HIAS-E-125

# Cambodian Refugees

Katsuo Kogure[(a)], Masahiro Kubo[(b)]

*(a) University of Aizu*
*(b) Brown University*

November 2022 (revised October 28, 2023)

Hitotsubashi Institute for Advanced Study, Hitotsubashi University
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan
tel:+81 42 580 8668 http://hias.hit-u.ac.jp/

HIAS discussion papers can be downloaded without charge from: https://hdl.handle.net/10086/27202
https://ideas.repec.org/s/hit/hiasdp.html

# Cambodian Refugees *

Katsuo Kogure[†] Masahiro Kubo[‡]
University of Aizu Brown University

October 28, 2023

**Abstract**

We examine the consequences of forced displacement for Cambodian refugees during the Cambodian conflict (1978–1991). Taking into account the political economy of the Indochina refugee crisis under the Cold War, our analyses reveal that returnees from Thailand attained higher levels of education—while those from Vietnam, by contrast, attained lower levels of education—than stayers. On the other hand, the two groups both experienced worse labor market outcomes, with employment shifts from the primary sector to the immature tertiary sector. Adverse displacement impacts can be attributed to congested labor markets resulting from limited access to available agricultural land.

JEL Codes: O15, J24, D74, N35
Keywords: conflict, forced displacement, refugees, repatriation, Cambodia

*We are grateful to Dany Bahar, Sascha O. Becker, Felipe Valencia Caicedo, Keisuke Hirano, Yoko Kijima, Satoru Kobayashi, Yuya Kudo, Takashi Kurosaki, Hisaki Kono, Tomohiro Machikita, Yasuyuki Sawada, Stelios Michalopoulos, Louis Putterman, Akira Shibanuma, Yoshito Takasaki, Kiyoyasu Tanaka, and participants at various conferences/seminars, including the 2nd Hitotsubashi Summer Institute 2016, PacDev 2019, the Japanese Economic Association Spring Meeting 2019, and the North American/Australian/Asian Econometric Society Summer Meeting 2021, for helpful comments on earlier versions of this work. We also gratefully acknowledge H. E. San Sy Than, H. E. Hang Lina, and Fumihiko Nishi for providing us with the 1998 Census microdata. We also thank the Cambodian Mine Action and Victim Assistance Authority (CMAA) for sharing landmine and UXO contamination and clearance data. This research is supported by Grants-in-Aid for Scientific Research No. 15K17044 and No. 20K01610, the Japan Society for the Promotion of Science, and the Refugee Studies Junior Scholar Awards, Shinnyo-en, Shinnyo Educational Trust.

[†]School of Computer Science and Engineering, University of Aizu, Tsuruga, Ikki-machi, Aizuwakamatsu, Fukushima 965-8580, Japan. Tel: +81-242-37-2500. Email: katsukogure@gmail.com.

[‡]Department of Economics, Brown University, 64 Waterman Street, Providence, RI 02912. E-mail: masahiro_kubo@brown.edu.

# 1 Introduction

Recent years have seen growing interest in the consequences of forced displacement due to wars, civil conflicts, or persecution (see Becker and Ferrara 2019, Becker 2022, Ruiz and Vargas-Silva 2013, Verme and Schuettler 2021 for reviews). One historical turning point for responses to forced displacement is the Cold War between the West and the East (UNHCR 2000).[1] Despite its significance for the history of forced displacement, due to a lack of large-scale empirical data, we know little about how forced displacement during the Cold War affected refugees and how those refugees have been reintegrated since their repatriation in the context of developing countries. In this paper, we employ the complete count Census microdata, which can explicitly identify returning refugees, to help address these issues.

Our study considers the consequences of the Indochina refugee crisis (1975–1995), one of the largest population shifts in history due to the Cold War and the rise of communism in the region (Robinson 1998).[2] Our focus is the two major groups of Cambodian returnees from the neighboring countries of Thailand and Vietnam, who experienced distinct repatriation processes in terms of timing and humanitarian assistance. The former stayed in refugee camps with humanitarian assistance until early 1990s, while the latter went without it until around early 1980s. The two groups were under the control of different great powers, the West and the East, respectively, during the 1980s, when the Cold War intensified.[3] With political economy considerations, our study provides unique comparative research, albeit focusing on a single country.

[1] While proxy wars caused massive movements of refugees, particularly in Africa, Asia, and Central America, the era also saw the expansion of humanitarian assistance in camps; moreover, the United Nations organized and implemented repatriation and peacebuilding operations around the end of the war. In 1975–1980, UNHCR's budget increased from US$ 76 million to US$ 510 million (UNHCR 2000, p. 7). In 1988–1994, the UN conducted 21 repatriation and peacebuilding operations, despite there having been only 13 such operations in the previous 40 years (UNHCR 2000, p.133).

[2] More than 3 million out of 56 million people in the region became refugees (UNHCR 2000, p. 79).

[3] "[R]efugees were used as pawns in geopolitical games to destabilize regimes and to encourage insurgency in their countries of origin" (UNHCR 2000, p. 7).

Our study has policy relevance for contemporary refugee assistance in developing countries. Although it is now common to construct and manage refugee camps with a wide range of services (e.g., vocational training programs, education programs), as well as to assist with repatriation and the rebuilding of refugees' lives after their displacement, these activities were initiated during the Indochina refugee crisis: Historically, assistance to the Indochinese refugees, especially Cambodian refugees, became a turning point for UNHCR (UNHCR 2000, pp. 7–8).[4] Since Cambodian refugees in Thailand (West) received such assistance and those in Vietnam (East) did not, our results offer implications for the effectiveness of refugee assistance.

Our analysis employs the complete count 1998 Cambodian Population Census microdata, which contain basic information regarding individual and household socioeconomic characteristics. Combining migration information with historical facts, we identify well-defined former refugees aged 20–60 who returned from Thailand and Vietnam and then select other individuals, ones who have never migrated before (hereafter, "stayers"), for their comparison samples.[5] For those aged 34–60, we also look at their children aged 6–19 (born after the event), if any, to explore the intergenerational impacts of forced displacement. Our examination of such older and younger returnees from Thailand and Vietnam enables us to grasp the whole picture of the consequences of the Cambodian refugees. Our evaluation focuses mainly on educational and labor market outcomes.

Our analysis faces the potential endogeneity concerns due to the non-random assignment of refugees and stayers in our observational data. We have some anecdotal evidence regarding who fled and why they fled (e.g., Reynell 1989, Greve 1987). First, people wanted to escape the oppressive Pol Pot regime. Second, the Pol Pot regime

[4] UNHCR was established in 1950, the onset of the Cold War. In its early days, UNHCR focused on facilitating the resettlement of refugees fleeing communist regimes in Europe. In the 1960s and the early 1970s, UNHCR became involved in facilitating repatriations in Africa and Bangladesh, but its involvement ended soon afterward (UNHCR 2000, pp. 1–8). During the Indochina refugee crisis of 1975–1980, UNHCR expanded the scope of its activities thanks to a budget increase from US$ 76 million to US$ 510 million (UNHCR 2000, p. 7).

[5] We look at those cohorts because they were born prior to the refugee crisis.

devastated people's lives, families, and property rights, weakening their ties to their homeland. Third, pervasive fear persisted under the Pol Pot regime and intensified following the 1978–1979 Vietnamese invasion. Fourth, the conflict between the Vietnamese army and the Khmer Rouge (officially the Communist Party of Kampuchea) forced people to flee. In our context, since nearly everyone was forcibly engaged in agricultural work under primitive conditions, with the abolition of all private property through collectivization, factors such as income and ownership of assets (e.g., housing, land, livestock) had no bearing on the migration decisions of refugees.

Taking into account these historical contexts, we first carefully use a selection-on-observables assumption to identify the impacts of forced displacement on refugees. This includes the consideration of the matched samples of refugees and stayers based on the Imbens-Rubin approach (Imbens and Rubin 2015), generally exactly balancing the observed covariates. To address potential concerns about selection-on-unobservables, for the limited matched samples, we construct more balanced samples, further balancing pre-treatment covariates. To address the remaining concerns, we also consider omitted variable bias due to unobserved confounders through Rosenbaum's sensitivity analysis (Rosenbaum 2002). For the full samples, we employ the machine learning-based instrumental variables method, following Windmeijer et al. (2019), and also conduct sensitivity analysis, following Oster (2019). These extensive analyses confirm the robustness of our base results.

Our analysis, on one hand, reveals a sharp contrast between the returnees from Thailand, who received education in camps, and those from Vietnam, who did so in Cambodia right after repatriation: The former attained higher levels of education—and the latter, by contrast, attained lower levels of education—than the stayers. This sharp contrast is likely to be due to the educational environment being relatively better in the Thai/Cambodian border refugee camps (West) than in Cambodia (East) (see, e.g., Rogge 1990, p. 47). The latter returnees' human capital accumulation might have been disrupted due to repatriation, because it generally takes a substan-

tial amount of time to readjust to a new environment. Our analysis, on the other hand, also reveals that both groups experienced worse labor market outcomes, with employment shifts from the primary sector to an immature tertiary sector. Such adverse displacement impacts are relatively stronger for later returnees.

These findings imply that higher education levels do not necessarily lead to better labor market outcomes after repatriation. Because the primary sector was a major industry and other sectors were not developed at the time in Cambodia (Vickery 1999), returnees were likely forced to engage in low-skilled work in other sectors (especially the tertiary, i.e., informal, sector), being pushed out of agricultural labor markets due to limited access to agricultural land—something exacerbated by the high contamination of landmines and UXOs. In other words, other sectors might not have matured enough to adequately absorb the available labor force of the returnees at the time. Thus, the adverse displacement impacts can be attributed to congested labor markets resulting from limited access to agricultural land. We provide strong suggestive evidence to support this mechanism, additionally using the large-scale nationally representative 2004 Cambodia Socio-Economic Survey (CSES) microdata and nationwide geospatial data about areas contaminated with landmines and UXOs during the Cambodian conflict.

The rest of this paper is organized as follows. Section 2 clarifies our contributions. Section 3 provides relevant historical background. Section 4 describes our research design and methods. Section 5 reports the estimation results, including the main results and the heterogeneity of the results. Section 6 explores the mechanisms underlying the results. Finally, Section 7 concludes.

## 2 Related Literature

Our study contributes to a recent growing literature on forced migration (see Becker and Ferrara 2019, Becker 2022, Ruiz and Vargas-Silva 2013, Verme and Schuettler 2021 for reviews). In particular, our study has three distinctive features as compared

to the existing literature. First, using the complete count Census microdata, our study extensively explores the heterogeneous impacts of forced displacement on returnees from two different countries (Thailand and Vietnam), as well as among their various subpopulations, taking into account political and social contexts. Becker and Ferrara (2019, p. 15) argue that "[m]any studies looking at mass expulsions just treat forced migrants as coming from a macro region or country without regard for heterogeneity with respect to urban or rural origin, or different geographic or political conditions within region/country of origin that can yield additional insights into the difficulty or ease of forced migrants to assimilate at their destination." Our results highlight the importance of considering potential heterogeneity in the impacts of forced displacement among migrants facing different political and social constraints.

Second, our study considers long-term displacement impacts in the context of developing countries.[6] In terms of long-term displacement impacts, the relevant existing works study forced migration situations in Europe or the United States (i.e., developed countries) (e.g., Arellano-Bover 2022, Bauer et al. 2013, Becker et al. 2020, Deryugina et al. 2018, Nakamura et al. 2022, Sarvimäki et al. 2022; also see Ruiz and Vargas-Silva 2013 for relevant discussion), and generally find positive displacement impacts on labor market outcomes, mainly due to the re-optimization of job and location choices through forced displacement (especially for those who are or may become engaged in the primary sector).[7] In contrast, our study of Cambodia (i.e., a develop-

[6] Extant works focusing on developing countries examine the short- or mid-term impacts of forced displacement (see Ruiz and Vargas-Silva 2013 for a review). The literature has considered internally displaced persons (IDPs) in Northern Uganda (e.g., Fiala 2015), IDPs in Colombia (e.g., Ibáñez and Vélez 2008, Ibáñez and Moya 2010), Burundian refugees (e.g., Fransen et al. 2017, Fransen et al. 2018), Rwandan refugees (e.g., Kondylis 2008), Bosnian refugees and IDPs (Kondylis 2010), and Mozambican refugees and IDPs (Chiovelli et al. 2021). Kondylis (2008) and Fransen et al. (2017) look at returnees, but do not distinguish between those from different destination countries. Fransen et al. (2018) study Burundian returnees from Tanzania who received education in camps and provide suggestive evidence that the forced displacement led to improved educational outcomes. A recent work, Chiovelli et al. (2021), examines the impacts of multiple forced displacement trajectories on human capital investments and occupational choices, focusing mainly on different types of IDPs in Mozambique, and find indications that conflict-driven human capital accumulation may spur structural transformation. They also look at long-term displacement impacts, using additional survey data collected in a major city that received a large number of IDPs.

[7] A change in preferences for education also leads to better labor market outcomes (Becker et al.

ing country) finds negative displacement impacts, although forced migrants may be able to re-optimize their job and location choices. The distinct results might partly be attributed to the systematic differences that forced migrants face in labor market structure and conditions (the extent of labor market distortions/frictions) between developed and developing countries: Job opportunities may be more limited in the latter than the former when forced migrants leave the primary sector. Other sectors in the latter countries may not always be mature enough to adequately absorb the available labor force.[8]

Lastly, to our knowledge, our study conducts the first econometric attempt to evaluate the historical legacies of the Indochina refugee crisis during the Cold War (West vs. East), which became a historical turning point in the expansion of refugee protection.[9] Blattman and Miguel (2010, p. 42) argue that "[t]he social and institutional legacies of conflict are arguably the most important but least understood of all war impacts." Our results provide key lessons regarding the effectiveness of and constraints on humanitarian assistance for refugees: Improved education in camps does not necessarily lead to better labor market outcomes after repatriation, and securing access to agricultural land might be a key consideration in agrarian countries.

2020).

[8] For example, in post-war Finland (1950), the proportion of people (except for forced migrants) engaged in "agriculture," "manufacture etc.," "construction," and "service etc." is respectively as follows: 0.38, 0.26, 0.09, and 0.27 (Sarvimäki et al. 2009), while in post-war Germany (1971), the proportion of male (female) natives (controls for first generation of migrants) engaged in "agriculture," "industry," and "services" is respectively as follows: 0.094 (0.199), 0.506 (0.274), and 0.398 (0.526) (Bauer et al. 2013). In contrast, in Cambodia (1998), the proportion of male (female) stayers aged 20–60 engaged in the primary, secondary, and tertiary sectors in the Full Samples (defined below) is respectively as follows: 0.839 (0.804), 0.013 (0.010), and 0.079 (0.041).

[9] There is a small but growing body of literature relevant to wars/conflicts/violence in Vietnam, Cambodia, and Laos (e.g., Miguel and Roland 2011, Kocher et al. 2011, Dell and Querubin 2018, Riaño and Valencia Caicedo 2020, de Walque 2006, Merrouche 2011, Islam et al. 2016, Lin 2022, Iwanowsky and Madestam 2019, Kogure and Takasaki 2016, Takasaki 2020); the latter seven study Cambodian contexts. Also, Cortes (2004) examines the difference in labor market outcomes between refugee and economic immigrants in the United States, the former of whom include refugees from Vietnam, Cambodia, and Laos.

## 3 Historical Background

This section describes relevant forced displacement situations in Cambodia during the Indochina refugee crisis (1975–1995). Appendix Section A.1 supplementarily provides political contexts during these periods.

**Forced Displacement.** The rise of communism in Indochina during the Cold War caused massive movements of refugees in Vietnam, Cambodia, and Laos (see, e.g., Robinson 1998). In Cambodia, three events caused large numbers of refugees. The first was the rule of the Khmer Rouge, led by Pol Pot, in 1975–1979.[10] Under threat of death due to persecution, around 170,000 and 34,000 people fled to Vietnam and Thailand, respectively, in 1975–1978 (UNHCR 2000, p. 92).

The second event was the collapse of the Pol Pot regime. Following the late 1978 Vietnamese invasion, around 138,000 people fled to Thailand before the end of 1979 (Rogge 1990, p. 31, Suenobu 1995, pp. 5–6). Late that year, in response to international pressure, Thailand adopted an "open door" policy toward refugees. With the large ensuing influx of refugees, however, the policy promptly changed to a "closed door" one in early 1980 (Rogge 1990, pp. 67–69, Suenobu 1995, pp. 9–10). Consequently, those who fled to the border regions after early 1980 could not cross into Thailand and instead had to stay in the border camps (Rogge 1990, p. 69, Suenobu 1995, p. 10).

The third event was the 1984–1985 Vietnamese dry-season offensive, which caused about 220,000 people in the border camps to finally cross the border (Rogge 1990, p. 49). We use the term "refugees" for forced migrants in the Thai/Cambodian border camps throughout this paper (as with other documents/reports), even though Thailand did not accede to the 1951 UN Refugee Convention and regarded all "refugees" as "illegal immigrants" from a legal and political perspective (see, e.g., Robinson 1994, p. 69).[11] Forced migrants in the UNHCR camps were granted de facto refugee

[10] Approximately two million people died from execution, disease, starvation, or exhaustion under the Pol Pot regime (Dy 2007, p. 69).

[11] The 1951 UN Refugee Convention defines a refugee as "someone who is unable or unwilling to

status and became eligible for third country resettlement.

**Refugee Camps.** Refugees in Vietnam stayed in camps under primitive conditions (Wiskari 1978). In contrast, refugees in the Thai/Cambodian border regions stayed in camps under relatively better living conditions (Suenobu 1995, Rogge 1990). In the Thai/Cambodian border regions, there were two types of refugee camps: UNHCR camps and border camps. The UNHCR camps were assisted and administered by UNHCR, whereas the border camps were assisted by the United Nations Border Relief Operation (UNBRO)[12] and administered by Anti-Vietnam political factions, including the Khmer Rouge. Basically, those who fled to Thailand before the end of 1979 stayed in the UNHCR camps, whereas those who fled after 1980 stayed in the border (UNBRO) camps (Suenobu 1995, p. 11). While both types of camps provided essential services, including food, water, shelter, health care, primary and secondary education, and vocational training, the UNHCR camps provided more elaborate services (Rogge 1990, p. 38, UNHCR 2000, p. 93). [13] This is partly because the politically affiliated border (UNBRO) camps suffered frequent attacks from Vietnamese troops into the mid 1980s, and services were temporarily stopped (Suenobu 1995, p. 3). The refugees in the camps ("illegal immigrants") were generally not allowed to leave.

**Repatriation.** Refugees in Vietnam after the first event mostly returned to Cambodia in 1979–1980, right after the collapse of the Pol Pot regime (Rogge 1990, pp. 92–93). The Kampuchean Red Cross was responsible for providing resettlement kits, food assistance (50 kg per family) and monitoring services (Robinson 1994, p. 6). Refugees in Thailand, mainly due to the second and third events, mostly returned to Cambodia in 1992–1993 through a large-scale repatriation program organized by

return to their country of origin owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion."

[12] The UNBRO was established in 1982; the United Nations Children's Fund (UNICEF) and the International Committee of the Red Cross (ICRC) provided services in the border camps from 1979, along with food supplies from the World Food Programme (WFP), until then.

[13] Khao I Dang, the only UNHCR camp after mid-1982, has often been called the most elaborately serviced refugee camp in the world during this period (e.g., Rogge 1990, p. 38).

UNHCR following the 1991 Paris Peace Agreement; those living in Thailand due to the first event mostly resettled to third countries (Rogge 1990, p. 33).

For refugee families in Thailand, UNHCR initially promised two hectares of farmland at the time of repatriation. However, UNHCR later found it impossible to fulfil this promise due to the limited availability of arable land; about 12% of Cambodian villages were highly contaminated by landmines and UXOs (Ministry of Planning 2003, p. 73). UNHCR then prepared several repatriation options, including agricultural land (but not necessarily in their area of choice) and cash ($50 per adult and $25 per child under age 12); both options also included a household/agricultural kit[14] and World Food Programme (WFP) food assistance (400 days) (see Robinson 1994, pp. 23–24 for details). Due to the uncertainty about the assignment of land areas, about 85% of the returnees chose the cash option (UNHCR 2000, p. 147). After repatriation, UNHCR also provided community development assistance and "quick impact projects," including the construction or rehabilitation of roads, schools, and health facilities (UNHCR 2000, p. 146).

# 4 Empirical Design

## 4.1 Data

Our analysis employs two main data sets: the complete count of the 1998 Cambodian Population Census microdata and the 2004 Cambodia Socio-Economic Survey microdata. The census data contain basic information regarding individual and household socioeconomic characteristics (e.g., age, education, occupation, home ownership, place of birth, migration). The information is available for all individuals and households, though three districts and one village are omitted for security reasons.[15] The 2004 Cambodia Socio-Economic Survey (CSES) data, the first large-scale, nationally representative survey data from the 1998 Census frame, contain more detailed information

[14] The household/agricultural kit included water buckets, mosquito nets, various hand tools, and a blue plastic sheet (Robinson 1994, p. 23).

[15] The estimated population in these areas is about 45,000 (National Institute of Statistics 2002); it may include returnees from Thailand.

regarding individual and household socioeconomic characteristics (e.g., land ownership, social capital, health, remittances), though its sample size is relatively small compared to that of the 1998 Census. We complementarily use the 2004 CSES data to explore the potential mechanisms underlying our main results, as well as further examine the long-term impacts of forced displacement on socioeconomic outcomes.

### 4.2 Samples

We construct two analysis samples, *Full Samples* and *Matched Samples*, from the complete count census data.

**Full Samples.** Using the 1998 Census data, Figure 1 provides the distribution of individuals aged 20–60 who previously resided in Thailand (panel A) and Vietnam (panel B) and migrated to their current residence in 1975–1998.[16] The figure clearly shows that migration from Thailand and Vietnam surges in 1992–1993 and 1979–1980, respectively, i.e., with exactly the same timing as their respective repatriations (mentioned above). We define the individuals aged 20–60 who migrated from Thailand and Vietnam in 1979–1998 as returning refugees from Thailand and Vietnam, respectively, and select those aged 20–60 who never migrated (stayers) to serve as comparison samples.

In selecting these samples, we impose certain conditions to reduce unobserved factors affecting outcomes (e.g., rural-born people, "Khmer" for mother tongue, "Buddhism" for religion) and deal with potential concerns about the infeasibility of finding appropriate control groups among stayers (see Appendix Table A1 for the complete sampling procedure). The resulting samples (*Full Samples*) consist of 36,760 returnees from Thailand, 15,548 returnees from Vietnam, and 1,968,687 stayers, with the same ethnicity, language, and religion.[17] Figure 2 provides the geographical distribution of the returnees from Thailand (panel A) and Vietnam (panel B) by district of birth. One limitation of our refugee samples is that they are limited to those who never

[16] We look at individuals aged 20–60 because they were born prior to the refugee crisis and their birth was not potentially affected by it (see Kogure 2022 for relevant discussion).

[17] Appendix Tables A2 and A3 report the descriptive statistics.

migrated after repatriation. Our robustness check addresses this point and considers those who experienced multiple migrations after repatriation.

**Matched Samples.** Since our samples of refugees and stayers are not randomly determined, we face the potential endogeneity concerns. Given our historical contexts, our analysis adjusts for basic individual characteristics, such as age, sex, religion, mother tongue, education, and regional differences (religion and mother tongue are already exactly matched); regional differences reflect both economic and political differences.[18] Prior to the analysis, we check the differences in age, education, and district of birth by sex between the two samples, as detailed in Appendix Section A.2.1. We find three distinctions. First, the returnees from Thailand and the stayers have a distinct age distribution. Second, the male returnees from Thailand are more educated than the male stayers. Third, those born in districts near Thailand and Vietnam are more likely to have fled to Thailand and Vietnam, respectively.[19]

The covariate imbalances can lead to unrobust estimates and/or imprecise inferences in evaluating the impacts of forced displacement; with no comparable units, estimates can rely heavily on extrapolation and may not be credible (see, e.g., Imbens and Rubin 2015). This is of particular concern when one extensively explores heterogeneous impacts among various subpopulations (see Lechner and Strittmatter 2019 for relevant discussion). We construct matched samples of refugees and stayers with "strong common support" using the Imbens-Rubin approach (Imbens and Rubin 2015). With a large number of stayers from the complete count census data, we construct the matched samples by age, sex, education (for age 34–60), and district of birth (the former two covariates are exactly matched, as are mother tongue and religion), as detailed in Appendix Section A.2.2. The resulting matched samples

[18] One concern is that while fearful individuals might tend to have been refugees, we cannot adjust for this characteristic due to the limited availability of our data. In Section 6.2, we provide evidence that this potential concern is unlikely to drive our results.

[19] For the latter two findings, similar patterns occur in other contexts; more able Bosnians tended to leave the country (Kondylis 2010), while Burundians and Mozambicans in border regions tend to have become refugees (Fransen et al. 2017, Chiovelli et al. 2021).

(*Matched Samples*) consist of 36,012 returnees from Thailand, 5,145 returnees from Vietnam, and a corresponding number of stayers.[20] We confirm the greatly improved covariate balance in the matched samples.

### 4.3 Treatments

We regard forced displacement for Cambodian refugees as a bundled treatment, which encompasses not only forced displacement itself but also subsequent humanitarian assistance, duration of stay in camps, and the repatriation and integration process. While data limitations do not allow us to decompose the effects of the bundled treatment into each dimension, our interest is how this entire historical displacement event affected the refugees themselves in the long-run. Given that the nature of the bundled treatment differed between returnees from Thailand and Vietnam, primarily in terms of humanitarian assistance, duration of stay in camps, and the timing of repatriation, we evaluate the displacement impacts separately for these two groups. In Section 5.3, to dig deeper into the bundled treatment, we examine the heterogeneity with respect to age, timing of repatriation, refugee camps, and destinations after repatriation.

### 4.4 Outcomes

**Age 20-60.** We consider ten labor market outcome measures and one home ownership outcome measure: two indicator variables for participating in the labor market (*Labor Force*) and being employed (*Employed*), employment period (months) during the past 12 months (*Months Worked*), three indicator variables for being employed in the primary sector (*Primary Sector*), secondary sector (*Secondary Sector*), or tertiary sector (*Tertiary Sector*), four indicator variables for engaging in high-skilled work (*High-skilled Work*), middle-skilled work (*Middle-skilled Work*), low-skilled work (*Low-skilled Work*), or the armed forces (*Armed Forces*), and an indicator variable for owning the dwelling in which one lives (*Home Ownership*). For younger returnees aged 20–33, we also evaluate three educational outcome measures: two indicator variables

[20] See Appendix Tables A2 and A3 for the descriptive statistics.

for having attended school (*Some Education*) and having completed primary school (*Primary School*), and years of schooling (*Years of Schooling*).

We base the four labor market outcome measures on occupational status on the latest International Standard Classification of Occupation 2008 (ISCO-08) (ILO 2012).[21] While our census data lack information on wages, each occupational status broadly reflects the economic circumstances of refugees and stayers. According to the Cambodia Labour Force Survey 2012, the average monthly wages are approximately $151 for high-skilled work, $129 for middle-skilled work, $107 for low-skilled work, and $114 for armed forces (National Institute of Statistics 2013).[22] Our choice to evaluate engagement in the armed forces is motivated by the fact that the three political factions that administered the border camps all ran isolated "hidden camps" with the full support of the Royal Thai Army. Civilian populations in the border camps might have been recruited for the resistance army (Rogge 1990). This military service might have adverse impacts on the human capital and labor market outcomes of young people (Blattman and Annan 2010).

**Age 6-19.** For older returnees aged 34–60, we also evaluate the educational and labor market outcomes of their children aged 15–19, 12–14, and 6–11, if any; each analysis is limited to those actually living with each cohort. We use household-level outcome measures—the proportion or average among cohorts aged 15–19, 12–14, or 6–11 within households. For the first cohort, we focus on the same outcome measures described above, with an additional educational outcome measure, namely the proportion attending school (*School Participation*). Although this cohort already fin-

[21] Our criteria are as follows: high-skilled work – "managers," "professionals," and "technicians and associate professionals" (skill levels 3 and 4), middle-skilled work – "clerical support workers," "services and sales workers," "skilled agricultural, forestry and fishery workers," "craft and related trades workers," and "plant and machine operators, and assemblers" (skill level 2), low-skilled work – "elementary occupations" (skill level 1), and armed forces – "armed forces occupations" (skill levels 1, 2, and 4).

[22] We calculate the average wages for occupation groups falling under high-skilled work, middle-skilled work, low-skilled work, and armed forces using the Cambodia Labour Force Survey 2012 data. Previous labor force surveys conducted in 2000 and 2001 did not provide wage information for individual occupation groups.

ished their nine-year compulsory education in the 1998 Cambodian education system, some may still have been in school due to delayed entry, temporary dropout, or grade retention.

For the latter two cohorts, who were still receiving compulsory education (assuming they were receiving any education at all), we consider three educational outcome measures and one labor market outcome measure at the time of the 1998 Census: the proportion having attended school (*Some Education*), the proportion attending school (*School Participation*), the average grade progression (*Grade Progression*), and the proportion participating in the labor market (*Child Labor*). We measure the grade progression of each member of these cohorts by $Grade - (Age - 5)$, which takes 0 if he/she progresses from any grade to the next one and a negative value otherwise. We look at the two cohorts separately because the time of having entered school corresponds to that of repatriation for many of the cohort aged 12–14 for the returnees from Thailand, and their human capital accumulation might have been disrupted heterogeneously.

### 4.5 Analysis

#### 4.5.1 Full Samples

Our analysis based on the Full Samples by sex estimates the following equation using OLS with robust standard errors:

$$Y_{id} = \alpha + \gamma Refugee_i + X_i'\delta + \pi_d + \epsilon_{id}, \quad (1)$$

where $Y_{id}$ is the outcome of individual $i$ born in district $d$, $Refugee_i$ is an indicator variable equal to 1 if individuals are returnees from Thailand or Vietnam and 0 otherwise, $X_i$ is a vector of individual characteristics (age, age squared, years of schooling (only for age 34–60)) and $\pi_d$ denotes district of birth fixed effects. A parameter of interest is $\gamma$, which captures the effects of forced displacement on the outcome. To save space, we report the results based on the Full Samples, along with those of robustness checks, in Section 5.2. They are consistent with the results based on the

Matched Samples, the details of which we present below.

### 4.5.2 Matched Samples

Our analysis based on the Matched Samples follows a potential outcomes framework (Neyman 1923, Rubin 1974). The parameter of interest is the average treatment effect on the treated (ATT), which addresses the question of how returnees' outcomes would differ if they were stayers; as noted, the ATT captures the overall (total) impacts of forced displacement for returning refugees relative to stayers. Policy makers are generally interested in the economic situations facing those who became refugees themselves.

To identify the ATT, we impose assumptions of conditional independence or unconfoundedness and common support/overlap, as commonly used in causal studies (see, e.g., Angrist and Pischke 2008, Imbens and Rubin 2015). The former asserts that the two potential outcomes of refugees and stayers are independent conditional on the observed covariates. The latter ensures sufficient overlap in the covariate distributions of refugees and stayers. The analysis of the intergenerational impacts generally estimates the "net treatment difference" (NTD), as defined by Rosenbaum (1984), a parameter estimated conditionally on the observed values of the post-treatment variable because the existence of individuals aged 6–19 born after the refugee crisis (i.e., household formation) may have been affected by the event (Appendix Table A4 examines the impacts of forced displacement on sociodemographic outcomes).[23]

To estimate the ATT, we mainly use the bias-corrected version of the nearest-neighbor matching method (Abadie and Imbens 2011, Imbens 2015) (our robustness checks consider alternative estimation methods). Although simple matching estimators can be biased when the matching is not exact, matching (we use one-to-one

[23] We find evidence suggesting that male returnees from Thailand and Vietnam have fewer children aged 15–19, 12–14, and 6–11 (only for those from Vietnam) than comparable stayers, whereas female returnees from Thailand have fewer children aged 15–19 and more children aged 12–14 and 6–11 than comparable stayers; the results of the intergenerational impacts should be taken with some caution.

covariate matching) with replacement, in combination with regression adjustments ("bias adjustments") within the matched pairs, can produce estimators with little remaining bias (Abadie and Imbens 2011). To grasp the overall patterns of the impacts, we first estimate the ATT based on the aggregated samples of cohorts aged 34–60 and 20–33 by sex. We use the basic set of covariates (age, years of schooling (only for age 34–60), and district of birth fixed effects) in the bias adjustments.

Since we evaluate many outcome measures at the same time, we potentially face the multiple hypothesis testing problem: The probability of at least one Type I error increases with the number of tests, and significant impacts may emerge by chance, despite a lack of displacement impact (Anderson 2008). To address this concern, we use the Benjamini-Hochberg procedure to control the false discovery rate (FDR) of a family of all hypothesis tests (Benjamini and Hochberg 1995).[24] With $N$ hypothesis tests, we first sort and rank the $p$-values, giving the smallest $p$-value rank 1, the next smallest rank 2, and the largest rank $N$, then adjust each $p$-value by multiplying $N$ and dividing its assigned rank. We use this adjusted $p$-value to construct 95% confidence intervals, as well as to conduct hypothesis tests.

# 5 Results

## 5.1 Main Results

Figure 3 plots the point estimates of the impacts of forced displacement on selected key outcomes and their 95% confidence intervals (adjusted with the Benjamini-Hochberg procedure), along with the mean for stayers, for male and female returnees aged 34–60 and 20–33 from Thailand and Vietnam and for the male and female children, aged 15–19, 12–14, and 6–11, of the male returnees and stayers aged 34–60. Appendix Figures A4 and A5 respectively provide the complete results for all outcomes for returnees aged 20–60 and for the children, aged 6–19, of the returnees aged 34–60.

[24] FDR is the expected proportion of incorrectly rejected null hypotheses among all rejected null hypotheses. We also consider the classical Bonferroni and Holm multiple testing procedures (Bonferroni 1935, Holm 1979); these are overly conservative in our case, where many outcomes are mutually related. The results are available from the authors upon request.

**Age 34-60.** We find evidence that the displacement had adverse impacts on labor market outcomes and home ownership for the returnees from both Thailand and Vietnam (panel A of Figure 3 and Appendix Figure A4). For instance, the point estimates suggest that the displacement decreased the proportion of employment by 2.5 and 9.5 percentage points for the male and female returnees from Thailand, respectively, and by 0.4 and 6.2 percentage points for the male and female returnees from Vietnam, respectively; except in the case of the male returnees from Vietnam, these impacts are statistically significant at conventional levels. For all four groups of returnees, the displacement decreased the proportion of employment in the primary sector by more than 20 percentage points and increased the proportion of employment in the tertiary sector by more than 10 percentage points. At the same time, the displacement increased the proportion of engagement in low-skilled work by more than 5 percentage points and decreased the proportion of engagement in middle-skilled work by more than 8 percentage points. The displacement also decreased the proportion of home ownership for all four groups of returnees by 1.3–3.4 percentage points and increased the proportion of armed forces engagement for the male returnees from Thailand by 6.3 percentage points.

**Age 20-33.** In the evaluation of educational outcomes, we find a sharp contrast between the returnees from Thailand and those from Vietnam (panel B of Figure 3 and Appendix Figure A4). Note that this cohort received education after the refugee crisis, and that most returnees from Thailand received it in camps while those from Vietnam received it in Cambodia after repatriation. The displacement had positive impacts on the educational outcomes of the male returnees from Thailand and negative impacts on those of the male and female returnees from Vietnam. For instance, the point estimates suggest that the displacement increased years of schooling by 1.060 years for the male returnees from Thailand and decreased years of schooling by 0.848 years for the male returnees from Vietnam.

In an evaluation of labor market and home ownership outcomes, however, there

are no such contrasting impacts. The displacement had adverse impacts on the labor market outcomes of the returnees from both Thailand and Vietnam, spurring their employment shifts from the primary sector to the tertiary one.[25] The estimated adverse impacts are relatively more modest for the latter, who experienced worse educational outcomes. For instance, the point estimates for the male returnees from Thailand and Vietnam suggest that the displacement decreased the proportion of employment in the primary sector by 23.3 and 16.3 percentage points, respectively, and increased the proportion of employment in the tertiary sector by 19.4 and 12.8 percentage points, respectively. At the same time, the displacement increased the proportion of engagement in low-skilled work by 7.4 and 3.1 percentage points, respectively, and decreased the proportion of engagement in middle-skilled work by 14.7 and 1.5 percentage points, respectively. Meanwhile, the displacement decreased the proportion of home ownership by 4.5 and 2.3 percentage points, respectively, and increased the proportion of engagement in the armed forces by 4.2 percentage points for the male returnees from Thailand.

**Age 6-19.** For the cohort aged 15–19, the results are qualitatively similar to those for the returnees aged 20–33 (panel C of Figure 3 and Appendix Figure A5). For the cohorts aged 12–14 and 6–11, the displacement had some negative impacts on the educational outcomes of the cohort aged 12–14 of returnees from Thailand and Vietnam and on those of the cohort aged 6–11 of returnees from Vietnam (panels D and E of Figure 3 and Appendix Figure A5). The point estimates show that the estimated adverse impacts are relatively strong for the children of returnees from Vietnam.

### 5.2 Robustness Checks

We next check the robustness of our main results. We report the complete results in Appendix Section A.3, and only summarize the results here.

[25] One exception is that the displacement significantly increased the proportion of engagement in high-skilled work for the male returnees from Thailand. We confirm that these results are mainly driven by regional differences in 1998 (i.e., destination choices).

**Alternative Estimation Methods.** We first confirm the robustness of our base results based on the Matched Samples by comparing them to an alternative estimation method, namely blocking on the estimated propensity score in combination with regression adjustments within the blocks, following Imbens and Rubin (2015). The results from OLS regressions based on the Full Samples are also consistent with those based on the Matched Samples.

**Threats to Identification.** We next address potential threats to identification. Although we balanced the key determinants of refugees, unobserved characteristics might not be balanced between refugees and stayers, and the resulting estimates might be biased even in the Matched Samples. In particular, we are concerned that since refugees and stayers were exposed to conflict and violence under the Pol Pot regime (Kiernan 2008), the former, who experienced multiple migrations before the refugee event, might have been exposed to them differently (see Ruiz and Vargas-Silva 2013, pp. 773–774 for relevant discussion). We construct more balanced samples, further balancing the level of conflict and/or violence to which individuals were exposed before the refugee event, and show the robustness of our base results. In addition, to address any potential remaining concerns, for our matched samples, we consider omitted variable bias due to unobserved confounders through sensitivity analysis, following Rosenbaum (2002); for our full samples, meanwhile, we conduct sensitivity analysis by following Oster (2019). We confirm that both results are robust to omitted variable bias.

For the full samples, we also employ an instrumental variables strategy in conjunction with machine-learning techniques, following Windmeijer et al. (2019). We employ the machine learning-based, rather than standard, instrumental variables approach, because finding valid instrumental variables in advance is not feasible in our context.[26] We consider district of birth dummy variables (key determinants of refugees) as potential candidates for instruments, assuming that some instruments work and others

[26] For example, distance to border regions can correlate with the level of regional development and thus is not a valid instrument.

do not. Having no prior knowledge of which instruments are potentially valid, we use the adaptive Lasso approach of Windmeijer et al. (2019), building on Kang et al. (2016) to assess the validity of the potential instruments. We can identify the causal effect of forced displacement on outcomes when the proportion of invalid instruments is less than 50%.[27] Due to the limited number of potential instruments for returnees from Vietnam, this robustness check is feasible only for those from Thailand. Due to endogeneity concerns or conceptual issues, we do not consider intergenerational impacts. We confirm that the results that satisfy the identification assumption are generally consistent with the original results.

Lastly, we address the potential concern about resettlement selection. Our samples of former refugees are limited to those who returned to Cambodia, thus excluding those who resettled in third countries. If returning refugees tend to have lower abilities, then the differences in unobserved ability between returnees and stayers might not be adequately balanced and might partly drive our results. We separately consider the returnees from the UNHCR and UNBRO camps (see Appendix Section A.3.2 for the construction of the samples); since Cambodian refugees in Thailand could be resettled to third countries, including the United States, only via the UNHCR camps, the sample selection problem, if any, is limited in the latter samples. We confirm that the results for the returnees from the UNBRO camps are consistent with the original results.

**External Validity.** Our robustness checks also address the potential threats to external validity in the Cambodian context. Because our samples of returning refugees are limited to those who did not migrate after repatriation, the displacement impacts might systematically differ for this group. We construct alternative samples based on an alternative definition of refugees, including those who experienced multiple migrations after repatriation, and show that our base results serve as conservative estimates of the displacement impacts: Returnees with multiple migrations

[27] To avoid computational errors, we limit the potential instruments to the dummy variables for districts where the proportion of the returnees is between 0.1 and 0.9.

experienced far worse educational, labor market, and home ownership outcomes.

**Additional Discussion.** To further mitigate concerns about omitted variable bias, we present additional reasons to believe that omitted variable bias is unlikely to be significant to alter our main results. First, as noted, the male returnees aged 34-60 from Thailand tend to have higher levels of education than the male stayers aged 34-60 (panel A-1 of Appendix Figure A2). Since ability is generally positively correlated with levels of education, the former might have higher ability than the latter in the Matched Sample, if any difference. As a result, our base results for the former should serve as conservative estimates of the displacement impacts on their labor market outcomes. Second, below, we provide results on heterogeneity in the displacement impacts with respect to age, timing of repatriation, refugee camps, and destinations. It is unlikely that unobserved characteristics systematically drive all of these results.

### 5.3 Heterogeneity

Given the robustness of our base results, we next examine the heterogeneity in displacement impacts with respect to age, timing of repatriation, refugee camps, and destinations. While returnees' choices of the timing of repatriation and destination are not random, we regard these choices (post-treatment variables) as interesting potential channels to drive our results. The first analysis is based on subsamples matched by age and sex, and the latter three analyses are based on the aggregated samples.

**Age.** Appendix Figure A12 examines heterogeneity with respect to age. While the results are broadly consistent with those based on the aggregated samples, we note the following distinction. The estimated impacts on educational outcomes are positive for the female returnees aged 20–26 from Thailand (the majority of whom had to start primary school in camps), but negative for those aged 27–33 (panel A2) (the majority of whom had to start primary school in Cambodia prior to forced displacement). These results suggest that the positive and negative impacts offset each other in the aggre-

gated sample (Figure 3). We may need to pay special attention to girls who are not in school, because they might have less access to educational opportunities in camps.

**Timing of Repatriation.** Focusing on four key outcomes, Figures 4 and 5 consider the heterogeneity among the returnees aged 34–60 and 20–33 from Thailand who returned in 1979–1987, 1988–1990, 1991, 1992, 1993, and 1994–1998 and among those aged 34–60 and 20–33 from Vietnam who returned in 1979, 1980, and 1981–1998, respectively (see also Appendix Figure A13). Although the timing of repatriation is not random and some caution is hence needed,[28] the analyses exhibit a sharp contrast: The later a given refugee returned, the worse their labor market outcomes, with employment shifts into the tertiary sector. In terms of employment shifts, we find some gaps between the male returnees from Thailand who returned in 1991 and those who returned in 1992. Given that the latter generally joined the repatriation program, they might have been more likely to have received assistance for accessing agricultural land. There are no such differences for the female returnees, because women are generally less likely to engage in agricultural work. For the returnees aged 20–33 from Thailand, late returnees experienced better educational outcomes but worse labor market outcomes.

Figure 6 provides results for the children aged 15–19 of the male returnees and stayers aged 34–60 (Appendix Figure A14 shows the results for the children aged 6–19). We find that the above pattern (i.e., late returnees experienced better educational outcomes but worse labor market outcomes) is relatively weak for age 15–19. For age 12–14 and 6–11, we find evidence that the displacement had adverse impacts on the educational outcomes of the cohorts who had to receive education around the time

[28] Appendix Section A.4.1 examines individual and regional characteristics correlated with early return migration decisions (before 1992) for the returnees aged 34–60 from Thailand. We focus on this age cohort because they seem to have been old enough to make migration decisions independently at the time. We find that younger, male, and less educated refugees, those from districts away from border regions, those from districts more contaminated with landmines and UXOs, and those who stayed in the UNBRO camps tend to have become early returnees. In a different context, Beaman et al. (2022) study return migration decisions for Syrian refugees stemming from the "Arab Spring" in Jordan, Lebanon, and Iraq from 2011 to 2018. They find that security and access to utilities in Syria played crucial roles in facilitating their return migration decisions.

of repatriation; these results suggest that the adverse displacement impacts in the aggregated samples (Figure 3) are largely driven by those cohorts. The results are also consistent with those for the returnees aged 20–30 from Vietnam, the majority of whom had to receive education right after repatriation.

**Refugee Camps.** Appendix Figures A8 and A9 examine the heterogeneity between the returnees from the UNHCR and UNBRO camps. Assuming no sample selection problem for the returnees from the UNHCR camps, the following findings might be worthy of note. First, the returnees aged 34–60 from the UNHCR camps experienced relatively worse labor market outcomes. Given that the UNHCR camps provided more elaborate services, this finding is surprising. Second, this pattern also occurs for the cohort aged 15–19, although returnees of this age from the UNHCR camps experienced relatively better educational outcomes. Third, the estimated impacts on engagement in the armed forces are relatively strong for the male returnees aged 34–60 from the UNBRO camps. Appendix Table A9 examines these findings using OLS, finding many statistically significant differences.

These findings might be interpreted as follows. Since the UNBRO camps were affiliated with political factions, those who stayed in them might have found it relatively easy to secure jobs through the networks built in the camps or through the exchange of information, though the men might have faced an increased risk of being recruited for the armed forces. In contrast, because most of those who stayed in the UNHCR (neutral) camps hoped for third country resettlement into the early 1990s (Suenobu 1995, p.3), they might have tended to not have much of a social network, thus finding it relatively difficult to secure jobs. Taking all this information into account, social networks built in camps might have significantly affected subsequent outcomes. This interpretation is consistent with the case of Japanese American interment during WWII (Arellano-Bover 2022): Those displaced to internment camps could re-optimize their job and location choices after internment through the exchange of information and skills in these camps.

**Destinations.** We have anecdotal evidence that social networks with relatives and village leaders at the time of repatriation were indeed important for access to agricultural land and local knowledge (Black and Koser 1999). To gain further insights into the role of social networks in the repatriation and reintegration process, Appendix Figures A15 and A16 consider heterogeneity for returnees who came back to their birth village and district. We assume that such returnees had relatively better access to social networks at the time of repatriation. Although returnees' choice of destination is not random and we should again be cautious,[29] the estimated adverse impacts on labor market outcomes are relatively modest for the returnees aged 34–60 and 20–33 from both Thailand and Vietnam; on the other hand, we confirm that the differences in local labor markets are unlikely to drive our results. The repatriation and reintegration process may need to carefully consider the availability of social networks.

## 6 Mechanisms

This section empirically explores the potential mechanisms underlying the adverse displacement impacts on labor market outcomes, additionally using the 2004 Cambodia Socio-Economic Survey (CSES) microdata for returning refugees and stayers aged 26-66, which corresponds to age 20-60 at the time of the 1998 Census. Appendix Section A.5.1 confirms the consistency of the 1998 Census and the 2004 CSES data, showing that the adverse displacement impacts continue in 2004.

We first examine whether the adverse displacement impacts are driven by congested labor markets resulting from limited access to agricultural land. Then, we consider other potential channels, including discrimination, health, and remittance networks. Due to the lack of specific information about previous residence (e.g., Thailand, Vietnam) in the 2004 CSES data, our analysis, based on this data, defines

[29] Appendix Section A.4.2 examines individual and regional characteristics correlated with return migration decisions in favor of birth districts among the returnees aged 34–60 from Thailand and Vietnam. For both groups of returnees, we find that less educated refugees and those from districts near border regions tend to have returned to their birth districts.

returning refugees as those who have lived abroad before and migrated to their current residence in 1979-1998. Since the great majority of returnees from Thailand and Vietnam live in the former Northwest, West, and North zones and in the Southwest and East zones, respectively, our analysis also restricts the samples to those living in these regions. These limited samples should largely capture the returnees from Thailand and Vietnam.

### 6.1 Congestion

After the collapse of the Pol Pot regime, the People's Republic of Kampuchea (PRK), supported by Vietnamese troops, was established in 1979. The PRK adopted socialism, respected basic human rights, and emphasized agriculture as the primary industry. In 1979, they introduced "solidarity groups" (*krom samakki*), consisting of 10–15 families who worked cooperatively and shared their production. This system arguably determined subsequent farmland ownership in Cambodia (Amakawa 2001), despite the establishment of de jure private property rights in 1989.[30] Indeed, the retrospective 2004 CSES data show that a large number of Cambodian households began using their current agricultural land in 1979 (see Appendix Figure A17).[31]

Given this, our empirical findings may imply that returning refugees, especially those who came back later, tend to have lacked access to agricultural land, something exacerbated by high levels of contamination by landmines and UXOs during the conflict, and to have been pushed out of agricultural labor markets. As a result, they might tend toward engagement in other sectors, especially the tertiary sector. At the same time, because other sectors had not at this time been developed in Cambodia (Vickery 1999), they might tend to have engaged in low-skilled work in the sectors

[30] People's Republic of Kampuchea, Council of Ministers, Sub-Decree No.25 on Providing House Ownership to the Cambodian Population, 22 April 1989; People's Republic of Kampuchea, Council of Ministers, Instruction No.03 on Implementation of Land use and Management Policy, 3 June 1989.

[31] Land ownership is less likely to have affected the migration decisions of refugees (i.e., reverse causality is unlikely). This is because migration decisions from refugees generally occurred under the Pol Pot regime or when Vietnamese troops invaded Cambodia in 1978–1979, before the establishment of *krom samakki* under the new socialist regime. In addition, during and after the Pol Pot regime in 1975–1989, people were not allowed to own land.

(i.e., informal sector), unlike in developed countries. In other words, these other sectors might not yet have matured enough to absorb the available labor force. In sum, the adverse displacement impacts might be attributable mainly to the congested labor markets in all sectors, resulting directly or indirectly from limited access to agricultural land. This subsection provides supporting evidence for the congestion mechanism.

**Access to Agricultural Land.** First, focusing on household heads aged 26–66 from the 2004 CSES data, Table 1 directly tests whether returning refugees are indeed less likely to have access to agricultural land. It's worth noting that the report based on the field survey conducted in 1989 (i.e., before the limited availability of arable land became apparent for refugees in Thailand) documents that the majority are likely to engage in agriculture work after repatriation (Lynch 1989). The dependent variable is an indicator variable equal to 1 if households had access to agricultural land ("owned it," "rented it," or "had access some other way") by 1998 and 0 otherwise (*Access to Agricultural Land*).

Adjusting for age, age squared, and a dummy variable for female, column 1 reveals that the returning refugees had access to agricultural land by 1998 45.8% less often than stayers (the mean is 74.7%). When we add a control for years of schooling in column 3, the coefficient of refugee status remains similar and highly significant. When we additionally adjust for district fixed effects in column 5, the magnitude of the coefficient decreases to -27.7%, implying that access to agricultural land is substantially affected by regional differences or geographic characteristics. Nevertheless, the returning refugees still have far less access to agricultural land. Lastly, when we limit the samples to those who live in the former Northwest, West, and North zones and in the Southwest and East zones in columns 7 and 9, respectively, we see consistent results. This suggests returning refugees are indeed less likely to have access to agricultural land.

Given that the availability of agricultural land potentially differs across regions due

to landmine and UXO contamination, in the even columns, we additionally employ nationwide geospatial data about landmine and UXO contaminated areas in 1992 (before clearance started). Appendix Figure A18 shows the spatial distribution, which reveals that the Northwest zone, with a large number of returnees from Thailand, is particularly heavily contaminated. Using a Geographic Information System (GIS), we construct a congestion measure and estimate the following equation via OLS with robust standard errors clustered by village:

$$Y_{ivd} = \alpha + \gamma_1 Refugee_i + \gamma_2 Refugee_i \times Congestion_v + X'_{iv}\beta + \mu_d + \epsilon_{ivd}, \quad (2)$$

where $Congestion_v$ is the village-level congestion measure, defined as stayer density per non-contaminated area ($km^2$) within a 3.0 km buffer zone around each village point before clearance started (see Appendix Figure A19 for the distribution), $X_{iv}$ is a vector of individual and village characteristics (age, age squared, a dummy variable for female, years of schooling, $Congestion_v$, and the logarithmic value of village population aged 20–60 in 1998), and $\mu_d$ denotes district fixed effects; the demographic variable is based on the 1998 Census data.[32] Note that the analysis samples are limited to household heads residing in villages with complete information about village points (contamination) and demographics.

We also have the following key finding: Returning refugees, especially those who live in more congested areas, are less likely to have access to agricultural land. This is true for the Northwest, West, and North zones, but not for the Southwest and East zones. The probable reason this pattern does not apply in the latter zones is because many returnees in these zones are from Vietnam, and thus came back early and had access to agricultural land by joining *krom samakki*. Indeed, the 2004 CSES data show that 35.6% of the plots of returnees in the Southwest and East zones and 20.3% of the plots of returnees in the Northwest, West, and North zones were "given

[32] We confirm that the results are consistent with those for alternative buffer sizes (1.0 km and 2.0 km) and alternative congestion measures, namely the logarithmic values of one plus $Congestion_v$ (i.e., $ln(1 + Congestion_v)$) for the three buffer sizes.

by the state," while the average year in which returnees started to use agricultural land in the Southwest and East zones and in the Northwest, West, and North zones is 1989 and 1992, respectively; this timing is consistent with the period that saw the establishment of de jure private property rights in 1989 and the organization of the repatriation program in 1992–1993.

**Labor Market and Home Ownership Outcomes.** Second, using the aggregated Matched Samples of the returnees and stayers aged 20–60, columns 1–5 in panel A of Table 2 estimate equation (2) for five key labor market and home ownership outcomes (employed, primary sector, tertiary sector, low-skilled work, and home ownership), additionally adjusting for district of birth fixed effects. To gain more insights, columns 6 and 7 restrict the samples to those engaged in the tertiary sector and consider two additional outcomes (low-skilled work and paid employee). The latter is an indicator variable for work as a wage and salary worker (*Paid Employee*) (i.e., in a better quality job). We find evidence that returning refugees, especially those who live in more congested areas, experience worse labor market and home ownership outcomes. Along with the findings in Table 1, these results imply that returnees lacking access to agricultural land due to congested agricultural land markets experience relatively worse labor market and home ownership outcomes.

Lastly, panel B of Table 2 examines the heterogeneity in displacement impacts among those who returned in 1979, 1980, 1981–1991, 1992, 1993, and 1994–1998, adjusting for both district of birth fixed effects and district fixed effects (Appendix Tables A13 and A14 consider the returnees and stayers aged 34–60 and 20–33 and the children aged 6–19, respectively). Given the above findings, we expect later returnees to experience worse labor market outcomes, because they tend to lack access to agricultural land. And indeed, the results strongly support this hypothesis. Importantly, since district fixed effects are conditioned on, the difference in destination choices is unlikely to drive our results. In sum, the results in Tables 1 and 2 strongly support the congestion channel.

### 6.2 Other Potential Channels

Motivated by the relevant literature and social contexts, Appendix Section A.5.2 also considers other potential channels, including discrimination, health, and remittance networks (see, e.g., Brell et al. 2020, Currie and Madrian 1999, Chami et al. 2005 for relevant discussions). There is a possibility that the adverse displacement impacts might be driven by potential discrimination against returnees, poor health for returnees, or better/worse remittance networks for returnees (e.g., relatives or friends abroad). We find no evidence that these potential channels are likely to drive our results.

## 7 Conclusions

This paper examined the consequences of forced displacement for Cambodian refugees. Focusing on returnees from Thailand and Vietnam and on their children, we mainly evaluated their educational and labor market outcomes. Our analyses revealed that the returnees from Thailand attained higher levels of education and that those from Vietnam, by contrast, attained lower levels of education through forced displacement, consistently with the availability of humanitarian assistance in Thailand and Vietnam. On the other hand, the two groups both experienced worse labor market outcomes, with employment shifts from the primary sector to the immature tertiary sector (informal sector). The adverse displacement impacts were relatively stronger for later returnees. We then empirically explored the potential mechanisms underlying the main results and provided strong suggestive evidence that congested labor markets resulting from limited access to agricultural land, exacerbated by high contamination by landmines and UXOs during the conflict, likely drive the results. Based on our empirical findings, we also discussed policy implications during camps, repatriation, and reintegration.

Focusing on the Cambodian refugee crisis and extensively exploring its heterogeneous impacts among various subpopulations of refugees, our study provides insights

into the complex relationships among political and social structures, social situations, and economic behavior. Our results highlight the importance of considering potential heterogeneity in behavior among people facing different social and political constraints, and they demonstrate that forced displacement, or more generally civil conflict, can be a potential source of future misallocation. Such considerations can lead to effective policy design, as well as a better understanding of human behavior and the detailed process of economic change in conflict-affected societies.

## References


**Abadie, Alberto and Guido W. Imbens**, "Bias-Corrected Matching Estimators for Average Treatment Effects," *Journal of Business and Economic Statistics*, 2011, *29* (1), 1–11.

**Amakawa, Naoko**, "System and Structure of Farmland Ownership: Restructuring Process after Pol Pot Regime," in Naoko Amakawa, ed., *Reconstruction and Development in Cambodia*, Chiba: Institute of Developing Economies, 2001, chapter 4, pp. 151–211. Written in Japanese.

**Anderson, Michael L.**, "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

**Angrist, Joshua D. and Jörn-Steffen Pischke**, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press, December 2008.

**Arellano-Bover, Jaime**, "Displacement, Diversity, and Mobility: Career Impacts of Japanese American Internment," *Journal of Economic History*, 2022, *82* (1), 126–174.

**Bauer, Thomas K., Sebastian Braun, and Michael Kvasnicka**, "The Economic Integration of Forced Migrants: Evidence for Post-War Germany," *Economic Journal*, 2013, *123* (571), 998–1024.

**Beaman, Lori, Harun Onder, and Stefanie Onder**, "When Do Refugees Return Home? Evidence from Syrian Displacement in Mashreq," *Journal of Development Economics*, 2022, *155* (C), 102802.

**Becker, Sascha O.**, "Forced Displacement in History: Some Recent Research," *Australian Economic History Review*, 2022, *62* (1), 2–25.

. **and Andreas Ferrara**, "Consequences of Forced Migration: A Survey of Recent Findings," *Labour Economics*, 2019, *59*, 1–16.

**, Irena Grosfeld, Pauline Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya**, "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers," *American Economic Review*, 2020, *110* (5), 1430–1463.

**Benjamini, Yoav and Yosef Hochberg**, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1995, *57* (1), 289–300.

**Black, Richard and Khalid Koser**, *The End of the Refugee Cycle?: Refugee Repatriation and Reconstruction*, Vol. 4, Berghahn Books, 1999.

**Blattman, Christopher and Edward Miguel**, "Civil War," *Journal of Economic Literature*, 2010, *48* (1), 3–57.

**and Jeannie Annan**, "The Consequences of Child Soldiering," *Review of Economics and Statistics,*, 2010, *92* (4), 882–898.

