



UPPSALA
UNIVERSITET

STS

Examensarbete 30 hp
Januari 2016

Sambandsanalys av sociala konsekvenser vid utbyggnad av transportsystem

En granskning av statistisk modellering
för nyttobedömning av höghastighetsjärnväg

Albin Engholm
Gabriel Johansson



UPPSALA
UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
<http://www.teknat.uu.se/student>

Abstract

Analysis of relationships between investments in transport infrastructure and social consequences

Albin Engholm and Gabriel Johansson

Sweden is currently planning for one of its largest investments in transport infrastructure ever. A cornerstone in the project is a new high speed railroad between Stockholm, Gothenburg and Malmö. A considerable portion of the funding will consist of co-funding from municipalities, regions and private actors that will benefit from the infrastructure. The amount of co-funding is based on the total benefits that each actor will receive. Therefore these benefits must be able to be quantified.

This report aims to study relationships between transport infrastructure and a number of social consequences that previous research has indicated being interesting. The method used is based on econometric methods that enable identification of statistical relationships using a simple linear model. This requires measures of transport infrastructure and indicators for social well-being which are chosen through a study of literature and established models quantifying wider economic benefits from transport infrastructure. The report suggests that a gravitybased accessibility measure is suitable for quantifying transport infrastructure in this context. The results from the econometrical models suggest that there are several significant relationships between accessibility and social well-being. In general accessibility seems to correlate with positive social consequences. However it is likely that endogeneity is biasing the results and make it difficult to draw any conclusions on causal relationships. It is stated that further research studying these relationships more carefully with better adapted data for the purpose is required to be able to obtain quantified estimates that can be used in co-funding of transport infrastructure.

Handledare: Anna Modin
Ämnesgranskare: Anders Karlström
Examinator: Elísabet Andrésdóttir
ISSN: 1650-8319, UPTec STS 16 008

Sammanfattning

Sverige planerar för en av de största infrastrukturinvesteringarna någonsin i Sveriges historia. I fokus står en ny höghastighetsjärnväg mellan Stockholm, Göteborg och Malmö som ska möjliggöra resor på två till två och en halv timme mellan städerna. Runt denna investering finns en ambition att samla andra samhällsbyggande åtgärder, som ökat bostadsbyggande och förbättringar i lokal och regional kollektivtrafik, för att genom en samordnad planering få ut så stora nyttor som möjligt. För detta ändamål har regeringen tillsatt den statliga utredningen Sverigeförhandlingen. En viktig del av deras arbete är att få fram en finansieringslösning för den nya statliga transportinfrastrukturen där medfinansiering från kommuner, regioner och andra aktörer står för en betydande andel. Idén kring medfinansiering är baserad på att aktörerna ska bidra till finansieringen som motsvarar de nyttor transportinfrastrukturen bidrar till. Alltså måste det gå att uppskatta att på förhand hur mycket nyttor som uppstår och vad de är värda. Sverigeförhandlingen har valt klassificera nyttorna i sex kategorier där ”sociala nyttor” är en av dessa. Däremot finns ingen etablerad metodik för att uppskatta dessa.

Det här examensarbetet undersöker vilken social påverkan transportinfrastrukturen ger upphov till. Detta görs genom att studera samband mellan tillgänglighet, alltså hur transportinfrastrukturen ger möjlighet för människor att nå olika målpunkter, och ett antal sociala förhållanden. Tillgänglighetsmättet valdes utifrån studier av tidigare forskning samt etablerade modeller för uppskattning av andra nyttor av transportinfrastruktur. För att veta vilka sociala förhållanden som är intressanta att undersöka har en studie av Svensk och internationell litteratur på området genomförts. Sambandsanalysen bygger på statistiska metoder där data för Sveriges kommuner används. Metoden ger möjlighet till att identifiera samvariation mellan tillgänglighet och sociala förhållanden men är inte tillräcklig för att säkerställa att tillgänglighet är en orsak till de sociala förhållandena.

Resultaten tyder på att tillgänglighet har statistiskt signifikanta samband med ett antal indikatorer för sociala förhållanden. Exempel på två av dessa är ungdomsidrottande och andelen barn som går i förskola. Det visade sig att de flesta av de sociala förhållanden som har signifikanta samband med tillgänglighet har en samvariation som tyder på att tillgänglighet korrelerar med positiva sociala förhållanden. Detta tyder på att tillgänglighet kan ha samband med faktorer som kan tolkas som sociala nyttor. Dock kan inget sägas om tillgänglighet är en orsak till detta eller om det är någon icke studerad orsak som ger upphov till detta. Dessutom påverkas tillgänglighetsmättet av andra faktorer än enbart tillgång till transportinfrastruktur och därför går det inte utifrån dessa samband att säkert säga att de beror på transportsystemet.

För att sociala nyttor ska kunna kvantifieras och ligga till grund för medfinansiering krävs vidare studier. En viktig uppgift är att med hjälp av bättre data och mer

avancerade metoder få fram bättre uppskattningar av sambanden samt orsaksrelationerna. Det krävs också studier för hur sociala nyttor ska värderas i monetära termer ur ett samhällsekonomiskt perspektiv.

.

Förord

Detta examensarbete är det slutliga arbetet för civilingenjörsutbildningen inom System i Teknik och Samhälle på Uppsala universitet. Vi som skrivit arbetet valde att göra det inom transportsystem, ett för oss helt nytt område som inte har varit inkluderat i vår utbildning men som vi båda var intresserade av och kände att vi ville ta chansen att dyka in i. Det var en enorm utmaning att angripa ett helt nytt fält men det visade sig enormt lärorikt och vi tar med oss en stor portion ödmjukhet och beundran inför den expertis och erfarenhet vi stött på under arbetets gång.

Först och främst vill vi passa på att tacka vår mycket proffsiga och skickliga handledare Anna Modin och hennes kollegor på Sverigeförhandlingen för att vi fick chansen att göra vårt examensarbete hos dem. Vi är tacksamma för alla de möten som vi hade med Anna från vilka vi alltid kom ut ifrån med ny inspiration och motivation.

Sedan vill vi också tacka vår ämnesgranskare Anders Karlström som gav oss många viktiga tips som var till stor hjälp för oss i vårt examensarbete. Han guidade oss med fast hand in i ett vetenskapsfält i vilket vi till en början kände oss väldigt vilsna inom. Med hans hjälp kunde vi snabbt navigera oss inom fältet och få en uppfattning om vad vi kunde tillföra.

Sist men inte minst vill vi tacka Elísabet Andrésdóttir som funnits till hand och gett oss snabba svar på frågor rörande det administrativa.

Innehållsförteckning

1. Inledning	1
1.1 Syfte	2
1.2 Disposition	2
2. Bakgrund.....	4
2.1 Nya finansieringsmodeller för transportinfrastruktur	4
2.2 2013 års Stockholmsförhandling och Sverigeförhandlingen	6
2.2.1 Nyttobedömningar i Sverigeförhandlingen	9
2.3 Samhällsekonomiska analyser	11
2.4 Lokaliserings- och tillväxtmodeller	13
2.5 Vad är social nytta?	15
2.6 Tidigare forskning	18
3. Metod.....	27
3.1 Tillgänglighet som mått på transportinfrastrukturens funktion	28
3.2 En enkel modell för samband mellan transportinfrastruktur och sociala förhållanden	33
3.3 Linjär regression	36
3.3.1 Pearsons korrelationskoefficient	37
3.3.2 Logaritmisk transformering	37
3.4 Paneldataanalys	38
3.4.1 Between effects regression/ tvärsnittsregression	40
3.5 Kommunspecifika effekter	40
3.5.1 Regression med fixa effekter	41
3.5.2 Regression med slumpmässiga effekter	42
3.6 Tid- och kommunspecifika fixa effekter	43
3.7 Metodproblem.....	44
3.7.1 Kausalitet	44
3.7.2 Endogenitet	46
3.7.3 Heteroskedasticitet	48
3.7.4 Robusta standardavvikelser	49
3.7.5 Kolinjäritet	49
3.7.6 Spatial korrelation	49
3.8 Programvara	50
3.9 Arbetsgång	50
4. Data.....	52
4.1 Geografisk indelning på kommunnivå	52
4.2 Oberoende variabler	53

4.2.1	Tillgänglighet	54
4.2.2	Socioekonomiska faktorer	55
4.2.3	Kolinjäritet	56
4.2.4	Observationernas fördelning	58
4.3	Beroende variabler – Sociala förhållanden	59
4.3.1	Förväntad livslängd	61
4.3.2	Deltagande i senaste riksdagsval	61
4.3.3	Sjukpenningtalet	62
4.3.4	Lönegap mellan kvinnor och män	62
4.3.5	Vård av Barn (VAB) och Föräldrapenning	63
4.3.6	Gymnasiefrekvens	63
4.3.7	Förvärsarbetande	63
4.3.8	Deltagartillfällen i idrottsföreningar	63
4.3.9	Bidragshushåll	64
4.3.10	Barn inskrivna i förskola och skolbarnsomsorg	64
4.3.11	Eftergymnasial utbildning.	65
4.3.12	Relevanta sociala effekter som inte ingår i undersökningen	65
4.3.13	Förhållanden mellan de beroende variablerna	65
4.4	Bortfall av observationer	69
4.5	Datasetets korrelationer	69
4.6	Observationernas heterogenitet	70
4.7	Datastruktur	72
5.	Resultat	74
5.1	Mått på transportinfrastruktur	74
5.2	Kvantifiering av sociala nyttor	74
5.3	Tvårsnittsestimering – Between effects	74
5.4	Estimering med kommunspecifika effekter	78
5.5	Estimering med kommunspecifika och tidsspecifika effekter - Fixed effects	81
5.6	Sammanfattande resultat av paneldatanalys	85
5.7	Osäkerhet och felkällor	86
5.7.1	Tillgänglighetsdata	86
5.7.2	Endogenitet	87
5.7.3	Generaliserade modeller	88
5.7.4	Hantering av stora datastrukturer	88
5.7.5	Offentlig statistik	88
6.	Diskussion	89
6.1	Relation till tidigare forskning	89
6.2	Uppslag för vidare studier	90

