



Regular Article

Language training, refugees' healthcare integration, and the next generation's health[☆]Thang Dang^{a,b,c,1} ^a INVEST Research Flagship Centre, University of Turku, Turku, Finland^b Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway^c University of Economics Ho Chi Minh City, Ho Chi Minh City, Viet Nam

ARTICLE INFO

JEL classification:

J10
J24
J30
J60
J61
J70
E64
I21
I30

Keywords:

Refugee
Language training
Healthcare integration
Intergenerational spillovers
Health at birth
Norway

ABSTRACT

Exploiting a policy reform mandating participation in a 300-h language training course in Norway, I investigate the multigenerational health effects of a comprehensive language training program. I document the significant positive impacts of the program, not only on the long-term integration of refugees into the healthcare system but also on the health of the next generation. The program improves the cumulative use of primary care services for 12 years after arrival by 44%. This increase is proposed to partially work through refugees' increased labor market attachment and economic success. The program also results in spillover benefits for the health at birth of the children of refugees with 21–49% decreases in premature birth, low birth weight, and short birth length. These intergenerational effects probably operate via significant improvements in the home environment and maternal care during pregnancy.

1. Introduction

Many more refugee-receiving countries around the world have included language training programs into their integration policies to assist resettled refugees with the acquisition of standard language proficiency (OECD, 2023). The share of the Council of Europe member states that have introduced formal language training requirements for

permanent residence or citizenship has increased gradually, from 58% in 2007 to 78% in 2013 and to 83% in 2018 (Rocca et al., 2020). English-speaking countries, including Australia, Canada, the United Kingdom, and the United States, have also made access to permanent resident status conditional on passing a language test (McLeod, 2023). Despite this policy trend, little is known about the degree to which language training enables refugees to integrate into the host society. The

[☆] This research was funded by the Research Council of Norway through its Centers of Excellence funding scheme (project number 262700), and the Research Council of Finland (project numbers 355153 and 345546).

E-mail address: thang.dang@thangdang.org.

¹ I am grateful to Thomas Cornelissen and Emma Tominey for the valuable discussions and the feedback. I would like to thank Sara Abrahamsson, Linea Hasager, Martin Flatø, Thor Indseth, Jonathan Wörn, and participants in seminars at the Norwegian Institute of Public Health, the Ragnar Frisch Centre for Economic Research, the Norwegian Institute for Social Research, the Norwegian University of Science and Technology, Curtin University, the Norwegian Directorate for Integration and Diversity (IMDi), the BeNA 20 Year Jubilee Conference, and the Finnish Health Economics Seminar Series (FINHESS) for their comments. All errors are mine.

available evidence is mostly from correlational studies that show that language training or language proficiency is positively correlated with integration outcomes for refugees such as labor market income (Clausen et al., 2009; Chin and Cortes, 2015; Evans and Fitzgerald, 2017; Auer, 2018). While knowledge about the causality of the relationship is vitally important for evaluating the effectiveness of language training programs in refugees' integration, especially in the context of the increasing number of refugee immigrants across the globe (Chiswick and Miller, 2015),² the problem has been understudied (Dustmann et al., 2017; Hatton, 2017; Li and Sah, 2019; Brell et al., 2020).

Two key challenges obstruct efforts to produce reliable causal estimates of the link between language training and refugees' integration, namely the unavailability of high-quality data and the lack of credible identification strategies. First, the existing studies typically draw on relatively small and/or cross-sectional samples of refugees from survey data (Chiswick and Miller, 2015). The use of such data reduces the precision of the estimates and equally importantly does not allow researchers to observe how refugees integrate into the host country over a long period. Second, identification difficulties affect attempts to isolate the impact of language training from the influence of other traits on refugee accomplishment in destination countries. For example, refugees who invest in language training likely possess some characteristics, such as greater self-awareness, which are conducive to labor market attachment even in the absence of language training, but those traits are seldom observable. Therefore, the endogeneity obtained between language skills and integration outcomes makes it difficult to disentangle the causal impact of language training without a randomized controlled trial (RCT). Such RCTs are typically costly and have limited external validity (Borjas, 1994; Chiswick and Miller, 1995).

In this paper, I make progress in the effort to tackle these data and identification problems in the context of Norway, providing new insights into the causal impacts of a language training program on

refugees' integration into the healthcare system and on the health of the next generation. My research questions—whether and how a language training program improves refugees' use of healthcare services and their children's health early in life—are strongly motivated by the established fact that language barriers have long been recognized as a key determinant of limited healthcare access among minority groups, particularly immigrants (Narayan, 2013; Black and Kunz, 2024). Language skills play a dual role in health outcomes. They influence health investments both directly, by improving access to resources such as income, and indirectly, by enhancing the efficiency of these investments through better information processing and communication. This hypothesis aligns with the Grossman model, which highlights the role of human capital—such as language skills—in optimizing health outcomes (Grossman, 1972). Moreover, language skills are intergenerationally correlated, contributing to persistent socioeconomic disparities among children born to parents with limited local language proficiency. Empirical evidence supports this correlation, indicating that children raised in households with language deficiencies often face lower healthcare and health outcomes (Narayan, 2013). Medical studies have underscored the detrimental effects of inadequate communication skills on in-utero and neonatal health through poor knowledge of health-related behaviors and suboptimal use of healthcare services (Obregon et al., 2019; Palau et al., 2019). However, establishing a causal

² As of 2024, the number of refugee immigrants who have been forced to flee their home countries is nearly 37.9 million, an increase of 94% over the last decade, from 19.5 million in 2014 (UNHCR, 2024b). These figures are projected to rise in the years to come due to increasingly widespread persecution, conflict, violence, human rights violations, and other shocks that lead to serious societal disorders in many parts of the world (OECD, 2016; Devictor et al., 2021; UNHCR, 2021). As of November 2024, the 2022 Russian invasion of Ukraine had resulted in over 6.7 million Ukrainian refugees fleeing to various countries (UNHCR, 2024a).

relationship between language skills and health outcomes remains a significant challenge, as much of the existing literature relies on small-scale, observational, and exploratory studies. A limited but growing body of research provides causal evidence of the impact of language skills on child health outcomes. For instance, Auer and Kunz (2024) exploit a natural experiment in Switzerland to demonstrate that exogenous assignment of local language proficiency significantly improves child well-being as early as the prenatal stage, as reflected in health at birth, a key metric for examining the intergenerational transmission of inequality in the modern economic literature (Almond and Currie, 2011; Almond et al., 2018).

By addressing these gaps in the literature, my study aims to contribute to a more nuanced understanding of how targeted language training programs can mitigate barriers to healthcare access for refugees and generate positive intergenerational health impacts. I first leverage extremely rich administrative data from Norwegian registers that cover the entire population, including refugees, to produce parameter estimates with sufficient statistical power and to improve the precision of the estimates. The extent to which language training helps refugees improve their connection with the healthcare system of the host country is fundamental to evaluating the effectiveness of such policy interventions, but the causal evidence from the economics literature is extremely scant. In this paper, I use the Norwegian primary care registry, a distinct source of data that enables me to provide the first analysis of the impact of language training on the use of healthcare services among refugees. Equally importantly, the register contains related information that was recorded for 12 years after the first year of settlement in Norway, which allows me to study refugees' integration into the local healthcare system over the long term.

In addition, the data contains identifiers that link family members, in particular children and parents, which enables me to explore the intergenerational spillover effects of language training on the health at birth of the children's generation; these effects have not been studied before. The question of whether the impacts of language training span more than one generation is crucially important, given that the family is the potential peer group within which the benefits of language training as a form of human capital investment are likely to spill over due to strong social interactions (see Kuziemko, 2014; Bau and Fernandez, 2023). Moreover, the entries for each refugee are linked across registers through a unique personal identifier, which provides rich information on demographic, family, economic, and healthcare behaviors. I exploit these rich sources to examine the mechanisms that may plausibly underlie the main effects on both generations and then to tell a comprehensive story about the *non-economic* impacts of language training programs on refugees.

Norway provides a suitable policy setting for the causal evaluation of language training programs.³ I particularly exploit a Norwegian policy change to identify exogenous exposure to language training. Regulated by the Introduction Act, refugees and their family members have been obliged to participate in 300 class hours of Norwegian language training from September 1, 2005 (hereafter "the reform").⁴ This reform was for the first time highly significant in increasing the probability of refugees attending Norwegian language courses for a sufficient period to fulfill the language requirement for their asylum being granted. The reform offers a plausible quasi-experimental strategy for ruling out the causal effect of language training from that of other traits that are likely to

³ Over the last few decades, along with receiving an increasing number of refugees, Norway has implemented a series of policies and reforms to promote refugees' integration. Norway has also received similar types of refugees as other advanced European countries. The lessons from the Norwegian context can therefore be relevant to policy makers in other countries that have a similar institutional setting.

⁴ I will often mention September 1, 2005 as "the cutoff date" in the remainder of the paper.

Table 1
Summary statistics.

	(1)	(2)
	Mean	SD
(a) Directly affected refugees		
Treated	0.404	0.491
<i>Primary care services</i>		
All consultations	53.971	44.730
Musculoskeletal consultations	14.870	18.006
Psychological consultations	5.216	12.021
Digestive consultations	4.799	7.139
<i>Economic outcomes</i>		
Labor market earnings	194.934	159.754
Sick leave benefits	9.606	17.273
Disposable income	171.698	112.861
(b) Refugee children		
Treated	0.391	0.488
<i>Health at birth</i>		
Preterm birth	0.056	0.230
Low birth weight	0.049	0.215
Short birth length	0.131	0.338
<i>Prenatal conditions</i>		
Parent's income	14.650	14.489
Parent's marriage	0.573	0.495
Mother's primary care consultations	10.758	8.910
Mother's prenatal checkups	4.564	4.459
Mother's prenatal supplements	0.415	0.493

Notes: This table shows summary statistics for the outcome variables for the directly affected generation (4,965 observations) in panel (a) and the children's generation (4,892 observations) in panel (b). Mean and s.d. values are reported for the total sample obtained using a bandwidth of ± 360 days for the distance between the arrival date and September 1, 2005. The children's sample include children born in 2005–2018 to the first parent arriving in Norway. See Section 3.2 for details on how these variables are constructed.

Data source: Own calculations based on the data provided by Statistics Norway.

drive the integration outcomes of interest. The unforeseeable and discontinuous implementation of the reform enables me to establish the causality of the link between exposure to language training and the outcome of interest. Using a regression discontinuity (RD) design, I compare an outcome such as the number of doctor visits among refugees who arrived just before and just after the cutoff date.⁵ I further provide evidence on density checks, balance, and placebo tests to confirm the validity of the identification strategy.

I present two sets of findings. First, I find that language training has significant and positive direct effects on refugees' integration into the healthcare system. Baseline RD estimates indicate that exposure to the language training program increases the total number of primary care doctor visits by 19.2 over the 12-year observation period. Specifically, I find an increase of 3.1 visits for musculoskeletal consultations, 2.9 visits for psychological consultations, and 2.2 visits for digestive consultations. Compared to the pre-reform mean values, these effects are equivalent to increases of 44%, 29.6%, 84.6%, and 48.9%, respectively, for total, musculoskeletal, psychological, and digestive visits. Heterogeneity analysis reveals that the effects on total, musculoskeletal, and digestive visits are

concentrated in women. In contrast, the effect for psychological consultations is significantly greater in men than in women by 2.9 visits. In examining the potential channels that underlie these effects, I show that the program improves the labor market attachment and economic success among refugees. Specifically, the accumulation of 11 years of labor market earnings increases by Norwegian Krone (NOK) 199,110 (10.7%), that of sick leave benefits claimed increases by NOK 19,050

⁵ Though the curriculum has a small social knowledge element, I still interpret the effect of the reform as being equivalent to the effect of language training.

(33.7%) and disposable income over the same period increases by NOK 138,050 (8.5%).

Second, I document the significant spillover effects of language training that accrue to the next generation in virtue of considerable improvements in health at birth. Baseline estimates show that children of treated parents are less likely to be born prematurely by 1.9 percentage points (33.9%), less likely to have low birth weight by 4.4 percentage points (48.8%), and less likely to have low birth height by 3.2 percentage points (21.2%). I also shed light on the potential mechanisms behind these intergenerational effects by showing that prenatal conditions in the home environment and maternal care improve considerably as a result of parental exposure to language training. Improvements in the home environment, in particular, are obtained as a result of an increase of NOK 17,460 (11.7%) in the disposable income of the parents and in virtue of the 14.9 percentage point increase (29.1%) in the probability that the parents are married during pregnancy. Improvements in maternal care behaviors are as follows: an increase in the incidence of primary care consultations by 5.4 visits (58.5%), an increase in the intake of prenatal checks by 2.3 visits (61.5%), and an increase in the probability that the mother takes prenatal supplements of 4.6 percentage points (11.7%). The effects are robust to a battery of robustness checks and sensitivity tests.⁶

I make two distinct contributions to the scant literature on the causal effects of language training for refugees.⁷ First, I advance the extant knowledge of the *non-economic* impacts of language training or language proficiency by providing the first evidence on the integration of refugees into the healthcare system of the host country. Refugees typically experience traumatic journeys, making them the most vulnerable group of immigrants in terms of health issues (Kalt et al., 2013; Mishori et al., 2017).⁸ Although understanding how refugees integrate into the healthcare system of the destination country is important for evaluating integration policies (Satinsky et al., 2019), research on causal determinants of that integration is lacking (Dustmann et al., 2017; Mishori et al., 2017). My findings indicate

that language training improves the connection between refugees and the local healthcare system by promoting the use of primary care services.

My findings on the impacts of language training on labor market earnings, sick leave benefits, and disposable income as the key transmission channels complement the evidence on its economic impacts which is an important dimension of an integration policy. Since refugees typically exhibit lower economic integration than other types of

⁶ In terms of the magnitude, these effects are also in line with previous findings from the economics literature, which indicates that language training has relatively large impacts on both refugees and their children (see Section 9 for a related discussion).

⁷ I discuss the papers that focus strictly on refugees rather than on other types of immigrants. This is so because refugees are special and differ considerably from other types of immigrants such as economic immigrants. Refugees do not plan to move, and they do not acquire linguistic capabilities in advance. Several papers examine the causal role of language training for the labor market integration among other types of immigrants (see Sarvimäki and Hämäläinen 2016; Lochmann et al., 2019; Lang, 2022; Heller and Mumma, 2023; Pont-Grau et al., 2023).

⁸ Epidemiological and medical studies have documented the high prevalence of mental health disorders (Fazel et al., 2005; Lindert et al., 2009; Bryant et al., 2023), physical health problems (Hadgkiss and Renzaho, 2014), infectious diseases (Clark and Mytton, 2007; Eiset and Wejse, 2017), and non-communicable diseases (Yun et al., 2012; WHO, 2019) among refugees. Relative to other types of immigrants, refugees have the worst health status (see Lindert et al., 2009; Lebao et al., 2020 for reviews from the public health and social medicine literature). For example, the prevalence rates of depression and anxiety among refugees are 44% and 40%, respectively, while the corresponding figures for economic immigrants are 20% and 21% (Lindert et al., 2009).

immigrants (Ruiz and Vargas-Silva, 2018; Fasani et al., 2022),⁹ promoting self-sufficiency through work in the local labor market is the foremost objective of all integration policies. This is so not only for fiscal reasons but also because of the need to improve the economic well-being of refugees and their families (Clemens et al., 2018; Marbach et al., 2018). My findings of positive effects on the labor market and economic outcomes are consistent with those of recent studies that have shown that language training has large effects on employment, earnings, and job quality (Foged et al., 2022; Foged and van der Werf, 2023).¹⁰

Second, I contribute to a sparse but important strand of the literature that documents the *spillover* effects of language training on the outcomes of refugee children. The only paper that provides causal evidence on this matter is by Foged et al. (2023), who showed that language training for refugee parents improves children's schooling outcomes while reducing children's involvement in illegal activities during adolescence. My paper is the first to examine the health of refugees' children early in life. Given that health at birth has significant impacts on long-term human capital and life cycle achievements (see, for example, Smith, 2009; Fletcher et al., 2010; Currie, 2011), the improvement in health at birth induced by language training that parents receive plays a role as an important mediating channel between that training and children's longer-term achievements, such as school performance and social behaviors during adolescence, as documented by Foged et al. (2023). Ignoring spillover effects would therefore result in an underestimation of the benefits of language training. This insight is particularly relevant to the numerous countries worldwide that have received refugees and asylum seekers. Overall, my paper is the first comprehensive study of the multigenerational causal impacts of a universal language training program for refugees.

2. Institutional background

2.1. Norway's introduction program and language training reform

Norwegian Introduction Program. With a longstanding tradition of receiving refugees, Norway has implemented various policies to promote refugees' economic and social integration (Bratsberg et al., 2016; Hardoy and Zhang, 2019). In the late 1990s, integration courses were introduced for refugees and their family members to enhance their labor market outcomes and broader social integration. However, these courses were highly decentralized, with municipalities independently determining implementation

without consistent national guidelines or a legal framework. Consequently, their impact was limited, as many participants remained reliant on welfare rather than entering the labor market.

