

Lunds Universitet
Nationalekonomiska institutionen
NEKH03 - Kandidatuppsats
Augusti 2018



LUND UNIVERSITY

School of Economics and Management

Uppväxt i fel område?

- en studie om vilka faktorer som påverkar människors uppnådda utbildningsnivån i sydöstra Europa

Madeleine Fraser
19900805-0628

Handledare:
Margareta Dackehag

Abstract

Several studies have been made regarding if there is discrimination towards the Roma on the South Eastern European labour market. The main result in previous studies has shown that the salary difference between the majority population and the Roma is to a great extent explained by difference in attained education. This study looks at what factors effects education attainment with the aim to explain why the Roma in general have a lower level of education than the majority population. Using data from UNDP, three different OLS regressions have been calculated. The theory of neighbourhood effects has been used to analyse if the result shows that segregation is one of the factors that effect the difference. Other factors that is expected to effect education attainment is controlled for such as family income, parents' education and number of household members. The result show that there is a neighbourhood effect and that family income, parents' education as well as number of household members effect the education attainment. The primary common denominator that explains the difference in the regression analysis is however the different ethnicities. *The main conclusion from the study is therefore that in order to fully explain the difference, more in-depth research would need to be done.*

Keywords: neighbourhood effects, Roma, Europe, education, segregation

Innehållsförteckning

1. Introduktion	1
1.1 Syfte	3
1.2 Frågeställning	3
1.3 Begränsningar	3
2. Teori	5
2.1 Innan arbetsmarkandsdiskriminering	5
2.1.1 Områdeseffekten	5
3. Tidigare forskning	7
3.1 Lönediskriminering	7
3.2 Områdeseffekter	9
4. Data	11
4.1 Deskriptiv statistik	13
5. Metod	15
5.1 Regressionsanalys	15
5.2 Modell	15
5.3 Beskrivning av variabler	15
5.3.1 Skolår	15
5.3.2 Ålder	16
5.3.3 Familjens inkomst	16
5.3.4 Förälders utbildningsnivå	16
5.3.5 Huvudsakligt språk som talas hemma	16
5.3.6 Typ av område för bosättning	16
5.3.7 Förälders arbetssituation	16
5.3.8 Kön	16
5.3.9 Land	17
5.4 Heteroskedasticitet och grupperade/robusta standardfel	17
6. Resultat	18
6.1 Separata regressioner	19
7. Diskussion och Analys	22
8. Sammanfattning	25
Litteraturförteckning	27

1. Introduktion

Utbildningsnivå korrelerar med både arbetsmarknadsdeltagande och lönenivå (Borjas, 2016). Utbildning framhålls därför ofta som vägen ut ur fattigdom och som en lösningen i länder med hög arbetslöshet. Men vilka faktorer påverkar vilken utbildningsnivå människor uppnår? Till vilken grad avgör vart vi växer upp, föräldrars utbildningsnivå eller vilken etnicitet vi tillhör vår framtid? I USA var 2013 skillnaden i uppnådd utbildningsnivå mellan kvinnor och män mycket liten men desto större mellan olika etniciteter (Borjas, 2016). 2013 hade 31% av den latinamerikanska etnicitetsgruppen inte avslutat den obligatoriska grundskolan i USA medan motsvarande siffra endast var 6.3% bland de vita (Borjas, 2016). I Europa så är romerna är en av de största etniska minoriteterna. Den romska folkgruppen är spridda över hela Europa men 2006 uppskattades 68% av dem vara bosatta i sydöstra Europa (Ivanov, et al., 2006). I en studie från 2004 så var det stora skillnader mellan majoritetsbefolkningen och den romska minoriteten i sydöstra Europa när det gäller uppnådd utbildningsnivå. 38% av de romska barnen avslutar inte grundskolan medan motsvarande siffra för majoritetsbefolkningen endast var 4% (Ivanov, et al., 2006).

Historierna om varför och hur romerna kom till Europa är många och mytomspunna och en av de mest populära är att de härstammar från Indien. Denna slutsats baseras på att språket romani, det gemensamt talade språket bland romer, är besläktat med sanskrit (Achim, 1998). Romers ställning har varierat genom historien. Under det Ottomanska riket kunde de röra sig relativt fritt då de varken var en del av den kristna gruppen eller den muslimska som riket då var indelat i. I Österrike-Ungern hade de till en början också möjlighet att röra sig fritt, fram till 1800-talet då de inleddes försök att göra dem till bofasta och assimilerade (Ivanov, et al., 2006). I Rumänien var det slavar fram till så sent som 1848 (Achim, 1998) och under andra världskriget dog många i koncentrationsläger. Efter kommunismens fall i Östra Europa har det konstaterats att romerna var de stora förlorarna på övergången till marknadsekonomin. Även om det fanns stora problem med bristande resurser under kommunismen innebar det för de flesta romer att de hade ett jobb att gå till och mat för dagen - vilken många av dem inte har idag (FRA- European Union Agency for Fundamental Rights, 2014). I en undersökning från 2011 som genomfördes av UNDP och Europeiska unionens byrå för grundläggande rättigheter så beräknades 90% av de tillfrågade romerna leva under fattigdomsgränsen, fler än

två tredje delar var arbetslösa och hälften uppgav att de upplevt diskriminering på grund av att de är romer (United Nations Development Programme, 2018).

Olika initiativ har tagits för att förbättra romernas utsatthet och 2005 påbörjades något som kallades ”Ett decennium av romsk inkludering”. Detta var ett samarbete mellan 12 europeiska länder för att motverka diskriminering samt minska avståndet mellan majoritetsbefolkningen och den romska folkgruppen (Brüggemann & Friedman, 2017), ett avstånd som påtagligt finns vilket vi inte minst kan se vad det gäller uppnådd utbildningsnivå. Rapporter har skrivits och vetenskapliga studier genomförts på romernas utsatthet och ifall de utsatts för diskriminering. Inom nationalekonomi har bland annat Milcher (2011) gjort en studie på ifall det förekommer lönediskriminering i fem länder i sydöstra Europa och O'Higgins, (2010) har studerat romers arbetsmarkandsresultat i nio länder i sydöstra Europa. Resultaten är inte helt entydiga och Milchers (2011) primära slutsats är att skillnader i lönenivåer beror på skillnader i utbildningsnivå, det vill säga skillnad i ackumulerat humankapital. O'Higgins (2010) resultat visar på att romer i genomsnitt får en sämre avkastning på investeringar i humankapital vilket kan innebära att romer har ett lägre incitament att investera i utbildning och detta kan därmed vara en förklaring till skillnaden i utbildningsnivå.

Men även om utbildningsnivå skulle vara den primära orsaken till löneskillnader mellan romer och majoritetsbefolkningen i Sydöstra Europa, blir den väsentliga frågan hur deras förutsättningar till utbildning skiljer sig åt? Orsakar fattigdom, segregation och diskriminering att romer inte har möjlighet eller kanske till och med väljer att inte utbilda sig?