7. Slutsatser	92
Referenser	94
Appendix A.....	100
Appendix B.....	102

1. Inledning

”Det blir den största järnvägsutbyggnaden sedan 1800-talet”. Så uttryckte sig Liberalernas¹ partiledare Jan Björklund i en artikel i Ny Teknik när han talade om utbyggnaden av höghastighetsjärnväg mellan Stockholm, Göteborg och Malmö (Nohrstedt, 2014). En investering som av Trafikverket under slutet av 2015 beräknas kosta mellan 190 – 320 miljarder kronor och vara färdigställd runt år 2035 (Trafikverket, 2015 a). Det var under den andra halvan av 1800-talet som Sverige byggde ut de nuvarande stambanorna. Utbyggnaden var då präglad utav industrialismen med den stora drivkraften att binda samman städer och möjliggöra snabbare och billigare godstransporter. Med den nya järnvägen kunde sträckan Stockholm – Göteborg avverkas på 14 timmar, en resa som tidigare, med häst, tagit en hel vecka. En annan följd av utbyggnaden av järnväg var att Sverige fick gemensam tid för hela landet, då det krävdes att tiden var densamma på alla järnvägsstationer. Järnvägen drev också på utvecklingen inom andra områden, exempelvis skapades ett samarbete så att telegrafverket kunde använda sig utav stolplinjerna längs järnvägarna då det ändå behövdes telegrafförbindelse mellan järnvägsstationerna (Mild, 2009). Enligt Kim Salomon, professor i historia vid Lunds universitet, bidrog järnvägen till en nationell samhörighet då den band samman olika delar av det svenska samhället både fysiskt och mentalt (Salomon, 2012). Folk från olika samhällsklasser reste tillsammans, i samma hastighet om än med olika klass på vagnarna.

Nu är det alltså dags för en ny utbyggnad av det Svenska järnvägssystemet, den här gången med höghastighetsjärnväg mellan Stockholm, Göteborg och Malmö. Vilken ska möjliggöra tågresor mellan Stockholm och Göteborg på två timmar och mellan Stockholm och Malmö på två och en halv timmar. Idag är syftet främst att skapa en ökad tillgänglighet mellan olika regioner och på så sätt uppnå en ökad tillväxt och hållbart resande. Planeringen av utbyggnaden handlar mycket om att maximera de nyttor som uppstår ut av en transportinfrastrukturutbyggnad. Inte bara de nyttor som uppstår direkt utav järnvägen i sig, utan också indirekta nyttor (Näringsdepartementet, 2014 a). I samband med utbyggnaden av höghastighetsjärnväg finns en politisk vilja att samordna detta med bostadsbyggande och andra samhällsbyggnadsåtgärder. En stor skillnad mellan utbyggnaden av stambanorna på 1800-talet och dagens utbyggnad är att det då ansågs som enbart ett infrastrukturbygge medan det idag, med alla de olika aspekter som tas hänsyn till, är mer av ett samhällsbygge. För att detta samhällsbygge ska ske effektivt och för att så stora nyttor som möjligt ska skapas tillsatte regeringen Sverigeförhandlingen den 1 juli 2014. Sverigeförhandlingen är en kommitté med uppdrag att hitta en lösning för utbyggnaden som ger; maximal samhällsekonomisk lönsamhet, ökad tillgänglighet i transportsystemet, ökat bostadsbyggande i storstäderna samt en finansieringsform där medfinansiering från näringsliv, kommuner och regioner står för en betydande andel (Näringsdepartementet, 2014 a). Detta är en relativt ny

¹ Dåvarande Folkpartiet

finansierings- och planeringsform som ställer krav på metoder och modeller för att uppskatta dessa nyttor. Sverigeförhandlingen delar upp nyttorna i sju olika områden: Arbetsmarknadsnytta, Bostadsmarknadsnytta, Restidsnytta, Näringslivsnytta, Miljönytta och Social nytta. Under Sverigeförhandlingens inledande skede ligger fokus på att undersöka vilka nyttor som skapas genom de infrastrukturprojekt som förhandlingen omfattar. Bedömningen av hur stora nyttor infrastrukturutbyggnaden skapar för respektive part ligger till grund för hur finansieringen ska fördelas. förhandlingsarbetet. Det finns etablerade metoder för att uppskatta alla dessa nyttor utom social nytta. I samarbete med Sverigeförhandlingen utreder det här arbetet om det är möjligt att uppskatta social nytta med liknande metod som används i de redan etablerade metoderna för de övriga nyttorna. Förhoppningen är att därigenom ta ett steg mot att underlätta för att skapa tydliga och väl underbyggda bedömningar av den sociala nyttan som kan vägas in i förhandlingsarbetet.

1.1 Syfte

Arbetet har som mål att med statistisk metod, kvantitativt undersöka samband mellan transportinfrastruktur och social nytta samt granska hur Sverigeförhandlingen bör förhålla sig till dessa samband i deras arbete. Därmed ämnas att öka förståelsen för transportsystemets inverkan på de sociala förhållandena i Sverige.

För att besvara syftet behöver följande frågeställningar besvaras

- Vad är ett lämpligt mått på transportinfrastruktur för att studera dess påverkan på sociala förhållanden?
- Hur kan begreppet social nytta tolkas för att möjliggöra statistisk undersökning?
- Vilka statistiska samband kan påvisas mellan transportinfrastruktur och social nytta?

1.2 Disposition

Den här uppsatsen inleds med ett bakgrundskapitel där läsaren får en inblick i relevanta områden. Här beskrivs vilka motiv som finns för en ny höghastighetsjärnväg och den finansieringsmodell som är tänkt att användas i samband med denna investering. Sedan ges en beskrivning av de traditionella samhällsekonomiska kalkylerna samt de lokaliserings- och tillväxtmodeller som används för att fånga så kallade *wider economic benefits*. Bakgrundskapitlet avslutas med att beskriva social nytta och ge en sammanställning av tidigare forskning som hanterat social nytta i samband med transportinfrastruktur. I metodkapitlet presenteras den ekonometriska modell och antaganden som används för att studera sambandet mellan utbyggnad av transportsystem och social nytta. Vidare presenteras de statistiska verktyg som används och problem som uppkommer i samband med dem. Avslutningsvis beskrivs hur arbetet har utförts rent praktiskt. Uppsatsen fortsätter sedan med ett datakapitel där den data som används i arbetet presenteras. Här diskuteras datans kvalité och struktur samt de justeringar och bortfall som gjorts av olika anledningar. I resultatavsnittet presenteras

resultaten av de regressioner som utförts. Sedan följer ett diskussionskapitel där de resultat som uppsatsen presenterar analyseras djupare. Här diskuteras också uppsatsens brister och förslag på framtida forskning. Avslutningsvis återfinns ett kapitel där de slutsatser som kan dras utifrån arbetet presenteras.

2. Bakgrund

De nuvarande stambanorna byggdes i mitten av 1800-talet. Sedan dess har Sverige upplevt en omfattande urbanisering och tillväxt av storstadsregioner. Goda kommunikationer har lett till att arbetsmarknadsregionerna har vidgats och gett möjlighet att bosätta sig allt längre ifrån arbetsplatsen. Detta har i sin tur lett till en ökad påfrestning på våra transportsystem, däribland järnvägen. För att möta en ökad befolkningstillväxt, som förväntas ske framför allt i storstäderna, framgår det i Utredningen för höghastighetsbanor (2009) att det krävs en ökad satsning på Sveriges infrastruktur och inte minst på järnvägen för att upprätthålla ett välfungerande transportsystem. Det första steget i utbyggnaden av höghastighetsjärnväg mellan Stockholm – Göteborg och Stockholm – Malmö togs i augusti 2012. Den dåvarande regeringen gav då förslag om att bygga Ostlänken mellan Järna och Linköping samt sträckan Mölnlycke till Bollebygd (Trafikverket, 2015 b).

Vidare skriver Utredningen för höghastighetsbanor (2009) att en utbyggnad av höghastighetståg i Sverige skulle skapa ett effektivare transportsystem för både gods- och persontrafik och bidra till landets utveckling. Detta på grund av både en snabbare transport på höghastighetsbanorna men också en frigjord kapacitet på de redan existerande järnvägarna. De förkortade restiderna skulle också leda till förstörade arbetsmarknadsregioner vilket skulle driva på tillväxt och utveckling. Det framgår också att en utbyggnad av höghastighetståg ses som ett samhällsbyggnadsprojekt då de skulle ha en stor effekt på samhället och dess strukturer i stort. Målet med utbyggnaden är att den ska bidra med samhällsekonomiskt effektiva och hållbara transportlösningar som ska förbättra kapacitet, framkomlighet och tillgänglighet.

2.1 Nya finansieringsmodeller för transportinfrastruktur

2009 antogs Sveriges gällande transportpolitiska mål av Riksdagen. I propositionen sammanfattas målbilden som ett övergripande mål och två delmål, ett så kallat funktionsmål och ett så kallat hänsynsmål, enligt följande (Näringsdepartementet, 2009):

- *Transportpolitikens mål ska vara att säkerställa en samhällsekonomiskt effektiv och långsiktigt hållbar transportförsörjning för medborgarna och näringslivet i hela landet.*
 - Funktionsmål: *Transportsystemets utformning, funktion och användning ska medverka till att ge alla en grundläggande tillgänglighet med god kvalitet och användbarhet samt bidra till utvecklingskraft i hela landet. Transportsystemet ska vara jämställt, dvs. Likvärdigt svara mot kvinnors respektive mäns transportbehov.*

- Hänsynsmål: *Transportsystemets utformning, funktion och användning ska anpassas till att ingen ska dödas eller skadas allvarligt samt bidra till att miljö kvalitetsmålen uppnås och till ökad hälsa.*

Målen är framtagna för att tjäna som den viktigaste utgångspunkten för regeringens val och planering av åtgärder och styrmedel inom transportområdet. I Statskontorets rapport från en undersökning av finansiering och planering av transportinfrastruktur (Wikhall, 2014) förklaras det som att det i praktiken innebär att staten, genom Trafikverket, är ansvarig för den långsiktiga planeringen av transportsystemet. Regeringen ger utifrån de transportpolitiska målen och övriga politiska ramar myndigheten direktiv om planeringen. Trafikverket utformar utifrån dessa förslag till långsiktiga nationella respektive regionala planer för transportinfrastruktur. Regeringen analyserar därefter planerna och fastställer planerna.

