Abstract: In the current literature, the positive associations between COVID-19 incidence rates and migrant populations are suggested. Within this research, this relationship is tested controlling for social determinants of health variables. It utilizes a logistic regression analysis whereby COVID-19 incidences are its dependent variable and gender, age, income, household size, and educational attainment its independent. Data provided for this analysis is derived from the Survey of Health, Ageing and Retirement in Europe and consists of respondents 50 years or older. There are 8104 valid cases based in 27 countries within Europe and Israel. In conclusion, there was no significant relationship between migrant background and the COVID-19 incidence rate observed. However, household size and total monthly average household income indicate a significant relationship in regard to COVID-19 incidences. Respectively, in the household size variable for each additional member of the household, the odds of testing positive on a COVID-19 test go up by 15.6%. For the total monthly average income in the household, it was found that high income is associated with an odds decrease of testing positive for covid-19 by 36,6%. These results indicate that other factors, than migrant background, are influential in determining COVID-19 incidence rates. These test results suggest that in migrant populations lower incomes and larger households the have higher odds of testing positive for COVID-19. Policies and interventions for future pandemics can be designed in order to tailor towards individuals with large household sizes and low-income households to better protect them from potential incidence rates.

university of groningen / faculty of spatial sciences / population research centre

How does having a migrant background influence the incidence levels of COVID-19 when controlling for gender, age, income, household size, and educational attainment?

Ake Jol van Benthem

S4570499

14-06-2024

Course: Bachelor's Project HGP

Supervisor: Adrien Remund

Introduction

In March 2020 the COVID-19 outbreak struck the world. As time passed research has explored the association between hospital admissions and COVID-19 incidences and suggests that they are not equally distributed across demographic groups. In this paper, the definition of migrant used is constructed by the United Nations (1998), which states that a migrant is someone who has or changes their usual residence for a different country. In a meta-analysis of 68 studies based in the United States researchers found a positive correlation between ethnic background and an increased positive testing on COVID-19 compared to non-migrant background. (Magesh et al., 2021). Similarly, a study in Norway suggests a higher incidence rate in migrant populations compared to non-migrants. (Indseth et al., 2021). In academia and beyond COVID-19, multitudes of studies have shown certain negative health disparities in migrant populations compared to the origin population. (Castañeda et al.,2015). Furthermore, a possible explanation for migrant health disparities could be the result of large household size, inefficient knowledge of the host language, low income, poor access to health care services and insufficient rates of health insurance. Consequently, they are less able to sustain good health and healthcare. (Feinberg, et al. 2021.) Furthermore, Social Determinants of Health frameworks could help in identifying the effect people's living situations have on their health. (Marmot, 2005). WHO (2008) defines social determinants of health as a combination of conditions of daily life and structural conditions. For instance, education, home situation, conditions of work, and access to health care in combination with social policy, politics and economic positions constitute social determinants of health. Social determinants of health have been studied in relation to covid-19 incidence rates, For example, In regard to income Baena et al (2020) suggest that lower income is linked with higher rates of Covid-19 incidence. Next, in regard to household size de Lusignan et al (2020) found that in households with two to eight people, the odds of a positive test result were lower than in households with one person. Lastly, Hawkins et al., (2020) suggest that low educational level is associated with higher rates of covid-19 incidence. The importance of investigating the incidence rates is emphasized by an example in the Netherlands. The head of the intensive care of the Amsterdam UMC noted that journalists did not want to report on the relatively higher risk of COVID-19 infection in migrant populations because of fear of stigmatization. Even though he and his colleagues at other hospitals saw the same correlation. (Start, 2021). On a different note, the WHO (2021) underlines the importance of coordinated responses to pandemics like COVID-19 for future policies. Whilst taking specific account of the situation of migrants, since they tend to be more exposed to the virus due to relatively poor social determinants of health in comparison with non-migrant populations. In academia, Krieger (2014) argues that within the existing literature on discrimination and health, there exists a large focus on biological determinants in individuals like the genetic cause of disease that translates itself to group rates of diseases. Therefore ignoring the social determinants of health. Additionally, Kapilashrami (2015) argues that researchers should expand their focus on an intersectional perspective that incorporates variables like gender and ethnicity in contrast to the emphasis on the socio-economic status of individuals. Similarly, Ingleby (2012) argues for an increased focus on gender, ethnic group and migrant status as categorical variables in contrast to one variable of socioeconomic status in relation to health.

This research aims to investigate the link between the presence of being a migrant and incidence levels of COVID-19 controlling for social determinant variables of health. Therefore the main research question is: What is the relationship between migrant background and COVID-19 infection rate whilst controlling for gender, age, income, household size and educational attainment? The secondary questions in this research entail:

1. Does an increased household size influence COVID-19 infection rates in migrant families?
2. Does the relationship between migrant background and COVID-19 infection rates vary across lower incomes and higher incomes?
3. Does lower educational attainment influence COVID-19 infection rates in migrants?

Theoretical framework

Theory regarding social determinants of health

The theory of "Fundamental Causes" is a theory developed by Link & Phelan (1995) that helps to explain health inequalities within society. It states that socioeconomic variables are able to function as a fundamental cause of diseases. Furthermore, this theory has its holding in the following reasoning: firstly it can have an effect on multiple diseases, secondly, access to resources can help people mitigate or evade diseases, thirdly it can affect multiple risk factors diseases and finally, fundamental causes of health are persistent over time in face of new mechanisms. Collecting from this theory the main takeaway could be the fact that despite advancements in technology some groups in society are "fundamentally" disadvantaged in their health due to an unequal distribution of resources socially and economically. Following this, a hypothesis can be derived from the theory, namely, that health outcomes vary by ethnicity. (Frishman, 2012). Furthermore, 'the social network theory' plays a role as a potential determinant of health in migrant populations. This theory proposes that the individual's behaviour is significantly explainable by the social structure in which the individual is situated. These social structures shape resources or information for these individuals. (Berkman, 2014). The overarching 'social disadvantage approach' might work as a framework to summarize social determinants of health. This approach is followed by researchers that establish that factors like, income, occupational grade, and educational level are linked with health outcomes across the life course. (Bharmal et al., 2015).

Migrant background

In existing literature, it has been found that migrant background made it more likely for individuals to be admitted to hospital, contract or pass away in regards to COVID-19. (Agyemang et al, 2021). Furthermore, within the Netherlands when identifying COVID-19 incidence amongst the migrant population during the first and second wave of infections it was discovered that especially during the second wave incidence was significantly higher in ethnic minority groups in comparison to those of Dutch origin. (Coyer et al, 2021.) On the other hand, the comparison of incidence and mortality rates between migrant populations and native populations varies significantly in regards to the disease studied. (Lanting et al, 2008; Rechel et al, 2013). Following this, a study in Italy on health inequalities found that the migrant population had higher hospitalization rates for some diseases and infections, but generally not in more cases of different diseases than for the host population. They contributed to the health inequalities for the positively correlated diseases mainly towards lower income, language difficulties and poor awareness of health status. (De Waure et al., 2015)

Income

Within contemporary society, migrants might face difficulties in obtaining the same level of income level as non-migrants. There are multiple explanations concerning this difficulty. First of all the employment market might hinder them. Barriers include discrimination which could reveal itself by not getting invited for job applications. Furthermore, language could put them in a disadvantaged position when employers require them to have a thorough understanding of the language of residence. Additionally, they obtained degrees gathered in their place of origin might not be recognised by the institutions of the place of residence. (Wrench, 2015) The factors mentioned above could contribute to disparities when migrants are forced to work in jobs that relatively speaking often pay less because they do not require extensive knowledge of language, do not require an application and do not require a recognized degree. (Careja, 2019) In literature, Gravelle & Sutton (2009) test whether individuals' health could be determined by their relative income in a given reference group this is also called “the relative income hypothesis”. Longitudinal data from 1979-2000 showed income equality had risen in the UK and self-assessed health decreased. Furthermore, one model suggested that the probability of good health increases when going from the lower to the upper quartiles in income.