To address this policy shortfall, the Norwegian government launched the Norwegian Introduction Program (NIP) in the early 2000s. The NIP is a training program designed to equip refugees and their family members with the skills and knowledge necessary for successful integration into the Norwegian workforce or educational system. The implementation of the NIP is the responsibility of municipalities, which follow the general guidelines of the government. Within three months of the first date of settlement, refugees are asked to participate in the NIP at the municipality they have settled. Online Appendix Figure A.1 shows the evolution of the NIP with key changes during the 2000s. The NIP was formally introduced on September 1, 2003 in some municipalities after several years of policy pilots and since then gradually rolled out across other Norwegian municipalities (Røed et al., 2019; Ferwerda and

Finseraas, 2024). On September 1, 2004, the NIP became mandatory for all municipalities, marking a time for universal implementation of the NIP across the country. The NIP is legally regulated by the Introduction Act.¹¹

In terms of the program's objectives, an average participant is expected to possess basic Norwegian language skills, and knowledge about Norwegian society and institutions, as well as to be prepared to obtain remunerative employment after completing the NIP. To obtain these objectives, the program is structured with three main components: (i) Norwegian language training, (ii) social studies, and

(iii) measures that prepare the participants for further education or labor market participation. A certificate is issued upon the completion of the program.

While the NIP is theoretically offered to all newly arrived foreign nationals, it chiefly targets refugees and their family members who are vulnerable due to lack employment skills and social interactions.¹² The Introduction Act provides a rigorous legal base for ensuring this targeting group's compliance of the program participation. In particular, refugees and their family members who are aged between 18 and 55 have the *right* and *obligation* to participate in a free-of-charge NIP course within two years¹³ after they have been settled in a municipality. By law, *right* refers to the obligation of municipalities to provide integration courses to all refugees who meet the established criteria while *obligation* denotes that individuals who choose not to participate in the program forfeit eligible welfare schemes. An exemption from participation in the NIP is only granted to those who have proven that they possess sufficient knowledge and skills to fully satisfy its requirements and to those with severe illnesses or disabilities. Also refugees who independently secure employment or education and can support themselves financially are exempt from participating in the program without facing any sanctions.

Furthermore, to make the program participation effectively compulsory, financial incentives were designed at the individual level to encourage full attendance (Hagelund, 2005; Djuve and Kavli, 2019).

Refugees and their family members enrolled in the NIP are entitled to an introduction benefit amounting to twice the basic amount in the National Insurance scheme (2G), in particular approximately NOK 121,398 or about EUR 15,000 in 2005 (Røed et al., 2019).¹⁴ However, this benefit is subject to reduction or withdrawal in instances of irregular absences. Missing classes without a legitimate reason resulted in a reduction of benefits. With limited alternative income opportunities upon arrival in Norway, refugees and their families were therefore incentivized to participate in the program (Hagelund and Kavli, 2009). These legal and financial incentives are expected to ensure high compliance with NIP participation among refugees and their family members. Statistics which is based on both the survey and administrative data actually confirm this expectation. An estimate from Djuve et al. (2017) indicates relatively low absenteeism rates, approximately 10% or less, primarily attributed to individuals with severe health challenges or those already possessing full language competence. A comprehensive policy report on the NIP carried out by Røed et al. (2019) draws on the registry data to show detailed summary statistics for refugees' participation rate of the IP over a period between 2000 and 2015. In particular, its Fig. 1.2 highlights a change in the participation rate around 2004, when the universal compulsory implementation of the NIP came into effect. Prior to this year, participation rates among

¹¹ See <https://lovdata.no/dokument/NLO/lov/2003-07-04-80> for the original Introduction Act.

¹² The citizens of the Nordic countries and the members of the European Economic Area are not subject to the Introduction Act.

¹³ In some cases and for special reasons, the program can be completed within three years.

¹⁴ See <https://www.skatteetaten.no/en/rates/national-insurance-scheme-basic-amount/> for a summary of the Norwegian National Insurance scheme basic amount overtime.

⁹ The estimates by Fasani et al. (2022) reveal that, in European countries, refugees are 11.6% less likely to be employed and 22% more likely to be unemployed than other immigrants of similar backgrounds.

¹⁰ Language similarity between the original and host countries also generates considerable economic benefits for refugees (for example, see Schmid, 2023) as well as the health of refugee children (for example, see Auer and Kunz, 2024).

refugees were relatively low, remaining below 20% until around 2002. However, between 2002 and 2004, a gradual yet notable surge occurred, with participation reaching nearly 90% by 2004. This steady rise likely reflects the staggered implementation of mandatory NIP participation across municipalities during this period. Following the formal introduction of the mandatory NIP on September 1, 2004, the participation rate for refugees upon arrival has stabilized at consistently high levels, fluctuating slightly but remaining above 90% throughout the rest of the observed period. This near-universal participation since 2004 underscores the effectiveness of the obligatory policy in ensuring the program engagement among refugees.

The 2005 reform. To further strengthen the program, a key reform was introduced on September 1, 2005 as shown in Figure A.1. This reform focused on the language training component of the NIP by establishing a standardized requirement of minimum 300 h of mandatory language training, emphasizing the importance of linguistic proficiency as a cornerstone of successful integration into Norwegian society. The reform was the first to establish the *standard* length of language training that refugees receive which was not never before. The reform particularly set 250 h up for the language class and remaining 50 h for social knowledge training.¹⁵

The general objective of the reform package is to help immigrants obtain a level of proficiency in the Norwegian language that enables them to utilize and develop their competence in education, work, and social life as well as to equip immigrants with general knowledge about Norwegian society (Djuve et al., 2017). The government introduced a national curriculum that was applied to all the training

centers across the country. The organization of Norwegian language courses are based on the Common European Framework of Reference for Languages (CEFR).¹⁶ Several courses are designed to help individuals achieve the B2 level under the CEFR. Social knowledge training covers seven key topics: immigrants in Norway; history, geography, and way of life; children and family; health; education and skills; working life; and democracy and the welfare society.

It is important to note that the primary change introduced by the 2005 reform was the establishment of a standard minimum duration for language training. Although Norwegian language training for refugees had been a mandatory component of the NIP since September 1, 2004, its duration was notably limited prior to the reform (Røed et al., 2019). This limitation arose from the absence of specific hour requirements, with participation in language training classes being largely voluntary across municipalities (Røed et al., 2019). While formal statistics on language training duration before the 2005 reform are unavailable, observational studies based on interviews and survey data suggest that sessions typically involved only a few hours per week over several weeks (Djuve, 2011a). In contrast, the reform enforced a minimum standard of 300 h, representing a significant increase in the time allocated to refugees' language training (Djuve, 2011a, 2011b).

The 2005 reform did not introduce changes to other regulations regarding eligibility, participation enforcement, or financial allowances, all of which had been in effect since September 1, 2004 (Joyce, 2017; Røed et al., 2019). Notably, the post-reform attendance rate remained consistently high at approximately 90%, similar to the level observed prior to September 2005, as discussed above.

2.2. Norwegian primary care and maternity care

Norwegian primary care. Primary care serves as the cornerstone of the Norwegian healthcare system, working in tandem with specialist care to form the foundation of the country's universal healthcare

¹⁵ See <https://lovdata.no/dokument/SFO/forskrift/2005-09-16-1055> for details about the curriculum.

¹⁶ See <https://www.coe.int/en/web/common-european-framework-reference-languages> for details about the CEFR.

service. This system is predominantly state-funded, with municipalities playing a central role in delivering primary care services to residents. While public funding covers the majority of examination and treatment costs, patients are required to pay user fees for general practitioner (GP) visits. These fees vary depending on the service but are capped annually by law to ensure affordability, for example about NOK 3,040 (EUR 268) in 2023.¹⁷

Primary care is largely based on the services that GPs provide.¹⁸ Under the GP scheme introduced in 2001, all residents registered with the National Population Register are entitled to a GP, who provides general health consultations, certifies sick leave, administers treatments, and acts as a gatekeeper to

specialist care by making referrals for serious health issues and providing follow-up care after specialist treatment.

Refugees access primary care services primarily through GPs, with each assigned a GP in their municipality. Despite equitable access, primary care utilization among refugees in Norway is low, possibly due to language barriers and cultural differences, highlighting the need for further research (Laue and Risør, 2017).

Norwegian maternity care. Pregnant women enrolled in the Norwegian Insurance Scheme receive free-of-charge healthcare services through the Norwegian Health Care Service.¹⁹ Regular pregnancy health checkups are a key part of antenatal care in Norway, with the first checkup ideally occurring in weeks 6–12 and an ultrasound offered around week 18. Pregnant women can choose to receive care from a GP, a midwife at a Maternity and Child Health Care Center, a private midwife, or a gynecologist, with service utilization left to their discretion.

Pregnant women in Norway can choose their birth location based on advice from their GP and midwife. Women with health complications during pregnancy or expected delivery issues are advised to give birth at a well-resourced hospital, while midwife-led units, hospital maternity wards, and specialist clinics are available for those without complications. Planned home births are rare and not provided by the Norwegian Health Care Service; however, healthy women with uncomplicated pregnancies may receive support from a private midwife if they choose a home birth.

3. Data

3.1. Data sources and sample selection

My analysis draws on several high-quality administrative registers that were provided by Statistics Norway. I particularly use the Central Population Register (CPR), the Income and Tax Register, the Primary Care Register, the Residence Register, the Marriage Register, and the Birth Register. These registers cover the entire population of Norway during the available period.²⁰ Importantly, each individual has a unique personal identifier which is used to link him across these registers and allowed me to construct various variables for the analysis.

I first used the CPR to construct variables of refugee status, the family members of a refugee, and the date of arrival in Norway. Refugee status

¹⁷ See <https://www.helsenorge.no/betaling-for-helsetjenester/betaling-hoslege/> for details about user fees for GP visits.

¹⁸ Primary care also includes services provided by other institutions such as emergency rooms, healthcare centers, out-of-hours medical services, nursing homes, home nursing, substance abuse programs, physiotherapists, occupational and speech therapists, chiropractors, wellness centers, and dentists.

¹⁹ See <https://www.helsenorge.no/en/pregnancy-and-maternity-care-in-norway/> for specific information about pregnancy and giving birth in Norway.

²⁰ The different registers cover different periods: the Central Population and the Birth Registers cover the period between 1967 and 2018, the Tax and Income Register covers the period between 1993 and 2017, the Primary Care Register covers the period between 2006 and 2019, and the Residence and Marriage Registers cover the period between 1975 and 2018.

is an indicator of immigration type, and I use it to differentiate between refugees and economic, family, and educational immigrants. I also constructed an indicator for family members, which indicates whether they are the wife, husband, or child of a refugee. The arrival date variable is based on information about one's first date of residence in Norway recorded in the CPR. I relied on these variables to construct the samples for the analysis. To investigate

first-generation impacts, I included individuals who meet three requirements in the overall sample of adult refugees. The requirements are that they (i) were refugees or family members thereof,²¹ (ii) were aged between 18 and 55 at arrival, and (iii) had an arrival date that fell between September 2004 and August 2006. The observations therefore fall within a bandwidth of ± 360 days around the cutoff date (September 1, 2005). The first and the second requirements are based on the criteria for eligibility for the IP under the reform package. The third requirement means that refugees in the sample are comparable in terms of their pre-determined characteristics, except for exposure to the reform. These restrictions resulted in 5,581 observations. Next, I eliminated 616 observations because the refugees in question had emigrated from Norway during the study period (the rate of attrition due to emigration is about 11%).²² Thus, the final sample covers 4,965 adult refugees as shown in panel (a) of Table 1.²³ There are 2,006 treated observations (40.4%).

To examine the reform effects on the early life health of the children's generation, I restrict the set of children who were born between 2005 and 2018 in Norway to the refugees²⁴ in the sample of adult refugees. I obtained a total sample of 4,892 children as shown in panel (b) of Table 1. The number of treated children is 1,911 (39.1%).

3.2. Construction of variables and summary statistics

Drawing the available registers, I constructed various sets of variables for the analysis. To examine the direct impacts of the reform on the healthcare integration of refugees, I constructed several measures of the use of primary care services for the main effects and some economic outcome variables for the potential channels. To estimate the inter-generational spillover effects of the reform, I constructed health at birth outcome variables of the children's generation for the main effects as well several measures for prenatal conditions, which I treat as the potential transmission mechanisms. I also created variables that are related to the refugees' pre-determined characteristics.

Primary care services. I exploited the information about individual consultations with a GP that is recorded in the Norwegian Control and Distribution of Health Reimbursement Database Register (KUHR for an acronym) to construct several measures of the utilization of primary care services. The key measure is the accumulated number of GP consultations over the 12 years after arrival. Using the second version of the International Classification of Primary Care (ICPC-2) to code specific types of consultations,²⁵ I constructed similar measures for the three most prevalent types of consultations,

namely (i) musculoskeletal consultations (L codes), (ii) psychological consultations (P codes), and (iii) digestive consultations (D codes). Panel (a) in Table 1 shows that a typical refugee attends nearly 54

consultations over the 12 years that follow the arrival year in Norway. The largest number of visits is for musculoskeletal consultations (14.9), which are followed by psychological (5.2), and digestive ones (4.8).

Economic outcomes. To understand the key mechanisms through which the healthcare integration effects run, I tracked the refugees across the administrative registers and collected related information to construct several proxy measures for labor market attachment and economic success. I exploited the Income and Tax Register to construct three outcome variables: labor market earnings, sick leave benefits, and disposable income. While I use labor market earnings and sick leave benefits as proxy measures for refugees' labor market attachment, I use disposable income as a proxy for refugees' economic well-being. Labor market earnings are defined as the sum of wage income and work-related payments such as payments for sick and paternity benefits. The sick leave benefit is the monetary amount one receives for paid leave due to illnesses. Disposable income is defined as the sum of wage income, work-related payments, capital income, pension income, and benefits income minus deductibles such as taxes. These variables are both measured over 11 years after arrival (in Norwegian Kroners (NOK) 10,000 in 2015 values).^{26,27} On average, a refugee obtained approximately NOK 1.95 million (EUR 214,500) in labor market earnings, NOK 96,000 (EUR 10,500) in sick leave benefits, and NOK

1.72 million (EUR 189,200) in disposable income, as reported in panel (a) of Table 1.

Children's health at birth. To study the effects of language training on the next generation, I relied on the Medical Birth Register of Norway (MBRN) to construct three main outcome variables for the health at birth of the children who were born in Norway to adult refugees in the sample. The outcomes in question are particularly preterm birth, low birth weight, and short birth length. *Preterm birth* is a dummy for having a gestational age of less than 37 weeks. *Low birth weight* is a dummy for having a birth weight of less than 2,500 g. *Short birth length* is a dummy for having a birth height below 48 cm. Panel (b) in Table 1 indicates that 5.6%, 4.9%, and 13.1% of the children of refugees were born prematurely, had a low birth weight, and had a low birth height, respectively.

Prenatal conditions. In order to examine the potential mechanisms behind the impacts of parental language training on refugee children's health at birth, I combined various registers to construct a set of proxy measures for prenatal conditions in the home environment and maternal care during pregnancy. As far as the home environment is concerned, I focus on the parent's income, and marriage. *Income* is measured by the average of the first parent's disposable income in the year before and in the birth year of the child (NOK 10,000 in 2015 values). The information about disposable income is from the Income and Tax Register. I consulted the Marriage Register to construct *marriage*, which is a dummy

for the parent being married during the year of the child's birth. The average refugee parent has an income of approximately NOK 146,500 (EUR 16,000) while their child was *in utero*. The share of married parents is nearly 57.3%.

Turning to maternal care during pregnancy, I focus on three variables: primary care consultations, prenatal checkups, and the use of prenatal supplements. I first drew on the KUHR Register to construct two variables: the use of *primary care consultations* and *prenatal checkups*. These variables are both measured by the total number of GP consultations over two years, the year before and the year of the child's birth. I also used the MBRN to construct *supplement take-ups*, which indicates whether the mother took any supplements, such as multivitamins and folic acid, during pregnancy. On average, the mothers in the sample had 10.8 primary care consultations and 4.6 prenatal checkups during pregnancy. The proportion of mothers who took any supplements is

²¹ The sample does not include asylum-seekers.

²² Online Appendix Table A.1 shows that there are no systematic differences in the attrition rate as well as in the pre-determined characteristics of those who leave between the treatment and the control group. Therefore, the results are unlikely to have been driven by attrition.

²³ It is important to note that this is an *overall* sample that is obtained using a bandwidth of ± 360 days around the cutoff date. A sub-sample for the RD regression was generated by using optimal bandwidth, as suggested by Calonico et al. (2019).

²⁴ The treatment of children born to two parents in the sample of adult refugees is defined by the parent who first arrived in Norway.

²⁵ See <https://www.ehelse.no/kodeverk-og-terminologi/ICPC-2/icpc-2e-english-version> for details about ICPC-2.