Data från UNDP och FRA (2004) används i den här studien för att undersöka varför utbildningsnivån skiljer sig mellan romer och majoritetsbefolkningen. I det relevanta urvalet så är medelinkomsten för de romska familjerna i urvalsgruppen 178 euro jämfört med 406 euro för familjerna som tillhörde majoritetsbefolkningen. Över hälften av de tillfrågade romerna bodde i ett område där majoriteten också var romer och så många som 65.2% av de romska barnen i den här studien hade en arbetslös förälder, medan motsvarande siffra för majoritetsbefolkningen var 26.5%. Så även om utbildningsnivå skulle förklara skillnaden i lönenivå så tyder mycket på att den romska gruppen och majoritetsbefolkningen i hög grad har olika förutsättningar att uppnå samma nivå på utbildning. Syftet med den här uppsatsen är att undersöka vilka faktorer som påverkar uppnådd utbildningsnivå för att se ifall det finns någon form av *innan arbetsmarknadsdiskriminering* mot romer i sydöstra Europa. Särskilt

fokus läggs på vilken inverkan segregation har och om det finns områdeseffekter som förklarar romernas lägre utbildningsnivå. Andra aspekter som kontrolleras för är bland annat familjens inkomstnivå, huvudsakligt språk som talas hemma och förälders utbildningsnivå.

1.1 Syfte

Syftet med den här uppsatsen är att undersöka vilka faktorer som kan förklara att romer har lägre utbildningsnivå jämfört med majoritetsbefolkningen i sydöstra Europa. Särskilt fokus ligger på att undersöka ifall segregation kan vara en av de bidragande orsakerna till nivåskillnaderna på utbildning och det undersöks ifall det finns några områdeseffekter.

1.2 Frågeställning

Vilka faktorer förklarar skillnaden i utbildningsnivå mellan romer och majoritetsbefolkningen i sydöstra Europa samt finns det en områdeseffekt som förklarar romernas lägre utbildningsnivå?

1.3 Begränsningar

Urvalet av länder i den här studien är Albanien, Bosnien, Bulgarien, Kroatien, Montenegro, Makedonien, Rumänien, Kosovo och Serbien. Att urvalet av länder är begränsat till sydöstra Europa är för att en majoritet av alla romer är bosatta i regionen, uppskattningsvis 68% år 2006 (Ivanov, et al., 2006).

Uppsatsen är baserad på ett antal regressioner som har genomförts med tvärsnittsdata från UNDP och FRA som 2004 genomförde en undersökning i elva länder i sydöstra Europa. I den här uppsatsen har Ungern och Tjeckien inte inkluderats då dataseten såg annorlunda ut för dessa länder jämfört med övriga länderna varvid en avvägning gjordes och jag valde att exkludera dem för att vara säker på att all data och allt material var enhetligt och jämförbart. Datamaterialet bestod av 34 116 observationer som var indelat i tre olika subgrupper, romer, majoritetsbefolkningen och flyktingar. I den här uppsatsen har urvalet begränsats till endast subgrupperna romer och majoritetsbefolkningen.

Uppsatsen har en del begränsningar på grund av det dataunderlag som fanns att tillgå. Datamaterialet är tvärsnittsdata och ej tidsseriedata vilket begränsar möjligheten att se hur utvecklingen har sett ut över tid samt att få en helhetsbild av exempelvis familjers

inkomstnivå. Uppgifter om inkomst finns bara från det specifika intervjutillfället och detta har använts som en proxy för familjens generella inkomst. För att kunna se vilka trender som funnits under en längre tidsperiod samt tydligare mönster för hur situationen ser ut så hade tidsseriedata kunnat ge ett mer komplett resultat. Vid studier av områdeseffekter hade det varit fördelaktigt att ha underlag kring genomsnittlig utbildningsnivå och genomsnittliga inkomster för varje enskilt område. Detta för att kunna jämföra hur förutsättningar skiljer sig åt mellan olika områden och vilket effekt det beräknas ha för varje område mer specifikt. Det fanns tyvärr inte som underlag så för att få en indikation om ifall det kan finnas områdeseffekter som påverkar utbildningsnivån har variabeln ”typ av område” använts. Denna variabel anger ifall individen är bosatt i ett mixat område, ett område där majoriteten är romer eller i ett område där majoriteten tillhör majoritetsbefolkningen.

2. Teori

2.1 Innan arbetsmarkandsdiskriminering

Innan arbetsmarkandsdiskriminering (pre-market diskriminering) är en samling av olika faktorer och teorier som påverkar individers humankapitalackumulation och ekonomiska position på arbetsmarknaden innan de ens har anträt den. Tillskillnad från andra arbetsmarkandsteorier som exempelvis den klassiska humankapitalteorin, som fokuserar på ifall det finns någon lönediskriminering på arbetsmarknaden med hänsyn till utbildningsnivå, så fokuserar dessa teorier på inverkan av saker som vilket område man växer upp i, samt vilken inverkan förebilder och nätverk har på individers möjligheter på arbetsmarknaden. Att sociala faktorer har en inverkan går att konstatera, det svåra är att mäta hur stor inverkan de har och att definiera dem. Jag har i den här uppsatsen valt att fokusera på hur uppväxtmiljöer påverkar uppnådd utbildningsnivå genom att studera områdeseffekten.

2.1.1 Områdeseffekten

Områdeseffekten syftar till den inverkan vilket område människor är bosatta och växer upp i har på deras framtid. Ett flertal studier av blanda annat Borjas (1995) har visat på att den här effekten existerar och det finns ett flertal möjliga orsaker som förklarar detta.

I USA påverkas exempelvis resurserna en skola har av områdets generella ekonomiska ställning på grund av hur regelverket ser ut när det gäller finansiering av kommunala skolor. Detta leder till att områden som generellt ligger sämre till ekonomiskt har skolor med bristande resurser och den utbildningskvalité som barn erbjuds påverkas av familjens ekonomiska förutsättningar (Durlauf, 2004).

Områdeseffekten handlar dock om mycket mer än bara ekonomiska förutsättningar - så som sociala och psykologiska faktorer. När vi växer upp påverkas våra beslut av hur andra i vår omgivning agerar och ofta imiterar vi personer vi ser upp till när vi tar beslut. Det kan vara förebilder, vänner och familj. Varför människor imiterar vissa beteende bottnar troligtvis i flera faktorer men delvis så har vi en undermedveten drivkraft att bete oss som andra för att känna att vi passar in i kombination med att vi har ett beroendeförhållanden till dem som finns i vår närhet. Detta har en inverkan när individer ska fatta beslut och att inte agera som majoriteten är förknippat med vissa kostnader, så väl materiella som sociala. Även informationsöverföring är influerad av dessa beroendeförhållanden vilket gör att inverkan på

information angående konsekvenserna av ett visst beteende påverkas av att se andra ändå bete sig så, exempelvis information om vad konsekvenserna av en avbruten skolgång blir förminskas av att andra i personens närhet har gjort det valet. Information om positiva effekter av att utbilda sig har större sannolikhet att göra starka avtryck hos personer som också har många i sin närhet med högre utbildning (Durlauf, 2004).

En annan viktig aspekt av områdeseffekten är så kallade *spillover* effekter av humankapital. Detta innebär att i områden med ett högt generellt humankapital troligtvis kommer höja alla individers nivå på humankapital då kunskap sprids och vi ofta delar med oss av våra erfarenheter till andra i vår närhet (Borjas, 1995). Detta innebär att i segregerade områden med en hög andel av av en viss etnicitet, där det sociala nätverket främst består av denna etnicitet, så har denna etnicitets generella nivå på ackumulerat humankapital en inverkan på dem som växer upp i området och deras nivå på ackumulerat humankapital (Borjas, 1995).

Sammanfattningsvis innebär dessa olika faktorer att en individ kommer finna val och beteenden mer attraktiva ifall personer i dess närhet beter sig likadant eller gjort samma val. Valet att till exempel slutföra skolan kommer sannolikt vara mer tilltalande för en individ i ett område där många har universitetsutbildning och vars vänner också stannar kvar i skolan (Durlauf, 2004). Människor lär sig även av varandra och delar med sig av sina erfarenheter vilket innebär att individers ackumulerade humankapital påverkas av vilken nivå på humankapital människor i deras omgivning har (Borjas, 1995).