In relation to COVID-19, a study in the United States exhibited that relative average income was a factor that could contribute to lower cases and mortality in COVID-19. They found that in counties with higher average wealth the incidences and mortality rates of COVID-19 were lower.(Amate-Fortes et al, 2023). In contrast, a study in Brazil concluded that incidence rates were higher in high-income areas. When assessing neighbourhoods in Rio de Janeiro on per capita income they noted that incidence was higher in the wealthiest neighbourhoods. They noted a possible explanation that constituted this finding towards the accessibility of testing facilities, which are more present in high-income neighbourhoods. (Rafael et al., 2020) For example, in Finland, researchers found that having low income in a household is a strong indicator of the severity and incidence of COVID-19. Furthermore, the differences in income between households are exacerbated if the household has a migrant background. (Saarinen et al, 2022).

Gender

There are similarities in the exposure rate to diseases between migrant females and males. (Anderson et al., 2021) Males more often work in essential jobs for instance as construction workers, factory workers or in farming. Females more often work in essential jobs like nursing, domestic workers or in teaching. (European Commission, 2009) In a systematic review on gender equality and health, researchers found that the difference in health effects varies according to the health outcome discussed in the papers and the control variables used. (King et al, 2020). In the case of COVID-19, there are several studies that suggest that there are disparities among women and men. For instance, in Pakistan, researchers found that men had a higher incidence rate of COVID-19 compared to women. These were specifically 3.6% of the population for women and 6.7 % for men. They noted that a multitude of factors influenced this difference like psychological factors, biological factors, smoking and lifestyle. (Waris et al, 2020) On the other hand, in a cross-national study Kim et al, (2021) found that within more than 100 countries males were disadvantaged when it came to incidence rates and mortality. Interestingly in countries where gender equality to a greater extent was observed the incidence and mortality rates were, additionally, more equal across men and women. This

last finding ties in with the so-called ‘convergence hypothesis’. This hypothesis states that an increase in gender-equality is positively correlated with less diverging health outcomes. Following this men and women experience a decrease in the gap in health outcomes because behaviours and exposures become more similar. (Backhans et al., 2007).

Age

In literature, researchers hypothesise about the “Healthy migrant effect”. In short, this entails the belief that low mortality, in some migrant groups compared to the host population, can be explained by the fact that strong, healthy and often young migrants at the time of migration. (Ichou et al., 2019). One explanation for this phenomenon is the “salmon bias”, which states that immigrants who are seriously ill return to their place of origin. (Razum et al., 2000). However, over time the healthy migrant effect might diminish when migrants are exposed to longitudinal stress due to a marginalized position in society. Literature suggests that ageing is associated with an increase in health problems. (Saß, 2010). Similarly, WHO (2015) explores the hypothesis that older people are more likely by at risk for developing or experiencing more than one chronic condition simultaneously. “Multimorbidity”, as it is known, is a risk factor in the sense that the interaction between different health conditions might be of influence on the trajectories of contracted diseases. Regarding COVID-19 a study in 19 states in the USA indicated that the incidence rate amongst young adults was significantly higher than in older adults. (Schneiderman et al., 2022) This finding is in contrast with an earlier study done in China. In this research, Zhang et al., (2020) noted that children aged 0-14 had a lower chance of COVID-19 incidence than adults aged 15-64. Furthermore, adults aged above 65 were more likely to be to be infected with COVID-19.

Household size

Migrant populations often live in larger households than non-migrant populations. This has several reasons. Firstly, financial constraints limit them from possessing multiple houses with the entire family. Secondly, multiple generations of the family often live in one house. (Jia et al., 2023) Lastly, culturally speaking leaving the parental home is often proceeded later in life in migrant populations. (Adugna, 2019). In existing literature, there is research that suggests that household size is linked to increased transmission count. For instance, a study by Marathe et al., (2011) investigated the different outcomes in two scenarios with one infected person in contact with the whole household and one infected person in contract with a caregiver, who then in turn is in contact with the entire household. They found that in both household scenarios, the amount of cases increases almost linearly with an increase in household size, especially for the moderate flu. In the case of COVID-19, a study conducted in the United States showed an increased risk of contracting COVID-19 for persons in large households. Additionally, they found that persons with low English proficiency were more likely to live in larger households than vice versa. (Giglio et al., 2022) In a similar fashion Liu et al., (2021) argue that individuals who reside in large households have a higher chance of getting infected with COVID-19. Furthermore, they found that measures to prevent COVID-19 like quarantine and social-distancing are more effective in large households in preventing infection rates. Lastly, a study in the UK compared single households, double households or three or more inhabitants with each other. They found that for a non-severe COVID-19 infection, single-person households had lower odds than for a double-person household. On the other hand,

severe COVID-19 odds were the highest in single-person households and three or more households. (Gillies et al., 2022)

Educational attainment

Within the educational system migrant populations might face difficulties attaining high degrees. There are different factors in play here. Per example: limited financial support to attend school or low host language proficiency. (Kaukko et al., 2022). These factors could contribute to lower higher education participation in migrant populations. Lower educational attainment is linked with work that exposes one more to health risks. (Hawkins et al., 2020). Literature suggests that adults with higher educational attainment are linked with longer and healthier lives than adults with lower educational attainment. (Zajacova, 2018) Similarly, Raghupathi & Raghupathi (2020) argue that higher educational attainment correlates with better health and lifespan. Additionally, they stress the fact that education is grounded in the Fundamental Cause Theory since education can improve abilities and resources that contribute to better health. In the case of COVID-19 Abedi et al. (2021) found that in counties with higher educational levels, amongst other variables like higher income and total population, the risk of COVID-19 infection was higher. In contrast, a study in the United States found that lower educational attainment was linked with higher incidences of COVID-19. (Hawkins et al., 2020). They noted that this could be due to essential work in the field being relatively low-paying and without a requirement for an advanced degree. Similarly, another study in the United States found that counties that had adults with less than a high school diploma in combination with a high percentage wise black Americans experienced more COVID-19 cases. (Khanijahani, 2021).

Within this conceptual model, the line in bold between migrant background (independent variable) and COVID-19 incidence (dependent variable) indicates the main effect this research tries to investigate. The lines coming from migrant background through Gender, Age, Household Income, Household Size and Educational Attainment towards COVID-19 incidence indicate the covariates that have a potential mediating effect on the main effect. These variables have been reported in different academic research that control for socioeconomic variables. As the “Fundamental Causes” suggests socioeconomic variables are able to function as a fundamental cause of diseases. In this research with the aid of the ‘social disadvantage approach’ the covariates in this conceptual model are identified as potential variables influencing the relationship between migrant background and COVID-19 incidence.