Vilka delar av transportinfrastrukturen som är statligt respektive kommunalt ansvar är relativt tydligt gällande järnvägar. Staten äger den största delen av järnvägssystemet och denna förvaltas av Trafikverket. Trafikverket är även ansvarig för byggandet, driften och underhållet av statens järnvägar. Stationsbyggnaderna däremot är kommunernas ansvarsområde men vissa av dessa ägs av det statliga företaget Jernhusen. Vissa regionala spår ägs av kommuner och landsting och det finns även exempel på järnvägar som ägs av staten men förvaltas av en annan aktör än trafikverket (Wikhall, 2014).

Från och med 2009 är det enligt lag tillåtet för kommuner att lämna bidrag till statliga investeringar inom väg och järnväg. Detta kallas för medfinansiering. Tidigare har det varit en stramare åtskillnad av ansvar och finansiering mellan statlig och kommunal transportinfrastruktur. I propositionen som föranledde lagstiftningen beskrivs det som att det finns ett stort intresse från lokala och regionala aktörer att påverka förändringar i statlig infrastruktur (Finansdepartementet, 2009). Dessutom är det tillåtet för andra aktörer än kommuner, så som exempelvis företag, att bidra till finansieringen.

Den statliga utredningen som lagstiftningen om medfinansiering är baserad på menar att det finns flera olika typer av tillfällen och motiv för att medfinansiering ska vara intressant. Ur det statliga perspektivet kan det vara av intresse dels för att öka volymen av infrastrukturinvesteringar men också för att genom samverkan med de övriga finansierande aktörerna kunna maximera den samhällsekonomiska nyttan av investeringar i transportinfrastrukturen (Medfinansieringsutredningen, 2011). De medfinansierande aktörernas och i synnerhet kommuners intresse för medfinansiering diskuteras i Wikhall (2014). Det kan handla om att få möjlighet att tidigarelägga investeringen eller anpassa investeringens utformning. Det är viktigt att poängtera att medfinansiering inte är ett verktyg för att genomföra infrastrukturåtgärder som inte ingår i de statliga planerna eller att driva igenom ändringar av prioriteringsordningen mellan objekt. Medfinansiering kan motiveras till följd av att den planerade transportinfrastrukturen genererar värde i form av framtida nyttor. Därigenom kan en medfinansierande aktör, exempelvis en kommun, utnyttja den framtida värdeökningen till att i förskott vara med och finansiera investeringen med ett belopp motsvarande de

förväntade framtida värdena. Detta är en situation där staten har ett incitament att utnyttja möjligheten till medfinansiering samtidigt som övriga aktörer får möjlighet att i och med sin medfinansiering kunna påverka utformningen och planeringen av infrastrukturen. Trafikverkets handledning om medfinansiering framhåller tydligt att medfinansieringens storlek i varje enskilt fall vara proportionell mot den nytta infrastrukturen genererar (Wikhall, 2014). Därigenom behöver nyttorna i någon mån kunna analyseras och kvantifieras.

I Sverigeförhandlingens är det bedömningar av de finansierande aktörernas nytta som utgör det huvudsakliga underlaget för förhandlingen. Eftersom Sverigeförhandlingen omfattar mer än bara enskilda transportinfrastrukturinvesteringar ställs också andra krav på nyttobedömningarna. För att bättre förstå Sverigeförhandlingens arbete och ansvar med nyttobedömningarna är det användbart att först diskutera 2013 års Stockholmsförhandling som kan ses som Sverigeförhandlingens föregångare.

2.2 2013 års Stockholmsförhandling och Sverigeförhandlingen

Den då sittande alliansregeringen tog beslut om att tillsätta Stockholmsförhandlingen i slutet av 2012. I direktivet från Näringsdepartementet (2013) ges utredningspersonen i uppdrag att förhandla fram en lösning att för bygga ut tunnelbanan och öka bostadsbyggandet för att upprätthålla fortsatt ekonomisk tillväxt. Direktivet hänvisar till analyser som fastslår att infrastruktur är den viktigaste faktorn för att möjliggöra ett bostadsbyggande i högre takt. Direktivet framhåller vikten av ett aktörsöverskridande samarbete eftersom att Stockholms läns landsting, och inte staten, förvaltar en betydande andel av kollektivtrafikinfrastrukturen. Därför ges förhandlingen i uppdrag att involvera samtliga berörda kommuner och Stockholms läns landsting men även synpunkter från berörda myndigheter och andra berörda ska inhämtas. I direktivet framgår att finansieringsformen bör innehålla höjda intäkter från trängselskatt och/eller andra former av medfinansiering för att överenskommelsen ska leda till största möjliga samhällsekonomiska nytta. Enligt Hans Rode², som är utredningssekreterare på Sverigeförhandlingen och som tidigare bland annat varit utredare på 2013 års Stockholmsförhandling och med mångårig erfarenhet från transportplanering inom Vägverket och Trafikverket, var en av anledningarna att staten ville ge kommunerna mer ansvar och inflytande i planeringsarbetet för att uppnå större nyttor med infrastrukturprojekt. Det fanns också, enligt Rode, en vilja att få igenom en politisk blocköverskridande överenskommelse så att stora projekt skulle kunna fortgå trots regeringsbyten.

I en delrapport från Stockholmsförhandlingen (2013 års Stockholmsförhandling, 2013) framgår att det under Stockholmsförhandlingen genomfördes ett antal undersökningar och analyser för att bedöma den samhällsekonomiska nyttan. Bland annat undersöktes

² Hans Rode, utredningssekreterare på Sverigeförhandlingen och som tidigare bland annat varit utredare på 2013 års Stockholmsförhandling och med mångårig erfarenhet från transportplanering inom Vägverket och Trafikverket. Intervju 2015-10-30.

social nytta i termer av socialt kapital. I 2013 års Stockholmsförhandling användes socialt kapital, vilket definierades som graden av tillit mellan människor samt tillit mellan människor och offentliga institutioner för att utvärdera hur en ny tunnelbanelinje skulle påverka sociala förhållanden i de berörda stadsdelarna. I den rapporten utgås det från att socialt kapital kan mätas i termer av bland annat tillit, trygghet, delaktighet i val och föreningslivsaktivitet men att data för dessa indikatorer kan vara bristfällig. Det finns även en ansats till att genomföra en kvantitativ bedömning av tunnelbanans påverkan på det sociala kapitalet utifrån hur den sammanbinder olika stadsdelar.

Stockholmsförhandlingen resulterade i en överenskommelse mellan Staten, Stockholms läns landsting, Stockholms stad, Nacka kommun, Solna stad och Järfälla kommun om en utbyggnad av fyra nya tunnelbanesträckningar och byggande av 78 000 nya bostäder (2013 års Stockholmsförhandling, 2014).

I Sverigeförhandlingens direktiv från Näringsdepartementet (2014 a) beskrivs hur regeringen ansåg att det fanns skäl att använda sig utav ett liknande arbetssätt som i 2013 års Stockholmsförhandling vid arbetet med att planera för de nya höghastighetsjärnvägarna då det arbetssättet ansågs vara lyckosamt. I kommittédirektivet står följande:

Syftet med förhandlingspersonens uppdrag är att dels möjliggöra ett snabbt genomförande av nya stambanor på ett sätt som maximerar deras samhällsekonomiska lönsamhet, dels identifiera kostnadseffektiva åtgärder som leder till en förbättrad tillgänglighet och ett ökat bostadsbyggande i framför allt storstäderna med fokus på resurseffektivitet, hållbarhet och förtätning. (Kommittédirektiv 2014:106)

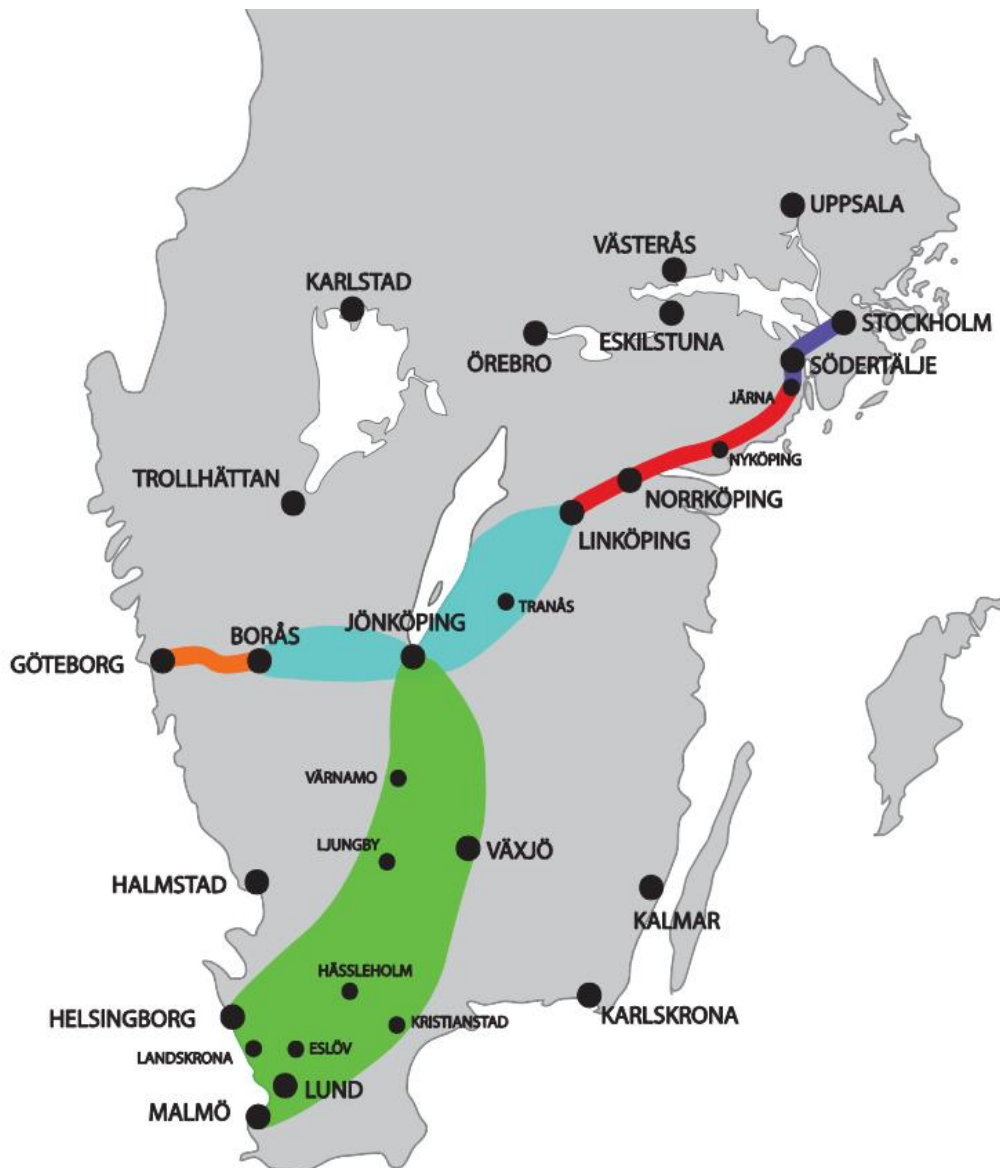
Delredovisningar av förhandlingen ska lämnas in med jämna mellanrum och uppdraget ska slutredovisas senast den 31 december 2017. Det framgår även i direktivet att de förslag och överenskommelser förhandlingen resulterar i ska återspegla de transportpolitiska målen.