3. Tidigare forskning

3.1 Lönediskriminering

Flera studier på lönediskriminering kopplat till etnicitets har gjorts, bland annat i USA. En studie från 1995 som studerade ifall det fanns någon lönediskriminering mot svarta visade att mellan 9.8%-13.4% av löneskillnaden förmodligen berodde på diskriminering. Hur stor andel av löneskillnaden som orsakades av diskriminering påverkades av vilka faktorer som kontrollerades för (Borjas, 2016).

Vad det gäller diskriminering mot romer på arbetsmarknaden i sydöstra Europa så har Milcher (2011) tidigare forskat på detta. Hon har genomfört studier med fokus på lönediskriminering med hänsyn till utbildningsnivå och arbetslivserfarenhet. Med data från UNDP 2004, som är samma data som används i den här studien, så genomförde Milcher regressionsanalyser utifrån en standard Blinder-Oaxaca dekomposition om relativa löneskillnader. Alla de elva länder som finns i datamaterialet var ej inkluderade i Milchers studie utan det urval som användes var länderna Albanien, Bulgarien, Kroatien, Kosovo och Serbien. I samtliga länder finns det skillnader i lönenivå mellan romer och majoritetsbefolkningen och som enligt studien till viss del beror på diskriminering. Diskrimineringen mättes genom en dekomponering om relativa löneskillnader där både en karaktärskoefficient och en diskrimineringskoefficient räknades fram. Signifikansen av dessa testades sedan genom en noll hypotesen att samtliga av dessa koefficienter var lika med 0. I Albanien och Kosovo förkastades noll hypotesen att diskrimineringskoefficienten var lika med noll medan noll hypotesen ej kunde förkastas i Bulgarien, Kroatien och Serbien (Milcher & Fischer, 2011). Utifrån det resultat som regressionsanalyserna och noll hypoteserna gav så kom Milcher fram till slutsatsen att löneskillnaderna i Albanien och Kosovo troligtvis beror på diskriminering medan det i Bulgarien, Kroatien och Serbien främst tyder på skillnader i ackumulerat humankapital (Milcher & Fischer, 2011) & (Milcher, 2011).

Resultatet av Milchers forskning visar inte på att det finns en utbredd lönediskriminering utan på att den primära förklaringen till skillnaden i lönenivå mellan romer och majoritetsbefolkningen är skillnad i utbildningsnivå. Milcher påpekar dock att även om det inte tyder på det finns någon utbredd lönediskriminering mot romer så kan det finnas andra former av diskriminering som påverkar deras ställning på arbetsmarknaden. Exempelvis kan det finnas diskriminering mot romer inom utbildningsväsendet som påverkar förvärvandet av

humankapital som i sin tur leder till att romerna generellt har lägre utbildningsnivå. Därmed kan det indirekt vara så att diskriminering trots allt är en betydande orsak till att romerna har lägre löner än majoritetsbefolkningen (Milcher & Fischer, 2011).

O'Higgins (2010) har forskat på varför romerna har en sämre ställning på arbetsmarknaden jämfört med majoritetsbefolkningen i sydöstra Europa. Även O'Higgins använde data från den undersökning som UNDP och FRA genomförde 2004. Forskningen utgjordes av ett flertal olika regressioner som exempelvis en OLS-regression för att undersöka vilka variabler som påverkar utbildningsnivån, en probit-modell för faktorer som påverkar sysselsättningsgraden samt en dekomponering om relativa löneskillnader. O'Higgins (2010) har gjort en samlad regression där nio länder som fanns i datamaterialet inkluderats samt använt LIML¹-metoder för att kontrollera för endogeniteten i utbildningsnivå samt två möjliga avvikelser i urvalet. Efter att detta kontrollerats för så visar hans resultat att den procentuella avkastning på utbildning generellt sett är samma för romer och majoritetsbefolkningen. Däremot när det gäller sysselsättningsgraden så leder inte investeringar i utbildning till jobb i lika hög grad för romer som det gör för majoritetsbefolkningen. När hänsyn tas till romers generellt låga löner samt låga sysselsättningsgrad så är den absoluta marginella avkastningen för ett extra år investerat i utbildning mycket lägre för romer än för majoritetsbefolkningen. Sammanfattningsvis är O'Higgins slutsats att det finns diskriminering mot romer på arbetsmarknaden och att det delvis kan anses vara ett rationellt val av romerna att inte investera i utbildning till samma grad som majoritetsbefolkningen då de inte kan förvänta sig att det ska ge samma avkastning (O'Higgins, 2010)

Beroende på vilken aspekt som står i fokus samt metod som använts så har slutsatserna om ifall romerna diskrimineras på arbetsmarknaden eller inte skiljt sig något åt. Men både arbetslöshet och lönenivå korrelerar med utbildningsnivå (Borjas, 2016), så även om diskriminering kan vara en orsak till att romer och majoritetsbefolkningen gör olika val kring hur många år det är värt att investera i utbildning så är det relevant att undersöka vilka andra faktorer som påverkar och kan förklara skillnaden i utbildningsnivå mellan romer och majoritetsbefolkningen.

¹ Limited information maximum likelihood

3.2 Områdeseffekter

Tidigare forskning på områdeseffekter har framförallt gjorts i USA och en av de mest kända på området är William Julius Wilson som med boken *The truly disadvantaged* lyfte segregerade bostadsområden som en viktig faktor i den stadigvarande fattigdomen som fanns i innerstadens ghetton (Durlauf, 2004). Manski (1993) poängterade i en artikel problematiken med att utföra studier som ska identifiera endogena sociala effekter, vilket områdeseffekter är ett exempel på, eftersom det i den typen av studier ofta riskerar att finnas en reflektionsproblematik. Reflektionsproblematiken uppstår i studier där en individs beteende mäts mot det generella beteendet av gruppen som individen är en del av vilket gör det svårt att identifiera vad som är orsak och verkan, det vill säga är individen i första hand med och skapar normerna eller främst påverkad av dem? Manski (1993) framhöll vikten av att det därför finns en referensgrupp när områdeseffekter mäts för att graden av tillförlitlighet i resultatet ska öka.

I den mängd studier som har genomförts på områdeseffekter så har resultaten varit motstridiga i vissa aspekter, även när studierna använt samma oberoende variabler. För att få fram ett mer generaliserbart resultat så genomförde Nieuwenhuis och Hooimeijer (2016) en meta-regression på 88 strategiskt utvalda studier som alla uppfyllde ett antal kriterier för att det totala urvalet skulle vara jämförbart. I deras studie använde de koefficienterna för områdesvariablerna som utfallsvariabler från de studier som de hade som underlag. Fyra olika typer av områdesegenskaper studerades som var fattigdom, utbildningsklimat, proportion av invandrargrupper/etniska grupper och graden av sociala problem i området (Nieuwenhuis & Hooimeijer, 2016). Studiens resultat visar att områdeseffekten på individens utbildningsresultat är en funktion av alla dessa områdesegenskaper men att vissa har en större inverkan än andra beroende på i vilket land studien utförts (Nieuwenhuis & Hooimeijer, 2016). Exempelvis så har områdesegenskapen ”bristfälligt utbildningsklimat” en mycket starkare negativ korrelation i USA på utbildningsresultat än vad den har i Europa.