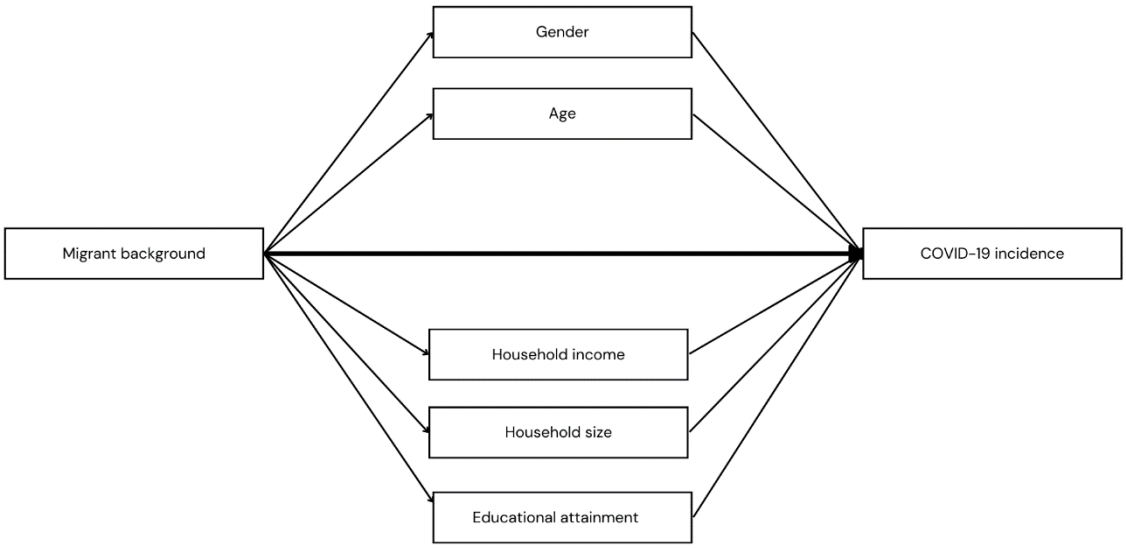

Figure 1: Conceptual model. (Constructed by author, 2024)

Methodology

Background data set

The data used in this research is extracted from the SHARE survey of HEALTH, AGING AND RETIREMENT IN EUROPE. SHARE is a multidisciplinary panel database that stretches across multiple countries in Europe. The data collected in this database consists of microdata on socio-economic variables, family and social networks and health. Their sample is derived from individuals aged 50 or older. Therefore their main objective is to discern the conditions of aging populations in Europe to inform policy makers on the characteristics/trends in the aging population of Europe. The data is collected every year and the base year is 2004. The main method of research gathering is face-to-face interviews. However during covid-19, due to restrictions, researchers resorted to the use of a telephone-administered interview (CATI) methodology in contrast to the computer-assisted personal interviewing (CAPI) method used in the previous waves of questionnaires. The complete dataset of SHARE consists of nine waves of questionnaires with an additional two corona-specific surveys. Furthermore, these corona-specific surveys contain a shortened list of existing variables from the original health data collection set whilst focusing on the pandemic and the living conditions of respondents. There are two versions of the questionnaires: one for respondents who will undertake the SHARE interview for the first time and the longitudinal questionnaire that has been designed for respondents who have done the interview before. This longitudinal questionnaire is designed because this way questions regarding education and country of birth are only asked once, therefore if you would want to investigate that in one wave, you would have to use the version of the questionnaire data that is filled in by the

first time respondents. Lastly, SHARE has received funding from the European Union and is supported by the German Federal Ministry of Education and Research.

Operationalization

The questionnaires of SHARE are divided into thematic blocks. Deriving from these thematic blocks there are 7 variables related to this research these are firstly, Male or Female (Gender) to determine the gender of the respondent. Secondly, Year of birth (yrbirth) to determine the age of the respondent. Thirdly, country of birth (dn004_) to establish whether the respondent was born in the country of interview and therefore falls under the description of migrant used in this research. Fourthly, Total household net income on a monthly basis (HH017e) to determine the income of the respondents household. Fifthly, Household size (hhsize) to ascertain the size of the respondents household. Sixthly, Highest school degree obtained (dn010_) to determine the educational attainment of the respondent. Lastly, COVID-19: respondent tested positive (cac005_01) to establish whether the respondent was infected with COVID-19 or not.

Data analysis method

This research makes use of a quantitative analysis method in the form of a logistic regression. The variables in this research are social determinants of health that are designed into survey questions, which are suitable for regression analysis. These variables are used in a logistic regression model to determine the association between the dependent variable and the independent variable controlling for the covariates. Firstly, (cac005_01) is utilized as a dependent variable. Within this question there is a yes or a no, therefore this variable is binary and consequently makes the entire method of analysis into logistic regression. Secondly, the independent variable in this research is (dn004_), which has a similar coding with yes or no. The remaining variables are covariates that are added after the initial relationship between the dependent variable and the independent variable is made to determine the statistical explanatory power of the main effect. The variables are entered in the following blocks: block 1: independent variable and dependent variable *(*cac005_01) and. Block 2: addition of basic demographic variables (yrbirth) and (Gender). Block 3: addition of social determinants of health variables (HH017e), (dn010_) and (hhsize) to (dn004_).

Data handling

The regression has 8104 valid cases based in 27 countries within Europe and Israel. Furthermore, there are a lot of missing cases because of multiple reasons like, respondents passing away, respondents not being listed as household members in certain waves, households not participating in certain waves, specific follow-up questions, individuals moving outside of Europe and possible incarnation of respondents. A logistic regression requires all variables to be valid in order to include them in the analysis. For this reason, this research used a listwise method of deleting cases, which could lead to less statistical explanatory power. Upon doing analysis the missing values were deleted. This method is

chosen above imputation.. However, the data set is still very valuable since it incorporates very specific microdata collected from thousands of respondents. Therefore the large sample size could compensate for its missing values. Nonetheless, I am aware of the consequences concerning potentially biased estimates and results upon choosing this method.

Certain variables have been recoded to accommodate analysis. First of all, (dn010_) has a country-specific answer construction which I recorded to low and high educational attainment. The border between low and high was established by looking at the mean value in the dataset, which constituted 6,705454. Therefore the decision has been made to indicate low education with 1 through 7 and high education with 8 through 25. Furthermore, the values for none and still in school have been put in the low category. In a similar fashion, (HH017e), has been recoded in low and high income. The border has been established by looking at the mean income level, which is 1339.817993. Similarly, low income is recorded as values 0 through 1339.817993 and high income as values 1340 through 134086,002762. On a different note, (yrbirth) has values that are lower than the age of year 50 due to the fact that the SHARE allowed new partners of its respondents to participate in the questionnaire, regardless of age. In order to not end up with outliers and anomalies in this variable, I choose to set a minimum age of 50 years old in this variable.

As mentioned earlier, the total number of waves is 9 with 2 additional Corona surveys. Within the database of SHARE an all-waves file with all valid cases on (yrbirth) and (gender) is present. Consequently, the 11 waves are merged with this all-waves file based on Mergeid as the common merge variable in order to get full coverage and potential for answers. This variable is unique for every respondent and does not change in any wave. Step-wise the variables, other than (yrbirth) and (gender), used in the regression has been filled by adding cases. The data was not merged with wave 3. This is due to the fact that waves 3 and 7 were part of the SHARELIFE data set. These waves were separate questionnaires that aimed to investigate the history in life stories of the respondents. Wave 7 does contain the CV_R, HH and DN thematic blocks that contain the relevant variables, which wave 3 does not have.