Sverigeförhandlingens omfång utifrån direktiven utgår ifrån att med utbyggnaden av höghastighetsbanan som utgångspunkt även genomföra anknytande samhällsbyggnadsåtgärder. Tanken är att förhandlingsarbetet ska fungera som ett effektivt arbetssätt för att ta ett helhetsperspektiv för planering och finansiering av infrastruktur och bostäder. Utifrån Sverigeförhandlingens direktiv (Näringsdepartementet, 2014 a, Näringsdepartementet, 2014 b) kan de åtgärder Sverigeförhandlingen arbetar med sammanfattas enligt följande.

- **Höghastighetsjärnväg.** Sverigeförhandlingen arbetar med att ta fram principer för finansiering, analysera kommersiella förutsättningar och föreslå en utbyggnadsstrategi för de nya höghastighetsjärnvägarna. Medfinansiering är tänkt att utgöra en betydande andel av finansieringslösningen med syfte att skyndsamt och effektivt kunna genomföra utbyggnaden. Det första förhandlingsmötet om höghastighetsjärnvägen är planerat att äga rum 1 februari 2016 (Sverigeförhandlingen, 2016 a). Figur 1 visar de ungefärliga korridorerna för höghastighetsjärnvägen. Den exakta sträckningen mellan Linköping-

Jönköping-Borås samt Jönköping-Malmö är ännu inte fastslagen. I den andra delrapporten, som behandlar höghastighetens kommersiella förutsättningar, skriver Sverigeförhandlingen (2016 b, s.12) att deras arbete med höghastighetsjärnvägen syftar till följande:

Den tolkning vi gör av syftet med utbyggnaden av höghastighetsjärnvägen är att den ska knyta de tre storstadsregionerna närmare varandra, bidra till utveckling i mellanliggande regioner och i övriga Sverige, bidra till mindre koldioxidutsläpp från trafiken och bidra till ett ökat bostadsbyggande. Utbyggnaden ska ske snabbt, maximera den samhällsekonomiska lönsamheten och vara kostnadseffektiv.



Figur 1. Korridorer för Höghastighetsjärnvägen (Sverigeförhandlingen, 2016 b)

- **Storstäder.** Åtgärder för ökat bostadsbyggande, förbättrad kollektivtrafik och förbättrad tillgänglighet i Stockholmsregionen, Göteborgsregionen samt Malmö, Lund och Helsingborg. Förhandlingarna för storstadsåtgärder inleds 8 februari 2016.
- **Cykel.** Cykelåtgärder som kan öka tillgängligheten och främja cykling.
- **Östlig förbindelse.** Sverigeförhandlingen ska analysera och pröva möjliga finansieringslösningar för en väganslutning öster om Stockholms innerstad mellan Norra länken och Södra länken.
- **Järnväg i norr.** Förutsättningarna för att stärka järnvägsförbindelserna i norra Sverige och stärka kopplingen mellan de nya höghastighetsjärnvägarna och det befintliga järnvägsnätet ska analyseras av Sverigeförhandlingen.
- **Ny Danmarksförbindelse.** Sverigeförhandlingen har i uppdrag att ta initiativ med företrädare för Danmark för att undersöka möjligheterna till en ny fast förbindelse mellan Sverige och Danmark. Dessutom ska Sverigeförhandlingen tillsammans med Trafikverket pröva vilka åtgärder som kan vidtas runt Öresundsbron för att öka kapaciteten.

2.2.1 Nyttobedömningar i Sverigeförhandlingen

Sverigeförhandlingens stora omfång innefattar en samordnad planering av infrastruktur och bostadsbyggande och andra samhällsutvecklande åtgärder. Det innebär att alla de nyttor som genereras inte kan bedömas enbart utifrån Trafikverkets standardmetodik för bedömning av transportinfrastrukturåtgärder. För att bredda nyttobegreppet har Sverigeförhandlingen valt att klassificera de nyttor som ska ligga till grund för medfinansiering i sex kategorier. Nedan följer en sammanfattning av hur Sverigeförhandlingen i sitt material som tillhandahålls intressenterna inför deras nyttobedömningar presenterar de olika nyttorna samt vilka metoder som anses lämpliga för att bedöma dem (Sverigeförhandlingen, 2015).

- **Bostadsnyttor.** Nya höghastighetsjärnvägar tillsammans med förbättringar av kollektivtrafiken i storstäderna kommer rimligtvis leda till ökad avkastning för markanvändning för bostäder och lokaler. Bedömningen av dessa ska göras med avseende på en tidshorisont till 2035 och två olika metoder rekommenderas. Den första är den så kallade nuvärdesmetoden som är baserad på att bedöma det ekonomiska värdet av bostadsbyggande till följd av höghastighetsjärnvägarna i de nya stationernas influenszoner. Den andra är ortsprismetoden vilken utgår ifrån att i största möjliga mån jämföra tilltänkt bostadsbyggande med erfarenheter från bostadsmarknaden om liknande bostäder.
- **Restidsvinster.** Effekter på pendling och regionförstoring till följd av förkortade restider är enligt tidigare erfarenheter störst inom stationernas influenszoner men de kan även ha större geografisk utsträckning. Då höghastighetstågen till stor del kommer att framföras på de nya, för ändamålet dedikerade, järnvägarna frigörs spårkapacitet på befintlig järnväg för annan passagerartrafik och godstrafik

vilket minskar restiden även för dessa. Trafikverket har huvudansvaret för att bedöma denna nytta för de nya höghastighetsjärnvägarna medan respektive kommun/region ansvarar för kollektivtrafikinvesteringar. Viktigt att poängtera är att indirekta effekter av restidsvinster så som exempelvis ökat socialt välbefinnande till följd av möjligheten att spendera mer tid med familj och vänner eller bättre matchning på arbetsmarknaden inte inkluderas i denna kategori.

- **Arbetsmarknadsnyttor.** Förbättrad tillgänglighet, exempelvis till följd av minskade restider, får på sikt konsekvenser på lokaliseringsmönster och agglomerationseffekter. Detta kan i sin tur påverka den ekonomiska utvecklingen för riket i stort. Arbetat med att bedöma dessa effekter ansvarar Trafikverket för och bedömningarna görs främst med modellverktygen Dynlok och Samlok som prognosticerar regionalekonomiska effekter till följd av infrastrukturinvesteringars påverkan på tillgängligheten.
- **Miljönyttor.** Till följd av ökat tågresande till följd av de nya höghastighetsjärnvägarna kommer användningen av andra transportslag med högre klimatpåverkan att minska. Trafikverket har i sina samhällsekonomiska kalkyler uppskattat påverkan på utsläppen av växthusgaser men respektive kommun/region uppmanas att göra egna kvantifierande bedömningar.
- **Näringslivsnyttor.** På samma sätt som transportinfrastrukturen påverkar arbetsmarknadens förutsättningar förändras även näringslivets förutsättningar. Bland annat kan företagen nå större marknader, öka kunskapsutbytet och nå mer specialiserade underleverantörer. Metoden som rekommenderas för att bedöma denna nytta utgår ifrån att kartlägga hur den påverkar fastighetsägare för näringsfastigheter. Detta uppmanas kommunerna ansvara för.
- **Sociala nyttor.** Sverigeförhandlingen är tydliga med att det inte finns någon etablerad metod för att bedöma sociala nyttor av ny transportinfrastruktur. Därför uppmanas kommunerna och de andra aktörerna att efter egen förmåga och ingång i ämnet bedöma de sociala nyttorna. Om de sociala nyttorna inte anses kunna bedömas kvantitativt kan de istället beskrivas kvalitativt. Metoden att undersöka socialt kapital, som användes under Stockholmsförhandlingen där horisontell och vertikal tillit bedömdes, ges som förslag på en ingång för att göra en uppskattning av de sociala effekterna. Dessutom nämns så kallad social konsekvensanalys som ett par kommuner utvecklat tillsammans med trafikverket. Dock poängteras att detta snarare är ett ramverk för diskussion än en fullständig bedömningsmetod.

Det råder alltså skillnader mellan de olika kategorierna i flera avseenden. Dels ligger ansvaret för bedömningarna hos olika aktörer för de olika nyttorna och dels finns skillnader i hur vida det finns etablerade och accepterade metoder för att genomföra bedömningarna. Inom området för de sociala nyttorna är det särskilt tydligt att det saknas beprövade metoder för att kunna kvantifiera dessa. Dessutom kan det vara svårt att definiera och särskilja sociala nyttor från de övriga kategorierna. Om inte detta kan

göras på ett tillfredställande sätt finns en uppenbar risk att vissa nyttor dubbelräknas eller missas.