**Bonferroni, Carlo Emilio**, *Il calcolo delle assicurazioni su gruppi di teste*, Tipografia del Senato, 1935.

**Brell, Courtney, Christian Dustmann, and Ian Preston**, "The Labor Market Integration of Refugee Migrants in High-Income Countries," *Journal of Economic Perspectives*, 2020, *34* (1), 94–121.

**Chami, Ralph, Connel Fullenkamp, and Samir Jahjah**, "Are Immigrant Remittance Flows a Source of Capital for Development?," *IMF Staff Papers*, 2005, *52* (1), 55–81.

**Chiovelli, Giorgio, Stelios Michalopoulos, Elias Papaioannou, and Sandra Sequeira**, "Forced Displacement and Human Capital: Evidence from Separated Siblings," 2021. Working Paper 29589, National Bureau of Economic Research.

**Cortes, Kalena E.**, "Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States," *Review of Economics and Statistics*, 2004, *86* (2), 465–480.

**Currie, Janet and Brigitte C. Madrian**, "Health, Health Insurance and The Labor Market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3, Elsevier, 1999, chapter 50, pp. 3309–3416.

**de Walque, Damien**, "The Socio-Demographic Legacy of the Khmer Rouge Period in Cambodia," *Population Studies*, 2006, *60* (2), 223–231.

**Dell, Melissa and Pablo Querubin**, "Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies," *Quarterly Journal of Economics*, 2018, *133* (133), 701–764.

**Deryugina, Tatyana, Laura Kawano, and Steven Levitt**, "The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns," *American Economic Journal: Applied Economics*, 2018, *10* (2), 202–233.

**Dy, Khamboly**, *A History of Democratic Kampuchea (1975-1979)*, Phnom Penh: Documentation Center of Cambodia, 2007.

**Fiala, Nathan**, "Economic Consequences of Forced Displacement," *Journal of Development Studies*, 2015, *51* (10), 1275–1293.

**Fransen, Sonja, Carlos Vargas-Silva, and Melissa Siegel**, "The Impact of Refugee Experiences on Education: Evidence from Burundi," *IZA Journal of Development and Migration*, 2018, *8* (6).

_ **, Isabel Ruiz, and Carlos Vargas-Silva**, "Return Migration and Economic Outcomes in the Conflict Context," *World Development*, 2017, *95*, 196–210.

**Greve, Hanne Sophie**, "Kampuchean Refugees between the Tiger and the Crocodile: International Law and the Overall Scope of One Refugee Situation," 1987. Unpublished PhD dissertation, University of Bergen, Bergen, Norway.

**Holm, Sture**, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, 1979, *6* (2), 65–70.

**Ibáñez, Ana María and Andrés Moya**, "Vulnerability of Victims of Civil Conflicts: Empirical Evidence for the Displaced Population in Colombia," *World Development*, 2010, *38* (4), 647–663.

_ **and Carlos Eduardo Vélez**, "Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia," *World Development*, 2008, *36* (4), 659–676.

**ILO**, *International Standard Classification of Occupations 2008 (ISCO-08): Structure, Group Definitions and Correspondence Tables*, Geneva: National Institute of Statistics, Ministry of Planning, 2012.

**Imbens, Guido W.**, "Matching Methods in Practice: Three Examples," *Journal of Human Resources*, 2015, *50* (2), 373–419.

_ **and Donald B. Rubin**, *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, New York: Cambridge University Press, 2015.

**Islam, Asadul, Chandarany Ouch, Russell Smyth, and Liang Choon Wang**, "The Long-Term Effects of Civil Conflicts on Education, Earnings, and Fertility: Evidence from Cambodia," *Journal of Comparative Economics*, 2016, *44* (3), 800–820.

**Iwanowsky, Mathias and Andreas Madestam**, "State Repression, Exit, and Voice: Living in the Shadow of Cambodia's Killing Fields," 2019. unpublished manuscript, University of Munich and Stockholm University.

**Kang, Hyunseung, Anru Zhang, T. Tony Cai, and Dylan S. Small**, "Instrumental Variables Estimation With Some Invalid Instruments and its Application to

Mendelian Randomization," *Journal of the American Statistical Association*, 2016, *111* (513), 132–144.

**Kiernan, Ben**, *The Pol Pot Regime: Race, Power, and Genocide in Cambodia under the Khmer Rouge, 1975-79*, 3rd ed., London: Yale University Press, 2008.

**Kocher, Matthew Adam, Thomas B. Pepinsky, and Stathis N. Kalyvas**, "Aerial Bombing and Counterinsurgency in the Vietnam War," *American Journal of Political Science*, 2011, *55* (2), 201–218.

**Kogure, Katsuo**, "Some Remarks on the Causal Inference for Historical Persistence with Contemporary Microdata," 2022. unpublished manuscript.

_ **and Yoshito Takasaki**, "Conflict, Institutions, and Economic Behavior: Legacies of the Cambodian Genocide," 2016. HIAS Discussion Paper E-39, Hitotsubashi University.

**Kondylis, Florence**, "Agricultural Outputs and Conflict Displacement: Evidence from a Policy Intervention in Rwanda," *Economic Development and Cultural Change*, 2008, *57* (1), 31–66.

_ , "Conflict Displacement and Labor Market Outcomes in Post-War Bosnia and Herzegovina," *Journal of Development Economics*, 2010, *93* (2), 235–248.

**Lechner, Michael and Anthony Strittmatter**, "Practical Procedures to Deal with Common Support Problems in Matching Estimation," *Econometric Reviews*, 2019, *38* (2), 193–207.

**Lin, Erin**, "How War Changes Land: Soil Fertility, Unexploded Bombs, and the Underdevelopment of Cambodia," *American Journal of Political Science*, 2022, *66* (1), 222–237.

**Lynch, James F.**, "Border Khmer: A Demographic Study of the Residents of Site 2, Site B, and Site 8," 1989. Ford Foundation: Bangkok.

**Merrouche, Ouarda**, "The Long Term Educational Cost of War: Evidence from Landmine Contamination in Cambodia," *Journal of Development Studies*, 2011, *47* (3), 399–416.

**Miguel, Edward and Gerard Roland**, "The Long-run Impact of Bombing Vietnam," *Journal of Development Economics*, 2011, *96* (1), 1–15.

**Ministry of Planning**, *Cambodia Millennium Development Goals Report 2003*, Phnom Penh: Ministry of Planning, 2003.

**Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson**, "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," *Review of Economic Studies*, 2022, *89* (3), 1557–1592.

**National Institute of Statistics**, *General Population Census of Cambodia 1998: Final Census Results*, 2nd ed., Phnom Penh: National Institute of Statistics, Ministry of Planning, 2002.

. , *Cambodia Labour Force and Child Labour Survey 2012: Labour Force Report*, Phnom Penh: National Institute of Statistics, Ministry of Planning, ORC Macro., 2013.

**Neyman, Jerzy**, "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9.," Translated in *Statistical Science*, 1923, *5* (4), 465–480, 1990.

**Oster, Emily**, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

**Reynell, Josephine**, *Political Pawns: Refugees on the Thai-Kampuchean Border*, Oxford: Refugee Studies Programme, 1989.

**Riaño, Juan Felipe and Felipe Valencia Caicedo**, "Collateral Damage: The Legacy of the Secret War in Laos," 2020. CEPR Discussion Paper No. DP15349.

**Robinson, W. Courtland**, ""Something Like Home Again" the Repatriation of Cambodian Refugee," 1994. The U.S. Committee for Refugees, Washington D.C.

. , *Terms of Refuge: The Indo-Chinese Exodus and the International Response*, London: Zed Books, 1998.

**Rogge, John R.**, "Return to Cambodia: The Significance and Implications of Past, Present and Future Spontaneous Repatriations," 1990. The Intertect Institute.

**Rosenbaum, Paul R.**, "The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment," *Journal of the Royal Statistical Society, Series A (General)*, 1984, *147* (5), 656–666.

. , *Observational Studies*, 2nd ed., New York: Springer, 2002.

**Rubin, Donald B.**, "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 1974, *66* (5), 688–701.

**Ruiz, Isabel and Carlos Vargas-Silva**, "The Economics of Forced Migration," *Journal of Development Studies*, 2013, *49* (6), 772–784.

**Sarvimäki, Matti, Roope Uusitalo, and Markus Jäntti**, "Long-Term Effects of Forced Migration," 2009. IZA DP No. 4003.

**. , . , and .** , "Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations," *Journal of the European Economic Association*, 2022, *20* (6), 2497–2539.

**Suenobu, Yumiko**, "Management of Education Systems in Zones of Conflict-Relief Operations: A Case-Study in Thailand," 1995. UNESCO Principal Regional Office for Asia and the Pacific.

**Takasaki, Yoshito**, "Impacts of Disability on Poverty: Quasi-experimental Evidence from Landmine Amputees in Cambodia," *Journal of Economic Behavior and Organization*, 2020, *180*, 85–107.

**UNHCR**, *The State of The World's Refugees*, Geneva: UNHCR, 2000.

**Verme, Paolo and Kirsten Schuettler**, "The Impact of Forced Displacement on Host Communities: A Review of the Empirical Literature in Economics," *Journal of Development Economics*, 2021, *150*, 102606.

**Vickery, Michael**, *Cambodia 1975-1982*, Boston: South End Press, 1999.

**Windmeijer, Frank, Helmut Farbmacher, Neil Davies, and George Davey Smith**, "On the Use of the Lasso for Instrumental Variables Estimation with Some Invalid Instruments," *Journal of the American Statistical Association*, 2019, *114* (527), 1339–1350.

**Wiskari, Werner**, "Vietnam CopesWith Cambodian Refugees, Who Sound Thankful Despite Their Privations," *The New York Times*, April 1978, p. 16.

Table 1: Access to Agricultural Land in 1998

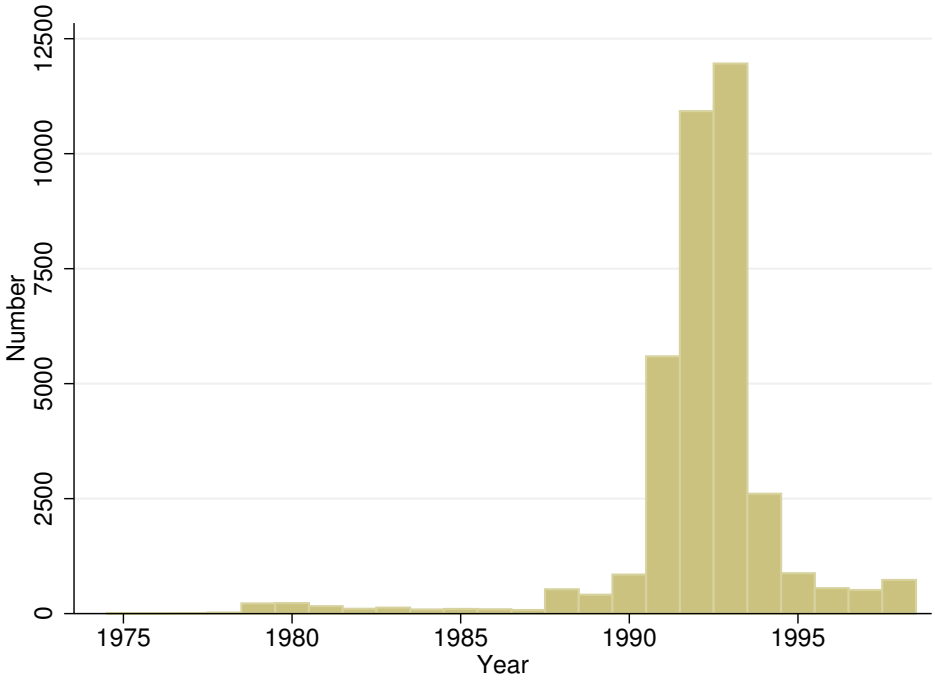
| Dependent Variable: | Access to Agricultural Land | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample: | All | All | All | All | All | All | NW West North | NW West North | SW East | SW East |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Refugee | -0.4581 | -0.2857 | -0.4457 | -0.2857 | -0.2769 | -0.2334 | -0.3290 | -0.1561 | -0.1769 | -0.2173 |
| | (0.0434) | (0.0498) | (0.0431) | (0.0498) | (0.0445) | (0.0601) | (0.0640) | (0.0739) | (0.0572) | (0.1031) |
| Refugee×Congestion | | -0.0038 | | -0.0038 | | -0.0035 | | -0.0186 | | 0.0097 |
| | | (0.0051) | | (0.0051) | | (0.0058) | | (0.0051) | | (0.0063) |
| Congestion | | 0.0071 | | 0.0071 | | 0.0092 | | 0.0142 | | 0.0073 |
| | | (0.0014) | | (0.0014) | | (0.0019) | | (0.0031) | | (0.0024) |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Years of schooling | No | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | No | No | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean (Dep. Var.) | 0.747 | 0.745 | 0.747 | 0.745 | 0.747 | 0.745 | 0.730 | 0.727 | 0.773 | 0.769 |
| Mean (Congestion) | | 9.838 | | 9.838 | | 9.838 | | 7.075 | | 11.781 |
| SD (Congestion) | | 8.759 | | 8.759 | | 8.759 | | 7.455 | | 9.174 |
| Observations | 6,087 | 5,364 | 6,087 | 5,364 | 6,087 | 5,364 | 2,245 | 1,938 | 2,752 | 2,394 |
| R-squared | 0.070 | 0.139 | 0.073 | 0.139 | 0.231 | 0.229 | 0.270 | 0.288 | 0.228 | 0.209 |

*Notes:* The table reports OLS estimates where the unit of observation is the household head. It reports robust standard errors, adjusted for clustering by village, in parentheses. Regressions use data about household heads aged 26–66 from the Cambodia Socio-Economic Survey 2004; age 26–66 corresponds to age 20–60 at the time of the 1998 Census. Columns 7-8 and columns 9-10 limit the analysis samples to the household heads residing in the former Northwest (NW), West, and North zones and in the former Southwest (SW) and East zones, respectively (see Appendix Figure A3). "Access to Agricultural Land" is an indicator variable equal to 1 if households had access to agricultural land (owned it, rented it, or had access in some other way) by 1998 and 0 otherwise. "Refugee" is an indicator variable equal to 1 if household heads lived abroad before and migrated to their current residence in 1979–1998 and 0 otherwise. "Congestion" is the village-level congestion measure, defined as stayer density per non-contaminated area ($km^2$) within a 3.0 km buffer zone around each village point before clearance started. Base controls include age, age squared, a dummy variable for female, and the logarithmic value of village population aged 20–60 in 1998. The analysis samples in the even columns are limited to the household heads residing in villages with complete information about village points (contamination) and demographics.

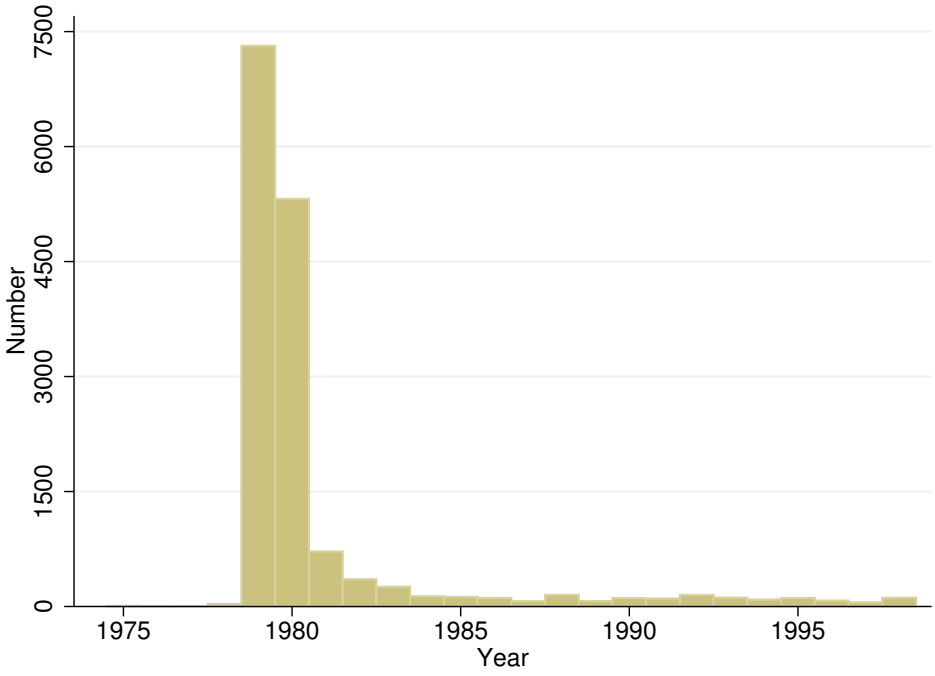
Table 2: Congestion Channel

| Sample: | All | | | | | Tertiary Sector | |
|---|---|---|---|---|---|---|---|
| Dependent Variable: | Emp-loyed | Primary Sector | Tertiary Sector | Low-skill. Work | Home Own. | Low-skill. Work | Paid Empl. |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | | | | A. Congestion | | | |
| Refugee | 0.0055 | -0.0519 | 0.0519 | 0.0387 | -0.0094 | 0.0905 | -0.1480 |
| | (0.0086) | (0.0123) | (0.0067) | (0.0050) | (0.0025) | (0.0218) | (0.0220) |
| Refugee×Congestion | -0.0025 | -0.0050 | 0.0022 | 0.0018 | -0.0007 | 0.0004 | 0.0016 |
| | (0.0007) | (0.0009) | (0.0006) | (0.0004) | (0.0002) | (0.0013) | (0.0013) |
| Congestion | 0.0018 | 0.0049 | -0.0029 | -0.0015 | 0.0008 | -0.0001 | -0.0019 |
| | (0.0007) | (0.0009) | (0.0005) | (0.0004) | (0.0002) | (0.0012) | (0.0013) |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Years of schooling | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean (Dep. Var.) | 0.856 | 0.703 | 0.145 | 0.058 | 0.973 | 0.317 | 0.440 |
| Mean (Congestion) | 8.903 | 8.903 | 8.903 | 8.903 | 8.903 | 7.394 | 7.394 |
| SD (Congestion) | 9.592 | 9.592 | 9.592 | 9.592 | 9.592 | 10.047 | 10.047 |
| Observations | 77,576 | 77,576 | 77,576 | 77,576 | 77,576 | 11,232 | 11,232 |
| R-squared | 0.128 | 0.236 | 0.190 | 0.084 | 0.040 | 0.261 | 0.338 |
| | | | | B. Timing of Repatriation | | | |
| 1979 | 0.053 | -0.005 | 0.053 | 0.018 | 0.013 | 0.028 | -0.064 |
| | (0.012) | (0.018) | (0.011) | (0.007) | (0.005) | (0.033) | (0.036) |
| 1980 | 0.016 | -0.034 | 0.046 | 0.016 | 0.018 | 0.050 | -0.114 |
| | (0.022) | (0.026) | (0.011) | (0.009) | (0.007) | (0.034) | (0.032) |
| 1981-1991 | -0.015 | -0.126 | 0.096 | 0.069 | -0.016 | 0.100 | -0.156 |
| | (0.008) | (0.013) | (0.008) | (0.007) | (0.004) | (0.020) | (0.021) |
| 1992 | -0.023 | -0.111 | 0.077 | 0.058 | -0.014 | 0.096 | -0.118 |
| | (0.010) | (0.013) | (0.008) | (0.006) | (0.003) | (0.019) | (0.018) |
| 1993 | -0.038 | -0.127 | 0.085 | 0.062 | -0.022 | 0.089 | -0.109 |
| | (0.009) | (0.012) | (0.007) | (0.005) | (0.004) | (0.017) | (0.017) |
| 1994-1998 | -0.050 | -0.141 | 0.088 | 0.066 | -0.055 | 0.104 | -0.086 |
| | (0.010) | (0.018) | (0.013) | (0.010) | (0.006) | (0.022) | (0.024) |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Years of schooling | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean (Dep. Var.) | 0.850 | 0.687 | 0.154 | 0.058 | 0.969 | 0.303 | 0.447 |
| Observations | 82,314 | 82,314 | 82,314 | 82,314 | 82,314 | 12,637 | 12,637 |
| R-squared | 0.136 | 0.257 | 0.198 | 0.078 | 0.057 | 0.253 | 0.328 |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. It reports robust standard errors, adjusted for clustering by village, in parentheses. Regressions use data about returnees from Thailand and from Vietnam and stayers aged 20–60 from the Matched Samples. In columns 6 and 7, the analysis samples are limited to those engaged in the tertiary sector. For definitions of the dependent variables, see the main text. For variable definitions for panel A, see the note to Table 1. In panel B, "19XX(–19XX)" is an indicator variable equal to 1 if individuals returned from Thailand or Vietnam in 19XX(–19XX) and 0 otherwise; stayers are the base group. Base controls include age, age squared, a dummy variable for female, and the logarithmic value of village population aged 20–60 in 1998 (panel A only).



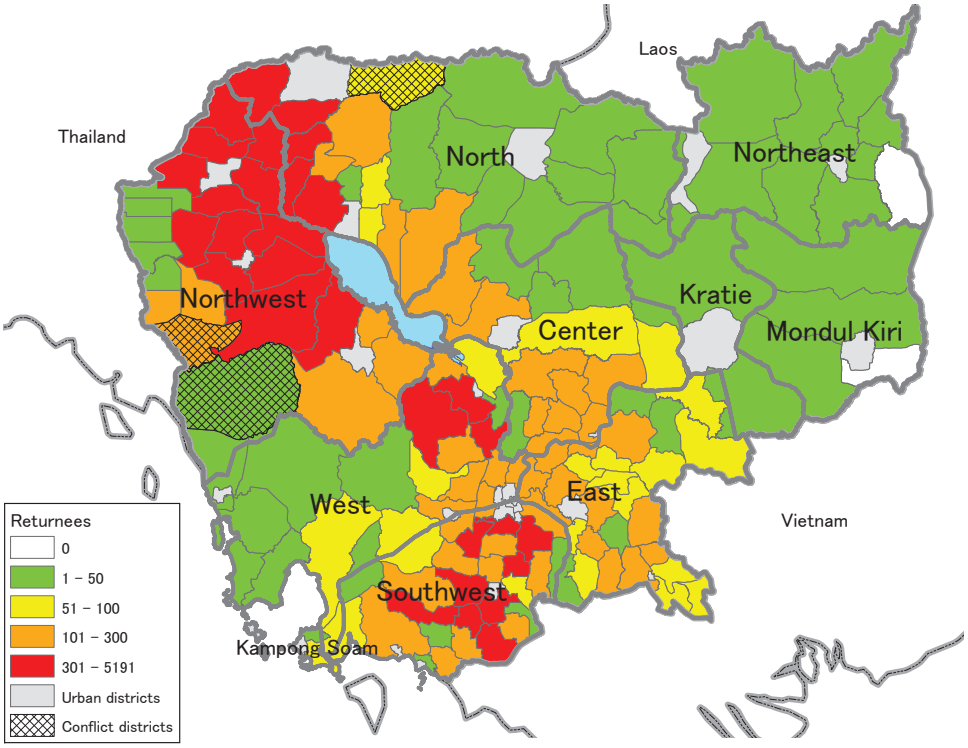

A. Thailand



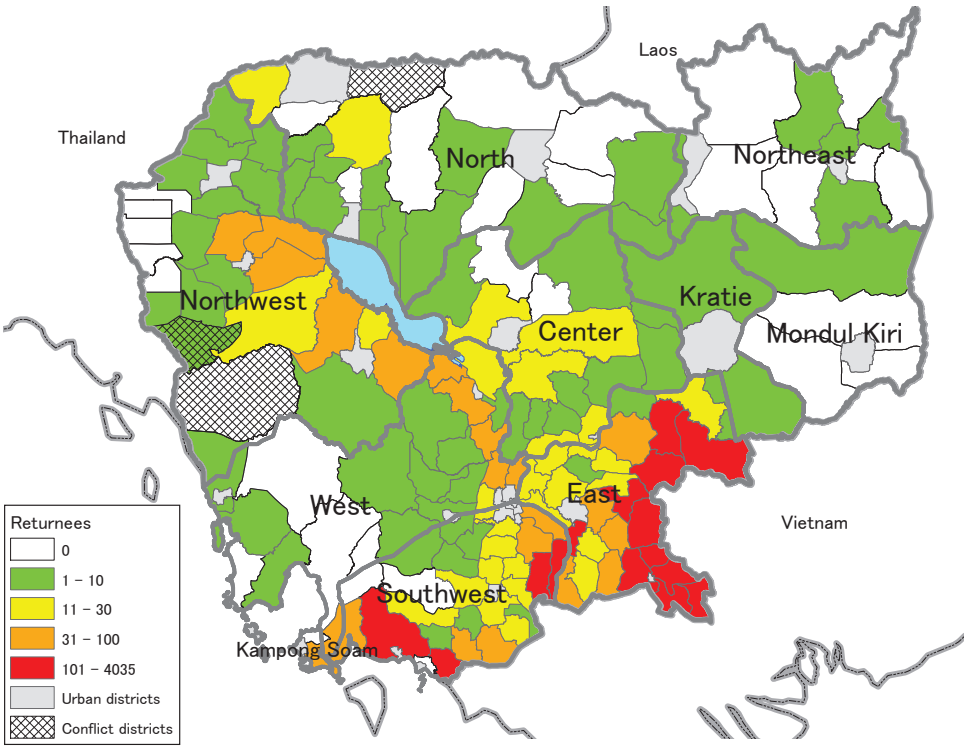

B. Vietnam

Figure 1: Distribution of Migrants from Thailand and Vietnam in 1975-1998

*Note*: The figure shows the distribution of individuals aged 20-60 who previously resided in Thailand (panel A) and Vietnam (panel B) and migrated to their current residence in 1975-1998.




A. Thailand




B. Vietnam

Figure 2: Geographic Distribution of Returnees from Thailand and Vietnam

*Notes*: The figure shows geographic distribution of returnees from Thailand (panel A) and from Vietnam (panel B) aged 20-60 by district of birth in the Full Samples. It also shows the 1977 administrative zones of the Pol Pot regime and the 1998 districts.

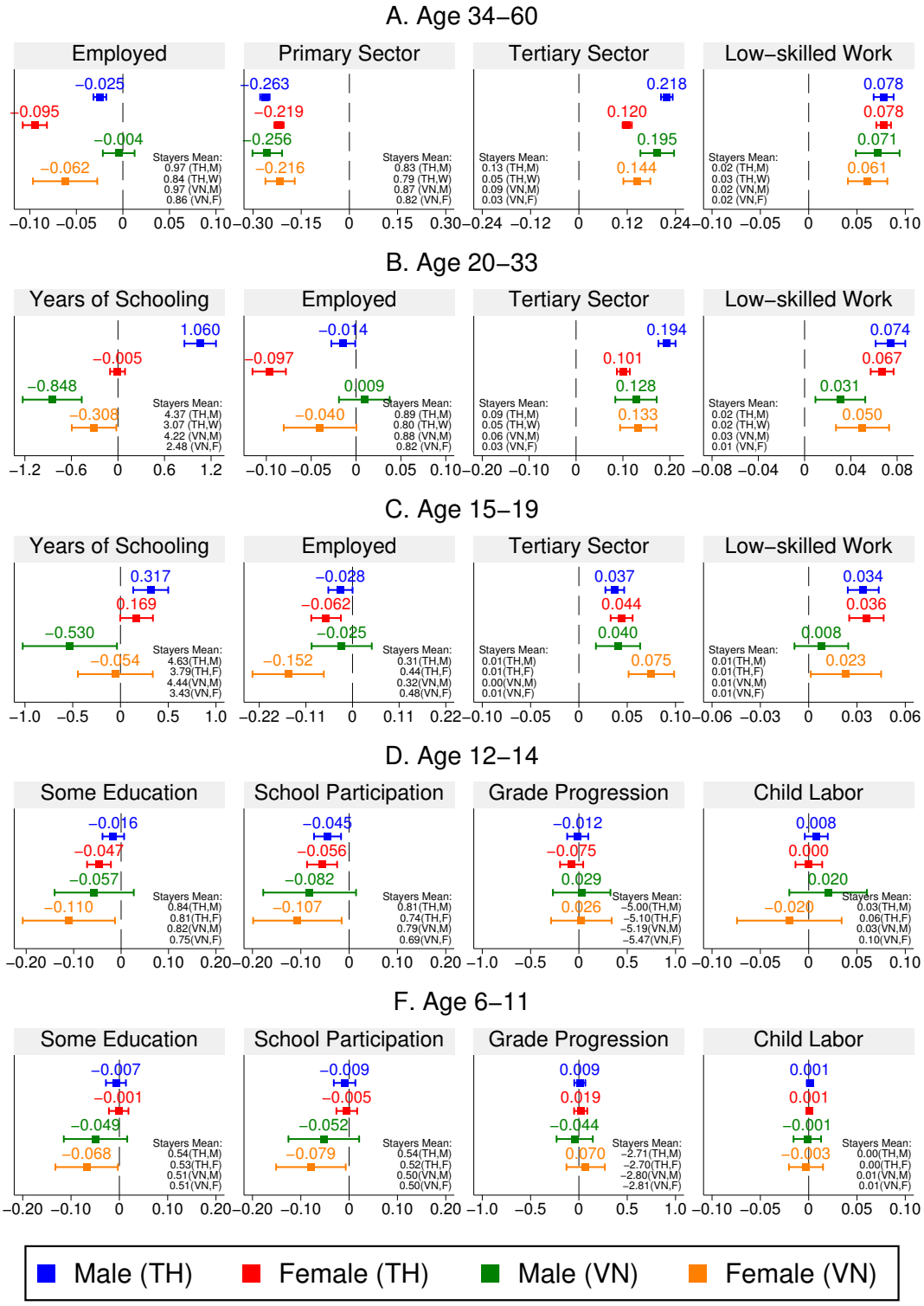



Figure 3: Impacts of Displacement on Educational and Labor Market Outcomes

*Notes*: The figure plots the point estimates of the impacts of forced displacement on educational and labor market outcomes (selected) and their 95% confidence intervals along with the stayers means for male and female returnees aged 20-60 from Thailand (TH) and Vietnam (VN) and for the male and female children aged 6-19 of the male returnees and stayers aged 34-60. The estimates are based on the Matched Samples and are from the bias-corrected version of the nearest-neighbor matching method (Abadie and Imbens 2011). We adjust the 95% confidence intervals using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995).

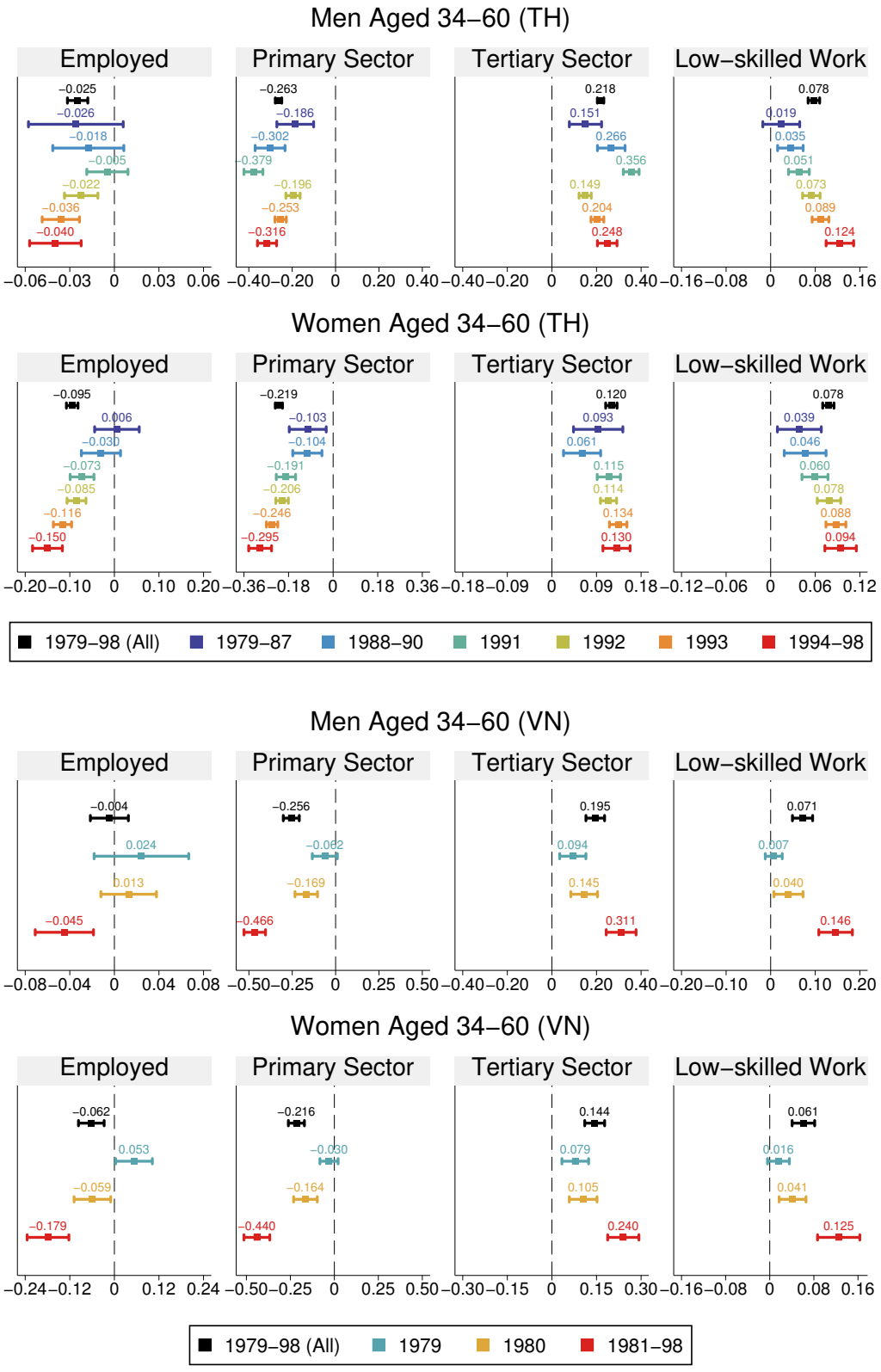



Figure 4: Heterogeneity – Timing of Repatriation (Age 34-60)

*Notes*: The figure plots the point estimates of the impacts of forced displacement on labor market outcomes (selected) and their 95% confidence intervals for the male and female returnees aged 34-60 from Thailand (TH) and Vietnam (VN) who returned in different years. For the estimation method, see the notes to Figure 3 and the main text.

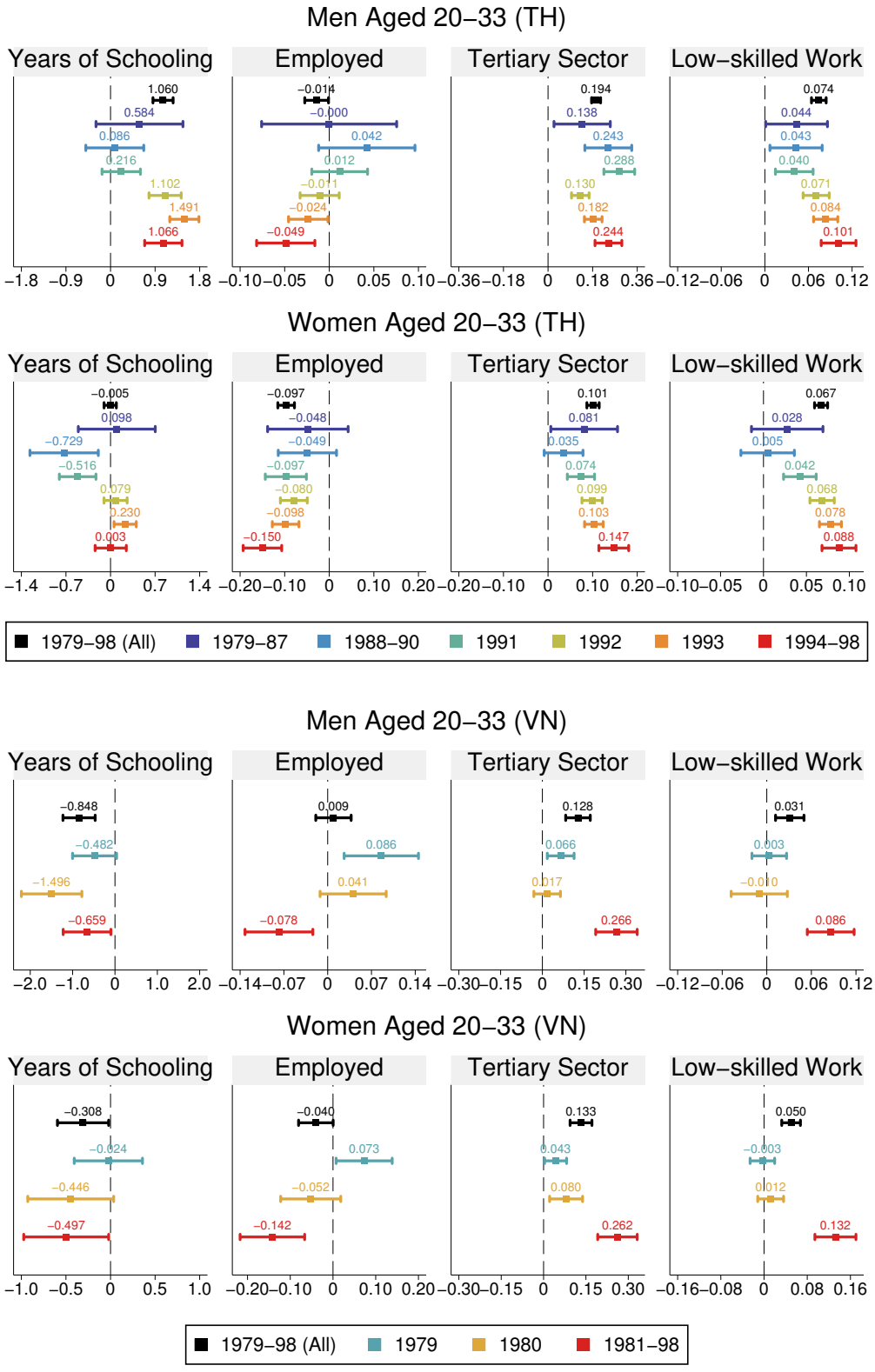



Figure 5: Heterogeneity – Timing of Repatriation (Age 20-33)

*Notes*: The figure plots the point estimates of the impacts of forced displacement on educational and labor market outcomes (selected) and their 95% confidence intervals for the male and female returnees aged 20-33 from Thailand (TH) and Vietnam (VN) who returned in different years. For the estimation method, see the notes to Figure 3 and the main text.

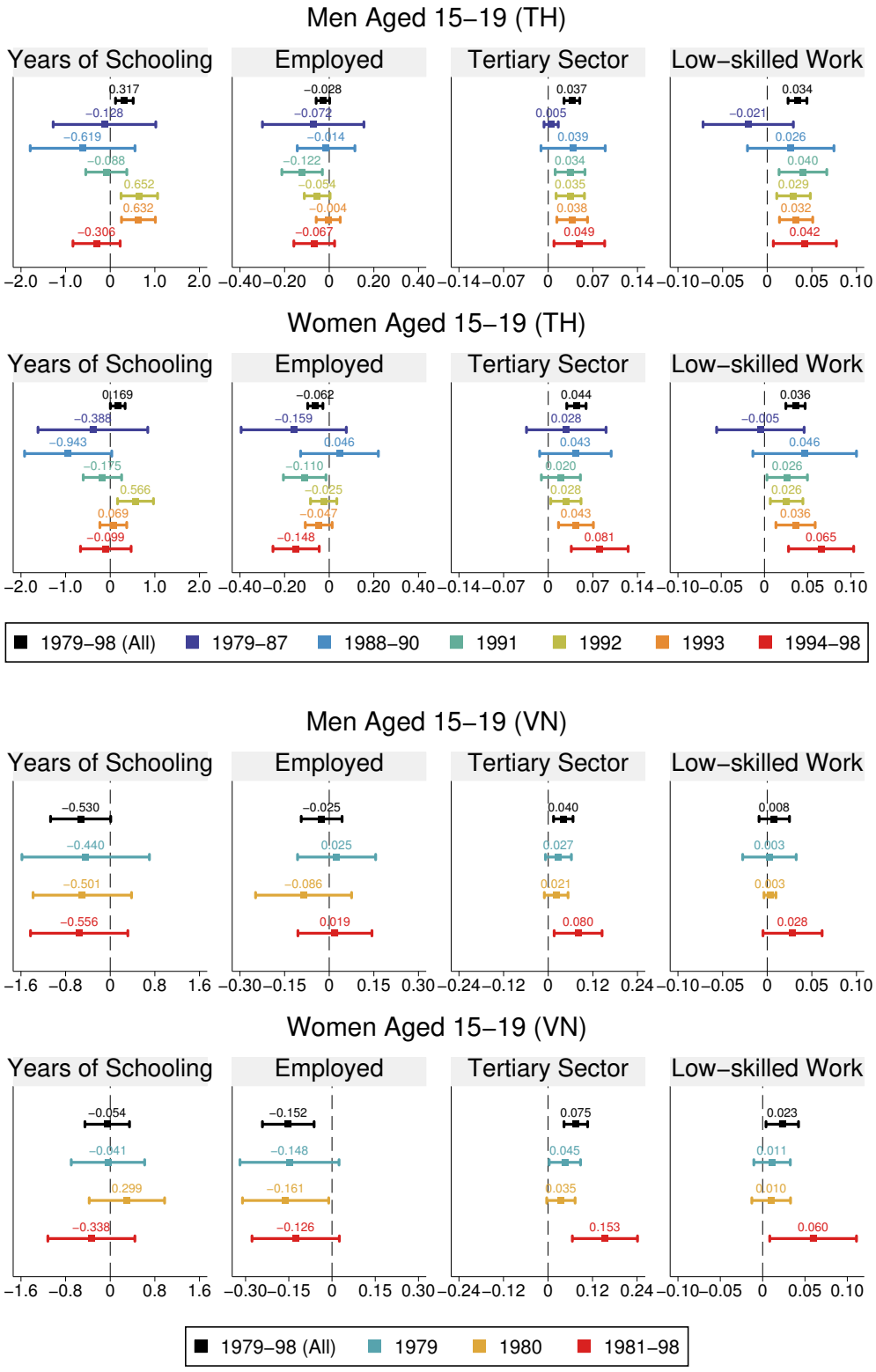



Figure 6: Heterogeneity – Timing of Repatriation (Age 15-19)

*Notes*: The figure plots the point estimates of the impacts of forced displacement on educational and labor market outcomes (selected) and their 95% confidence intervals for the male and female children aged 15-19 of the male returnees aged 34-60 from Thailand (TH) and Vietnam (VN) who returned in different years. For the estimation method, see the notes to Figure 3 and the main text.

## For Online Publication

## Contents



## List of Tables



## List of Figures

## A.1 Historical Background

**Political Context.** After the collapse of the Pol Pot regime, the People's Republic of Kampuchea (PRK), backed by Vietnamese troops, was established in 1979; Vietnam eventually withdrew its forces from Cambodia in 1989. The leaders were former Khmer Rouge cadres—including then-current prime minister Hun Sen, who had defected to Vietnam in 1977–1978—and old revolutionaries who had been in Vietnam during the Pol Pot era. The PRK adopted socialism and respected basic human rights, which had been denied during the Pol Pot era (Vickery 1986); they also reintroduced markets, money currency, Buddhism culture, and formal school education. There was an emphasis on agriculture as the primary industry to rebuild the economy and the introduction of "solidarity groups" (*krom samakki*) consisting of 10–15 families who would work cooperatively and share their production; all land still belonged to the state in 1979–1989. The PRK was supported by Vietnam, the Soviet Union, and other Eastern Bloc countries.

The Khmer Rouge, on the other hand, fled to the Thai/Cambodian border regions following the 1978–1979 Vietnamese invasion. Rebuilding the military, they continued guerilla warfare against the new government (PRK) army until the 1990s. In opposition to the "Vietnamese-installed" PRK, two other political factions also arose: the Khmer People's National Liberation Front (KPNLF) and the National United Front for an Independent, Neutral, Peaceful, and Cooperative Cambodia (FUNCINPEC).[1] The three political factions formed an anti-Vietnamese coalition government, the Coalition Government of Democratic Kampuchea (CGDK), in 1982 and administered the border camps in the 1980s. The CGDK was supported by Thailand, China, and the Western Bloc countries, including the United States, and held Cambodia's seat at the United Nations until 1990.

Following the 1991 Paris Peace Agreement among the four political factions (PRK, KPNLF, FUNCINPEC, and the Khmer Rouge), UNHCR organized a repatriation program between March 30, 1992 and April 30, 1993, and a national election took place under the supervision of the United Nations Transitional Authority in Cambodia (UNTAC) in May 1993, though the Khmer Rouge in the end boycotted the election and refused to demobilize their forces. FUNCINPEC, led by Prince Ranariddh, became the leading party, while the Cambodian People's Party,[2] led by Hun Sen, became the second party. Forming a new coalition government, Prince Ranariddh and Hun Sen became Cambodia's first and second prime ministers. Hun Sen later overthrew Prince

[1] The KPNLF was led by Son Saan, formerly prime minister under Prince Sihanouk from 1967 to 1968, whereas FUNCINPEC was founded by Prince Sihanouk and subsequently led by his son, Prince Ranariddh.

[2] The name of the party changed over time as follows: People's Republic of Kampuchea (1979–1989), the State of Cambodia (1989–1991), and Cambodia People's Party (1991–present).

Ranariddh in a 1997 coup. The Khmer Rouge, meanwhile, continued to fight against the newly elected government, though the leaders defected or were arrested in the late 1990s.

## A.2 Empirical Design

### A.2.1 Full Samples

We check the differences in the observed individual characteristics between refugees and stayers in the Full Samples. Figures A1, A2, and A3 consider the covariate balances for age, education, and district of birth, respectively. Figure A1 provides age distribution for both sexes. Figure A2 plots the point estimates and 95% confidence intervals of the coefficients of years of schooling for cohorts aged 34–60 by age and sex, adjusting for district of birth fixed effects using ordinary least squares (OLS); we presume years of schooling is a variable determined after the refugee crisis (i.e., post-treatment variable) for cohorts aged 20–33. Figure A3 provides geographic distribution by district of birth for both sexes.[3]

We note three distinctions. First, the returnees from Thailand and the stayers have a distinct age distribution (panel A-1 of Figure A1), implying that those aged around 35–40 are most likely to have become refugees. Second, the male returnees from Thailand are more educated than the male stayers (panel A-1 of Figure A2). Third, those born in districts near Thailand and Vietnam are more likely to have fled to Thailand and Vietnam (panels A-1 and B-1 of Figure A3), respectively. For the latter two findings, similar patterns occur in other contexts; more able Bosnians tended to leave the country (Kondylis 2010), while Burundians and Mozambicans in border regions tend to have become refugees (Fransen et al. 2017, Chiovelli et al. 2021).

The following interpretations of these three findings might be plausible in our contexts; while being consistent with anecdotal evidence, they provide further insights into who fled and why they fled. First, those aged around 35–40 might have found it relatively easy to migrate, because many were still single around the collapse of the Pol Pot regime and might have been able to make migration decisions based on their own preferences. Second, given that they were targeted by the Khmer Rouge, better educated people—especially males, due to their relatively high levels of education—might tend to have fled to Thailand and its camps to flee persecution. Third, those who were near the border regions might have found it easier to flee. A relatively large number of people from the Southwest zone also fled to Thailand, because this zone was the heartland of

[3]Men and women display similar patterns for both age and geographic distribution by district of birth; the results showing these are available from the authors upon request.

the Pol Pot regime (Vickery 1999, pp. 93–107) and many might have fled along with the Khmer Rouge or have been forced to flee (Reynell 1989, p.32) as the Khmer Rouge were pushed out by Vietnamese troops. In four districts along the border with Vietnam, a great majority of residents are returnees. This is probably because there was a large-scale purge in the East zone in 1977–1978 (Kiernan 2008), during which many might have fled to Vietnam.

### A.2.2 Matched Samples

We construct matched samples of refugees and stayers with "strong common support" using the Imbens-Rubin approach (Imbens and Rubin 2015). With a large number of stayers from the complete count census data, we construct the matched samples by age, sex, education (for age 34–60), and district of birth (the former two covariates are exactly matched, as are mother tongue and religion); because some pairs are potentially mismatched for education and district of birth, we rely on propensity scores to balance these two covariates and thus not to decrease the number of observations, although our resulting matched samples also nearly exactly match them.[4] Prior to the work, we exclude the returnees from Vietnam whose districts of birth correspond to the four districts bordering Vietnam, mentioned above, (red and white background in panels B-1 and B-2 of Figure A3, respectively), because finding appropriate comparison samples (stayers) is not feasible due to the small number of observations. Their inclusion leads to imbalances in the covariate distributions.

For each subsample by age and sex, we estimate the propensity score using a logistic regression model, with a specification that includes years of schooling, district of birth dummy variables, and a subset of their second-order terms (quadratic and interaction), selected based on the Imbens-Rubin algorithm (Imbens and Rubin 2015, pp. 285–288). For each cohort aged 20–33, we use district of birth dummy variables as a fixed set of covariates. We finally construct a matched sample by matching each returnee unit to the stayer unit with the closest estimated propensity score without replacement. If there are ties, we select a match (a stayer unit) randomly. The resulting matched samples (*Matched Samples*) consist of 36,012 returnees from Thailand, 5,145 returnees from Vietnam, and a corresponding number of stayers.[5][6] Figures A1, A2, and A3 confirm the greatly improved covariate balance in the matched samples.[7]

[4] We confirm that our base results are similar to those derived using exactly balanced data for age, sex, education (for age 34–60), and district of birth; the latter results are available from the authors upon request.

[5] See Appendix Tables A2 and A3 for the descriptive statistics.

[6] Some returnees are also dropped systematically in each propensity score estimation due to the lack of covariate overlap making it impossible to calculate the propensity score.

[7] The distribution of estimated propensity score generally exactly overlaps between returnees and stayers in each

## A.3 Robustness Checks

This section checks the robustness of our main results to alternative estimation methods and threats to identification, as well as external validity.

### A.3.1 Alternative Estimation Methods

Tables A5 and A6 consider the robustness of the results to alternative estimation methods for the returnees aged 20–60 (born prior to the refugee crisis). For comparison purposes, we also report the original results and those from OLS based on the Full Samples. The alternative estimation methods include blocking on the estimated propensity score in combination with regression adjustments within the blocks and an instrumental variables strategy in conjunction with machine-learning techniques (only for the returnees from Thailand). The former uses the Matched Samples and follows the approach proposed by Imbens and Rubin (2015); we construct the blocks based on the Imbens and Rubin algorithm, which gives the optimal number of blocks (Imbens and Rubin 2015, pp. 290-294), and the regressions adjust for the basic set of covariates (age, years of schooling (only for the cohorts aged 34–60), and district of birth fixed effects). The latter uses the Full Samples and follows the approach proposed by Windmeijer et al. (2019) (detailed in the main text). We confirm that the results from the alternative estimation methods, as well as OLS are generally consistent with the original results.

### A.3.2 Threats to Identification

We next consider the robustness of the results to potential threats to identification (we present sensitivity analysis separately below). Since our census data offers limited covariates, unobserved characteristics might not be balanced between returnees and stayers and the resulting estimates might be biased. In particular, since refugees and stayers were exposed to conflict and violence under the Pol Pot regime (Kiernan 2008), the former, who experienced multiple migrations before the refugee event, might have been exposed to them differently; conditioning on district of birth fixed effects might not adequately balance the level of conflict and/or violence to which such individuals were exposed before the refugee event, though refugee status is largely determined by birth place (Figure A3). In this case, the estimated impacts of forced displacement might be contaminated by the impacts of conflict and/or violence under the Pol Pot regime (see Ruiz and Vargas-Silva 2013, pp. 773-774 for relevant discussions).

sample matched by age and sex; the results showing this are available from the authors upon request.

To mitigate this potential concern, we consider two alternative matched samples, constructed as follows. First, using the samples of returnees and all stayers who had their children in 1975-1983 in Cambodia, we reestimate the propensity score, additionally controlling for different patterns of having children in 1975–1983 and for their district of birth;[8] we can further balance local and family-related characteristics and exposure to conflict and/or violence before the refugee event. Second, limiting the above samples to married couples, we reestimate propensity score, further additionally controlling for the characteristics of both husbands and wives (age, education, and district of birth). Although the sample size decreases due to limited overlap, these two alternative matched samples are better balanced between returnees and stayers. Unfortunately, this approach is feasible only for returnees aged 34–60 from Thailand.

Figures A6 and A7 present the results for returnees aged 34-60 from Thailand and for their children aged 6–19, respectively. To compare the results with the original ones, we sequentially report them based on five samples: *Matched Samples* (MS), *Limited Matched Samples I* (LMS-I) (constructed based on the treated units included in both MS and AMS-I (defined next)), *Alternative Matched Samples I* (AMS-I) (constructed based on the first alternative specifications for propensity score), *Limited Matched Samples II* (LMS-II) (constructed based on the treated units included in both MS and AMS-II (defined next)), *Alternative Matched Samples II* (AMS-II) (constructed based on the second alternative specifications for propensity score). For the limited samples, we confirm that the results based on the alternative matched samples are consistent with the original results, based on the baseline specifications of the propensity score.

For the returnees from Thailand, we further check the robustness of the results to another potential threat to identification. As noted, the returnee samples are limited to those who returned to Cambodia, excluding those who resettled in third countries. Indeed, a substantial number of refugees in Thailand eventually moved to third countries, including the United States, through the UNHCR camps; such decisions were made nonrandomly. We have some anecdotal evidence that refugees with families or close relatives in third countries were likely to have been selected for resettlement to facilitate the reunification of refugee families (e.g., Rogge 1990). If returning refugees tend to have lower ability, then the differences in unobserved ability between returnees and stayers might not be adequately balanced and may partly drive our results.

[8]Due to the limited overlap (the impossibility of calculating the propensity score), we can only construct the alternative matched samples for the returnees who had one or two children during this period. For the two groups, we first limit the samples to those who had their child or children born in the same district (58.1% of the returnee samples) and then estimate the propensity score, additionally adjusting when they had children.

To address this potential concern, we separately consider the returnees from the UNHCR and UNBRO camps. The sample selection problem, if any, will be limited in the latter samples. With no direct information to identify those returnees, we assume that returnees aged 34–60 from Thailand, whose children were born in Thailand in 1975–1983 and in Cambodia in 1980-1983, to be from the UNHCR and UNBRO camps, respectively. This exploits the fact that those who fled before early 1980 basically stayed in the UNHCR camps, located in Thailand, while those who fled afterward stayed in the UNBRO camps, located in Cambodia, until 1983, due to the Thai government's change in refugee policy in early 1980. The number of returnees from the UNHCR and UNBRO camps is 5,522 and 3,932, respectively; analysis is not feasible for the returnees aged 20–33 because they were generally single in 1975–1983. Figures A8 and A9 present the results for the returnees from each camp and those who returned from each camp in 1992–1993. We confirm that the results for returnees from the UNBRO camps are consistent with the original results.

### A.3.3 Threats to External Validity

We next consider the robustness of the results to potential threats to external validity in the context of Cambodia. As noted above, one limitation of our samples is that returning refugees are limited to those who did not migrate again after repatriation. Some returnees might have migrated after repatriation to find better jobs or to join family members or relatives (Robinson 1994), and displacement impacts might systematically differ for this group. To address this potential concern, we consider alternative samples based on a different definition of refugees, though this is feasible only for returnees aged 34–60 from Thailand. We define returning refugees from Thailand as those whose household members were born in Thailand after 1975.[9] We construct the alternative matched samples following the same procedures. Importantly, these samples include returnees who experienced multiple migrations after repatriation.

Figures A10 and A11 report the results for the returnees aged 34–60 from Thailand and for their children aged 6–19, respectively. To compare these results with the originals, we sequentially report them based on four samples: *Matched Samples* (MS), *Alternative Matched Samples III* (AMS-III) (constructed based on the alternative definition of refugee), *Limited Alternative Matched Samples III-A* (LAMS-III-A) (constructed based on the treated units included in both MS and AMS-III), and *Limited Alternative Matched Samples III-B* (LAMS-III-B) (constructed based on

[9]This sample construction is consistent with our base samples for age 34–60 and can provide us with more observations. We also confirm similar results when limiting the samples to those who had children (conditioning on parent-child relationships); the results are available from the authors upon request.

the treated units included in AMS-III, but not in MS). The returnees in LAMS-III-B are those who experienced multiple migrations after repatriation. We confirm that the estimated adverse impacts are generally relatively strong for these returnees, implying that our original results serve as conservative estimates of the displacement impacts: Returnees with multiple migrations experienced far worse educational, labor market, and home ownership outcomes.

### A.3.4 Sensitivity Analysis

**Matched Samples.** Following Rosenbaum's approach for matched pairs (Rosenbaum 2002), Tables A5 and A6 assess the robustness of our main estimates to hidden bias to address concerns about remaining omitted variable bias. Assuming an unobserved binary confounder, we rely on the sensitivity parameter, $\Gamma$, or the ratio of the respective odds of being a refugee for two matched units. This parameter also determines an lower and upper bound on inference quantities, $p$-values, for the null hypothesis of no displacement impact to reflect uncertainty due to hidden bias, manipulating a range of probabilities of "success," or the occurrence of the event that an outcome value differs between two matched units.

For example, if $\Gamma = 1$, then the two matched units have the same chance of being refugees, as in a randomized experiment (a study free from hidden bias); the lower and upper bounds of the probabilities of success, defined by $1/(1+\Gamma)$ and $\Gamma/(1+\Gamma)$, respectively, for the case where an outcome value differs between them, are both an equal $1/2$, which leads to the usual significance level (i.e., no range of significance level due to no hidden bias). If $\Gamma = 2$, then one may be twice as likely as another to have been a refugee due to an unobserved confounder (although the two units appear similar in terms of the observed covariates); the lower and upper bounds of the probabilities of success then become $1/3$ and $2/3$, respectively, thereby leading to a range of significance levels that reflect uncertainty due to the hidden bias resulting from the unobserved confounder. As $\Gamma$ increases, the interval becomes wider and eventually uninformative, with large $p$-values; the null hypothesis is less likely to be rejected. As such, our estimates are sensitive if the original conclusions change for a $\Gamma$ just barely larger than 1. We can report a $\Gamma$ corresponding to maximum multiplicity-adjusted $p$-values less than 0.05, and can thus confirm that our key findings are generally robust to hidden bias.

**Full Samples.** To assess the sensitivity of the results based on the Full Samples, we conduct a sensitivity analysis, following the approach proposed by Oster (2019). This approach considers both coefficient movements and $R$-squared movements when covariates are included, assuming that

both observed and unobserved covariates explain the same amount of variability (variance) in the outcome variable in a regression model. Based on Oster's results, we consider the following bias-adjusted impacts of forced displacement:

$$\gamma^* \approx \hat{\gamma}_{BS} - \delta(\hat{\gamma}_{RS} - \hat{\gamma}_{BS})\frac{R^2_{max} - R^2_{BS}}{R^2_{BS} - R^2_{RS}},$$

where $\hat{\gamma}_{BS}$ and $\hat{\gamma}_{RS}$ are the estimates based on the baseline specification (age, age squared, years of schooling (only for age 34-60), district of birth fixed effects) and restricted specification (no covariates), respectively.