²⁶ The consumer price index data for deflating monetary values to 2015 was obtained from Statistics Norway, see <https://www.ssb.no/en/statbank/list/kpi>.

²⁷ For clarity of exposition, in some cases, I present measurements in euros by using the 2015 exchange rate of 0.11 NOK/EUR.

Table 2
Balance tests.

	(1)	(2)	(3)	(4)	(5)
	Total sample		RDD estimate		
	Mean	SD	β	SE	p-value
Male	0.508	0.500	-0.009	0.057	0.881
Age	31.514	8.562	-1.482	1.225	0.227
Married	0.571	0.495	0.008	0.070	0.907
Children	0.555	0.497	0.082	0.071	0.251
Refugee	0.903	0.296	0.045	0.049	0.355
Oslo and Akershus	0.187	0.390	-0.003	0.057	0.952
South Eastern	0.174	0.379	-0.048	0.049	0.337
Agder and Rogaland	0.177	0.381	0.011	0.044	0.807
Russia	0.129	0.336	-0.048	0.037	0.201
Myanmar	0.124	0.330	0.003	0.030	0.910
Somali	0.109	0.312	0.044	0.049	0.371
Afghanistan	0.089	0.285	-0.069	0.049	0.161
HDI	0.548	0.136	0.006	0.018	0.755
Children: Girl	0.493	0.500	0.005	0.048	0.911

Notes: The coefficient in column 3 is the β coefficient for RDD estimate in Equation (1). The estimation sample is based within the optimal bandwidth which is calculated using the triangular kernel and the procedure proposed by Calonico et al. (2019). Column 4 shows robust standard errors.

Data source: Own calculations based on the data provided by Statistics Norway.

about 41.5%.

Pre-determined characteristics. I combined different registers to construct an additional set of variables for refugees' pre-determined characteristics, which were measured in the year of arrival. I used these variables to run relevant tests or checks as well as to establish a set of control variables that are added to the estimation model. For the sample of adult refugees, the pre-determined variables consist of age, indicators of gender (male) and marital status (married), an indicator of having children, indicators of the refugees' regions of initial settlement in Norway (Oslo and Akershus, the South Eastern region, Agder and Rogaland), and indicators of several major countries of origin (Russia, Myanmar, Somalia, and Afghanistan). Turning to the children's sample, I included indicators of the child's gender (girl) and birth cohorts. Moreover, I used a continuous score for the Human Development Index (HDI), measured at the country level, as a proxy for the overall characteristics of the refugees' countries of origin.²⁸ Columns 1–2 of Table 2 display the summary statistics for these pre-determined characteristics. Approximately 50.8% of the refugees were male, and the average age at arrival was 31.5. The shares of married refugees of those who had at least one child at arrival were 57.1% and 55.5%, respectively. In the sample of adults, refugees account for 90.3% of observations, while the percentage of family members is only 9.7%. As far as the regions of first settlement are concerned, Oslo and Akershus account for 18.7% of the total. They are followed by Agder and Rogaland (17.7%) and by the South Eastern region (17.4%). The major countries of origin in the sample are Russia (12.9%), Myanmar (12.4%), Somalia (10.9%), and Afghanistan (8.9%). The mean HDI score is 0.55, indicating a low level of development in the countries of origin of the refugees. In the children's sample, the share of girls is about 49.3%.

²⁸ The HDI data is obtained from <https://hdr.undp.org/data-center/human-development-index/#/indicies/HDI>. The HDI is a composite index measuring the achievements of a country in three key dimensions of human development: health, education, and living standards. The HDI score ranges from 0 to 1 with a higher value indicating a higher level of development. The level of development based on the HDI score is classified as low (<0.550), medium (0.550–0.699), high (0.700–0.799), and very high (≥ 0.800). To construct a stable HDI, I took an average score for the 2003–2007 period.

4. Empirical strategy

4.1. Identification and estimation

Participation in the 300-h language training program is mandatory for refugees who arrived in Norway on or after September 1, 2005. I use this breakpoint in an RD design to estimate the causal effect of exposure to the reform on integration outcomes of interest. In particular, I compare refugees who arrived on or just after the cutoff date (the treatment group) to those who arrived just before the cutoff date (the comparison group). The key identifying assumption is that treated and untreated refugees do not exhibit any systematic differences other than exposure to the reform. As long as this assumption holds, the strategy serves as an assignment of exposure to language training that is as good as random.

My baseline RD specification estimates a regression of the following type:

$$Y_i = \alpha + \beta D_i + \gamma(t_i - T) + \delta D_i \times (t_i - T) + \epsilon_i \quad (1)$$

where Y_i is the outcome of interest for individual i . D_i is the treatment indicator, and it is defined as

$D_i = 1\{t_i \geq T\}$ where t_i and T are the arrival date and the cutoff date, respectively. The difference measured in days between these two days ($t_i - T$) is the running variable. The inclusion of $(t_i - T)$ and $D_i \times (t_i - T)$ allows the model to account for the linear dependence of the outcome on the distance between the arrival date and the reform date, as suggested by Gelman and Imbens (2019). I include no control variables for pre-determined characteristics in the baseline estimation. If such control variables are orthogonal to the treatment, they are “bad” controls because they would make standard errors bigger without changing the magnitudes of the coefficients. The coefficient β is the RD estimate, and it indicates the causal effect of exposure to language training on Y_i . The intercept α captures the pre-reform mean at the boundary point. I use it to re-scale the estimated coefficients into percentage effects.

In the regression model, I assign higher weights to the observations that are closer to the cutoff point by using the weighted least squares with a triangular kernel. I follow the procedure suggested by Calonico et al. (2019) with a triangular kernel to select optimal bandwidths, which results in different bandwidths for different dependent variables. The sample sizes vary across regressions due to corresponding variations in mean squared error (MSE) optimal bandwidths.²⁹ I use heteroskedasticity-robust standard errors in these regressions, as recommended by Lee and Lemieux (2010).

I interpret the parameter of interest β as the causal impact of exposure to language training rather than of the actual take-up. As described in Section 2.1, the reform package generates strong incentives for participation that are both legal (i.e., compulsory attendance as regulated by the Introduction Act) and economic (i.e., large financial benefits are made conditional on attendance). If the rate of policy compliance were in principle 100%, the β coefficient would be treated as the average treatment effect.

(ATE). However, despite the actual rate of participation being relatively high up to 90% (Røed et al., 2019), some refugees and their family members had severe illnesses or disability, which means that they were exempted from the attendance requirement. For that reason, I interpret the β coefficient as an intention-to-treat (ITT) estimate to arrive at a conservative interpretation. It is also important to note that I do not estimate the treatment effect of take-up or actual language proficiency due to the unavailability of data for the first-stage estimation. Given the rigorous organization of the program as well as the high participation rate, it is reasonable to expect that exposure to the program may have had a considerable impact on refugees' integration.

²⁹ The sample sizes are around 1,000 observations.

To evaluate the robustness of the baseline specification, I conduct several robustness checks. First, I replace the linear function of the running variable from the baseline specification with the quadratic polynomial function. [Pei et al. \(2022\)](#) argued that the choice of an optimal polynomial should depend on the underlying data-generating process; therefore, higher-order polynomials should be used in some circumstances. Second, I add a set of control variables to the specification. The control variables include several pre-determined characteristics at the time of arrival: age, age squared, indicators for gender (male) and marital status (married), an indicator of whether the refugee has children, indicators for the regions of initial settlement in Norway (Oslo and Akershus, the South Eastern region, Agder and Rogaland), and indicators for several major countries of origin (Russia, Myanmar, Somalia, and Afghanistan). The additional controls for the child's characteristics, which pertain to the intergenerational effects, include a dummy for gender (boy) and birth cohort fixed effects. Third, I exclude the weights for the distance between the arrival date and the cutoff date from the baseline regressions. Fourth, I use a uniform kernel to select the optimal bandwidth rather than a triangular kernel as in the baseline specification. In order to conduct robustness checks for the effects on children's health at birth, I also run regressions in which I control for the child's order of birth, relative to either all of their other siblings or only the siblings who were born in Norway. The results from these exercises confirm the robustness of the baseline estimates.

4.2. Assessing the validity of the identification strategy

The validity of my identification strategy depends on the strong assumption that no manipulations affected the timing of refugees' arrivals around September 1, 2005. In Norway, such manipulations are theoretically impossible for both Norwegian immigration authorities and asylum seekers.

First, as the authority that processes asylum applications, the Norwegian Directorate of Immigration (UDI) follows a rigorous procedure that is based on law and independent of the reform. For that reason the possibility that the UDI manipulated the date of arrival of a group of asylum seekers around the time of the reform is remote ([Djuve, 2011b](#); [Djuve et al., 2017](#)). Furthermore, the processing of asylum seekers' documents by the UDI is independent of the role of municipalities in implementing the IP as well as the language training program, which also makes selection into the IP at the cutoff improbable.

Second, it is unlikely that the asylum seekers adjusted the dates on which they applied for asylum status to affect the dates on which they received which would change the probability of being exposed to the reform ([Ferwerda and Finseraas, 2024](#)). It typically takes between 10 and 16 months for an asylum application to be reviewed and processed. This long period makes it difficult for asylum seekers to predict policy changes and to adjust their behaviors strategically. Asylum seekers are unlikely to have known about the reform before submitting their asylum applications.

To check whether there was a manipulation and selection into exposure to the reform formally, I conduct several assessments. First, I check whether there were unusual changes in the number of refugees who were granted asylum around the cutoff date. [Fig. 1](#) shows the density of the observations within a window of ± 360 days, using 30-day bins, for the adult sample. The raw distribution of the number of observations in panel (a) demonstrates that the number of refugees was equally distributed for some months around September 2005. However, in the whole sample, the distribution is skewed towards the early months before the cutoff point. This pattern is likely driven by cyclical and seasonal trends in asylum admissions. To adjust for these trends, I calculate the demeaned share which is measured as the deviation of the share of the observations for a month in my sample from the average share for that month over five years between 2003 and 2007. The graph in panel (b) indicates that the demeaned shares are relatively small and tend to fluctuate around zero for both the periods before and after the

cutoff point. This pattern suggests that there are no systematic differences around the cutoff date and thus no evidence of manipulation into the treatment status.

Second, I perform balance tests to check whether refugees' other individual characteristics, apart from the treatment status of being exposed to the reform, evolve smoothly around September 1, 2005. I estimate the baseline specification Equation (1) for the pre-determined covariates on arrival including gender (male), age, marital status (married), having at least one child, refugee status, region of first settlement (Oslo and Akershus, the South Eastern region, and Agder and Rogaland), country of origin (Russia, Myanmar, Somalia, and Afghanistan), HDI score, and gender (girl) for the children's sample. The RD estimates that are presented in columns (3) and (4) of [Table 2](#) show that exposure to the reform does not influence these pre-determined covariates because p -values in column (4) are all statistically insignificant at the 10% level. [Fig. 2](#) further demonstrates graphically that these covariates fluctuate smoothly across the cutoff point.

I arrive at similar results for the parents' characteristics when I examine the children's sample, as shown in Online Appendix [Figure A.2](#) for the density checks and in [Figure A.3](#) for the balance tests. Furthermore, using the raw samples that are potentially to some extent driven by seasonal factors I find no evidence of strategic manipulation in the running variable at the cutoff as presented in Online Appendix [Figure A.4](#), which demonstrates the results of density tests following the procedure proposed by [McCrary \(2008\)](#). These tests are conducted on both directly affected refugees and refugee children. Both figures confirm that the running variable's density remains continuous at the cutoff across all the samples, again supporting the validity of my identifying assumption. In sum, the absence of unusual

changes in the number of observations around the cutoff date and the balancing of the pre-determined traits support the assumption that there were no manipulations for selection into the treatment and that the *only* difference between the treated and the untreated individuals was exposure to language training.

To assess the validity of the identification strategy further, I do several placebo tests in which I estimate the effects on the outcomes of interest by using artificial reforms as well as an unaffected group. The results in Section 7.3 also support the argument that my empirical strategy is valid.

Finally, regarding the concern that the language training program may have influenced the selection into treatment by altering the pool of asylum applications around the cutoff date, several reasons suggest this is unlikely. The decision to seek asylum in Norway, or any other country, is primarily influenced by factors beyond the individual's control, such as geopolitical conditions in their country of origin, domestic asylum policies and priorities of host nations, the timing of applications, access to legal assistance, and the completeness of documentation. These determinants far outweigh the impact of specific integration policies like language training reforms. Asylum seekers often make decisions under extreme duress, with their primary focus on immediate resettlement in a safe country. Under such circumstances, it is implausible that detailed integration policies, such as language training programs, play a significant role in their choice of destination. Furthermore, language training reforms are only one element of broader refugee policy frameworks. As such, Norway's language training program was a minor factor and unlikely to affect its overall attractiveness as a host country. Considering these points, it is highly improbable that the introduction of the language training program altered the composition or volume of asylum applications to Norway around the cutoff date.

5. Direct effects on the integration of refugees into the healthcare system

Integration into the healthcare system of a host country is of crucial importance for integration policies on refugees. In this section, I present the empirical results for the direct reform effects on adult refugees' long-

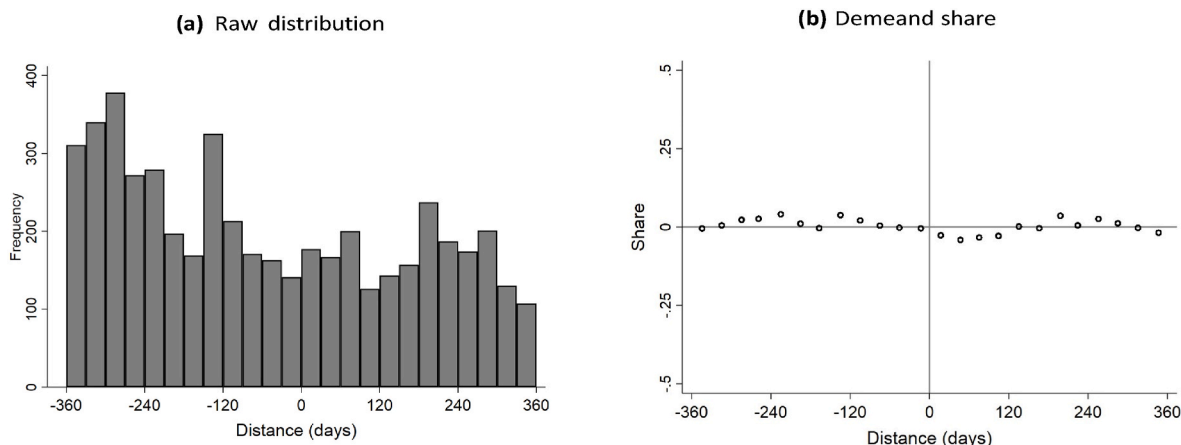


Fig. 1. Density.
Notes: This figure graphically illustrates the density of the observations in the adult sample across the distance between the arrival date and September 1, 2005 for the directly affected individual sample. Each graph plots the mean value using 30-day bins for the distance.
Data source: Statistics Norway.

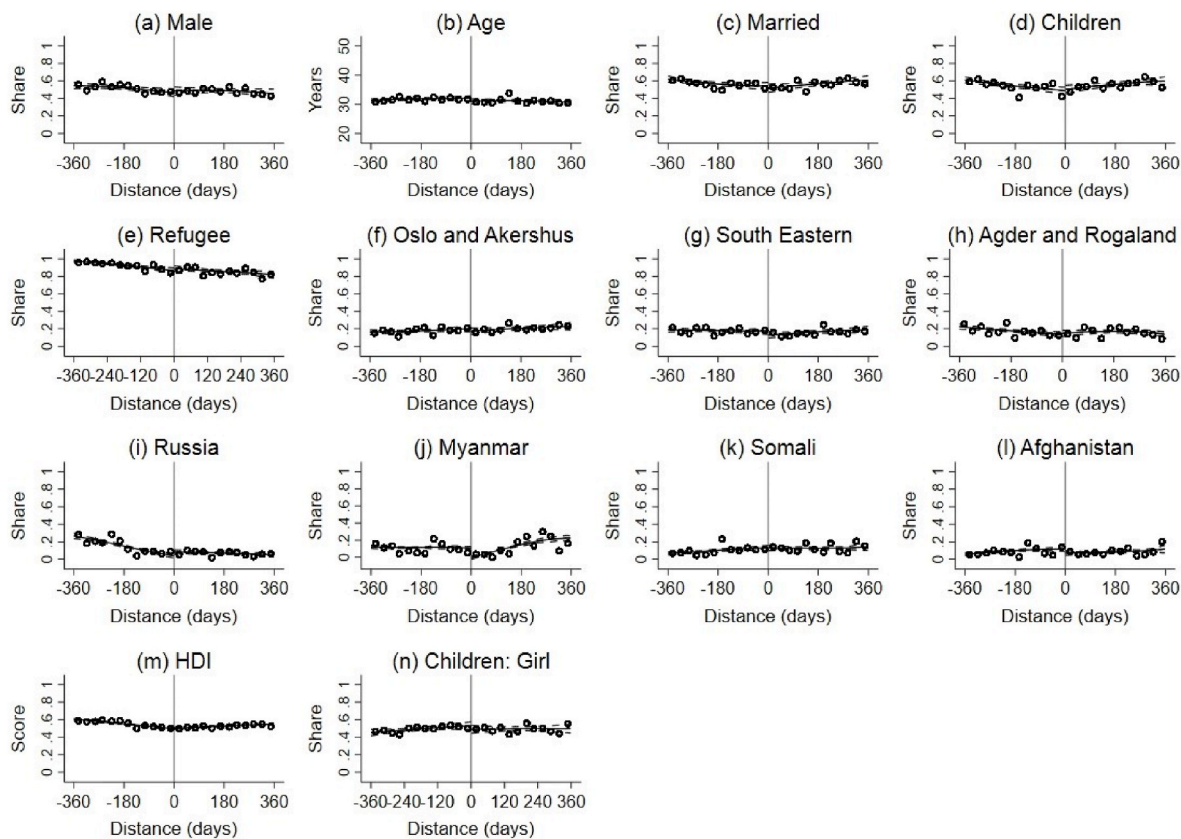


Fig. 2. Balancing of pre-determined covariates.
Notes: This figure graphically illustrates the smoothness of pre-determined characteristics across the distance between the arrival date and September 1, 2005 for the directly affected individual sample (panel a–m) and of the children’s sample (panel n). Each graph plots the mean value using 30-day bins for the distance.
Data source: Statistics Norway.

term use of primary care services. I show the baseline estimates before the results for the robustness checks. To examine the key mechanisms behind the main reform impacts, I concentrate on the economic effect of language training.