Data protection and ethical considerations

Based on confidentiality rules my username and password are hidden from vulnerable places and the data is stored in personal x- drives that are on the network of the university of Groningen, which is a protected network that is only accessible by myself. Since data in this data set has factual anonymity the only purpose of its usage, allowed by SHARE, is that of a scientific nature. The data presented additionally has no intrinsic identifying information about its participants. As I am studying at a university my purpose is to engage in science by means of the scientific method, which is utilized in this paper. The data collected by SHARE is subdued to the Respect Code of Practice for Socio-Economic Research, which is a set of ethical guidelines on how to collect data for Socio-Economic Research and the 'Declaration of Helsinki' which is a set of ethical principles guiding the medical community when engaging in human experimentation. As a user of the data by SHARE I have studied the Respect Code of Practice for Socio-Economic Research and the 'Declaration of Helsinki' upon doing research.

Therefore I practice the following principles: being compliant with the law, upholding scientific standards and avoiding social and personal harm. Furthermore, I respect the privacy of the respondents, my purpose solely is to examine relationships in order to inform policymakers and academics, the choice of researching migrant populations is underpinned by a large body of research therefore indicating its priority and finally conclusions made with this research should be received with extreme caution regarding generalization and validity.

Results

Block 1: Main effect between depended variable and the independent variable. (Constructed by author, 2024)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Born in the country of interview	-,168	,099	2,895	1	,089	,846
	Constant	-1,400	,094	223,366	1	<,001	,247

a. Variable(s) entered on step 1: Born in the country of interview.

Within this analysis, 8104 valid cases were included from which 6688 selected not and 1416 selected yes to a positive test on COVID-19. Overall there is no significant relationship between Born in the country of interview and tested positive for COVID-19 as it does not go under the conventional 5% threshold. (P,089 > P ,05). Additionally, the odds of testing positive on a COVID-19 test decrease by a factor of 0.846 when the respondent is born in the country of interview. Specifically, being born in the country of the interview is associated with a 15,6% reduction in the odds ratio of testing positive on a COVID-19 test. The insignificant result could still be partially explained by the “Salmon Bias” described by (Razum et al., 2000), since this data set is based on people aged 50 years or older the likelihood of a severe COVID-19 case could be increased. This could lead to less reporting of COVID-19 if they left their country of residence for their country of origin.

Block 2: Relationship between depended variable, independent variable, age and gender. (Constructed by author, 2024)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Born in the country of interview	-,176	,099	3,156	1	,076	,839
	Year of birth	,027	,003	62,833	1	<,001	1,028
	Male or female	,013	,061	,043	1	,835	1,013
	Constant	-54,981	6,761	66,128	1	<,001	,000

a. Variable(s) entered on step 1: Born in the country of interview, Year of birth, Male or female.

The inclusion of year of birth and male or female in the model has slightly altered the main dependent variable. The difference is seen in a stronger negative impact on the odds, however, the significance has not dropped below the critical a-level of 0,05 therefore making it not a statistically significant result. On the other hand, the year of birth is statistically significant. (P ,001 < P,05) Furthermore, for each year added in year of birth the odds of a positive covid-19 increase by 2,8%, which entails that adults closer to 50 years old tend to answer more yes on positive covid-19 tests. Lastly, the male or female variable is not significant (p,,835 > P,05) and shows a minimal coefficient indicating that the odds ratio indicates no substantial difference between males and females. The significant result in year of birth variable and the corresponding coefficient is in accordance with research conclusions by Schneiderman et al. (2022.). However is in contrast with research conclusions by Zhang et al., (2020). A possible explanation for the significant result could be the effect “Multimorbidity”, posed by WHO (2015). Elderly people might be less likely to attend big gatherings or engage in large groups when posed with underlying health conditions that could exacerbate the effect of COVID-19. However in migrant populations due to multigenerational living, this effect might not be that strong if an entire family could transmit COVID-19 to the elderly living in the household.

Block 3: Relationship between depended variable and all other covariates. (Constructed by author, 2024)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Born in the country of interview	-,154	,100	2,397	1	,122	,857
	Year of birth	,023	,004	40,063	1	<,001	1,023
	Male or female	-,014	,062	,051	1	,821	,986
	Household size	,145	,026	29,943	1	<,001	1,156
	Highest school degree obtained	,081	,070	1,356	1	,244	1,085
	Total income received by all hh members an average month last year	-,455	,061	54,841	1	<,001	,634
	Constant	-46,369	7,112	42,510	1	<,001	,000

a. Variable(s) entered on step 1: Born in the country of interview, Year of birth, Male or female, Household size, Highest school degree obtained, Total income received by all hh members an average month last year.

The inclusion of household size, highest high school degree obtained and average household income monthly have made the main independent variable less significant compared to blocks 1 and 2 indicated. Therefore you could suggest that the inclusion of the latest covariates could function as potential mediators that have explanatory power in the relationship between COVID-19 incidence and migrant background. Additionally, household size has a significant test result (P,001 < P,05). Interpreting the EXP(B) gives that for each additional member of the household the odds of testing positive on a COVID-19 test go up by 15.6%. Furthermore, the total monthly income of a given household gives a statistically significant result. (P,001 < P,05). High income is associated with an odds decrease of testing positive for COVID-19 of 36,6% indicating that higher incomes tend to test in less quantity positively on COVID-19. As mentioned earlier, year of birth is significant (P,001 < P,05) and suggests that the odds of a positive COVID-19 test increase by 2,3% for every year subtracted towards 50 from the oldest age in the data set, which is 100.

The significant result in in household size and corresponding coefficient is in accordance with research conclusions by Giglio et al., (2022) and Liu et al., (2021). A possible explanation could be the fact that transmission count is linked to household size. (Marathe et al., 2011). Transmission was seen to linearly increase with an increase in household size even accounting for the transmission by just one caregiver. Additionally, in migrant populations, the effect might be exacerbated due to cultural, financial and multigenerational living. This could be in line with the associations between educational attainment and job prospects. Since migrants struggle with getting higher education they tend to occupy more jobs in essential sectors that continued in the pandemic. In contrast, higher-educated jobs tend to be more in

the service sector, which could be done remotely. Their professions influence their exposure rate and could lead to more COVID-19 cases in households.

The significant result in household monthly income and the corresponding coefficient is in accordance with research conclusions by Amate-Fortes et al, 2023). and Saarinen et al., (2022). A possible explanation could be that in line with "the relative income hypothesis" relative to high-income, low-income households have less financial means to live in more spacious homes with fewer people. This income effect is definitely amplified in migrant populations. Migrants tend to struggle with obtaining higher levels of income in the labour market due to discrimination, language proficiency and unrecognised degrees. Consequently, this could lead to a disproportional amount of migrants working in exposure risk lower-paying jobs that are often in essential sectors, like construction, retail and healthcare.