Our interest is in how coefficient estimates change due to unobserved confounders when we adjust the covariates. $\delta$ is the proportional degree of selection. Following Oster's suggestion, we assume equal selection: The ratio of the coefficient movement is the same as that of the $R$-squared movement. To allow for the over- or underestimation of the true displacement impacts, we consider two cases, $\delta = 1$ and $\delta = -1$: The former assumes the same amount of selection going in the same direction, whereas the latter assumes the same amount of selection going in the opposite direction. $R^2_{BS}$ and $R^2_{RS}$, respectively, are the $R$-squared from the baseline specification and the restricted specification. $R^2_{max}$ is the $R$-squared from a regression that controls for all observed and unobserved covariates. Although $R^2_{max}$ is unobserved, we know that $R^2_{max}$ is bounded by the upper bound 1 ($R^2_{max} = 1$), which gives the most conservative estimate of the displacement impacts, $\gamma^*$. An $R^2_{max}$ below 1 should be considered in empirical works, based on Oster's recommendation (she derives a cutoff value of 1.3 as a multiplier for the $R$-squared from restricted regression models).

We consider two cases: (1) $R^2_{max} = 1.3 \times R^2_{BS}$ and (2) $R^2_{max} = 1$ (see Table A7). In the first case, ($R^2_{max} = 1.3 \times R^2_{BS}$), regardless of the direction of the unobserved selection, the displacement impacts are consistent with those in Tables A5 and A6. For some outcomes, this holds true even for the most conservative case ($R^2_{max} = 1$). These results suggest that omitted variable bias is unlikely to be significant enough to alter our conclusions.

## A.4 Auxiliary Analyses

This section considers individual and regional characteristics correlated with returnees' early return migration decisions (only for returnees from Thailand) and their return migration decisions to go back to birth regions.

### A.4.1 Early Return Migration Decisions

Focusing on returnees aged 34–60 from Thailand among the Matched Samples, we examine individual and regional characteristics correlated with their early return migration decisions. We focus on this age cohort because they seem to have been old enough to make migration decisions independently at the time. We estimate the following equation using OLS with robust standard errors clustered by district of birth:

$$\begin{aligned} Y_{idp} = \alpha &+ X_i'\delta + \beta_1 ln(Distance\,to\,Thai\,border)_d + \beta_2 Prop\,of\,contamination_d \\ &+ \beta_3 ln(Areas)_d + \lambda_p + \epsilon_{idp}, \end{aligned}$$

where $Y_{idp}$ is an indicator variable equal to 1 if returnee $i$ born in district $d$, province $p$, returned to Cambodia before 1992 and 0 otherwise, $X_i$ is a vector of individual characteristics (age, a female dummy, years of schooling), $ln(Distance\,to\,Thai\,border)_d$ is the logarithmic value of the shortest distance ($km$) from the centroid of district $d$ to the Thai border, $Prop\,of\,contamination_d$ is the proportion of contaminated areas in the total village buffer zone areas (3.0 km radius) in district $d$, $ln(Areas)_d$ is the logarithmic value of the total village buffer zone areas (3.0 km radius) in district $d$, and $\lambda_p$ is province of birth fixed effects. We create $ln(Distance\,to\,Thai\,border)_d$ and $ln(Areas)_d$ using a GIS.

Table A8 reports the results. Column 1 reports results without adjusting for province of birth fixed effects. We find that younger, male, and worse educated refugees, those from districts away from border regions, and those from districts more contaminated with UXOs and landmines tend to be have become early returnees. When we additionally adjust for province of birth fixed effects in column 2, the results are similar to those in column 1, though the significant difference for $ln(Distance\,to\,Thai\,border)_d$ becomes weak, as expected. Columns 3 and 4 limit the sample to the returnees from the UNHCR and UNBRO camps whom we can identify (see Section A.3.2). After adding an indicator variable for those from the UNHCR camps ($UNHCR$), we do the same exercises as in columns 1 and 2 and additionally find that refugees in the UNHCR camps are more likely to be have become late returnees.

### A.4.2 Return Migration Decisions to Birth Regions

We next examine individual and regional characteristics correlated with return migration decisions to birth districts for returnees aged 34–60 from Thailand and Vietnam in the Matched Samples. Columns 1 and 2 and columns 5 and 6 of Table A10 estimate the above equation,

replacing the dependent variable with an indicator variable for the returnees from Thailand and Vietnam, respectively, who returned to their birth districts. Columns 3 and 4 and columns 7 and 8 focus on the returnees from Thailand and Vietnam who returned in 1992–1993 and 1979–1980, respectively. For the returnees from Vietnam, given that the majority returned in 1979–1980, we do not consider $Prop\,of\,contamination_d$ and $ln(Areas)_d$. Also, in columns 5–9, we consider $ln(Distance\,to\,Vietnam\,border)_d$, the logarithmic value of the shortest distance ($km$) from the centroid of district $d$ to the Vietnamese border, instead of $ln(Distance\,to\,Thai\,border)_d$. For both groups of returnees, we find that less educated refugees and those from districts near border regions tend to have returned to their birth districts.

## A.5 Mechanisms

### A.5.1 Consistency

In terms of its consistency with the 1998 Census data, we have two concerns about the 2004 CSES data. First, the refugee status might be contaminated, because these data can only identify those who have lived abroad. To address this concern, we limit the samples to those who live in the former Northwest, West, and North zones and in the Southwest and East zones. Since the great majority of returnees from Thailand and Vietnam live in these regions (see Figure A3), such limited samples should largely capture the returnees from Thailand and Vietnam. Second, with no information about districts of birth, our regression analysis based on the 2004 CSES data cannot adjust for district of birth fixed effects. To address this point, using the Full Samples of the returnees from Thailand, returnees from Vietnam, and stayers aged 20–60, Appendix Table A11 considers the displacement impacts on nine outcomes (evaluated below) without controlling for district of birth fixed effects via OLS, confirming that the results are generally consistent with the base results.

Using the 2004 CSES data, Appendix Table A12 considers the further long-term displacement impacts on the nine educational, labor market, and home ownership outcomes. While the magnitudes of some coefficients differ, the results are qualitatively similar to those in Appendix Table A11. This suggests that the adverse displacement impacts on labor market and home ownership outcomes lasted into 2004. Given the consistency between the 1998 Census data and the 2004 CSES data, we examine the potential mechanisms underlying these results.

### A.5.2 Other Potential Channels

Table A15 considers other potential channels, including discrimination, health, and remittance networks.

**Discrimination.** As in other forced migration situations, there is the possibility that discrimination against returnees might drive the adverse displacement impacts (see, e.g., Brell et al. 2020 for relevant discussion). With no direct information about discrimination, we consider this possibility by examining the relationship with neighborhood trust; the variable for neighborhood trust is an indicator variable equal to 1 if household heads feel safe from crime and violence in their neighborhood (*Neighborhood Trust*). If the discrimination channel exists, then returnees might exhibit lower levels of trust in neighbors; we can expect to find a negative relationship (see, e.g., Smith 2010 between trust and neighborhood). Panel A of Table A15 looks at this relationship. The results show no statistically significant relationships among the three samples.[10] Thus, the discrimination channel is unlikely to drive our results.

**Health.** We next consider whether the adverse displacement impacts are driven by poor health for returnees (see, e.g., Currie and Madrian 1999 for the relationship between health and labor market outcomes). This mechanism could be likely in our context because most survivors of the Pol Pot regime suffered from long-term mental health disorders, such as post-traumatic stress disorder (PTSD) (e.g., Beth et al. 2011), and health conditions might have deteriorated in forced displacement or camps, or fearful individuals might tend to have been refugees (Marshall et al. 2005). Also, returnees might be more likely to have fallen victim to landmines.

Panel B of Table A15 examines the relationship with three health outcomes: an indicator variable equal to 1 if individuals report that their health is "very good" or "good" and 0 otherwise (*Health Status I*), an indicator variable equal to 1 if individuals report that their health is "much better" or "somewhat better" than others of the same age and 0 otherwise (*Health Status II*), and an indicator variable equal to 1 if individuals have any disability and 0 otherwise (*Disability*). Again, none of these variables have statistically significant relationships in the three samples. Thus, the health channel is also unlikely to drive our results.

[10] As an alternative robustness check, focusing on stayers, we also examine the relationship between the proportion of returning refugees in villages and neighborhood trust; the variable for the proportion of returnees is a village-level measure constructed based on returnees from Thailand or Vietnam and stayers aged 20–60 from the 1998 Census data. In the analyses, we also adjust for the village population aged 20–60. If the discrimination channel exists, then stayers living in villages with a larger proportion of returnees might exhibit lower levels of trust in neighbors. Again, we can confirm no statistically significant relationships among the three samples. These results are available from the authors upon request.

**Remittance Networks.** Lastly, we consider whether the adverse displacement impacts are driven by the difference in remittance networks between returnees and stayers. Due to forced displacement, returnees might have different social networks, having built new social networks and/or disrupted existing ones (see Sarvimäki et al. 2022 for relevant discussions). In our context, in particular, since returnees from Thailand stayed in camps for a long time, they might tend toward better remittance networks (e.g., relatives or friends abroad; however, worse remittance networks could also happen) and thus might tend to receive more remittances. In this case, they might also lack motivation to work (Chami et al. 2005). To examine this possibility, panel C of Table A15 looks at the relationship with four outcomes related to domestic and international remittances: an indicator variable equal to 1 if households receive remittances from relatives or others in Cambodia (or from abroad) (*Receipt*) and the remittance amount in dollars from relatives or others in Cambodia (or from abroad) (*Amount (USD)*). Again, we find no evidence that the difference in motivation to work drives our results.

## References


**Abadie, Alberto and Guido W. Imbens**, "Bias-Corrected Matching Estimators for Average Treatment Effects," *Journal of Business and Economic Statistics*, 2011, *29* (1), 1–11.

**Benjamini, Yoav and Yosef Hochberg**, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1995, *57* (1), 289–300.

**Beth, V. Schaak, Reicherter Daryn, and Chhang Youk**, *Cambodia's Hidden Scars: Trauma Psychology in the Wake of the Khmer Rouge*, Phnom Penh: Documentation Center of Cambodia, 2011.

**Brell, Courtney, Christian Dustmann, and Ian Preston**, "The Labor Market Integration of Refugee Migrants in High-Income Countries," *Journal of Economic Perspectives*, 2020, *34* (1), 94–121.

**Chami, Ralph, Connel Fullenkamp, and Samir Jahjah**, "Are Immigrant Remittance Flows a Source of Capital for Development?," *IMF Staff Papers*, 2005, *52* (1), 55–81.

**Chiovelli, Giorgio, Stelios Michalopoulos, Elias Papaioannou, and Sandra Sequeira**, "Forced Displacement and Human Capital: Evidence from Separated Siblings," 2021. Working Paper 29589, National Bureau of Economic Research.

**Currie, Janet and Brigitte C. Madrian**, "Health, Health Insurance and The Labor Market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3, Elsevier, 1999, chapter 50, pp. 3309–3416.

**Fransen, Sonja, Isabel Ruiz, and Carlos Vargas-Silva**, "Return Migration and Economic Outcomes in the Conflict Context," *World Development*, 2017, *95*, 196–210.

**Imbens, Guido W. and Donald B. Rubin**, *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, New York: Cambridge University Press, 2015.

**Kiernan, Ben**, *The Pol Pot Regime: Race, Power, and Genocide in Cambodia under the Khmer Rouge, 1975-79*, 3rd ed., London: Yale University Press, 2008.

**Kondylis, Florence**, "Conflict Displacement and Labor Market Outcomes in Post-War Bosnia and Herzegovina," *Journal of Development Economics*, 2010, *93* (2), 235–248.

**Marshall, Grant N., Terry L. Schell, Marc N. Elliott, Megan Berthold, and Chi-Ah Chun**, "Mental Health of Cambodian Refugees 2 Decades After Resettlement in the United States," *Journal of American Medical Association*, 2005, *294* (5), 571–579.

**Oster, Emily**, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

**Reynell, Josephine**, *Political Pawns: Refugees on the Thai-Kampuchean Border*, Oxford: Refugee Studies Programme, 1989.

**Robinson, W. Courtland**, ""Something Like Home Again" the Repatriation of Cambodian Refugee," 1994. The U.S. Committee for Refugees, Washington D.C.

**Rogge, John R.**, "Return to Cambodia: The Significance and Implications of Past, Present and Future Spontaneous Repatriations," 1990. The Intertect Institute.

**Rosenbaum, Paul R.**, *Observational Studies*, 2nd ed., New York: Springer, 2002.

**Ruiz, Isabel and Carlos Vargas-Silva**, "The Economics of Forced Migration," *Journal of Development Studies*, 2013, *49* (6), 772–784.

**Sarvimäki, Matti, Roope Uusitalo, and Markus Jäntti**, "Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations," *Journal of the European Economic Association*, 2022, *20* (6), 2497–2539.

**Smith, Sandra Susan**, "Race and Trust," *Annual Review of Sociology*, 2010, *36*, 453–475.

**Vickery, Michael**, *Kampuchea: Politics, Economics and Society*, 1st ed., London: Rances Pinter, 1986.

_ , *Cambodia 1975-1982*, Boston: South End Press, 1999.

**Windmeijer, Frank, Helmut Farbmacher, Neil Davies, and George Davey Smith**, "On the Use of the Lasso for Instrumental Variables Estimation with Some Invalid Instruments," *Journal of the American Statistical Association*, 2019, *114* (527), 1339–1350.

Table A1: Construction of Full Samples

| No. | Description of Conditions | Observations |
|---|---|---|
| (0) | Total number of individuals in the complete set of 1998 Census microdata | 11,435,097 |
| (1) | They are 20–60 years old. | 4,650,448 |
| (2) | They speak Khmer (Cambodian) as their mother tongue. | 4,432,053 |
| (3) | They are Buddhists. | 4,396,895 |
| (4) | They were born in rural areas in Cambodia. | 4,000,123 |
| (5) | "Employment period" is not missing, whether they are employed or unemployed. | 3,991,582 |
| (6) | The highest grade of school that they completed is not missing or other. | 3,986,991 |
| (7) | The information on housing (light, fuel, water, and toilet) conditions is not missing. | 3,878,455 |
| | Returnees from Thailand | |
| (8) | They were born in Cambodia and previously resided in Thailand. | 37,023 |
| (9) | The information about their "duration of stay" in their current residence is not missing. | 36,853 |
| (10) | Duration of stay in their current residence is less than 20 years. | 36,760 |
| | Returnees from Vietnam | |
| (8) | They were born in Cambodia and previously resided in Vietnam. | 15,669 |
| (9) | The information about their "duration of stay" in their current residence is not missing. | 15,619 |
| (10) | Duration of stay in their current residence is less than 20 years. | 15,548 |
| (11) | They were born in districts other than Chantrea, Kampong Rou, Rumduol, and Svay Teab in Svay Rieng. | 5,219 |
| | Stayers | |
| (8) | They have never migrated before. | 1,968,687 |

*Notes:* The table reports the detailed procedures for developing the base sample used for our analysis. The sample is constructed from the complete set of the 1998 Census microdata, with 11,435,097 individuals. Columns 2 and 3 describe the conditions and the number of individuals that satisfy these conditions, respectively.

Table A2: Descriptive Statistics (Mean) – Age 20-60

| | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
| | Full Sample | | Matched Sample | | Full Sample | | Matched Sample | |
| Group: | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer |
| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | | | | A-1. Age 34-60 (TH) | | | | |
| Age | 42.678 | 44.467 | 42.555 | 42.555 | 42.675 | 44.532 | 42.625 | 42.625 |
| Years of schooling | 4.217 | 3.788 | 4.197 | 4.242 | 1.892 | 1.888 | 1.884 | 1.908 |
| Labor force | 0.973 | 0.982 | 0.973 | 0.979 | 0.788 | 0.880 | 0.789 | 0.865 |
| Employed | 0.943 | 0.974 | 0.942 | 0.966 | 0.745 | 0.865 | 0.746 | 0.843 |
| Months worked | 10.409 | 10.741 | 10.402 | 10.630 | 7.847 | 9.217 | 7.865 | 9.037 |
| Primary sector | 0.572 | 0.875 | 0.570 | 0.831 | 0.574 | 0.822 | 0.575 | 0.795 |
| Secondary sector | 0.044 | 0.013 | 0.044 | 0.015 | 0.016 | 0.009 | 0.016 | 0.007 |
| Tertiary sector | 0.341 | 0.090 | 0.342 | 0.126 | 0.171 | 0.040 | 0.171 | 0.052 |
| High-skilled work | 0.058 | 0.041 | 0.057 | 0.062 | 0.014 | 0.007 | 0.013 | 0.009 |
| Middle-skilled work | 0.724 | 0.914 | 0.723 | 0.879 | 0.642 | 0.845 | 0.643 | 0.820 |
| Low-skilled work | 0.099 | 0.016 | 0.100 | 0.020 | 0.103 | 0.018 | 0.103 | 0.025 |
| Armed forces | 0.075 | 0.006 | 0.075 | 0.011 | 0.002 | 0.000 | 0.002 | 0.000 |
| Home ownership | 0.959 | 0.991 | 0.960 | 0.990 | 0.956 | 0.990 | 0.956 | 0.990 |
| Observations | 11,324 | 362,881 | 10,954 | 10,954 | 13,057 | 598,647 | 12,819 | 12,819 |
| | | | | B-1. Age 34-60 (VN) | | | | |
| Age | 45.360 | 44.467 | 44.788 | 44.788 | 45.115 | 44.532 | 44.654 | 44.654 |
| Years of schooling | 4.159 | 3.788 | 3.565 | 3.562 | 1.629 | 1.888 | 1.498 | 1.491 |
| Labor force | 0.983 | 0.982 | 0.971 | 0.975 | 0.898 | 0.880 | 0.814 | 0.875 |
| Employed | 0.976 | 0.974 | 0.958 | 0.969 | 0.890 | 0.865 | 0.798 | 0.858 |
| Months worked | 10.893 | 10.741 | 11.148 | 11.102 | 9.562 | 9.217 | 9.095 | 9.639 |
| Primary sector | 0.752 | 0.875 | 0.614 | 0.868 | 0.785 | 0.822 | 0.608 | 0.824 |
| Secondary sector | 0.032 | 0.013 | 0.069 | 0.009 | 0.007 | 0.009 | 0.014 | 0.006 |
| Tertiary sector | 0.195 | 0.090 | 0.279 | 0.093 | 0.100 | 0.040 | 0.180 | 0.034 |
| High-skilled work | 0.040 | 0.041 | 0.036 | 0.046 | 0.004 | 0.007 | 0.006 | 0.004 |
| Middle-skilled work | 0.877 | 0.914 | 0.822 | 0.902 | 0.850 | 0.845 | 0.715 | 0.842 |
| Low-skilled work | 0.041 | 0.016 | 0.087 | 0.017 | 0.037 | 0.018 | 0.080 | 0.018 |
| Armed forces | 0.020 | 0.006 | 0.018 | 0.006 | 0.001 | 0.000 | 0.001 | 0.001 |
| Home ownership | 0.981 | 0.991 | 0.960 | 0.987 | 0.986 | 0.990 | 0.976 | 0.989 |
| Observations | 4,226 | 362,881 | 1,526 | 1,526 | 4,850 | 598,647 | 1,630 | 1,630 |
| | | | | A-2. Age 20-33 (TH) | | | | |
| Age | 27.557 | 26.122 | 27.555 | 27.555 | 27.708 | 26.385 | 27.703 | 27.703 |
| Some education | 0.823 | 0.783 | 0.823 | 0.774 | 0.635 | 0.657 | 0.635 | 0.646 |
| Primary school | 0.472 | 0.383 | 0.472 | 0.363 | 0.201 | 0.207 | 0.202 | 0.201 |
| Years of schooling | 5.421 | 4.518 | 5.417 | 4.373 | 3.048 | 3.140 | 3.059 | 3.067 |
| Labor force | 0.931 | 0.928 | 0.931 | 0.932 | 0.754 | 0.872 | 0.755 | 0.840 |
| Employed | 0.875 | 0.886 | 0.875 | 0.893 | 0.694 | 0.834 | 0.695 | 0.796 |
| Months worked | 9.422 | 9.651 | 9.428 | 9.543 | 7.214 | 8.818 | 7.233 | 8.417 |
| Primary sector | 0.562 | 0.807 | 0.562 | 0.798 | 0.545 | 0.786 | 0.544 | 0.753 |
| Secondary sector | 0.047 | 0.014 | 0.048 | 0.015 | 0.020 | 0.012 | 0.020 | 0.007 |
| Tertiary sector | 0.280 | 0.069 | 0.281 | 0.087 | 0.146 | 0.042 | 0.147 | 0.047 |
| High-skilled work | 0.052 | 0.024 | 0.052 | 0.029 | 0.017 | 0.012 | 0.017 | 0.012 |
| Middle-skilled work | 0.689 | 0.842 | 0.689 | 0.838 | 0.606 | 0.812 | 0.606 | 0.776 |
| Low-skilled work | 0.096 | 0.018 | 0.097 | 0.022 | 0.084 | 0.016 | 0.085 | 0.018 |
| Armed forces | 0.052 | 0.006 | 0.052 | 0.010 | 0.003 | 0.000 | 0.003 | 0.000 |
| Home ownership | 0.936 | 0.979 | 0.936 | 0.981 | 0.939 | 0.977 | 0.938 | 0.980 |
| Observations | 5,318 | 423,886 | 5,260 | 5,260 | 7,061 | 583,273 | 6,979 | 6,979 |
| | | | | B-2. Age 20-33 (VN) | | | | |
| Age | 26.771 | 26.122 | 26.625 | 26.625 | 26.819 | 26.385 | 26.928 | 26.928 |
| Some education | 0.794 | 0.783 | 0.570 | 0.733 | 0.636 | 0.657 | 0.450 | 0.545 |
| Primary school | 0.499 | 0.383 | 0.278 | 0.365 | 0.230 | 0.207 | 0.140 | 0.144 |
| Years of schooling | 5.286 | 4.518 | 3.371 | 4.220 | 3.213 | 3.140 | 2.174 | 2.483 |
| Labor force | 0.959 | 0.928 | 0.944 | 0.920 | 0.912 | 0.872 | 0.830 | 0.848 |
| Employed | 0.931 | 0.886 | 0.888 | 0.878 | 0.888 | 0.834 | 0.781 | 0.822 |
| Months worked | 10.309 | 9.651 | 10.318 | 10.097 | 9.496 | 8.818 | 8.907 | 9.251 |
| Primary sector | 0.777 | 0.807 | 0.648 | 0.811 | 0.797 | 0.786 | 0.602 | 0.786 |
| Secondary sector | 0.027 | 0.014 | 0.064 | 0.014 | 0.008 | 0.012 | 0.019 | 0.008 |
| Tertiary sector | 0.132 | 0.069 | 0.188 | 0.061 | 0.086 | 0.042 | 0.164 | 0.031 |
| High-skilled work | 0.025 | 0.024 | 0.015 | 0.024 | 0.008 | 0.012 | 0.012 | 0.009 |
| Middle-skilled work | 0.872 | 0.842 | 0.817 | 0.832 | 0.856 | 0.812 | 0.709 | 0.803 |
| Low-skilled work | 0.026 | 0.018 | 0.057 | 0.026 | 0.026 | 0.016 | 0.064 | 0.014 |
| Armed forces | 0.013 | 0.006 | 0.011 | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 |
| Home ownership | 0.974 | 0.979 | 0.956 | 0.979 | 0.976 | 0.977 | 0.959 | 0.973 |
| Observations | 2,888 | 423,886 | 971 | 971 | 3,584 | 583,273 | 1,018 | 1,018 |

*Notes:* The table shows the means of variables for returnees from Thailand (TH) and Vietnam (VN) and for stayers aged 20-60 in the Full Samples and Matched Samples. For the variable definitions, see the main text.

Table A3: Descriptive Statistics (Mean) – Age 6-19

| | Age 34-60 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | | Women | | | | | |
| | Both Sexes | | Boys | | Girls | | Both Sexes | | Boys | | Girls | |
| Group: | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer |
| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | A-1. Age 15-19 (TH) | | | | | | | | | | | |
| Some education | 0.826 | 0.807 | 0.865 | 0.849 | 0.788 | 0.770 | 0.813 | 0.789 | 0.854 | 0.836 | 0.776 | 0.744 |
| School participation | 0.436 | 0.439 | 0.517 | 0.532 | 0.353 | 0.340 | 0.406 | 0.407 | 0.482 | 0.502 | 0.332 | 0.317 |
| Primary school | 0.351 | 0.319 | 0.423 | 0.380 | 0.284 | 0.259 | 0.340 | 0.301 | 0.410 | 0.362 | 0.276 | 0.241 |
| Years of schooling | 4.415 | 4.197 | 4.911 | 4.625 | 3.964 | 3.792 | 4.313 | 4.036 | 4.808 | 4.490 | 3.879 | 3.588 |
| Labor force | 0.435 | 0.460 | 0.388 | 0.388 | 0.486 | 0.537 | 0.463 | 0.481 | 0.412 | 0.408 | 0.509 | 0.550 |
| Employed | 0.338 | 0.373 | 0.302 | 0.313 | 0.378 | 0.436 | 0.365 | 0.396 | 0.324 | 0.334 | 0.403 | 0.453 |
| Months worked | 3.457 | 3.890 | 3.096 | 3.235 | 3.861 | 4.584 | 3.793 | 4.158 | 3.376 | 3.479 | 4.185 | 4.771 |
| Primary sector | 0.279 | 0.361 | 0.254 | 0.304 | 0.304 | 0.420 | 0.299 | 0.384 | 0.273 | 0.326 | 0.319 | 0.437 |
| Secondary sector | 0.022 | 0.009 | 0.017 | 0.005 | 0.029 | 0.014 | 0.022 | 0.007 | 0.018 | 0.004 | 0.027 | 0.011 |
| Tertiary sector | 0.049 | 0.011 | 0.043 | 0.008 | 0.058 | 0.014 | 0.057 | 0.015 | 0.047 | 0.011 | 0.070 | 0.020 |
| High-skilled work | 0.004 | 0.001 | 0.003 | 0.001 | 0.004 | 0.001 | 0.004 | 0.001 | 0.003 | 0.000 | 0.005 | 0.001 |
| Middle-skilled work | 0.305 | 0.373 | 0.269 | 0.311 | 0.343 | 0.437 | 0.322 | 0.395 | 0.285 | 0.331 | 0.354 | 0.453 |
| Low-skilled work | 0.039 | 0.008 | 0.038 | 0.006 | 0.044 | 0.009 | 0.050 | 0.010 | 0.045 | 0.008 | 0.056 | 0.013 |
| Armed forces | 0.003 | 0.000 | 0.005 | 0.000 | 0.000 | 0.000 | 0.002 | 0.000 | 0.004 | 0.000 | 0.001 | 0.000 |
| Observations | 4,585 | 5,980 | 2,630 | 3,617 | 2,686 | 3,695 | 6,116 | 6,507 | 3,418 | 3,943 | 3,646 | 3,954 |
| | B-1. Age 15-19 (VN) | | | | | | | | | | | |
| Some education | 0.673 | 0.774 | 0.721 | 0.832 | 0.628 | 0.717 | 0.657 | 0.730 | 0.710 | 0.794 | 0.609 | 0.686 |
| School participation | 0.431 | 0.441 | 0.511 | 0.548 | 0.351 | 0.329 | 0.400 | 0.385 | 0.489 | 0.493 | 0.322 | 0.297 |
| Primary school | 0.270 | 0.307 | 0.318 | 0.375 | 0.235 | 0.242 | 0.265 | 0.294 | 0.313 | 0.377 | 0.224 | 0.226 |
| Years of schooling | 3.467 | 3.925 | 3.801 | 4.443 | 3.178 | 3.429 | 3.462 | 3.779 | 3.920 | 4.437 | 3.048 | 3.287 |
| Labor force | 0.397 | 0.480 | 0.361 | 0.387 | 0.447 | 0.571 | 0.432 | 0.534 | 0.388 | 0.445 | 0.480 | 0.606 |
| Employed | 0.317 | 0.400 | 0.287 | 0.321 | 0.358 | 0.481 | 0.349 | 0.449 | 0.303 | 0.367 | 0.397 | 0.524 |
| Months worked | 3.583 | 4.457 | 3.271 | 3.587 | 4.036 | 5.331 | 3.942 | 5.019 | 3.464 | 4.059 | 4.451 | 5.857 |
| Primary sector | 0.246 | 0.389 | 0.232 | 0.314 | 0.277 | 0.465 | 0.268 | 0.446 | 0.237 | 0.360 | 0.298 | 0.523 |
| Secondary sector | 0.014 | 0.008 | 0.018 | 0.004 | 0.009 | 0.014 | 0.016 | 0.002 | 0.017 | 0.000 | 0.015 | 0.003 |
| Tertiary sector | 0.059 | 0.005 | 0.041 | 0.004 | 0.077 | 0.007 | 0.070 | 0.007 | 0.052 | 0.009 | 0.092 | 0.007 |
| High-skilled work | 0.002 | 0.001 | 0.004 | 0.001 | 0.002 | 0.002 | 0.004 | 0.001 | 0.004 | 0.002 | 0.006 | 0.000 |
| Middle-skilled work | 0.294 | 0.394 | 0.273 | 0.317 | 0.330 | 0.475 | 0.321 | 0.446 | 0.280 | 0.357 | 0.360 | 0.523 |
| Low-skilled work | 0.023 | 0.007 | 0.014 | 0.005 | 0.030 | 0.009 | 0.029 | 0.010 | 0.022 | 0.011 | 0.038 | 0.010 |
| Armed forces | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Observations | 811 | 943 | 479 | 580 | 496 | 600 | 826 | 789 | 463 | 474 | 522 | 480 |

*Continue*

Table A3: Descriptive Statistics (Mean) – Age 6-19

| | Age 34-60 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | | Women | | | | | |
| | Both Sexes | | Boys | | Girls | | Both Sexes | | Boys | | Girls | |
| Group: | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer | Refugee | Stayer |
| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | | | | | | A-2. Age 12-14 (TH) | | | | | | |
| Some education | 0.802 | 0.827 | 0.835 | 0.840 | 0.770 | 0.810 | 0.791 | 0.803 | 0.820 | 0.831 | 0.763 | 0.773 |
| School participation | 0.736 | 0.777 | 0.779 | 0.811 | 0.694 | 0.741 | 0.716 | 0.744 | 0.766 | 0.789 | 0.668 | 0.694 |
| Grade progression | -5.066 | -5.043 | -4.967 | -4.999 | -5.170 | -5.103 | -5.154 | -5.179 | -5.064 | -5.098 | -5.249 | -5.287 |
| Child labor | 0.053 | 0.045 | 0.043 | 0.032 | 0.062 | 0.058 | 0.062 | 0.058 | 0.046 | 0.040 | 0.080 | 0.077 |
| Observations | 6,017 | 6,556 | 3,290 | 3,614 | 3,379 | 3,632 | 7,078 | 6,471 | 3,872 | 3,599 | 3,940 | 3,519 |
| | | | | | | B-2. Age 12-14 (VN) | | | | | | |
| Some education | 0.695 | 0.785 | 0.729 | 0.816 | 0.660 | 0.748 | 0.687 | 0.753 | 0.744 | 0.773 | 0.627 | 0.738 |
| School participation | 0.648 | 0.744 | 0.688 | 0.793 | 0.610 | 0.688 | 0.646 | 0.695 | 0.709 | 0.732 | 0.574 | 0.661 |
| Grade progression | -5.402 | -5.324 | -5.291 | -5.188 | -5.526 | -5.467 | -5.445 | -5.340 | -5.306 | -5.330 | -5.622 | -5.341 |
| Child labor | 0.066 | 0.066 | 0.046 | 0.034 | 0.084 | 0.097 | 0.074 | 0.072 | 0.058 | 0.056 | 0.094 | 0.089 |
| Observations | 822 | 880 | 428 | 477 | 477 | 490 | 766 | 766 | 404 | 429 | 433 | 411 |
| | | | | | | A-3. Age 6-11 (TH) | | | | | | |
| Some education | 0.530 | 0.538 | 0.529 | 0.541 | 0.520 | 0.528 | 0.526 | 0.532 | 0.523 | 0.534 | 0.511 | 0.510 |
| School participation | 0.524 | 0.534 | 0.523 | 0.537 | 0.513 | 0.524 | 0.519 | 0.528 | 0.517 | 0.530 | 0.504 | 0.505 |
| Grade progression | -2.687 | -2.707 | -2.699 | -2.711 | -2.687 | -2.697 | -2.810 | -2.783 | -2.804 | -2.809 | -2.803 | -2.725 |
| Child labor | 0.004 | 0.003 | 0.004 | 0.002 | 0.004 | 0.003 | 0.005 | 0.005 | 0.006 | 0.004 | 0.004 | 0.006 |
| Observations | 8,534 | 8,356 | 5,809 | 5,779 | 5,702 | 5,564 | 8,849 | 7,979 | 5,926 | 5,412 | 5,822 | 5,211 |
| | | | | | | B-3. Age 6-11 (VN) | | | | | | |
| Some education | 0.461 | 0.510 | 0.436 | 0.505 | 0.445 | 0.506 | 0.453 | 0.478 | 0.457 | 0.508 | 0.436 | 0.438 |
| School participation | 0.451 | 0.508 | 0.431 | 0.503 | 0.432 | 0.504 | 0.448 | 0.475 | 0.455 | 0.506 | 0.428 | 0.435 |
| Grade progression | -2.827 | -2.817 | -2.830 | -2.799 | -2.833 | -2.813 | -2.952 | -2.875 | -2.927 | -2.778 | -2.916 | -2.975 |
| Child labor | 0.009 | 0.006 | 0.007 | 0.006 | 0.013 | 0.008 | 0.015 | 0.006 | 0.016 | 0.004 | 0.017 | 0.008 |
| Observations | 1,001 | 1,116 | 682 | 766 | 638 | 722 | 862 | 850 | 579 | 586 | 536 | 527 |

Table A4: Additional Results – Sociodemographic Outcomes (Age 20-60)

| | A-1. Returnees Aged 34-60 from Thailand | | | | | | B-1. Returnees Aged 34-60 from Vietnam | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | Women | | | Men | | | Women | | |
| | | Stayers | | | Stayers | | | Stayers | | | Stayers | |
| | ATT | Mean | Γ | ATT | Mean | Γ | ATT | Mean | Γ | ATT | Mean | Γ |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Household head | -0.006 | 0.899 | | 0.087 | 0.225 | 1.5 | -0.004 | 0.907 | | 0.038 | 0.215 | 1.0 |
| | (0.005) | | | (0.006) | | | (0.013) | | | (0.015) | | |
| Marital status | | | | | | | | | | | | |
| Married | 0.018 | 0.956 | 1.5 | 0.027 | 0.758 | 1.1 | 0.013 | 0.951 | | 0.036 | 0.721 | 1.0 |
| | (0.003) | | | (0.005) | | | (0.009) | | | (0.015) | | |
| Divorced | -0.001 | 0.009 | | 0.008 | 0.048 | 1.1 | 0.001 | 0.009 | | 0.001 | 0.052 | |
| | (0.001) | | | (0.003) | | | (0.003) | | | (0.008) | | |
| Widowed | -0.004 | 0.014 | 1.1 | 0.009 | 0.126 | 1.0 | 0.002 | 0.012 | | -0.011 | 0.148 | |
| | (0.002) | | | (0.004) | | | (0.004) | | | (0.012) | | |
| Demographics | | | | | | | | | | | | |
| Household size | -0.331 | 6.577 | 1.4 | -0.237 | 6.038 | 1.2 | -0.162 | 6.624 | | 0.153 | 5.871 | |
| | (0.033) | | | (0.029) | | | (0.087) | | | (0.086) | | |
| Grandfather | -0.007 | 0.018 | 1.2 | -0.005 | 0.014 | 1.2 | -0.005 | 0.018 | | 0.002 | 0.016 | |
| | (0.002) | | | (0.001) | | | (0.007) | | | (0.005) | | |
| Grandmother | -0.011 | 0.043 | 1.2 | -0.017 | 0.060 | 1.3 | 0.006 | 0.033 | | -0.006 | 0.067 | |
| | (0.003) | | | (0.003) | | | (0.008) | | | (0.009) | | |
| Children aged 15-19 | -0.137 | 0.561 | 1.6 | -0.040 | 0.533 | 1.1 | -0.070 | 0.634 | 1.2 | 0.015 | 0.515 | |
| | (0.007) | | | (0.006) | | | (0.017) | | | (0.017) | | |
| Male children aged 15-19 | -0.097 | 0.339 | 1.5 | -0.045 | 0.320 | 1.2 | -0.047 | 0.390 | 1.1 | -0.010 | 0.307 | |
| | (0.006) | | | (0.006) | | | (0.018) | | | (0.016) | | |
| Female children aged 15-19 | -0.093 | 0.346 | 1.4 | -0.030 | 0.325 | 1.1 | -0.062 | 0.402 | 1.1 | 0.021 | 0.313 | |
| | (0.006) | | | (0.006) | | | (0.017) | | | (0.016) | | |
| No. of children aged 15-19 | -0.244 | 0.825 | 1.9 | -0.105 | 0.769 | 1.3 | -0.160 | 0.965 | 1.3 | 0.028 | 0.747 | |
| | (0.011) | | | (0.010) | | | (0.031) | | | (0.029) | | |
| No. of male children aged 15-19 | -0.121 | 0.407 | 1.9 | -0.059 | 0.383 | 1.5 | -0.071 | 0.480 | 1.2 | 0.003 | 0.362 | |
| | (0.008) | | | (0.007) | | | (0.024) | | | (0.021) | | |
| No. of female children aged 15-19 | -0.123 | 0.418 | 1.8 | -0.046 | 0.386 | 1.4 | -0.089 | 0.485 | 1.3 | 0.025 | 0.385 | |
| | (0.008) | | | (0.007) | | | (0.023) | | | (0.022) | | |

*Continue*

Table A4: Additional Results – Sociodemographic Outcomes (Age 20-60)

| | A-1. Returnees Aged 34-60 from Thailand | | | | | | B-1. Returnees Aged 34-60 from Vietnam | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | Women | | | Men | | | Women | | |
| | Stayers | | | Stayers | | | Stayers | | | Stayers | | |
| | ATT | Mean | Γ | ATT | Mean | Γ | ATT | Mean | Γ | ATT | Mean | Γ |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Demographics | | | | | | | | | | | | |
| Children aged 12-14 | -0.053 | 0.612 | 1.2 | 0.041 | 0.526 | 1.1 | -0.039 | 0.591 | | 0.001 | 0.494 | |
| | (0.007) | | | (0.006) | | | (0.019) | | | (0.017) | | |
| Male children aged 12-14 | -0.034 | 0.337 | 1.1 | 0.018 | 0.293 | 1.0 | -0.045 | 0.323 | 1.0 | -0.018 | 0.277 | |
| | (0.007) | | | (0.006) | | | (0.018) | | | (0.016) | | |
| Female children aged 12-14 | -0.024 | 0.339 | 1.1 | 0.030 | 0.286 | 1.1 | 0.001 | 0.329 | | 0.019 | 0.264 | |
| | (0.007) | | | (0.006) | | | (0.018) | | | (0.016) | | |
| No. of children aged 12-14 | -0.064 | 0.745 | 1.2 | 0.051 | 0.640 | 1.1 | -0.065 | 0.729 | 1.1 | -0.001 | 0.587 | |
| | (0.010) | | | (0.009) | | | (0.026) | | | (0.023) | | |
| No. of male children aged 12-14 | -0.037 | 0.372 | 1.3 | 0.017 | 0.326 | 1.0 | -0.049 | 0.358 | 1.2 | -0.018 | 0.301 | |
| | (0.008) | | | (0.007) | | | (0.021) | | | (0.018) | | |
| No. of female children aged 12-14 | -0.028 | 0.374 | 1.2 | 0.034 | 0.314 | 1.0 | -0.015 | 0.371 | | 0.017 | 0.287 | |
| | (0.008) | | | (0.007) | | | (0.021) | | | (0.018) | | |
| Children aged 6-11 | 0.009 | 0.782 | | 0.054 | 0.663 | 1.2 | -0.080 | 0.754 | 1.2 | -0.009 | 0.577 | |
| | (0.006) | | | (0.006) | | | (0.017) | | | (0.016) | | |
| Male children aged 6-11 | 0.000 | 0.540 | | 0.033 | 0.448 | 1.1 | -0.059 | 0.515 | 1.1 | -0.010 | 0.394 | |
| | (0.007) | | | (0.006) | | | (0.019) | | | (0.017) | | |
| Female children aged 6-11 | 0.007 | 0.520 | | 0.039 | 0.432 | 1.1 | -0.056 | 0.487 | 1.1 | 0.000 | 0.351 | |
| | (0.007) | | | (0.006) | | | (0.019) | | | (0.016) | | |
| No. of children aged 6-11 | 0.013 | 1.433 | | 0.086 | 1.184 | 1.1 | -0.164 | 1.343 | 1.2 | -0.029 | 0.980 | |
| | (0.014) | | | (0.013) | | | (0.037) | | | (0.034) | | |
| No. of male children aged 6-11 | 0.001 | 0.737 | | 0.042 | 0.606 | 1.0 | -0.091 | 0.700 | 1.1 | -0.022 | 0.526 | |
| | (0.012) | | | (0.010) | | | (0.030) | | | (0.026) | | |
| No. of female children aged 6-11 | 0.012 | 0.697 | | 0.044 | 0.578 | 1.0 | -0.073 | 0.644 | 1.1 | -0.007 | 0.455 | |
| | (0.012) | | | (0.010) | | | (0.028) | | | (0.023) | | |
| Observations (Returnees) | 10,954 | | | 12,819 | | | 1,526 | | | 1,630 | | |
| Observations (Stayers) | 10,954 | | | 12,819 | | | 1,526 | | | 1,630 | | |

*Continue*

Table A4: Additional Results – Sociodemographic Outcomes (Age 20-60)

| | A-2. Returnees Aged 20-33 from Thailand | | | | | | B-2. Returnees Aged 20-33 from Vietnam | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | Women | | | Men | | | Women | | |
| | | Stayers | | | Stayers | | | Stayers | | | Stayers | |
| | ATT | Mean | Γ | ATT | Mean | Γ | ATT | Mean | Γ | ATT | Mean | Γ |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Household head | 0.018 | 0.601 | 1.0 | 0.063 | 0.108 | 1.5 | -0.021 | 0.540 | | 0.016 | 0.081 | |
| | (0.008) | | | (0.006) | | | (0.020) | | | (0.012) | | |
| Marital status | | | | | | | | | | | | |
| Married | -0.004 | 0.734 | | 0.053 | 0.748 | 1.3 | -0.054 | 0.713 | 1.0 | 0.010 | 0.713 | |
| | (0.007) | | | (0.007) | | | (0.019) | | | (0.019) | | |
| Divorced | 0.000 | 0.009 | | 0.011 | 0.029 | 1.2 | -0.003 | 0.007 | | -0.010 | 0.039 | |
| | (0.002) | | | (0.003) | | | (0.003) | | | (0.008) | | |
| Widowed | -0.001 | 0.005 | | 0.009 | 0.022 | 1.2 | 0.004 | 0.003 | | -0.015 | 0.028 | |
| | (0.001) | | | (0.003) | | | (0.003) | | | (0.006) | | |
| Demographics | | | | | | | | | | | | |
| Household size | 0.032 | 5.361 | | 0.159 | 5.302 | 1.0 | 0.291 | 5.478 | 1.0 | 0.345 | 5.252 | 1.0 |
| | (0.048) | | | (0.035) | | | (0.113) | | | (0.107) | | |
| Grandfather | -0.001 | 0.014 | | 0.000 | 0.014 | | -0.009 | 0.024 | | 0.005 | 0.010 | |
| | (0.002) | | | (0.002) | | | (0.006) | | | (0.005) | | |
| Grandmother | -0.008 | 0.048 | | 0.006 | 0.040 | | -0.011 | 0.057 | | -0.008 | 0.045 | |
| | (0.004) | | | (0.003) | | | (0.010) | | | (0.009) | | |
| Observations (Returnees) | 5,260 | | | 6,979 | | | 971 | | | 1,018 | | |
| Observations (Stayers) | 5,260 | | | 6,979 | | | 971 | | | 1,018 | | |

*Notes:* Columns 1 and 4 (Columns 7 and 10) report the estimates of the impacts of forced displacement on the sociodemographic outcomes for male and female returnees aged 20–60 from Thailand (Vietnam). The estimates are based on the Matched Samples and are from the bias-corrected version of the nearest-neighbor matching method (Abadie and Imbens 2011). Robust Abadie-Imbens standard errors are reported in parentheses. Columns 2 and 5 (Columns 8 and 11) report the means of the corresponding male and female stayers aged 20–60, respectively, for male and female returnees aged 20–60 from Thailand (Vietnam). Columns 3 and 6 (Columns 9 and 12) report the results of Rosenbaum's sensitivity analysis (Rosenbaum 2002); the values of Γ are sensitivity parameters, which correspond to maximum multiplicity-adjusted $p$-values less than 5%. "Household head" is an indicator variable equal to 1 if individuals are household heads and 0 otherwise. "Married"/"Divorced"/"Widowed" is an indicator variable equal to 1 if individuals are married/divorced/widowed and 0 otherwise. "Household size" is the number of household members. "Grandfather"/"Grandmother" is an indicator variable equal to 1 if individuals live together with their grandfather/grandmother and 0 otherwise. "Children aged 15–19/12–14/6–11" is an indicator variable equal to 1 if individuals live together with children aged 15–19/12–14/6–11. "No. of children aged 15–19/12–14/6–11" is the number of children aged 15–19/12–14/6–11 living together.

Table A5: Robustness Check – Alternative Estimation Methods (Age 20-60, TH)

| | A. Returnees from Thailand Aged 34-60 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | | Women | | | | | |
| | Matched Sample | | | Full Sample | | | Matched Sample | | | Full Sample | | |
| Dependent Variable | Match (1) | Blocking (2) | Γ (3) | OLS (4) | Post-ALASSO (5) | Inv. IVs (6) | Match (7) | Blocking (8) | Γ (9) | OLS (10) | Post-ALASSO (11) | Inv. IVs (12) |
| Labor force | -0.006 | -0.005 | 1.1 | -0.008 | -0.051 | 19 | -0.074 | -0.077 | 1.6 | -0.081 | -0.043 | 7 |
| | (0.002) | (0.003) | | (0.002) | (0.014) | | (0.005) | (0.007) | | (0.004) | (0.010) | |
| Employed | -0.025 | -0.025 | 1.6 | -0.028 | -0.136 | 27 | -0.095 | -0.098 | 1.7 | -0.104 | -0.333 | 12 |
| | (0.003) | (0.004) | | (0.002) | (0.031) | | (0.005) | (0.007) | | (0.004) | (0.058) | |
| Months worked | -0.260 | -0.263 | 1.0 | -0.284 | -1.120 | 17 | -1.170 | -1.193 | 1.3 | -1.254 | -3.325 | 12 |
| | (0.041) | (0.055) | | (0.033) | (0.248) | | (0.058) | (0.080) | | (0.047) | (0.356) | |
| Primary sector | -0.263 | -0.266 | 3.7 | -0.289 | -0.446 | 21 | -0.219 | -0.222 | 2.7 | -0.235 | -0.185 | 10 |
| | (0.006) | (0.008) | | (0.005) | (0.039) | | (0.006) | (0.008) | | (0.005) | (0.014) | |
| Secondary sector | 0.029 | 0.029 | 2.5 | 0.030 | -0.031 | 28 | 0.009 | 0.009 | 1.8 | 0.008 | -0.021 | 15 |
| | (0.002) | (0.003) | | (0.002) | (0.007) | | (0.001) | (0.002) | | (0.001) | (0.007) | |
| Tertiary sector | 0.218 | 0.220 | 3.5 | 0.239 | 0.403 | 20 | 0.120 | 0.119 | 3.4 | 0.130 | 0.045 | 5 |
| | (0.006) | (0.007) | | (0.005) | (0.030) | | (0.004) | (0.005) | | (0.003) | (0.014) | |
| High-skilled work | 0.000 | -0.002 | | 0.008 | 0.003 | 9 | 0.005 | 0.004 | 1.5 | 0.007 | 0.004 | 5 |
| | (0.003) | (0.004) | | (0.002) | (0.005) | | (0.001) | (0.002) | | (0.001) | (0.002) | |
| Middle-skilled work | -0.156 | -0.159 | 2.7 | -0.178 | -0.314 | 18 | -0.175 | -0.177 | 2.3 | -0.188 | -0.162 | 9 |
| | (0.006) | (0.007) | | (0.004) | (0.040) | | (0.005) | (0.007) | | (0.004) | (0.013) | |
| Low-skilled work | 0.078 | 0.080 | 4.4 | 0.081 | 0.052 | 24 | 0.078 | 0.078 | 3.9 | 0.081 | -0.001 | 16 |
| | (0.003) | (0.004) | | (0.003) | (0.014) | | (0.003) | (0.004) | | (0.003) | (0.010) | |
| Armed forces | 0.063 | 0.065 | $> 5.0$ | 0.070 | 0.080 | 16 | 0.002 | 0.002 | 3.7 | 0.002 | 0.002 | 0 |
| | (0.003) | (0.004) | | (0.003) | (0.011) | | (0.000) | (0.001) | | (0.000) | (0.000) | |
| Home ownership | -0.031 | -0.030 | 3.1 | -0.031 | -0.053 | 24 | -0.034 | -0.035 | 3.6 | -0.035 | -0.004 | 12 |
| | (0.002) | (0.003) | | (0.002) | (0.016) | | (0.002) | (0.003) | | (0.002) | (0.008) | |
| Number of instruments | | | | | | 32 | | | | | | 24 |
| Obs. (Returnees) | 10,954 | 10,954 | | 11,324 | 11,324 | | 12,819 | 12,819 | | 13,057 | 13,057 | |
| Obs. (Stayers) | 10,954 | 10,954 | | 362,881 | 362,881 | | 12,819 | 12,819 | | 598,647 | 598,647 | |

*Continue*

Table A5: Robustness Check – Alternative Estimation Methods (Age 20-60, TH)

| | B. Returnees from Thailand Aged 20-33 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | | Women | | | | | |
| | Matched Sample | | | Full Sample | | | Matched Sample | | | Full Sample | | |
| Dependent Variable | Match (1) | Blocking (2) | Γ (3) | OLS (4) | Post-ALASSO (5) | Inv. IVs (6) | Match (7) | Blocking (8) | Γ (9) | OLS (10) | Post-ALASSO (11) | Inv. IVs (12) |
| Some education | 0.051 | 0.049 | 1.2 | 0.046 | 0.031 | 4 | -0.010 | -0.010 | | -0.004 | 0.072 | 5 |
| | (0.008) | (0.011) | | (0.006) | (0.021) | | (0.008) | (0.011) | | (0.006) | (0.020) | |
| Primary school | 0.111 | 0.110 | 1.4 | 0.107 | 0.021 | 5 | 0.001 | 0.000 | | -0.001 | -0.236 | 9 |
| | (0.009) | (0.013) | | (0.007) | (0.026) | | (0.007) | (0.009) | | (0.005) | (0.051) | |
| Years of schooling | 1.060 | 1.045 | 1.6 | 0.992 | 0.584 | 4 | -0.005 | -0.014 | | 0.004 | -2.972 | 10 |
| | (0.067) | (0.094) | | (0.053) | (0.178) | | (0.049) | (0.068) | | (0.038) | (0.432) | |
| Labor force | 0.002 | 0.001 | | -0.002 | -0.010 | 3 | -0.082 | -0.084 | 1.6 | -0.090 | -0.041 | 4 |
| | (0.005) | (0.007) | | (0.004) | (0.013) | | (0.007) | (0.009) | | (0.005) | (0.016) | |
| Employed | -0.014 | -0.015 | 1.1 | -0.019 | -0.061 | 5 | -0.097 | -0.099 | 1.6 | -0.107 | -0.563 | 8 |
| | (0.006) | (0.008) | | (0.005) | (0.039) | | (0.007) | (0.010) | | (0.006) | (0.075) | |
| Months worked | -0.084 | -0.100 | | -0.170 | -7.814 | 7 | -1.139 | -1.165 | 1.3 | -1.212 | -4.907 | 7 |
| | (0.075) | (0.104) | | (0.058) | (1.327) | | (0.082) | (0.114) | | (0.065) | (0.562) | |
| Primary sector | -0.233 | -0.234 | 2.8 | -0.245 | -0.449 | 9 | -0.204 | -0.206 | 2.3 | -0.212 | -0.122 | 5 |
| | (0.009) | (0.012) | | (0.007) | (0.132) | | (0.008) | (0.011) | | (0.006) | (0.020) | |
| Secondary sector | 0.032 | 0.032 | 2.5 | 0.034 | -0.037 | 3 | 0.013 | 0.013 | 2.0 | 0.012 | -0.014 | 7 |
| | (0.003) | (0.005) | | (0.003) | (0.006) | | (0.002) | (0.003) | | (0.002) | (0.004) | |
| Tertiary sector | 0.194 | 0.194 | 3.6 | 0.201 | 0.097 | 2 | 0.101 | 0.100 | 3.0 | 0.100 | -0.012 | 10 |
| | (0.007) | (0.010) | | (0.006) | (0.015) | | (0.005) | (0.007) | | (0.004) | (0.028) | |
| High-skilled work | 0.024 | 0.024 | 1.5 | 0.023 | 0.018 | 1 | 0.006 | 0.005 | 1.1 | 0.005 | -0.010 | 7 |
| | (0.004) | (0.005) | | (0.003) | (0.009) | | (0.002) | (0.003) | | (0.002) | (0.004) | |
| Middle-skilled work | -0.147 | -0.147 | 2.1 | -0.153 | -0.099 | 6 | -0.166 | -0.168 | 2.0 | -0.173 | -0.685 | 6 |
| | (0.008) | (0.011) | | (0.007) | (0.022) | | (0.007) | (0.010) | | (0.006) | (0.062) | |
| Low-skilled work | 0.074 | 0.074 | 3.6 | 0.076 | -0.032 | 7 | 0.067 | 0.067 | 4.0 | 0.065 | -0.047 | 7 |
| | (0.005) | (0.006) | | (0.004) | (0.010) | | (0.004) | (0.005) | | (0.003) | (0.006) | |
| Armed forces | 0.042 | 0.042 | 3.7 | 0.045 | 0.024 | 0 | 0.002 | 0.002 | 1.8 | 0.002 | 0.001 | 10 |
| | (0.003) | (0.005) | | (0.003) | (0.004) | | (0.001) | (0.001) | | (0.001) | (0.002) | |
| Home ownership | -0.045 | -0.045 | 2.8 | -0.046 | 0.033 | 7 | -0.039 | -0.040 | 2.6 | -0.040 | 0.056 | 4 |
| | (0.004) | (0.005) | | (0.004) | (0.011) | | (0.004) | (0.005) | | (0.003) | (0.010) | |
| Number of instruments | | | | | | 12 | | | | | | 13 |
| Obs. (Returnees) | 5,260 | 5,260 | | 5,318 | 5,318 | | 6,979 | 6,979 | | 7,061 | 7,061 | |
| Obs. (Stayers) | 5,260 | 5,260 | | 423,886 | 423,886 | | 6,979 | 6,979 | | 583,273 | 583,273 | |

*Notes:* The table considers the robustness of the results to alternative estimation methods for the male and female returnees aged 34-60 (panel A) and 20-33 (panel B) from Thailand. For comparison purposes, columns 1 and 7 report the estimates based on the Matched Samples from the bias-corrected version of the nearest-neighbor matching method (Abadie and Imbens 2011). Columns 3 and 9 report the corresponding results of Rosenbaum's sensitivity analysis (Rosenbaum 2002); the values of Γ are sensitivity parameters, which correspond to maximum multiplicity-adjusted $p$-values less than 5%. Columns 2 and 8 report the estimates based on the Matched Samples from blocking on the estimated propensity score, in combination with regression adjustment within the blocks for the optimal number of blocks, selected based on the algorithm proposed by Imbens and Rubin (2015). Columns 4 and 10 report the OLS estimates based on the Full Samples. Columns 5 and 11 report the estimates based on the Full Samples from the machine learning-based instrumental variables (adaptive Lasso) approach, following Windmeijer et al. (2019), assessing instrumental validity in columns 6 and 12. Robust (Abadie-Imbens) standard errors are reported in parentheses (in columns 1 and 7). For the definitions of the dependent variables, see the main text.