5.1. Main effects on the use of primary care services

Before presenting the estimation results, I discuss the RD graphs that capture the use of primary care services to enable the reader to observe

whether there is a discontinuity at the cutoff of the running variable. In the graphs in Fig. 3, I plot the accumulated number of doctor visits used over the 12 years after arrival for all consultations (graph a), musculoskeletal consultations (graph b), psychological consultations (graph c), and digestive consultations (graph d). I observe an obvious discontinuity at the cutoff point for all these outcomes. Refugees and their family members marginally to the right of

the cutoff (the treated ones) visited their GPs more frequently than those marginally to the left of the cutoff (the untreated ones).

To confirm whether the discontinuities in these graphs are statistically significant, Table 3 reports the baseline estimates of the reform effects on these outcomes. The estimate for overall utilization which is displayed in column (1) indicates that the reform increased the total number of GP visits by.

19.2. The effect is strongly statistically significant at the 1% level. Relative to the pre-reform mean, this effect corresponds to an increase of about 44%.

The remaining columns in Table 3 present the baseline estimates of the effects on specific types of GP consultations including musculoskeletal problems (column 2), psychological problems (column 3), and digestive problems (column 4). The treated individuals attend more consultations for such health problems than the untreated ones; this finding is similar to the result for the total number of consultations. The increases in the number of doctor visits that were induced by the language training program amount to 3.1, 2.9, and 2.2 visits for musculoskeletal, psychological, and digestive problems, respectively. The effects are statistically significant at the 10% level for musculoskeletal problems and at the 5% level for the other types of problems. Compared to the pre-reform means, these effects are equivalent to an increase of 29.6% for musculoskeletal consultations, an increase of 84.6% for psychiatric consultations, and an increase of 48.9% for digestive consultations.

It is important to clarify that I interpret the effect on GP visits as that on the use of primary care services. Assuming that two refugees have similar characteristics, including pre-determined health status, but except for exposure to the language training program, the treated one would have stronger contact with the local healthcare system by visiting their GP more frequently than the untreated one. My interpretation is therefore somewhat different from the common interpretation of the health status per se for native or general populations. The studies, which typically use the same KUHR data, treat a higher number of GP diagnoses as indicative of poorer health status (for example, see Bhuller et al., 2023; Butikofer et al., 2023 for mental disorder diagnoses).

The RD estimates for the robustness checks in Online Appendix Table A.3 show that the baseline RD estimates are strongly robust to estimating an alternative specification with the second-order polynomial function of the running variable (panel a), to adding a set of control variables (panel b), to excluding the regression weights (panel c), and to using the uniform kernel for the selection of optimal bandwidth (panel d). The results for the overall use of primary care services, which are presented in column (1), reveal that the estimated coefficients are all statistically significant in line with the baseline estimates in terms of both the direction and the magnitude. Almost all of the estimates for the other measures in columns (2)–(4) are statistically significant. The exceptions are the coefficients for musculoskeletal and digestive consultations in panel (c), which are marginally statistically insignificant although the effects are still positive and economically similar to the corresponding baseline estimates. On the whole, the results that were presented in this section provide evidence of the robust, strong, and positive effects of exposure to language training on refugees' overall use of primary care and on their

utilization of specific types of consultations. These effects suggest that the language training program improves refugees' integration into the Norwegian healthcare system. The effect on the utilization of psychological consultations is the largest. This finding likely reflects the fact that mental disorders are among the most severe health problems that

refugees face due to their terrible experiences during their journeys from their home countries to their destination countries.³⁰

5.2. Labor market attachment and economic well-being as potential mechanisms

Several mediating factors may link language training to the integration of refugees into the host country's healthcare system. First, language training improves refugees' knowledge of the healthcare system, particularly of the means of accessing primary care services, through, for example, improvements in their ability to read and understand the relevant materials in the host language. Second, language training likely improves refugees' efficiency in communicating with their GPs and thus encourages them to use those services more frequently. Third, language training may conduce labor market attachment and gains in economic power among refugees. While I cannot test the first two channels due to the data unavailability, I can provide empirical evidence of the existence of the third channel.

Labor market attachment. I use labor market earnings and sick leave benefits as two proxy measures for labor market attachment among refugees. The higher the earnings that a refugee obtains from the labor market, the deeper their attachment to their workplace. A workplace likely serves as a 'platform' on which refugees exchange information and acquire knowledge from natives and long-settled immigrant co-workers. That knowledge may include information about the primary care system of the host country as well as practical insights into doctor visits. In addition, higher labor market attachment offers refugees valuable opportunities to practice Norwegian in dynamic, real-world settings such as workplaces, where regular interactions with colleagues and clients expose them to diverse vocabulary and communication styles beyond formal classes. Improved language skills, in turn, may facilitate more frequent doctor visits by enhancing refugees' ability to understand the healthcare system and communicate their health problems effectively. In the Norwegian context, sick leave benefits may convey information about labor market attachment for two reasons. The first reason is that only those who have been employed over the last four weeks can claim sick leave benefits. Therefore, the amount of sick leave benefits that one receives is likely to be indicative of the intensity of labor market attachment. The second reason is that the existence of the health condition that results in the claim for sick leave must be certified by a GP. A refugee who has more sick leave benefits tends to visit their family doctor more frequently.

I first inquire whether there is a discontinuity at the cutoff for these outcomes. The RD graphs in Fig. 4 depict a clear discontinuity around the threshold for both labor market earnings (graph a) and sick leave benefits (graph b). When each of these outcomes moves from the left to the right of

the zero point on the horizontal axis of the running variable, there is an obvious jump at the cutoff. The baseline estimates that are reported in columns (1) and (2) of Table 4 indicate that the reform has a positive effect on refugees' labor market earnings (statistically significant at the 5% level) and sick leave benefits (statistically significant at the 1% level) accumulated 11 years after their arrival. Relative to the corresponding measures at the cutoff of the counterfactual individuals, the earnings amount of the treated ones is higher by NOK 199,110 (EUR 21,902), and sick leave benefits are higher by NOK 19,050 (EUR 2,096). In percentage terms, these increases correspond to an increase of 10.7% in earnings and an increase of 33.7% in sick leave benefits, respectively.

Economic well-being. Disposable income is used as a proxy measure for refugees' economic prosperity. As indicated in Section 2.2,

³⁰ The epidemiological and medical pieces of literature have documented the prevalence of mental health problems among refugees extensively (Fazel et al., 2005; Chey et al., 2007; Kirmayer et al., 2011; Silove et al., 2017; Bryant et al., 2023).

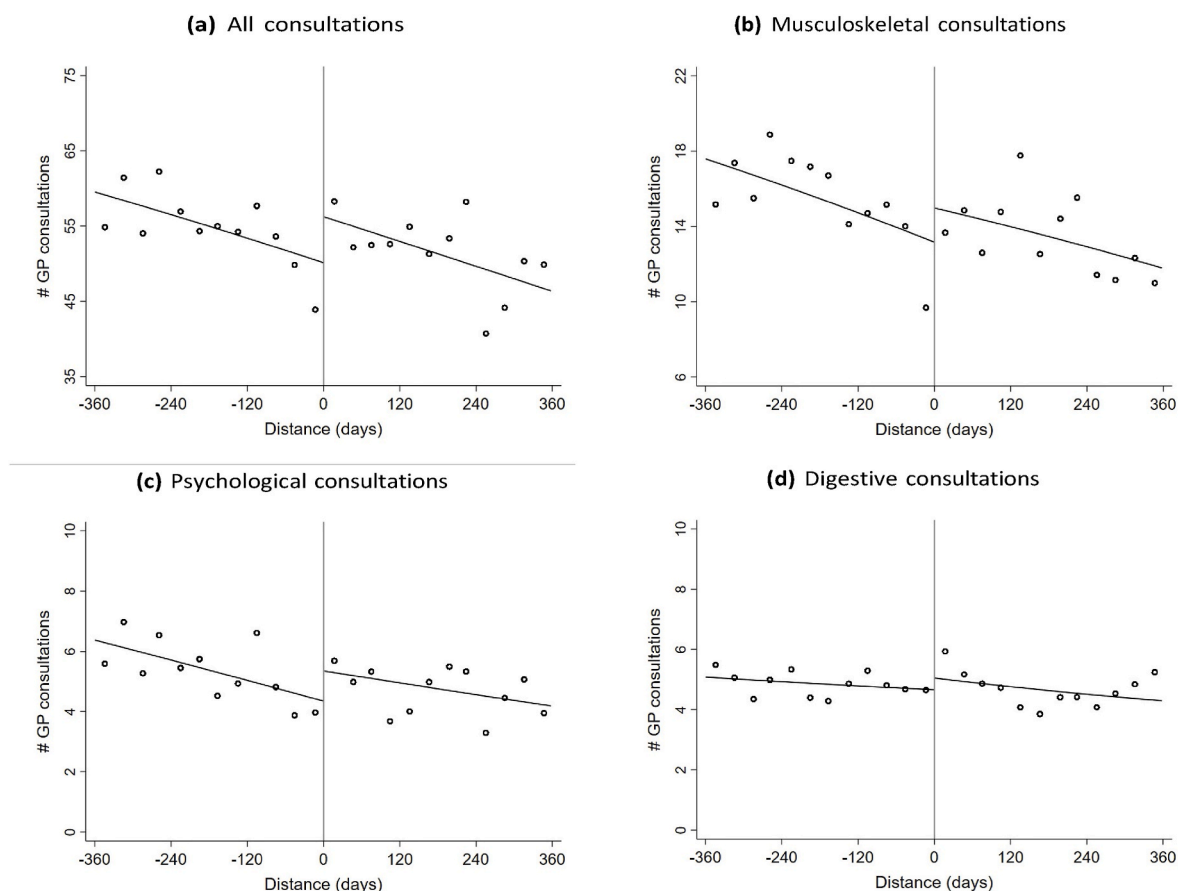


Fig. 3. RD graphs for refugees' use of primary care services.

Notes: This figure graphically illustrates the smoothness of primary care services utilization across the distance between the arrival date and September 1, 2005 for the directly affected individual sample. Each graph plots the mean value using 30-day bins for the distance.

Data source: Statistics Norway.

Table 3

Effect on refugees' use of primary care services.

	(1)	(2)	(3)	(4)
	All consultations	Musculoskeletal	Psychological	Digestive
Treated	19.179*** (6.587)	3.114* (1.695)	2.929** (1.252)	2.235** (1.098)
Pre-reform mean	43.561	10.512	3.462	4.570
Bandwidth	76.581	138.082	98.547	91.682
Observations	864	1,666	1,127	1,042
Mean of dep. var	51.007	13.900	4.807	5.012

Notes: The dependent variables are measured with a 12-year accumulation of the number of primary care consultations between year 2 and year 13 since arrival. Treated shows RDD estimate, which is the β coefficient in Equation (1). Pre-reform mean is the mean of the outcome for the untreated group measured at the cutoff, which is the α coefficient in Equation (1). Robust standard errors are in parentheses below coefficients. Bandwidth is calculated using the triangular kernel and the optimal bandwidth procedure proposed by Calonico et al. (2019). Observations and the mean of dependent variables are the estimation sample within the optimal bandwidth.

** $p < 0.01$, * $p < 0.05$, $p < 0.1$.

Data source: Own calculations based on the data provided by Statistics Norway.

patients must pay a user fee for using primary care services (around EUR 200 per year). For natives and skilled immigrants, this payment is small relative to their overall income. However, the fee is a significant financial barrier to access to primary care services among refugees, who

are the most economically disadvantaged group in society. For that reason, an improvement in economic prosperity, that is in disposable income, likely enables refugees to use more primary care services.

As is evident from graph (c) in Fig. 4, there is a jump in disposable income at the cutoff point. The RD estimate that is reported in column (3) of Table 4 also indicates a positive effect of exposure to language training on refugees' disposable income (statistically significant at the 5% level). The disposable income of the treated observations is higher than that of the untreated ones by approximately NOK 138,050 (EUR 15,186), corresponding to an 8.5-% increase.

The estimates from the robustness checks in Online Appendix Table A.4 are highly similar to the baseline estimates of all measures of labor market attachment and economic success when an alternative specification is estimated by using the second-order polynomial function of the running variable when control variables are added, when the regression weights are excluded, and when the uniform kernel is used to select optimal bandwidth. Therefore, one may be confident in the conclusion that language training plays an important role in refugees' integration into the local labor market as well as in their economic success, which likely eliminates a barrier to access to primary care services (Debesay et al., 2019).

5.3. Additional outcomes: Family formation and fertility

Deeper integration into Norwegian healthcare and labor markets may promote assimilation into the social and cultural norms of the host society. This process has the potential to shape preferences regarding family formation and childbearing practices, which often contrast

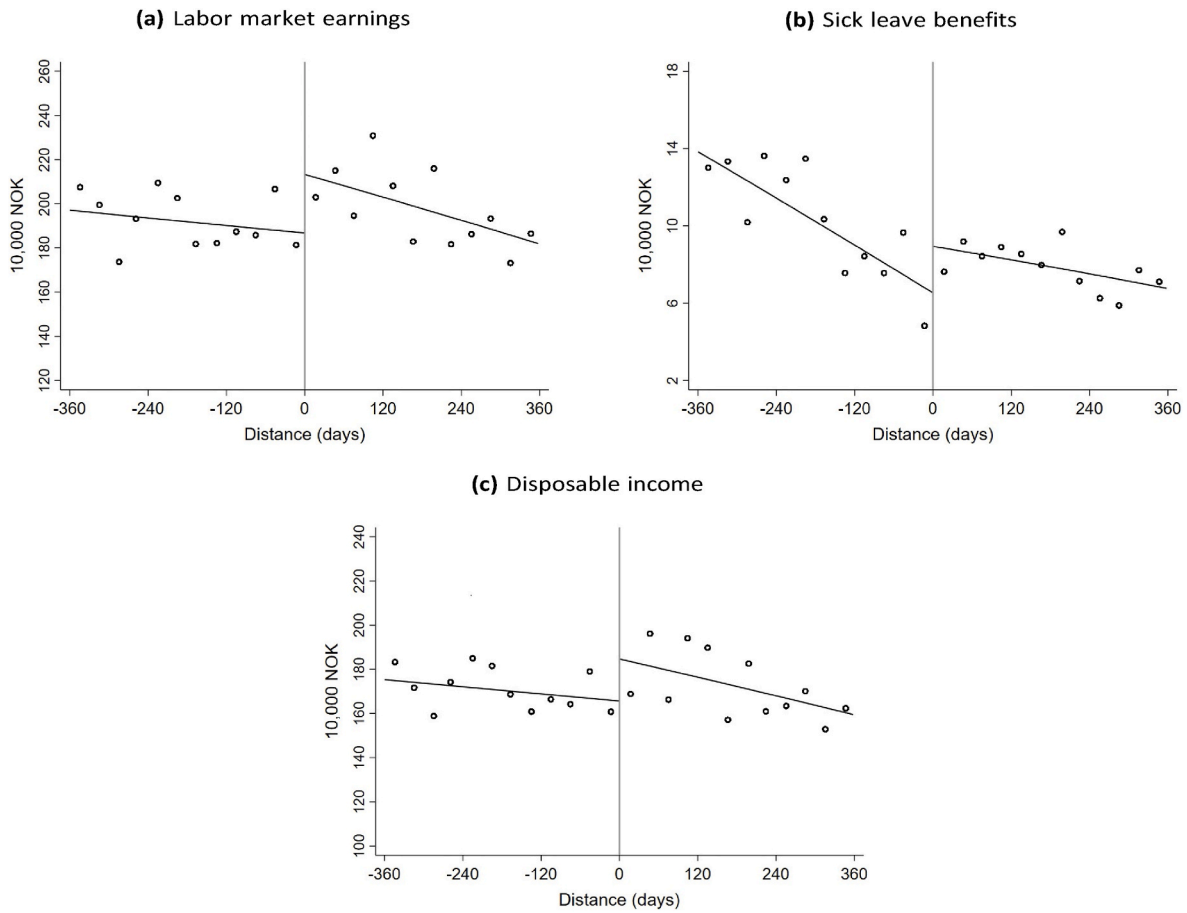


Fig. 4. RD graphs for refugees' economic outcomes.