Interaction effects

An exploratory analysis on interaction effects identified an interaction effect with a migrant background, household size tested significantly with a P-value of,026 which is below the critical p-value of ,05. Furthermore, this indicates that the effect of household size on positive COVID-19 test results varies depending on migrant background status. However, the effect is noted to be minimal through the Partial Eta Squared test result being,002. Nonetheless, it is necessary to employ an AIC and BIC analysis for model fit to determine the trade-off between goodness of fit and model complexity. Following this two goodness-of-fit tables are produced. The first model consists of the dependent variable and the independent variables while the second model includes the interaction term of household size * Born in the country of interview. What comes apparent from comparing model 2 with model 1 is that for both the AIC (3106,672 < 3162,109) and the BIC (3155,673 < 3652,117) the values are lower in model 1. Notably, the difference between models 1 and 2 in regards to BIC is larger than in the AIC. This could be the result of the way BIC is calculated since the BIC penalizes harder for model complexity. With the addition of this interaction effect, the model gets more saturated. Therefore the addition of model complexity does not outweigh the potential implications derived from the interaction effect. In conclusion, model 1 has lower values for AIC and BIC and therefore fits the data better.

Tests of Between-Subjects Effects

Dependent Variable: COVID-19: respondent tested positive

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	10,739[a]	16	,671	4,688	<,001	,009
Intercept	8,488	1	8,488	59,282	<,001	,007
dn004_	,707	1	,707	4,936	,026	,001
hhsize	6,355	9	,706	4,931	<,001	,005
dn004_ * hhsize	2,055	6	,342	2,392	,026	,002
Error	1157,846	8087	,143			
Total	1416,000	8104				
Corrected Total	1168,584	8103				

a. R Squared = ,009 (Adjusted R Squared = ,007)

Table 1: Interaction effect between household size and migrant background. (Constructed by author, 2024.)

Goodness of Fit[a]

	Value	df	Value/df
Deviance	1741,306	1556	1,119
Scaled Deviance	1741,306	1556	
Pearson Chi-Square	1651,874	1556	1,062
Scaled Pearson Chi-Square	1651,874	1556	
Log Likelihood[b]	-1511,054		
Akaike's Information Criterion (AIC)	3162,109		
Finite Sample Corrected AIC (AICC)	3163,346		
Bayesian Information Criterion (BIC)	3652,117		
Consistent AIC (CAIC)	3722,117		

Dependent Variable: COVID-19: respondent tested positive
Model: (Intercept), Male or female, Year of birth, Household size, Born in the country of interview, Highest school degree obtained, Total income received by all hh members an average month last year

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Model 2: Goodness of fit with the dependent variable and the independent variables with interaction effect (Constructed by author, 2024.)

Goodness of Fit[a]

	Value	df	Value/df
Deviance	1811,869	1619	1,119
Scaled Deviance	1811,869	1619	
Pearson Chi-Square	1761,864	1619	1,088
Scaled Pearson Chi-Square	1761,864	1619	
Log Likelihood[b]	-1546,336		
Akaike's Information Criterion (AIC)	3106,672		
Finite Sample Corrected AIC (AICC)	3106,686		
Bayesian Information Criterion (BIC)	3155,673		
Consistent AIC (CAIC)	3162,673		

Dependent Variable: COVID-19: respondent tested positive
Model: (Intercept), Male or female, Year of birth, Household size, Born in the country of interview, Highest school degree obtained, Total income received by all hh members an average month last year

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Model 1: Goodness of fit with the dependent variable and the independent variables (Constructed by author, 2024.)

Differences between logistic regression test results for non-migrant sample selection and migrant sample selection.

To determine differences between migrant and non-migrant samples it is possible to identify differences in both groups by comparing their individual logistic test results. In table 2 output the individuals who responded yes in the variable Born in the country of interview are selected. There are 7386 respondents in this analysis, of which 6112 selected not and 1274 selected yes to a positive test for COVID-19. Table 3 describes the individuals who responded no to Born in the country of an interview with 718 respondents, which had 575 selected not and 114 selected yes to a positive test for COVID-19. The most interesting differences are as follows: Firstly, when assessing significant results in both sample selections, the non-migrant sample selection shows in the year of birth variable a significant P value of,001. However, in the migrant selection sample, this is not significant with a P value of ,521. This indicates that in the migrant selection age does not have a significant statistical effect on whether a respondent answers yes on positive COVID-19. As mentioned earlier the “Salmon Bias” could still be an explanation as you would see fewer reported cases of COVID-19 if people adults 50 or older fell really ill, consequently returning to their destination of origin for care. Secondly, the sign of the highest school degree obtained changes when going from table 2 to

table 3. In the non-migrant sample selection, the odds of a positive COVID-19 test increase 10,2% when having higher education, in contrast, the result is not statistically significant. On the other hand, in the migrant selection has a coefficient of -0,92 indicating a minor negative relationship between a higher high school degree obtained and the log odds of testing positive on a COVID-19 test. Similarly, in the gender variable, there is a positive sign change going from table 2 to table 3. In the non-migrant sample selection the odds of a positive COVID-19 test decrease by 3,3% when being a woman. On the contrary, In the migrant sample selection the odds of a positive COVID-19 test increase by 21,8% when being a woman. The results in gender and education variables are however not statistically significant. In conclusion, the only significant difference between the two sample groups is the fact that age has a significant effect on covid-19 positive test results in the non-migrant sample group compared to its migrant counterpart.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Year of birth	,025	,004	41,101	1	<,001	1,025
	Male or female	-,034	,065	,274	1	,600	,967
	Household size	,131	,028	22,113	1	<,001	1,140
	Highest school degree obtained	,097	,073	1,753	1	,185	1,102
	Total income received by all hh members an average month last year	-,451	,064	49,063	1	<,001	,637
	Constant	-49,528	7,495	43,662	1	<,001	,000

a. Variable(s) entered on step 1: Year of birth, Male or female, Household size, Highest school degree obtained, Total income received by all hh members an average month last year.

Table 2: Logistic regression output individuals who responded yes in the variable Born in the country of interview. (Constructed by author, 2024.)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Year of birth	,008	,012	,412	1	,521	1,008
	Male or female	,198	,206	,919	1	,338	1,218
	Household size	,289	,087	10,998	1	<,001	1,336
	Highest school degree obtained	-,092	,232	,157	1	,691	,912
	Total income received by all hh members an average month last year	-,503	,208	5,866	1	,015	,605
	Constant	-16,310	22,858	,509	1	,476	,000

a. Variable(s) entered on step 1: Year of birth, Male or female, Household size, Highest school degree obtained, Total income received by all hh members an average month last year.

Table 3: Logistic regression output individuals who responded no in the variable Born in the country of interview. (Constructed by author, 2024.)

Discussion

In this research gender, age, income, household size and educational attainment are used as covariates that could contribute to the main effect, however in theory there could be more variables that influence this relationship. Therefore this model might not capture reality as well as it could.

Within the data handling, assigning borders to educational attainment and income based on the mean might not capture the actual population parameters of having lower/higher education or lower/higher income than the average. This can lead to skewed data selection. Additionally, educational attainment has 25 country-specific categories that are not uniform across all countries and, therefore could be biased in some countries. Furthermore, the Total household net income on a monthly basis is based on household data and not on individual data. Therefore it could distort the association with individual filled-in data on the dependent variable, which is not a household case.

Replacing the missing values with deletion has to be met with extreme caution. It leads to data loss and when it's not assessed on its missing completely at random results it can lead to bias in the data. Therefore it reduces statistical power and can potentially lead to misleading results.

The data sample is comprised of only respondents who are 50 years or older. Therefore it does not represent society as a whole. Within this study, it is not possible to compare with other ages and can only be cross-examined when looking at other studies limiting the analysis abilities. Furthermore, data be removed when very old respondents pass away leading to potentially fewer cases and missing longitudinal data.