Sammanfattningsvis så är deras slutsats att alla fyra områdesegenskaper har en inverkan på utbildningsresultat men att vilken modell som används samt att rätt kontrollvariabler inkluderas är av största vikt för att inte områdeseffekterna ska riskeras att över- eller underskattas (Nieuwenhuis & Hooimeijer, 2016). Även Durlauf (2004) framhåller vikten av rätt kontrollvariabler tillämpas när studier på områdeseffekter görs. Betydelsen av att

kontrollera för familjens förutsättningar betonas särskilt då områdeseffekten annars riskeras att överskattas.

4. Data

Den data och använt som empiri är tvärsnittsdata från UNDP och FRA som 2004 genomförde en survey i elva länder i sydöstra Europa. På grund av att det finns ett stigma kopplat till att vara rom samt att identifiering med att tillhöra den romska etniciteten ibland förväxlas med nationalitet och medborgarskap så antas det finnas en underrepresentation i folkräkningen för hur många hur många romer det finns i sydöstra Europa. Den officiella statistiken på hur många romer som är bosatta i Europa skiljer sig ifrån vad ett flertal experter uppskattar att antalet är. På grund av detta gjorde UNDP bedömningen att folkräkningen riskerade att inte ge ett representativt urval (Ivanov, et al., 2006). Så istället för ett helt slumpmässigt urval för intervjuerna så användes en form av 'pyramid urval'. Olika nivåer fanns i pyramiden där det uppskattade totala antalet romer var i botten av pyramiden, folkräkningen en nivå i mitten och toppen utgjordes av en hypotetisk situation att ingen som är rom själv skulle uppge sig vara det. Befolkningsräkningen som var det mittersta skiktet i pyramiden utgjorde ett underlag för befolkningsstrukturen och den förväntade underrepresentationen i folkräkningen antogs vara jämnt fördelat över landet vilket antogs vara tillräckligt för att ge ett väntevärdesriktigt representativt stratifierat urval (Ivanov, et al., 2006). Urvalet för den romska populationen representerades av 'Romska bosättningar eller områden med en stor andel romsk befolkning'. Områdena definierades utifrån att andelen romer i området var samma eller större som den totala andelen romer i landet. Utifrån detta definierades sedan urvals kluster och inom dessa kluster användes sedan en 'slumpvandring' som metod att få fram det slutliga urvalet (Ivanov, et al., 2006).

För att urvalet av majoritetsbefolkningen skulle anses vara en jämförbar grupp så valdes urvalet för majoritetsbefolkningen genom att de levde nära den romska gruppen. Den mer exakta benämningen för den här populationen skulle vara 'icke-romer som lever i nära anslutning till det sårbara urvalet (romerna)' (Ivanov, et al., 2006). I slutsatserna och analysen av resultatet är det viktigt att beakta att resultatet för majoritetsbefolkningen inte är representativt för hela den subgruppen utan just de som valts som kontrollgrupp för den romska gruppen. Eftersom urvalet för majoritetsbefolkningen valts utifrån kriteriet att de ska vara ett jämförbart urval skulle därför skillnaderna mot majoritetsbefolkningen möjligtvis kunna vara ännu större.

Urvalet baseras på områden som har en genomsnittlig eller hög romsk befolkning. Detta gör att urvalet inte heller är helt representativt för alla romer eftersom de romer som lever i områden där de är i en klar minoritet inte omfattas av urvalet samt att romer som har blivit helt assimilerade och själva inte identifierar sig som romer hamnar utanför det här forskningsunderlaget. Urvalet beräknas dock representera 85% av den romska befolkningen vilket är tillräckligt för att ge ett resultat som kan ge starka indikationer på hur romernas situation ser ut i de undersökta länderna (Ivanov, et al., 2006).

Totala antalet observationer i datamaterialet var 34 116 varvid 4 372 var relevanta för den här studien. Strukturen för datamaterialet var att de var indelat i familjer där samtliga familjemedlemmar var en observation, det fanns ingen nedre eller övre åldersgräns. Urvalet som har använts i den här uppsatsen är de observationer som är angivna som söner/döttrar i familjen. Alla observationer under 15 år har tagits bort på grund av att de ska ha haft möjlighet att avsluta grundskolan för att vara relevanta i regressionsanalysen. Datasetet var indelat i tre olika subgrupper som var romer, majoritetsbefolkningen samt flyktingar. Endast de två första grupperna har använts i regressioner då subgruppen flyktingar delvis låg utanför syftet med uppsatsen men också att det fanns begränsat med observationer för subgruppen. Att det fanns begränsat med observationer gjorde det svårt att få ett tillräckligt stort urval för gruppen flyktingar att göra en beräkning på vars resultat hade varit jämförbart med de två andra subgrupperna. Att endast de som var markerade som söner/döttrar har använts för beräkningen var för att förälders utbildningsnivå var en av de relevanta kontrollvariablerna. Förälders utbildningsnivå fanns inte observerat som en variabel i datamaterialet varvid denna behövde konstrueras. Därför har alla observationer som var söner/döttrar i datamaterialet använts som observationer för regressionen och kontrollvariabeln har skapats genom att använda den angivna utbildningsnivån från deras föräldrar. Observationer som var markerade som föräldrarna har därmed inte kunnat användas som observationer i regressionen utan endast för att konstruera kontrollvariabeln.

4.1 Deskriptiv statistik

Tabellen nedan visar deskriptiv statistik för variablerna som finns i regressionen.

Tabell 1

	Romer	Majoritetsbefolkningen
Medelinkomst, angivet i euro	178	406
Medelålder	21.4	21.2
Medelvärde antal utbildningsår	5.8	11.4
Medel förälders utbildningsår	5.5	11.5
Medel antal familjemedlemmar	6.7	4.9
Andel bosatta i romskt område	53.6	5.7
Andel bosatta i mixat område	32.4	25.6
Andel bosatta i majoritetsområde	14	68.7
Andel som talar landets officiella språk hemma	42.4	99.2
Andel som talar romani hemma	57.6	0.8
Andel med en arbetslösförälder	65.2	26.5
Män	62.3	55.9
Kvinnor	37.8	44.1

Från den deskriptiva statistiken kan vi se att medelinkomsten i romska familjer är avsevärt lägre än för majoritetsbefolkningen, den är mindre än hälften. Medelvärdet för antal familjemedlemmar är högre i romska familjer och medelvärdet för antalet utbildningsår är även lägre för både det romska urvalet och deras föräldrar. Att det finns en utbredd segregation kan vi också se eftersom över 53.6% av romerna är bosatta i områden där majoriteten också är romer och 68.7% av majoritetsbefolkningen är bosatta i områden där en majoritet är en del av majoritetsbefolkningen. En mycket stor andel av den romska urvalsgruppen är även arbetslösa, 65.2% av urvalets föräldrar är arbetslösa.

För att de två grupperna ska vara jämförbara är det en fördel ifall åldersspridningen inte skiljer sig för mycket i de olika grupperna. Histogrammen nedan visar åldersspridningen för de två olika urvalsgrupperna. Vi kan se att fördelningen är någorlunda lika för båda urvalen.