Notes: This figure graphically illustrates the smoothness of economic outcomes across the distance between the arrival date and September 1, 2005 for the directly affected individual sample. Each graph plots the mean value using 30-day bins for the distance.

Data source: Statistics Norway.

Table 4
Effect on refugees' economic outcomes.

	(1)	(2)	(3)
	Labor market earnings	Sick leave benefits	Disposable income
Treated	19.911** (8.159)	1.905*** (0.675)	13.805** (5.830)
Pre-reform mean	185.241	5.649	161.535
Bandwidth	126.174	124.819	113.127
Observations	1,456	1,391	1,293
Mean of dep. var	200.183	8.326	174.103

Notes: Treated shows RDD estimate, which is the β coefficient in Equation (1). Pre-reform mean is the mean of the outcome for the untreated group measured at the cutoff, which is the α coefficient in Equation (1). Robust standard errors are in parentheses below coefficients. Bandwidth is calculated using the triangular kernel and the optimal bandwidth procedure proposed by Calonico et al. (2019). Observations and the mean of dependent variables are the estimation sample within the optimal bandwidth. Labor market earnings, sick leave benefit and disposable income (NOK 10,000, 2015 values) are measured with the 11-year accumulation since arrival.

** $p < 0.01$, *** $p < 0.05$, * $p < 0.1$.

Data source: Own calculations based on the data provided by Statistics Norway.

sharply with those prevalent in refugees' countries of origin. To investigate this, I analyze whether exposure to the language training program affects marriage and fertility outcomes among refugees compared to those without such exposure. For both marriage and fertility, I examine cumulative behaviors over a 12-year period following arrival, focusing

on two primary outcomes measured at both extensive and intensive margins.

Marriage outcomes are assessed based on whether an individual has at least one year in marriage and the total number of years in marriage. Fertility outcomes are evaluated by whether an individual has at least one child and the total number of children, accounting for both Norwegian-born children and all children.

The estimates presented in Online Appendix Table A.5 reveal no statistically significant differences in marriage and fertility behaviors between treated and untreated refugees. These statistically insignificant effects persist across both the extensive and intensive margins. These results are in line with those found by Foged et al. (2022) for the Danish case. These findings suggest that participation in the language training program has no statistically significant effect on the size, structure, or composition of refugee families. This implies that the observed impacts on healthcare integration are more plausibly attributed to economic channels, as previously identified, rather than to changes in family dynamics or composition.

6. Spillover effects on the next generation's health

Having established the robust positive effects of language training on healthcare and economic integration among adult refugees, I now examine the existence and operation of spillover effects on the early-life health of the next generation. This part of the paper is motivated by the positive impacts of the reform on the first generation of refugees and their family members, as described in the previous section as well as by a

well-established body of the economics literature that demonstrates the influence of parental background on infant health (for example, see [Almond and Currie, 2011](#); [Hoynes et al., 2015](#)). For the main effects, I examine the effects of parental exposure to the reform on the likelihood of a child having health issues at birth, including preterm birth, low birth weight, and short birth length. Given the assumption that treated and untreated parents have similar characteristics except for exposure to the reform, as supported by the various checks and tests in Section 4.2, it is likely that in the absence of the reform, the children who were born into these groups would have had similar health outcomes early in life. The observed differences in children's health at birth, if any, should only be attributable to differences in exposure to the reform. I furthermore explore several suggestive channels for the main effects by estimating the effects of the reform on prenatal conditions including the home environment and maternal care during pregnancy.

I focus on the Norwegian-born children of refugee parents. Those parents feature in the sample of adults and were the first of the two parents to arrive in Norway (for children with two parents). To produce the baseline estimates, I re-estimate a model that is similar to the baseline specification of Equation (1). The terms on the right-hand side are defined on the basis of the parent, and the dependent variables on the left-hand side are the child's health outcomes.

6.1. Main effects on children's health at birth

To provide the reader with a visual illustration, I first present the discontinuous reductions around the cutoff for these three health outcomes of the children, as shown in [Fig. 5](#). Birth to parents whose running variables are marginally above the zero point on the horizontal axis is associated with a sudden drop in the share of premature birth (graph a), low birth weight (graph b), and short birth length (graph c).

The corresponding RD estimates in [Table 5](#) suggest that there is an improvement in all measures of children's health at birth thanks to parental exposure to language training. Parental exposure to the reform reduces the child's probability of having a health problem at birth: preterm birth decreases by 1.9 percentage points (column 1, statistically significant at the 5% level), low birth weight decreases by 4.4 percentage points (column 2, statistically significant at the 5% level), and short birth length decreases by 3.2 percentage points (column 3, statistically significant at the 1% level). In terms of the magnitude, these effects are comparable because, relative to the pre-reform means, they correspond to reductions of 33.9%, 45.8%, and 21.2% for premature birth, low birth weight, and short birth length, respectively.

It is also important to recall that exposure to the language training program has no statistically significant effects on adult refugees' child-bearing behaviors at both the extensive and intensive margins among Norwegian-born children and among all children as discussed in Section 5.3. These findings are helpful in clearing out a concern regarding fertility selection which can to some extent affect the estimates of the intergenerational impacts. In addition, the finding that the reform has no effect on the number of children would suggest no potential drivers which likely stem from the quantity-quality trade-offs.

I inquire how robust the baseline effects are by using the second-order function of the running variable, by adding control variables, by excluding the weights from the regressions, and by using the uniform kernel to select a bandwidth. Moreover, I run two additional regressions in which I also control for the child's birth order relative to Norwegian-born siblings and relative to all siblings. The corresponding estimates that are reported in Online Appendix [Table A.6](#) are highly similar to the baseline RD estimates in terms of the direction and the size of the effect. Each estimated coefficient is statistically significant at least at the 10% level. Importantly, while previous studies have shown birth order has an influence on the health of newborns (see for example [Brenøe and Molitor, 2018](#); [Lehmann et al., 2018](#); [Coffey and Spears, 2021](#); [Pruckner et al., 2021](#)), the estimates in panels (c) and (d) of [Table A.6](#) indicate that the baseline estimates are not driven by the child's birth order, be it

among the Norwegian-born children of the parent (panel c) or among all of the parents' children (panel d). In light of these results, I am confident in concluding that parental language training improves the health of the next generation early in life.

6.2. Prenatal conditions as potential mechanisms

It is highly important to understand the key channels through which the refugee parent's language training affects the child's health at birth positively, as displayed in [Table 5](#) (see [Fig. 6](#) for the RD graphs). The economics literature has well-documented the roles of prenatal conditions and family backgrounds for infant health (see [Almond and Currie, 2011](#); [Almond and Mazumder, 2011](#)). In this section, I explore several factors that are related to the home environment and maternal care during pregnancy as suggestive mechanisms.³¹ **Home environment.** In columns (1) and (2) in [Table 6](#), I examine the reform effects on two measures that are proxies for the quality of the family environment, namely the parent's income and marriage status. The RD estimates are indicative of considerable improvements in these outcomes that are due to the parent's exposure to the language training program. Disposable income during pregnancy, a key economic outcome for parents who were exposed to the reform, is NOK 17,460 higher than that which untreated refugee parents obtain (equivalent to an increase of 11.7%). The home environment is also improved by an increase in the probability of the parent being married of

14.9 percentage points (corresponding to an increase of 29.1%). The presence of two parents during pregnancy likely boosts parental investment, which can subsequently enhance the child's health at birth. Given the well-established evidence on the importance of economic conditions and financial resources (see [Dehejia and Lleras-Muney, 2004](#); [Lindo, 2011](#); [Hoynes et al., 2015](#); [van den Berg et al., 2020](#)) and parental family formation (see [Buckles and Price, 2013](#); [Ayll'on and Ferreira-Batista, 2015](#)) for birth outcomes, these findings suggest that an improved home environment serves as a channel that links parental language training to children's health early in life.

Maternal care. Finally, to rule out the maternal care pathway, I further investigate the effects of the reform on the mother's access to primary care, prenatal checkups, and the use of prenatal supplements during pregnancy. Previous work has shown that maternal care conditions during pregnancy matter for child development outcomes ([Evans and Lien, 2005](#); [von Hinke et al., 2022](#)). The estimates in columns (3)–(5) in [Table 6](#) demonstrate an improvement for each of the outcomes: primary care consultations increase by about 5 GP visits (54.3%), prenatal checkups increase 2.2 GP visits (58.7%), and the use of prenatal supplements increases by 5.5 percentage points (14.3%). Such positive maternal care behaviors can be treated as prenatal investments because they are likely to be favorable to children's outcomes later in life such as educational achievements and labor market success (see [Buckles and Kolka, 2014](#); [Lehmann et al., 2018](#)). In my view, seeking primary care, attending prenatal check-ups, and taking prenatal supplements during pregnancy serve as additional pathways for the positive effects of language training on health at birth among the children of refugees.

The estimates in Online [Appendix Table A.7](#) suggest that the effects on prenatal conditions are strongly robust to several changes in the specification of the estimation. These robustness checks are similar to those that were conducted for the main effects.

7. Sensitivity, heterogeneity, and placebo tests

7.1. Sensitivity

To conduct a sensitivity analysis, I first examine the effects on the outcomes of interest by using an alternative measurement of the

³¹ See Section 3.2 for details on the definitions of these variables.

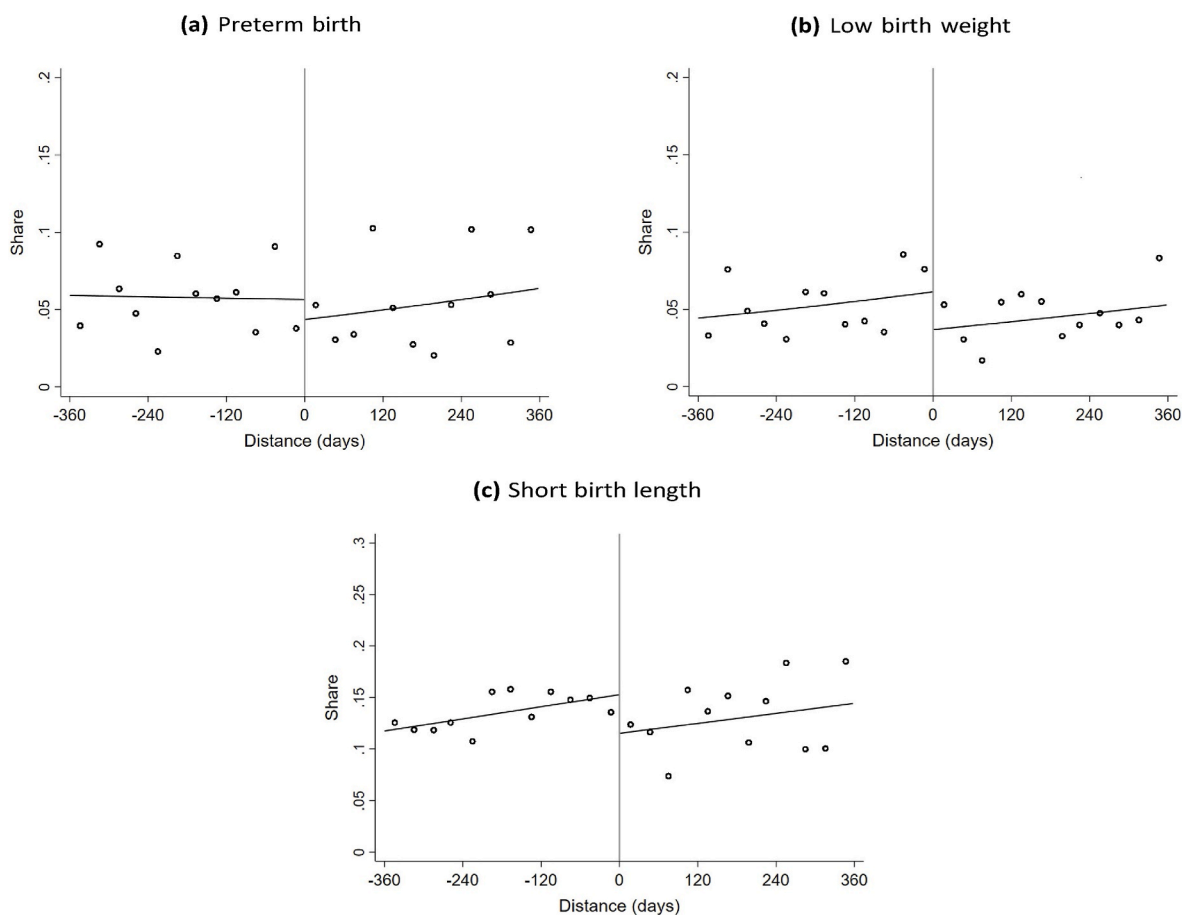


Fig. 5. RD graphs for children’s health at birth.

Notes: This figure graphically illustrates the smoothness of children’s health at birth across the distance between the arrival date and September 1, 2005 for the directly affected individual sample. Each graph plots the mean value using 30-day bins for the distance.

Data source: Statistics Norway.

Table 5
Effect on children’s health at birth.

	(1)	(2)	(3)
	Preterm birth	Low birth weight	Short birth length
Treated	-0.019** (0.009)	-0.044** (0.017)	-0.032*** (0.012)
Pre-reform mean	0.056	0.096	0.151
Bandwidth	130.935	104.557	120.852
Observations	1,613	1,289	1,463
Mean of dep. var	0.058	0.049	0.133

Notes: Treatment is language exposure of the first parent who arrived in Norway among the parents (for children with two parents). Treated shows RDD estimate, which is the β coefficient in Equation (1). Pre-reform mean is the mean of the outcome for the untreated group measured at the cutoff, which is the α coefficient in Equation (1). Robust standard errors are in parentheses below coefficients. Bandwidth is calculated using the triangular kernel and the optimal bandwidth procedure proposed by Calonico et al. (2019). Observations and the mean of dependent variables are the estimation sample within the optimal bandwidth.

** $p < 0.01$, * $p < 0.05$, $p < 0.1$.

Data source: Own calculations based on the data provided by Statistics Norway.

dependent variables and an analytical sample that is constructed in a different way. I also provide the results for the effects as estimated with various bandwidths rather than the optimal bandwidth for the baseline effect.

Healthcare integration. For the use of primary care services, I first exclude prenatal and genital consultations from the overall utilization in

order to arrive at a measure that is more amenable to comparisons between men and women. The RD estimate in column (1) in panel (a) of Online Appendix Table A.8 shows a positive effect (statistically significant at the 1% level). The effect is an increase of

16.8 consultations (42%) relative to the baseline effect of 19.2 consultations (44%). This effect is only 2 percentage points smaller than the baseline effect, suggesting that the use of a new measure of overall doctor visits does not change the main effect.

Second, I calculate the total use of primary care services by reference to a fixed period 2007–2019. The RD estimates in panel (b) of Online Appendix Table A.8 demonstrate the statistically significant and positive effects which are highly similar to those from the baseline estimates. The effect on overall utilization is an increase of about 21.2 visits (44.6%). The effects on specific types of consultations include an increase of nearly 3.7 visits for musculoskeletal consultations (31.5%), of approximately 3 visits for psychiatric consultations (82%), and of almost 2.5 visits for digestive consultations (52.5%). This alternative measure causes the corresponding effects to be larger than the baseline effects by a mere 0.6 percentage points for overall consultations, by 1.9 percentage points for musculoskeletal consultations, by 2.6 percentage points for psychiatric consultations, and by 3.6 percentage points for digestive consultations.³²

Online Appendix Figure A.7 shows that the magnitude and the

³² Online Appendix Figure A.6 illustrates a discontinuity in these measures around the cutoff.