Table A6: Robustness Check – Alternative Estimation Methods (Age 20-60, VN)

| | A. Returnees from Vietnam Aged 34-60 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | Women | | | | |
| | Matched Sample | | | Full Sample (OLS) | | Matched Sample | | | Full Sample (OLS) | |
| | Match | Blocking | Γ | Limited | All | Match | Blocking | Γ | Limited | All |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Labor force | 0.003 | -0.003 | | -0.005 | 0.001 | -0.064 | -0.061 | 1.3 | -0.065 | -0.038 |
| | (0.008) | (0.009) | | (0.005) | (0.005) | (0.013) | (0.012) | | (0.010) | (0.009) |
| Employed | -0.004 | -0.011 | | -0.010 | -0.001 | -0.062 | -0.060 | 1.2 | -0.067 | -0.037 |
| | (0.008) | (0.010) | | (0.006) | (0.006) | (0.013) | (0.013) | | (0.010) | (0.010) |
| Months worked | 0.118 | 0.040 | | 0.041 | 0.110 | -0.565 | -0.544 | 1.0 | -0.633 | -0.357 |
| | (0.108) | (0.132) | | (0.071) | (0.073) | (0.157) | (0.153) | | (0.122) | (0.115) |
| Primary sector | -0.256 | -0.262 | 3.4 | -0.273 | -0.256 | -0.216 | -0.215 | 2.4 | -0.226 | -0.182 |
| | (0.015) | (0.020) | | (0.012) | (0.012) | (0.015) | (0.014) | | (0.012) | (0.011) |
| Secondary sector | 0.060 | 0.062 | 3.9 | 0.062 | 0.060 | 0.008 | 0.008 | | 0.009 | 0.008 |
| | (0.007) | (0.010) | | (0.007) | (0.007) | (0.004) | (0.003) | | (0.003) | (0.003) |
| Tertiary sector | 0.195 | 0.193 | 2.9 | 0.203 | 0.195 | 0.144 | 0.146 | 4.1 | 0.149 | 0.136 |
| | (0.013) | (0.018) | | (0.011) | (0.011) | (0.010) | (0.010) | | (0.010) | (0.009) |
| High-skilled work | -0.006 | -0.009 | | 0.001 | 0.000 | 0.003 | 0.002 | | 0.003 | 0.003 |
| | (0.007) | (0.009) | | (0.005) | (0.005) | (0.002) | (0.002) | | (0.002) | (0.002) |
| Middle-skilled work | -0.081 | -0.083 | 1.6 | -0.092 | -0.082 | -0.129 | -0.127 | 1.7 | -0.133 | -0.098 |
| | (0.013) | (0.017) | | (0.010) | (0.010) | (0.014) | (0.014) | | (0.011) | (0.011) |
| Low-skilled work | 0.071 | 0.073 | 3.0 | 0.070 | 0.068 | 0.061 | 0.062 | 2.8 | 0.062 | 0.056 |
| | (0.008) | (0.011) | | (0.008) | (0.007) | (0.007) | (0.007) | | (0.007) | (0.006) |
| Armed forces | 0.014 | 0.012 | 1.6 | 0.012 | 0.013 | 0.001 | 0.001 | | 0.001 | 0.001 |
| | (0.004) | (0.005) | | (0.003) | (0.003) | (0.001) | (0.001) | | (0.001) | (0.001) |
| Home ownership | -0.027 | -0.028 | 1.6 | -0.030 | -0.027 | -0.013 | -0.013 | 1.1 | -0.012 | -0.012 |
| | (0.006) | (0.008) | | (0.005) | (0.005) | (0.005) | (0.005) | | (0.004) | (0.004) |
| Obs. (Returnees) | 1,526 | 1,526 | | 1,574 | 4,226 | 1,630 | 1,630 | | 1,650 | 4,850 |
| Obs. (Stayers) | 1,526 | 1,526 | | 362,821 | 362,881 | 1,630 | 1,630 | | 598,411 | 598,647 |

*Continue*

Table A6: Robustness Check – Alternative Estimation Methods (Age 20-60, VN)

| | B. Returnees from Vietnam Aged 20-33 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | Women | | | | |
| | Matched Sample | | | Full Sample (OLS) | | Matched Sample | | | Full Sample (OLS) | |
| | Match | Blocking | Γ | Limited | All | Match | Blocking | Γ | Limited | All |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Some education | -0.163 | -0.164 | 1.6 | -0.161 | -0.121 | -0.095 | -0.096 | 1.2 | -0.106 | -0.078 |
| | (0.020) | (0.019) | | (0.015) | (0.011) | (0.020) | (0.020) | | (0.015) | (0.012) |
| Primary school | -0.087 | -0.087 | 1.2 | -0.074 | -0.056 | -0.004 | -0.004 | | -0.021 | -0.008 |
| | (0.020) | (0.020) | | (0.014) | (0.012) | (0.015) | (0.015) | | (0.011) | (0.010) |
| Years of schooling | -0.848 | -0.852 | 1.5 | -0.792 | -0.592 | -0.308 | -0.312 | 1.4 | -0.384 | -0.254 |
| | (0.151) | (0.146) | | (0.113) | (0.091) | (0.120) | (0.119) | | (0.092) | (0.075) |
| Labor force | 0.025 | 0.025 | | 0.015 | 0.052 | -0.016 | -0.018 | | -0.042 | -0.008 |
| | (0.012) | (0.011) | | (0.008) | (0.008) | (0.016) | (0.016) | | (0.012) | (0.009) |
| Employed | 0.009 | 0.009 | | 0.000 | 0.045 | -0.040 | -0.041 | | -0.057 | -0.012 |
| | (0.014) | (0.014) | | (0.010) | (0.010) | (0.018) | (0.017) | | (0.013) | (0.010) |
| Months worked | 0.222 | 0.224 | | 0.135 | 0.557 | -0.330 | -0.342 | | -0.543 | -0.120 |
| | (0.173) | (0.172) | | (0.124) | (0.113) | (0.210) | (0.205) | | (0.155) | (0.120) |
| Primary sector | -0.163 | -0.163 | 1.8 | -0.170 | -0.080 | -0.182 | -0.183 | 1.9 | -0.197 | -0.112 |
| | (0.020) | (0.019) | | (0.015) | (0.013) | (0.020) | (0.019) | | (0.015) | (0.012) |
| Secondary sector | 0.049 | 0.049 | 2.3 | 0.054 | 0.040 | 0.011 | 0.011 | | 0.012 | 0.009 |
| | (0.009) | (0.008) | | (0.008) | (0.006) | (0.005) | (0.005) | | (0.004) | (0.003) |
| Tertiary sector | 0.128 | 0.128 | 2.3 | 0.124 | 0.090 | 0.133 | 0.132 | 3.4 | 0.127 | 0.091 |
| | (0.015) | (0.014) | | (0.013) | (0.010) | (0.013) | (0.012) | | (0.012) | (0.008) |
| High-skilled work | -0.008 | -0.008 | | -0.008 | -0.009 | 0.003 | 0.003 | | 0.003 | 0.003 |
| | (0.006) | (0.006) | | (0.004) | (0.004) | (0.005) | (0.004) | | (0.003) | (0.002) |
| Middle-skilled work | -0.015 | -0.015 | | -0.019 | 0.030 | -0.092 | -0.093 | 1.3 | -0.104 | -0.044 |
| | (0.017) | (0.017) | | (0.013) | (0.011) | (0.019) | (0.018) | | (0.014) | (0.011) |
| Low-skilled work | 0.031 | 0.031 | 1.3 | 0.029 | 0.021 | 0.050 | 0.050 | 2.3 | 0.043 | 0.030 |
| | (0.009) | (0.009) | | (0.008) | (0.006) | (0.009) | (0.008) | | (0.008) | (0.005) |
| Armed forces | 0.007 | 0.007 | | 0.006 | 0.007 | 0.000 | 0.000 | | 0.000 | 0.000 |
| | (0.004) | (0.004) | | (0.003) | (0.003) | (0.000) | (0.000) | | (0.000) | (0.000) |
| Home ownership | -0.023 | -0.024 | 1.2 | -0.020 | -0.014 | -0.015 | -0.015 | 1.0 | -0.016 | -0.015 |
| | (0.008) | (0.008) | | (0.007) | (0.005) | (0.008) | (0.008) | | (0.006) | (0.004) |
| Obs. (Returnees) | 971 | 971 | | 974 | 2,888 | 1,018 | 1,018 | | 1,021 | 3,584 |
| Obs. (Stayers) | 971 | 971 | | 423,383 | 423,886 | 1,018 | 1,018 | | 582,622 | 583,273 |

*Notes:* The table considers the robustness of the results to alternative estimation methods for the male and female returnees aged 34-60 (panel A) and 20-33 (panel B) from Vietnam. For comparison purposes, columns 1 and 6 report the estimates based on Matched Samples from the bias-corrected version of the nearest-neighbor matching method (Abadie and Imbens 2011). Columns 3 and 8 report the corresponding results of Rosenbaum's sensitivity analysis (Rosenbaum 2002); the values of Γ are sensitivity parameters, which correspond to maximum multiplicity-adjusted $p$-values less than 5%. Columns 2 and 7 report the estimates based on the Matched Samples from blocking on the estimated propensity score, in combination with regression adjustment within the blocks for the optimal number of blocks, selected based on the algorithm proposed by Imbens and Rubin (2015). Columns 4 and 9 (Columns 5 and 9) report the OLS estimates based on the Full Samples, without (with) those who live in the four districts along the border with Vietnam, the great majority of whom are returnees. Robust (Abadie-Imbens) standard errors are reported in parentheses (in columns 1 and 6). For the definitions of the dependent variables, see the main text.

Table A7: Sensitivity Analysis – (Age 20-60)

| | A-1. Returnees Aged 34-60 from Thailand | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | Women | | | | |
| | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | |
| | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Labor force | -0.008 | -0.007 | -0.008 | 0.060 | -0.076 | -0.081 | -0.078 | -0.085 | 0.034 | -0.197 |
| | (0.002) | | | | | (0.004) | | | | |
| Employed | -0.028 | -0.027 | -0.029 | 0.163 | -0.219 | -0.104 | -0.099 | -0.109 | 0.058 | -0.265 |
| | (0.002) | | | | | (0.004) | | | | |
| Months worked | -0.284 | -0.269 | -0.299 | -0.054 | -0.514 | -1.254 | -1.218 | -1.289 | -0.498 | -2.009 |
| | (0.033) | | | | | (0.047) | | | | |
| Primary sector | -0.289 | -0.283 | -0.294 | -0.149 | -0.428 | -0.235 | -0.231 | -0.239 | -0.125 | -0.345 |
| | (0.005) | | | | | (0.005) | | | | |
| Secondary sector | 0.030 | 0.030 | 0.031 | 0.004 | 0.057 | 0.008 | 0.008 | 0.008 | 0.022 | -0.006 |
| | (0.002) | | | | | (0.001) | | | | |
| Tertiary sector | 0.239 | 0.235 | 0.244 | 0.144 | 0.334 | 0.130 | 0.129 | 0.130 | 0.106 | 0.154 |
| | (0.005) | | | | | (0.003) | | | | |
| High-skilled work | 0.008 | 0.005 | 0.011 | -0.061 | 0.077 | 0.007 | 0.007 | 0.007 | 0.014 | 0.001 |
| | (0.002) | | | | | (0.001) | | | | |
| Middle-skilled work | -0.178 | -0.173 | -0.182 | 0.020 | -0.375 | -0.188 | -0.183 | -0.193 | -0.048 | -0.328 |
| | (0.004) | | | | | (0.004) | | | | |
| Low-skilled work | 0.081 | 0.080 | 0.082 | -0.007 | 0.168 | 0.081 | 0.080 | 0.083 | -0.064 | 0.227 |
| | (0.003) | | | | | (0.003) | | | | |
| Armed forces | 0.070 | 0.070 | 0.069 | 0.114 | 0.025 | 0.002 | 0.002 | 0.002 | 0.021 | -0.017 |
| | (0.003) | | | | | (0.000) | | | | |
| Home ownership | -0.031 | -0.031 | -0.031 | -0.007 | -0.055 | -0.035 | -0.036 | -0.035 | -0.179 | 0.109 |
| | (0.002) | | | | | (0.002) | | | | |
| Obs. (Returnees) | 11,324 | | | | | 13,057 | | | | |
| Obs. (Stayers) | 362,881 | | | | | 598,647 | | | | |

*Continue*

Table A7: Sensitivity Analysis – (Age 20-60)

| | B-1. Returnees Aged 34-60 from Vietnam | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | Women | | | | |
| | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | |
| | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Labor force | -0.005 | -0.003 | -0.006 | 0.319 | -0.328 | -0.065 | -0.064 | -0.065 | -0.047 | -0.082 |
| | (0.005) | | | | | (0.010) | | | | |
| Employed | -0.010 | -0.008 | -0.011 | 0.365 | -0.384 | -0.067 | -0.067 | -0.067 | -0.058 | -0.075 |
| | (0.006) | | | | | (0.010) | | | | |
| Months worked | 0.041 | -0.070 | 0.152 | -1.632 | 1.714 | -0.633 | -0.784 | -0.483 | -3.800 | 2.533 |
| | (0.071) | | | | | (0.122) | | | | |
| Primary sector | -0.273 | -0.275 | -0.272 | -0.322 | -0.225 | -0.226 | -0.229 | -0.223 | -0.304 | -0.149 |
| | (0.012) | | | | | (0.012) | | | | |
| Secondary sector | 0.062 | 0.064 | 0.061 | 0.362 | -0.238 | 0.009 | 0.011 | 0.008 | 0.109 | -0.090 |
| | (0.007) | | | | | (0.003) | | | | |
| Tertiary sector | 0.203 | 0.204 | 0.201 | 0.247 | 0.159 | 0.149 | 0.152 | 0.147 | 0.270 | 0.029 |
| | (0.011) | | | | | (0.010) | | | | |
| High-skilled work | 0.001 | 0.002 | 0.001 | 0.014 | -0.011 | 0.003 | 0.004 | 0.002 | 0.038 | -0.032 |
| | (0.005) | | | | | (0.002) | | | | |
| Middle-skilled work | -0.092 | -0.090 | -0.094 | 0.011 | -0.196 | -0.133 | -0.133 | -0.133 | -0.132 | -0.134 |
| | (0.010) | | | | | (0.011) | | | | |
| Low-skilled work | 0.070 | 0.070 | 0.071 | -0.030 | 0.171 | 0.062 | 0.062 | 0.062 | 0.021 | 0.103 |
| | (0.008) | | | | | (0.007) | | | | |
| Armed forces | 0.012 | 0.012 | 0.013 | -0.018 | 0.043 | 0.001 | 0.001 | 0.001 | 0.010 | -0.008 |
| | (0.003) | | | | | (0.001) | | | | |
| Home ownership | -0.030 | -0.029 | -0.031 | 0.159 | -0.220 | -0.012 | -0.011 | -0.013 | 0.283 | -0.308 |
| | (0.005) | | | | | (0.004) | | | | |
| Obs. (Returnees) | 1,574 | | | | | 1,650 | | | | |
| Obs. (Stayers) | 362,821 | | | | | 598,411 | | | | |

*Continue*

Table A7: Sensitivity Analysis – (Age 20-60)

| | A-2. Returnees Aged 20-33 from Thailand | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | Women | | | | |
| | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | |
| | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Some education | 0.046 | 0.048 | 0.044 | 0.114 | -0.022 | -0.004 | 0.001 | -0.010 | 0.167 | -0.176 |
| | (0.006) | | | | | (0.006) | | | | |
| Primary school | 0.107 | 0.112 | 0.102 | 0.274 | -0.061 | -0.001 | 0.001 | -0.002 | 0.057 | -0.059 |
| | (0.007) | | | | | (0.005) | | | | |
| Years of schooling | 0.992 | 1.018 | 0.965 | 1.612 | 0.371 | 0.004 | 0.033 | -0.025 | 0.700 | -0.692 |
| | (0.053) | | | | | (0.038) | | | | |
| Labor force | -0.002 | -0.004 | -0.001 | -0.067 | 0.062 | -0.090 | -0.081 | -0.098 | 0.226 | -0.406 |
| | (0.004) | | | | | (0.005) | | | | |
| Employed | -0.019 | -0.021 | -0.016 | -0.091 | 0.053 | -0.107 | -0.097 | -0.117 | 0.172 | -0.387 |
| | (0.005) | | | | | (0.006) | | | | |
| Months worked | -0.170 | -0.153 | -0.188 | 0.181 | -0.522 | -1.212 | -1.093 | -1.331 | 1.399 | -3.822 |
| | (0.058) | | | | | (0.065) | | | | |
| Primary sector | -0.245 | -0.244 | -0.245 | -0.237 | -0.252 | -0.212 | -0.203 | -0.221 | 0.004 | -0.428 |
| | (0.007) | | | | | (0.006) | | | | |
| Secondary sector | 0.034 | 0.034 | 0.034 | 0.046 | 0.021 | 0.012 | 0.013 | 0.010 | 0.113 | -0.090 |
| | (0.003) | | | | | (0.002) | | | | |
| Tertiary sector | 0.201 | 0.197 | 0.205 | -0.170 | 0.572 | 0.100 | 0.098 | 0.101 | -0.094 | 0.294 |
| | (0.006) | | | | | (0.004) | | | | |
| High-skilled work | 0.023 | 0.021 | 0.024 | -0.700 | 0.745 | 0.005 | 0.005 | 0.005 | -0.044 | 0.054 |
| | (0.003) | | | | | (0.002) | | | | |
| Middle-skilled work | -0.153 | -0.154 | -0.153 | -0.168 | -0.139 | -0.173 | -0.163 | -0.183 | 0.087 | -0.433 |
| | (0.007) | | | | | (0.006) | | | | |
| Low-skilled work | 0.076 | 0.076 | 0.077 | 0.011 | 0.141 | 0.065 | 0.064 | 0.066 | -0.102 | 0.232 |
| | (0.004) | | | | | (0.003) | | | | |
| Armed forces | 0.045 | 0.044 | 0.045 | -0.005 | 0.095 | 0.002 | 0.002 | 0.002 | 0.000 | 0.005 |
| | (0.003) | | | | | (0.001) | | | | |
| Home ownership | -0.046 | -0.047 | -0.045 | -0.277 | 0.185 | -0.040 | -0.040 | -0.040 | -0.134 | 0.054 |
| | (0.004) | | | | | (0.003) | | | | |
| Obs. (Returnees) | 5,318 | | | | | 7,061 | | | | |
| Obs. (Stayers) | 423,886 | | | | | 583,273 | | | | |

*Continue*

Table A7: Sensitivity Analysis – (Age 20-60)

| | B-2. Returnees Aged 20-33 from Vietnam | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | | | Women | | | | |
| | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | | OLS | $R^2_{max} = 1.3 \times R^2_{BS}$ | | $R^2_{max} = 1$ | |
| | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ | (BS) | $\delta = 1$ | $\delta = -1$ | $\delta = 1$ | $\delta = -1$ |
| Dependent Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Some education | -0.161 | -0.146 | -0.177 | 0.470 | -0.793 | -0.106 | -0.076 | -0.136 | 0.869 | -1.080 |
| | (0.015) | | | | | (0.015) | | | | |
| Primary school | -0.074 | -0.065 | -0.083 | 0.212 | -0.360 | -0.021 | -0.008 | -0.034 | 0.515 | -0.557 |
| | (0.014) | | | | | (0.011) | | | | |
| Years of schooling | -0.792 | -0.686 | -0.898 | 1.660 | -3.245 | -0.384 | -0.209 | -0.558 | 3.793 | -4.561 |
| | (0.113) | | | | | (0.092) | | | | |
| Labor force | 0.015 | 0.015 | 0.016 | 0.005 | 0.025 | -0.042 | -0.042 | -0.042 | -0.041 | -0.043 |
| | (0.008) | | | | | (0.012) | | | | |
| Employed | 0.000 | 0.000 | 0.001 | -0.016 | 0.017 | -0.057 | -0.058 | -0.056 | -0.081 | -0.034 |
| | (0.010) | | | | | (0.013) | | | | |
| Months worked | 0.135 | -0.024 | 0.293 | -2.989 | 3.258 | -0.543 | -0.728 | -0.358 | -4.569 | 3.484 |
| | (0.124) | | | | | (0.155) | | | | |
| Primary sector | -0.170 | -0.173 | -0.168 | -0.275 | -0.065 | -0.197 | -0.200 | -0.194 | -0.280 | -0.114 |
| | (0.015) | | | | | (0.015) | | | | |
| Secondary sector | 0.054 | 0.055 | 0.054 | 0.167 | -0.058 | 0.012 | 0.013 | 0.010 | 0.129 | -0.106 |
| | (0.008) | | | | | (0.004) | | | | |
| Tertiary sector | 0.124 | 0.125 | 0.123 | 0.280 | -0.032 | 0.127 | 0.128 | 0.126 | 0.318 | -0.064 |
| | (0.013) | | | | | (0.012) | | | | |
| High-skilled work | -0.008 | -0.007 | -0.008 | 0.113 | -0.128 | 0.003 | 0.004 | 0.002 | 0.478 | -0.472 |
| | (0.004) | | | | | (0.003) | | | | |
| Middle-skilled work | -0.019 | -0.017 | -0.021 | 0.063 | -0.101 | -0.104 | -0.104 | -0.104 | -0.102 | -0.106 |
| | (0.013) | | | | | (0.014) | | | | |
| Low-skilled work | 0.029 | 0.026 | 0.032 | -0.349 | 0.407 | 0.043 | 0.041 | 0.045 | -0.220 | 0.306 |
| | (0.008) | | | | | (0.008) | | | | |
| Armed forces | 0.006 | 0.006 | 0.006 | 0.097 | -0.085 | 0.000 | 0.000 | 0.000 | 0.023 | -0.024 |
| | (0.003) | | | | | (0.000) | | | | |
| Home ownership | -0.020 | -0.018 | -0.021 | 0.329 | -0.368 | -0.016 | -0.016 | -0.017 | 0.120 | -0.152 |
| | (0.007) | | | | | (0.006) | | | | |
| Obs. (Returnees) | 974 | | | | | 1,021 | | | | |
| Obs. (Stayers) | 423,383 | | | | | 582,622 | | | | |

*Notes:* The table considers the coefficient bounds based on Oster's approach. See Section A.1 for a detailed explanation.

Table A8: Auxiliary Analysis – Early Return Migration Decisions (Age 34-60, TH)

| Sample: | All | | Limited Sample | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Age | -0.005 | -0.004 | -0.005 | -0.005 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Female | -0.018 | -0.017 | -0.024 | -0.024 |
| | (0.007) | (0.007) | (0.008) | (0.008) |
| Years of schooling | -0.004 | -0.004 | -0.004 | -0.004 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| UNHCR | | | -0.164 | -0.163 |
| | | | (0.021) | (0.023) |
| ln (Distance to Thai border) | 0.077 | 0.066 | 0.084 | 0.058 |
| | (0.019) | (0.040) | (0.015) | (0.035) |
| Prop. of contamination | 1.493 | 1.313 | 1.433 | 1.495 |
| | (0.653) | (0.564) | (0.538) | (0.548) |
| ln (Areas) | 0.022 | 0.026 | 0.027 | 0.040 |
| | (0.022) | (0.017) | (0.021) | (0.017) |
| Province of birth FE | No | Yes | No | Yes |
| Mean (Dep. Var.) | 0.247 | 0.247 | 0.224 | 0.224 |
| Observations | 23,773 | 23,773 | 10,407 | 10,407 |
| R-squared | 0.023 | 0.041 | 0.059 | 0.077 |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. Robust standard errors, adjusted for clustering by district of birth, are reported in parentheses. Regressions use data about returnees aged 34-60 from Thailand from the Matched Samples. In columns 3 and 4, the analysis samples are limited to those who returned from the UNHCR and UNBRO camps whom we can identify. The dependent variable is an indicator variable equal to 1 if returnees returned to Cambodia before 1992 and 0 otherwise. "UNHCR" is an indicator variable for returnees from the UNHCR camp. "ln(Distance to Thai border)" is the logarithmic value of the shortest distance ($km$) from the centroid of each district to the Thai border. "Prop of Contamination" is the proportion of contaminated areas among the total village buffer zone areas (3.0 km radius) in each district. "ln (Areas)" is the logarithmic value of the total village buffer zone areas (3.0 km radius) in each district.

Table A9: Mechanism – Returnees from UNHCR vs. UNBRO Camps (Age 34-60, 6-19)

| Dependent Variable: | Years of Schooling | | Employed | | Primary Sector | | Tertiary Sector | | Low-skilled Work | | Armed Forces | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | | | | | | A. Age 34-60 | | | | | | |
| | | | | | | Men | | | | | | |
| UNHCR | | | -0.038 | -0.030 | -0.263 | -0.130 | 0.205 | 0.087 | 0.098 | 0.065 | 0.047 | 0.017 |
| | | | (0.006) | (0.007) | (0.016) | (0.014) | (0.014) | (0.013) | (0.009) | (0.008) | (0.006) | (0.006) |
| UNBRO | | | -0.023 | -0.023 | -0.268 | -0.119 | 0.217 | 0.073 | 0.065 | 0.050 | 0.078 | 0.014 |
| | | | (0.006) | (0.007) | (0.018) | (0.015) | (0.018) | (0.014) | (0.009) | (0.009) | (0.012) | (0.007) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | | | 0.028 | 0.339 | 0.781 | 0.478 | 0.502 | 0.337 | 0.001 | 0.201 | 0.013 | 0.735 |
| Mean (Dep. Var.) | | | 0.965 | 0.965 | 0.714 | 0.714 | 0.235 | 0.235 | 0.058 | 0.058 | 0.039 | 0.039 |
| Observations | | | 8,486 | 8,486 | 8,486 | 8,486 | 8,486 | 8,486 | 8,486 | 8,486 | 8,486 | 8,486 |
| R-squared | | | 0.038 | 0.069 | 0.222 | 0.340 | 0.203 | 0.311 | 0.062 | 0.123 | 0.058 | 0.196 |
| | | | | | | Women | | | | | | |
| UNHCR | | | -0.133 | -0.062 | -0.282 | -0.148 | 0.138 | 0.081 | 0.095 | 0.072 | 0.000 | -0.001 |
| | | | (0.013) | (0.014) | (0.017) | (0.016) | (0.011) | (0.010) | (0.008) | (0.008) | (0.001) | (0.001) |
| UNBRO | | | -0.094 | -0.050 | -0.209 | -0.126 | 0.117 | 0.080 | 0.059 | 0.048 | 0.003 | 0.002 |
| | | | (0.013) | (0.014) | (0.018) | (0.016) | (0.011) | (0.010) | (0.008) | (0.009) | (0.002) | (0.001) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | | | 0.004 | 0.368 | 0.000 | 0.120 | 0.097 | 0.941 | 0.000 | 0.013 | 0.115 | 0.093 |
| Mean (Dep. Var.) | | | 0.811 | 0.811 | 0.694 | 0.694 | 0.118 | 0.118 | 0.068 | 0.068 | 0.001 | 0.001 |
| Observations | | | 10,422 | 10,422 | 10,422 | 10,422 | 10,422 | 10,422 | 10,422 | 10,422 | 10,422 | 10,422 |
| R-squared | | | 0.079 | 0.151 | 0.174 | 0.303 | 0.103 | 0.163 | 0.055 | 0.094 | 0.009 | 0.152 |
| | | | | | | B. Age 15-19 | | | | | | |
| | | | | | | Household Heads | | | | | | |
| UNHCR | 0.524 | 0.422 | -0.024 | 0.004 | -0.083 | -0.031 | 0.049 | 0.037 | 0.047 | 0.045 | 0.002 | 0.001 |
| | (0.075) | (0.079) | (0.014) | (0.013) | (0.014) | (0.012) | (0.005) | (0.006) | (0.005) | (0.006) | (0.001) | (0.001) |
| UNBRO | 0.090 | 0.214 | -0.003 | 0.003 | -0.056 | -0.028 | 0.040 | 0.030 | 0.028 | 0.028 | 0.004 | 0.002 |
| | (0.085) | (0.088) | (0.016) | (0.015) | (0.016) | (0.015) | (0.006) | (0.006) | (0.005) | (0.006) | (0.001) | (0.001) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | 0.000 | 0.012 | 0.175 | 0.928 | 0.064 | 0.844 | 0.190 | 0.320 | 0.002 | 0.008 | 0.069 | 0.160 |
| Mean (Dep. Var.) | 4.269 | 4.269 | 0.376 | 0.376 | 0.336 | 0.336 | 0.035 | 0.035 | 0.027 | 0.027 | 0.001 | 0.001 |
| Observations | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 | 15,546 |
| R-squared | 0.211 | 0.253 | 0.124 | 0.369 | 0.131 | 0.348 | 0.043 | 0.085 | 0.041 | 0.078 | 0.008 | 0.021 |
| | | | | | | Parents | | | | | | |
| UNHCR | 0.532 | 0.458 | -0.018 | 0.010 | -0.070 | -0.017 | 0.043 | 0.031 | 0.042 | 0.040 | 0.001 | 0.000 |
| | (0.073) | (0.083) | (0.015) | (0.014) | (0.015) | (0.013) | (0.006) | (0.006) | (0.005) | (0.006) | (0.001) | (0.001) |
| UNBRO | 0.073 | 0.206 | -0.011 | 0.002 | -0.058 | -0.025 | 0.035 | 0.027 | 0.027 | 0.028 | 0.005 | 0.003 |
| | (0.090) | (0.098) | (0.018) | (0.017) | (0.017) | (0.017) | (0.006) | (0.007) | (0.005) | (0.007) | (0.002) | (0.001) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | 0.000 | 0.008 | 0.703 | 0.572 | 0.487 | 0.586 | 0.306 | 0.509 | 0.025 | 0.091 | 0.031 | 0.037 |
| Mean (Dep. Var.) | 4.330 | 4.330 | 0.365 | 0.365 | 0.333 | 0.333 | 0.029 | 0.029 | 0.022 | 0.022 | 0.001 | 0.001 |
| Observations | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 | 12,801 |
| R-squared | 0.271 | 0.299 | 0.146 | 0.383 | 0.159 | 0.367 | 0.065 | 0.112 | 0.058 | 0.101 | 0.018 | 0.040 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Education controls (panel B) | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| District FE | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |

*Continue*

Table A9: Mechanism – Returnees from UNHCR vs. UNBRO Camps (Age 34-60, 6-19)

| Dependent Variable: | Some Education | | | School Participation | | | Grade progression | | | Child labor | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | | | | | | C. Age 12-14 | | | | | | |
| | | | | | | Household Heads | | | | | | |
| UNHCR | 0.006 | -0.004 | -0.008 | -0.013 | -0.034 | -0.040 | 0.074 | -0.026 | -0.069 | -0.002 | 0.009 | 0.011 |
| | (0.012) | (0.013) | (0.013) | (0.013) | (0.014) | (0.014) | (0.060) | (0.063) | (0.062) | (0.006) | (0.007) | (0.007) |
| UNBRO | -0.032 | -0.035 | -0.024 | -0.034 | -0.050 | -0.039 | -0.091 | -0.045 | -0.027 | -0.003 | 0.008 | 0.009 |
| | (0.015) | (0.016) | (0.015) | (0.016) | (0.017) | (0.016) | (0.068) | (0.071) | (0.070) | (0.007) | (0.008) | (0.008) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | 0.011 | 0.028 | 0.274 | 0.185 | 0.278 | 0.972 | 0.017 | 0.782 | 0.543 | 0.895 | 0.935 | 0.807 |
| Mean (Dep. Var.) | 0.808 | 0.808 | 0.807 | 0.746 | 0.746 | 0.745 | -5.135 | -5.135 | -5.144 | 0.053 | 0.053 | 0.053 |
| Observations | 10,942 | 10,942 | 10,453 | 10,942 | 10,942 | 10,453 | 10,942 | 10,942 | 10,453 | 10,942 | 10,942 | 10,453 |
| R-squared | 0.108 | 0.138 | 0.162 | 0.105 | 0.134 | 0.156 | 0.186 | 0.229 | 0.247 | 0.056 | 0.070 | 0.070 |
| | | | | | | Parents | | | | | | |
| UNHCR | 0.000 | -0.009 | -0.013 | -0.025 | -0.044 | -0.048 | 0.080 | -0.025 | -0.072 | 0.001 | 0.013 | 0.013 |
| | (0.012) | (0.014) | (0.014) | (0.013) | (0.015) | (0.015) | (0.061) | (0.066) | (0.066) | (0.007) | (0.008) | (0.008) |
| UNBRO | -0.030 | -0.039 | -0.029 | -0.035 | -0.057 | -0.047 | -0.092 | -0.060 | -0.050 | -0.002 | 0.010 | 0.011 |
| | (0.016) | (0.018) | (0.017) | (0.017) | (0.020) | (0.019) | (0.072) | (0.080) | (0.077) | (0.007) | (0.009) | (0.009) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | 0.068 | 0.075 | 0.340 | 0.568 | 0.466 | 0.935 | 0.022 | 0.647 | 0.781 | 0.718 | 0.798 | 0.756 |
| Mean (Dep. Var.) | 0.817 | 0.817 | 0.816 | 0.758 | 0.758 | 0.757 | -5.090 | -5.090 | -5.098 | 0.050 | 0.050 | 0.049 |
| Observations | 9,344 | 9,344 | 8,940 | 9,344 | 9,344 | 8,940 | 9,344 | 9,344 | 8,940 | 9,344 | 9,344 | 8,940 |
| R-squared | 0.142 | 0.165 | 0.187 | 0.137 | 0.161 | 0.181 | 0.235 | 0.270 | 0.286 | 0.074 | 0.089 | 0.089 |
| | | | | | | D. Age 6-11 | | | | | | |
| | | | | | | Household Heads | | | | | | |
| UNHCR | -0.002 | -0.026 | -0.025 | -0.005 | -0.029 | -0.028 | -0.001 | -0.077 | -0.071 | 0.000 | 0.000 | 0.000 |
| | (0.012) | (0.013) | (0.013) | (0.012) | (0.013) | (0.013) | (0.025) | (0.027) | (0.027) | (0.001) | (0.001) | (0.001) |
| UNBRO | -0.039 | -0.041 | -0.034 | -0.043 | -0.043 | -0.037 | -0.081 | -0.109 | -0.097 | 0.003 | 0.003 | 0.003 |
| | (0.015) | (0.017) | (0.016) | (0.015) | (0.017) | (0.016) | (0.030) | (0.032) | (0.032) | (0.002) | (0.002) | (0.002) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | 0.016 | 0.338 | 0.556 | 0.015 | 0.336 | 0.519 | 0.010 | 0.299 | 0.410 | 0.094 | 0.079 | 0.085 |
| Mean (Dep. Var.) | 0.528 | 0.528 | 0.525 | 0.524 | 0.524 | 0.521 | -2.779 | -2.779 | -2.786 | 0.003 | 0.003 | 0.003 |
| Observations | 18,212 | 18,212 | 17,415 | 18,212 | 18,212 | 17,415 | 18,212 | 18,212 | 17,415 | 18,212 | 18,212 | 17,415 |
| R-squared | 0.228 | 0.253 | 0.257 | 0.223 | 0.249 | 0.252 | 0.572 | 0.588 | 0.590 | 0.014 | 0.024 | 0.025 |
| | | | | | | Parents | | | | | | |
| UNHCR | -0.006 | -0.031 | -0.030 | -0.010 | -0.035 | -0.033 | -0.009 | -0.082 | -0.075 | 0.001 | 0.001 | 0.001 |
| | (0.013) | (0.014) | (0.014) | (0.013) | (0.014) | (0.014) | (0.026) | (0.029) | (0.029) | (0.001) | (0.001) | (0.001) |
| UNBRO | -0.035 | -0.038 | -0.028 | -0.039 | -0.041 | -0.032 | -0.088 | -0.117 | -0.098 | 0.001 | 0.002 | 0.002 |
| | (0.016) | (0.018) | (0.017) | (0.016) | (0.018) | (0.018) | (0.031) | (0.035) | (0.035) | (0.002) | (0.002) | (0.002) |
| $p$-value ($H_0$: $\gamma_{HC} = \gamma_{BR}$) | 0.084 | 0.710 | 0.940 | 0.079 | 0.705 | 0.956 | 0.016 | 0.301 | 0.509 | 0.872 | 0.669 | 0.772 |
| Mean (Dep. Var.) | 0.531 | 0.531 | 0.528 | 0.527 | 0.527 | 0.524 | -2.749 | -2.749 | -2.754 | 0.003 | 0.003 | 0.003 |
| Observations | 15,782 | 15,782 | 15,117 | 15,782 | 15,782 | 15,117 | 15,782 | 15,782 | 15,117 | 15,782 | 15,782 | 15,117 |
| R-squared | 0.261 | 0.280 | 0.284 | 0.256 | 0.275 | 0.279 | 0.583 | 0.596 | 0.598 | 0.026 | 0.037 | 0.040 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Access to schools | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |
| District FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. Robust standard errors, adjusted for clustering by village, are reported in parentheses. Regressions in panel A use data about the returnees from the UNHCR and UNBRO camps and about stayers aged 34–60 from the Matched Samples. Regressions in panels B, C, and D use data about the children aged 15–19, 12–14, and 6–11, respectively, of returnees from UNHCR and UNBRO camps and of stayers, all aged 34–60; the samples are constructed based on the Matched Samples. The samples in panels B, C, and D for "Parents" are limited to the children of married couples. "UNHCR" ("UNBRO") in panel A is an indicator variable equal to 1 if individuals are returnees from UNHCR (UNBRO) camps and 0 otherwise; stayers are the base group. "UNHCR" ("UNBRO") in panels B, C, and D is an indicator variable equal to 1 if household heads and/or spouses are returnees from UNHCR (UNBRO) camps and 0 otherwise; children of stayers are the base group. Base controls in panel A include age, age squared, and years of schooling. Base controls in panels B, C, and D for "Household Heads" ("Parents") include age, a dummy variable for female, age of household head (father's age, mother's age), years of schooling for household head (years of schooling for father and mother), and an indicator variable for female household head. District of birth FE for "Household Heads" ("Parents") include district of birth fixed effects for household head (district of birth fixed effects for father and mother). Education controls (panel B only) include years of schooling and an indicator variable for attending school. Access to schools (panels C and D only) includes two continuous variables for the distance (km) to the nearest primary (panels C and D) and secondary (panel C) schools; the analysis samples are limited to individuals residing in villages with information about village points.

Table A10: Auxiliary Analysis – Return Migration Decisions to Birth Regions (Age 34-60)

| Sample: | Returnees from Thailand | | | | Returnees from Vietnam | | | |
|---|---|---|---|---|---|---|---|---|
| | All | | 1992-1993 | | All | | 1979-1980 | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Age | 0.003 | 0.002 | 0.001 | 0.000 | -0.005 | -0.003 | -0.004 | -0.002 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.002) |
| Female | 0.018 | 0.015 | 0.014 | 0.012 | -0.022 | -0.005 | -0.044 | -0.032 |
| | (0.011) | (0.011) | (0.016) | (0.016) | (0.022) | (0.019) | (0.028) | (0.026) |
| Years of schooling | -0.014 | -0.013 | -0.018 | -0.017 | -0.027 | -0.015 | -0.025 | -0.016 |
| | (0.004) | (0.003) | (0.004) | (0.004) | (0.006) | (0.005) | (0.009) | (0.006) |
| ln (Distance to Thai/Vietnam border) | -0.179 | -0.136 | -0.194 | -0.168 | -0.178 | -0.245 | -0.197 | -0.199 |
| | (0.039) | (0.078) | (0.044) | (0.093) | (0.040) | (0.067) | (0.056) | (0.084) |
| Prop. of contamination | -1.175 | -0.908 | -2.082 | -1.741 | | | | |
| | (1.615) | (1.658) | (1.782) | (1.897) | | | | |
| ln (Areas) | -0.041 | 0.008 | -0.029 | 0.041 | | | | |
| | (0.040) | (0.043) | (0.046) | (0.049) | | | | |
| Province of birth FE | No | Yes | No | Yes | No | Yes | No | Yes |
| Mean (Dep. Var.) | 0.392 | 0.392 | 0.470 | 0.470 | 0.593 | 0.593 | 0.733 | 0.733 |
| Observations | 23,773 | 23,773 | 14,766 | 14,766 | 3,156 | 3,156 | 1,994 | 1,994 |
| R-squared | 0.104 | 0.153 | 0.113 | 0.168 | 0.191 | 0.322 | 0.191 | 0.321 |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. Robust standard errors, adjusted for clustering by district of birth, are reported in parentheses. Regressions use data about returnees aged 34-60 from Thailand/Vietnam from the Matched Samples. In columns 3 and 4 (columns 7 and 8), the analysis samples are limited to the returnees from Thailand (Vietnam) who returned in 1992-1993 (1979-1980). The dependent variable is an indicator variable equal to 1 if returnees returned to their birth districts and 0 otherwise. "ln(Distance to Thai/Vietnam border)" is the logarithmic value of the shortest distance ($km$) from the centroid of each district to the Thai/Vietnamese border. "Prop of Contamination" is the proportion of contaminated areas among the total village buffer zone areas (3.0 km radius) in each district. "ln (Areas)" is the logarithmic value of the total village buffer zone areas (3.0 km radius) in each district.

Table A11: Consistency Check – Census 1998 (Age 20-60)

| Sample: | All | NW West North | SW East | All | NW West North | SW East |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable: | Years of Schooling | | | Employed | | |
| Refugee | 0.204 | 0.722 | -0.620 | -0.085 | -0.081 | -0.057 |
| | (0.083) | (0.088) | (0.259) | (0.007) | (0.008) | (0.010) |
| Mean (Dep. Var.) | 3.721 | 3.333 | 4.087 | 0.879 | 0.866 | 0.900 |
| Observations | 1,020,278 | 296,645 | 529,622 | 2,008,857 | 581,828 | 1,052,767 |
| R-squared | 0.056 | 0.053 | 0.073 | 0.040 | 0.044 | 0.035 |
| Dependent Variable: | Primary Sector | | | Secondary Sector | | |
| Refugee | -0.264 | -0.252 | -0.249 | 0.020 | 0.017 | 0.027 |
| | (0.014) | (0.016) | (0.026) | (0.002) | (0.002) | (0.004) |
| Mean (Dep. Var.) | 0.813 | 0.789 | 0.836 | 0.012 | 0.013 | 0.013 |
| Observations | 2,008,857 | 581,828 | 1,052,767 | 2,008,857 | 581,828 | 1,052,767 |
| R-squared | 0.021 | 0.030 | 0.014 | 0.001 | 0.002 | 0.001 |
| Dependent Variable: | Tertiary Sector | | | High-skilled Work | | |
| Refugee | 0.168 | 0.163 | 0.165 | 0.010 | 0.011 | 0.011 |
| | (0.008) | (0.010) | (0.016) | (0.002) | (0.002) | (0.003) |
| Mean (Dep. Var.) | 0.060 | 0.071 | 0.054 | 0.018 | 0.019 | 0.019 |
| Observations | 2,008,857 | 581,828 | 1,052,767 | 2,008,857 | 581,828 | 1,052,767 |
| R-squared | 0.020 | 0.033 | 0.015 | 0.008 | 0.008 | 0.010 |
| Dependent Variable: | Middle-skilled Work | | | Low-skilled Work | | |
| Refugee | -0.189 | -0.187 | -0.139 | 0.077 | 0.078 | 0.059 |
| | (0.010) | (0.012) | (0.017) | (0.005) | (0.006) | (0.008) |
| Mean (Dep. Var.) | 0.844 | 0.825 | 0.867 | 0.019 | 0.024 | 0.016 |
| Observations | 2,008,857 | 581,828 | 1,052,767 | 2,008,857 | 581,828 | 1,052,767 |
| R-squared | 0.024 | 0.030 | 0.018 | 0.007 | 0.015 | 0.002 |
| Dependent Variable: | Armed Forces | | | Home Ownership | | |
| Refugee | 0.025 | 0.027 | 0.013 | -0.033 | -0.033 | -0.032 |
| | (0.003) | (0.003) | (0.002) | (0.003) | (0.003) | (0.004) |
| Mean (Dep. Var.) | 0.003 | 0.006 | 0.002 | 0.983 | 0.983 | 0.985 |
| Observations | 2,008,857 | 581,828 | 1,052,767 | 2,008,857 | 581,828 | 1,052,767 |
| R-squared | 0.009 | 0.015 | 0.004 | 0.004 | 0.005 | 0.003 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | No | No | No | No | No | No |
| Years of schooling | No | No | No | No | No | No |
| District FE | No | No | No | No | No | No |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. Robust standard errors, adjusted for clustering by village, are reported in parentheses. Regressions use data about individuals (returnees from Thailand and Vietnam and stayers) aged 20–60 from the Full Samples. In columns 2 and 5 (columns 3 and 6), the analysis samples are limited to the individuals residing in the former Northwest (NW), West, and North zones (Southwest (SW) and East zones), and the data used in the regressions for "Years of Schooling" are limited to the individuals aged 20–33. For the definitions of the dependent variables, see the main text. "Refugee" is an indicator variable equal to 1 if individuals are returnees from Thailand or Vietnam and 0 otherwise; stayers are the base group. Base controls include age, age squared, and a dummy variable for female. Years of schooling and district FE (district fixed effects) are not controlled in all regressions.

Table A12: Consistency Check – CSES 2004 (Age 26-66)

| Sample: | All | NW West North | SW East | All | NW West North | SW East |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable: | Years of Schooling | | | Employed | | |
| Refugee | 1.204 | 1.695 | 0.490 | -0.064 | -0.105 | -0.017 |
| | (0.509) | (0.757) | (0.664) | (0.021) | (0.029) | (0.024) |
| Mean (Dep. Var.) | 4.069 | 3.967 | 4.347 | 0.923 | 0.921 | 0.928 |
| Observations | 6,386 | 2,208 | 2,943 | 13,378 | 4,718 | 6,203 |
| R-squared | 0.065 | 0.071 | 0.084 | 0.053 | 0.059 | 0.053 |
| Dependent Variable: | Primary Sector | | | Secondary Sector | | |
| Refugee | -0.373 | -0.482 | -0.218 | 0.029 | 0.043 | 0.013 |
| | (0.044) | (0.049) | (0.069) | (0.016) | (0.023) | (0.025) |
| Mean (Dep. Var.) | 0.716 | 0.717 | 0.723 | 0.046 | 0.042 | 0.051 |
| Observations | 13,378 | 4,718 | 6,203 | 13,378 | 4,718 | 6,203 |
| R-squared | 0.022 | 0.057 | 0.006 | 0.01 | 0.011 | 0.013 |
| Dependent Variable: | Tertiary Sector | | | High-skilled Work | | |
| Refugee | 0.276 | 0.326 | 0.188 | 0.101 | 0.125 | 0.068 |
| | (0.037) | (0.044) | (0.063) | (0.024) | (0.031) | (0.038) |
| Mean (Dep. Var.) | 0.160 | 0.161 | 0.153 | 0.037 | 0.038 | 0.036 |
| Observations | 13,378 | 4,718 | 6,203 | 13,378 | 4,718 | 6,203 |
| R-squared | 0.025 | 0.043 | 0.019 | 0.023 | 0.031 | 0.023 |
| Dependent Variable: | Middle-skilled Work | | | Low-skilled Work | | |
| Refugee | -0.235 | -0.305 | -0.138 | 0.053 | 0.065 | 0.030 |
| | (0.035) | (0.039) | (0.056) | (0.019) | (0.024) | (0.034) |
| Mean (Dep. Var.) | 0.835 | 0.830 | 0.841 | 0.043 | 0.045 | 0.044 |
| Observations | 13,378 | 4,718 | 6,203 | 13,378 | 4,718 | 6,203 |
| R-squared | 0.021 | 0.041 | 0.015 | 0.006 | 0.009 | 0.005 |
| Dependent Variable: | Armed Forces | | | Home Ownership | | |
| Refugee | 0.012 | 0.003 | 0.025 | -0.077 | -0.048 | -0.110 |
| | (0.007) | (0.008) | (0.014) | (0.021) | (0.019) | (0.042) |
| Mean (Dep. Var.) | 0.006 | 0.008 | 0.006 | 0.976 | 0.976 | 0.976 |
| Observations | 13,378 | 4,718 | 6,203 | 13,378 | 4,718 | 6,203 |
| R-squared | 0.008 | 0.01 | 0.011 | 0.011 | 0.008 | 0.017 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | No | No | No | No | No | No |
| Years of schooling | No | No | No | No | No | No |
| District FE | No | No | No | No | No | No |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. Robust standard errors, adjusted for clustering by village, are reported in parentheses. Regressions use data about individuals (returnees from abroad and stayers) aged 26–66 from the Cambodia Socio-economic Survey (CSES) 2004; age 26–66 corresponds to age 20-60 at the time of the 1998 Census. In columns 2 and 5 (columns 3 and 6), the analysis samples are limited to the individuals residing in the former Northwest (NW), West, and North zones (Southwest (SW) and East zones), and the analysis samples for "Years of Schooling" are limited to the individuals aged 26–39; age 26–39 corresponds to age 20–33 at the time of the 1998 Census. For the definitions of dependent variables, see the main text. "Refugee" is an indicator variable equal to 1 if individuals have lived abroad before and migrated to their current residence in 1979–1998 and 0 otherwise; stayers are the base group. Base controls include age, age squared, and a dummy variable for female. Years of schooling and district FE (district fixed effects) are not controlled in all regressions.

Table A13: Mechanism – Adjustment for Residence in 1998 (Age 20-60)

| Dep. Var.: | Employed | | Primary Sector | | Tertiary Sector | | Low-skilled Work | | Home Ownership | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Timing of Repat. | | | | | A-1. Age 34-60 (TH) | | | | | |
| 1979-1987 | -0.015 | 0.010 | -0.140 | -0.046 | 0.113 | 0.047 | 0.027 | 0.030 | -0.031 | -0.015 |
| | (0.014) | (0.014) | (0.021) | (0.019) | (0.018) | (0.016) | (0.010) | (0.010) | (0.009) | (0.009) |
| 1988-1990 | -0.020 | -0.018 | -0.184 | -0.111 | 0.158 | 0.082 | 0.042 | 0.073 | -0.022 | -0.023 |
| | (0.015) | (0.015) | (0.020) | (0.021) | (0.017) | (0.017) | (0.013) | (0.012) | (0.007) | (0.008) |
| 1991 | -0.040 | -0.033 | -0.268 | -0.157 | 0.220 | 0.107 | 0.058 | 0.079 | -0.014 | -0.015 |
| | (0.012) | (0.011) | (0.017) | (0.017) | (0.015) | (0.012) | (0.011) | (0.010) | (0.004) | (0.005) |
| 1992 | -0.062 | -0.031 | -0.217 | -0.121 | 0.138 | 0.079 | 0.079 | 0.061 | -0.026 | -0.011 |
| | (0.010) | (0.010) | (0.018) | (0.013) | (0.011) | (0.008) | (0.007) | (0.006) | (0.003) | (0.003) |
| 1993 | -0.081 | -0.044 | -0.258 | -0.139 | 0.167 | 0.092 | 0.089 | 0.066 | -0.035 | -0.017 |
| | (0.009) | (0.009) | (0.016) | (0.012) | (0.010) | (0.007) | (0.007) | (0.006) | (0.004) | (0.003) |
| 1994-1998 | -0.099 | -0.057 | -0.301 | -0.151 | 0.185 | 0.084 | 0.109 | 0.076 | -0.074 | -0.048 |
| | (0.011) | (0.011) | (0.023) | (0.020) | (0.016) | (0.013) | (0.013) | (0.010) | (0.008) | (0.006) |
| Mean (Dep. Var.) | 0.871 | 0.871 | 0.702 | 0.702 | 0.162 | 0.162 | 0.059 | 0.059 | 0.975 | 0.975 |
| Observations | 50,702 | 50,702 | 50,702 | 50,702 | 50,702 | 50,702 | 50,702 | 50,702 | 50,702 | 50,702 |
| R-squared | 0.099 | 0.130 | 0.157 | 0.240 | 0.162 | 0.210 | 0.050 | 0.081 | 0.025 | 0.064 |
| | | | | | B-1. Age 34-60 (VN) | | | | | |
| 1979 | 0.023 | 0.052 | -0.103 | 0.003 | 0.113 | 0.047 | 0.028 | 0.007 | -0.001 | 0.004 |
| | (0.013) | (0.013) | (0.027) | (0.021) | (0.018) | (0.015) | (0.009) | (0.009) | (0.005) | (0.006) |
| 1980 | -0.034 | 0.004 | -0.191 | -0.042 | 0.132 | 0.039 | 0.044 | 0.013 | -0.003 | 0.008 |
| | (0.020) | (0.020) | (0.035) | (0.026) | (0.022) | (0.013) | (0.011) | (0.011) | (0.006) | (0.007) |
| 1981-1998 | -0.099 | -0.017 | -0.444 | -0.167 | 0.276 | 0.121 | 0.130 | 0.077 | -0.056 | -0.036 |
| | (0.013) | (0.016) | (0.029) | (0.029) | (0.021) | (0.021) | (0.018) | (0.015) | (0.009) | (0.009) |
| Mean (Dep. Var.) | 0.899 | 0.899 | 0.794 | 0.794 | 0.099 | 0.099 | 0.029 | 0.029 | 0.988 | 0.988 |
| Observations | 30,085 | 30,085 | 30,085 | 30,085 | 30,085 | 30,085 | 30,085 | 30,085 | 30,085 | 30,085 |
| R-squared | 0.096 | 0.109 | 0.142 | 0.184 | 0.152 | 0.184 | 0.046 | 0.082 | 0.021 | 0.046 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |

*Continue*

Table A13: Mechanism – Adjustment for Residence in 1998 (Age 20-60)

| Dep. Var.: | Years of Schooling | | Employed | | Primary Sector | | Tertiary Sector | | Low-skilled Work | | Home Ownership | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Timing of Repat. | | | | | | A-2. Age 20-33 (TH) | | | | | | |
| 1979-1987 | 0.069 | 0.147 | -0.015 | 0.018 | -0.120 | -0.029 | 0.098 | 0.047 | 0.036 | 0.037 | -0.029 | -0.018 |
| | (0.283) | (0.203) | (0.025) | (0.021) | (0.031) | (0.026) | (0.020) | (0.019) | (0.015) | (0.014) | (0.014) | (0.014) |
| 1988-1990 | -0.424 | 0.424 | 0.001 | 0.026 | -0.127 | -0.049 | 0.123 | 0.067 | 0.031 | 0.044 | -0.016 | -0.013 |
| | (0.213) | (0.165) | (0.021) | (0.019) | (0.028) | (0.027) | (0.021) | (0.018) | (0.014) | (0.014) | (0.008) | (0.010) |
| 1991 | -0.112 | 0.350 | -0.044 | -0.002 | -0.208 | -0.094 | 0.161 | 0.089 | 0.043 | 0.061 | -0.023 | -0.024 |
| | (0.146) | (0.156) | (0.022) | (0.017) | (0.026) | (0.024) | (0.016) | (0.015) | (0.010) | (0.011) | (0.008) | (0.009) |
| 1992 | 0.523 | 0.317 | -0.055 | -0.010 | -0.200 | -0.096 | 0.123 | 0.075 | 0.072 | 0.060 | -0.031 | -0.014 |
| | (0.103) | (0.106) | (0.013) | (0.012) | (0.022) | (0.017) | (0.012) | (0.010) | (0.009) | (0.008) | (0.005) | (0.005) |
| 1993 | 0.734 | 0.470 | -0.081 | -0.028 | -0.236 | -0.110 | 0.139 | 0.078 | 0.080 | 0.064 | -0.048 | -0.027 |
| | (0.112) | (0.105) | (0.014) | (0.013) | (0.023) | (0.018) | (0.012) | (0.010) | (0.008) | (0.008) | (0.007) | (0.007) |
| 1994-1998 | 0.558 | 0.254 | -0.096 | -0.039 | -0.294 | -0.139 | 0.187 | 0.105 | 0.090 | 0.057 | -0.084 | -0.053 |
| | (0.119) | (0.111) | (0.013) | (0.014) | (0.028) | (0.023) | (0.021) | (0.017) | (0.014) | (0.013) | (0.009) | (0.008) |
| Mean (Dep. Var.) | 3.812 | 3.812 | 0.808 | 0.808 | 0.672 | 0.672 | 0.128 | 0.128 | 0.052 | 0.052 | 0.960 | 0.960 |
| Observations | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 | 26,467 |
| R-squared | 0.172 | 0.211 | 0.102 | 0.146 | 0.148 | 0.251 | 0.095 | 0.169 | 0.051 | 0.086 | 0.032 | 0.072 |
| | | | | | | B-2. Age 20-33 (VN) | | | | | | |
| 1979 | -0.283 | -0.477 | 0.067 | 0.091 | -0.025 | 0.043 | 0.075 | 0.044 | 0.012 | 0.000 | 0.008 | 0.016 |
| | (0.204) | (0.206) | (0.019) | (0.020) | (0.031) | (0.026) | (0.020) | (0.015) | (0.011) | (0.010) | (0.007) | (0.008) |
| 1980 | -0.947 | -1.170 | -0.010 | 0.025 | -0.095 | 0.009 | 0.072 | 0.022 | 0.014 | -0.016 | 0.004 | 0.010 |
| | (0.266) | (0.246) | (0.041) | (0.044) | (0.055) | (0.049) | (0.023) | (0.015) | (0.015) | (0.013) | (0.011) | (0.012) |
| 1981-1998 | -0.511 | -1.143 | -0.114 | 0.005 | -0.415 | -0.126 | 0.245 | 0.106 | 0.093 | 0.029 | -0.066 | -0.031 |
| | (0.201) | (0.226) | (0.022) | (0.026) | (0.032) | (0.035) | (0.024) | (0.021) | (0.017) | (0.015) | (0.012) | (0.013) |
| Mean (Dep. Var.) | 3.485 | 3.485 | 0.838 | 0.838 | 0.757 | 0.757 | 0.075 | 0.075 | 0.025 | 0.025 | 0.977 | 0.977 |
| Observations | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 | 16,217 |
| R-squared | 0.177 | 0.201 | 0.109 | 0.127 | 0.142 | 0.204 | 0.082 | 0.182 | 0.046 | 0.095 | 0.029 | 0.068 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Years of schooling | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| District FE | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. Robust standard errors, adjusted for clustering by village, are reported in parentheses. Regressions in panel A (panel B) use data about the returnees from Thailand (Vietnam) and stayers, all aged 20–33 from the Matched Samples. For the definitions of the dependent variables, see the main text. "19XX(-19XX)" in panel A (panel B) is an indicator variable equal to 1 if individuals are returnees from Thailand (Vietnam) who returned in 19XX(-19XX); stayers are the base group. Base controls include age, age squared, a dummy variable for female, and district of birth fixed effects.