Diagram 1

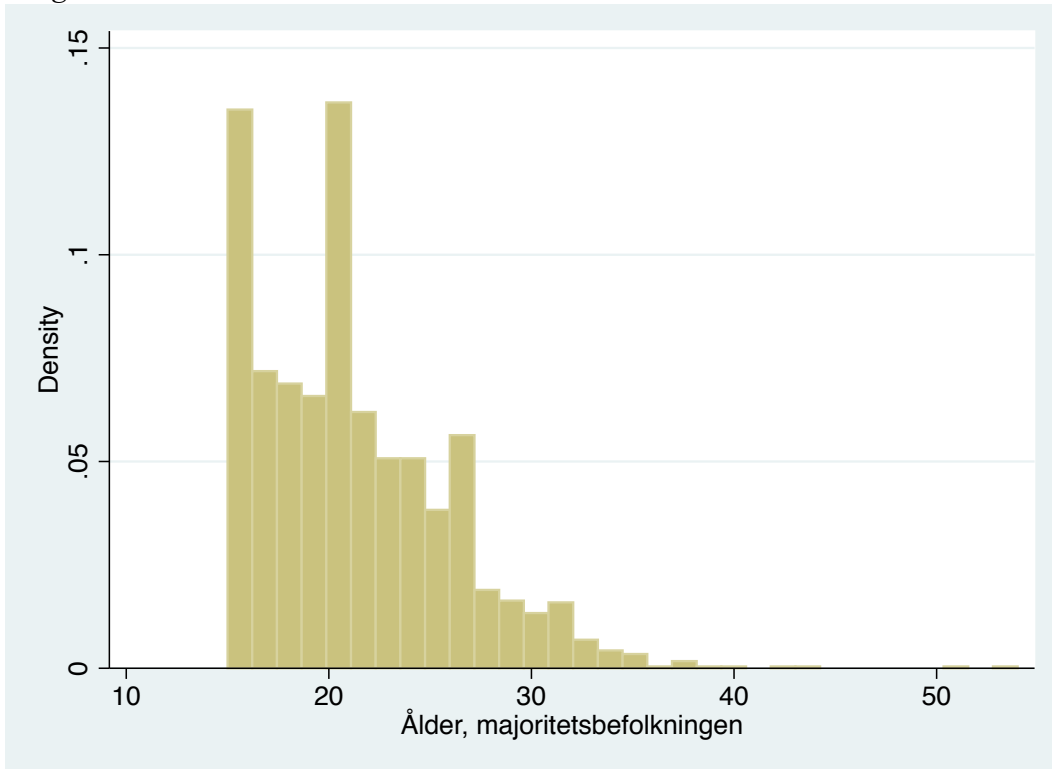
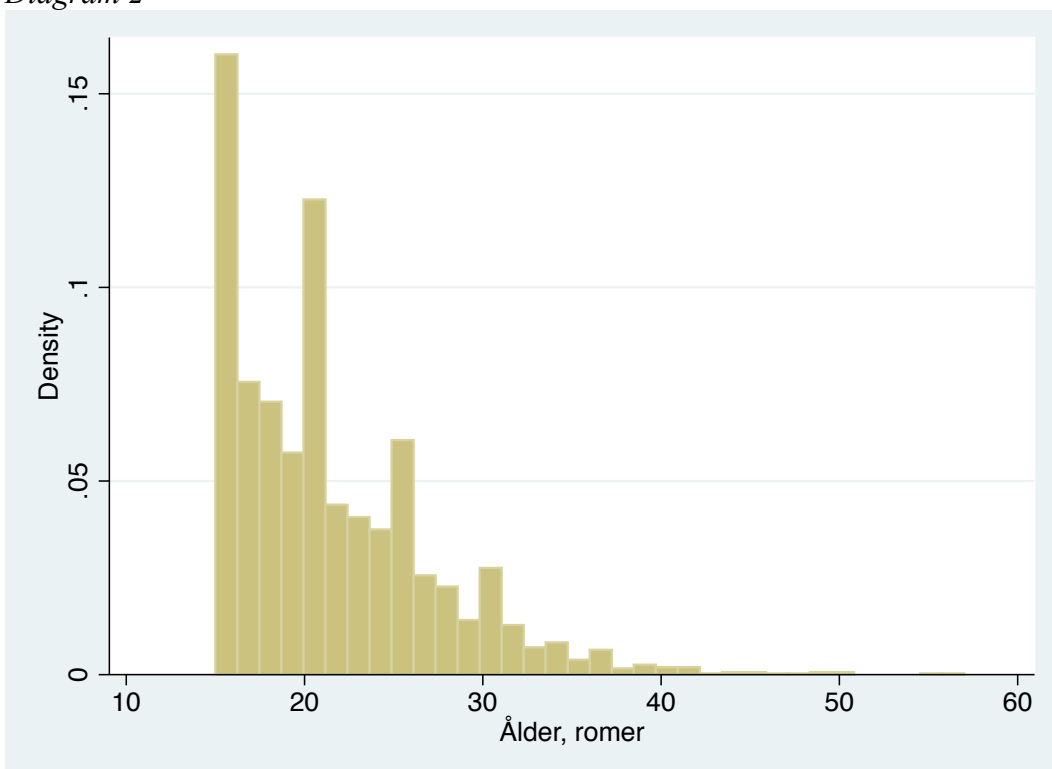


Diagram 2



5. Metod

5.1 Regressionsanalys

Tre olika regressionsanalyser har gjorts för att skatta vilka variabler som påverkar uppnådd utbildningsnivå, här mätt som antalet skolår. En samlad regression har gjorts för hela urvalet för att få ett generellt resultat samt två separata regressioner för båda etnicitetsgrupperna för att mäta ett antal faktorer som främst påverkar den ena gruppen.

5.2 Modell

Den modell som använts för regressionen är en linjär regressionsmodell med antal skolår som beroende variabel. Skattningarna har gjorts med minstakvadratmetoden, OLS. I skattningen för vad som påverkar den beroende variabeln har ett flertal oberoende kontrollvariabler använts. Modellen är baserad på en artikel skriven av Manski (1993) om endogena sociala effekter på individers beteende.

Modellen går att sammanfatta enligt följande:

$$\text{Skolår} = \alpha + \beta_1 \text{familjens inkomst} + \beta_2 \text{förälders utbildningsnivå} + \beta_3 \text{antal personer i hushållet} + \partial_1 \text{huvudsakligt språk som talas hemma} + \partial_2 \text{bosatt i mixat område} + \partial_3 \text{bosatt i majoritetsområde} + \partial_4 \text{förälders arbetsstatus} + \partial_5 \text{kön} + \partial_6 \text{land} + \varepsilon_i$$

Alla variabler är inte inkluderade i samtliga regressionerna. Detta beror på att exempelvis språkvariabeln endast är relevant för den romska gruppen.

Datasetet bestod som tidigare nämnt av 34 116 observationer som alla är indelade efter familj. Beräkningen har alltså gjorts på de som har varit söner/döttrar i familjen och alla individer under 15 år har tagits bort på grund av att de ska ha haft möjlighet att avsluta grundskolan för att vara relevanta i regressionsanalysen.

5.3 Beskrivning av variabler

5.3.1 Skolår

Antal skolår som individen har gått i skolan vid intervjutillfället.

5.3.2 Ålder

Uppnådd ålder den senaste födelsedagen vid intervjutillfället.

5.3.3 Familjens inkomst

Familjens inkomstnivå per månad angivet i euro.

5.3.4 Förälders utbildningsnivå

Den person som har definierats som ”Familjeöverhuvud” i familjen har använts som proxy för förälders utbildningsnivå. I majoriteten av fallen är det faderns utbildning men i ett fåtal fall är det moderns.