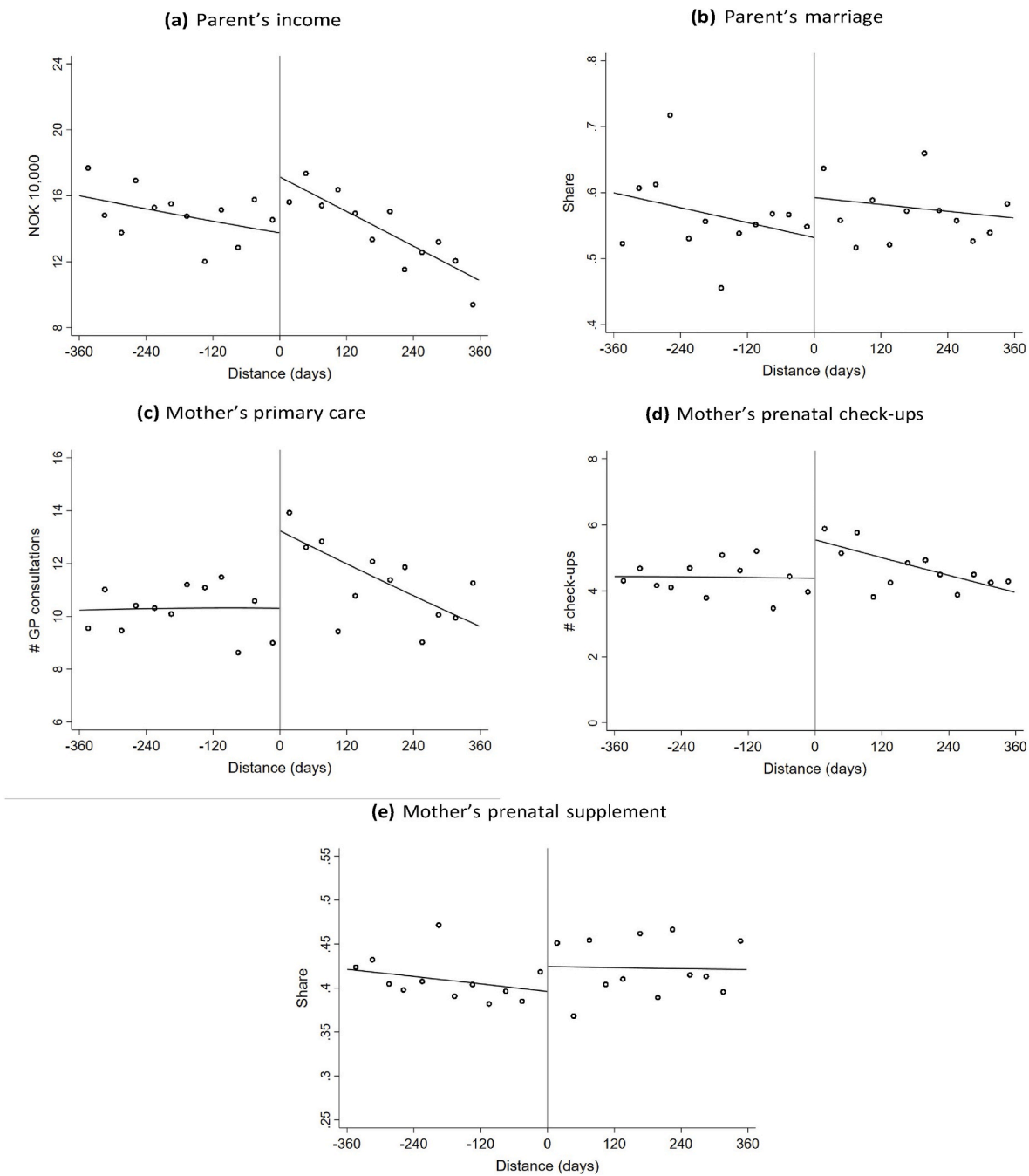


Fig. 6. RD graphs for children's prenatal conditio.

Notes: This figure graphically illustrates the smoothness of children's health at birth across the distance between the arrival date and September 1, 2005 for the directly affected individual sample. Each graph plots the mean value using 30-day bins for the distance.

Data source: Statistics Norway.

Table 6
Effect on children's prenatal conditions.

	(1)	(2)	(3)	(4)	(5)
	Home environment		Maternal care		
	Income	Marriage	Primary care consultations	Prenatal checkups	Prenatal supplements
Treated	1.746** (0.795)	0.149** (0.070)	5.408*** (1.251)	2.328*** (0.598)	0.046*** (0.017)
Pre-reform mean	14.968	0.512	9.249	3.783	0.391
Bandwidth	122.646	78.867	108.909	129.079	149.600
Observations	1,480	980	795	988	1,863
Mean of dep. var	15.445	0.568	11.321	4.801	0.410

Notes: Treatment is language exposure of the *first* parent who arrived in Norway among the parents (for children with two parents). Treated shows RDD estimate, which is the β coefficient in Equation (1). Pre-reform mean is the mean of the outcome for the untreated group measured at the cutoff, which is the α coefficient in Equation (1). Bandwidth is calculated using the triangular kernel and the optimal bandwidth procedure proposed by Calonico et al. (2019). Observations and the mean of dependent variables are the estimation sample within the optimal bandwidth. Disposable income is measured with the two-year average before the child's birth (10,000 NOK, 2015 values). Marriage is an indicator for being married in the year the child was born. Rich area is an indicator for the parent living in an area of the 5th income quintile in the year of child's birth. Primary healthcare services and prenatal checkups are measured by the accumulated number of consultations of two years one before and one in the year of child's birth. Prenatal supplement is an indicator for taking any type of supplements such as multivitamin and folate during pregnancy. Robust standard errors are in parentheses below coefficients.

** $p < 0.01$, * $p < 0.05$, $p < 0.1$.

Data source: Own calculations based on the data provided by Statistics Norway.

statistical significance of the RD estimates for the effect on doctor visits are highly robust to the choice of bandwidth for all four available measures. The estimates are highly stable for all outcomes from the baseline estimate to the estimate with a bandwidth of ± 360 days although the size of the effect diminishes slightly. The estimates are larger for the observations that are close to the cutoff in particular for the bandwidths of

± 30 days or ± 60 days but the standard errors that correspond to these estimates are also larger. This is likely to be the case because these samples are relatively small.

Children's health at birth. To understand whether the main effects are driven by the mother's or the father's exposure to the reform, I re-estimate the baseline Equation (1) separately for sub-samples that only comprise mothers or fathers. The estimates that are reported in Online Appendix Table A.9 indicate that both the only mother's (panel a) and the only father's (panel b) exposure to the reform has negative impacts on children's health at birth. All estimates remain statistically significant at least

at the 10% level, except for the one for the father's exposure on preterm birth (column 1, panel b). In terms of magnitude, while the effects on preterm birth and short birth length are larger for the mother's exposure than for the father's exposure, those on low birth weight show a contrast pattern. These findings suggest that the baseline effects are driven by the mother's exposure for preterm birth and short birth length whereas they are driven by the father's exposure for low birth weight.

Online Appendix Figure A.10 suggests that in general, the RD estimate of the reform effect on children's health at birth outcomes is highly robust to the choice of bandwidth for estimation in terms of both the

magnitude of the effect and its statistical significance. The size of the effect declines gradually across the bandwidths from ± 30 days to ± 360 days for preterm birth and low birth weight. It decreases more pronouncedly for bandwidths of between ± 30 days and ± 90 days and then becomes highly stable for low birth height. Overall, the estimates are largest for the bandwidths that are close to the cutoff, although the corresponding standard errors are large, especially for short birth length.

7.2. Heterogeneity

Next, I examine the heterogeneity of the reform effect across various groups. Specifically, I explore the direct effects on the integration into the healthcare system across genders and across the HDI levels of the countries of origin of the refugees. For the spillover effects on offspring's health at birth, I examine the child's genders and the parent's HDI level of the country of origin. The indicators for genders consist of a dummy for a male (or a boy for the children's generation) and a dummy for a female (or a girl for the children's generation). The indicators for the HDI levels include a dummy for a high HDI level which is defined as an HDI score above the median of the sample, and a dummy for a low HDI level which is defined as an HDI score below that median. I estimate a specification which is a modified version of the one from the baseline model, by adding *Treated* (D_i) and its interactions with the following dummies: male versus female, and a high versus a low HDI level of the countries of origin. The coefficient of the interaction (for example *Treated* \times *Male*) measures how exposure to the reform affects the outcome of interest for that group (for example males).

Healthcare integration. Online Appendix Table A.10 presents heterogeneous effects on the use of primary care services. For gender heterogeneity, which is presented in panel (a), the effects on total consultations are stronger for females than for males and the corresponding estimated coefficients are both statistically significant at the 1% and 5% level, respectively. The ratio of these effects is 1.43. For specific types of health consultations, the effects on musculoskeletal and digestive consultations are existent only for females (statistically significant at the 10% and 5% level, respectively) whereas the effects on psychological consultations are in evidence only among males (statistically significant at the 1% level). Turning to the coefficients for gender differences in the third row of panel (a) of Table A.10, only the coefficient for psychological consultations is statistically significant at the 5% level, which suggests that the number of psychiatric visits which is induced by the reform is larger for men than for

women by 2.9. Since the other coefficients for gender differences are statistically insignificant I cannot reject the hypothesis that the effects on the other outcomes are equal between genders.

In panel (b) of Table A.10, I also find heterogeneity in the reform effects across the high and low HDI groups. The effects on total consultations (column 1) and psychological consultations (column 3) are larger for the high HDI group than for the low HDI group by a fraction of 1.45 and 1.66, respectively. The estimated coefficients for both groups of refugees are statistically significant at the 5% or the 1% level. The effects on musculoskeletal consultations (column 2) and digestive consultations (column 4) are only existent for the low HDI group (statistically significant at the 10% and the 5% level, respectively). Despite these heterogeneous effects, the differences are not statistically different from zero because the corresponding coefficients in the third row are not statistically significant.

Children's health at birth. Regarding the generational persistence of the effect of language training across generations, I explore who

benefits the most as between boys and girls, and as between the high and low HDI groups of parental countries of origin. Online Appendix Table A.11 presents the effects across children's genders (panel a) and parental HDI levels (panel b). For gender heterogeneity, the effect on premature birth (column 1) only obtains for girls (a reduction of 2.1 percentage points, statistically significant at the 5% level). The effects on the other outcomes are surviving for both genders. The estimates are statistically significant at the 1% and the 10% level for boys' and girls' low birth weight (column 2), respectively, and at the 5% level for both genders' short birth length (column 3). The estimated coefficients reveal an effect that is more favorable for boys than for girls by a factor of nearly 1.7 for low birth weight and of almost 1.2 for short birth length. However, the coefficients for gender differences in the third row of the panel are not statistically significant.

The estimated coefficients in panel (b) demonstrate that the effects are generally more pronounced among children who were born to parents from the high HDI group than among children who were born to parents from the low HDI group. The effects on preterm birth are only alive for the high HDI group with a reduction of 3.8 percentage points (statistically significant at the 1% level). The effects on the other outcomes are existent for both groups, with those for the high HDI group being larger than those for the low HDI group by a factor of nearly 1.3 for low birth weight and of 1.8 for short birth length. However, the only coefficient for the difference as shown in the third row of the panel that is statistically significant is that for low birth weight (at the 5% level). It suggests that the effect gap on preterm birth between the high and the low HDI group is 2.4 percentage points.

7.3. Placebo tests

To validate my identification strategy further, I perform two sets of placebo tests to investigate whether the baseline estimates for the direct and the spillover effects *only* originate from the *true* reform. For the first set of placebo tests, I estimate the effects of several *artificial* reforms by using the samples of refugees and their family members, which are constructed in the same way as in the main analysis. I particularly move the cutoff date 3–6 years back (or 1–3 years forward) before (or after) September 1, 2005 to create cutoff dates for the placebo reforms. I use six artificial reforms, dated September 1 for the 2000–2002 and the 2006–2008 period. The corresponding overall sample for the artificial reform includes refugees and their family members who were granted asylum status within ± 360 days of the cutoff date. I then estimate the baseline specification from Equation (1) based on similar outcomes of interest for these artificial reforms. Because of data availability, the measurement of the dependent variables is slightly different from the baseline outcomes.

As far as the period over which primary care consultations accumulate are concerned, I use years 7–18 for the 2000 reform, years 6–17 for the 2001 reform, years 5–16 for the 2000 reform, years 1–12 for the 2006 reform, years 1–11 for the 2007 reform, and years 1–10 for the 2008 reform. I find none of these fake reforms have an impact on the measures of primary care services utilization as shown in Online Appendix, panels (a)–(f) of Table A.12.

Turning to the effects of the artificial reforms on children's health at birth, the samples of children are constructed similarly to the main sample. The dependent variables for health at birth (preterm birth, low birth weight, and short birth length) are measured as in the main analysis. The results in Online Appendix, panels (a)–(f) of Table A.13 indicate that none of the artificial reforms have a statistically significant impact on the health at birth outcomes of the offspring's generation.

Second, I exploit the original reform through the cutoff date of September 1, 2005, but this time I focus on individuals who were not actually affected by the true reform. I select the immigrants from the Nordic countries and the European Economic Area (EEA) country members as an ideal unaffected group because, under the Introduction Act, they are not required to participate in the Norwegian language

training program. I estimate the baseline specification Equation (1) based on similar outcomes of interest for refugees' use of primary care services and children's health at birth. The estimates in panel (g) of Table A.12 and Table A.13 confirm that the reform does not affect any of the outcomes among the immigrants from the Nordic and the EEA countries.

7.4. Multiple hypothesis testing

Given that I estimate the effects of being exposed to the language training program on a series of potentially correlated outcomes within each generation, I adopt the stepwise multiple testing procedure developed by Romano and Wolf (2005) to address the potential correlations among these outcomes. This approach, designed to maintain control over the family-wise error rate, operates through an iterative rejection or acceptance mechanism for a fixed level of significance. I therefore employ 1,000 block-bootstrap replications to compute the adjusted *p*-values for the estimates. I particularly test hypotheses for the estimates presented in separately each table and for those jointly combined for each generation. The total number of hypotheses tested simultaneously corresponds to the number of estimates in each table and the combined estimates for each generation—seven for the directly affected generation and eight for the next generation.

Online Appendix Table A.15 and Table A.16 respectively present *p*-values for the estimates on directly affected refugees and refugee children, in particular those for the baseline estimates (column 1) and those for the estimates after accounting for multiple hypothesis testing (column 2 for the estimates in each table, and column 3 for all the estimates of each generation). Although the Romano-Wolf step-down adjusted *p*-values in both columns (2) and (3) in both tables are to some extent larger than the baseline *p*-values, they remain statistically significant at least the 10% level. These results suggest that all the estimates for the effect on both generations are statistically different from zero after adjusting the inference for multiple hypothesis testing.

8. Discussion

8.1. Magnitude of the estimates and economic meanings

The effects of language training that emerged from this study are typically large and thus highlight the value of language training not only for refugees and their family members but also for the next generation in a receiving country. It is also highly important to discuss the possible causes of these significant impacts. Refugees typically lack linguistic capabilities before they are granted protection because they seldom plan to leave their home countries; instead, it is wars, conflicts, and the risk of serious human rights violations and persecution, among others, that forced them to flee. Usually, refugees do not know what their final destinations will be. Learning the language of the host country during such a journey is almost impossible. Refugees' inability to communicate effectively with residents in receiving society prevents them from joining the local labor market and from using essential services. For that reason, it is understandable that language training programs are highly conducive to refugees' successful integration.

To obtain a more general idea of the size of the effects of language training that are documented in the literature, I propose to examine estimates from recent causal studies. It emerges that the effects in this paper are in line with previous findings; this is true of both the direct and the intergenerational effects. For example, a similar language training program in Denmark was found to improve long-run labor market earnings by 34% and the probability of an affected refugee obtaining a complex job increased by 123% (see Foged et al., 2022). That program was also found to reduce the probability of such a refugee being convicted of a violent crime by 82.3%; the likelihood of affected refugees being charged with a property crime by 62.4%; and that of a member of the next generation being charged with any crime during adolescence by

42% (see Foged et al., 2023). Settlement in an area where the spoken language is similar to that of the refugee's country of origin has been reported to raise the probability of the labor market participation by more than 100% in the first five years of residence in Switzerland (see Schmid, 2023). Although the outcomes of interest in those studies differ from the ones that were examined in this paper, their findings highlight the significant effects of language training and serve to contextualize my own.

8.2. Norwegian training take-up and language proficiency

A fundamental question for evaluating any language training program is whether it effectively improves the language proficiency of its participants. While assessing whether refugees' Norwegian language skills are directly driven by their participation in the program is crucial, this paper cannot provide causal RDD estimates for this outcome due to the unavailability of proficiency data in the datasets used. However, I draw on the data from Statistics Norway's Living Conditions Survey for Immigrants 2016, which offers descriptive evidence on the relationship between participation in Norwegian language training and language proficiency (Statistics Norway, 2022), to provide some insights into this relationship.³³

I particularly exploit information about immigrants' language to construct an indicator for Norwegian language training participation, coded as 1 if a respondent attended any training and 0 otherwise. To assess language proficiency, I create six indicators: one for overall proficiency and five for specific skills—understanding radio and TV news, reading newspapers, talking to people, writing a job application, and discussing health issues with a doctor. Proficiency in each domain is measured on a 5-point scale based on the question: "Would you say that your Norwegian proficiency is (1) very good, (2) quite good, (3) okay, (4) quite poor, or (5) extremely poor?" A positive proficiency is defined as a response of (1), (2), or (3). For language use, I construct two additional indicators. Speaking Norwegian at work is based on the question: "How much of the working day do you speak Norwegian?" with responses.