2.3 Samhällsekonomiska analyser

Den klassiska metoden för att bedöma nyttan av investeringar i transportsystemet bygger på samhällsekonomiska analyser. Börjesson och Eliasson (2015) förklarar konceptet som ett ramverk för systematisk jämförelse och metod för att sammanfatta effekter (nyttor) och kostnader av en investering under hela dess livscykel. I regel görs detta genom att uppskatta såväl nyttor som kostnader i samma monetära enhet. En styrka med metodiken är att samtliga värderingar av nyttor och kostnader uttrycks explicit. Därigenom kan, om samma värderingar och viktningar används, en objektiv bedömning av nyttan för olika åtgärdsförslag genomföras. Utifrån värderingen av kostnader och nyttor kan två enkla mått på den samhällsekonomiska nyttan beräknas. Nettonuvärdet (NNV) som är nuvärdesnyttan (NB), alltså nyttan diskonterad till nuvärde, subtraherat med nuvärdeskostnaden (NC), alltså kostnaden diskonterad till nuvärde, se ekvation (1). Detta är alltså det totala nettovärdet av investeringen diskonterat över kalkylperioden.

$$NNV = NB - NC \quad (1)$$

Nettonuvärdeskvoten (NNK) är nettonuvärdet dividerat med investeringskostnaderna (NIC), se ekvation (2). Detta värde kan tolkas som den avkastning varje investerad krona genererar. Alltså en NNK på 0.3 innebär att en investering på en krona ger en avkastning på 0.3 kronor.

$$NNK = \frac{NNV}{NIC} \quad (2)$$

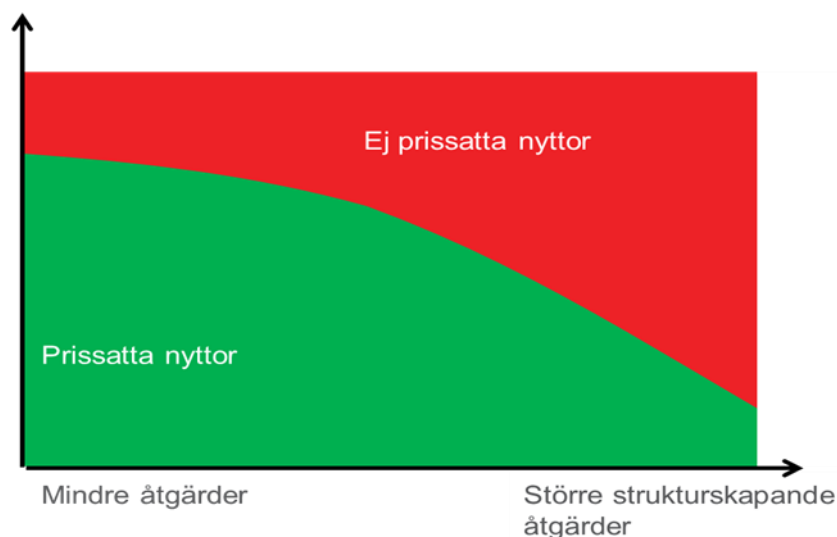
För att beräkna samhällsekonomiska effekter behöver nyttor som inte naturligt har något monetärt värde, alltså som inte regleras av någon ekonomisk marknad och därmed har ett marknadspris, värderas ekonomiskt för att inkluderas i kalkylen (von Sydow, 2011). I Trafikverkets planeringsarbete för investeringar i transportinfrastruktur utgör samhällsekonomiska kalkyler grunden för beslutsfattandet. Ansvar för att utveckla de principer som Trafikverket använder sig av vid samhällsekonomiska kalkyler ligger hos den myndighetsgemensamma samrådsgruppen ASEK³. I en publikation från Trafikverket (2012) beskriver ASEK-gruppen de grundantaganden kalkylmetoden vilar på. Metodiken bygger på antagandet att allt som produceras har ett värde för användarna som kan bedömas utifrån marknadsekonomiska principer. När summan av värdet för alla ”tjänster” och ”produkter” som samhället producerar är större än kostnaderna för att producera och upprätthålla dessa uppnås så kallad samhällsekonomisk effektivitet. Detta innebär att summan av värdet, eller välfärden, för de som gynnas av en specifik åtgärd måste vara större än summan av de negativa effekter de som missgynnas av åtgärden drabbas av. De nyttor och onyttor en åtgärd skapar bedöms alltså utifrån hur de

³ Arbetsgruppen för samhällsekonomiska kalkyl- och analysmetoder inom transportområdet.

subjektivt upplevs och värderas av individer. För att värdera effekter som inte har ett naturligt monetärt marknadsvärde används betalningsvilja som mått. Därigenom kan allt som en individ är villig att betala för värderas utan att det har ett marknadspris. Då inkomsten, och därmed de tillgängliga resurserna, inte är oändlig innebär det per definition att om en individ väljer att allokera en del av sina resurser för en kostnad återstår mindre resurser att lägga på något annat. Betalningsviljan beskriver därför alternativvärdet på en produkt, vad en person är villig att uppoffra för att få produkten.

Börjesson et al. (2013) diskuterar vad den samhällsekonomiska kalkylen faktiskt innefattar. De samhällsekonomiska nyttor som kalkylen avser att fånga sträcker sig bortom enbart rent ekonomiska effekter så som exempelvis ökad ekonomisk aktivitet. Istället är det alla effekter som får konsekvenser för medborgarnas välfärd som ämnas att bli bedömda. Exempel på denna typ av effekter är luftkvalitet, trafiksäkerhet, möjlighet till aktiviteter och liknande. Vissa av dessa tar sig, efter mer eller mindre tid, uttryck i förändringar av BNP, andra inte. Däremot ska alla beståndsdelar av BNP som är relaterat till välfärd inkluderas i kalkylen. I praktiken är detta dock inte fallet eftersom att det som faktiskt beräknas är det så kallade konsumentöverskottet vilket är den nytta som tillfaller resenärerna. Denna nytta motsvarar i princip den kategori Sverigeförhandlingen kallar för restidsvinster.

Sweco (Sandberg, 2015) har genomfört en granskning av några av metoderna för nyttobedömningar åt Sverigeförhandlingen. I denna förs resonemanget om att ju större och samhällsförändrande en infrastrukturåtgärd är desto mindre andel av nyttorna som uppstår kan prissättas med hjälp av den samhällsekonomiska kalkylen. Vid större infrastrukturåtgärder kan det därför behövas kompletterande modeller för att fånga in de nyttor som inte fångas in av den samhällsekonomiska kalkylen. Detta illustreras i figur 2. Då Sverigeförhandlingen definitivt omfattar åtgärder som i allra högsta grad kan anses vara strukturskapande inses behovet av att använda kompletterande bedömningsmetoder utöver den ordinarie samhällsekonomiska kalkylen.



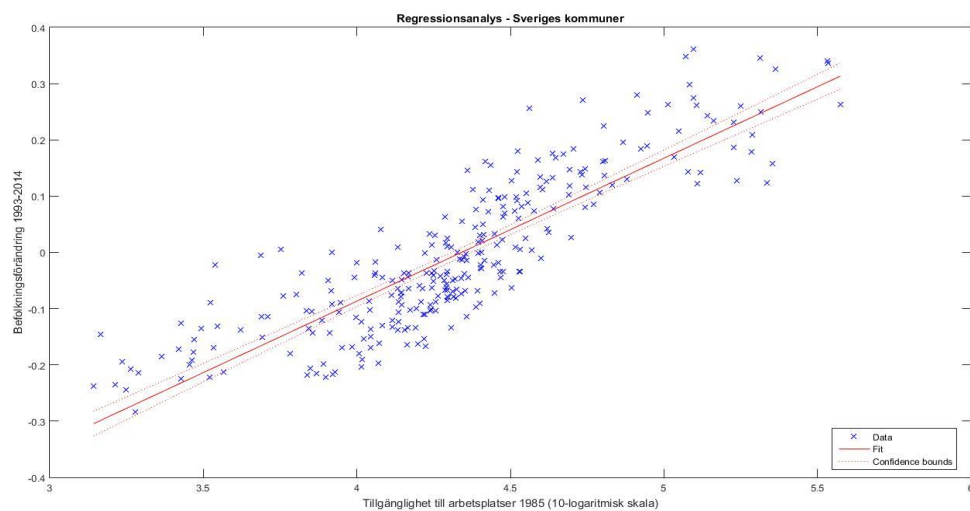
Figur 2. Grön färg illustrerar de nyttor den samhällsekonomiska kalkylen värderar och med röd färg de nyttor som modellerna inte fångar in. Ju större effekt en investering har på samhället desto större del av nyttorna utelämnas.

De BNP effekter som inte ingår i konsumentöverskottet, och därmed inte i den samhällsekonomiska kalkylen, kallas wider economic benefits. I frågan om hur dessa ska och bör kvantifieras råder ingen konsensus. I Sverige finns ett antal modeller som försöker att fånga vissa av dessa. Främst är det så kallade lokaliseringseffekter och agglomerationseffekter som dessa beräknar vilka kan sägas motsvara det Sverigeförhandlingen benämner arbetsmarknadsnyttor. Dessa modeller diskuteras närmare i avsnitt 2.4. Resultatet av modellernas uppskattningar inkluderas inte som en nytta i den traditionella samhällsekonomiska kalkylen utan Trafikverket väljer att istället presentera de som ett komplement i det som kallas samlad effektbedömning (Börjesson et al., 2013).

2.4 Lokaliserings- och tillväxtmodeller

Den traditionella formen för samhällsekonomiska analyser fångar alltså vissa, men inte alla, nyttor som kan tänkas genereras av åtgärderna Sverigeförhandlingen omfattar. Några av de nyttor som inte täcks in av den samhällsekonomiska kalkylen kan istället beräknas med hjälp av lokaliserings- och tillväxtmodeller. I Sverige används fyra olika modeller för att räkna på lokaliseringseffekter och tillväxt från en förändring i transportsystem. Dessa fyra är DYNLOK, RUT, SAMLOK och en modell från ÅF/Infraplan. Modellen från ÅF/Infraplan räknar endast på tillväxt utifrån förändrat antal pendlare. DYNLOK, som är framtagen av Internationella handelshögskolan i Jönköping, skattar lokaliserings- och tillväxteffekter utifrån restidsförändringar i bil- och kollektivtrafik. DYNLOK beräknar lokaliseringseffekter för arbetstillfällen och arbetskraftutbud till följd av åtgärder i transportsystemet som påverkar tillgängligheten. Tillgänglighetsmått som används i modellen är baserade på både bil- och kollektivresetider ifrån Trafikverkets Sampers, som modellerar resande på nationell och regional nivå med olika reseslag. Utgångspunkten är ett antagande att människor flyttar

mellan regioner för att maximera sin vinst i respektive nytta. Ett samband som används i modellen är att antalet arbetstillfällen som skapas i en kommun är en funktion av tillgängligheten på arbetskraft inom kommunen, inom arbetsmarknadsregionen och utanför arbetsmarknadsregionen. Tillgänglighetsmått baseras på arbetstillfällen/arbetskraftsutbud och restider (WSP analys och strategi, 2008). I figur 3 illustreras sambandet mellan tillgänglighet och befolkningsutveckling genom en regressionsanalys.