Conclusion

In conclusion, the logistic regression output indicates that there is no significant relationship between born in the country of interview, as a proxy for migrant background, and COVID-19: respondent tested positive, as a proxy for COVID-19 incidence rate for the critical P =,05. This suggests that other factors are influential in determining COVID-19 incidence rates. In fact, household size and total monthly average household income indicate a significant relationship between migrant background. For household size, the effect is that for each additional member of the household, the odds of testing positive on a covid-19 test go up by 15.6%. This suggests that that an increased household size is positively correlated with a positive COVID-19 test. Possible explanations for this phenomenon in migrant populations could be multigenerational living, financial constraints and later average migration from parental home. For total monthly average household income, the effect is that high income is associated with an odds decrease of testing positive for covid-19 of 36,6%. This suggests that in lower incomes the odds of testing positive for COVID-19 are higher. Possible explanations

for this phenomenon in migrant populations could be barriers to well paying labour market jobs that force migrants to work in low paying and elevated exposure to COVID-19 jobs. However, the statistical significance for the highest school degree obtained was below the threshold of P=,05. Therefore no conclusions can be made about whether lower education influences COVID-19 rates. Further research, should focus on different variables that could influence the relationship between COVID-19 incidence rate and migrant background like access to healthcare, language proficiency and social capital. Additionally, the interaction effect between household size and migrant background should be investigated.

References

Abedi, V., Olulana, O., Avula, V., Chaudhary, D., Khan, A., Shahjouei, S., ... & Zand, R. (2021). Racial, economic, and health inequality and COVID-19 infection in the United States. *Journal of racial and ethnic health disparities*, *8*, 732-742.

Agyemang, C., Richters, A., Jolani, S., Hendriks, S., Zalpuri, S., Yu, E., B. Pijls, M. Prins, K. Stronks, M. P. Zeegers (2021). Ethnic minority status as social determinant for COVID-19 infection, hospitalisation, severity, ICU admission and deaths in the early phase of the pandemic: a meta-analysis. BMJ global health 6(11).

Amate-Fortes, I., Guarnido-Rueda, A., Martínez-Navarro, D., & Oliver-Márquez, F. J. (2023). Income inequality and COVID-19 in the USA. Critical Public Health, 33(5), 814–827.

Anderson, B., Poeschel, F., & Ruhs, M. (2021). Rethinking labour migration: Covid-19, essential work, and systemic resilience. Comparative Migration Studies, 9(1), 45.

Adugna, G. (2019). Migration patterns and emigrants' transnational activities: Comparative findings from two migrant origin areas in Ethiopia. *Comparative Migration Studies*, *7*(1), 5.

Backhans, M. C., Lundberg, M., & Månsdotter, A. (2007). Does increased gender equality lead to a convergence of health outcomes for men and women? A study of Swedish municipalities. *Social science & medicine*, *64*(9), 1892-1903.

Beller, A. H. (1982). Occupational segregation by sex: Determinants and changes. *Journal of Human Resources*, 371-392.

Berkman, L. F., Kawachi, I., & Glymour, M. M. (Eds.). (2014). *Social epidemiology*. Oxford University Press.

Bharmal, N., Derose, K. P., Felician, M., & Weden, M. M. (2015). Understanding the upstream social determinants of health. *California: RAND*, 1-18.

Baena-Díez, J. M., Barroso, M., Cordeiro-Coelho, S. I., Díaz, J. L., & Grau, M. (2020). Impact of COVID-19 outbreak by income: hitting hardest the most deprived. *Journal of Public Health*, *42*(4), 698-703.

Careja, R. (2019). Immigrants' labor market outcomes: Contributions from multilevel studies. *Kölner Zeitschrift Für Soziologie Und Sozialpsychologie*, *71*(Supplement 1), 187-220.

Castañeda, H., Holmes, S. M., Madrigal, D. S., Young, M. E. D., Beyeler, N., & Quesada, J. (2015). Immigration as a social determinant of health. *Annual review of public health*, *36*, 375-392.

Coyer, L., Boyd, A., Schinkel, J., Agyemang, C., Galenkamp, H., Koopman, A. D., ... & Prins, M. (2021). Ethnic disparities in incident SARS-CoV-2 infections became wider during the second wave of SARS-CoV-2 in Amsterdam, the Netherlands: a population-based longitudinal study. *medRxiv*, 2021-07.

De Waure, C., Bruno, S., Furia, G., Di Sciullo, L., Carovillano, S., Specchia, M. L., ... & Ricciardi, W. (2015). Health inequalities: an analysis of hospitalizations with respect to migrant status, gender and geographical area. *BMC international health and human rights*, *15*, 1-10.

European Commission. (2009). Gender segregation in the labour market: root causes, implications and policy responses in the EU Luxembourg: Publications Office of the European Union

Feinberg, I., O'Connor, M. H., Owen-Smith, A., & Dube, S. R. (2021). Public health crisis in the refugee community: little change in social determinants of health preserve health disparities. *Health education research*, *36*(2), 170-177.

Frishman, N. (2012). The contribution of three social psychological theories: Fundamental cause theory, stress process model, and social cognitive theory to the understanding of health disparities--a longitudinal comparison.

Giglio, M. E., Pelton, M., Yang, A. L., Patel, A., Buzzelli, L. K., Schick, J., ... & Fredrick, B. (2022). COVID-19 contact tracing highlights disparities: household size and low-English proficiency. *Health Equity*, *6*(1), 330-333.

Gillies, C. L., Rowlands, A. V., Razieh, C., Nafilyan, V., Chudasama, Y., Islam, N., ... & Khunti, K. (2022). Association between household size and COVID-19: A UK Biobank observational study. *Journal of the Royal Society of Medicine*, *115*(4), 138-144.

Gravelle, H., & Sutton, M. (2009). Income, relative income, and self-reported health in Britain 1979–2000. *health economics*, *18*(2), 125-145.

Hawkins, R. B., Charles, E. J., & Mehaffey, J. H. (2020). Socio-economic status and COVID-19–related cases and fatalities. *Public health*, *189*, 129-134.

Indseth, T., Grøsland, M., Arnesen, T., Skyrud, K., Kløvstad, H., Lamprini, V., ... & Kjøllesdal, M. (2021). COVID-19 among immigrants in Norway, notified infections, related hospitalizations and associated mortality: a register-based study. *Scandinavian journal of public health*, *49*(1), 48-56.

Ingleby, D. (2012). Ethnicity, migration and the 'social determinants of health'agenda. *Psychosocial Intervention*, *21*(3), 331-341.

Ichou, M., & Wallace, M. (2019). The healthy immigrant effect. *Demographic Research*, *40*, 61-94.

Jia, F., Liu, X., Wang, Y., & Ma, M. (2023). The effect of housing tenure on health status of migrant populations in China: are health service utilization and social integration mediating factors?. *Archives of Public Health*, *81*(1), 200.

Kapilashrami, A., Hill, S., & Meer, N. (2015). What can health inequalities researchers learn from an intersectionality perspective? Understanding social dynamics with an inter-categorical approach?. *Social Theory & Health*, *13*, 288-307.

Kaukko, M., Alisaari, J., Heikkola, L. M., & Haswell, N. (2022). Migrant-background student experiences of facing and overcoming difficulties in Finnish comprehensive schools. *Education Sciences*, *12*(7), 450.