5.3.5 Huvudsakligt språk som talas hemma

Huvudsakligt språk som talas hemma är en dummyvariabel som jämför de som talar romani hemma med de som talar landets officiella språk. De flesta som har romani som huvudsakligt språk är romer men ett fåtal är av majoritetsetniciteten. Dock är antalet av majoritetsbefolkningen som talar romani så få att de inte går att göra en tillförlitlig beräkning på.

5.3.6 Typ av område för bosättning

Typ av område för bosättning är en grupp dummyvariabler som jämför hur utbildningsnivån uppskattas att påverkas av vilket bostadsområde urvalet är bosatt i. Områdena är indelade i tre kategorier som är områden där majoriteten är romer, mixade områden eller områden där majoriteten är en del av majoritetsbefolkningen. Variabeln ’romskt bostadsområde’ är referensvariabel i regressionen.

5.3.7 Förälders arbetssituation

Förälderns arbetssituation är en dummyvariabel som jämför det beräknade utfallet av utbildningsnivå ifall föräldern är arbetslös jämfört med referensvariabeln arbetande förälder.

5.3.8 Kön

En dummyvariabel som beräknar ifall män har högre eller lägre utbildning jämfört med kvinnor som är referensvariabeln.

5.3.9 Land

Dummyvariabler som visar på landspecifika skillnader i utbildningsnivå. Albanien är referensland.

5.4 Heteroskedasticitet och grupperade/robusta standardfel

För att skattningarna av koefficienterna i regressionen ska vara konsistenta och effektiva är det viktigt att feltermerna inte är heteroskedastiska, det vill säga att variationen i feltermerna beräknas vara densamma.

Samtliga regressioner har testats för heteroskedastiska feltermar och nollhypotesen har inte kunnat förkastas i någon av regressionerna, alltså har heteroskedasticitet varit förekommande. Detta har hanterats genom att använda grupperade standardfel ("clustered standard errors"). Detta tillåter feltermerna att vara korrelerade på individnivå men de tillåts inte vara korrelerade mellan olika individer. Skattningarna i regressionen är på grund av detta konsistenta och effektiva.

6. Resultat

Som tidigare nämnt har tre regressioner gjorts, en samlad regression med etnicitet som dummyvariabel och två separata, en för respektive etnicitetsgrupp. En samlad regression med interaktionsvariabler undersöktes först som modell men för många variabler behövde uteslutas på grund av en för hög grad av multikollinearitet för att modellen skulle vara relevant. Ett flertal av variablerna uppvisade en hög korrelation med den romska etnicitetsdummin i modellen med interaktionsvariabler vilket indikerar på att ett flertal av faktorerna främst påverkar den romska gruppen. Exempel på detta är de olika områdesvariablerna samt språkvariabeln, så detta kommer primärt analyseras för den romska gruppen.

Tabell 2 – samlad regression

Utfallsvariabel: Antal skolår	Koefficient	Robusta Standardfel	p-värde
Intercept	4.131191	.3761689	0.000
Romsk etnicitet	-2.694373	.1696353	0.000
Ålder	.1422153	.0108199	0.000
Familjens inkomst	.0010297	.0002004	0.000
Förälders utbildning	.355759	.0191582	0.000
Antal familjemedlemmar	-.1944773	.0260208	0.000
Mixat område	.3140389	.1631143	0.054
Majoritetsområde	.338408	.1674814	0.044
Arbetslös förälder	-.3843096	.1311341	0.003
Man	.2262629	.0897317	0.012
<i>Bulgarien</i>	.5959282	.1987011	0.003
<i>Bosnien-Hercegovina</i>	-.1478267	.2334016	0.527
<i>Kroatien</i>	.7018495	.2584502	0.007
<i>Makedonien</i>	1.212402	.1863063	0.000
<i>Serbien</i>	.5678125	.2229005	0.011
<i>Montenegro</i>	-.0797444	.3012276	0.791
<i>Rumänien</i>	.3894582	.2154201	0.071
<i>Kosovo</i>	.247148	.2034323	0.225
R²	0.6038		
Antal observationer	4372		

*Referensvariabel för "romsk etnicitet" är majoritetsbefolkning. Referensvariabel för "mixat område" samt "majoritetsområde" är dummyvariabeln "romskt område". Referensvariabel för dummyvariabeln "arbetslös förälder" är "förälder med anställning". Referensland för de olika landdummyvariablerna är Albanien.

Det som kan utläsa ifrån den samlade regressionen är att jämfört med majoritetsbefolkningen så har den romska subgruppen lägre utbildningsnivå, den beräknade generella skillnaden är så stor som 2.7 år. Ålder, familjens inkomst och förälders utbildningsnivå är faktorer som enligt beräkningen ökar utbildningsnivån. Av variabeln 'antal familjemedlemmar' kan vi se att det påverkar utbildningsnivån negativt att vara många familjemedlemmar. Jämfört med att bo i ett 'romskt område' så har både bosättning i 'mixat område' och 'majoritetsområde' en positiv effekt på utbildningsnivån. En viktig sak att notera angående områdesvariablerna är att endast effekten av bosättning i 'majoritetsområde' är på en 95% signifikansnivå. P-värdet för 'mixat område' är 0.054 och är därmed endast signifikant på en 90% signifikansnivå. Regressionen visar även på att det har en negativ inverkan på utbildningsnivån att ha en arbetslös förälder. Jämfört med en arbetande förälder så har personer med en arbetslös förälder 0.38 år lägre utbildning. Vi kan även utläsa från den samlade regressionen att män har generellt 0.23 år längre utbildning jämfört med kvinnor.

6.1 Separata regressioner

Tabell 3- Regression för den romska urvalsgruppen

Utfallsvariabel: Antal skolår	Koefficient	Robusta Standardfel	p-värde
Intercept	1.708363	.4456292	0.000
Ålder	.0938163	.013594	0.000
Familjens inkomst	.0016452	.0005486	0.003
Förälders utbildning	.4053574	.0272384	0.000
Antal familjemedlemmar	-.1973837	.0325814	0.000
Landets officiella språk talas hemma	.7040619	.1914494	0.000
Mixat område	.3631578	.1909597	0.058
Majoritetsområde	.2615038	.2314357	0.259
Arbetslös förälder	-.551977	.1881024	0.003
Man	.6082203	.1311742	0.000
<i>Bulgarien</i>	.7495202	.2995184	0.013
<i>Bosnien-Hercegovina</i>	-.4609722	.3786499	0.224
<i>Kroatien</i>	1.005198	.4107129	0.015
<i>Makedonien</i>	1.481394	.3240343	0.000
<i>Serbien</i>	.7859886	.3611684	0.030
<i>Montenegro</i>	-.5463818	.611128	0.372
<i>Rumänien</i>	.5466165	.3154803	0.084
<i>Kosovo</i>	-.4731838	.3846747	0.219
R²	0.3480		
Antal observationer	2466		

*Referensvariabel för "mixat område" samt "majoritetsområde" är dummyvariabeln "romskt område". Referensvariabel för dummyvariabeln "landets officiella språk talas hemma" är "romani talas hemma". Referensvariabel för dummyvariabeln "arbetslösförälder" är "förälder med anställning". Referensland för de olika landdummyvariablerna är Albanien.