Figur 3. Regressionsanalys för sambandet mellan tillgänglighet till arbetsplatser 1985 och befolkningsutvecklingen 1993 till 2014 för Sveriges samtliga kommuner.

WSP har tagit fram två modeller, RUT och SAMLOK som båda, precis som DYNLOK, baserar sina tillgänglighetsmått på Sampers. I RUT-modellen skattas lokaliseringseffekter för sysselsättning och en separat beräkning görs för tillväxteffekter. (WSP analys och strategi, 2008). SAMLOK beskrivs på följande sätt I WSP analys och strategi (2015, s.14): *"Samlok är en modell för att uppskatta hur åtgärder i transportsystemet genom förändringar av tillgänglighet påverkar befolkning, sysselsättning, inkomster och markvärden på kommunal nivå."*

Likt DYNLOK utgår SAMLOK ifrån att företag lokaliseras utifrån tillgänglighet till arbetskraft och att både arbetskraftens lokalisering och deras inkomster påverkas av tillgängligheten till arbetsplatser. Både arbetsplatsernas och arbetskraftens lokalisering antas också påverkas av restid till regionalt centrum och närmaste storstad (WSP analys och strategi, 2015).

Gemensamt för alla modellerna utom ÅF/Infraplans modell är att de använder sig av ett tillgänglighetsmått för att beräkna tillväxt- och agglomerationseffekter. Det är alltså den tillgänglighetsförändring (i regel en förbättring) som en infrastrukturåtgärd ger upphov till och inte ett mått på själva infrastrukturförändringen i sig som dessa modeller baserar sina beräkningar på.

För att bedöma de nyttor som kommer uppstå utav en höghastighetsjärnväg använder sig Sveriges kommuner utav dessa modelleringsverktyg som underlag för Sverigeförhandlingen. I figur 4, som är baserad på visas vilka modeller som kan användas för att beräkna de olika nyttorna som används som underlag i förhandlingen. Det framgår att det finns etablerade metoder (DYNLOK, SAMLOK och Sampers) för att uppskatta alla nyttor som uppstår utom för sociala nyttor.

Bostadsnytta	• DYNLOK • SAMLOK
Restidsvinster	• Sampers
Arbetsmarknadsnytta	• DYNLOK • SAMLOK
Miljönytta	• Sampers
Näringslivsnytta	• DYNLOK • Sampers
Social nytta	• -

Figur 4. En sammanställning över vilka modeller som kan användas för att kvantifiera de olika nyttorna i Sverigeförhandlingen.

För att kunna göra en samhällsekonomisk kalkyl av olika nyttor måste först alla de samband som finns mellan infrastruktur och en viss nytta förstås. För sociala nyttor har inte dessa samband kartlagts tillräckligt för att en kvantifierande modell ska kunna utvecklas. Detta kan bero på att det är svårt att tolka de komplexa sambanden mellan en transportutbyggnad och social nytta men också för att det kan vara svårt att definiera vad social nytta består av.

2.5 Vad är social nytta?

Det finns i Sverige ingen given begreppsapparat, metod eller modell för att beskriva eller kvantifiera sociala konsekvenser av transportinfrastrukturutbyggnad. Det är heller inte uppenbart hur tillvägagångssättet att avgöra ifall dessa konsekvenser utgör så kallad social nytta och hur stor denna i så fall är. Den tidigare nämnda ingången som presenteras i underlaget för nyttobedömningarna i samband med Sverigeförhandlingen är långt ifrån en allmänt vedertagen sådan och begreppet är nytt i svensk

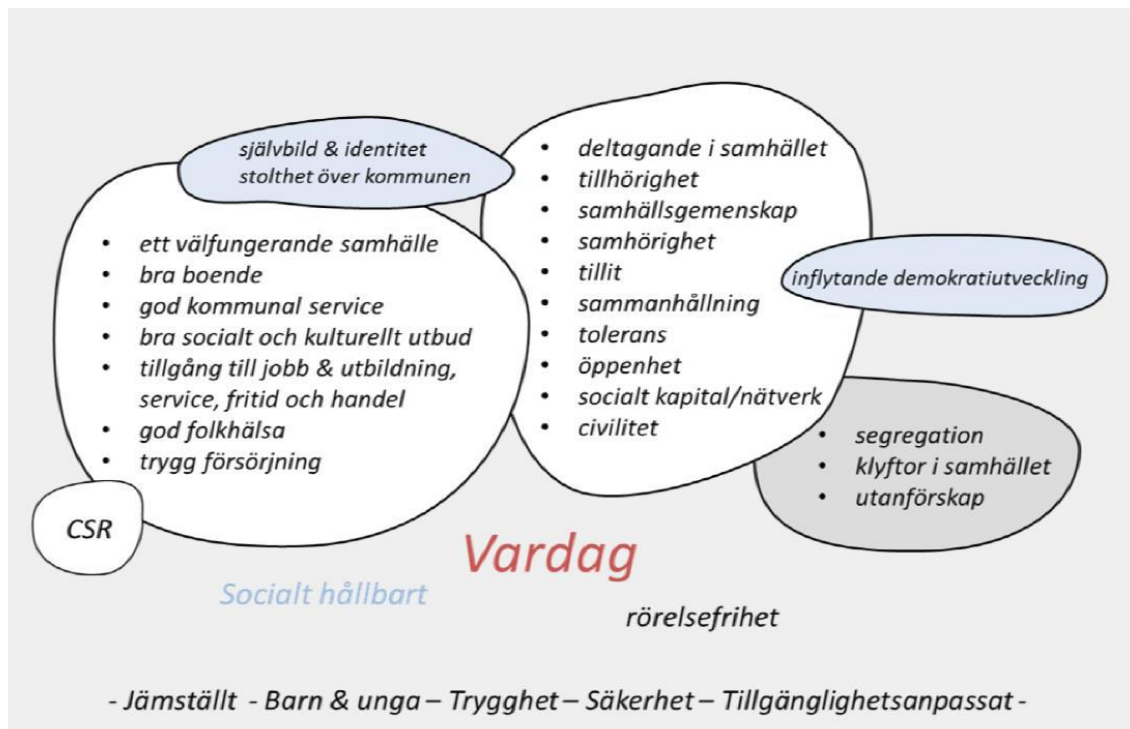
transportplanering. Under hösten 2015 gavs samtliga aktörer intresserade av att delta i Sverigeförhandlingen i uppdrag att genomföra nyttoberäkningar där de tidigare nämnda nyttorna ska redovisas. Dessa är tänkta att fungera som förhandlingsunderlag och därmed är deras tolkningar begreppet sociala nyttor högst relevant för att förstå hur begreppet kommer att användas under förhandlingen. Industriadoktorand Karin Winter har för Sverigeförhandlingens räkning skrivit en rapport om just aktörernas nyttoberäkningar för sociala nyttor.

Winter (2015) diskuterar, innan aktörernas tolkningar av begreppet presenteras, den mångdimensionella karaktären av sociala konsekvenser av transportförändringar och framhåller livskvalitet, gemenskap, inflytande, jämlikhet, jämställdhet och rättvisa som centrala frågeställningar inom ämnet. I rapporten används ett försök till definition av sociala konsekvenser av transporter av Geurs et al. (2009, s.71) som fångar detta samt dynamiken mellan individ, grupp och samhällsnivå.

”...changes in transport sources that (might) positively or negatively influence the preferences, well-being, behavior or perception of individuals, groups, social categories and society in general (in the future)”

Winter (2015) tolkar Sverigeförhandlingens hållning till social nytta som ett försök att förena det växande intresset för sociala hållbarhetsfrågor med en breddad syn på det klassiska samhällsekonomiska nyttobegreppet i transportplaneringssammanhang. Som tidigare nämnt syftar den samhällsekonomiska bedömningen till att värdera positiva och negativa nyttor i kronor som jämförs med investeringens kostnad för att värdera dess lönsamhet. Däremot behöver inte dessa, i monetära termer, värderade nyttor vara direkt kopplade till en aktör som kan kapitalisera på nyttan. Istället kan den generera välfärd men ingen uppenbar direkt ekonomisk vinst. Winter (2015) menar att Sverigeförhandlingens förhållande till i mångt och mycket utgår ifrån att koppla kapitaliserbar nytta till specifika aktörer.

Hur har då de olika kommuner, regioner och övriga aktörer som skickat in nyttoberäkningar inför Sverigeförhandlingen förhållit sig till begreppet sociala nyttor? Av totalt 59 stycken inskickade nyttoberäkningar har Winter (2015) undersökt de 35 stycken som explicit rubricerat sociala nyttor. Generellt för dessa är att de förhåller sig till sociala nyttor som att de utgörs av grundläggande samhälleliga välfärdsfunktioner som boende, kommunal service, kulturellt utbud, folkhälsa, tillgång till utbildning och jobb. Många lyfter även frågor om segregation, tillit, samhällsgemenskap och liknande. I figur 5 finns en sammanställning över intressenternas tolkningar av social nytta från Winter (2015)



Figur 5. En sammanställning av hur intressenter i Sverigeförhandlingen förhåller sig till begreppet social nytta (Winter, 2015).