Table A14: Mechanism – Adjustment for Residence in 1998 (Age 6-19)

| Controls: | Household Heads | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dependent Variable: | Years of Schooling | | Employed | | Primary Sector | | Tertiary Sector | | Low-skilled Work | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Timing of Repatriation | A-1-1. Age 15-19 (TH) | | | | | | | | | |
| 1979-1987 | 0.149 | 0.250 | -0.038 | -0.024 | -0.064 | -0.039 | 0.029 | 0.024 | 0.005 | 0.013 |
| | (0.232) | (0.207) | (0.032) | (0.026) | (0.032) | (0.026) | (0.011) | (0.011) | (0.007) | (0.008) |
| 1988-1990 | -0.623 | -0.029 | 0.040 | 0.033 | -0.006 | -0.008 | 0.039 | 0.037 | 0.036 | 0.050 |
| | (0.179) | (0.194) | (0.037) | (0.033) | (0.039) | (0.032) | (0.015) | (0.017) | (0.015) | (0.016) |
| 1991 | 0.101 | 0.222 | -0.024 | 0.009 | -0.067 | -0.014 | 0.039 | 0.028 | 0.032 | 0.038 |
| | (0.120) | (0.138) | (0.023) | (0.022) | (0.022) | (0.021) | (0.008) | (0.009) | (0.010) | (0.012) |
| 1992 | 0.492 | 0.375 | -0.009 | 0.005 | -0.067 | -0.034 | 0.043 | 0.033 | 0.039 | 0.037 |
| | (0.093) | (0.094) | (0.018) | (0.015) | (0.018) | (0.015) | (0.006) | (0.007) | (0.006) | (0.007) |
| 1993 | 0.523 | 0.361 | 0.013 | 0.032 | -0.050 | -0.006 | 0.050 | 0.038 | 0.043 | 0.041 |
| | (0.089) | (0.091) | (0.018) | (0.016) | (0.019) | (0.017) | (0.007) | (0.006) | (0.006) | (0.007) |
| 1994-1998 | 0.026 | -0.069 | -0.094 | -0.060 | -0.174 | -0.099 | 0.052 | 0.031 | 0.042 | 0.035 |
| | (0.161) | (0.136) | (0.021) | (0.021) | (0.019) | (0.020) | (0.011) | (0.011) | (0.011) | (0.010) |
| Mean (Dep. Var.) | 4.183 | 4.183 | 0.383 | 0.383 | 0.353 | 0.353 | 0.026 | 0.026 | 0.021 | 0.021 |
| Observations | 26,086 | 26,086 | 26,086 | 26,086 | 26,086 | 26,086 | 26,086 | 26,086 | 26,086 | 26,086 |
| R-squared | 0.210 | 0.236 | 0.131 | 0.372 | 0.135 | 0.351 | 0.037 | 0.069 | 0.034 | 0.062 |
| | A-1-2. Age 15-19 (VN) | | | | | | | | | |
| 1979 | 0.063 | -0.017 | -0.078 | -0.050 | -0.119 | -0.079 | 0.040 | 0.027 | 0.005 | -0.002 |
| | (0.194) | (0.192) | (0.047) | (0.045) | (0.045) | (0.045) | (0.014) | (0.012) | (0.009) | (0.007) |
| 1980 | -0.145 | -0.396 | -0.105 | -0.092 | -0.134 | -0.101 | 0.039 | 0.023 | 0.011 | 0.000 |
| | (0.216) | (0.213) | (0.038) | (0.042) | (0.036) | (0.039) | (0.015) | (0.012) | (0.010) | (0.010) |
| 1981-1998 | -0.509 | -0.847 | -0.052 | -0.020 | -0.194 | -0.115 | 0.121 | 0.081 | 0.049 | 0.039 |
| | (0.195) | (0.223) | (0.033) | (0.038) | (0.037) | (0.043) | (0.025) | (0.021) | (0.013) | (0.014) |
| Mean (Dep. Var.) | 4.075 | 4.075 | 0.385 | 0.385 | 0.369 | 0.369 | 0.016 | 0.016 | 0.010 | 0.010 |
| Observations | 19,644 | 19,644 | 19,644 | 19,644 | 19,644 | 19,644 | 19,644 | 19,644 | 19,644 | 19,644 |
| R-squared | 0.227 | 0.241 | 0.147 | 0.381 | 0.148 | 0.364 | 0.038 | 0.085 | 0.021 | 0.043 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Education controls | - | - | No | Yes | No | Yes | No | Yes | No | Yes |
| District FE | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |

*Continue*

Table A14: Mechanism – Adjustment for Residence in 1998 (Age 6-19)

| Controls: | Parents | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Dependent Variable: | Years of Schooling | | Employed | | Primary Sector | | Tertiary Sector | | Low-skilled Work | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Timing of Repatriation | B-1-1. Age 15-19 (TH) | | | | | | | | | |
| 1979-1987 | 0.088 | 0.237 | -0.064 | -0.040 | -0.084 | -0.048 | 0.021 | 0.016 | 0.002 | 0.008 |
| | (0.255) | (0.225) | (0.034) | (0.029) | (0.033) | (0.028) | (0.010) | (0.011) | (0.007) | (0.009) |
| 1988-1990 | -0.662 | -0.072 | 0.034 | 0.050 | -0.030 | -0.017 | 0.043 | 0.050 | 0.045 | 0.064 |
| | (0.204) | (0.244) | (0.041) | (0.038) | (0.042) | (0.037) | (0.018) | (0.023) | (0.017) | (0.021) |
| 1991 | 0.166 | 0.416 | -0.034 | 0.008 | -0.066 | 0.005 | 0.030 | 0.018 | 0.025 | 0.027 |
| | (0.131) | (0.142) | (0.025) | (0.024) | (0.024) | (0.024) | (0.009) | (0.010) | (0.010) | (0.011) |
| 1992 | 0.496 | 0.384 | -0.008 | 0.008 | -0.061 | -0.025 | 0.036 | 0.027 | 0.034 | 0.033 |
| | (0.100) | (0.105) | (0.019) | (0.016) | (0.019) | (0.016) | (0.006) | (0.007) | (0.006) | (0.007) |
| 1993 | 0.431 | 0.284 | 0.030 | 0.041 | -0.026 | 0.012 | 0.045 | 0.034 | 0.035 | 0.034 |
| | (0.090) | (0.095) | (0.020) | (0.018) | (0.020) | (0.018) | (0.008) | (0.008) | (0.007) | (0.008) |
| 1994-1998 | -0.044 | -0.174 | -0.079 | -0.054 | -0.149 | -0.082 | 0.050 | 0.030 | 0.044 | 0.039 |
| | (0.164) | (0.151) | (0.024) | (0.023) | (0.022) | (0.023) | (0.012) | (0.012) | (0.011) | (0.012) |
| Mean (Dep. Var.) | 4.240 | 4.240 | 0.370 | 0.370 | 0.347 | 0.347 | 0.021 | 0.021 | 0.016 | 0.016 |
| Observations | 21,946 | 21,946 | 21,946 | 21,946 | 21,946 | 21,946 | 21,946 | 21,946 | 21,946 | 21,946 |
| R-squared | 0.258 | 0.276 | 0.144 | 0.378 | 0.151 | 0.361 | 0.051 | 0.085 | 0.043 | 0.076 |
| | B-1-2. Age 15-19 (VN) | | | | | | | | | |
| 1979 | -0.007 | -0.113 | -0.089 | -0.059 | -0.117 | -0.081 | 0.026 | 0.021 | -0.004 | -0.003 |
| | (0.188) | (0.190) | (0.052) | (0.050) | (0.051) | (0.049) | (0.012) | (0.011) | (0.005) | (0.006) |
| 1980 | -0.065 | -0.316 | -0.109 | -0.101 | -0.126 | -0.098 | 0.031 | 0.019 | 0.007 | 0.003 |
| | (0.229) | (0.230) | (0.041) | (0.043) | (0.036) | (0.038) | (0.017) | (0.011) | (0.012) | (0.010) |
| 1981-1998 | -0.432 | -0.606 | -0.049 | 0.001 | -0.189 | -0.087 | 0.114 | 0.067 | 0.031 | 0.025 |
| | (0.204) | (0.235) | (0.036) | (0.042) | (0.039) | (0.049) | (0.026) | (0.018) | (0.012) | (0.013) |
| Mean (Dep. Var.) | 4.139 | 4.139 | 0.372 | 0.372 | 0.358 | 0.358 | 0.014 | 0.014 | 0.009 | 0.009 |
| Observations | 17,373 | 17,373 | 17,373 | 17,373 | 17,373 | 17,373 | 17,373 | 17,373 | 17,373 | 17,373 |
| R-squared | 0.263 | 0.273 | 0.154 | 0.382 | 0.158 | 0.368 | 0.062 | 0.108 | 0.043 | 0.070 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Education controls | - | - | No | Yes | No | Yes | No | Yes | No | Yes |
| District FE | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |

*Continue*

Table A14: Mechanism – Adjustment for Residence in 1998 (Age 6-19)

| Controls: | Household Heads | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Some Education | | | School Participation | | | Grade Progression | | | Child Labor | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Timing of repatriation | A-2-1. Age 12-14 (TH) | | | | | | | | | | | |
| 1979-1987 | -0.036 | -0.036 | -0.028 | -0.028 | -0.042 | -0.034 | -0.127 | -0.123 | -0.151 | -0.001 | 0.004 | -0.001 |
| | (0.026) | (0.026) | (0.026) | (0.027) | (0.027) | (0.028) | (0.128) | (0.125) | (0.130) | (0.015) | (0.015) | (0.016) |
| 1988-1990 | -0.012 | 0.020 | 0.033 | -0.053 | -0.044 | -0.033 | -0.283 | 0.101 | 0.151 | 0.019 | 0.025 | 0.017 |
| | (0.028) | (0.028) | (0.029) | (0.029) | (0.030) | (0.032) | (0.129) | (0.135) | (0.144) | (0.015) | (0.018) | (0.018) |
| 1991 | 0.026 | 0.003 | 0.004 | 0.027 | -0.031 | -0.030 | -0.091 | 0.014 | -0.003 | -0.011 | 0.013 | 0.009 |
| | (0.017) | (0.017) | (0.018) | (0.018) | (0.020) | (0.020) | (0.081) | (0.085) | (0.088) | (0.007) | (0.010) | (0.010) |
| 1992 | -0.002 | -0.011 | -0.015 | -0.019 | -0.038 | -0.042 | 0.122 | 0.030 | 0.010 | 0.009 | 0.017 | 0.018 |
| | (0.015) | (0.015) | (0.013) | (0.015) | (0.015) | (0.014) | (0.070) | (0.069) | (0.064) | (0.007) | (0.008) | (0.008) |
| 1993 | -0.017 | -0.025 | -0.014 | -0.045 | -0.064 | -0.055 | 0.112 | 0.005 | 0.024 | 0.018 | 0.027 | 0.028 |
| | (0.014) | (0.015) | (0.013) | (0.015) | (0.015) | (0.014) | (0.066) | (0.065) | (0.061) | (0.008) | (0.009) | (0.009) |
| 1994-1998 | -0.070 | -0.057 | -0.047 | -0.086 | -0.084 | -0.069 | -0.241 | -0.288 | -0.249 | -0.004 | 0.004 | 0.001 |
| | (0.028) | (0.024) | (0.025) | (0.026) | (0.024) | (0.024) | (0.119) | (0.112) | (0.115) | (0.008) | (0.008) | (0.009) |
| Mean (Dep. Var.) | 0.806 | 0.806 | 0.805 | 0.743 | 0.743 | 0.743 | -5.145 | -5.145 | -5.152 | 0.058 | 0.058 | 0.057 |
| Observations | 23,116 | 23,116 | 22,137 | 23,116 | 23,116 | 22,137 | 23,116 | 23,116 | 22,137 | 23,116 | 23,116 | 22,137 |
| R-squared | 0.101 | 0.123 | 0.148 | 0.103 | 0.125 | 0.147 | 0.178 | 0.207 | 0.225 | 0.049 | 0.058 | 0.060 |
| | A-2-2. Age 12-14 (VN) | | | | | | | | | | | |
| 1979 | -0.014 | -0.019 | 0.001 | 0.000 | -0.003 | 0.017 | 0.269 | 0.174 | 0.303 | -0.024 | -0.017 | -0.021 |
| | (0.037) | (0.038) | (0.041) | (0.037) | (0.039) | (0.042) | (0.176) | (0.180) | (0.191) | (0.023) | (0.025) | (0.025) |
| 1980 | -0.062 | -0.087 | -0.027 | -0.066 | -0.087 | -0.022 | -0.115 | -0.255 | 0.066 | -0.001 | 0.010 | -0.003 |
| | (0.040) | (0.044) | (0.039) | (0.040) | (0.045) | (0.040) | (0.181) | (0.189) | (0.199) | (0.022) | (0.024) | (0.026) |
| 1981-1998 | -0.083 | -0.109 | -0.063 | -0.095 | -0.120 | -0.078 | -0.245 | -0.366 | -0.166 | -0.016 | -0.001 | -0.003 |
| | (0.033) | (0.037) | (0.041) | (0.033) | (0.037) | (0.041) | (0.157) | (0.154) | (0.173) | (0.017) | (0.021) | (0.022) |
| Mean (Dep. Var.) | 0.805 | 0.805 | 0.805 | 0.751 | 0.751 | 0.752 | -5.155 | -5.155 | -5.152 | 0.055 | 0.055 | 0.055 |
| Observations | 16,145 | 16,145 | 15,467 | 16,145 | 16,145 | 15,467 | 16,145 | 16,145 | 15,467 | 16,145 | 16,145 | 15,467 |
| R-squared | 0.119 | 0.131 | 0.159 | 0.112 | 0.124 | 0.148 | 0.194 | 0.209 | 0.232 | 0.051 | 0.057 | 0.060 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Access to schools | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |
| District FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

*Continue*

Table A14: Mechanism – Adjustment for Residence in 1998 (Age 6-19)

| Controls: | Parents | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Some Education | | | School Participation | | | Grade Progression | | | Child Labor | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Timing of repatriation | B-2-1. Age 12-14 (TH) | | | | | | | | | | | |
| 1979-1987 | -0.029 | -0.034 | -0.030 | -0.025 | -0.041 | -0.037 | -0.134 | -0.126 | -0.207 | 0.004 | 0.006 | 0.000 |
| | (0.029) | (0.029) | (0.030) | (0.031) | (0.031) | (0.032) | (0.138) | (0.134) | (0.140) | (0.016) | (0.015) | (0.016) |
| 1988-1990 | -0.007 | 0.026 | 0.041 | -0.025 | -0.003 | 0.012 | -0.277 | 0.174 | 0.206 | 0.016 | 0.018 | 0.005 |
| | (0.029) | (0.031) | (0.032) | (0.032) | (0.034) | (0.036) | (0.133) | (0.154) | (0.161) | (0.017) | (0.020) | (0.020) |
| 1991 | 0.033 | 0.007 | 0.005 | 0.028 | -0.034 | -0.034 | -0.026 | 0.117 | 0.085 | -0.006 | 0.014 | 0.010 |
| | (0.016) | (0.018) | (0.018) | (0.017) | (0.021) | (0.020) | (0.084) | (0.097) | (0.102) | (0.008) | (0.011) | (0.011) |
| 1992 | -0.005 | -0.016 | -0.021 | -0.022 | -0.042 | -0.046 | 0.141 | 0.034 | 0.009 | 0.007 | 0.015 | 0.015 |
| | (0.015) | (0.016) | (0.014) | (0.015) | (0.016) | (0.015) | (0.073) | (0.074) | (0.069) | (0.008) | (0.008) | (0.008) |
| 1993 | -0.024 | -0.036 | -0.026 | -0.058 | -0.080 | -0.073 | 0.099 | -0.028 | -0.021 | 0.018 | 0.027 | 0.028 |
| | (0.015) | (0.016) | (0.015) | (0.016) | (0.017) | (0.016) | (0.070) | (0.070) | (0.067) | (0.008) | (0.009) | (0.009) |
| 1994-1998 | -0.071 | -0.069 | -0.065 | -0.088 | -0.094 | -0.083 | -0.276 | -0.343 | -0.310 | -0.006 | 0.006 | 0.004 |
| | (0.029) | (0.025) | (0.025) | (0.027) | (0.025) | (0.025) | (0.123) | (0.114) | (0.115) | (0.008) | (0.009) | (0.010) |
| Mean (Dep. Var.) | 0.816 | 0.816 | 0.815 | 0.759 | 0.759 | 0.758 | -5.093 | -5.093 | -5.099 | 0.052 | 0.052 | 0.052 |
| Observations | 19,728 | 19,728 | 18,922 | 19,728 | 19,728 | 18,922 | 19,728 | 19,728 | 18,922 | 19,728 | 19,728 | 18,922 |
| R-squared | 0.124 | 0.141 | 0.164 | 0.122 | 0.138 | 0.158 | 0.214 | 0.236 | 0.252 | 0.055 | 0.064 | 0.066 |
| | B-2-2. Age 12-14 (VN) | | | | | | | | | | | |
| 1979 | -0.022 | -0.024 | -0.003 | -0.002 | -0.006 | 0.017 | 0.351 | 0.258 | 0.370 | -0.022 | -0.019 | -0.025 |
| | (0.039) | (0.042) | (0.045) | (0.040) | (0.042) | (0.045) | (0.194) | (0.196) | (0.208) | (0.024) | (0.026) | (0.027) |
| 1980 | -0.059 | -0.086 | -0.031 | -0.063 | -0.098 | -0.041 | -0.094 | -0.224 | 0.076 | -0.001 | 0.010 | -0.001 |
| | (0.042) | (0.047) | (0.042) | (0.043) | (0.049) | (0.045) | (0.181) | (0.194) | (0.197) | (0.022) | (0.024) | (0.026) |
| 1981-1998 | -0.088 | -0.100 | -0.051 | -0.100 | -0.111 | -0.074 | -0.283 | -0.364 | -0.121 | -0.012 | -0.008 | -0.006 |
| | (0.039) | (0.041) | (0.046) | (0.039) | (0.043) | (0.047) | (0.185) | (0.174) | (0.195) | (0.018) | (0.022) | (0.023) |
| Mean (Dep. Var.) | 0.813 | 0.813 | 0.813 | 0.763 | 0.763 | 0.763 | -5.112 | -5.112 | -5.107 | 0.051 | 0.051 | 0.051 |
| Observations | 14,530 | 14,530 | 13,952 | 14,530 | 14,530 | 13,952 | 14,530 | 14,530 | 13,952 | 14,530 | 14,530 | 13,952 |
| R-squared | 0.137 | 0.148 | 0.173 | 0.127 | 0.138 | 0.160 | 0.226 | 0.238 | 0.258 | 0.057 | 0.065 | 0.069 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Access to schools | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |
| District FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

*Continue*

Table A14: Mechanism – Adjustment for Residence in 1998 (Age 6-19)

| Controls: | Household Heads | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Some Education | | | School Participation | | | Grade Progression | | | Child Labor | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Timing of repatriation | A-3-1. Age 6-11 (TH) | | | | | | | | | | | |
| 1979-1987 | -0.031 | -0.028 | -0.027 | -0.030 | -0.025 | -0.024 | -0.078 | -0.082 | -0.087 | 0.006 | 0.006 | 0.004 |
| | (0.026) | (0.025) | (0.025) | (0.027) | (0.025) | (0.025) | (0.047) | (0.045) | (0.046) | (0.005) | (0.005) | (0.005) |
| 1988-1990 | -0.072 | -0.026 | -0.028 | -0.074 | -0.024 | -0.026 | -0.111 | -0.011 | -0.030 | 0.008 | 0.009 | 0.007 |
| | (0.028) | (0.023) | (0.023) | (0.028) | (0.024) | (0.024) | (0.052) | (0.053) | (0.054) | (0.004) | (0.005) | (0.005) |
| 1991 | -0.015 | -0.019 | -0.017 | -0.017 | -0.020 | -0.017 | -0.066 | -0.084 | -0.074 | 0.001 | 0.003 | 0.003 |
| | (0.020) | (0.019) | (0.020) | (0.020) | (0.018) | (0.019) | (0.034) | (0.037) | (0.038) | (0.001) | (0.003) | (0.003) |
| 1992 | -0.001 | -0.011 | -0.015 | -0.005 | -0.015 | -0.019 | 0.002 | -0.036 | -0.039 | 0.000 | 0.001 | 0.001 |
| | (0.015) | (0.014) | (0.013) | (0.015) | (0.014) | (0.014) | (0.028) | (0.027) | (0.027) | (0.002) | (0.002) | (0.002) |
| 1993 | 0.004 | -0.014 | -0.005 | 0.000 | -0.018 | -0.009 | 0.009 | -0.062 | -0.054 | 0.000 | 0.000 | 0.000 |
| | (0.014) | (0.014) | (0.013) | (0.014) | (0.014) | (0.013) | (0.027) | (0.026) | (0.026) | (0.001) | (0.001) | (0.001) |
| 1994-1998 | -0.041 | -0.054 | -0.052 | -0.045 | -0.058 | -0.058 | -0.083 | -0.157 | -0.161 | 0.000 | 0.001 | 0.001 |
| | (0.022) | (0.019) | (0.020) | (0.022) | (0.019) | (0.019) | (0.043) | (0.039) | (0.039) | (0.002) | (0.002) | (0.002) |
| Mean (Dep. Var.) | 0.519 | 0.519 | 0.517 | 0.515 | 0.515 | 0.512 | -2.770 | -2.770 | -2.775 | 0.004 | 0.004 | 0.004 |
| Observations | 41,103 | 41,103 | 39,403 | 41,103 | 41,103 | 39,403 | 41,103 | 41,103 | 39,403 | 41,103 | 41,103 | 39,403 |
| R-squared | 0.222 | 0.241 | 0.244 | 0.216 | 0.236 | 0.239 | 0.573 | 0.585 | 0.585 | 0.009 | 0.013 | 0.017 |
| | A-3-2. Age 6-11 (VN) | | | | | | | | | | | |
| 1979 | -0.024 | -0.038 | -0.043 | -0.027 | -0.041 | -0.047 | 0.083 | 0.023 | 0.016 | 0.007 | 0.009 | 0.009 |
| | (0.038) | (0.040) | (0.041) | (0.038) | (0.039) | (0.040) | (0.092) | (0.091) | (0.094) | (0.009) | (0.010) | (0.010) |
| 1980 | -0.023 | -0.030 | -0.029 | -0.021 | -0.029 | -0.028 | 0.000 | -0.049 | -0.044 | -0.002 | -0.001 | -0.002 |
| | (0.035) | (0.037) | (0.038) | (0.035) | (0.037) | (0.038) | (0.078) | (0.081) | (0.086) | (0.005) | (0.006) | (0.006) |
| 1981-1998 | -0.005 | -0.026 | -0.011 | -0.017 | -0.035 | -0.020 | 0.043 | -0.092 | -0.080 | -0.002 | 0.001 | 0.001 |
| | (0.031) | (0.031) | (0.032) | (0.030) | (0.030) | (0.031) | (0.060) | (0.062) | (0.066) | (0.003) | (0.004) | (0.005) |
| Mean (Dep. Var.) | 0.519 | 0.519 | 0.519 | 0.516 | 0.516 | 0.516 | -2.756 | -2.756 | -2.755 | 0.004 | 0.004 | 0.004 |
| Observations | 28,209 | 28,209 | 27,055 | 28,209 | 28,209 | 27,055 | 28,209 | 28,209 | 27,055 | 28,209 | 28,209 | 27,055 |
| R-squared | 0.234 | 0.244 | 0.246 | 0.230 | 0.240 | 0.242 | 0.574 | 0.579 | 0.578 | 0.012 | 0.013 | 0.014 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Access to schools | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |
| District FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

*Continue*

Table A14: Mechanism – Adjustment for Residence in 1998 (Age 6-19)

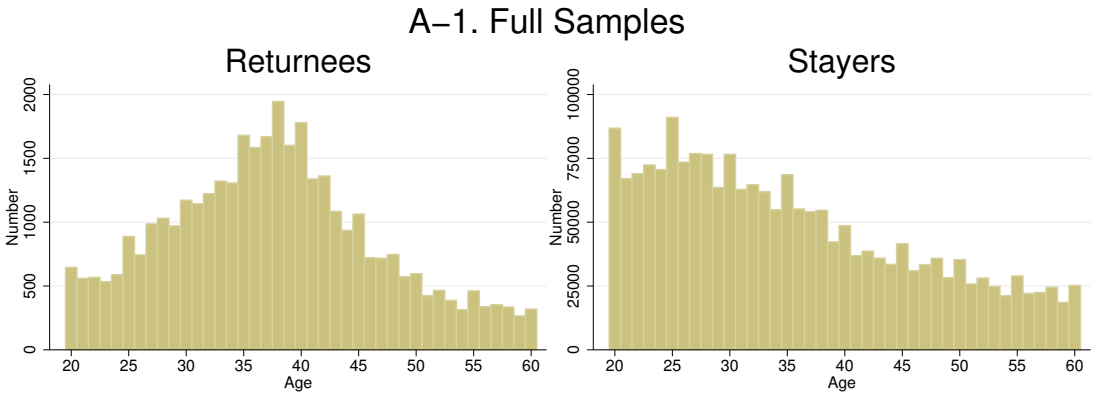
| Controls: | Parents | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Some Education | | | School Participation | | | Grade Progression | | | Child Labor | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Timing of repatriation | B-3-1. Age 6-11 (TH) | | | | | | | | | | | |
| 1979-1987 | -0.039 | -0.032 | -0.031 | -0.038 | -0.030 | -0.029 | -0.088 | -0.093 | -0.102 | 0.008 | 0.007 | 0.005 |
| | (0.028) | (0.027) | (0.027) | (0.029) | (0.027) | (0.028) | (0.049) | (0.051) | (0.052) | (0.006) | (0.005) | (0.005) |
| 1988-1990 | -0.041 | -0.002 | 0.002 | -0.042 | 0.001 | 0.005 | -0.069 | 0.024 | 0.018 | 0.000 | 0.000 | -0.002 |
| | (0.029) | (0.027) | (0.027) | (0.030) | (0.027) | (0.027) | (0.055) | (0.060) | (0.061) | (0.002) | (0.003) | (0.003) |
| 1991 | -0.007 | -0.011 | -0.007 | -0.009 | -0.012 | -0.006 | -0.062 | -0.082 | -0.073 | 0.001 | 0.004 | 0.003 |
| | (0.019) | (0.020) | (0.021) | (0.019) | (0.020) | (0.021) | (0.035) | (0.039) | (0.040) | (0.002) | (0.003) | (0.003) |
| 1992 | 0.000 | -0.011 | -0.015 | -0.004 | -0.015 | -0.018 | 0.010 | -0.032 | -0.032 | 0.002 | 0.002 | 0.002 |
| | (0.015) | (0.015) | (0.014) | (0.015) | (0.015) | (0.014) | (0.029) | (0.029) | (0.029) | (0.002) | (0.002) | (0.002) |
| 1993 | 0.010 | -0.012 | -0.001 | 0.006 | -0.016 | -0.005 | 0.012 | -0.066 | -0.059 | 0.000 | 0.001 | 0.000 |
| | (0.014) | (0.015) | (0.014) | (0.014) | (0.015) | (0.014) | (0.028) | (0.027) | (0.027) | (0.001) | (0.001) | (0.001) |
| 1994-1998 | -0.043 | -0.056 | -0.055 | -0.049 | -0.062 | -0.060 | -0.095 | -0.180 | -0.179 | 0.000 | 0.000 | 0.001 |
| | (0.021) | (0.021) | (0.021) | (0.022) | (0.021) | (0.022) | (0.042) | (0.042) | (0.043) | (0.001) | (0.002) | (0.002) |
| Mean (Dep. Var.) | 0.523 | 0.523 | 0.521 | 0.519 | 0.519 | 0.517 | -2.743 | -2.743 | -2.746 | 0.004 | 0.004 | 0.003 |
| Observations | 35,528 | 35,528 | 34,088 | 35,528 | 35,528 | 34,088 | 35,528 | 35,528 | 34,088 | 35,528 | 35,528 | 34,088 |
| R-squared | 0.243 | 0.257 | 0.259 | 0.237 | 0.252 | 0.254 | 0.582 | 0.590 | 0.590 | 0.016 | 0.024 | 0.028 |
| | B-3-2. Age 6-11 (VN) | | | | | | | | | | | |
| 1979 | -0.029 | -0.040 | -0.048 | -0.030 | -0.042 | -0.050 | 0.067 | 0.020 | 0.012 | 0.004 | 0.004 | 0.004 |
| | (0.041) | (0.042) | (0.044) | (0.041) | (0.042) | (0.044) | (0.096) | (0.093) | (0.097) | (0.005) | (0.006) | (0.006) |
| 1980 | -0.035 | -0.045 | -0.046 | -0.035 | -0.045 | -0.047 | -0.026 | -0.059 | -0.070 | -0.002 | -0.002 | -0.003 |
| | (0.035) | (0.037) | (0.038) | (0.035) | (0.037) | (0.038) | (0.074) | (0.074) | (0.077) | (0.005) | (0.005) | (0.006) |
| 1981-1998 | 0.007 | 0.007 | 0.028 | -0.006 | -0.003 | 0.018 | 0.061 | -0.003 | 0.031 | -0.003 | -0.002 | -0.002 |
| | (0.033) | (0.034) | (0.036) | (0.031) | (0.032) | (0.034) | (0.062) | (0.065) | (0.070) | (0.003) | (0.004) | (0.005) |
| Mean (Dep. Var.) | 0.520 | 0.520 | 0.520 | 0.517 | 0.517 | 0.517 | -2.739 | -2.739 | -2.738 | 0.004 | 0.004 | 0.003 |
| Observations | 25,688 | 25,688 | 24,659 | 25,688 | 25,688 | 24,659 | 25,688 | 25,688 | 24,659 | 25,688 | 25,688 | 24,659 |
| R-squared | 0.253 | 0.259 | 0.261 | 0.248 | 0.254 | 0.256 | 0.582 | 0.586 | 0.585 | 0.018 | 0.021 | 0.023 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| District of birth FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Access to schools | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |
| District FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

*Notes:* The table reports OLS estimates where the unit of observation is the individual. Robust standard errors, adjusted for clustering by village, are reported in parentheses. Regressions use data about the children aged 15–19, 12–14, and 6–11 of the returnees from Thailand (Vietnam) and of stayers, all aged 34–60, in panels A/B-1/2/3-1 (A/B-1/2/3-2), respectively; the samples are constructed based on the Matched Samples. The samples in panels A/B-1/2/3-2 are limited to the children of married couples. For the definitions of the dependent variables, see the main text. "19XX(-19XX)" is an indicator variable equal to 1 if household heads and/or spouses are returnees from Thailand (panels A/B-1/2/3-1)/ Vietnam (panels A/B-1/2/3-2) who returned in 19XX(-19XX); children of stayers are the base group. Base controls in A/B-1/2/3-1 (panels A/B-1/2/3-2) include age, a dummy variable for female, age of household head (father's age, mother's age), years of schooling for household head (years of schooling for father and mother), and an indicator variable for female head. District of birth FE in A/B-1/2/3-1 (panels A/B-1/2/3-2) include district of birth fixed effects for household head (district of birth fixed effects for father and mother). Access to schools (panels A/B-2/3-1/2 only) includes two continuous variables of the distance (km) to the nearest primary (panels A/B-2/3-1/2) and secondary (panels A/B-2-1/2) schools; the analysis samples are limited to the individuals residing in villages with information about village points.

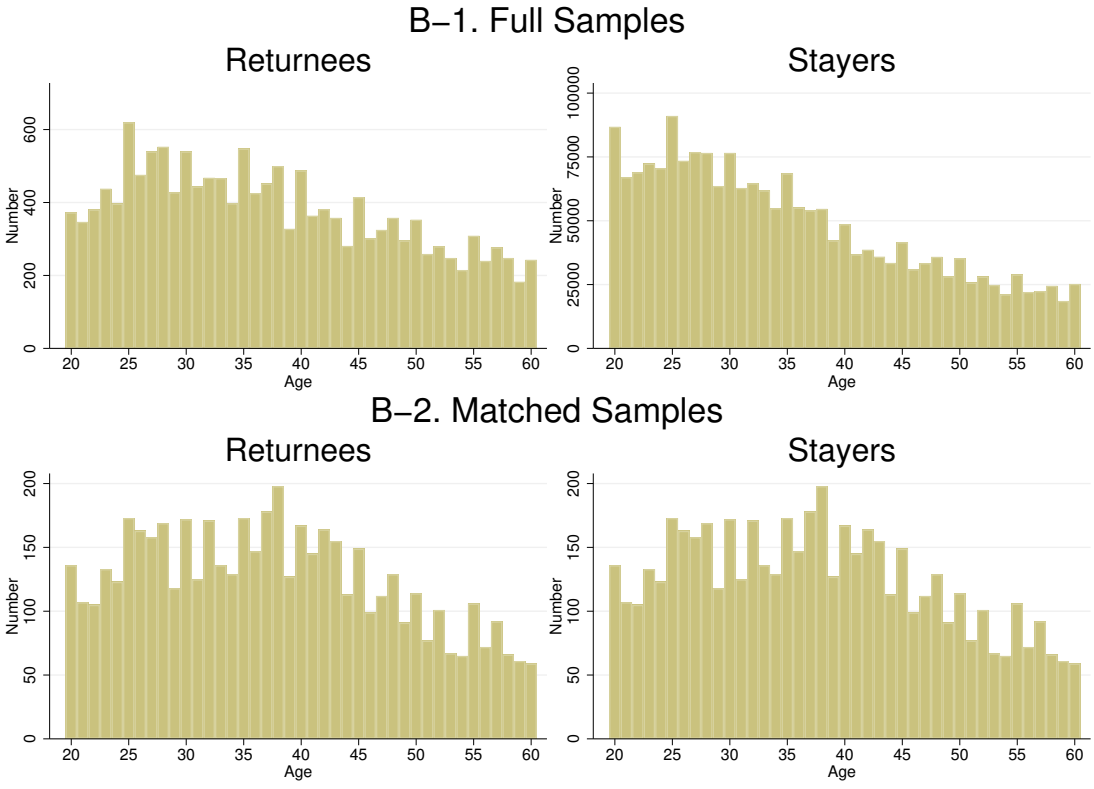
Table A15: Mechanism – Other Potential Channels

| Sample: | All | NW West North | SW East | All | NW West North | SW East |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Channel: | A. Discrimination | | | B. Health | | |
| Dependent Variable: | Neighborhood Trust | | | Health Status I | | |
| Refugee | 0.024 | 0.024 | 0.016 | -0.020 | 0.039 | -0.074 |
| | (0.048) | (0.053) | (0.095) | (0.035) | (0.049) | (0.050) |
| Mean (Dep. Var.) | 0.499 | 0.530 | 0.473 | 0.129 | 0.092 | 0.150 |
| Observations | 6,086 | 2,245 | 2,752 | 13,378 | 4,718 | 6,203 |
| R-squared | 0.206 | 0.272 | 0.187 | 0.093 | 0.092 | 0.087 |
| Channel: | B. Health | | | | | |
| Dependent Variable: | Health Status II | | | Disability | | |
| Refugee | -0.020 | 0.022 | -0.062 | 0.008 | 0.037 | -0.024 |
| | (0.029) | (0.033) | (0.053) | (0.022) | (0.029) | (0.036) |
| Mean (Dep. Var.) | 0.132 | 0.098 | 0.153 | 0.073 | 0.072 | 0.074 |
| Observations | 13,378 | 4,718 | 6,203 | 13,378 | 4,718 | 6,203 |
| R-squared | 0.077 | 0.090 | 0.070 | 0.077 | 0.097 | 0.069 |
| Channel: | C. Motivation | | | | | |
| | Domestic Remittances | | | | | |
| Dependent Variable: | Receipt | | | Amount (USD) | | |
| Refugee | 0.019 | 0.024 | 0.014 | 1.432 | 1.731 | 3.285 |
| | (0.028) | (0.038) | (0.039) | (4.205) | (6.034) | (5.381) |
| Mean (Dep. Var.) | 0.110 | 0.089 | 0.126 | 10.28 | 6.57 | 11.71 |
| Observations | 6,087 | 2,245 | 2,752 | 6,087 | 2,245 | 2,752 |
| R-squared | 0.123 | 0.124 | 0.112 | 0.045 | 0.069 | 0.038 |
| | International Remittances | | | | | |
| Dependent Variable: | Receipt | | | Amount (USD) | | |
| Refugee | 0.027 | 0.044 | 0.005 | 40.63 | 89.81 | -28.41 |
| | (0.027) | (0.035) | (0.046) | (35.34) | (56.85) | (25.06) |
| Mean (Dep. Var.) | 0.041 | 0.062 | 0.030 | 15.42 | 22.16 | 11.29 |
| Observations | 6,087 | 2,245 | 2,752 | 6,087 | 2,245 | 2,752 |
| R-squared | 0.087 | 0.085 | 0.082 | 0.042 | 0.056 | 0.042 |
| Base controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Years of schooling | Yes | Yes | Yes | Yes | Yes | Yes |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes:* The table reports OLS estimates where the unit of observation is the household head (panels A and C) and the individual (panel B). It reports robust standard errors, adjusted for clustering by village, in parentheses. Regressions use data about household heads/individuals aged 26–66 from the Cambodia Socio-Economic Survey 2004; age 26–66 corresponds to age 20–60 at the time of the 1998 Census. In columns 2 and 5 (columns 3 and 6), the analysis samples are limited to the household heads/individuals residing in the former Northwest (NW), West, and North zones (Southwest (SW) and East zones). "Neighborhood Trust" is an indicator variable equal to 1 if household heads feel safe from crime and violence in their neighborhood and 0 otherwise; one observation is missing in column 1 of panel A. "Health Status I" is an indicator variable equal to 1 if individuals report that their health is "very good" or "good" and 0 otherwise. "Health Status II" is an indicator variable equal to 1 if individuals report that their health is "much better" or "somewhat better" than others of the same age. "Disability" is an indicator variable equal to 1 if individuals have any disability and 0 otherwise. "Receipt" of domestic (international) remittances is an indicator variable equal to 1 if households receive remittances from relatives or others in Cambodia (from abroad). "Amount (USD)" of domestic (international) remittances is the remittance amount in dollars from relatives or others in Cambodia (from abroad). "Refugee" is an indicator variable equal to 1 if household heads/individuals have lived abroad before and migrated to their current residence in 1979–1998 and 0 otherwise. Base controls include age, age squared, and a dummy variable for female.



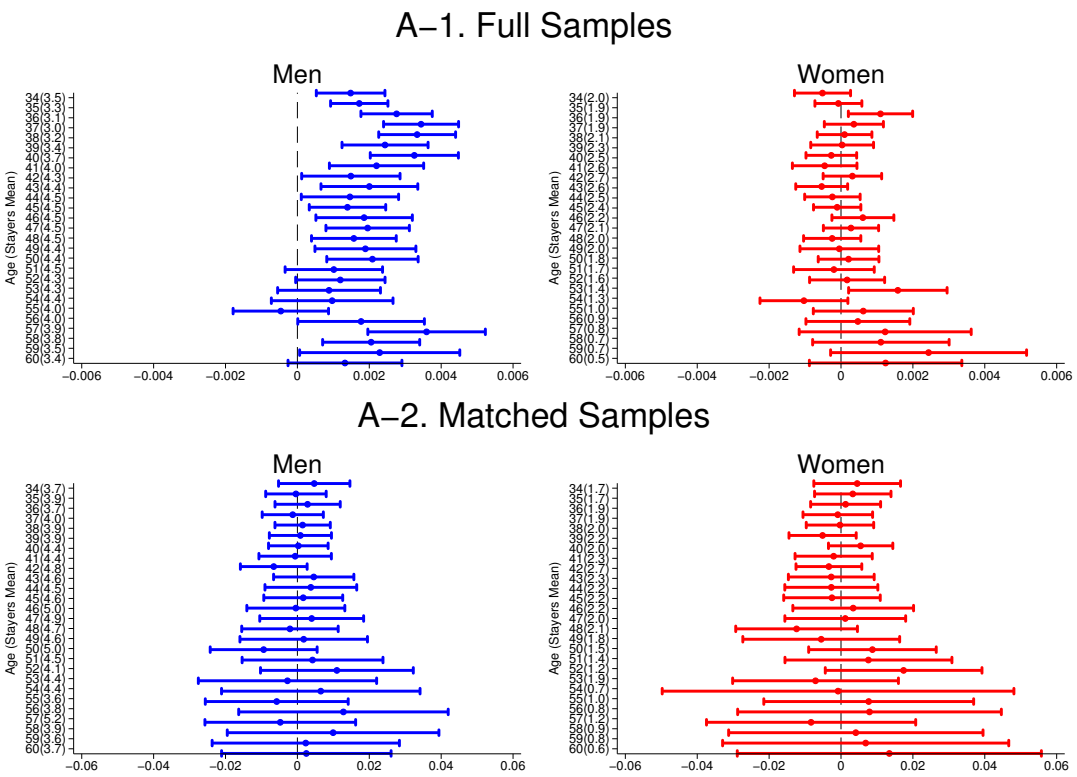

A. Thailand



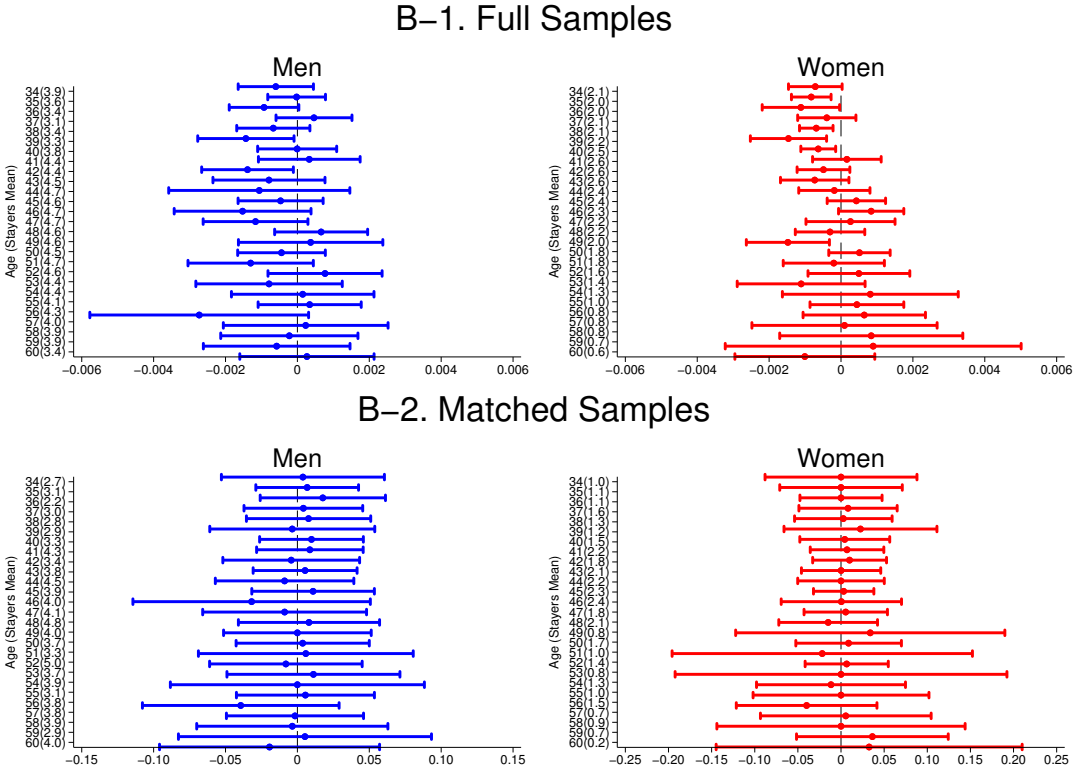

B. Vietnam

Figure A1: Age Distribution of Returnees and Stayers

*Note*: The figure shows age distribution for returnees from Thailand (panel A) and Vietnam (panel B) and for stayers aged 20-60 in the Full Samples and Matched Samples.




A. Thailand

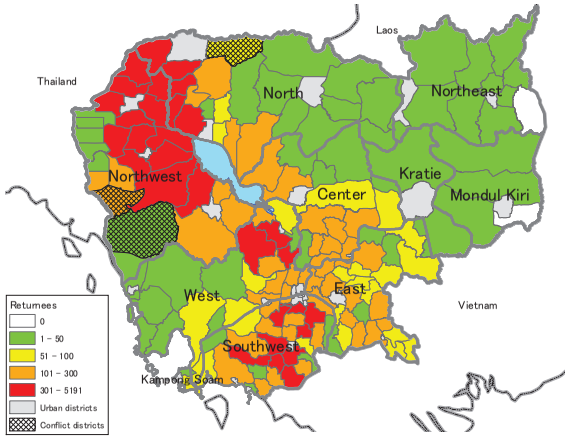



B. Vietnam

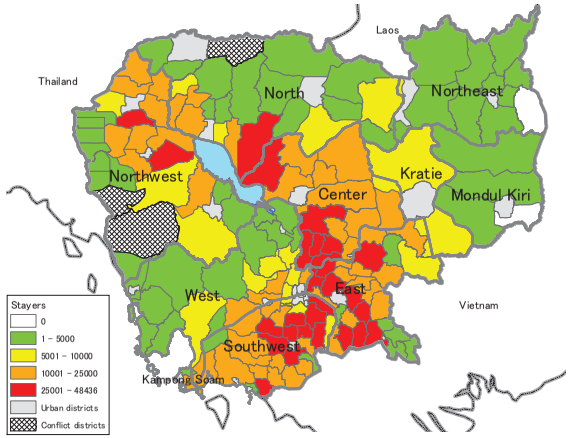
Figure A2: Difference in Education between Returnees and Stayers

*Note*: The figure plots the point estimates and 95% confidence intervals of the coefficients of years of schooling for cohorts (returnees from Thailand (panel A) and from Vietnam (panel B) and stayers) aged 34-60 by sex in the Full Samples and Matched Samples, adjusting for district of birth fixed effects using ordinary least squares (OLS); the stayers means are also shown.

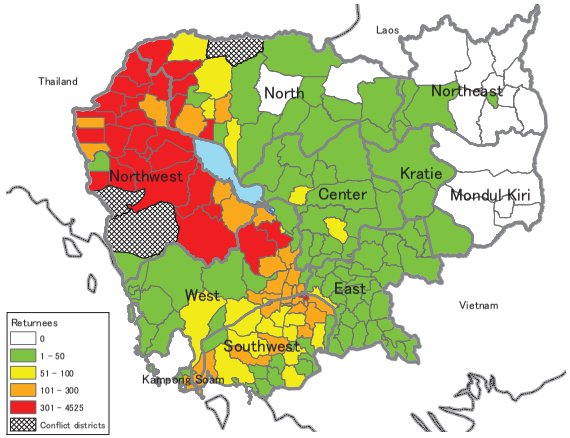
A-1. Full Samples




Birth Districts (Returnees)

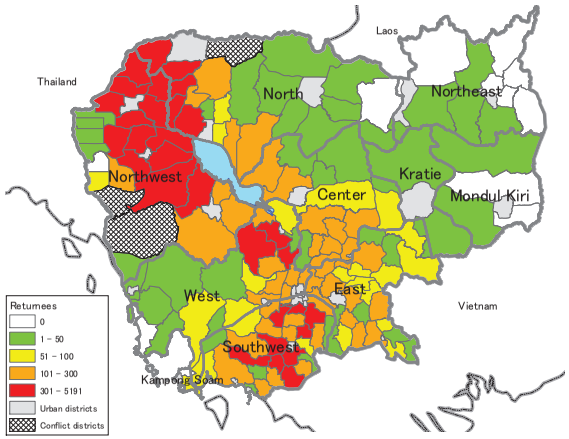



Birth Districts (Stayers)

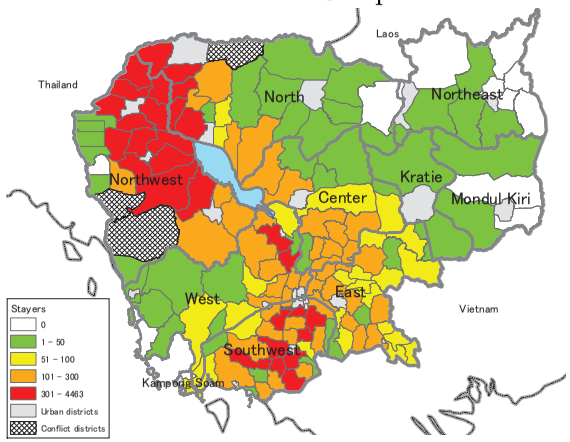



Residence in 1998 (Returnees)

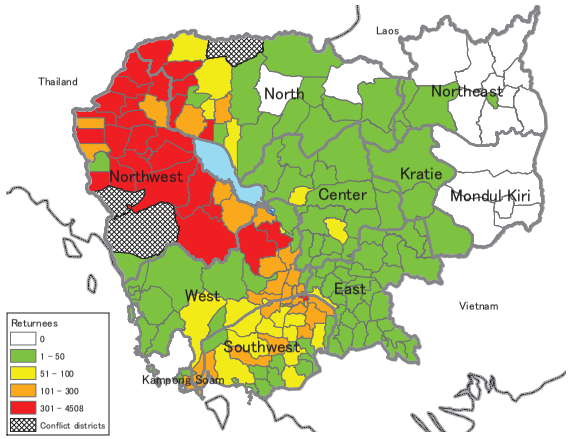
A-2. Matched Samples




Birth Districts (Returnees)




Birth Districts (Stayers)

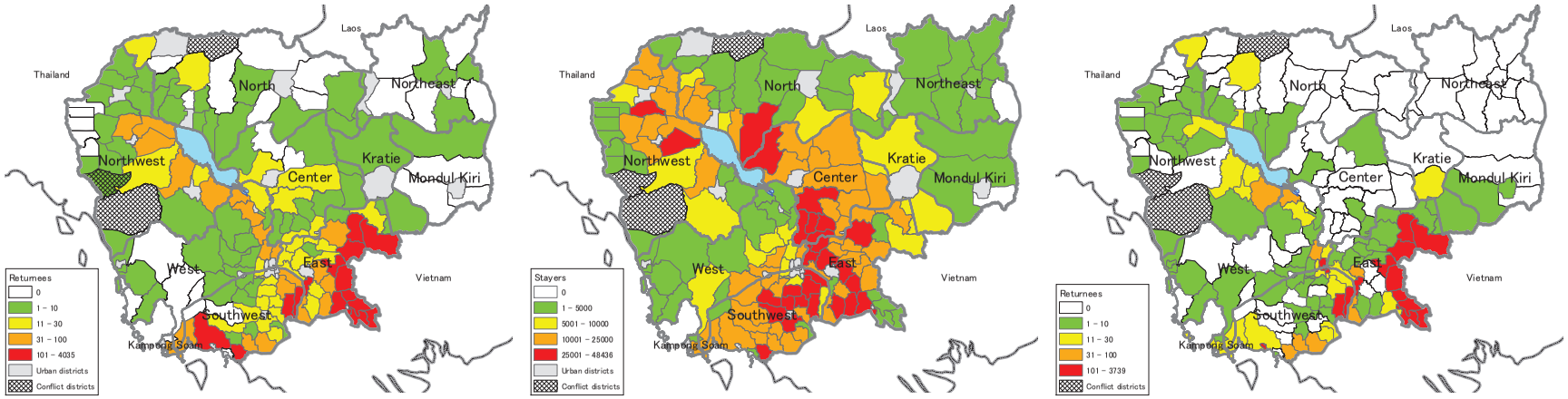



Residence in 1998 (Returnees)

A. Thailand

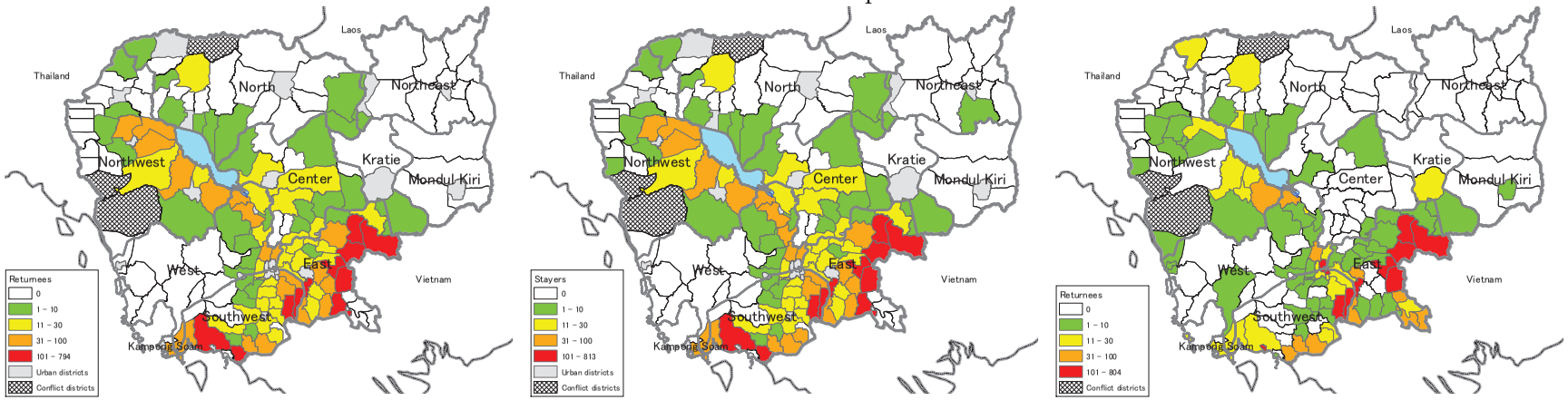
Figure A3: Geographic Distribution of Returnees and Stayers

B-1. Full Samples




Birth Districts (Returnees) | Birth Districts (Stayers) | Residence in 1998 (Returnees)

B-2. Matched Samples




Birth Districts (Returnees) | Birth Districts (Stayers) | Residence in 1998 (Returnees)

B. Vietnam

Figure A3: Geographic Distribution of Returnees and Stayers

*Notes*: The figure shows geographic distribution of returnees from Thailand (panel A) and from Vietnam (panel B) and stayers aged 20-60 by district of birth and residence in 1998 in the Full Samples and Matched Samples. It also shows the 1977 administrative zones of the Pol Pot regime and the 1998 districts.

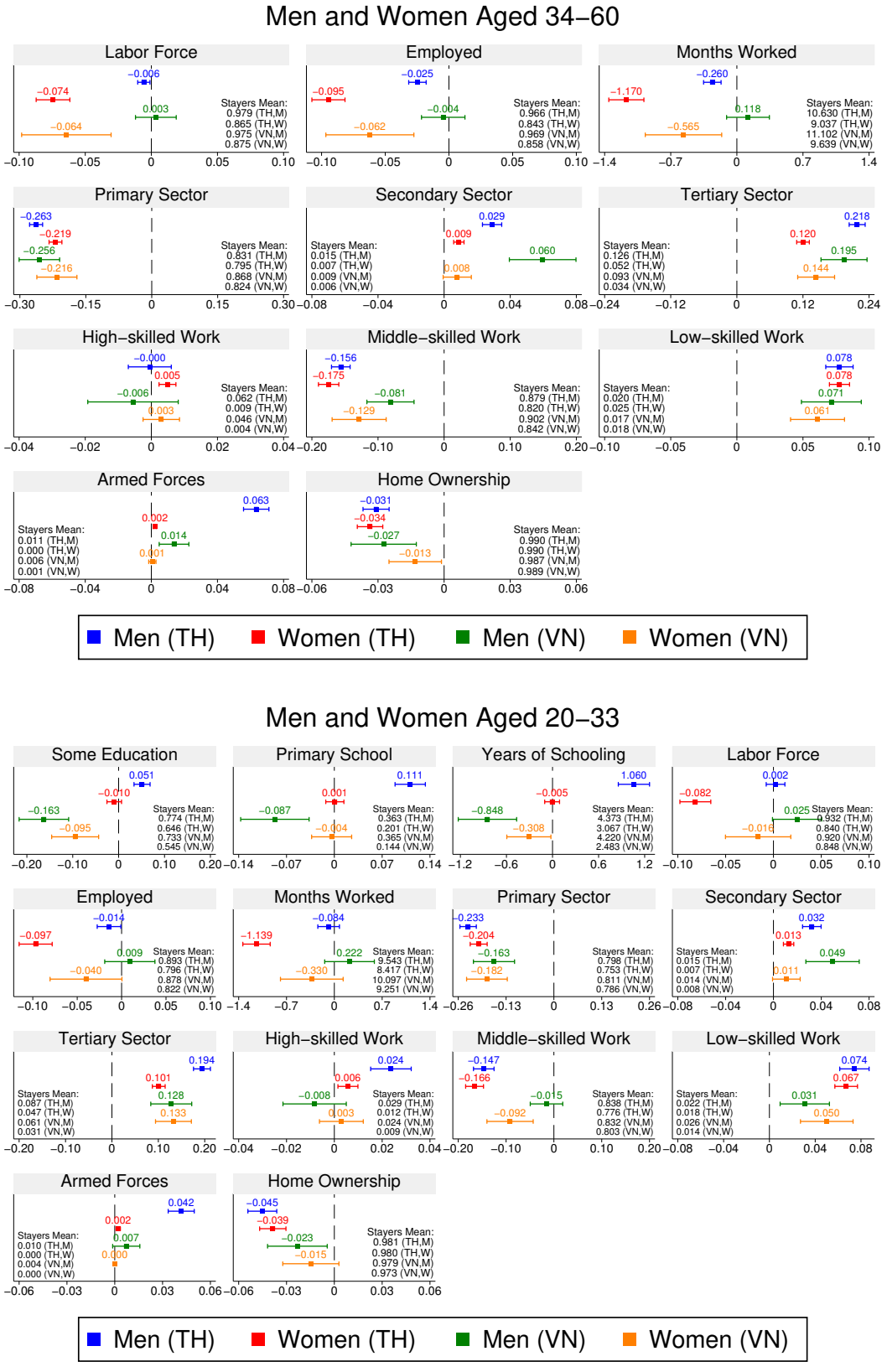
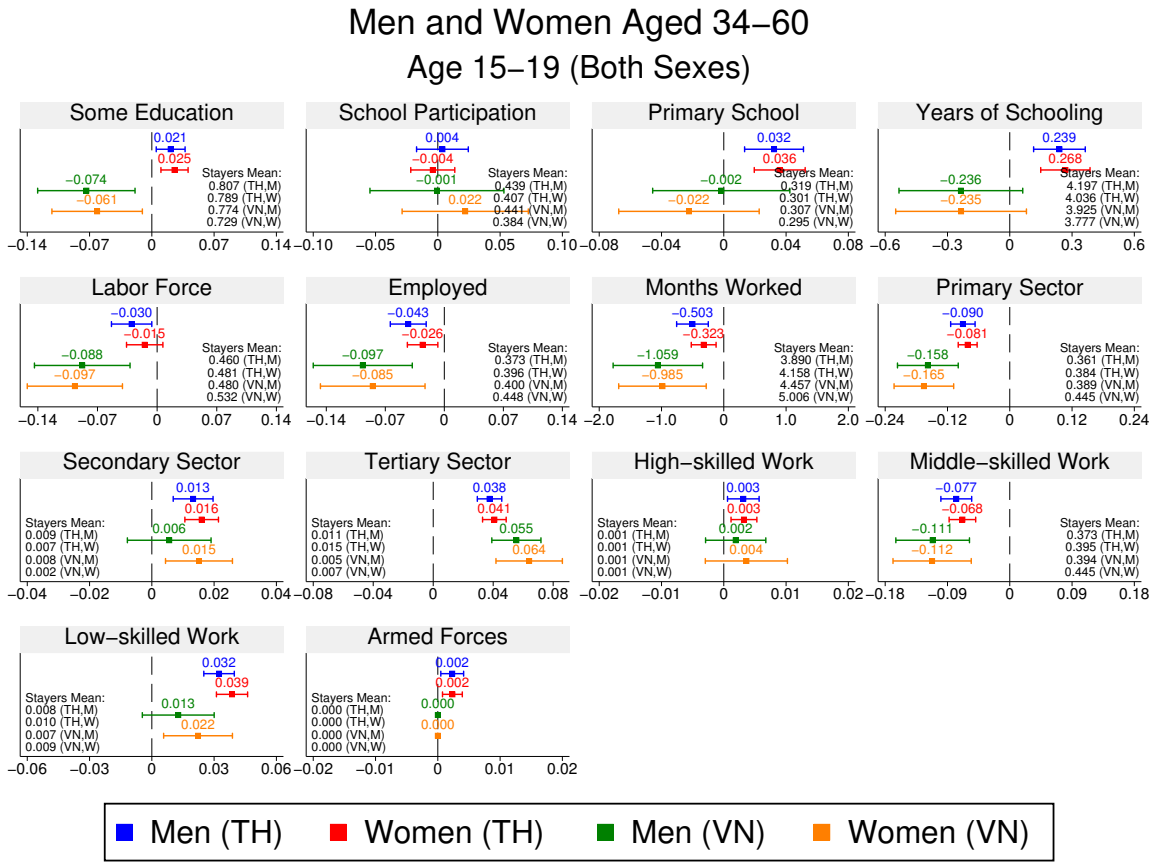



Figure A4: Main Results – Age 20-60

*Notes*: The figure plots the point estimates of the impacts of forced displacement on educational and labor market outcomes and home ownership and their 95% confidence intervals, along with the stayers mean, for male and female returnees aged 20–60 from Thailand (TH) and Vietnam (VN). The estimates are based on the Matched Samples and are from the bias-corrected version of the nearest-neighbor matching method (Abadie and Imbens 2011). The 95% confidence intervals are adjusted with the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995).