Tabell 4 - Regression för urvalsgruppen som tillhör majoritetsbefolkningen

Utfallsvariabel: Antal skolår	Koefficient	Robusta Standardfel	p-värde
Intercept	4.813245	.5319446	0.000
Ålder	.239338	.0138527	0.000
Familjens inkomst	.0009015	.0001841	0.000
Förälders utbildning	.2310703	.023981	0.000
Antal familjemedlemmar	-.2058702	.0366768	0.000
Mixat område	-.0002759	.2523599	0.999
Majoritetsområde	-.0852194	.2472283	0.730
Arbetslös förälder	-.2176667	.1454972	0.135
Kön (=man)	-.3042562	.1013049	0.003
<i>Bulgarien</i>	-.4934485	.2471077	0.046
<i>Bosnien-Hercegovina</i>	-.2774213	.192649	0.150
<i>Kroatien</i>	-.8063787	.3235451	0.013
<i>Makedonien</i>	.060038	.1961196	0.760
<i>Serbien</i>	-.1538558	.2120626	0.468
<i>Montenegro</i>	-.1867811	.2313033	0.420
<i>Rumänien</i>	-.7854152	.289991	0.007
<i>Kosovo</i>	-.1936585	.2027926	0.340
R²	0.3505		
Antal observationer	1906		

*Referensvariabel för "mixat område" samt "majoritetsområde" är dummyvariabeln "romskt område. Referensvariabel för dummyvariabeln "arbetslösförälder" är "förälder med anställning". Referensland för de olika landdummyvariablerna är Albanien.

Det vi kan utläsa ifrån de separata regressionerna är att det för den romska gruppen är positivt för antalet utbildningsår ifall landets officiella språk talas hemma jämfört med ifall romani talas hemma. Enligt regressionen så är både bosättning i 'mixat område' samt 'majoritetsområde' något som skulle kunna ha en positiv inverkan på utbildningsnivån, men viktigt att notera är de höga p-värdena för dessa variabler som gör att på en 95% konfidensnivå så är dessa resultat inte signifikanta. Resultatet för 'mixat område' är signifikant på en 90% signifikansnivå och resultatet för 'majoritetsområde' är inte statistiskt signifikant alls.

I regressionen för majoritetsbefolkningen så är språkvariabeln exkluderad på grund av att det var ett för litet urval som inte talade landets officiella språk hemma. P-värdena för områdesvariablerna är så höga i den separata regressionen för majoritetsbefolkningen att de inte är signifikanta.

Gemensamt för de separata regressionerna och den samlade regressionen är att ålder, förälders utbildningsnivå samt familjens inkomst har en positiv effekt på utfallsvariabeln antal utbildningsår. Samtliga regressioner visar även att antal familjemedlemmar har en negativ effekt. Vad gäller dummyvariabeln arbetslös förälder så är resultatet signifikant i den samlade regressionen och i den separata regressionen för romer. I den separata regressionen för majoritetsbefolkningen är p-värdet för högt för att resultatet ska vara signifikant. Regressionerna med ett signifikant resultat visar på att en arbetslösförälder har en negativ inverkan på utbildningsnivån.

7. Diskussion och Analys

Resultatet visar att det finns en form av områdeseffekt eftersom det i den samlade regression samt i regressionen för den romska gruppen har en positiv inverkan på utbildningsnivån ifall de bor i ett mixat område eller i ett majoritetsområde jämfört med bosättning i ett romskt område. Regressionen för majoritetsbefolkningen kan dock inte bekräfta hypotesen eftersom resultatet i den regressionen inte är signifikant. Den här studien har ett ganska litet urval av majoritetsbefolkningen som är bosatta i ett område med majoriteten romer, så för ett statistiskt signifikant resultat hade urvalet förmodligen behövt vara större. Generellt är det problematiskt att områdesvariablerna hade höga p-värden, men på en 90% konfidensnivå kan vi ändå utläsa en områdeseffekt, primärt från resultatet i den samlade regressionen.

Områdeseffekter är något som är svårt att mäta eftersom det både kan överskattas men även underskattas. Eftersom andra faktorer även påverkar, vilket vi kan se i regressionerna att både inkomst, förälders utbildning och antal familjemedlemmar gör så blir frågan vilka andra faktorer som orsakar den skillnad som de skattade koefficienterna inte kan förklara, det vill säga varför är skillnaden i de skattade intercepten så stor för de olika grupperna och varför visar den romska etnicitetdummys ett så högt värde. Medelvärdet för utbildningsnivå är avsevärt högre för majoritetsbefolkningen än vad det är för den romska befolkningen och variablerna i regressionsanalysen har inte fullt ut kunnat förklara den skillnaden. Teorin om områdeseffekter har sin utgångspunkt i att vi påverkas av rådande normer i vår omgivning. Eftersom det finns en utbredd segregation så kan det innebära att olika normer råder i de olika bostadsområdena. Förutom den segregation som ger utslag genom vart människor är bosatta så kanske det även finns en form av 'social segregation', det vill säga att även de som är bosatta i områden med mixade etniciteter eller där majoriteten är av annan etnicitet än en själv ändå främst umgås med människor som tillhör samma etnicitet som en själv. Det skulle därmed ha kunnat bildats olika normer mellan de olika etniciteterna som förklarar varför romerna har en lägre utbildningsnivå än majoritetsbefolkningen. Även den *spillover* effekt som Borjas (1995) framhåller kan vara en av orsakerna som förklarar skillnaden. Om det finns en 'social segregation' kommer identifikationen förmodligen vara mycket starkare till den egna etniciteten än vad den är till de personer som bosatta närmst oss. Detta gör att de faktorer som teorin om områdeseffekter pekar på kanske visar sig i de olika etnicitetsvariablerna i högre grad än i områdesvariablerna ifall det är inom den egna etniciteten det sociala nätverket finns.

Områdeseffekterna handlar som tidigare nämnt om hur vi påverkas av de normer som finns runt oss och hur vi imiterar de beteenden som människor i vår närhet har. Att växa upp med föräldrar med ingen eller mycket låg utbildning och att det även är normen i andra familjer i ens sociala nätverk kommer troligtvis påverka vilka val en individ tar när det gäller utbildning. Delvis påverkar det rimligen synen på utbildning och vikten av den men också vilka förutsättningar en person har att klara av skolan. Att hoppa av skolan i unga år är inte lika kontroversiellt ifall det är mer en regel än ett undantag i ens omgivning. Ifall man kanske inte ens känner någon med en universitetsutbildning som barn så är det kanske inte lika troligt att det är något man tidigt börjar sträva efter. På grund av den segregation som existerar så skulle det kunna vara så att olika normer finns för majoritetsbefolkningen och den romska gruppen.

Förutom områdeseffekten så finns det fler aspekter från regressionen som påverkar den uppnådda utbildningsnivån. En viktig aspekt är exempelvis inverkan som vilket modersmål människor har. Även ifall romerna i de flesta fall är födda och uppvuxna i det land de är bosatta i så talar 57.6% av romerna i urvalet främst romani hemma. Detta kan komma att innebära en nackdel när skolundervisningen sker på vad som då blir deras andra språk. Enligt den separata regressionen för den romska gruppen så har romer som talar samma språk hemma som majoritetsbefolkningen 0.7 år kortare skolgång. Det visar inte på att den enskilda variabeln är avgörande för utbildningsnivån men det visar ändå att det är något som påverkar och att det kan innebära en nackdel för de som har ett annat modersmål än landets officiella språk.