Winter (2015) har också tittat på vilka nyttor aktörerna tror kan uppkomma till följd av de transportinfrastrukturåtgärder Sverigeförhandlingen omfattar. I rapporten är de kategoriserade och sammanfattade enligt följande

- Sociala effekter av ökad tillgänglighet. Dels till följd av vidgad arbetsmarknad men också att förbättrad kollektivtrafik ger billigare och därmed mer rättvisa möjligheter till resande. Vissa intressenter menar att större nytta skapas om socioekonomiskt svaga grupper får förbättrad tillgång till resmöjligheter i förhållande till välbeställda grupper.
- Social nytta genom fler möten. Möten i transportsystemet anses av många utgöra en nytta i sig och vissa aktörer hävdar att de kan få påverkan på sociala normer vilket utgör den egentliga effekten.
- Socialt hållbar stadsbyggnad. I samband med åtgärderna Sverigeförhandlingen omfattar menar många aktörer att stadsplaneringen kan bidra till att skapa goda sociala miljöer. Funktionsblandning, täthet och befolkade stadsrum framhålls som centrala positiva egenskaper. Även blandade upplåtelseformer för bostäder ses som en nytta.
- Bättre kopplingar ger social nytta. Både fysiska och mentala kopplingar mellan stadsdelar och platser som ny infrastruktur kan generera ses som sociala nyttor. Framförallt framhålls vikten av att områden med olika socioekonomisk status knyts ihop.
- Positiva psykologiska effekter. Vissa aktörer menar att infrastrukturens satsningar i sig kan ge ett socialt uppsving i form av att exempelvis en ny station kan ge

boende något positivt att identifiera sig med samt att den kan locka besökare. Dessutom menar ett antal aktörer att bättre trafikering och bra stationer skapar förtroende för samhället och ökar tilliten.

- Ökad jämställdhet och jämlikhet. Bygger i många avseenden på samma resonemang som för nyttor av ökad tillgänglighet. Billigare och fler icke-bilbaserade resalternativ skapar mer jämlika förutsättningar.
- Övriga effekter. Här lyfts exempelvis folkhälsoeffekter till följd av minskat bilåkande och att tillgång till sportaktiviteter kan förbättras.

Ett fåtal aktörer nämner även sociala onyttor. Exempel på onyttor som tas upp är buller från höghastighetståg, barriäreffekter och förändrad landskapsbild. Även potentiellt otrygga miljöer kring stationsområden nämns. Ett par aktörer nämner även olika former av segregation som kan förädlas av exempelvis alltför exklusiva nybyggnationer av bostäder eller att kommuner eller regioner som får en förhållandevis mindre ökning av tillgänglighet tappar i attraktivitet.

Winter (2015) visar alltså att tolkningen av sociala nyttor är bred och framhåller att enbart en aktör (Tranås kommun) har kvantifierat sociala nyttor i kronor.

2.6 Tidigare forskning

Då detta arbete undersöker hur bedömning av sociala konsekvenser av transportinfrastruktur kan bedömas genom statistiska metoder ligger fokus i forskningsöversikten på arbeten som rör sambandet mellan transportinfrastruktur och dess sociala konsekvenser. Den genomgångna litteraturen behandlar flera olika transportslag och är inte begränsade till enbart höghastighetsjärnväg. Litteraturstudien ligger som grund för att avgöra vilka sociala konsekvenser vars samband med transportinfrastruktur som bör vara relevanta att undersökas statistiskt i den här studien. Dessutom är dess funktion att ge en bredare bild av hur forskningsläget är gällande teoribildningar, metoder och tillämpningsområden och hur detta arbete relaterar till dessa.

Det finns mycket forskning som hanterar de samhällsekonomiska nyttor i samband med utbyggnad av transportsystem. Det finns en stor mängd litteratur på de rent ekonomiska nyttorna som ett infrastrukturprojekt för med sig men när det kommer till de sociala nyttorna är rapporterna färre (Center for Transportation and the Environment North Carolina State University, 2008). Något som ofta nämns är dock att genom integrering av sociala nyttor i planering av transportsystem kan utfallet av en utbyggnad effektiviseras och generera ökad livskvalitet och socialt välbefinnande. Även andra sociala och ekonomiska områden kan uppleva en positiv effekt om hänsyn tas till sociala nyttor i planeringsfasen (Jones och Lucas, 2013).

Social exkludering är en av de teoribildningar som använts för att undersöka transportsystemets sociala konsekvenser. I början av 2000-talet genomfördes tongivande arbete på området genom, den av Storbritanniens regering tillsatta gruppen,

”Social Exclusion Unit” (SEU) där sambandet mellan transporter och olika former av socialt utanförskap undersöktes (Social Exclusion Unit 2003). Resultatet av SEUs arbete var i praktiken ett politiskt dokument som belyste vikten av att genom politisk handling ta itu med socialt utanförskap. I en artikel av Lucas (2012) sammanfattas hur detta forskningsfält fortsatte utvecklas efter SEUs rapport och göra en bedömning av vilka svar och nya frågeställningar det gett upphov till. Först och främst diskuterar Lucas att det inte finns någon konsensus angående vad exakt social exkludering består av. Däremot finns en bred samsyn på att det utgör ett begrepp som sträcker sig bortom enbart fattigdom för att ur ett mer mångfacetterat och dynamiskt ta ett grepp på deprivation. Generellt sett är begreppets huvudsakliga tillämpningsområde inom policy-skapande och är, enligt författaren, ett användbart sådant. Främst för att det lägger fokus på det sociala och ekonomiska utfallet av människors icke fördelaktiga transportsituation snarare än deras upplevelse av den. Därigenom kan det användas för att studera de samband som finns mellan de kausala faktorer som finns hos individer, de som finns i lokalområdets struktur samt de som ligger i den nationella och globala ekonomin. Viktigt att poängtera är att en individ med låg tillgänglighet inte nödvändigtvis behöver innebära att individen är transportmässigt exkluderad. Lucas (2012) refererar till ett antal nationella enkätundersökningar genomförda i USA och Storbritannien som visar på signifikanta ojämlikheter i resmönster och tillgänglighet mellan låginkomsttagare och övriga socioekonomiska grupper. Låginkomsttagare reser både mer sällan och kortare sträckor än övriga. Även studier i Kanada och Australien tyder på liknande resultat. Lucas menar dock att det i denna typ av studier saknas tydliga slutsatser om i vilken omfattning en låg tillgänglighet för låginkomsttagare påverkar deras sociala kapital och välmående. Forskningsfältet har dock under senare år utvecklat de teoretiska koncepten för att tydligare kunna belysa just dessa kopplingar.

Lucas (2012) delar upp dessa framsteg i tre kategorier.

- *Tillgänglighetsperspektiv:* Grieco (2006) föreslår tre huvudkategorier för att genomföra analysarbete om social exkludering i samband med transportsystemsåtgärder. (1) platsbaserade undersökningar i personers närområde. (2) undersökningar baserade på olika sociala gruppers påverkan av åtgärden (3) individbaserade undersökningar om resmönster och liknande.
- *Socialt kapital och kapabilitetsperspektiv:* Här ingår mer dynamiska faktorer så som konceptet hypermobilitet, som diskuteras av, bland andra, Urry (2000). Kring detta tema diskuteras hur tillgång eller exkluderande från mobilitet kan skapa samhälleliga normer och värderingar samt att upprätthålla redan existerande sociala skiktningar mellan sociala grupper.
- *Tid-rumsliga perspektiv:* Inom denna gren ligger fokus på hur förändringar i den rumsliga organisationen av samhället kan exkludera vissa sociala grupper, främst arbetande kvinnor med barn. Det moderna arbetslivets krav på tigha scheman och hög rörlighet ställer stora krav på människors tillgång till transportsystemet.

Utöver de teoretiska utvecklingarna inom fältet för social exkludering menar Lucas (2012) att det sedan 2003 även skett en breddning av teoribildningens tillämpningsområde. De genomförda studierna kategoriseras i tre huvudkategorier.

- *Tillgänglighetsstudier i Storbritannien:* De viktigaste resultaten från dessa är att transportrelaterad exkludering i mycket hög grad är såväl kontext- som personspecifik. Detta har visat sig gälla för alla sociala grupper som är socialt exkluderade.
- *Spatiala- och sociala nätverksstudier i Europa:* I dessa var ansatsen att undersöka hur globalisering påverkar människors livsöden. Den generella slutsatsen var att den spatiala dynamiken av globaliseringsprocessen i hög utsträckning påverkat rörligheten av arbete och arbetskraft. Detta har bidragit till att förstärka redan existerande sociala skiktningar. Andra liknande studier, så som Frei et al. (2009), resonerar utifrån tanken att det sociala nätverket är viktigt för att förstå social exkludering eftersom att det är främst i detta som det sociala kapitalet är förankrat. Studien visade att de människor som hade hög mobilitet, i regel tack vare tillgång till bil, hade ett både större och starkare socialt nätverk och därmed ett mer stabilt socialt kapital.
- *Spatiala och strukturella modelleringsstudier i Australien och Kanada:* Här nämner Lucas (2012) studier av Currie et al. (2009,2010) som visat att det finns tydlig statistisk positiv samvarians mellan transport och social exkludering. Däremot fanns inget statistiskt signifikant samband mellan antal genomförda resor och självupplevelse av exkludering.

Sammanfattningsvis drar Lucas (2012) slutsatsen att transportrelaterad social exkludering har fått stor genomslagskraft inom policyskapande och forskning samt att den såväl teoretiska som metodologiska utvecklingen lett till positiva framsteg. Däremot saknas fortfarande ett tydligt svar på frågan om vilken grad av tillgänglighet till transportsystem som krävs för att minimera risken för transportrelaterad social exkludering. Slutligen tror författaren att den kritik som efter finanskrisen 2008 växt sig starkare mot den neoliberala ekonomiska och politiska världsbilden i grunden kan komma att förändra hur vi ser på frågor om social, spatial och miljömässig rättvisa. Därigenom behövs nya teoretiska och metodologiska perspektiv för att bättre förstå transportsystemets sociala dimension som når bortom den klassiska ”nedspridningsteorin”.

De stora dragen i Lucas forskning om social exkludering är att det trots bra tillgänglighet kan förekomma social exkludering. Detta beror på att även om tillgängligheten är god på en specifik plats så är det ofta olika sociala grupper som kan tillgodogöra sig denna tillgänglighet. Tillgänglighet kan därför leda till ökad social exkludering.