(1) "All or most," (2) "About half," or (3) "Less than half" coded as 1, and all others as 0. Speaking Norwegian at home is coded as 1 if the respondent answered "Yes" to using Norwegian at home, and 0 otherwise. The analysis sample includes 2,480 immigrants with complete data for these variables.

Descriptive statistics. The graphs in Online Appendix Figure A.11 present descriptive statistics on the proportion of immigrants achieving positive Norwegian language skills and usage behaviors, comparing those who participated in Norwegian training programs with those who did not. The results highlight the substantial benefits of such programs, with participation consistently linked to improvements in language proficiency ranging from 12 to 21 percentage points across various domains. Skills critical for integration—such as understanding media, engaging in conversations, writing job applications, and accessing healthcare—show particularly notable gains. For instance, panel (a) illustrates that 88% of Norwegian training participants demonstrated overall proficiency, compared to only 67% of non-participants, reflecting a significant 21-percentage-point advantage attributed to formal language instruction.

Panels (b)–(f) illustrate specific language skills applied to key life

activities, underscoring the positive impact of Norwegian training programs. Panel (b) shows that 88% of training participants were proficient in understanding radio and TV news, compared to 70% of non-participants. Similarly,

panel (c) reveals that 87% of participants could read Norwegian newspapers proficiently, while only 69% of non-participants achieved the same. Panel (d) highlights that 92% of individuals who underwent training were proficient in conversing with others, compared to 75% of non-participants. The ability to write a job application, shown in panel (e), was achieved by 79% of training participants, significantly higher than the 58% of non-participants. Finally, panel (f) emphasizes the health-related advantages of language proficiency, with 88% of participants able to discuss health concerns with a GP, compared to 67% of non-participants—a 21-percentage-point difference. These findings reinforce the critical role of Norwegian language training in equipping immigrants with the skills necessary for integration, particularly in accessing healthcare services and navigating essential daily activities.

Panel (g) highlights that 60% of training participants spoke Norwegian at work, compared to 54% of non-participants, indicating a modest 6-percentage-point improvement in workplace communication. Panel (h) reveals that 54% of training participants used Norwegian at home, compared to 42% of non-participants, reflecting a 12-percentage-point difference. These smaller gains suggest that language use in domestic and workplace settings may be shaped by factors beyond formal training, such as family dynamics, cultural preferences, or workplace environments. Nonetheless, the positive effects of Norwegian training on language usage behaviors remain evident, reinforcing its role in promoting integration across different spheres of daily life.

Correlational estimates. I next examine the correlational estimates for the relationship between Norwegian language training take-up and language proficiency and usage behaviors to check whether this relationship is statistically significant. I regress each of the eight outcomes defined above on participation in Norwegian language training without (panel a) or with conditioning on some control variables (panels b–c).³⁴

The estimates in Online Appendix Table A.14 present a strong positive correlation between a Norwegian language training take-up and Norwegian language proficiency and usage behaviors. The estimates are strongly statistically significant at the 1% level for any outcome and for any specification. Participation in a Norwegian training course are positively associated with positive language skills by 14–20 percentage points for over proficiency (column 1) and 13–21 percentage points for other specific skills (columns 2–6), and with Norwegian usage by 5–12 percentage points for speaking at work or at home (columns 7–8).

Both descriptive statistics and correlational estimates I have presented here support the hypothesis that Norwegian language training which is induced through exposure to the reform improves the language skills of refugees. The reform would therefore substantially improve refugees' ability to understand local media, strengthen their verbal communication skills, enhance their preparedness for the labor market, and improve their engagement with the healthcare system. These findings highlight

the critical role of structured language training in equipping refugees with the essential skills needed for active and meaningful participation in Norwegian society.

³³ This survey contains detailed information on the living conditions of approximately 5,500 respondents with immigrant backgrounds in Norway, covering individuals aged 16–74 from twelve major immigrant groups: Poland, Bosnia-Herzegovina, Kosovo, Turkey, Iraq, Iran, Afghanistan, Pakistan, Sri Lanka, Vietnam, Eritrea, and Somalia. The survey addresses topics such as background, housing, employment, income, language skills, social interactions, transnational ties, well-being, health, discrimination, citizenship, and belonging (Vrøalstad and Wiggen, 2017).

³⁴ The control variables include indicators for gender (male, female), fixed effects for age at arrival (0–6 years, 7–15 years, 16–19 years, 20 years or older), and fixed effects for the length of residence (2–3 years, 4–6 years, 7–10 years, 11–15 years, 16–20 years, 21 years or over), and fixed effects for countries of origin (which are grouped into several categories: Poland, Turkey, Bosnia and Herzegovina, Kosovo, Eritrea, Somalia, Afghanistan, Sri Lanka, Iraq, Iran, Pakistan, Vietnam).

8.3. Policy implications beyond Norway and developed countries

While the findings on the benefits of language training programs for refugees' healthcare integration and the health of the next generation in this study are drawn from Norway—a high-income country with an exceptional healthcare system—the policy implications can be extended far beyond this specific context. Refugees almost universally face linguistic barriers that hinder their integration and access to essential healthcare services, regardless of where they are hosted. This is because refugee immigrants frequently migrate from areas characterized by distinct linguistic roots with destination societies. For instance, Syrian refugees resettling in Germany or the United States must acclimate to environments dominated by German or English, languages that diverge substantially from Arabic. Similarly, Syrian refugees in Turkey confront analogous linguistic obstacles, as Arabic is fundamentally dissimilar to Turkish. Rohingya refugees from Myanmar grapple with similar challenges in Bangladesh, where Bengali predominates as the primary language. Likewise, Somali refugees in Kenya face impediments to integration due to significant linguistic divergence between Somali and the nation's principal languages, Swahili and English. Venezuelan refugees in Colombia encounter subtler yet consequential distinctions between Venezuelan and Colombian Spanish, complicating communication and social assimilation. These examples underscore that linguistic disparities between refugees' countries of origin and their host societies constitute a global problem for refugee integration, in particular healthcare access for refugees in receiving societies (World Health Organization, 2022, 2023).

Refugees essentially need considerable access to healthcare in host countries to address their immediate physical and mental health problems which are often exacerbated by the trauma of displacement, poor living conditions, and lack of care during their journeys from home to destination (World Health Organization, 2022, 2023; Awuah et al., 2023). Yet, there are many significant obstacles facing refugees, often leaving refugees with insufficient access to healthcare and necessary medical services. This stylized fact remains an urgent public health problem facing many hosting countries in the developing world (Zihindula et al., 2015; Kaya et al., 2018; Tayfur et al., 2019; Carreno-Calderon et al., 2020; King et al., 2022). Deficient healthcare conditions likely negatively affect not only refugees themselves but also the health of natives through the spread of diseases among other public health costs (Baez, 2011; Ib'añez et al., 2021). Anecdotal evidence and discussions indicate that refugees in developing countries still encounter linguistic barriers that hinder refugees' access to local healthcare services and undermine their health and well-being (Mangrio and Sjögren Forss, 2017; Chuah et al., 2018; World Health Organization, 2018; Assi et al., 2019; World Health Organization, 2022).

This problem is particularly exacerbated in destination countries in developing regions, which accommodate a significant proportion of the world's refugees and asylum-seekers and where language training and translation services are often insufficient (Devictor et al., 2021). As of November 2024, low- and middle-income countries accommodate approximately 71% of the world's forcibly displaced immigrants (UNHCR, 2024b).³⁵ Knowledge about the effectiveness of any policy intervention, including language training, is therefore very helpful in allocating resources to improve refugees' integration in these countries. For instance, previous studies have analyzed the roles of cash transfers for refugees' integration (Salti et al., 2022; Altundäg and O'Connell, 2023; Aygün et al., 2024). Given the importance of language training for refugees and their households in host societies which has been shown in several developed countries, very little is known about the benefits of

programs aiming to improve refugees' language proficiency in developing countries. My findings suggest that empowering refugees through the provision of language training programs is likely to yield significant positive outcomes for refugees. Notably, these benefits are expected to be even greater in developing countries, where support systems such as welfare schemes for more crowded refugee populations often operate less effectively than in high-income countries like Norway due to limited resources and institutionalized barriers (Asgary and Segar, 2011).

In sum, this paper studies the impact of a language training program that has been increasingly implemented in many host countries, making the findings broadly applicable. This focus on a common intervention, rather than a rare or atypical event, enhances the generalizability of the results. Furthermore, language barriers are increasingly identified as a critical obstacle to refugee integration across numerous countries, underscoring the relevance of this analysis in addressing a pervasive issue in global migration and integration policy.

9. Conclusion

Poor language skills have been identified as the most prominent barrier to social integration for refugees. Linguistic deficits prevent refugees from securing remunerative employment, accessing healthcare services, and seeking resources to improve their children's health, among others. In this paper, I have studied whether and how a well-designed and universal language training program affects refugees' integration into host-country healthcare systems and the health of their children early in life. To identify causal effects, I exploited a policy reform in Norway that made it compulsory for refugees and their family members to participate in at least 300 h of Norwegian language training after September 1, 2005. The discontinuous implementation of the reform enabled me to perform an RD analysis, through which I compared an outcome of interest across two groups of refugees and their family members who had similar characteristics with the exception of exposure to the reform.

I documented the large impact of language training on the refugees' integration into the Norwegian healthcare system. The refugees and their family members who were exposed to the Norwegian language training program used more primary care services than those who were not exposed to it. I suggested

that the effect in question operates through considerable improvements in labor market attachment and economic well-being that language training induces. I also found that language training is linked to improvements in the health of the children of refugees at birth. Parental language training is tied to significant reductions in the incidence of prematurity, low birth weight, and short birth length. To shed light on the key mechanisms that may underlie the intergenerational impacts, I discovered the positive effects of language training on several important measures of the home environment (income and marriage) and maternal care (primary care, prenatal checkups, and supplement take-ups) outcomes during pregnancy. Those improvements are known to conduce to the birth of healthy children. The estimates of the direct effects on healthcare utilization among refugees and the intergenerational effects are robust to key changes in the functional form for estimation as well as to some sensitivity checks. Overall, I have provided the first causal evidence of the multigenerational impacts of a language training program for refugees in a receiving country. My findings imply that, through both the long-run effects on the integration of the refugees into the healthcare system and the spillover effects across generations, enacting and enforcing compulsory language training programs for newly arrived refugees has high social returns.

Despite the significant contributions of this paper to the existing literature, it has certain limitations that warrant discussion. First, while the causal impact of exposure to the language training reform on refugees' actual language proficiency is a critical aspect to investigate, this study could not present the RD estimates for this outcome due to the unavailability of information about language skills in registry data. This

³⁵ Four of the top five countries that have hosted almost one-third of the world's refugees are from the developing world: Iran with 3.8 million, Turkey with 3.1 million, Colombia with 2.8 million, Germany with 2.7 million, and Uganda with 1.7 million (UNHCR, 2024b).

limitation is not unique to this study but is a common challenge in research relying on administrative data, as noted in similar works (see Foged et al., 2022; Foged et al., 2023). The absence of direct measures of language proficiency constrains the ability to examine the intermediary mechanisms that link language training to refugees' long-term healthcare outcomes. Addressing this gap in future research could involve integrating survey data, language assessments, or alternative data sources to complement registry data. Second, from a methodological perspective, the RD design employed in this study is recognized for its robustness and reliability in producing credible causal estimates. Compared to other quasi-experimental methods, RD estimators often provide greater internal consistency when assessing the impacts of policy interventions (Lee, 2008; Lee and Lemieux, 2010). In this context, the internal validity of the RD estimates is reinforced by focusing on a sub-sample of refugees narrowly distributed around the policy cutoff. These individuals share similar personal characteristics and institutional contexts, reducing potential biases and ensuring the estimates reflect the causal effects of the language training reform. Furthermore, the use of a complete population sample of refugees around the cutoff date enhances the external validity of the findings. This approach ensures that the results are representative of the population most directly impacted by the reform during the specified period. However, external validity may still be challenged when considering temporal variations. Refugees arriving at different times—particularly those who were far from the cutoff date—might encounter diverse contexts, including changes in policy environments, economic conditions, or societal attitudes.

These factors could interact with the language training program and influence healthcare outcomes. Consequently, isolating the specific effect of the language training program from other concurrent influences remains a complex endeavor. Therefore, while this study provides robust and valuable insights into the multigenerational healthcare benefits of language training for refugees, acknowledging these limitations helps delineate the scope of its contributions and highlights areas for further investigation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jdevco.2025.103470>.

Data availability

The authors do not have permission to share data.