Andra aspekter som den här studien inte har kunnat mäta är ifall det kan vara så att romer på något sätt blir diskriminerade inom utbildningsväsendet eller ifall det helt enkelt är ett rationellt val att investera kortare tid i utbildning. O'Higgins studie visade att romer får lägre avkastning för tid de investerat i utbildning jämfört med majoritetsbefolkningen. Detta kan bidra till en lägre motivation att utbilda sig och cementera den rådande normen att inte satsa på utbildning. Ifall romer på något sätt blir diskriminerade i skolan genom exempelvis negativ särbehandling på grund av fördomar eller ifall de systematiskt inte ges samma möjligheter har det inte funnits några lämpliga kontrollvariabler för och det är även något som är svårt att mäta och skulle därmed behöva en annan typ av studie. De 2.7 år i utbildningsnivå som skiljer romer och majoritetsbefolkningen åt kan dock alltså vara ett rationellt val av romerna att inte

investera i utbildning till samma nivå som majoritetsbefolkningen på grund av den lägre förväntade avkastningen som O'Higgings studie visat på.

Sammanfattningsvis så har vilket bostadsområde människor är bosatt i någon form av inverkan, även om regressionerna inte visar att det förklarar den stora skillnaden i utbildningsnivå mellan romer och majoritetsbefolkningen. Ekonomiska förhållanden har även en ganska så stor inverkan och vilka familjeförhållanden vi kommer ifrån påverkar också. Dessa faktorer kan även komma att påverkas av vart människor är bosatta. Vad exakt skillnaden mellan etniciteterna beror på som regressionen visar kan inte säkert fastställas med utgångspunkt i de regressioner som gjorts. Det kan vara så att det är ett rationellt val av romerna att investera kortare i utbildning på grund av en lägre förväntad avkastning. De kan också vara så att olika normer råder mellan de olika grupperna och att detta är sammankopplat med en form av 'social segregation'. Det är även mycket möjligt att de på flera plan har olika förutsättningar för att uppnå en högre utbildningsnivå och att det är detta som skillnaden bottnar i.

8. Sammanfattning

Den tidigare forskningen på lönediskriminering mot romer på arbetsmarknaden i sydöstra Europa visade den primära förklaringen till löneskillnader var skillnad i utbildningsnivå mellan romer och majoritetsbefolkningen.

Enligt teorin om områdeseffekten påverkas människor av valen som andra i deras närhet gör när det gäller utbildning eftersom vi rent psykologiskt påverkas av de normer som råder i vår närhet. Den mätbara effekten utifrån datamaterialet visade att romer bosatta i ett mixat område ökar utbildningsnivån med 0.36 år och att bosättning i ett majoritetsområde ökar utbildningsnivån med 0.26 år. Den effekten i sig visar inte på att områdeseffekten är en primär förklaring till skillnaderna i ackumulerat humankapital mellan romer och majoritetsbefolkningen, men det påverkar. Det kan även vara så att rådande normer inom den egna folkgruppen är starkare än normerna i vårt eget bostadsområde varvid normer och sociala kostnader kan vara aspekter som påverkar utfallet. När familj, släktingar och vänner har en låg utbildning eller kanske inte någon utbildning alls är ett avhopp från skolan inte nödvändigtvis förknippat med ett stigma eller andra sociala kostnader vilket kan vara fallet där andra normer är rådande.

För att kunna ge ett mer allomfattande svar på varför utbildningsnivån skiljer sig mellan romer och majoritetsbefolkningen hade det varit en fördel att göra en studie med tidsseriedata för att se vilka trender som kan utläsas över tid mellan de olika grupperna. Till exempel så hade det då varit intressant att se ifall utbildningsnivån ökar när familjens inkomst ökar i båda grupperna eller ifall grupperna uppvisar olika trender. Även en närmare studie av upplevd diskriminering av institutioner och hur det påverkar tilliten hade varit en möjlig framtida studie för att se ifall det mellan de etniska grupperna finns olika nivåer av förtroende till myndigheter och institutioner som eventuellt kan komma att påverka attityden till utbildning samt förtroende till utbildningsväsendet. För att kunna dra några mer definitiva slutsatser om områdeseffekter hade förmodligen både ett större och ett mer specificerat urval behövts för att undersöka på vilket sätt områdeseffekter påverkar skillnaden i utbildningsnivå mellan romer och majoritetsbefolkningen. Utifrån urvalet har vi kunnat se att områdeseffekten förmodligen har en viss inverkan på utbildningsnivå för den romska gruppen och att de sämre ekonomiska förutsättningarna de har ger en negativ effekt på uppnådd utbildningsnivå. En del av de faktorer som sammankopplas med områdeseffekter kan även ha tagit sig formen av att olika

normer har etablerats mellan de olika etnicitetsgrupperna samt att det kan finnas *spillover* effekter som har en koppling till etnicitet snarare än de olika bostadsområdena. Men det kan även vara så att de inte riktigt har samma förutsättningar att uppnå samma nivå på utbildning. Romerna är i stor utsträckning arbetslösa, har en avsevärt lägre medelinkomst än majoritetsbefolkningen och många lever segregat. Det skapar många negativa spiraler men den gemensamma nämnaren i beräkning är att de alla är romer. En djupare studie om både hur normer skiljer sig men även förutsättningar mellan de olika grupperna hade behövts göras för att kunna ge ett mer genomgripande svar på varför det är en så stor skillnad på uppnådd utbildningsnivån mellan grupperna.

Litteraturförteckning

Achim, V., 1998. *The Roma in Romanian History*. 1:a upplagan ed. Budapest: Central European University Press.

Borjas, G., 1995. Ethnicity, Neighborhoods, and Human-Capital Externalities. *THE AMERICAN ECONOMIC REVIEW*, June, pp. 365-390.

Borjas, G. J., 2016. *Labor Economics*. Seventh Edition ed. New York: McGraw-Hill Education.

Brüggemann, C. & Friedman, E., 2017. The Decade of Roma Inclusion: Origins, Actors, and Legacies. *European Education*, Volume 49, pp. 1-9.

Durlauf, S. N., 2004. Neighborhood effects. In: J. H. & J. Thisse, ed. *Handbook of Regional and Urban Economics*. 1:a upplagan ed. Elsevier B.V.: s.n., pp. 2173-2242.

FRA- European Union Agency for Fundamental Rights, 2014. *Roma survey - Data in focus. Poverty and employment: the situation of Roma in 11 EU Member States*, Luxembourg: European Union Agency for Fundamental Rights.

Ivanov, A. et al., 2006. *At risk: Roma and the Displaced in Southeast Europe*, Bratislava: United Nations Development Programme, Regional Bureau for Europe and the Commonwealth of Independent States.

MANSKI, C. F., 1993. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, Volume 60, pp. 531-542.

Milcher, S., 2011. DECOMPOSING INCOME DIFFERENTIALS BETWEEN ROMA AND NON-ROMA IN SOUTH EAST EUROPE. *Romanian Journal of Regional Science, The Journal of the Romanian Regional Science Association*, Summer, 5(1).

Milcher, S. & Fischer, M. M., 2011. On labour market discrimination against Roma in South East Europe. *Papers in Regional Science*, November, Volume 90.

Nieuwenhuis, J. & Hooimeijer, P., 2016. The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *P. J Hous and the Built Environ*, June, Volume 31, p. 321-347.

O'Higgins, N., 2010. "It's not that I'm a racist, it's that they are Roma" Roma discrimination and returns to education in South Eastern Europe. *International Journal of Manpower*, 31(2), pp. 163-187.

United Nations Development Programme, 2018. *Roma Data*. [Online]
Available at: <http://www.eurasia.undp.org/content/rbec/en/home/ourwork/sustainable-development/development-planning-and-inclusive-sustainable-growth/roma-in-central-and-southeast-europe/roma-data.html>
[Accessed 17 05 2018].