Khanijahani, A. (2021). Racial, ethnic, and socioeconomic disparities in confirmed COVID-19 cases and deaths in the United States: a county-level analysis as of November 2020. *Ethnicity & health*, *26*(1), 22-35.

Kim, H., Fox, A. M., Kim, Y., Kim, R., Bae, G., & Kang, M. (2021). Is the male disadvantage real? Cross-national variations in sex gaps in COVID-19 incidence and mortality. *Global Public Health*, *16*(12), 1793-1803.

King, T. L., Kavanagh, A., Scovelle, A. J., & Milner, A. (2020). Associations between gender equality and health: a systematic review. *Health promotion international*, *35*(1), 27-41.

Krieger, N. (2014). Discrimination and health inequities. *International journal of health services*, *44*(4), 643-710.

Lanting, L. C., Joung, I. M., Vogel, I., Bootsma, A. H., Lamberts, S. W. J., & Mackenbach, J. (2008). Ethnic differences in outcomes of diabetes care and the role of self-management behavior. *Patient Education and Counseling*, *72*(1), 146-154.

Link, B. G., & Phelan, J. C. (1995). Social conditions as fundamental causes of disease. Journal of Health and Social Behavior (extra issue), 80–94.

Liu, P., McQuarrie, L., Song, Y., & Colijn, C. (2021). Modelling the impact of household size distribution on the transmission dynamics of COVID-19. *Journal of the Royal Society Interface*, *18*(177), 20210036.

de Lusignan, S., Dorward, J., Correa, A., Jones, N., Akinyemi, O., Amirthalingam, G., ... & Hobbs, F. R. (2020). Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study. *The Lancet Infectious Diseases*, *20*(9), 1034-1042.

Magesh, S., John, D., Li, W. T., Li, Y., Mattingly-App, A., Jain, S., ... & Ongkeko, W. M. (2021). Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic review and meta-analysis. *JAMA network open*, *4*(11), e2134147-e2134147.

Marathe, A., Lewis, B., Chen, J., & Eubank, S. (2011). Sensitivity of household transmission to household contact structure and size. *PloS one*, *6*(8), e22461.

Marmot M. Social determinants of health inequalities. Lancet. 2005;365(9464):1099–104

Rafael, R. D. M. R., Neto, M., Depret, D. G., Gil, A. C., Fonseca, M. H. S., & Souza-Santos, R. (2020). Effect of income on the cumulative incidence of COVID-19: an ecological study. *Revista Latino-Americana de Enfermagem*, *28*, e3344.

Raghupathi, V., & Raghupathi, W. (2020). The influence of education on health: an empirical assessment of OECD countries for the period 1995–2015. *Archives of Public Health*, *78*, 1-18.

Razum O & Twardella D (2002) Time travel with Oliver Twist– towards an explanation foa a paradoxically low mortality among recent immigrants. Tropical Medicine and International Health 7, 4–10.

Razum, O., Zeeb, H., & Rohrmann, S. (2000). The 'healthy migrant effect'–not merely a fallacy of inaccurate denominator figures. *International journal of epidemiology*, *29*(1), 191-192.

Rechel, B., Mladovsky, P., Ingleby, D., Mackenbach, J. P., & McKee, M. (2013). Migration and health in an increasingly diverse Europe. *The Lancet*, *381*(9873), 1235-1245.

Saarinen, S., Moustgaard, H., Remes, H., Sallinen, R., & Martikainen, P. (2022). Income differences in COVID-19 incidence and severity in Finland among people with foreign and native background: A population-based cohort study of individuals nested within households. PLoS medicine, 19(8), e1004038.

Saß, A. C., Wurm, S., & Scheidt-Nave, C. (2010). Alter und Gesundheit-Eine Bestandsaufnahme aus Sicht der Gesundheitsberichterstattung. Bundesgesundheitsbl 2010. 53:404–416

Schneiderman, M., Rumain, B., Kaganovskiy, L., & Geliebter, A. (2022). Incidence and relative risk of COVID-19 in adolescents and youth compared with older adults in 19 US states, fall 2020. *JAMA Network Open*, *5*(7), e2222126-e2222126.

SHARE-ERIC (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w1.900

SHARE-ERIC (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w2.900

SHARE-ERIC (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 3 – *SHARELIFE*. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w3.900

SHARE-ERIC (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w4.900

SHARE-ERIC (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w5.900

SHARE-ERIC (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w6.900

SHARE-ERIC (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w7.900

SHARE-ERIC (2024). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8*. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w8.900

SHARE-ERIC (2024). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. COVID-19 Survey 1*. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w8ca.900

SHARE-ERIC (2024). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9*. Release version: 9.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w9.900

SHARE-ERIC (2024). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9. COVID-19 Survey 2*. Release version: 9.0.0. SHARE-ERIC. Data set.

Spallek J, Zeeb H & Razum O (2011) What do we have to know from migrants' past exposures to understand their health status? a life course approach. Emerging Themes in Epidemiology 8, 6.

Start, R. (2021, May 21). "Corona-problemen bij migranten hadden eerder erkend moeten worden." NOS. https://nos.nl/artikel/2381750-corona-problemen-bij-migranten-hadden-eerder-erkend-moeten-worden. Accessed: 28-02-2024

United Nations. (1998). Recommendations on statistics of international migration, revision 1. *Statistical Papers, Series M, No. 58, Rev. 1*.

Waris, A., Ali, M., Khan, A. U., Ali, A., Bangash, A. K., & Baset, A. (2020). COVID-19 incidence in Pakistan: gender disparity. *Iranian Journal of Psychiatry and Behavioral Sciences*, *14*(3), e105990.

WHO Commission on Social Determinants of Health, & World Health Organization. (2008). *Closing the gap in a generation: health equity through action on the social determinants of health: Commission on Social Determinants of Health final report*. World Health Organization.

World Health Organization (2021). Refugees and migrants in times of COVID-19: mapping trends of public health and migration policies and practices. BY-NC-SA 3.0 IGO

World Health Organization. (2015). *World report on ageing and health*. World Health Organization.

Wrench, J. (2015). Discrimination against immigrants in the labour market: An overview and a typology. *Routledge Handbook of Immigration and Refugee Studies*, 118-124.

Zajacova, A., & Lawrence, E. M. (2018). The relationship between education and health: reducing disparities through a contextual approach. *Annual review of public health*, *39*, 273-289.

Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W., Zhao, S., ... & Yu, H. (2020). Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*, *368*(6498), 1481-1486.