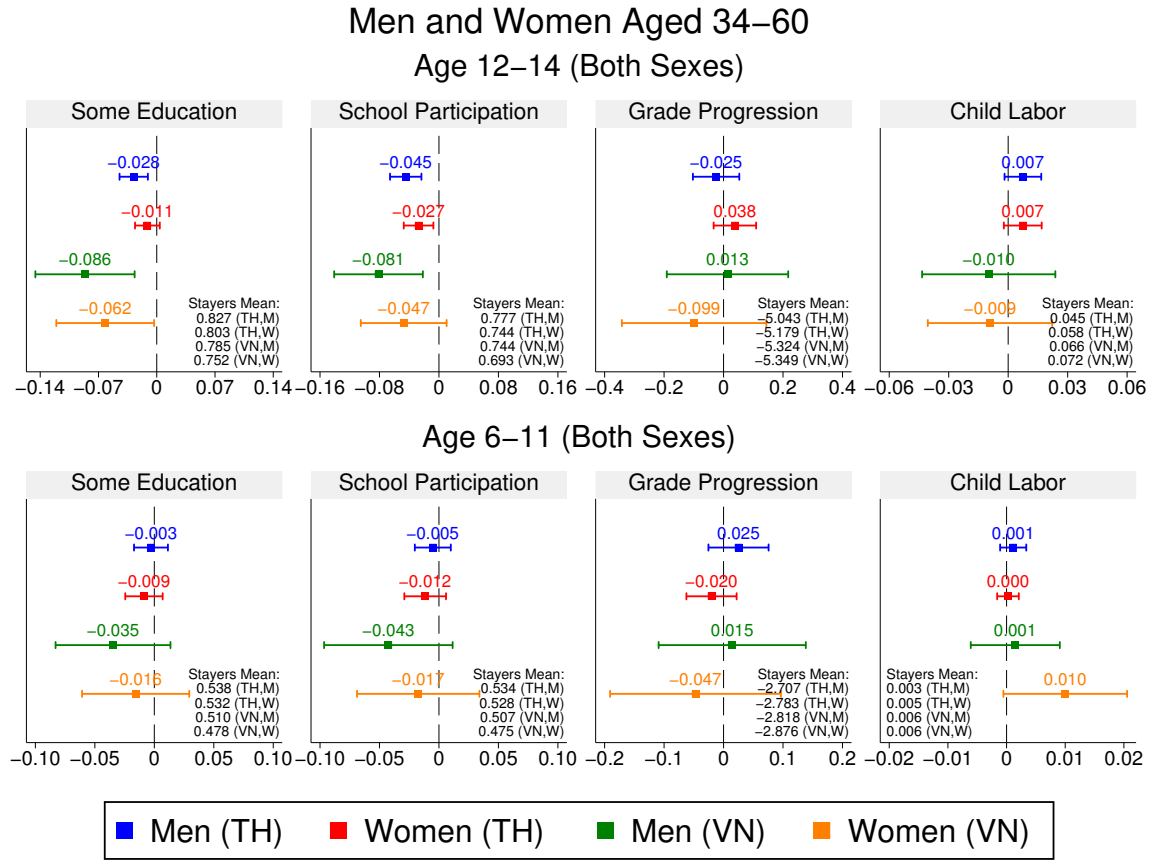
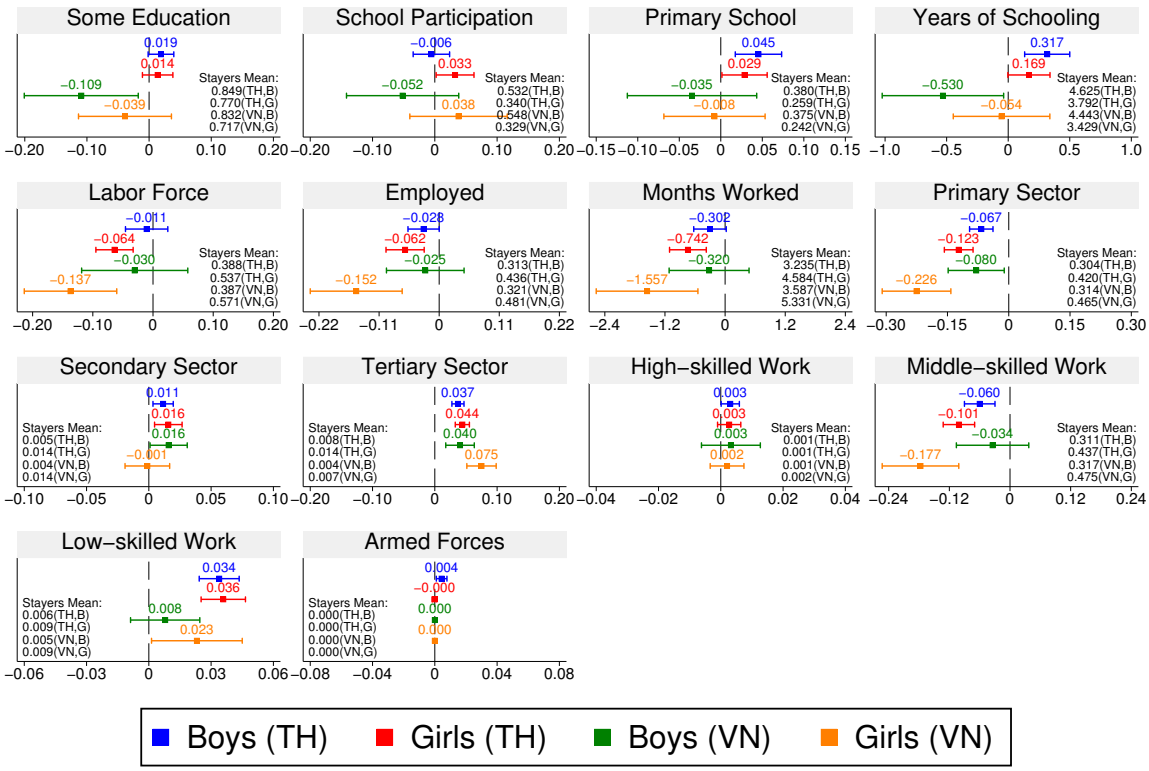




Figure A5: Main Results – Age 6-19



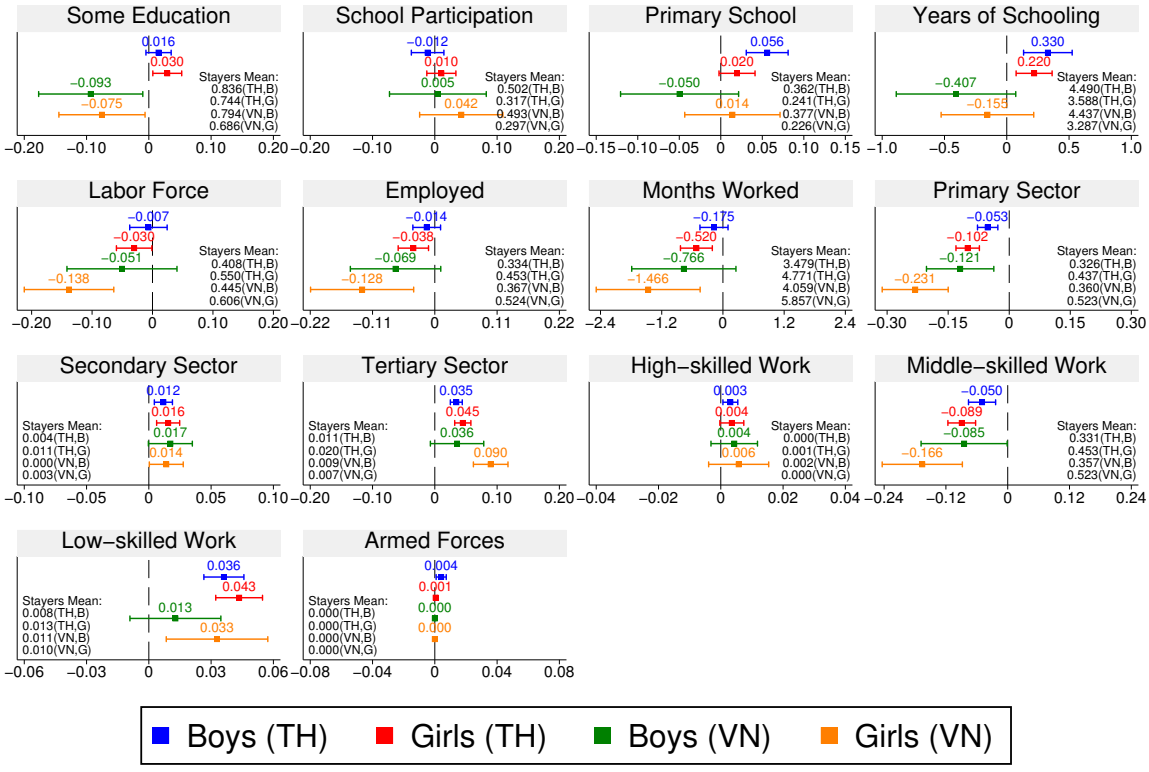
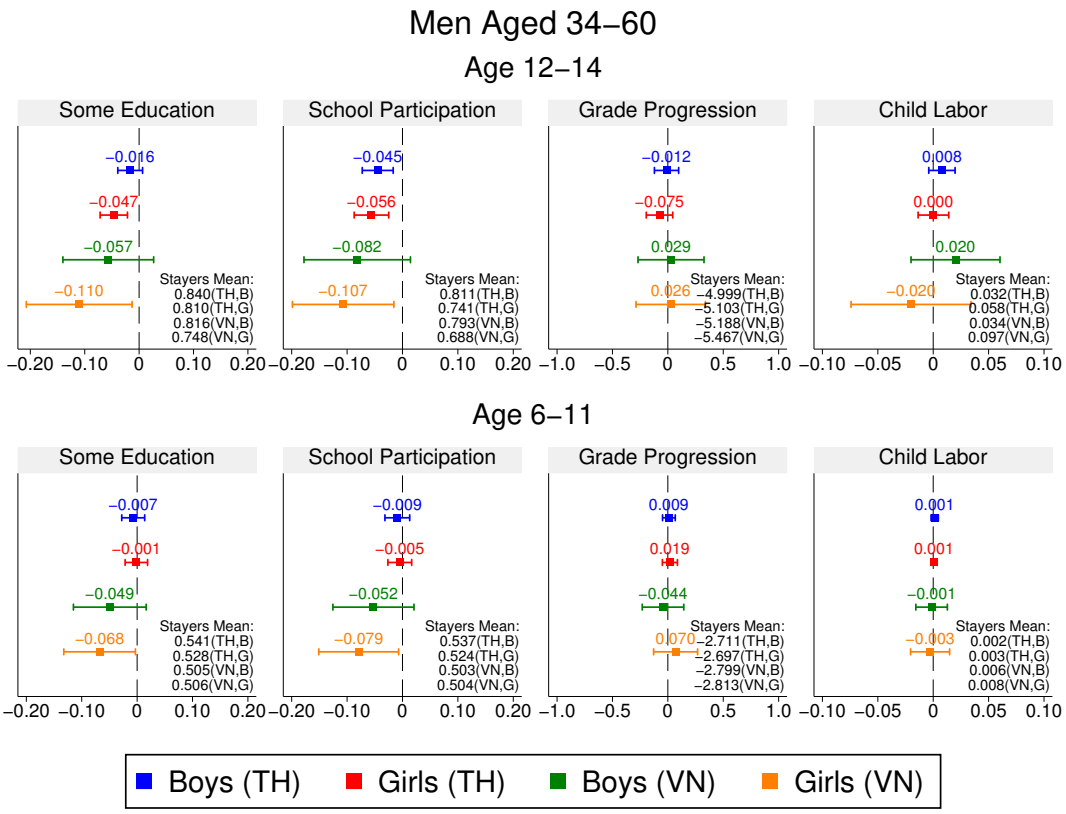




Figure A5: Main Results – Age 6-19



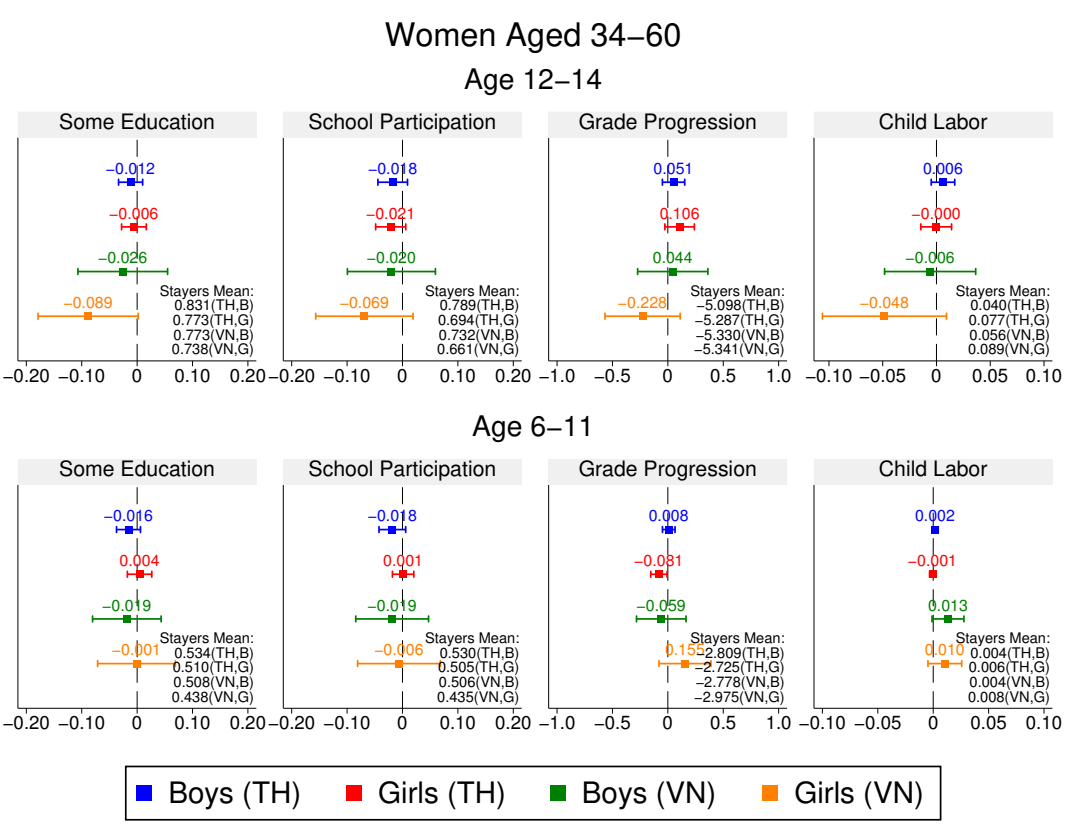
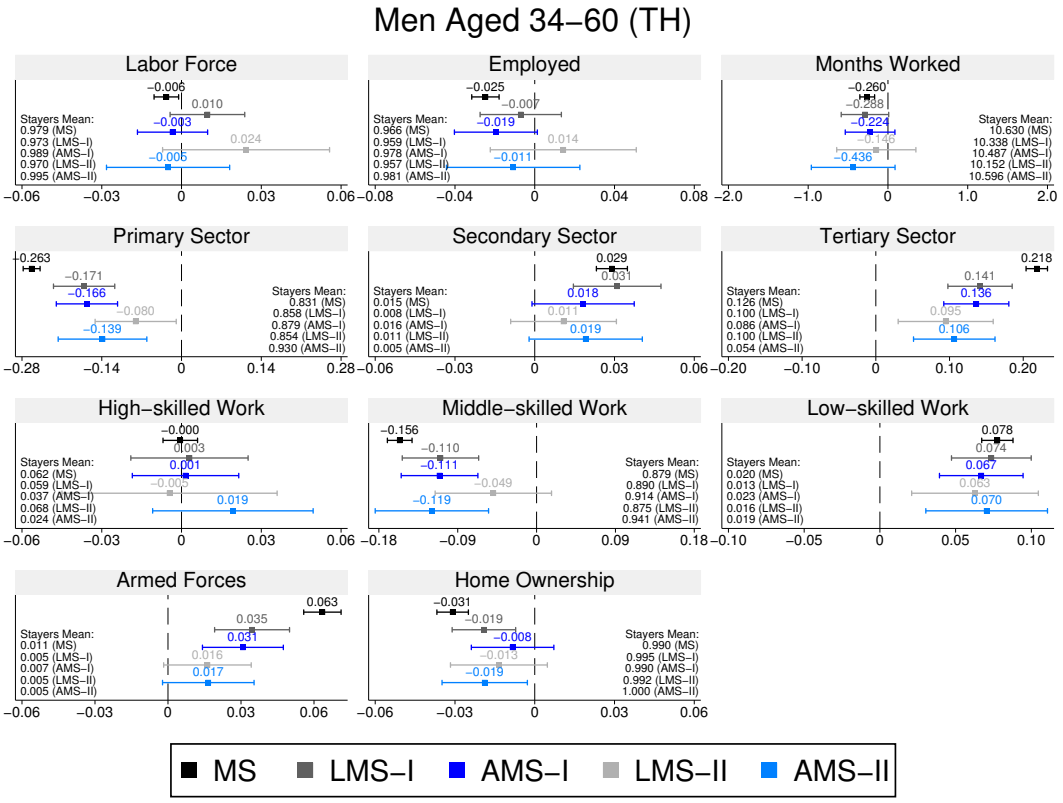




Figure A5: Main Results – Age 6-19

*Note*: The figure plots the point estimates of the impacts of forced displacement on educational and labor market outcomes and their 95% confidence intervals, along with stayers mean, for the children aged 6-19 (for both sexes, for boys, and for girls) of the male and female returnees aged 34–60 from Thailand (TH) and Vietnam (VN). For the estimation method, see the notes to Figure A4 and the main text.



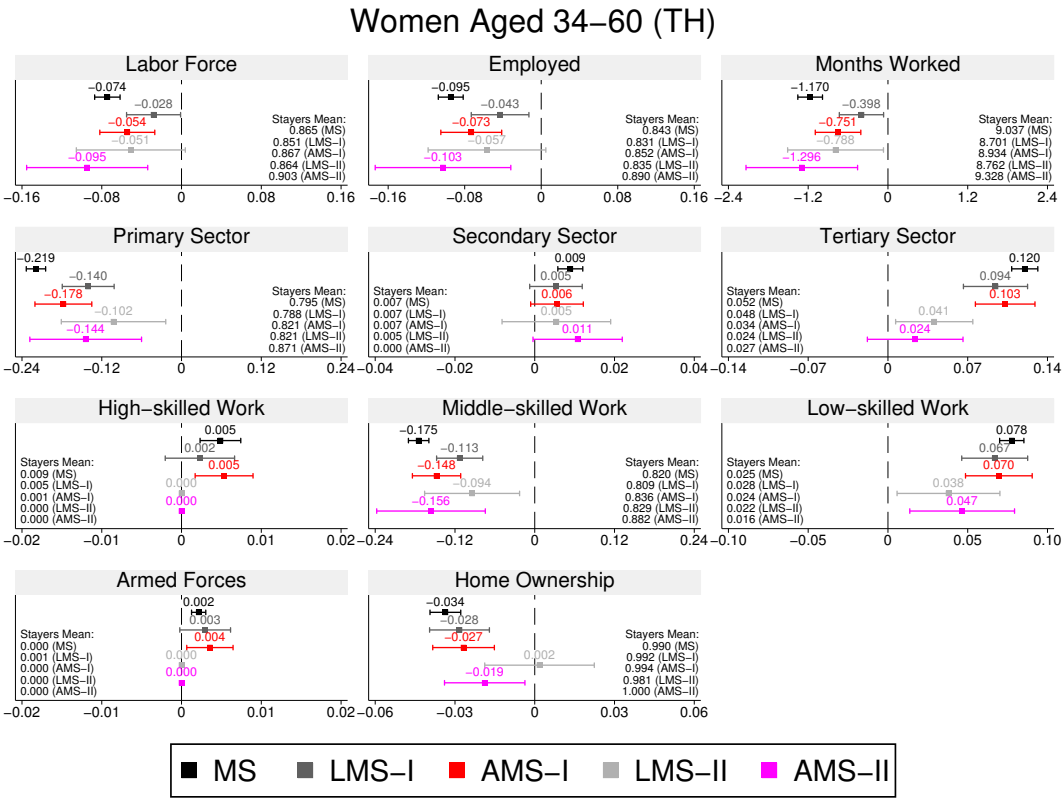




Figure A6: Robustness Check – Alternative Specifications (Age 34-60, TH)

*Notes*: MS – Matched Samples; LMS-I – Limited Matched Samples I (constructed based on the treated units included in both MS and AMS-I); AMS-I – Alternative Matched Samples I (constructed based on the alternative specifications for propensity score, with district of birth fixed effects for children born in 1975–1983; LMS-II – Alternative Matched Samples II (constructed based on the treated units included in both MS and AMS-II); and AMS-II – Alternative Matched Samples II (constructed based on the alternative specifications for propensity score, with the characteristics of both husbands and wives and district of birth fixed effects for children born in 1975–1983).

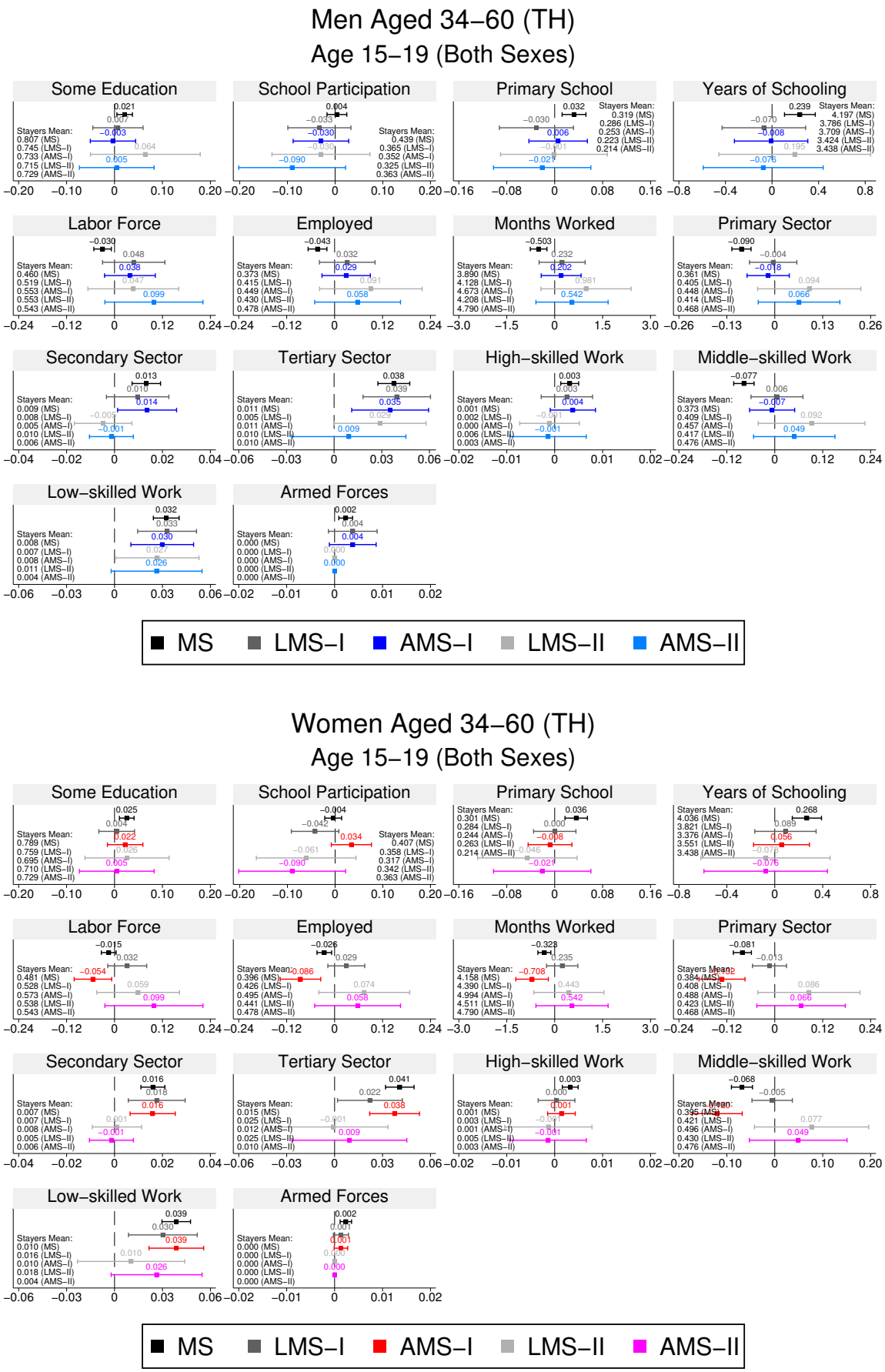
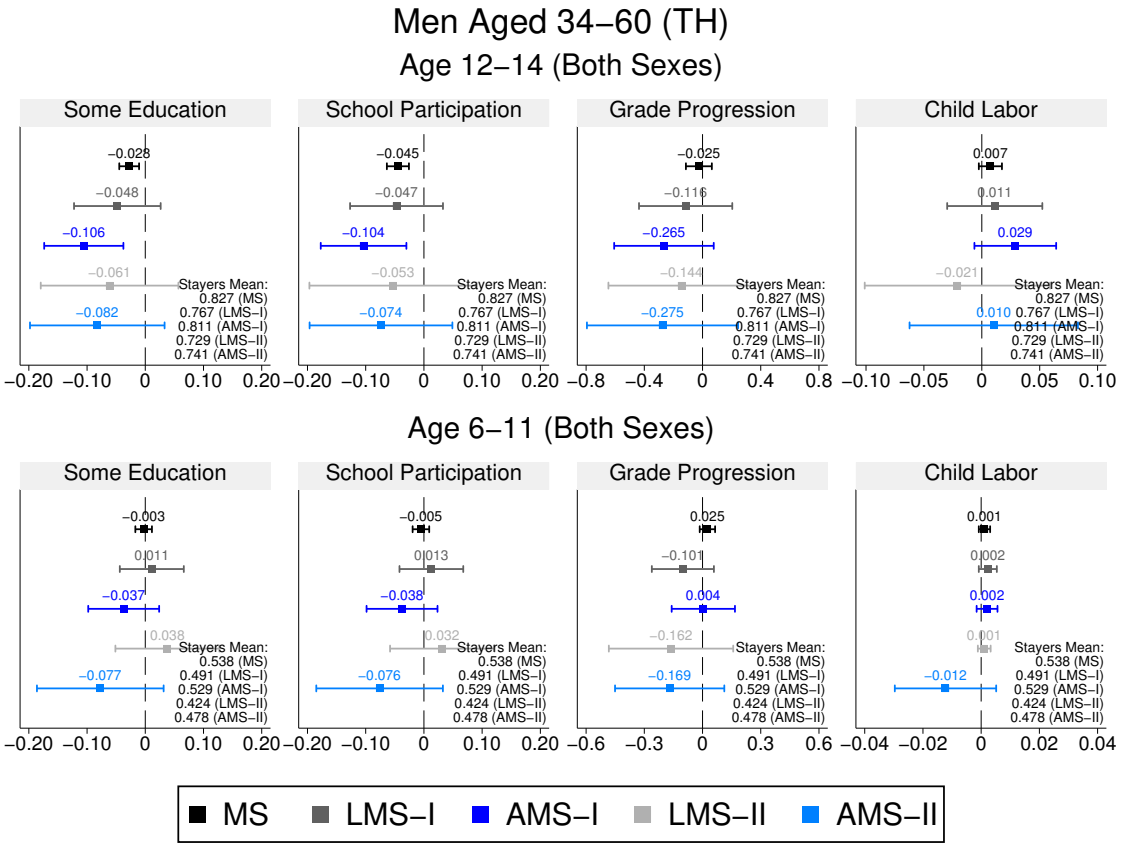



Figure A7: Robustness Check – Alternative Specifications (Age 6-19, TH)



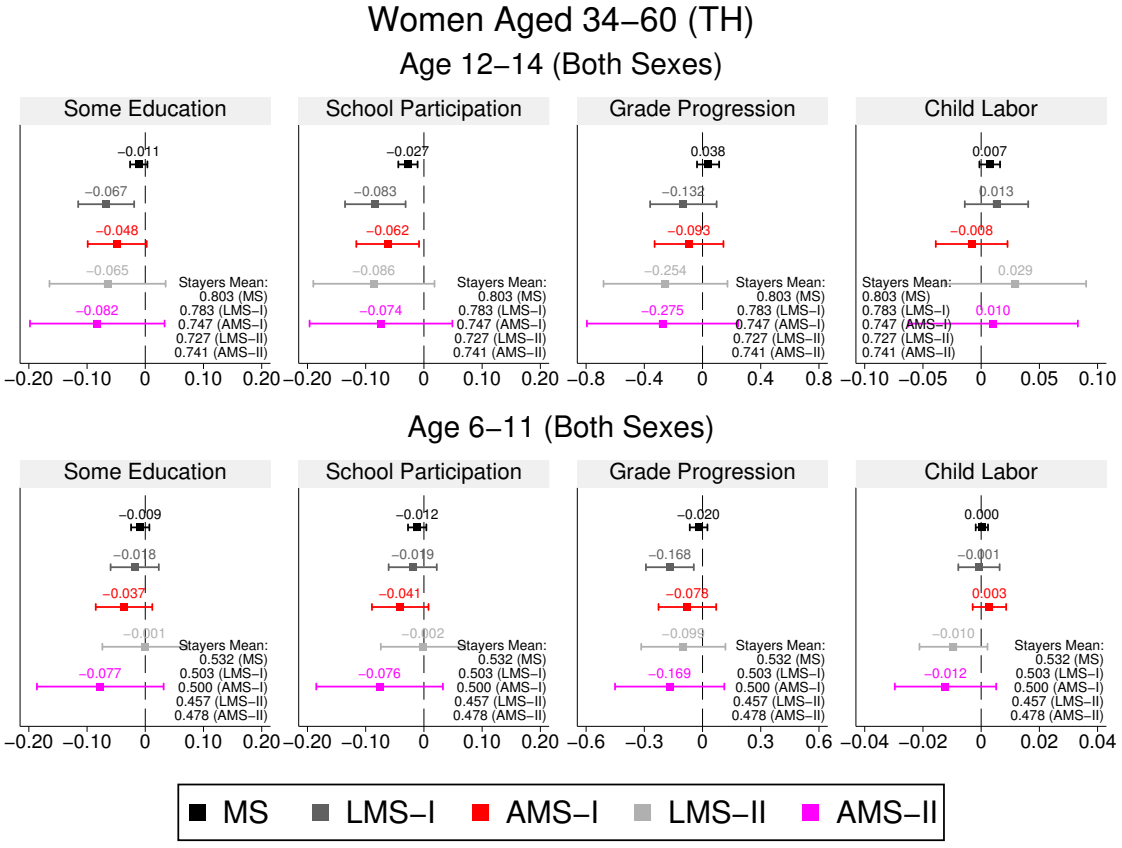




Figure A7: Robustness Check – Alternative Specifications (Age 6-19, TH)

*Note*: See the notes to Figure A6.

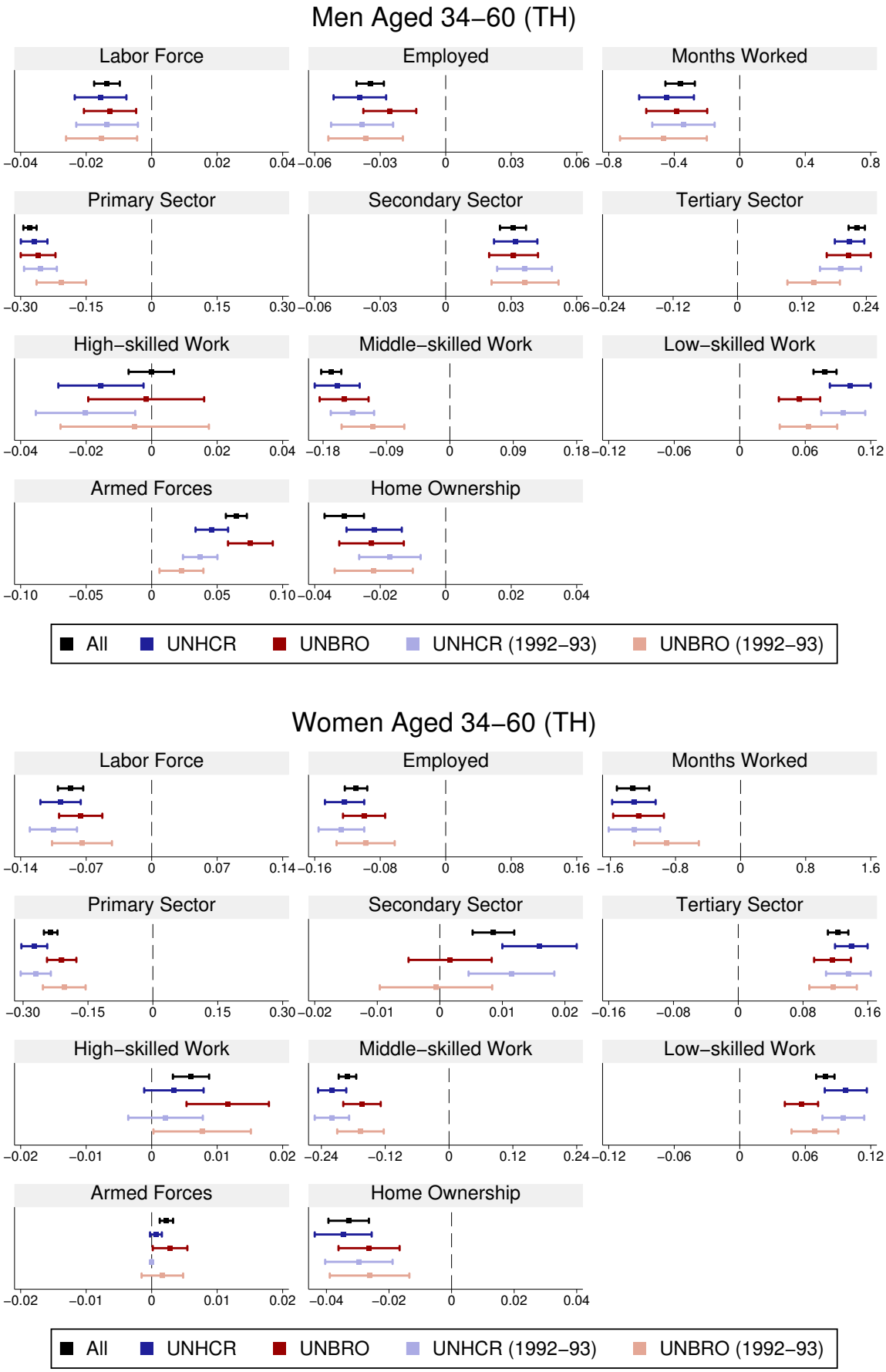
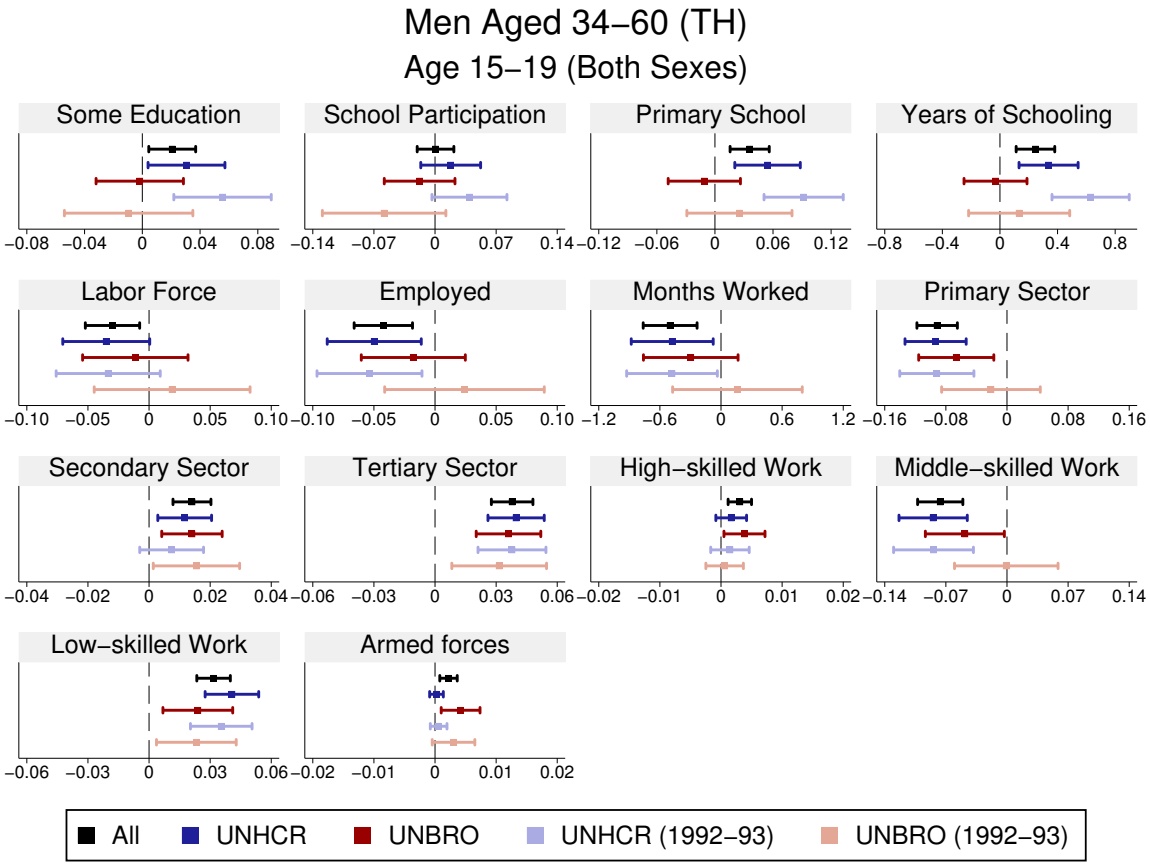



Figure A8: Robustness Check and Heterogeneity – Refugee Camps (Age 34-60, TH)



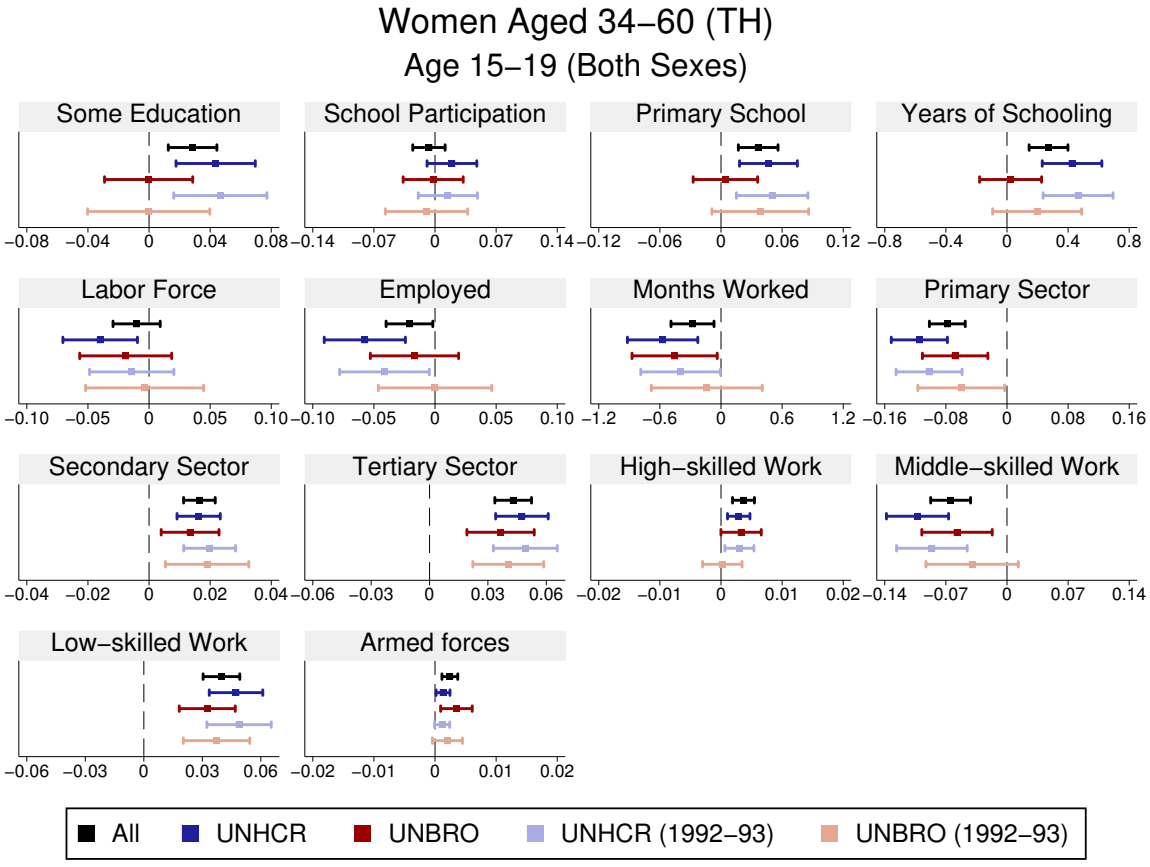
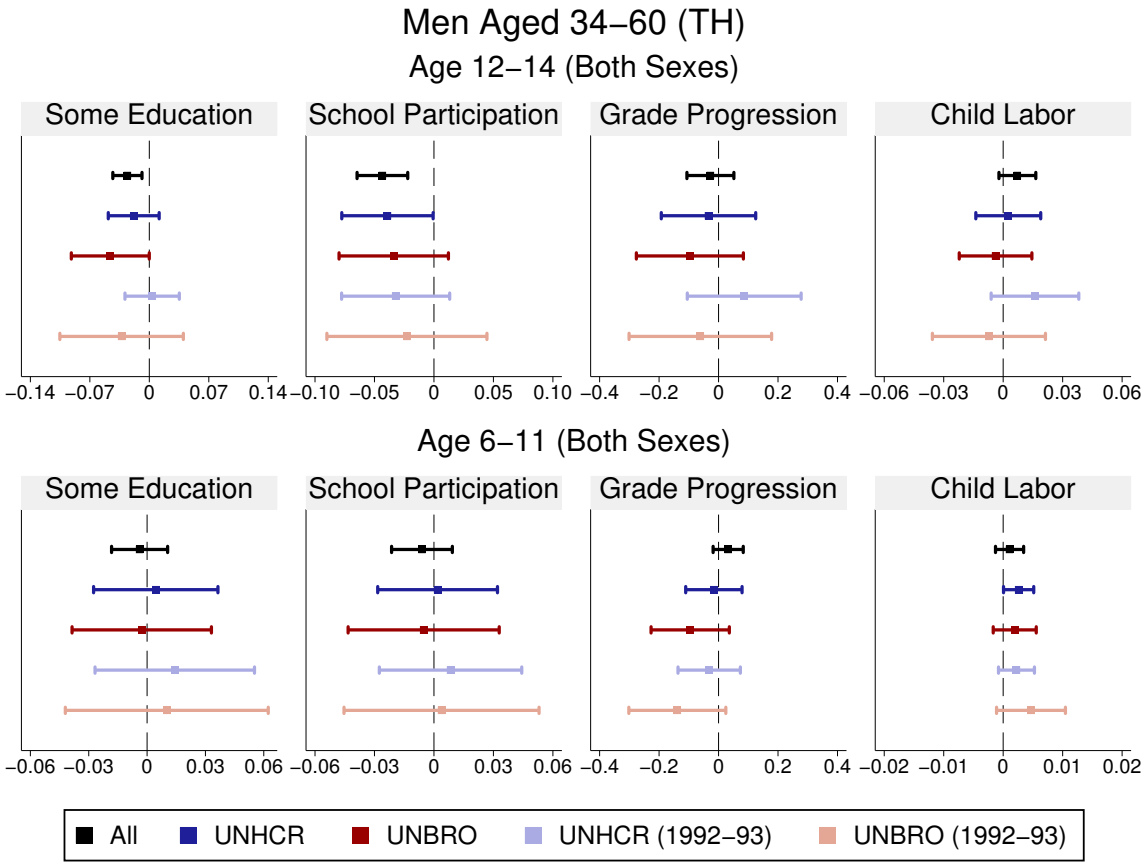




Figure A9: Robustness Check and Heterogeneity – Refugee Camps (Age 6-19, TH)



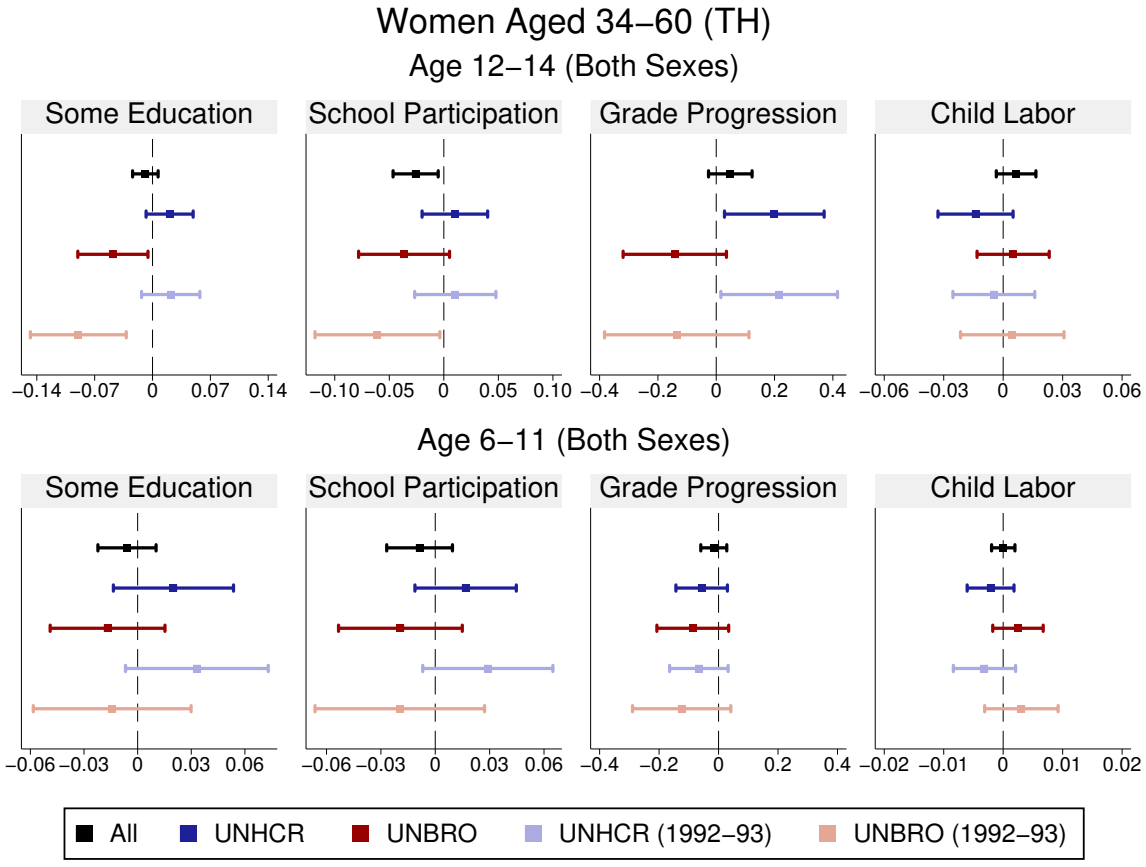




Figure A9: Robustness Check and Heterogeneity – Refugee Camps (Age 6-19, TH)

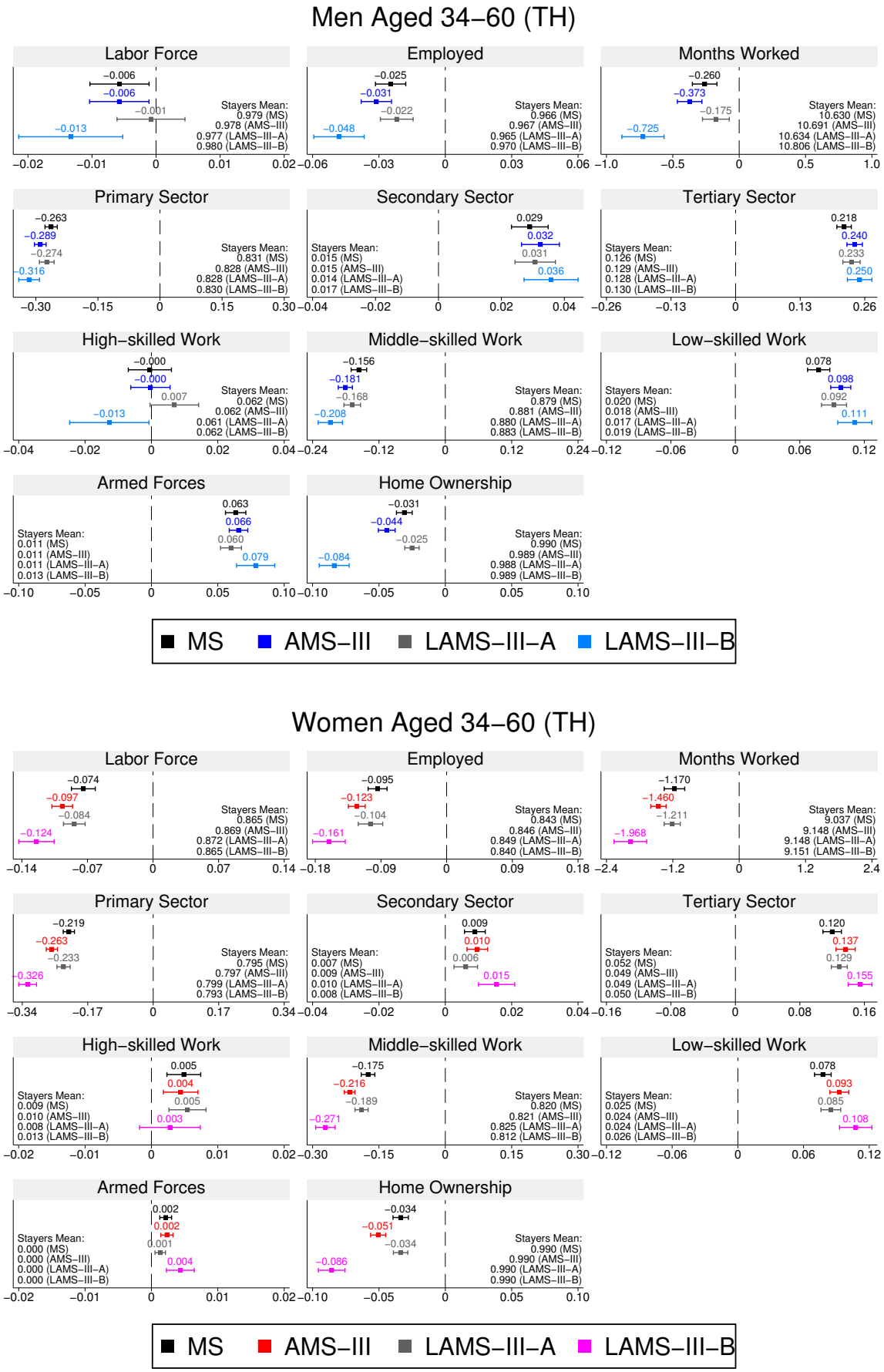
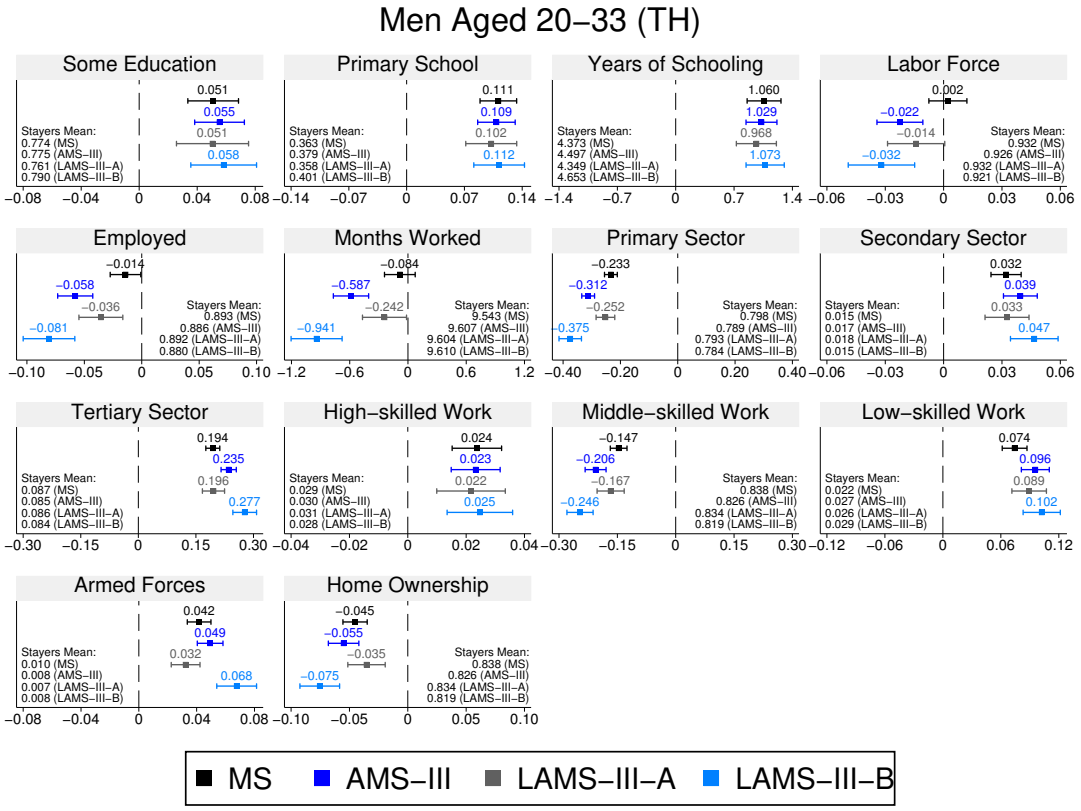



Figure A10: Robustness Check – Alternative Samples (Age 20-60, TH)



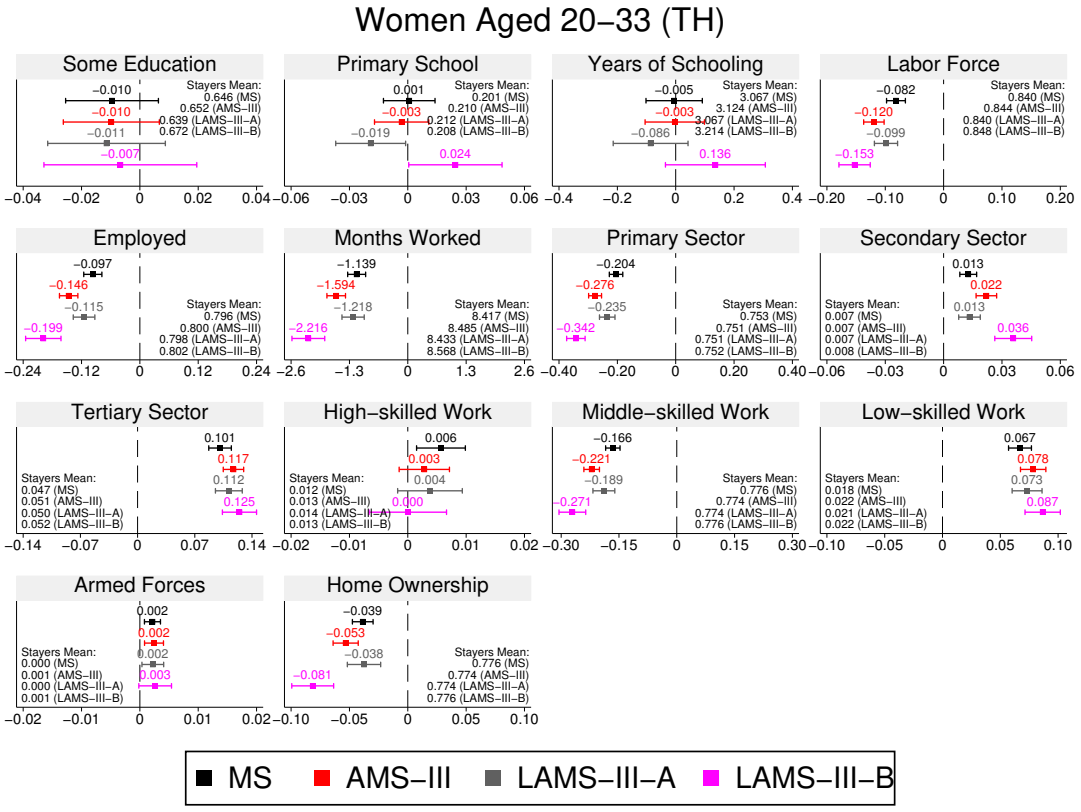
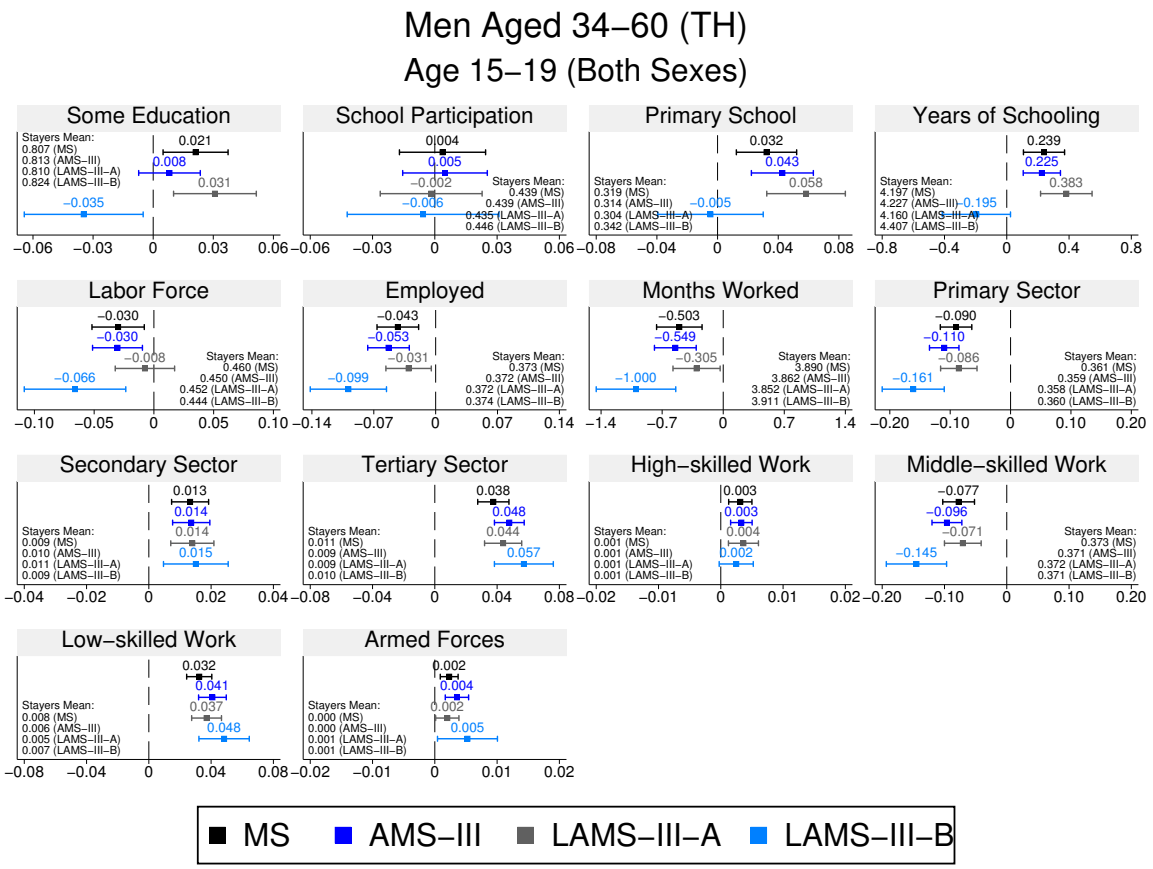




Figure A10: Robustness Check – Alternative Samples (Age 20-60, TH)

*Notes*: MS – Matched Samples; AMS-III – Alternative Matched Samples III (constructed based on an alternative definition of former refugees); LAMS-III-A – Limited Alternative Matched Samples III-A (constructed based on the treated units included in both MS and AMS-III); and LAMS-III-B – Limited Alternative Matched Samples III-B (constructed based on the treated units included in AMS-III, but not in MS).