References

- Almond, D., Currie, J., 2011. Killing me softly: the fetal origins hypothesis. *J. Econ. Perspect.* 25 (3), 153–172.
- Almond, D., Currie, J., Duque, V., 2018. Childhood circumstances and adult outcomes: Act II. *J. Econ. Lit.* 56 (4), 1360–1446.
- Almond, D., Mazumder, B., 2011. Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy. *Am. Econ. J. Appl. Econ.* 3 (4), 56–85.
- Altındağ, O., O'Connell, S.D., 2023. The short-lived effects of unconditional cash transfers to refugees. *J. Dev. Econ.* 160, 102942.
- Asgary, R., Segar, N., 2011. Barriers to health care access among refugee asylum seekers. *J. Health Care Poor Underserved* 22 (2), 506–522.
- Assi, R., Zger-Ilhan, S.O., Ilhan, M., 2019. Health needs and access to health care: the case of Syrian refugees in Turkey. *Public Health* 172, 146–152.
- Auer, D., 2018. Language roulette—the effect of random placement on refugees' labour market integration. *J. Ethnic Migrat. Stud.* 44 (3), 341–362.
- Auer, D., Kunz, J.S., 2024. Communication barriers and infant health: the intergenerational effect of randomly allocating refugees across language regions. *Am. Econ. J. Econ. Pol.* 00 (0), 000–000 <https://www.aeaweb.org/articles?id=10.1257/pol.20230220&from=f>.
- Auwah, W.A., Adebuseye, F.T., Ferreira, T., Azeem, S., Bharadwaj, H.R., Akpan, A.A., Wellington, J., Zia, M.R., Kumar, H., Khalid, A., et al., 2023. The Unmet Surgical Needs of Global Refugee Populations: A Perspective Review, vol. 11. SAGE Open Medicine, 20503121231204492.
- Aygün, A.H., Kırdar, M.G., Koyuncu, M., Stoeffler, Q., 2024. Keeping refugee children in school and out of work: evidence from the world's largest humanitarian cash transfer program. *J. Dev. Econ.* 168, 103266.
- Ayllo'n, S., Ferreira-Batista, N.N., 2015. 'mommy, i miss daddy'. the effect of family structure on children's health in Brazil. *Econ. Hum. Biol.* 19, 75–89.
- Baez, J.E., 2011. Civil wars beyond their borders: the human capital and health consequences of hosting refugees. *J. Dev. Econ.* 96 (2), 391–408.
- Bau, N., Fernández, R., 2023. Culture and the family. In: *Handbook of the Economics of the Family*, vol. 1. Elsevier, pp. 1–48.
- Bhuller, M., Dahl, G.B., Løken, K.V., Mogstad, M., 2023. Domestic violence reports and the mental health and well-being of victims and their children. *J. Hum. Resour.* 59 (S), S152–S186.
- Black, N., Kunz, J.S., 2024. The intergenerational effects of language proficiency on child health outcomes: evidence from survey-and census-matched health care records. *J. Econ. Behav. Organ.* 225, 136–152.
- Borjas, G.J., 1994. The economics of immigration. *J. Econ. Lit.* 32 (4), 1667–1717.
- Bratsberg, B., Raum, O., Roed, K., 2016. Labour market integration of refugees in Norway. Refugees and economic migrants: Facts, policies and challenges 37–54.
- Brell, C., Dustmann, C., Preston, I., 2020. The labor market integration of refugee migrants in high-income countries. *J. Econ. Perspect.* 34 (1), 94–121.
- Breneau, A.A., Molitor, R., 2018. Birth order and health of newborns: what can we learn from Danish registry data? *J. Popul. Econ.* 31, 363–395.
- Bryant, R.A., Nickerson, A., Morina, N., Liddell, B., 2023. Posttraumatic stress disorder in refugees. *Annu. Rev. Clin. Psychol.* 19, 413–436.
- Buckles, K., Kolka, S., 2014. Prenatal investments, breastfeeding, and birth order. *Soc. Sci. Med.* 118, 66–70.
- Buckles, K.S., Price, J., 2013. Selection and the marriage premium for infant health. *Demography* 50 (4), 1315–1339.
- Butikofer, A., Ginja, R., Løken, K.V., Landaud, F., 2023. Higher-achievement schools, peers and mental health. *Econ. J.* 133 (655), 2580–2613.
- Calonico, S., Cattaneo, M.D., Farrell, M.H., Titiunik, R., 2019. Regression discontinuity designs using covariates. *Rev. Econ. Stat.* 101 (3), 442–451.
- Carreno-Calderon, A., Cabieses, B., Correa-Matus, M.E., 2020. Individual and structural barriers to Latin American refugees and asylum seekers' access to primary and mental healthcare in Chile: a qualitative study. *PLoS One* 15 (11), e0241153.
- Chey, T., Brooks, R., Dominique le Touze, B., Bryant, R., 2007. The impact of the refugee decision on the trajectory of PTSD, anxiety, and depressive symptoms among asylum seekers: a longitudinal study. *American Journal of Disaster Medicine* 2 (6).
- Chin, A., Cortes, K.E., 2015. The refugee/asylum seeker. In: *Handbook of the Economics of International Migration*, vol. 1. Elsevier, pp. 585–658.
- Chiswick, B.R., Miller, P.W., 1995. The endogeneity between language and earnings: International analyses. *J. Labor Econ.* 13 (2), 246–288.
- Chiswick, B.R., Miller, P.W., 2015. International migration and the economics of language. In: *Handbook of the Economics of International Migration*, vol. 1. Elsevier, pp. 211–269.
- Chuah, F.L.H., Tan, S.T., Yeo, J., Legido-Quigley, H., 2018. The health needs and access barriers among refugees and asylum-seekers in Malaysia: a qualitative study. *Int. J. Equity Health* 17, 1–15.
- Clark, R., Mytton, J., 2007. Estimating infectious disease in UK asylum seekers and refugees: a systematic review of prevalence studies. *Journal of Public Health* 29 (4), 420–428.
- Clausen, J., Heinesen, E., Hummelgaard, H., Husted, L., Rosholm, M., 2009. The effect of integration policies on the time until regular employment of newly arrived immigrants: evidence from Denmark. *Labour Econ.* 16 (4), 409–417.
- Clemens, M., Huang, C., Graham, J., et al., 2018. The Economic and Fiscal Effects of Granting Refugees Formal Labor Market Access, vol. 496. Center for Global Development Working Paper.
- Coffey, D., Spears, D., 2021. Neonatal death in India: birth order in a context of maternal undernutrition. *Econ. J.* 131 (638), 2478–2507.
- Currie, J., 2011. Inequality at birth: some causes and consequences. *Am. Econ. Rev.* 101 (3), 1–22.
- Debesay, J., Arora, S., Bergland, A., 2019. Migrants' consumption of healthcare services in Norway: inclusionary and exclusionary structures and practices. In: *Inclusive Consumption: Immigrants' Access to and Use of Public and Private Goods and Services*. Universitetsforlaget, pp. 63–78.
- Dehejia, R., Lleras-Muney, A., 2004. Booms, busts, and babies' health. *Q. J. Econ.* 119 (3), 1091–1130.
- Devictor, X., Do, Q.-T., Levchenko, A.A., 2021. The globalization of refugee flows. *J. Dev. Econ.* 150, 102605.
- Djuve, A.B., 2011a. Introductory programs for immigrants. *Nordic Journal of Migration Research* 1 (3), 113–125.
- Djuve, A.B., 2011b. Introduksjonsordningen for nyankomne innvandrere. et integreringspolitisk paradigmeskifte. Dr. avhandling. Universitetet i Oslo.
- Djuve, A.B., Kavli, H.C., 2019. Refugee integration policy the Norwegian way—why good ideas fail and bad ideas prevail. *Transfer: European Review of Labour and Research* 25 (1), 25–42.
- Djuve, A.B., Kavli, H.C., Sterri, E.B., Brøten, B., 2017. Introduksjonsprogram og norskkopplærning. hva virker—for hvem. Rapport 17, 31.
- Dustmann, C., Fasani, F., Frattini, T., Minale, L., Schönberg, U., 2017. On the economics and politics of refugee migration. *Econ. Policy* 32 (91), 497–550.
- Eiset, A.H., Wejse, C., 2017. Review of infectious diseases in refugees and asylum seekers—current status and going forward. *Public Health Rev.* 38 (1), 1–16.
- Evans, W.N., Fitzgerald, D., 2017. The Economic and Social Outcomes of Refugees in the United States: Evidence from the ACS. National Bureau of Economic Research. Technical report.
- Evans, W.N., Lien, D.S., 2005. The benefits of prenatal care: evidence from the pat bus strike. *J. Econ. Econ.* 125 (1–2), 207–239.
- Fasani, F., Frattini, T., Minale, L., 2022. (the struggle for) refugee integration into the labour market: evidence from Europe. *J. Econ. Geogr.* 22 (2), 351–393.

- Fazel, M., Wheeler, J., Danesh, J., 2005. Prevalence of serious mental disorder in 7000 refugees resettled in western countries: a systematic review. *Lancet* 365 (9467), 1309–1314.
- Ferwerda, J., Finseraas, H., 2024. Do integration courses influence refugees' integration trajectories? evidence from Norway. *J. Polit.* 00 (0), 000–000 <https://www.journals.uchicago.edu/doi/10.1086/734247>.
- Fletcher, J.M., Green, J.C., Neidell, M.J., 2010. Long term effects of childhood asthma on adult health. *J. Health Econ.* 29 (3), 377–387.
- Foged, M., Hasager, L., Peri, G., Arendt, J.N., Bolvig, I., 2022. Language training and refugees' integration. *Rev. Econ. Stat.* 1–41.
- Foged, M., Hasager, L., Peri, G., Arendt, J.N., Bolvig, I., 2023. Intergenerational spillover effects of language training for refugees. *J. Publ. Econ.* 220, 104840.
- Foged, M., van der Werf, C., 2023. Access to language training and the local integration of refugees. *Labour Econ.*, 102366.
- Gelman, A., Imbens, G., 2019. Why high-order polynomials should not be used in regression discontinuity designs. *J. Bus. Econ. Stat.* 37 (3), 447–456.
- Grossman, M., 1972. On the concept of health capital and the demand for health. *J. Polit. Econ.* 80 (2), 223–255.
- Hadgkiss, E.J., Renzaho, A.M., 2014. The physical health status, service utilisation and barriers to accessing care for asylum seekers residing in the community: a systematic review of the literature. *Aust. Health Rev.* 38 (2), 142–159.
- Hagelund, A., 2005. Why it is bad to be kind. educating refugees to life in the welfare state: a case study from Norway. *Soc. Pol. Adm.* 39 (6), 669–683.
- Hagelund, A., Kavli, H., 2009. If work is out of sight. activation and citizenship for new refugees. *J. Eur. Soc. Pol.* 19 (3), 259–270.
- Hardoy, I., Zhang, T., 2019. The long and winding road: labour market integration of refugees in Norway. Technical report, Memorandum.
- Hatton, T.J., 2017. Refugees and asylum seekers, the crisis in Europe and the future of policy. *Econ. Policy* 32 (91), 447–496.
- Heller, B., Mumma, K.S., 2023. Immigrant integration in the United States: the role of adult English language training. *Am. Econ. J. Econ. Pol.* 15 (3), 407–437.
- Hoynes, H., Miller, D., Simon, D., 2015. Income, the earned income tax credit, and infant health. *Am. Econ. J. Econ. Pol.* 7 (1), 172–211.
- Ibáñez, A.M., Roza, S.V., Urbina, M.J., 2021. Forced migration and the spread of infectious diseases. *J. Health Econ.* 79, 102491.
- Joyce, P., 2017. Inspiration för integration—en esorapport om arbetsmarknadspolitik för nyanlända i fem länder. Rapport till Expertgruppen för Studier i Offentlig ekonomi (ESO) 7.
- Kalt, A., Hossain, M., Kiss, L., Zimmerman, C., 2013. Asylum seekers, violence and health: a systematic review of research in high-income host countries. *American Journal of Public Health* 103 (3), e30–e42.
- Kaya, E., Karadag Caman, O., Kilic, C., Uner, S., 2018. Refugees' access to and utilization of health services: challenges and solutions in Turkey. *Eur. J. Publ. Health* 28 (Suppl. 4), cky214–273.
- King, J., Prabhakar, P., Singh, N., Sulaiman, M., Greco, G., Mounier-Jack, S., Borghi, J., 2022. Assessing equity of access and affordability of care among south sudanese refugees and host communities in two districts in Uganda: a cross-sectional survey. *BMC Health Serv. Res.* 22 (1), 1165.
- Kirmayer, L.J., Narasiah, L., Munoz, M., Rashid, M., Ryder, A.G., Guzder, J., Hassan, G., Rousseau, C., Pottie, K., 2011. Common mental health problems in immigrants and refugees: general approach in primary care. *Canadian Medical Association Journal* 183 (12), E959–E967.
- Kuziemko, I., 2014. Human capital spillovers in families: do parents learn from or lean on their children? *J. Labor Econ.* 32 (4), 755–786.
- Lang, J., 2022. Employment effects of language training for unemployed immigrants. *J. Popul. Econ.* 35 (2), 719–754.
- Laue, J., Risør, T., 2017. Refugees and healthcare services. *Tidsskrift for Den norske legeforening* 137 (1). <https://tidsskriftet.no/en/2017/12/global-helse/refugees-and-healthcare-services>.
- Lebano, A., Hamed, S., Bradby, H., Gil-Salmeron, A., Dura-Ferrandis, E., Garces-Ferrer, J., Azzedine, F., Riza, E., Karnaki, P., Zota, D., et al., 2020. Migrants' and refugees' health status and healthcare in Europe: a scoping literature review. *BMC Public Health* 20 (1), 1–22.
- Lee, D.S., 2008. Randomized experiments from non-random selection in US house elections. *J. Econom.* 142 (2), 675–697.
- Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. *J. Econ. Lit.* 48 (2), 281–355.
- Lehmann, J.-Y.K., Nuevo-Chiquero, A., Vidal-Fernandez, M., 2018. The early origins of birth order differences in children's outcomes and parental behavior. *J. Hum. Resour.* 53 (1), 123–156.
- Li, G., Sah, P.K., 2019. Immigrant and Refugee Language Policies, Programs, and Practices in an Era of Change, vol. 325. *Routledge International Handbook of Migration Studies*.
- Lindert, J., von Ehrenstein, O.S., Priebe, S., Mielck, A., Brähler, E., 2009. Depression and anxiety in labor migrants and refugees—a systematic review and meta-analysis. *Soc. Sci. Med.* 69 (2), 246–257.
- Lindo, J.M., 2011. Parental job loss and infant health. *J. Health Econ.* 30 (5), 869–879.
- Lochmann, A., Rapoport, H., Speciale, B., 2019. The effect of language training on immigrants' economic integration: empirical evidence from France. *Eur. Econ. Rev.* 113, 265–296.
- Mangrio, E., Sjögren Forss, K., 2017. Refugees' experiences of healthcare in the host country: a scoping review. *BMC Health Serv. Res.* 17, 1–16.
- Marbach, M., Hainmueller, J., Hangartner, D., 2018. The long-term impact of employment bans on the economic integration of refugees. *Sci. Adv.* 4 (9), eaap9519.
- McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: a density test. *J. Econom.* 142 (2), 698–714.
- McLeod, M., 2023. English-language proficiency requirements for migration to Canada, Australia, the United Kingdom, and the United States, and the implications for language testing research. *Can. J. Appl. Ling.* 26 (1), 1–27.
- Mishori, R., Aleinikoff, S., Davis, D., 2017. Primary care for refugees: challenges and opportunities. *Am. Fam. Physician* 96 (2), 112–120.
- Narayan, L., 2013. Addressing language barriers to healthcare in India. *National Med J India* 26 (4), 236–238.
- Obregon, E., Martin, C.R., Frantz III, I.D., Patel, P., Smith, V.C., 2019. Neonatal intensive care unit discharge preparedness among families with limited English proficiency. *J. Perinatol.* 39 (1), 135–142.
- OECD, 2016. Perspectives on Global Development 2017 International Migration in a Shifting World. OECD Publishing.
- OECD, 2023. Indicators of Immigrant Integration: Overview and Challenges.
- Palau, M.A., Meier, M.R., Brinton, J.T., Hwang, S.S., Roosevelt, G.E., Parker, T.A., 2019. The impact of parental primary language on communication in the neonatal intensive care unit. *J. Perinatol.* 39 (2), 307–313.
- Pei, Z., Lee, D.S., Card, D., Weber, A., 2022. Local polynomial order in regression discontinuity designs. *J. Bus. Econ. Stat.* 40 (3), 1259–1267.
- Pont-Grau, A., Lei, Y.-H., Lim, J.Z., Xia, X., 2023. The effect of language training on immigrants' integration: does the duration of training matter? *J. Econ. Behav. Organ.* 212, 160–198.
- Pruckner, G.J., Schneeweis, N., Schober, T., Zweimüller, M., 2021. Birth order, parental health investment, and health in childhood. *J. Health Econ.* 76, 102426.
- Rocca, L., Carlsen, C.H., Deygers, B., 2020. Linguistic Integration of Adult Migrants: Requirements and Learning Opportunities. Council of Europe Publication, Strasbourg.
- Røed, M., Schone, P., Umblijs, J., 2019. Flyktingers deltagelse i arbeid og utdanning, før og etter introduksjonsordningen. Rapport—Institutt for samfunnsforskning. <http://hdl.handle.net/11250/2597916>.
- Romano, J.P., Wolf, M., 2005. Exact and approximate stepdown methods for multiple hypothesis testing. *J. Am. Stat. Assoc.* 100 (469), 94–108.
- Ruiz, I., Vargas-Silva, C., 2018. Differences in labour market outcomes between natives, refugees and other migrants in the UK. *J. Econ. Geogr.* 18 (4), 855–885.
- Salti, N., Chaaban, J., Moussa, W., Irani, A., Al Mokdad, R., Jamaluddine, Z., Ghattas, H., 2022. The impact of cash transfers on Syrian refugees in Lebanon: evidence from a multidimensional regression discontinuity design. *J. Dev. Econ.* 155, 102803.
- Sarvimäki, M., Hämäläinen, K., 2016. Integrating immigrants: the impact of restructuring active labor market programs. *J. Labor Econ.* 34 (2), 479–508.
- Satinsky, E., Fuhr, D.C., Woodward, A., Sondorp, E., Roberts, B., 2019. Mental health care utilisation and access among refugees and asylum seekers in Europe: a systematic review. *Health Policy* 123 (9), 851–863.
- Schmid, L., 2023. The impact of host language proficiency on migrants' employment outcomes. *Am. Econ. Rev.: Insights* 5 (4), 477–493.
- Silove, D., Ventevogel, P., Rees, S., 2017. The contemporary refugee crisis: an overview of mental health challenges. *World Psychiatry* 16 (2), 130–139.
- Smith, J.P., 2009. The impact of childhood health on adult labor market outcomes. *Rev. Econ. Stat.* 91 (3), 478–489.
- Statistics Norway, 2022. Living conditions survey among people with an immigrant background, 2016 (version 5). sikt - knowledge sector service provider [Dataset].
- Tayfur, I., Günaydin, M., Suner, S., 2019. Healthcare service access and utilization among Syrian refugees in Turkey. *Annals of Global Health* 85 (1).
- UNHCR, 2021. Global report 2021: The stories behind the numbers: Technical report. United Nations High Commissioner for Refugees.
- UNHCR, 2024a. Operational data portal - Ukraine refugee situation. United Nations High Commissioner for Refugees.
- UNHCR, 2024b. Refugee population statistics Database. United Nations High Commissioner for Refugees.
- van den Berg, G.J., Paul, A., Reinhold, S., 2020. Economic conditions and the health of newborns: evidence from comprehensive register data. *Labour Econ.* 63, 101795.
- von Hinke, S., Rice, N., Tominey, E., 2022. Mental health around pregnancy and child development from early childhood to adolescence. *Labour Econ.* 78, 102245.
- Vr̄alstad, S., Wiggen, K.S., 2017. Levek̄ar blant innvandrere i Norge 2016. Technical report, Statistisk sentralbyrå.
- WHO, 2019. United Nations Agency Briefs: Responding to the Challenge of Non-communicable Diseases. World Health Organization.
- World Health Organization, 2018. Report on the Health of Refugees and Migrants in the WHO European Region: No Public Health without Refugee and Migrant Health. World Health Organization, Regional Office for Europe. Technical report.
- World Health Organization, 2022. World Report on the Health of Refugees and Migrants: Summary. World Health Organization.
- World Health Organization, 2023. Promoting the Health of Refugees and Migrants: Experiences from Around the World. World Health Organization.
- Yun, K., Hebrank, K., Graber, L.K., Sullivan, M.-C., Chen, I., Gupta, J., 2012. High prevalence of chronic non-communicable conditions among adult refugees: implications for practice and policy. *J. Community Health* 37, 1110–1118.
- Zihindula, G., Meyer-Weitz, A., Akintola, O., 2015. Access to health care services by refugees in southern Africa: a review of literature. *Southern African Journal of Demography* 16 (1), 7–35.