. J Polit Econ 82(6):1119–1143
Guo G, Zhao H (2000) Multilevel modeling for binary data. Rev Sociol 26(1):441–462
Hall DB, Severini TA (1998) Extended generalized estimating equations for clustered data. J Am Stat Assoc 93(444):1365–1375
Hamerle A, Ronning G (1995) Panel analysis for qualitative variables pp 401–451
Hausman JA (1978) Specification tests in econometrics. Econometrica 46(6):1251–1272
Heckman JJ (1974) Shadow prices, market wages, and labor supply. Econometrica pp 679–694
Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47(1):153–161
Hersh ED (2013) Long-term effect of september 11 on the political behavior of victims' families and neighbors. Proc Nat Acad Sci USA 110(52):20959–20963
Hosmer DW Jr, Lemeshow S, Sturdivant RX (2013) Applied logistic regression, vol 398. Wiley, London
Huddy L, Feldman S (2011) Americans respond politically to 9/11: understanding the impact of the terrorist attacks and their aftermath. Am Psychol 66(6):455–467
Huddy L, Feldman S, Taber C, Lahav G (2005) Threat, anxiety, and support of antiterrorism policies. Am J Polit Sci 49(3):593–608
Jost JT, Glaser J, Kruglanski AW, Sulloway FJ (2003) Political conservatism as motivated social cognition. Psychol Bull 129(3):339–375
Karaca-Mandic P, Norton EC, Dowd B (2012) Interaction terms in nonlinear models. Health Serv Res 47:255–274
Kaushal N, Kaestner R, Reimers CW (2007) Labor market effects of September 11th on Arab and Muslim residents of the United States. J Hum Resour 42(2):275–308
Keane M, Moffitt R, Runkle D (1988) Real wages over the business cycle: estimating the impact of heterogeneity with micro data. J Polit Econ 96(6):1232–1266
Killingsworth MR (1983) Labor supply. Cambridge Press, Cambridge
Kleinman LC, Norton EC (2009) What's the risk? a simple approach for estimating adjusted risk measures from nonlinear models including logistic regression. Health Serv Res 44(1):288–302
Lancaster T (2000) The incidental parameter problem since 1948. J Econom 95(2):391–413
Manacorda M, Manning A, Wadsworth J (2012) The impact of immigration on the structure of wages: theory and evidence from Britain. J Eur Econ Assoc 10(1):120–151
Moffitt RA (1999) New developments in econometric methods for labor market analysis. In: Handbook of labor economics 3:1367–1397
Moraga JFH (2011) New evidence on emigrant selection. Rev Econ Stat 93(1):72–96
Mroz TA (1987) The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. Econometrica 55(4):765–799
Mulligan CB, Rubinstein Y (2008) Selection, investment, and women's relative wages over time. Q J Econ 123(3):1061–1110
Mundlak Y (1978) On the pooling of time series and cross section data. Econometrica 46(1):69–85
Munshi K (2003) Networks in the modern economy: Mexican migrants in the U. S. labor market. Qu J Econ 118(2):549–599
Neal D (2004) The measured black–white wage gap among women is too small. J Polit Econ 112(S1):S1–S28
Neuhaus JM, Kalbfleisch J (1998) Between- and within-cluster covariate effects in the analysis of clustered data. Biometrics 54(2):638–645
Neuhaus JM, McCulloch CE (2006) Separating between- and within-cluster covariate effects by using conditional and partitioning methods. J R Stat Soc B 68(5):859–872
Neyman J, Scott EL (1948) Consistent estimates based on partially consistent observations. Econometrica 16(1):1
Olivetti C, Petrongolo B (2008) Unequal pay or unequal employment?: a cross-country analysis of gender gaps. J Labor Econ 26(4):621–654
Ottaviano GIP, Peri G (2012) Rethinking the effects of immigration on wages. J Eur Econ Assoc 10(1):152–197
Pendergast JF, Gange SJ, Newton MA, Lindstrom MJ, Palta M, Fisher MR (1996) A survey of methods for analyzing clustered binary response data. Int Stat Rev 64(1):89–118

Pepper JV (2002) Robust inferences from random clustered samples: an application using data from the panel study of income dynamics. Econ Lett 75(3):341–345
Peri G (2007) Immigrants’ complementarities and native wages: Evidence from california. National Bureau of Economic Research
Peri G (2011) Rethinking the area approach: immigrants and the labor market in California. J Int Econ 84(1):1–14
Peri G (2014) Do immigrant workers depress the wages of native workers? The IZA World of Labor pp 1–42
Piore MJ (1980) Birds of passage. Cambridge University Press, Cambridge
Rabby F, Rodgers WM (2011) Post 9–11 U.S. Muslim labor market outcomes. Atl Econ J 39(3):273–289
Rodriguez G, Goldman N (2001) Improved estimation procedures for multilevel models with binary response. J R Stat Soc Ser A 164(2):339–355
Rooth DO, Saarela J (2007) Selection in migration and return-migration: evidence from micro data. Econ Lett 94(1):90–95
Ruist J (2013) Immigrant-native wage gaps in time series: complementarities or composition effects? Econ Lett 119(2):154–156
Schüller S (2016) The effects of 9/11 on attitudes toward immigration and the moderating role of education. Kyklos 69(4):604–632
Sheridan LP (2006) Islamophobia pre- and post-september 11th, 2001. J Interpers Violence 21(3):317–336
Singh A (2002) We are Not the Enemy: Hate Crimes Against Arabs, Muslims, and Those Perceived to be Arab Or Muslim After September 11, vol 14. Human Rights Watch
Smith JP, Ward M (1989) Women in the labor market and in the family. J Econ Perspect 3(1):9–23
Snijders TAB, Berkhof J (2008) Diagnostic checks for multilevel models. Handbook of Multilevel Analysis pp 141–175
Tobin J (1958) Estimation of relationships for limited dependent variables. Econometrica 26(1):24
Topa G (2001) Social interactions, local spillovers and unemployment. Rev Econ Stud 68(2):261–295
Widner D, Chicoine S (2011) It’s all in the name: employment discrimination against Arab Americans. Sociol Forum 26(4):806–823
Wooldridge JM (2003) Cluster-sample methods in applied econometrics. Am Econ Rev 93(2):133–138
Wooldridge JM (2010) Econometric analysis of cross section and panel data. MIT Press, Cambridge