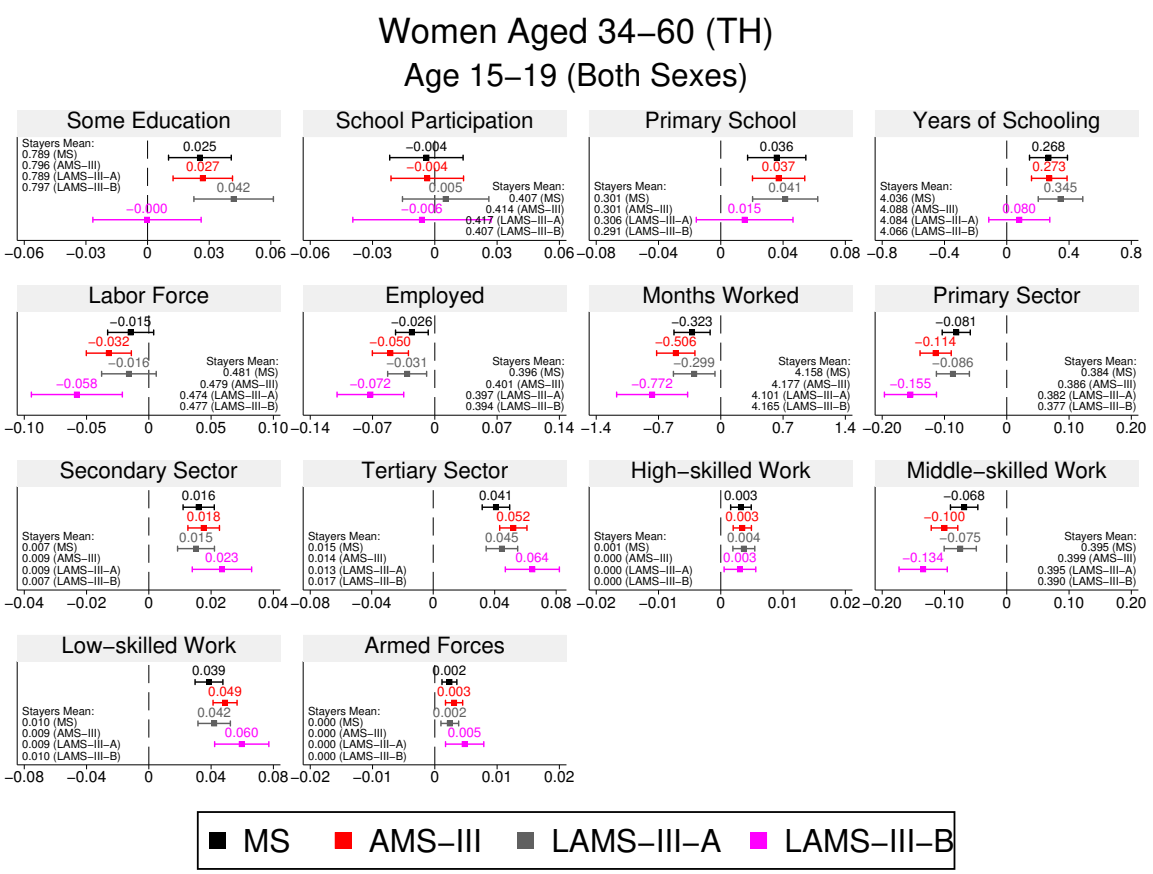




Figure A11: Robustness Check – Alternative Samples (Age 6-19, TH)

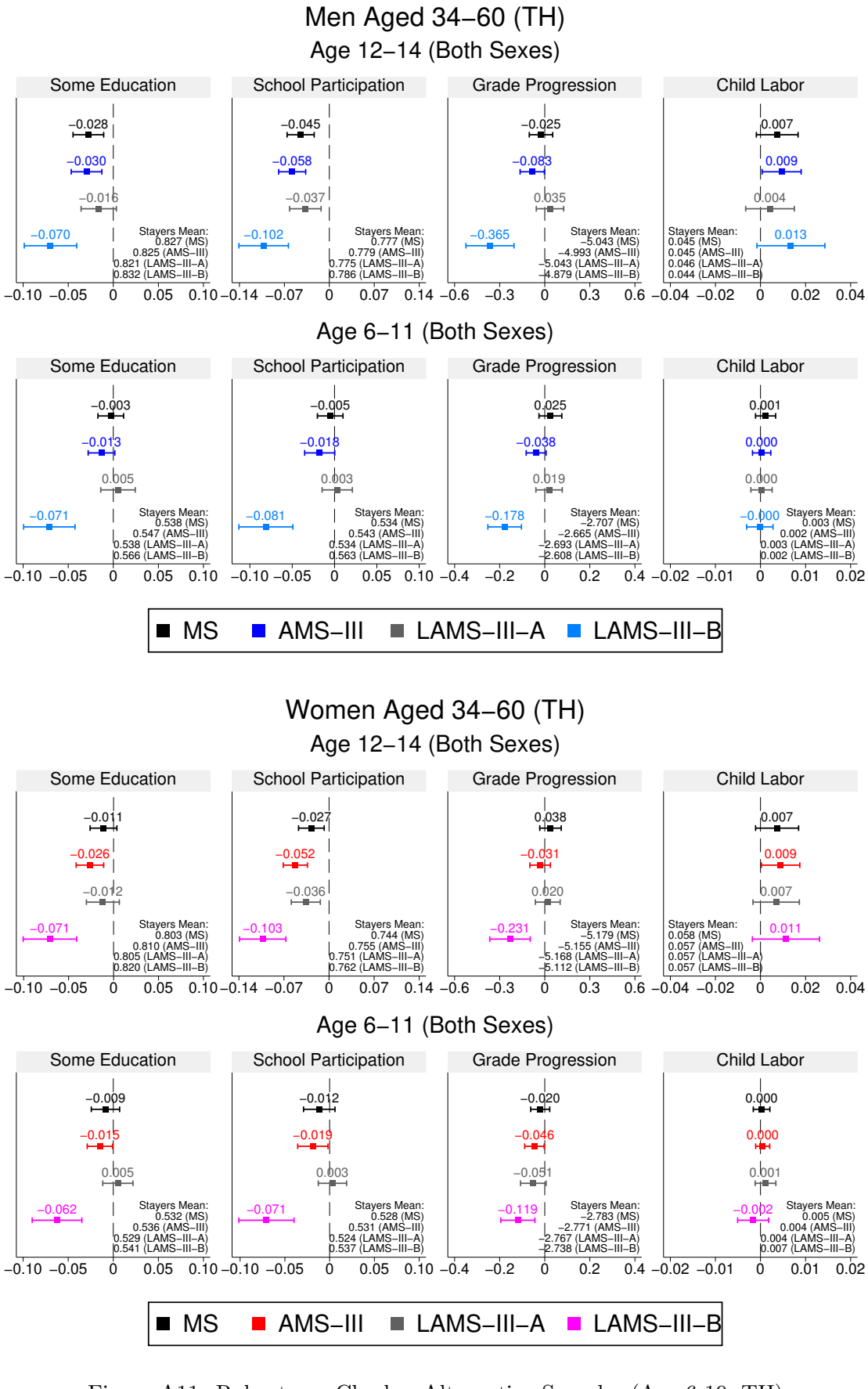
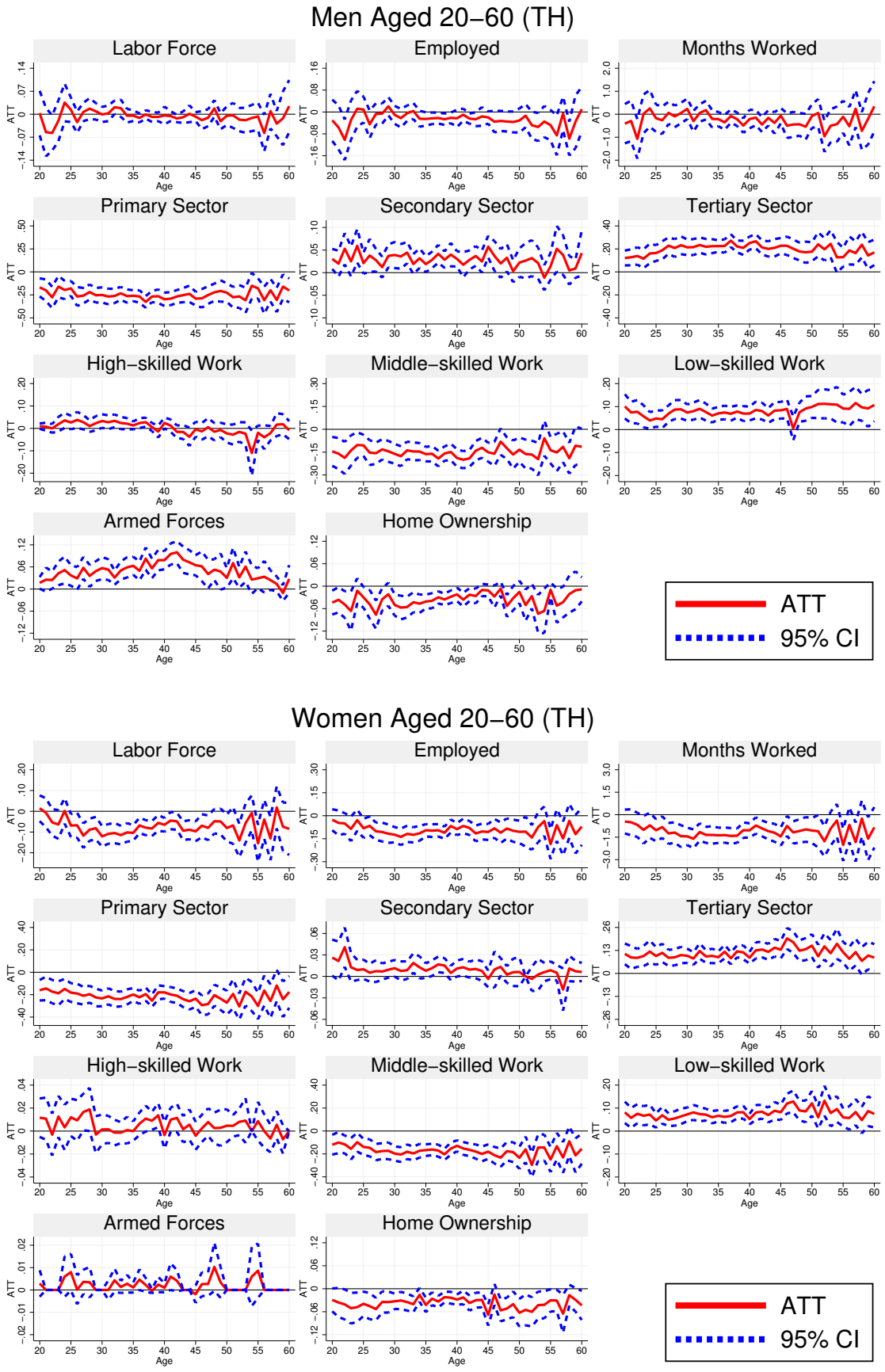


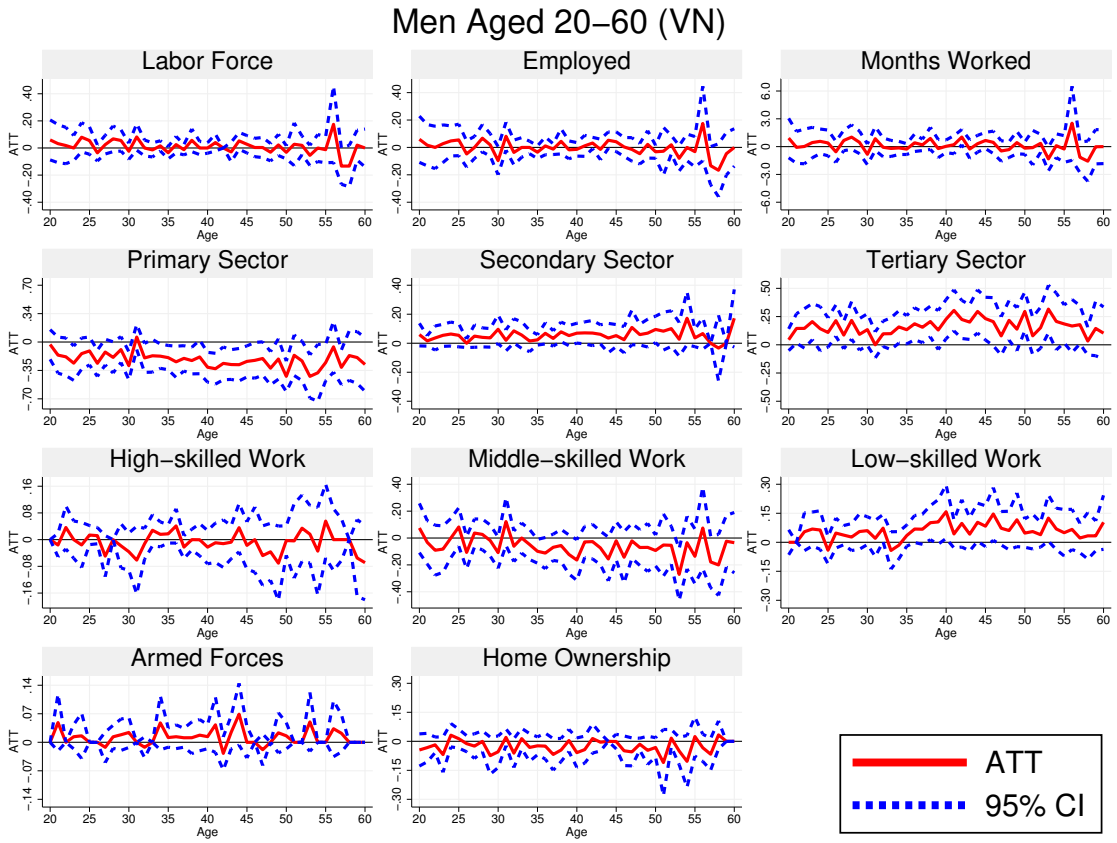

Figure A11: Robustness Check – Alternative Samples (Age 6-19, TH)

*Note*: See the notes to Figure A10.




A-1. Thailand

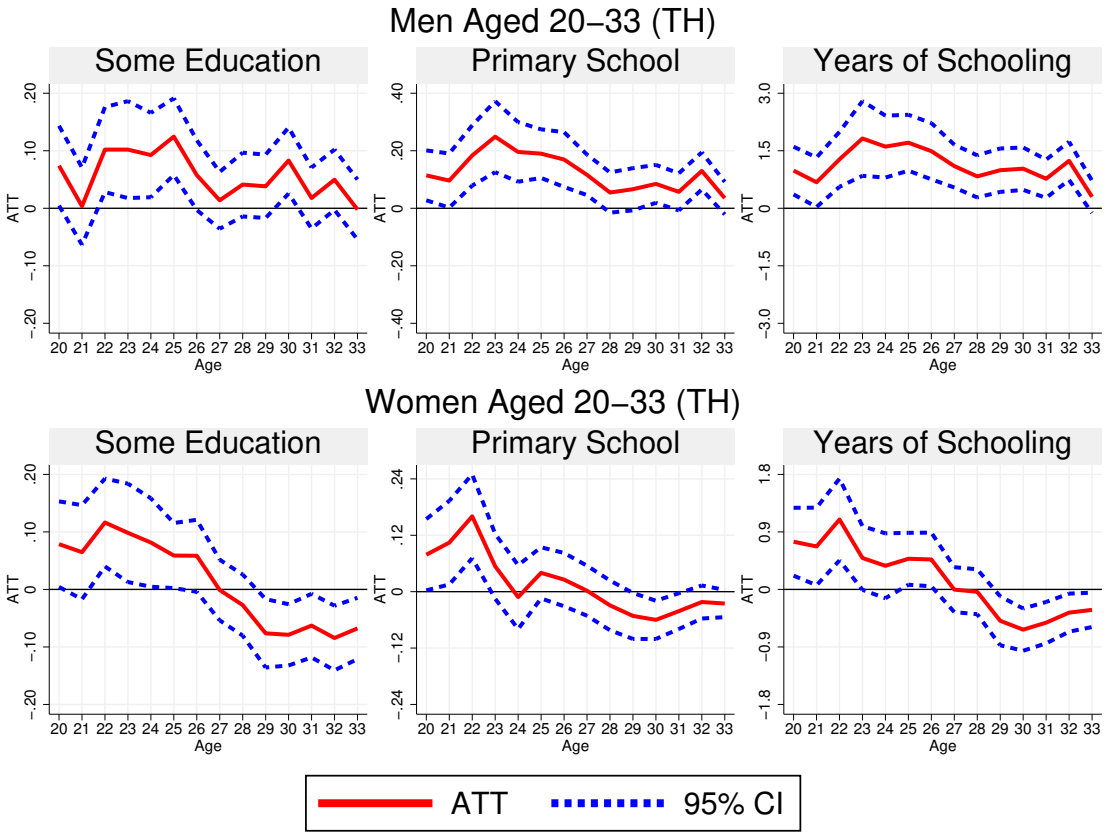
Figure A12: Heterogeneity – Age (Age 20-60)



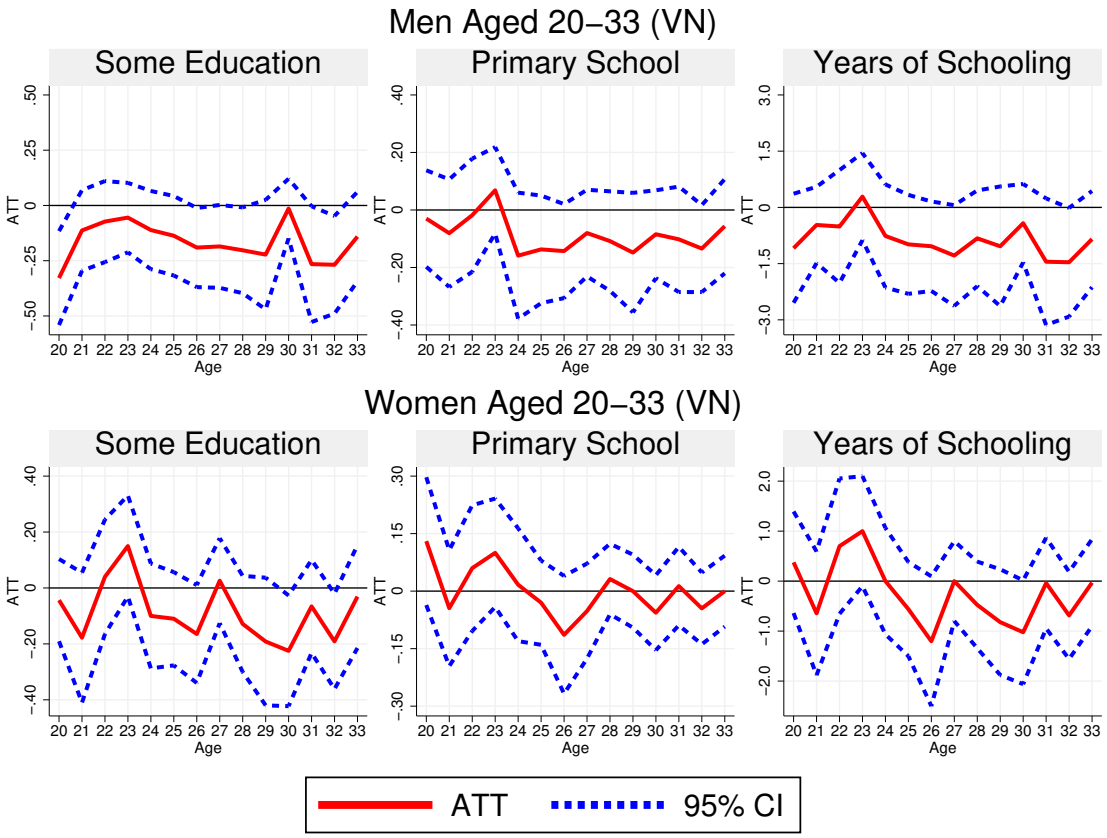


B-1. Vietnam

Figure A12: Heterogeneity – Age (Age 20-60)

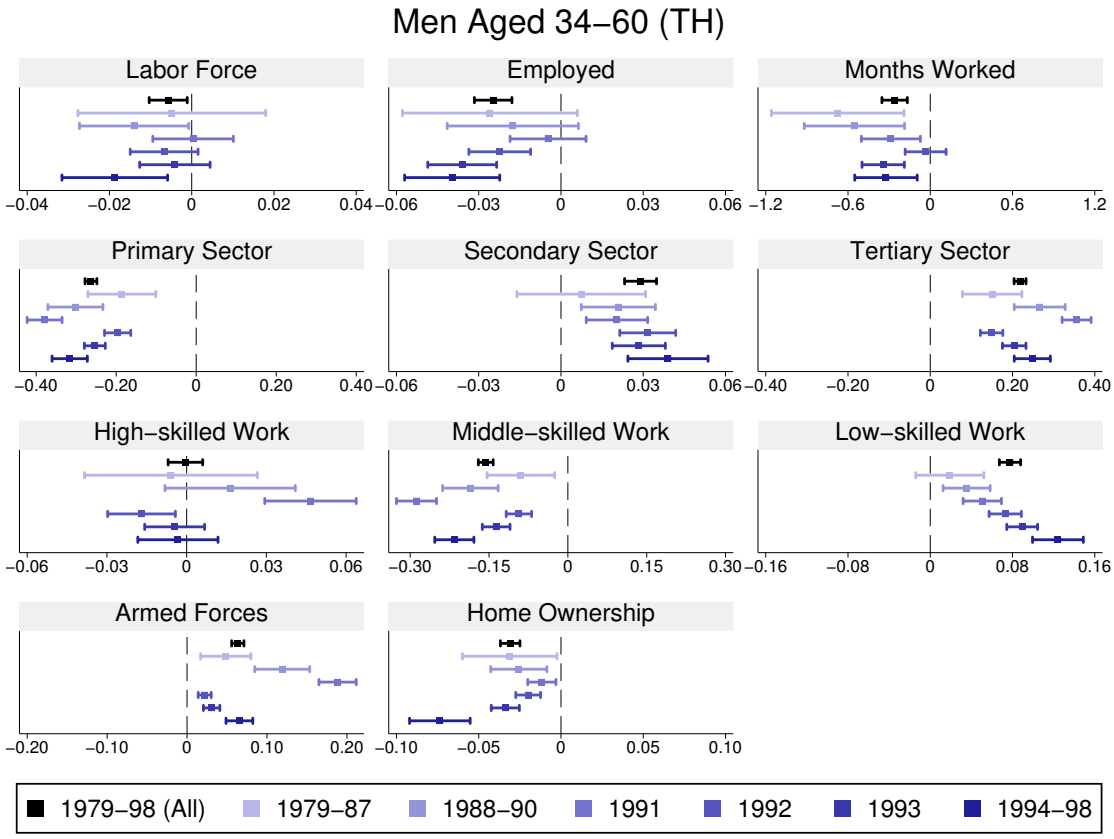



A-2. Thailand




B-2. Vietnam

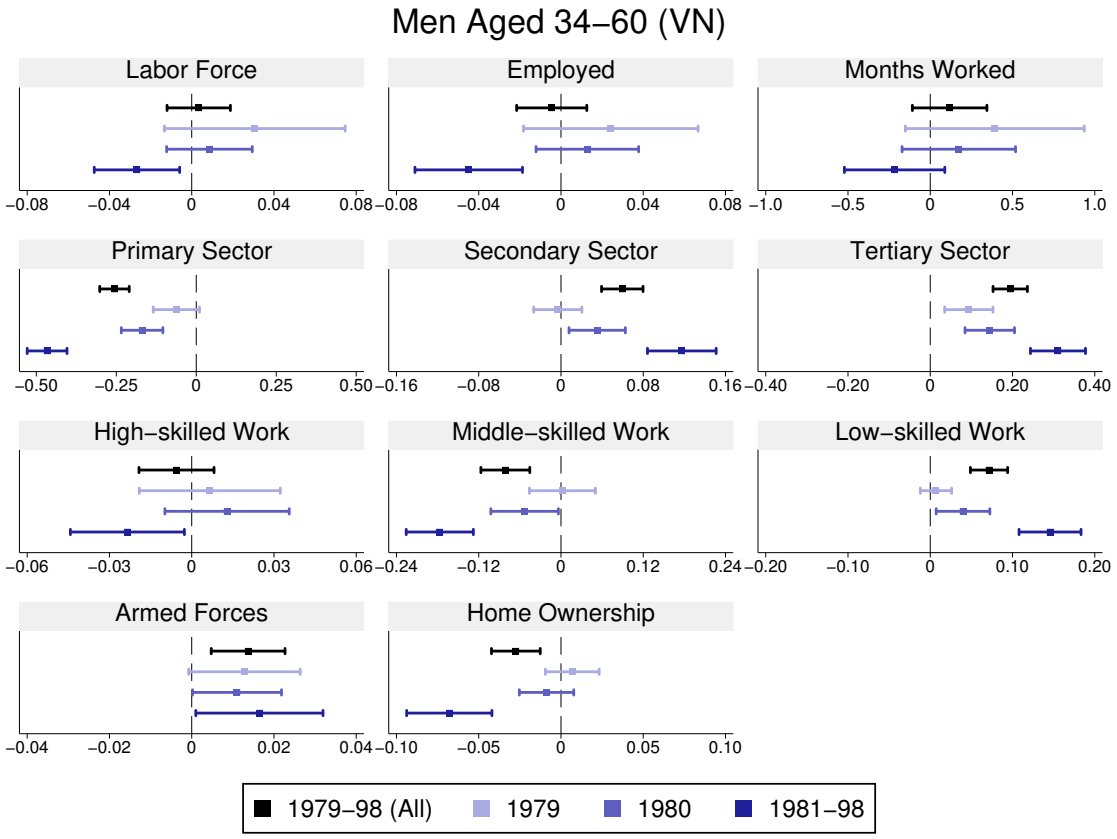
Figure A12: Heterogeneity – Age (Age 20-60)

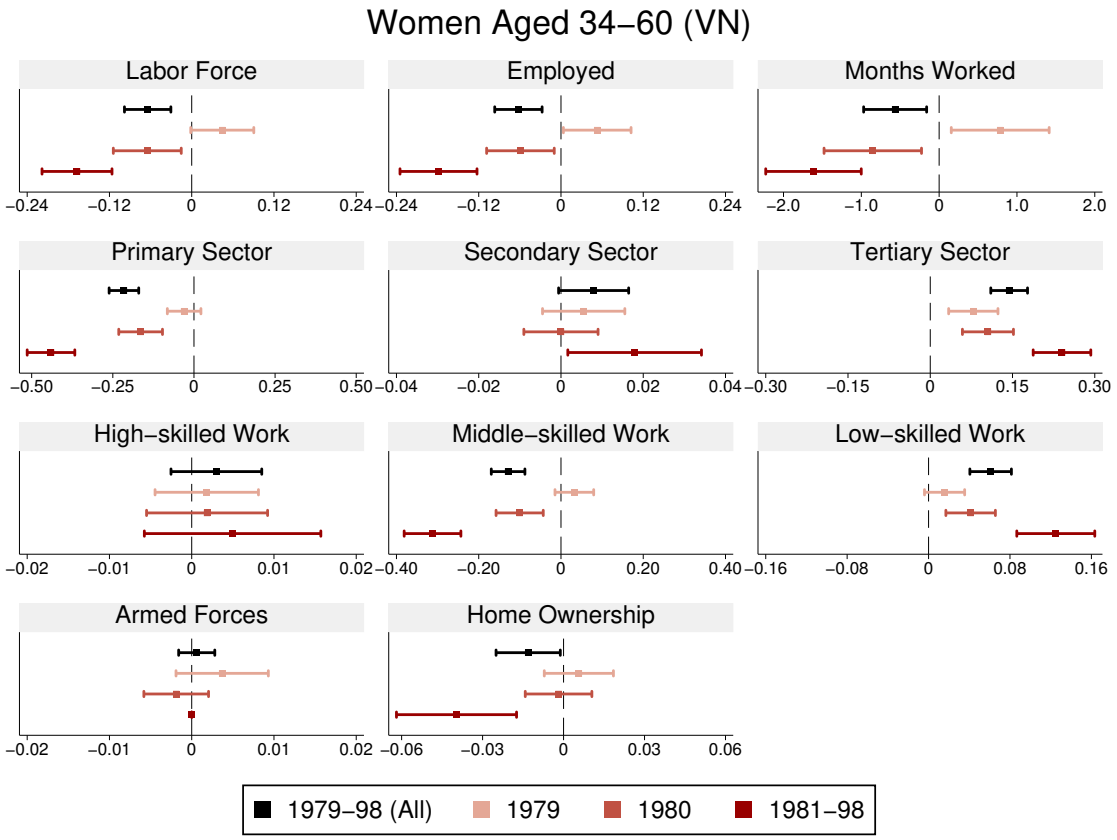




A-1. Thailand

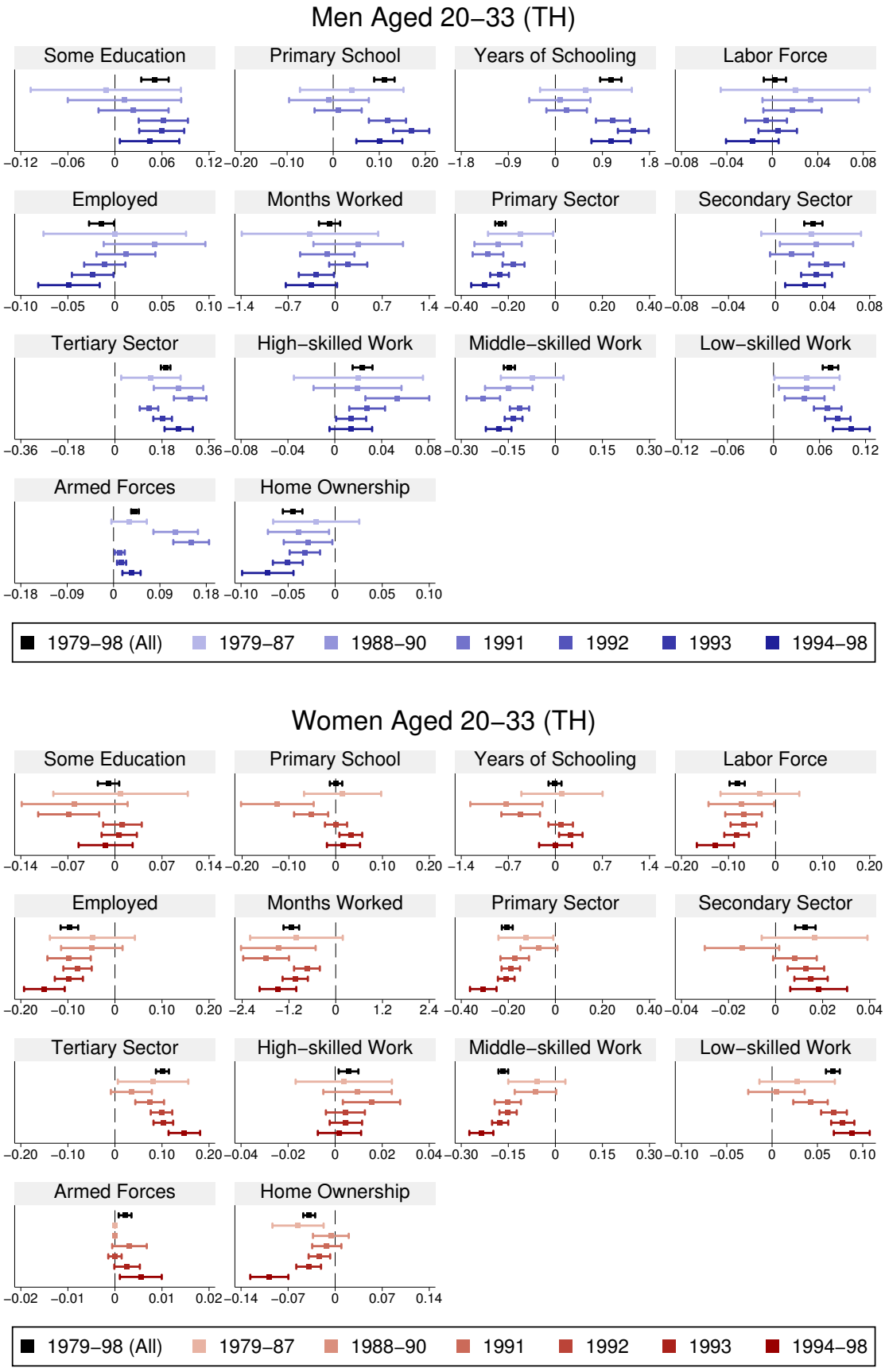
Figure A13: Heterogeneity – Timing of Repatriation (Age 20-60)



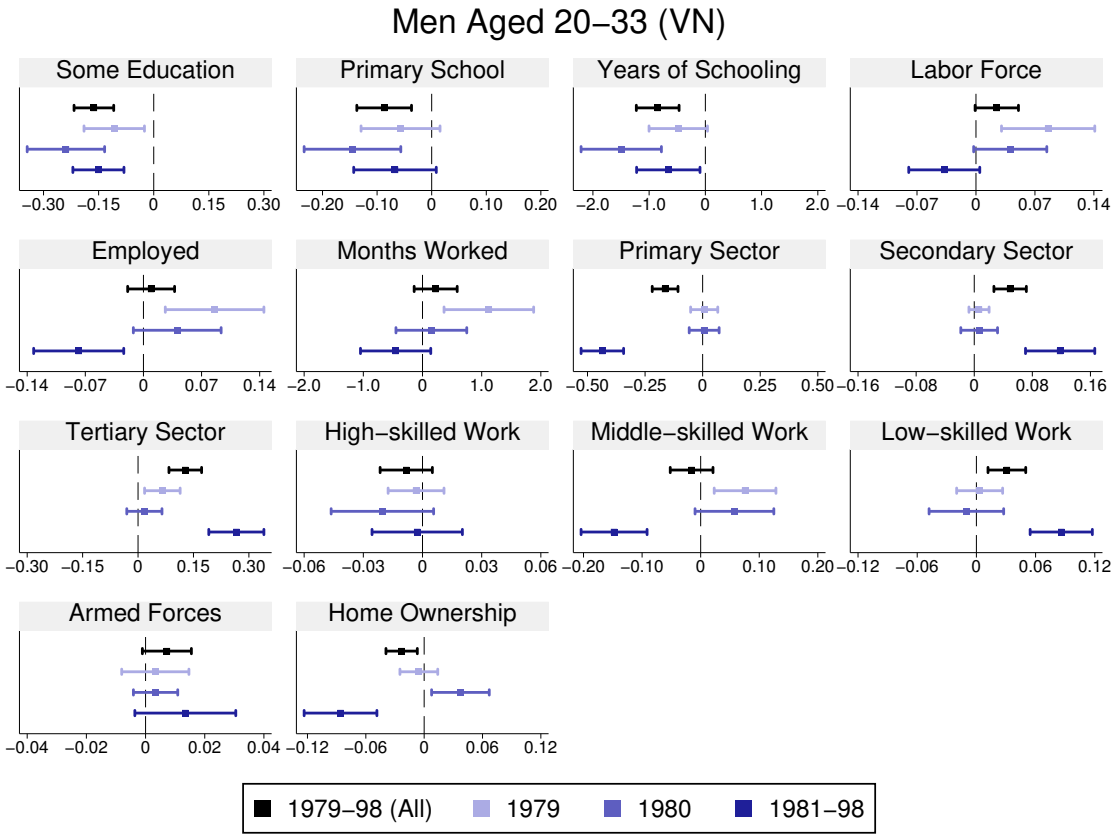

B-1. Vietnam

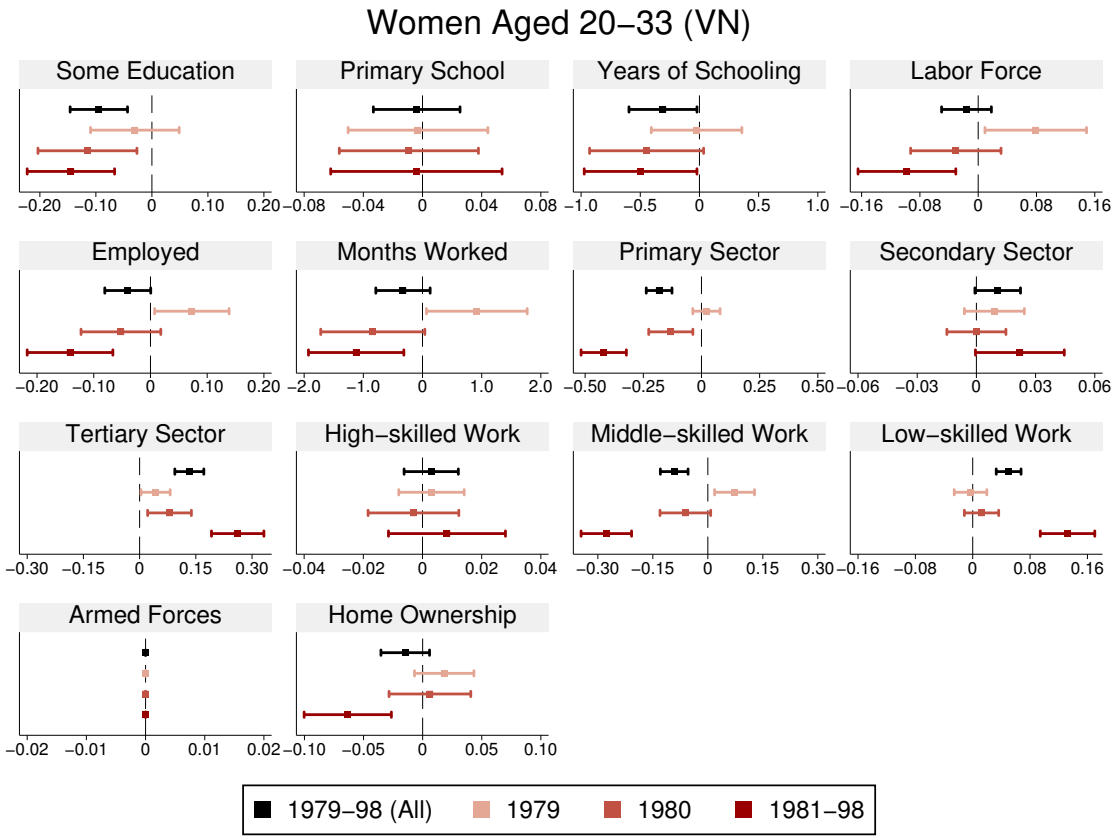
Figure A13: Heterogeneity – Timing of Repatriation (Age 20-60)




A-2. Thailand

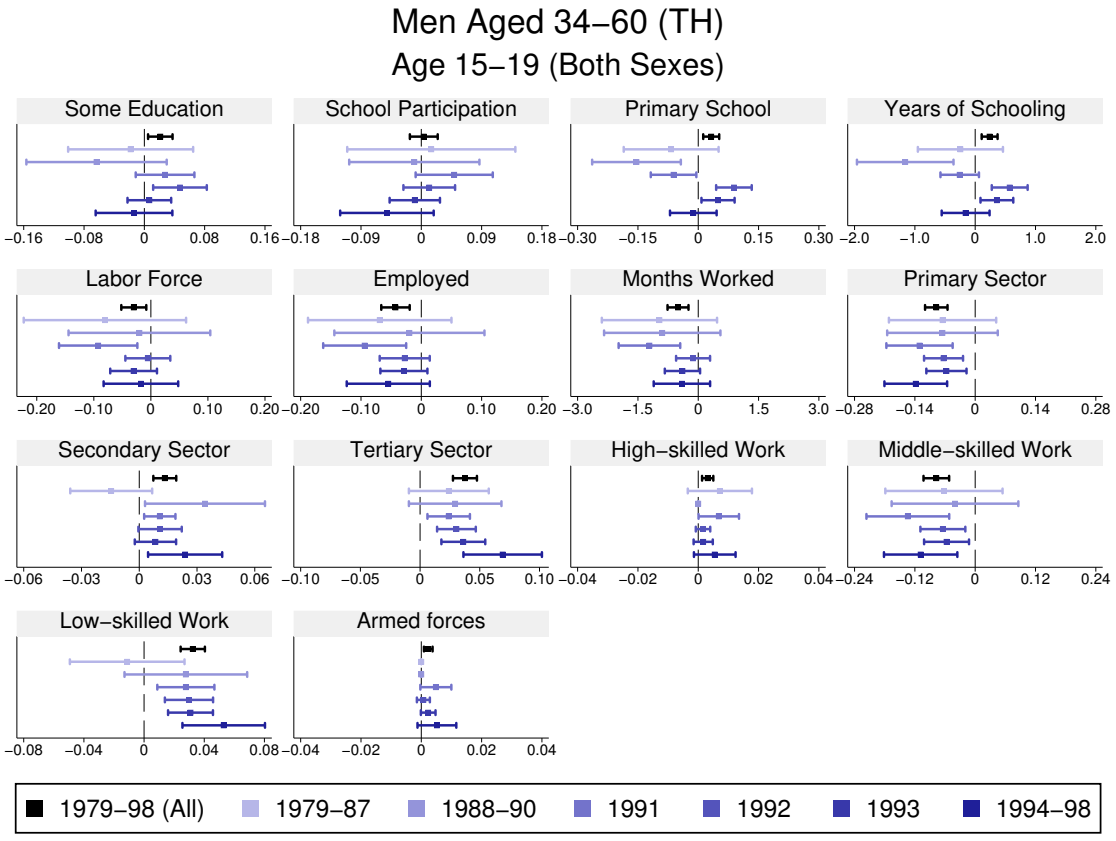
Figure A13: Heterogeneity – Timing of Repatriation (Age 20-60)

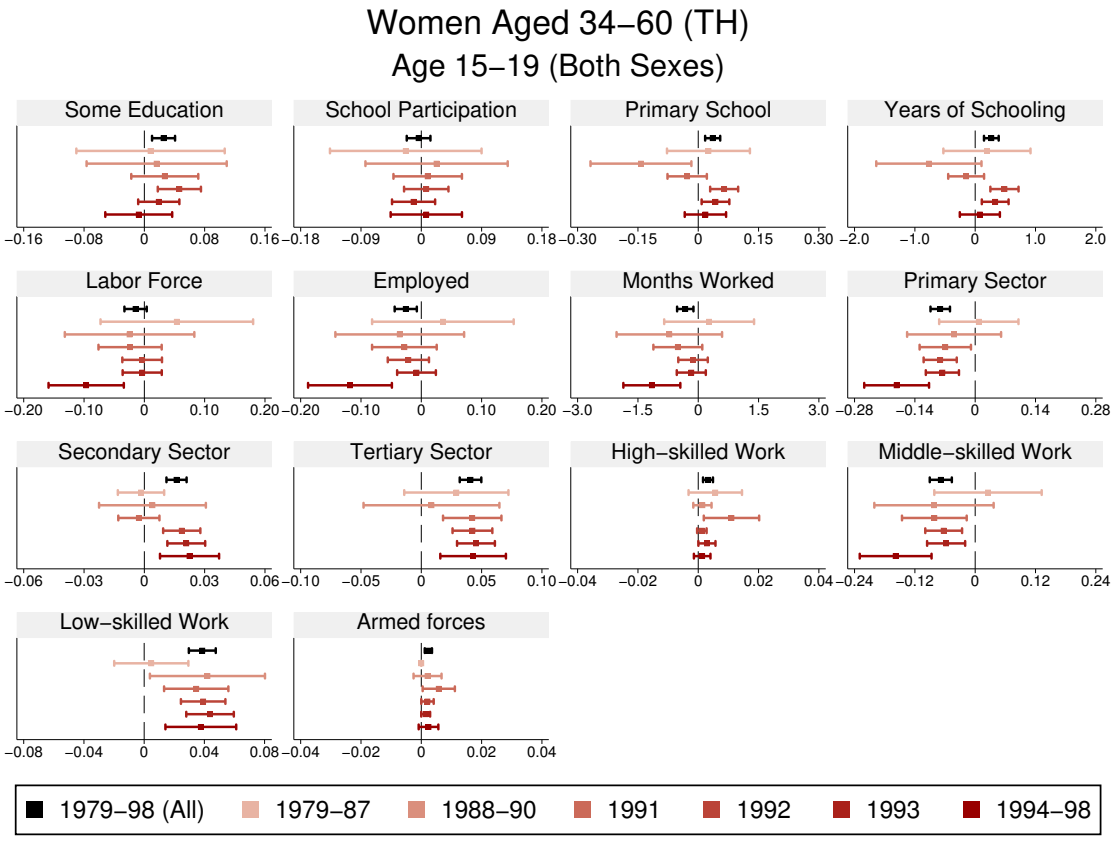



B-2. Vietnam

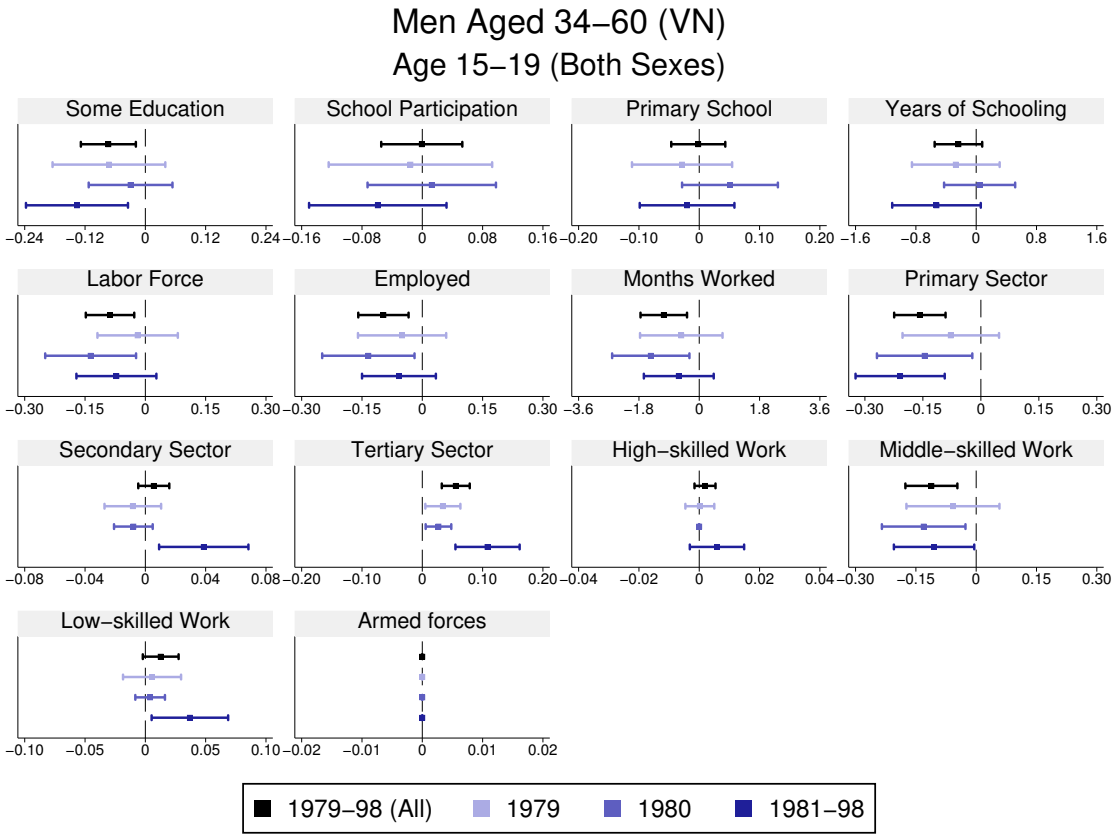
Figure A13: Heterogeneity – Timing of Repatriation (Age 20-60)

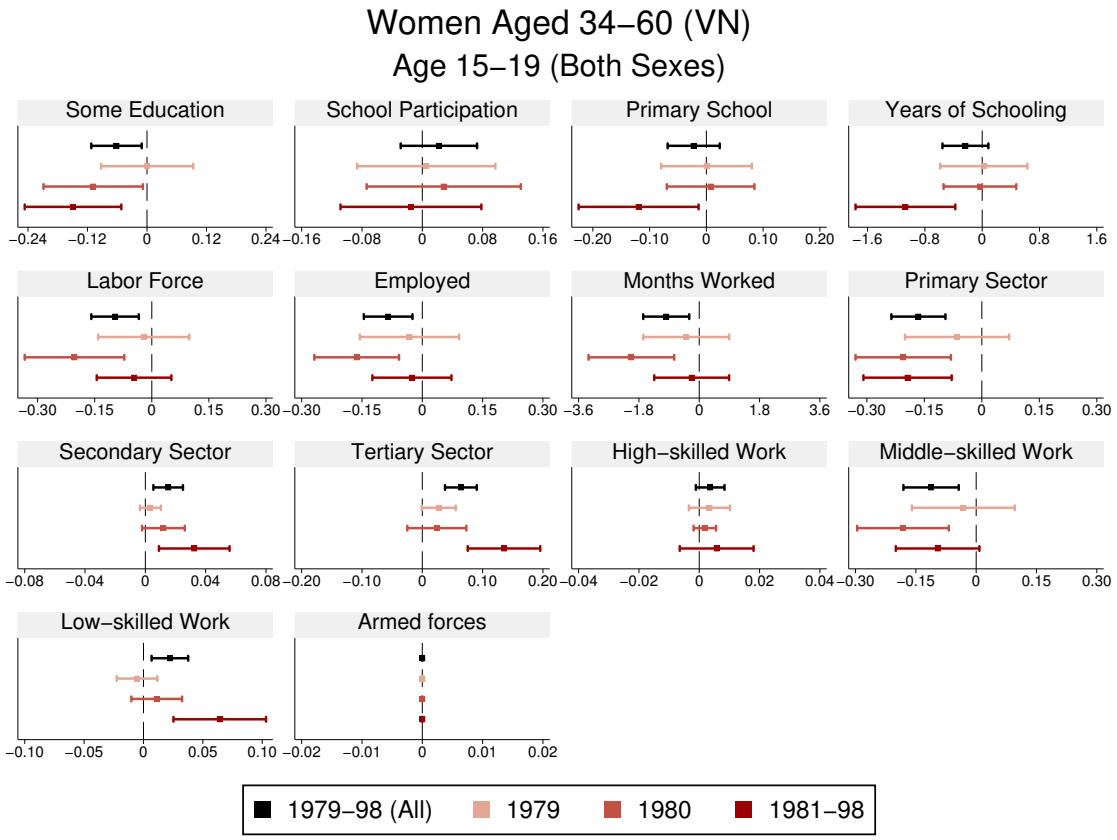



A-1. Thailand

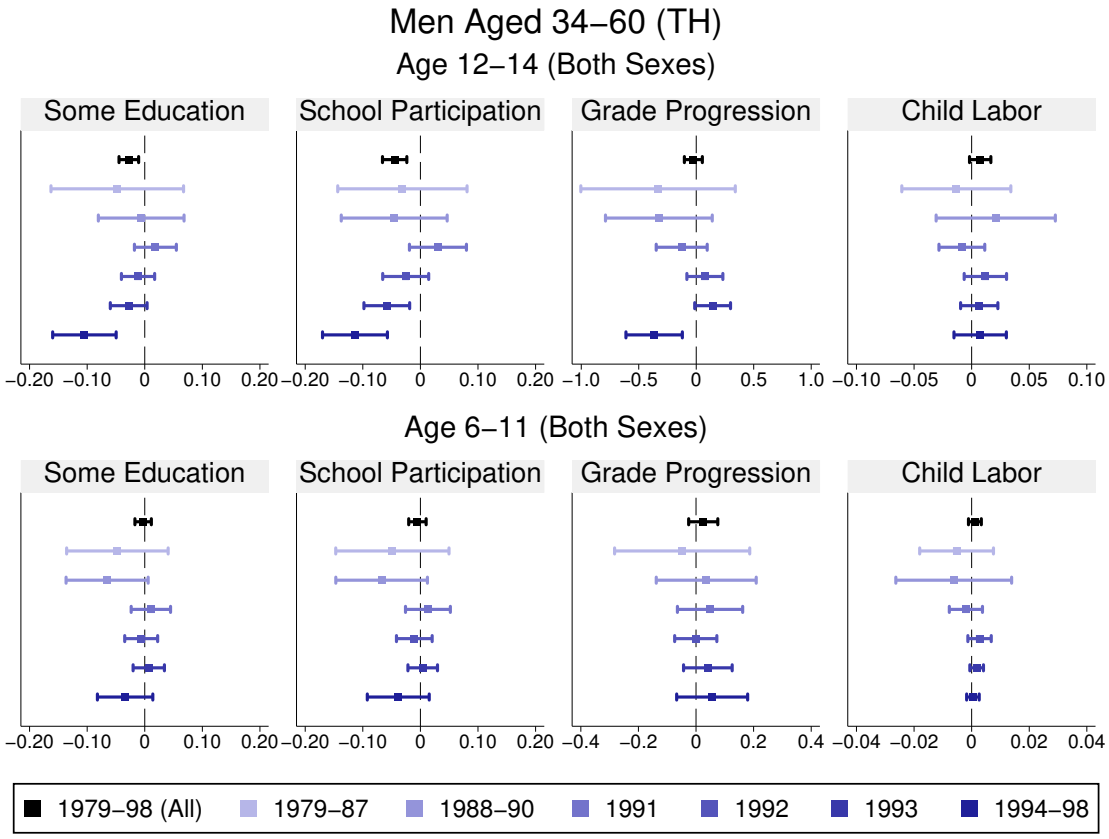
Figure A14: Heterogeneity – Timing of Repatriation (Age 6-19)

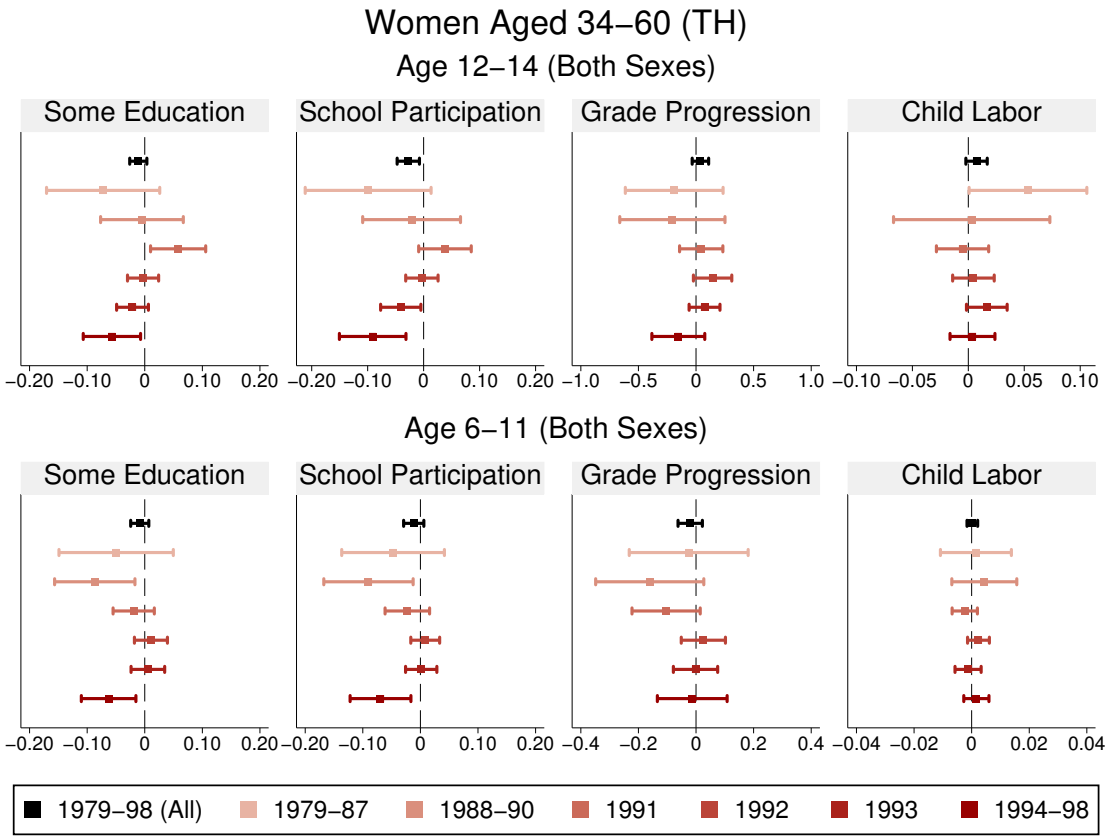



B-1. Vietnam

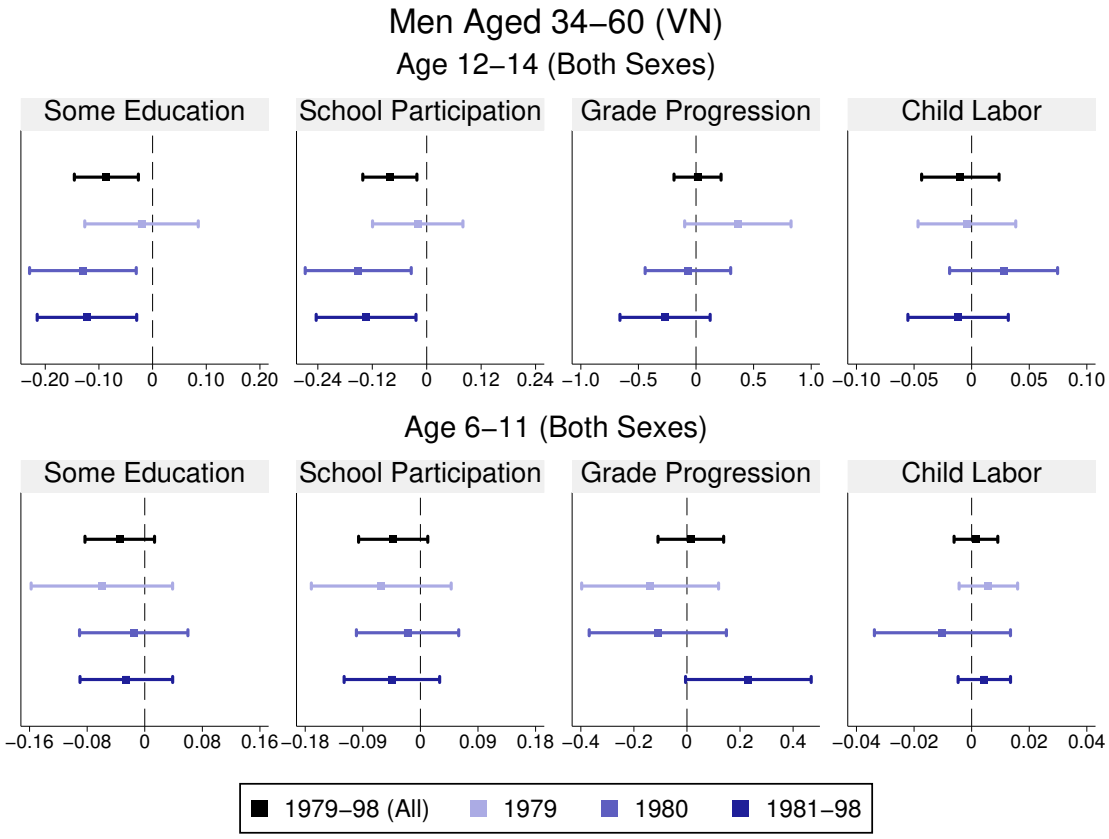
Figure A14: Heterogeneity – Timing of Repatriation (Age 6-19)