Appendices

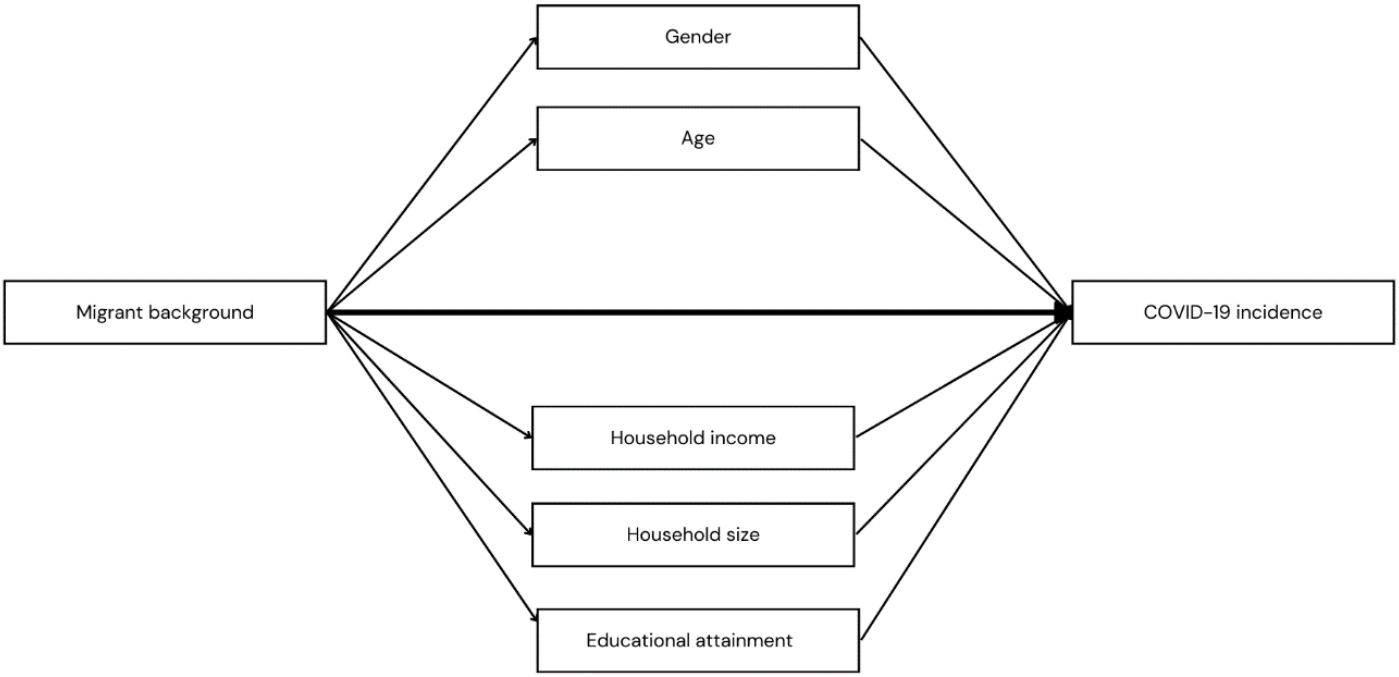

Figure 1: Conceptual model. (Constructed by author, 2024)

Block 1: Main effect between depended variable and the independent variable. (Constructed by author, 2024)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Born in the country of interview	-,168	,099	2,895	1	,089	,846
	Constant	-1,400	,094	223,366	1	<,001	,247

a. Variable(s) entered on step 1: Born in the country of interview.

Block 2: Relationship between depended variable, independent variable, age and gender. (Constructed by author, 2024)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Born in the country of interview	-,176	,099	3,156	1	,076	,839
	Year of birth	,027	,003	62,833	1	<,001	1,028
	Male or female	,013	,061	,043	1	,835	1,013
	Constant	-54,981	6,761	66,128	1	<,001	,000

a. Variable(s) entered on step 1: Born in the country of interview, Year of birth, Male or female.

Block 3: Relationship between depended variable and all other covariates. (Constructed by author, 2024)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Born in the country of interview	-,154	,100	2,397	1	,122	,857
	Year of birth	,023	,004	40,063	1	<,001	1,023
	Male or female	-,014	,062	,051	1	,821	,986
	Household size	,145	,026	29,943	1	<,001	1,156
	Highest school degree obtained	,081	,070	1,356	1	,244	1,085
	Total income received by all hh members an average month last year	-,455	,061	54,841	1	<,001	,634
	Constant	-46,369	7,112	42,510	1	<,001	,000

a. Variable(s) entered on step 1: Born in the country of interview, Year of birth, Male or female, Household size, Highest school degree obtained, Total income received by all hh members an average month last year.

Goodness of Fit[a]

	Value	df	Value/df
Deviance	1811,869	1619	1,119
Scaled Deviance	1811,869	1619	
Pearson Chi-Square	1761,864	1619	1,088
Scaled Pearson Chi-Square	1761,864	1619	
Log Likelihood[b]	-1546,336		
Akaike's Information Criterion (AIC)	3106,672		
Finite Sample Corrected AIC (AICC)	3106,686		
Bayesian Information Criterion (BIC)	3155,673		
Consistent AIC (CAIC)	3162,673		

Dependent Variable: COVID-19: respondent tested positive
Model: (Intercept), Male or female, Year of birth, Household size, Born in the country of interview, Highest school degree obtained, Total income received by all hh members an average month last year

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Model 1: Goodness of fit with the dependent variable and the independent variables (Constructed by author, 2024.)

Goodness of Fit[a]

	Value	df	Value/df
Deviance	1741,306	1556	1,119
Scaled Deviance	1741,306	1556	
Pearson Chi-Square	1651,874	1556	1,062
Scaled Pearson Chi-Square	1651,874	1556	
Log Likelihood[b]	-1511,054		
Akaike's Information Criterion (AIC)	3162,109		
Finite Sample Corrected AIC (AICC)	3163,346		
Bayesian Information Criterion (BIC)	3652,117		
Consistent AIC (CAIC)	3722,117		

Dependent Variable: COVID-19: respondent tested positive
Model: (Intercept), Male or female, Year of birth, Household size, Born in the country of interview, Highest school degree obtained, Total income received by all hh members an average month last year

a. Information criteria are in smaller-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Model 2: Goodness of fit with the dependent variable and the independent variables with interaction effect (Constructed by author, 2024.)

Tests of Between-Subjects Effects

Dependent Variable: COVID-19: respondent tested positive

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	10,739[a]	16	,671	4,688	<,001	,009
Intercept	8,488	1	8,488	59,282	<,001	,007
dn004_	,707	1	,707	4,936	,026	,001
hhsize	6,355	9	,706	4,931	<,001	,005
dn004_ * hhsize	2,055	6	,342	2,392	,026	,002
Error	1157,846	8087	,143			
Total	1416,000	8104				
Corrected Total	1168,584	8103				

a. R Squared = ,009 (Adjusted R Squared = ,007)

Table 1: Interaction effect between household size and migrant background. (Constructed by author, 2024.)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Year of birth	,025	,004	41,101	1	<,001	1,025
	Male or female	-,034	,065	,274	1	,600	,967
	Household size	,131	,028	22,113	1	<,001	1,140
	Highest school degree obtained	,097	,073	1,753	1	,185	1,102
	Total income received by all hh members an average month last year	-,451	,064	49,063	1	<,001	,637
	Constant	-49,528	7,495	43,662	1	<,001	,000

a. Variable(s) entered on step 1: Year of birth, Male or female, Household size, Highest school degree obtained, Total income received by all hh members an average month last year.

Table 2: Logistic regression output individuals who responded yes in the variable Born in the country of interview. (Constructed by author, 2024.)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Year of birth	,008	,012	,412	1	,521	1,008
	Male or female	,198	,206	,919	1	,338	1,218
	Household size	,289	,087	10,998	1	<,001	1,336
	Highest school degree obtained	-,092	,232	,157	1	,691	,912
	Total income received by all hh members an average month last year	-,503	,208	5,866	1	,015	,605
	Constant	-16,310	22,858	,509	1	,476	,000

a. Variable(s) entered on step 1: Year of birth, Male or female, Household size, Highest school degree obtained, Total income received by all hh members an average month last year.

Table 3: Logistic regression output individuals who responded no in the variable Born in the country of interview. (Constructed by author, 2024.)