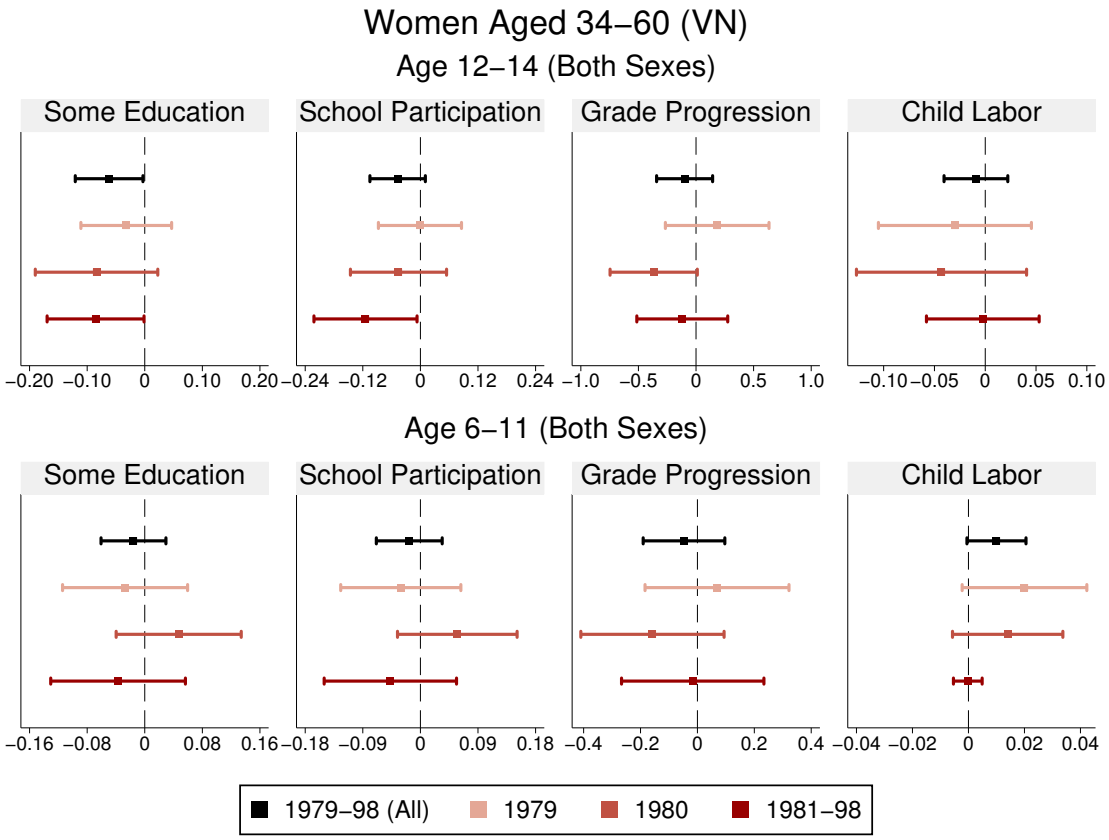



A-2. Thailand

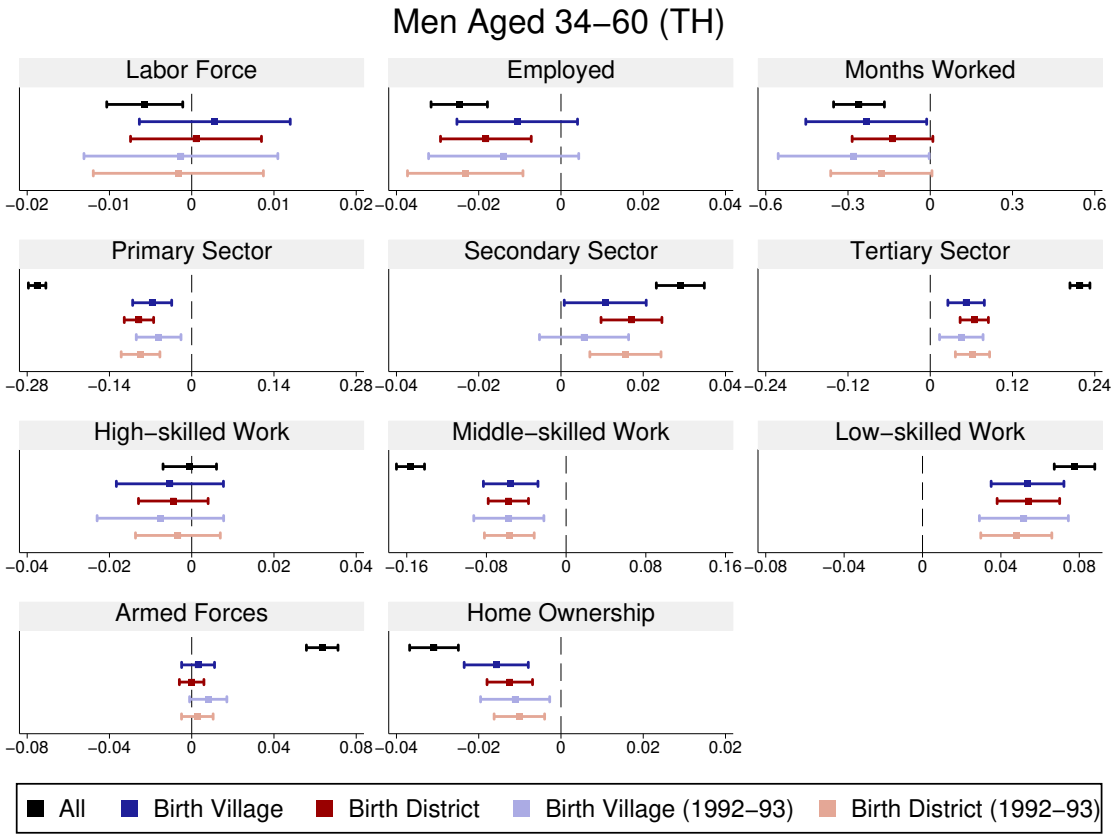
Figure A14: Heterogeneity – Timing of Repatriation (Age 6-19)

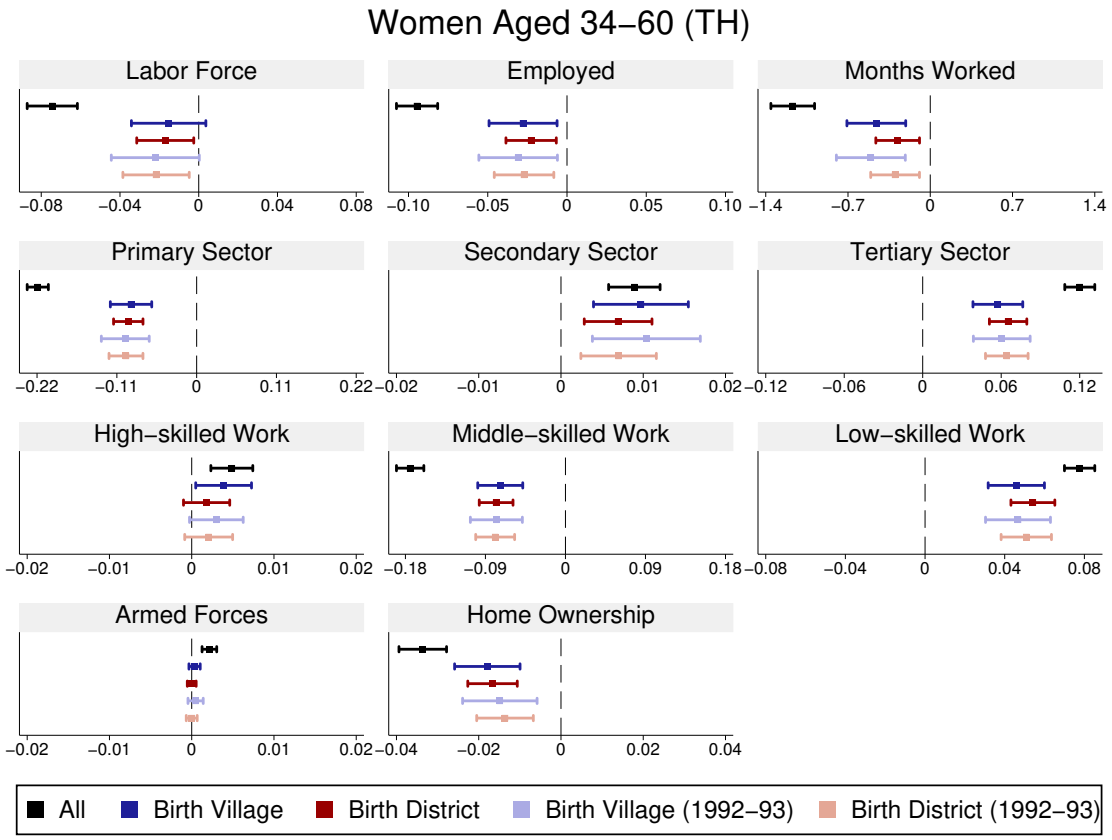



B-2. Vietnam

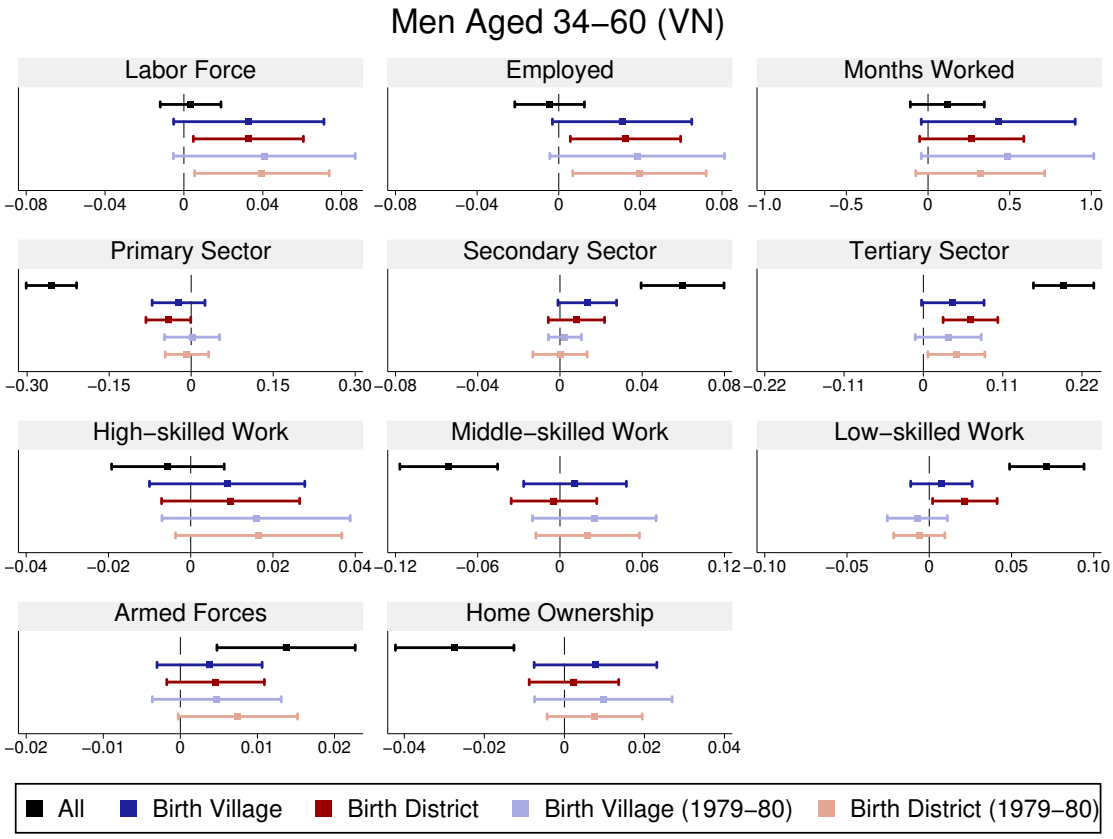
Figure A14: Heterogeneity – Timing of Repatriation (Age 6-19)

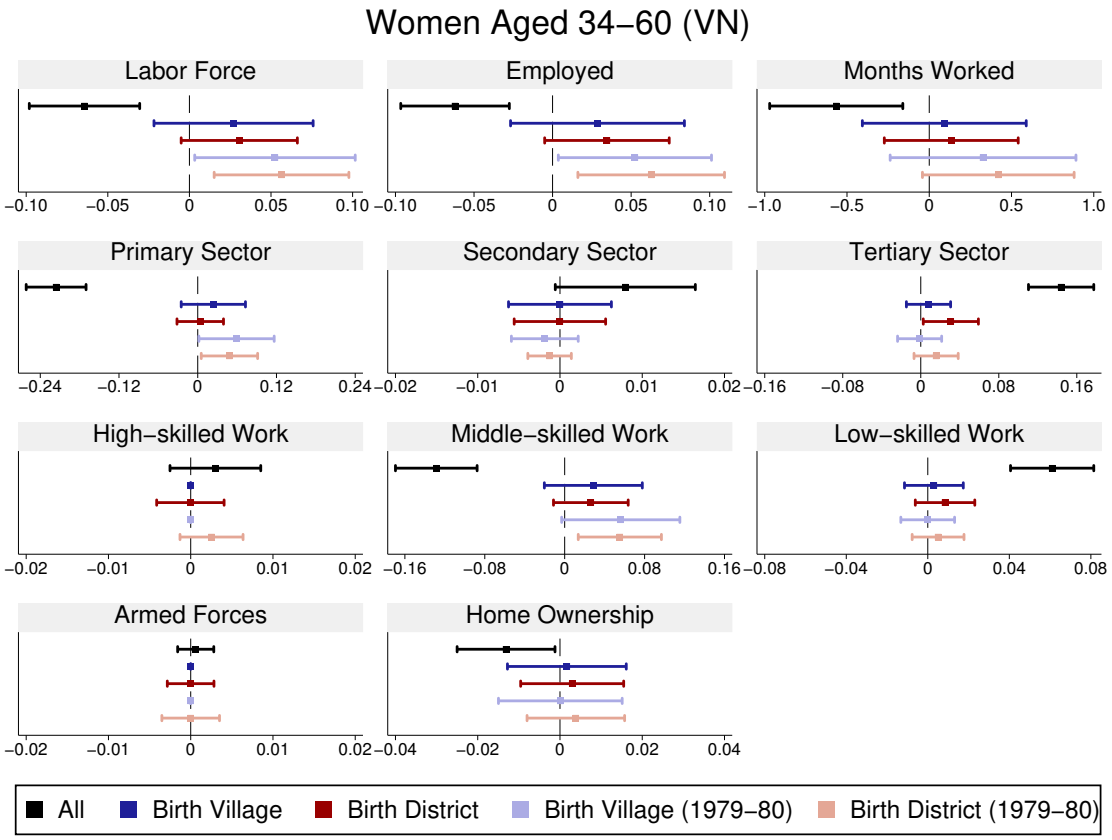



A-1. Thailand

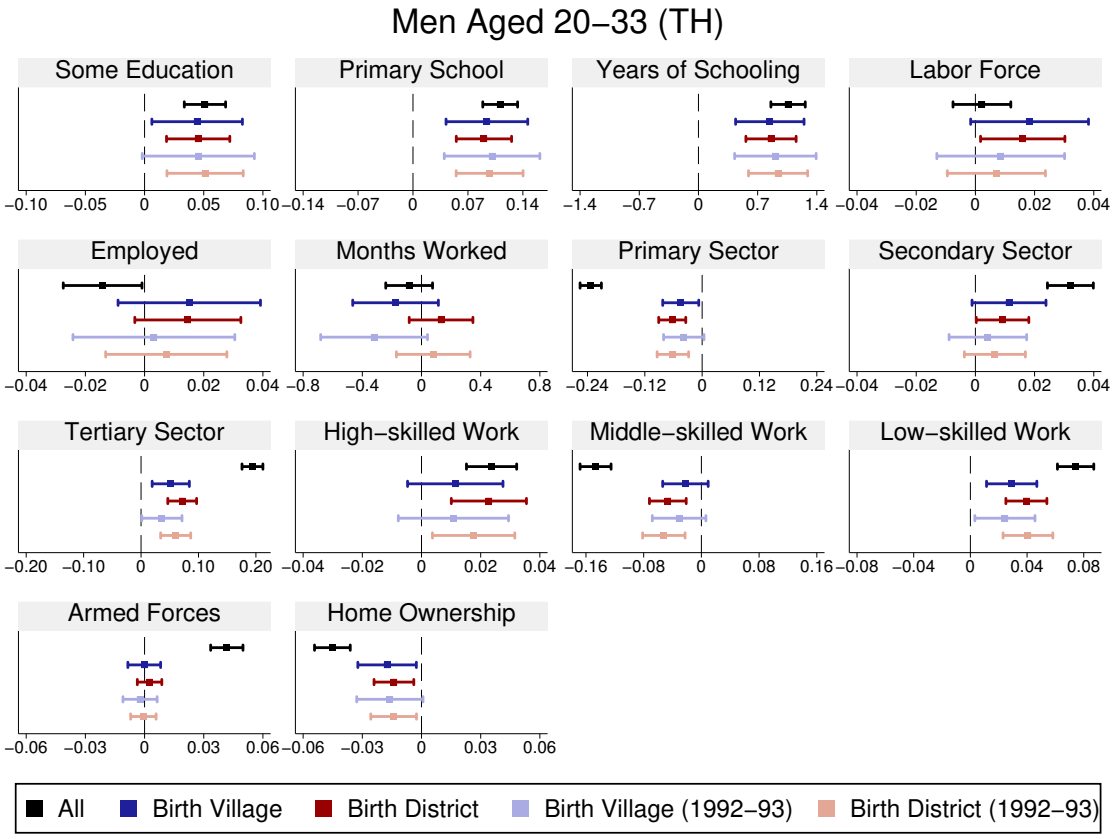
Figure A15: Heterogeneity – Place of Birth (Age 20-60)

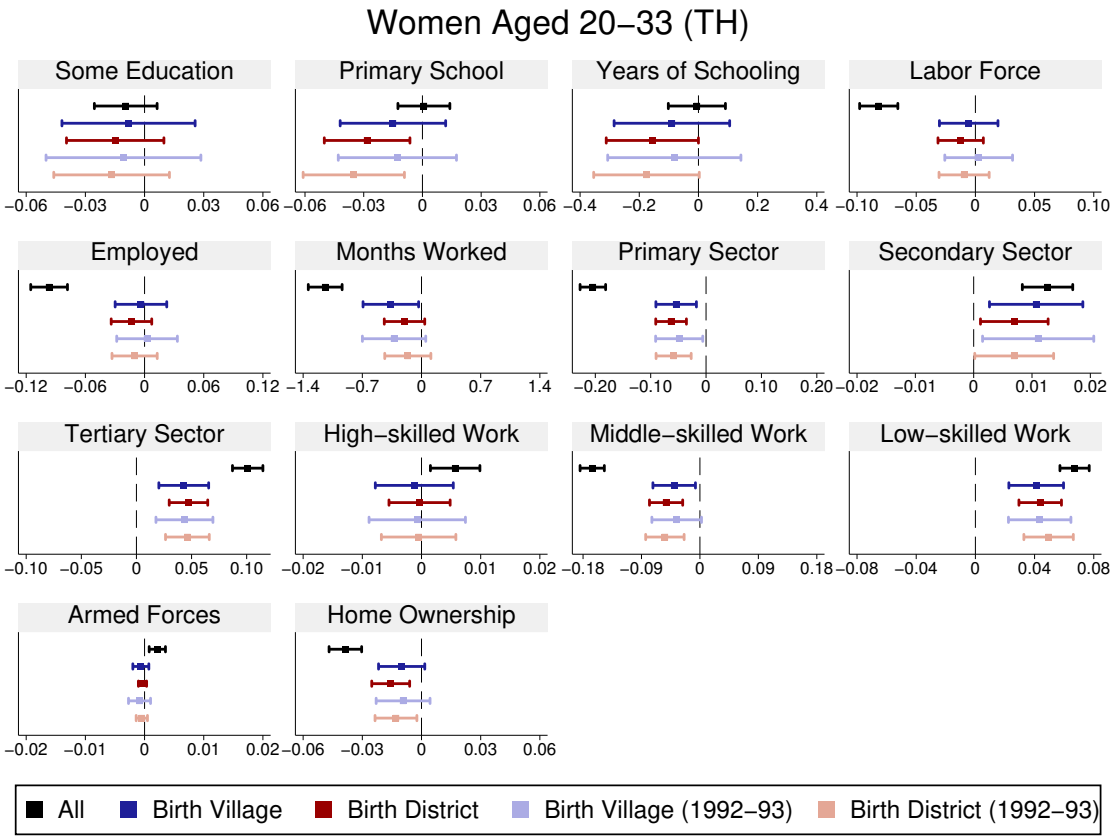



B-1. Vietnam

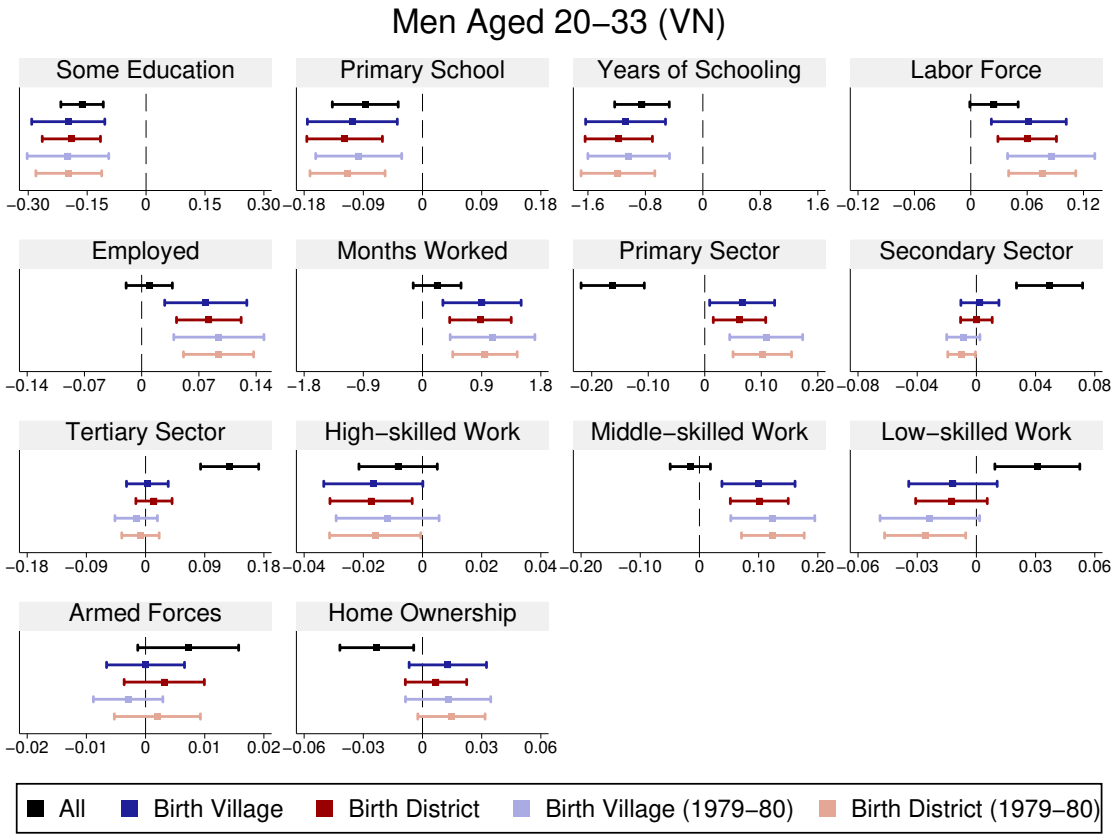
Figure A15: Heterogeneity – Place of Birth (Age 20-60)

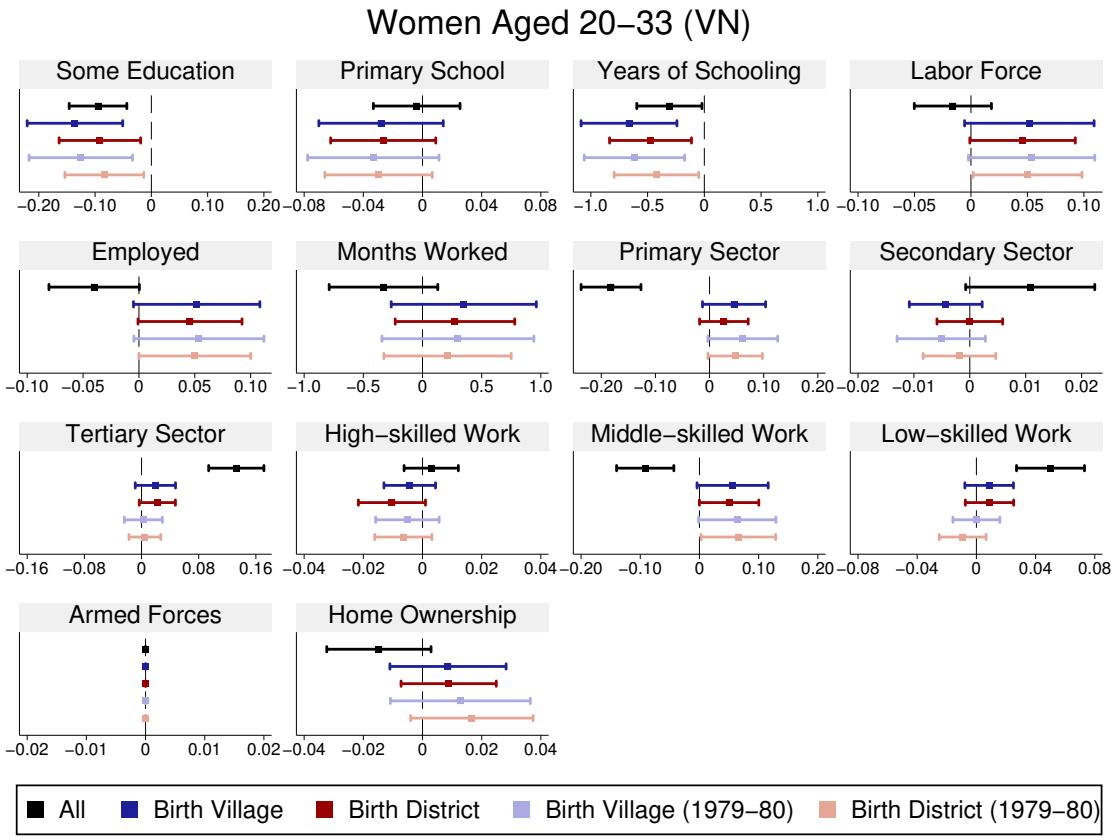



A-2. Thailand

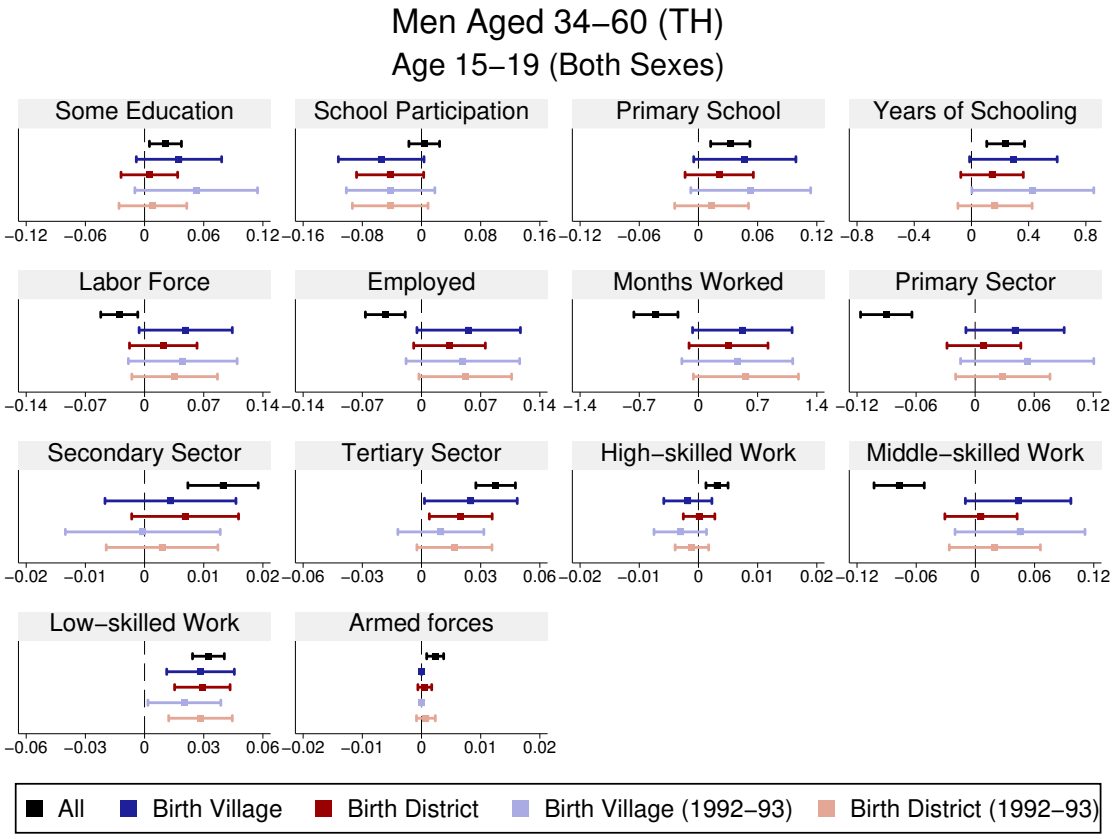
Figure A15: Heterogeneity – Place of Birth (Age 20-60)

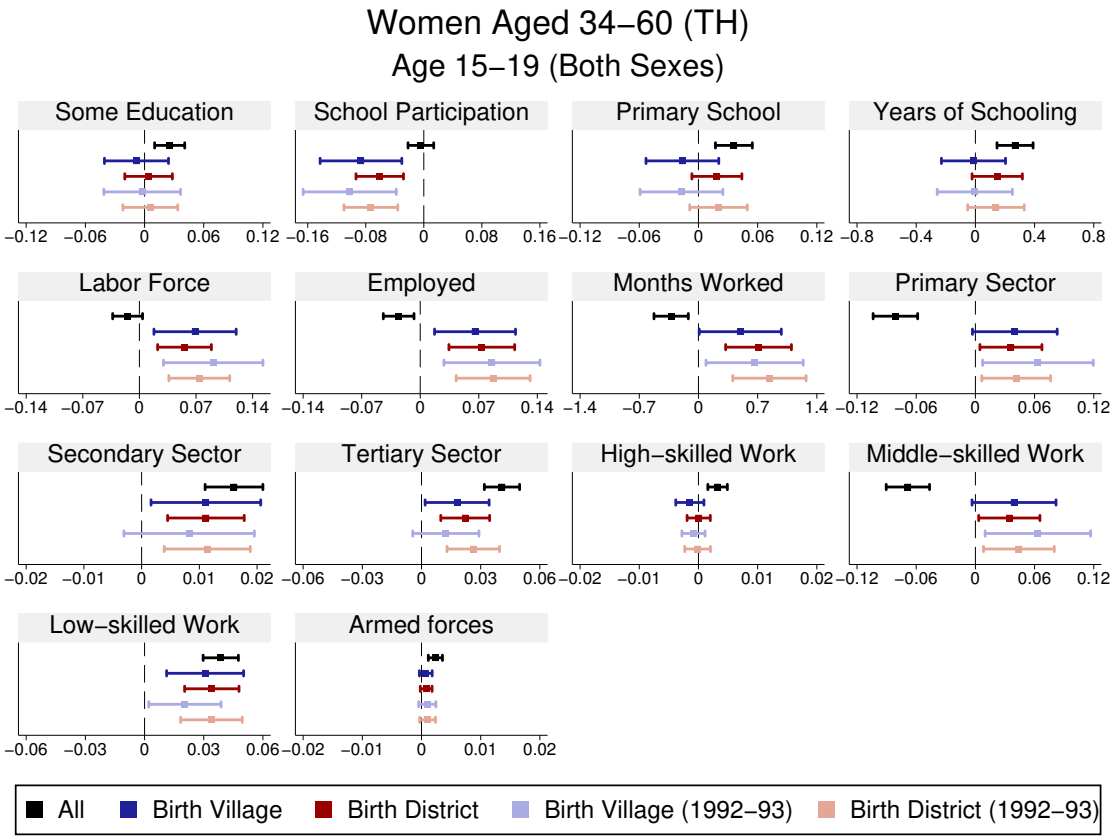



B-2. Vietnam

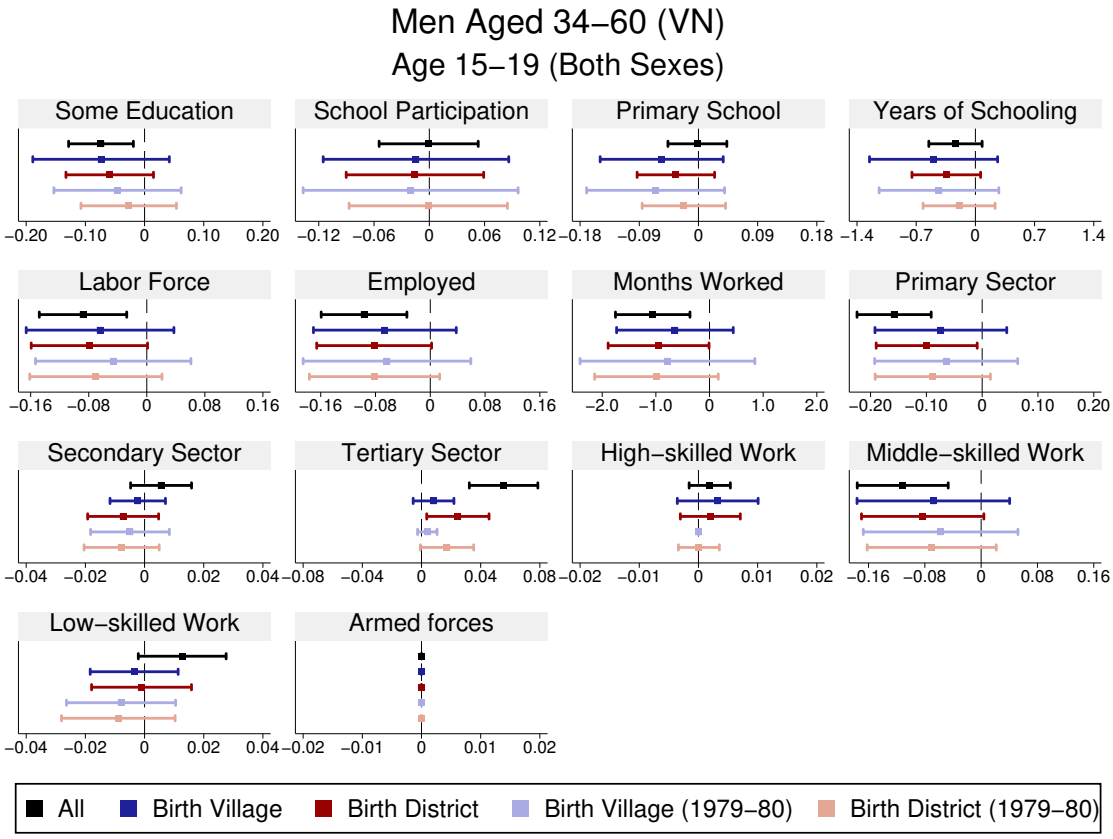
Figure A15: Heterogeneity – Place of Birth (Age 20-60)

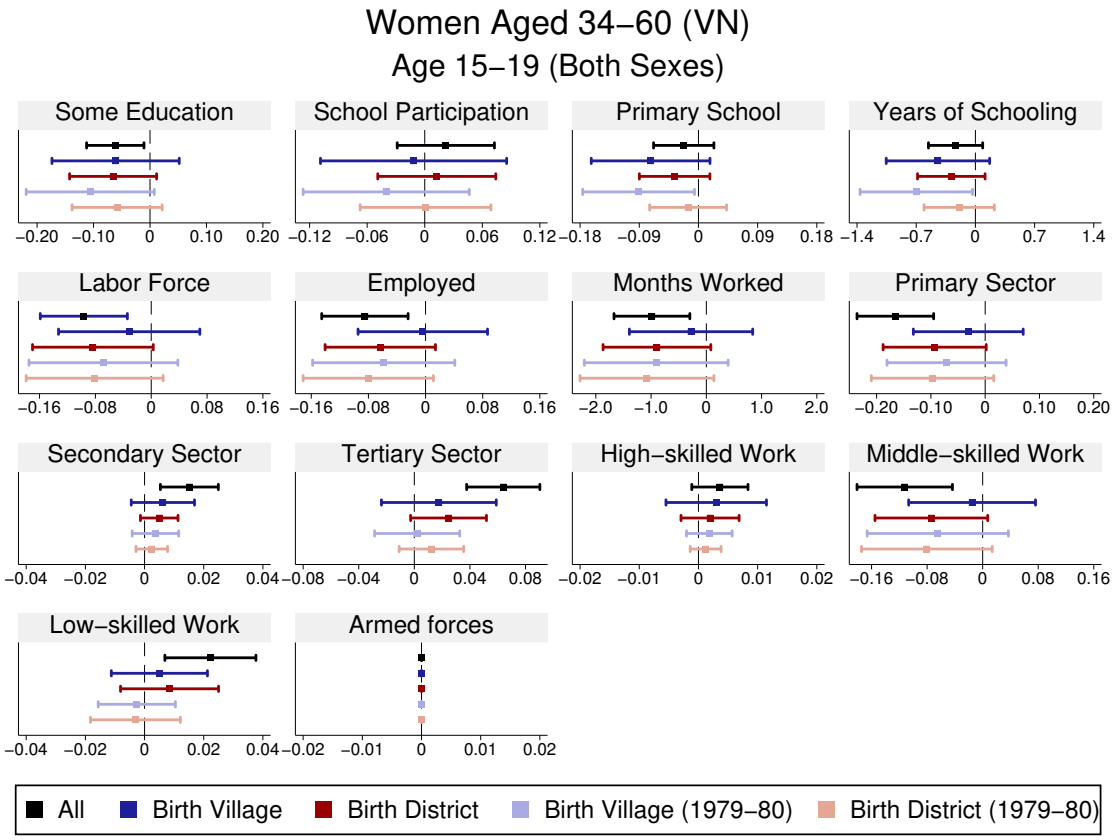



A-1. Thailand

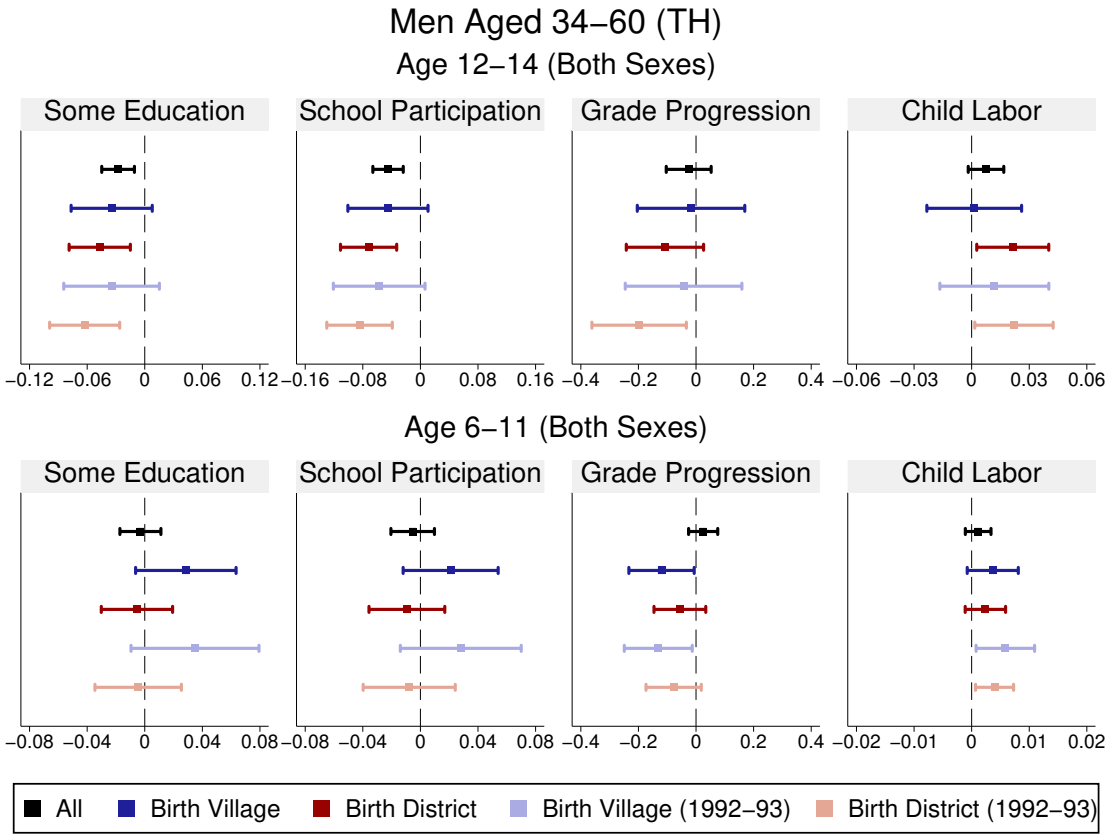
Figure A16: Heterogeneity – Place of Birth (Age 6-19)

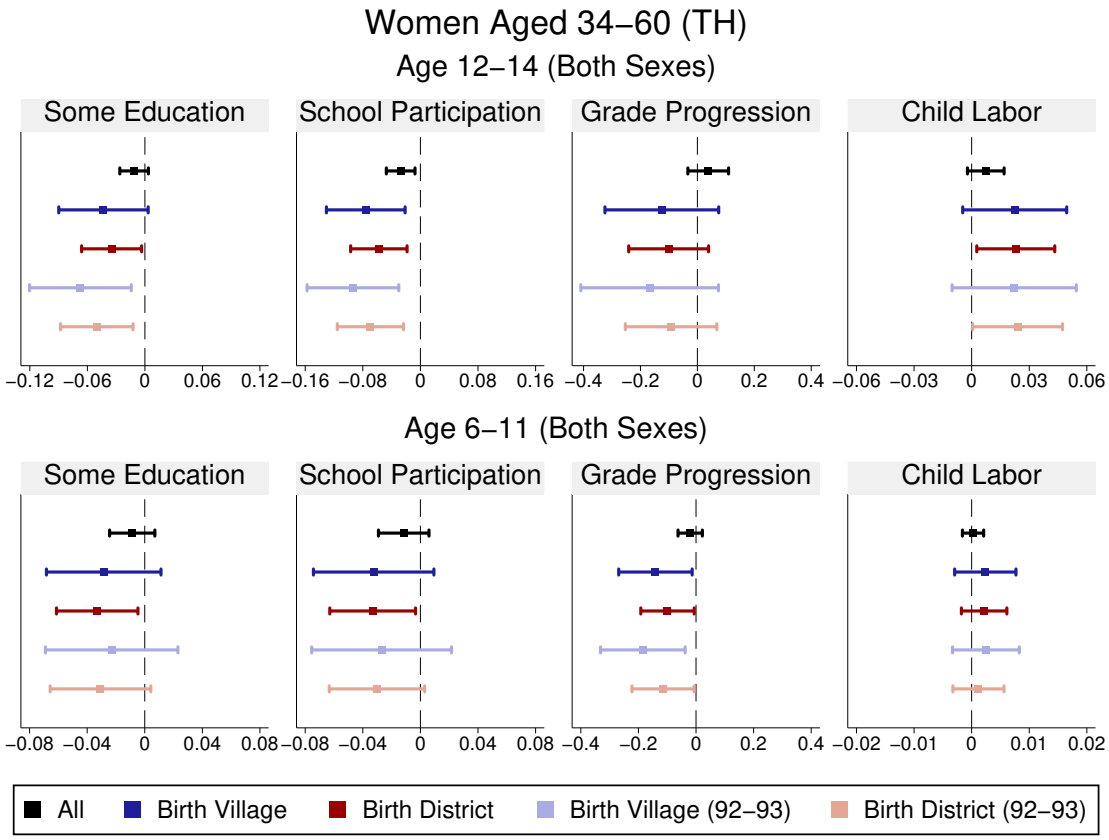



B-1. Vietnam

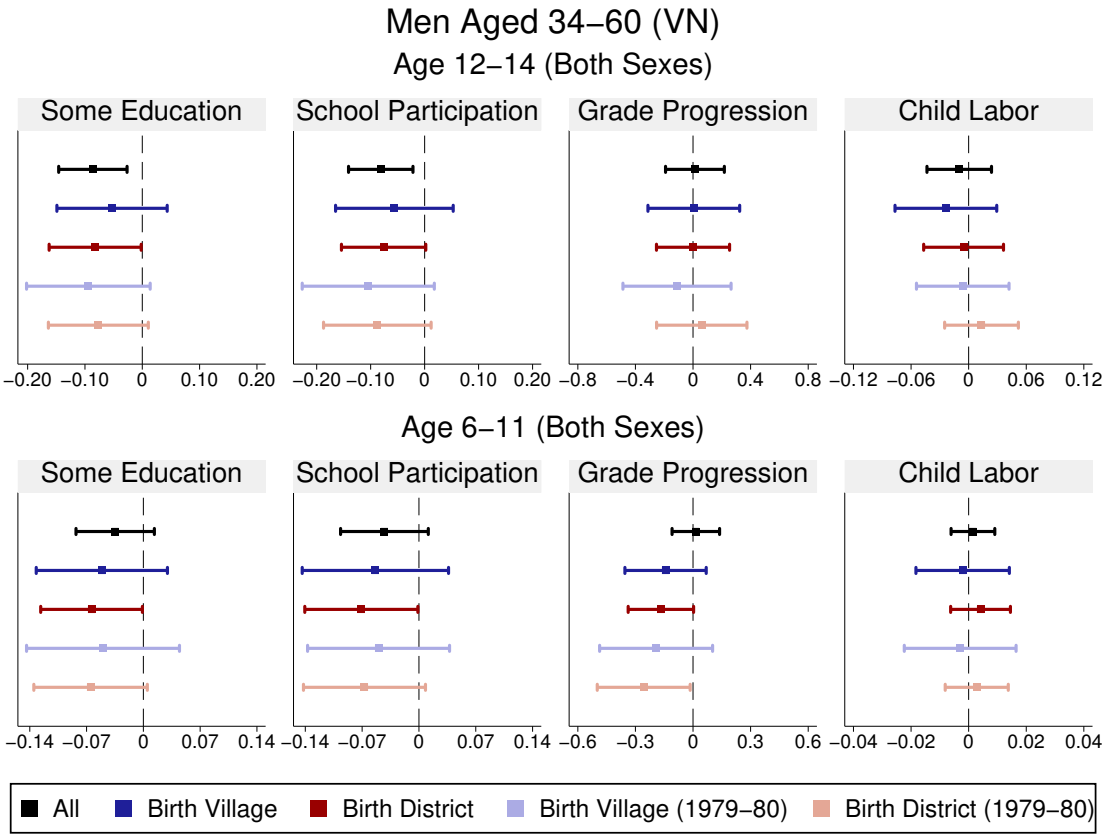
Figure A16: Heterogeneity – Place of Birth (Age 6-19)

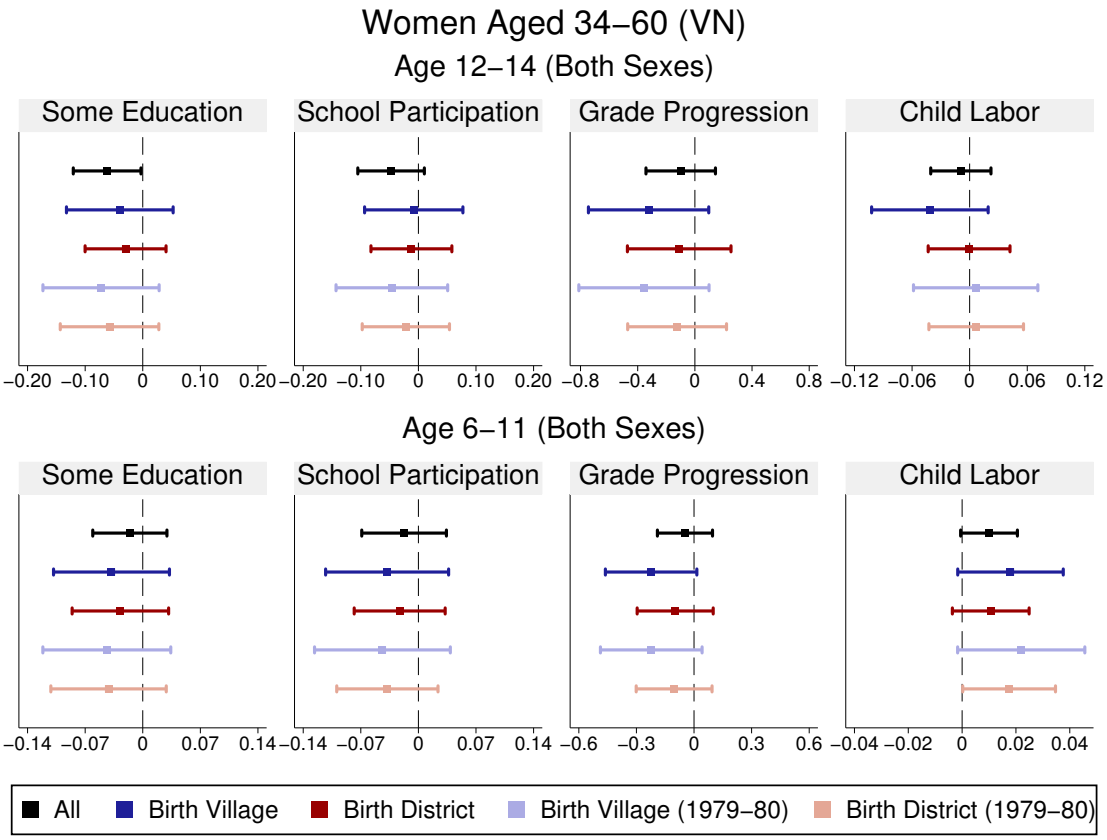



A-2. Thailand

Figure A16: Heterogeneity – Place of Birth (Age 6-19)




B-2. Vietnam

Figure A16: Heterogeneity – Place of Birth (Age 6-19)

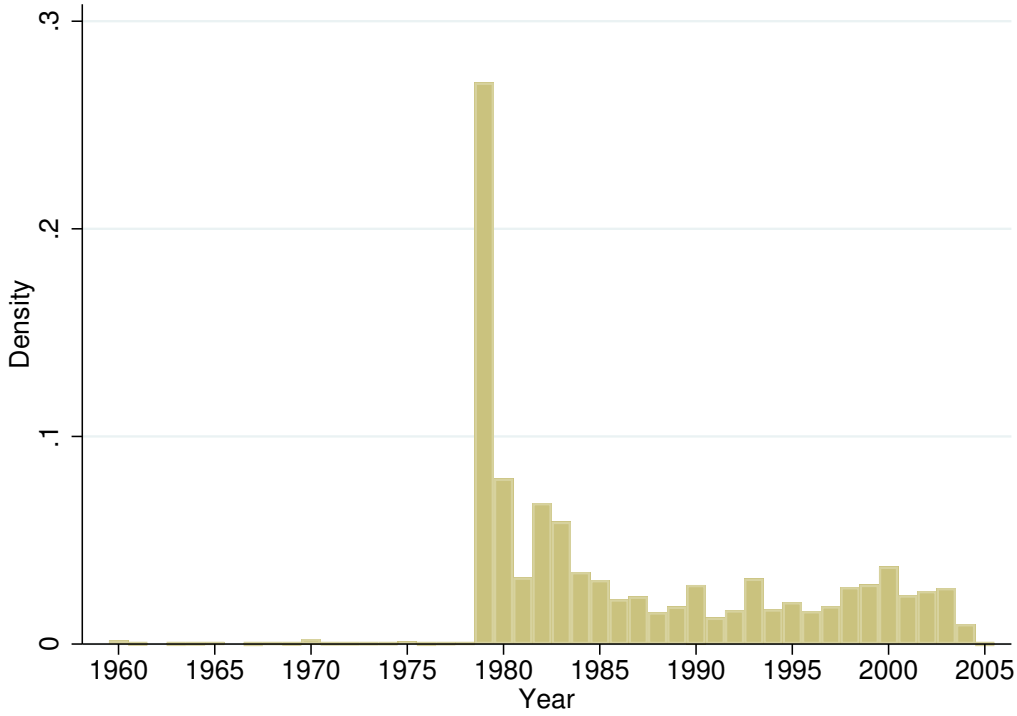



Figure A17: Timing of Access to Agricultural Land

*Note*: The figure shows the timing of access to agricultural land for Cambodian households in the 2004 Cambodia Socio-Economic Survey data.

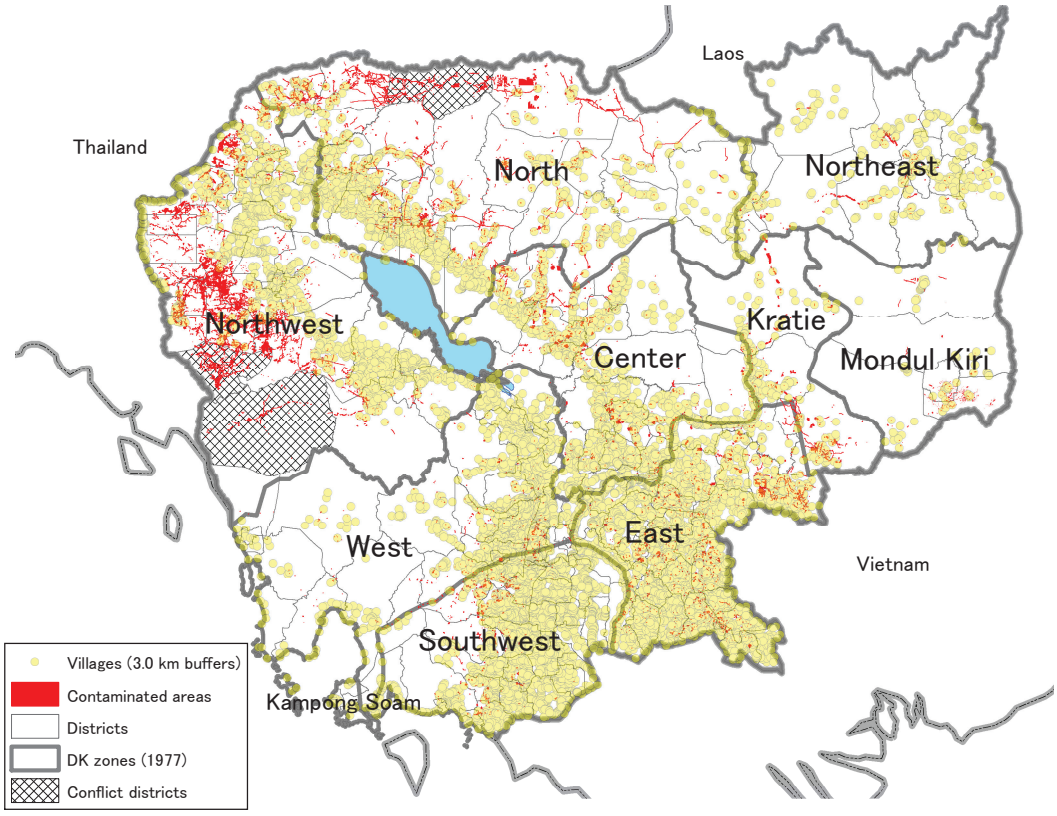



Figure A18: Areas Contaminated with Landmines and UXOs in 1992

*Notes*: The figure shows the landmine and UXO contaminated areas in 1992 (before clearance started), which are depicted based on the results of baseline survey (BLS) and clearance by the Cambodian Mine Action and Victim Assistance Authority (CMAA). The 1977 administrative zones of the Pol Pot regime (DK zones (1977)), the 1998 districts, conflict districts (not covered by the 1998 Census), and 3.0 km village buffers are also depicted.

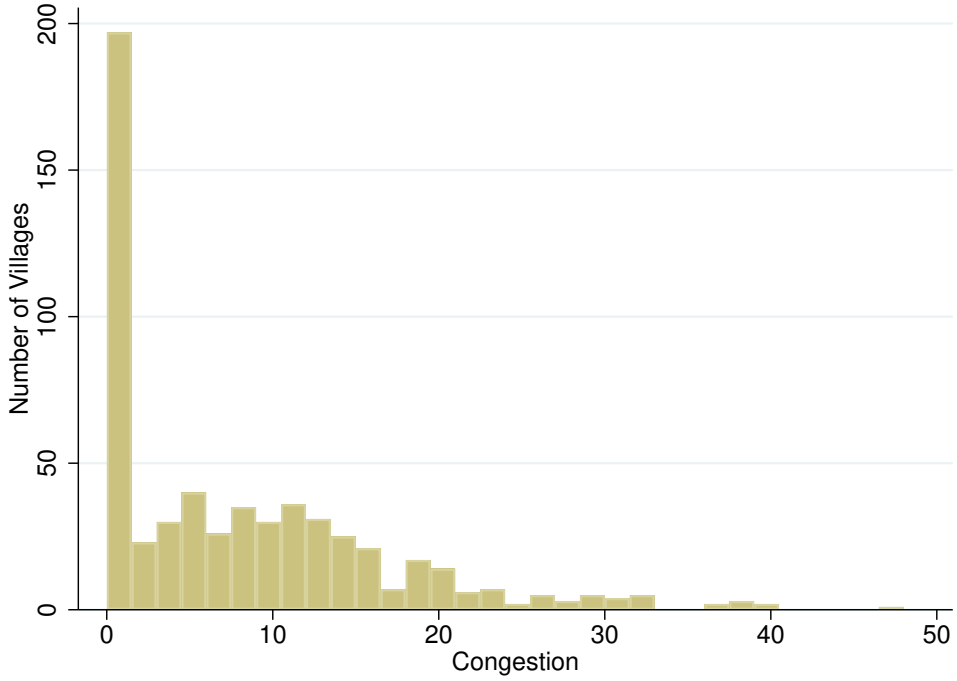



Figure A19: Distribution of Congestion Measure

*Note*: The figure shows the distribution of the village-level congestion measure used in the analyses based on the 2004 CSES data; the congestion measure is defined as stayer density per non-contaminated area ($km^2$) within a 3.0 km buffer zone around each village point before clearance started.