Jones och Lucas (2013) gör ett försök att definiera sociala effekter och vilka variabler som indikerar social inverkan från transportsystem. De delar upp social inverkan i kortsiktiga och långsiktiga effekter. De kortsiktiga effekterna är de mer direkta så som

tillgänglighet, rörelse och aktiviteter, hälso-, finans- och samhällsrelaterade effekter. De mer långsiktiga effekterna är hälsa, välmående och jämlikhet. Vidare anser de att det ofta är svårt att se kausala samband mellan transportsystem och långsiktiga sociala effekter så som hälsotillstånd, utanförskap, socialt kapital och allmänt välbefinnande. Jones och Lucas (2013) poängterar vikten av dels en särskiljning av de sociala konsekvenserna från ekonomiska och miljömässiga sådana dels att alla dessa tre faktorer kan få fördelningseffekter. Författarna menar att i policy-sammanhang beskrivs transportsystemets potentiellt icke-jämlika fördelning av nyttor och onyttor som en social konsekvens medan det snarare bör ses som en produkt av de sociala, ekonomiska och miljömässiga konsekvenserna. Med andra ord menar Jones och Lucas (2013) att sociala, ekonomiska och miljömässiga konsekvenser från ett transportsystem fördelas olika över olika grupper.

I en rapport sammanställd av Center for Transportation and the Environment North Carolina State University (2008) beskrivs social nytta i termer av samhälleligt och socialt välbefinnande. I rapporten föreslås att samhälleligt och socialt välbefinnande beror på tre områden: psykiskt välbefinnande, ekonomiskt välbefinnande och socialt kapital. Inom dessa tre områden påverkar faktorer som allmän nöjdhet, kriminalitet, långtidsarbetslöshet, social tillit, tid spenderad på att pendla, valdeltagande och migration. En annan rapport av Department for Transport (2013) tar ett mer nationalekonomiskt perspektiv då de försöker sätta ett monetärt värde på de sociala nyttorna från kollektivtrafik genom att använda sig av begreppen betalningsvilja och acceptansvilja. De definierar social nytta som värdet bussresenärer upplever genom att ha tillgång till tjänster som de annars inte hade haft tillgång till.

De flesta metoder och verktyg för att analysera sociala effekter av ett transportsystem studerar endast det rådande förhållandet medan ytterst få förutspår effekterna utav en planerad förändring (Forkenbrock, Benshoff och Weisbrodd, 2001).

Brown och Werner (1985), Weening et al. (1990) och Widgery (1982) är exempel på arbeten som använt sig av ett kvantitativt angreppssätt för att förklara vilka indikatorer som bör användas för att förklara livskvalitet och social samhörighet. De har använt sig av statistiska metoder, som principalkomponentsanalys, multivariat kovariansanalys och multipel linjär regression, för att analysera stora datamängder insamlade via enkätundersökningar. Weening et al. (1982) kom bland annat fram till att social samhörighet beror på socialt stöd, interaktion med andra och solidaritet. Widgery (1982) kom fram till att de viktigaste variablerna för att förutse livskvalitet inom ett område var tilltro till regeringen och det politiska systemet, nöjdhet med familj och vänner, antal år i området och optimism om området.

Denna forskning tyder på att några av de faktorer som är viktiga för social samhörighet och välbefinnande är socialt stöd, interaktion med andra och tilltro till regeringen.

I USA har det federala forskningsprogrammet NCHRP⁴ (2001 a) satt ihop en handbok för hur sociala och ekonomiska effekter från transportsystemprojekt ska kunna uppskattas på förhand. Detta för att underlätta beslutsfattande och planeringsarbete. NCHRP (2001 a) menar att det slutgiltiga målet med en förändring av ett transportsystem är att förbättra livskvaliteten för så många som möjligt. Vilket kan tolkas som att maximera den samhällsekonomiska nyttan. De utgår ifrån synen att det inte är någon mening med att försöka skapa något sammanvägt index för sociala och ekonomiska effekter, utan de bör studeras var och en för sig. Handboken definierar elva sociala och ekonomiska effekter som påverkar livskvaliteten och ger förslag på metoder, verktyg och tekniker så att trafikplanerare ska kunna undersöka och planera för dessa effekter. De delar in effekterna i två subgrupper, transportsystemeffekter samt sociala och ekonomiska effekter. Transportsystemeffekterna fokuserar på hur väl ett transportsystem fungerar för dess användare medan de sociala och ekonomiska effekterna riktar in sig på hur andra människor i samhället än användarna påverkas av ett transportsystem. Utöver dessa två subgrupper så undersöker de även fördelningseffekter vilket är något som de menar påverkar alla de andra grupperna. Det som undersöks här är hur olika effekter påverkar olika grupper i ett samhälle. Av de elva effekter handboken behandlar presenterar vi tre av dessa mer ingående; tillgänglighet, samhällelig samhörighet och val av transportsätt.

Tillgänglighet

NCHRP (2001 a) definierar tillgänglighet som den relativa enkelheten att ta sig till en önskad destination. Hur bra tillgänglighet en plats har beror på både tillgänglighet på regionnivå och på lokal nivå. Tillgänglighet påverkas av förändringar i restid, säkerhet, resekostnad och valmöjligheter av olika transportslag. Vid en förändring i transportsystemet är påverkan på tillgänglighet den kumulativa användarrelaterade effekten till följd av förändringen. Enligt NCHRP är god tillgänglighet till arbetsplatser, skola, allmänna samhällsfunktioner, vänner och familj och underhållning en fundamental dimension av livskvalitet.

Ett flertal modeller har utvecklats för att mäta tillgänglighet baserat på olika indikatorer. I forskningsunderlaget till handboken (NCHRP, 2001 b) argumenteras det för att de modeller som tar hänsyn till allmänna och personliga resurser i termer av rumslig jämlikhet är de mest lovande för att utvärdera hur ett projekt kommer påverka tillgänglighet. Ett antal studier har studerat det rumsliga avståndet till allmänna nyttor i samhället för att ge ett mått på tillgänglighet (exempelvis Talen och Anselin, 1998; Truelove, 1993 och Talen, 1998). Gutiérrez et al. (1998) utvecklade en tillgänglighetsindikator för att uppskatta tillgängligheten för varje nod i ett transportnätverk. Detta genom att beräkna ett viktat medelvärde av impedansen (hur väl systemet fungerar) och det rumsliga avståndet mellan en nod och ett aktivitetscentrum

⁴ NCHRP står för National Cooperative Highway Research Program som är ett federalt forskningsprogram med uppgift att utveckla planering, design och liknande av transportinfrastruktur i USA

som destination. Den ekonomiska aktiviteten på destinationerna används som vikt. Författarna menar att detta är ett bra verktyg för att fatta beslut angående investeringar i transportsystem.

En annan metod finns beskriven av Handy och Niemeier (1997) som delar in tillgänglighet i tre olika typer: kumulativ möjlighetsmått, potentialbaserade mått och nyttobaserade mått. De kumulativa möjlighetsmått baseras på hur många möjligheter en person har att utföra önskade funktioner inom en viss restid. Gravitationsbaserade mått viktar möjligheter utifrån impedans som ofta beräknas utifrån resetid och resekostnad. Nyttobaserade mått bedömer sannolikheten för att en person ska göra ett visst val utifrån nyttan av det valet jämfört med nyttan av alla andra valmöjligheter. Nyttan för ett val beräknas utifrån parametrar som beskriver varje vals egenskaper utifrån destinationens attraktivitet och individens socioekonomiska ställning.

Samhällelig samhörighet⁵

I NCHRP (2001 b) beskrivs integration som det sociala nätverkandet inom en samhällsgrupp. Till exempel kan integration innebära en känsla av närhet mellan invånare och en känsla av trygghet i ett område. Ett transportprojekt kan ha både negativa och positiva effekter på integrationen i ett område. För att mäta förändring i integration från ett transportprojekt så bör det rådande förhållandet i ett samhälles sociala nätverkande dokumenteras och sedan uppskatta hur förhållandet skulle förändras om projektet genomfördes. Transportsystem kan skapa både fysiska och psykiska barriärer inom ett samhälle men kan också fungera som en sammanbindande länk. Nya och större transportsystem kan bli till synliga gränser och skärma av en grupp människor i ett område från sina vänner och grannskap. Mindre projekt som gång och cykelvägar kan däremot ha en motsatt effekt och förbättra förbindelserna mellan invånare och samhällstjänster inom ett område.

Mycket av den tidigare forskningen på integration härstammar från samhällspsykologi. Forskare har här studerat integration i många olika samhällen med olika indikatorer. I regel är studierna försök att uppskatta integrationsgraden utifrån olika samhälls- eller grannskapskaraktärer. I en studie av Widgery (1982) användes multipel linjär regression för att studera hur väl 38 olika faktorer påverkar människors nöjdhet med deras livskvalitet. Några av de oberoende variabler som hade störst betydelse var, tillit till regeringen och det politiska systemet, nöjdhet med familj och vänner, optimism om samhället och nöjdhet med samhällsfunktioner. Cuba och Hummon (1993) genomförde en studie där de undersökte integration på olika stadsnivåer, från hemmet till samhället till regionen. På samhällsnivå visade sig deltagande i sociala aktiviteter, vänskap och bostadsrelaterade parametrar ha stor betydelse. På regionnivå var bra integration förknippat med deltagande i aktiviteter i samhällen utanför ens egen hemstad.

⁵ Från engelska community cohesion

Val av transportsätt

I NCHRP (2001 b) menas att val av fortskaffningsmedel beror av kvantiteten och kvaliteten på de transportalternativ som finns tillgängliga för invånare i ett specifikt område. Här fokuserar de på vilka alternativa transportmedel som finns gentemot personbil. Att ha flera transportalternativ påverkar människors livskvalitet på fyra sätt: Hjälper att uppnå jämlikhet, fungerar som reservalternativ för de som kör bil, ökar transportsystemets effektivitet och förbättrar levnadsförhållanden. De konstaterar att nya transportsystem har potentialen att förse individer med möjligheten att välja transportsätt som tidigare inte varit tillgängliga. Men ibland kan det också ta bort valmöjligheter för vissa individer. Majoriteten av de studier som gjorts på hur valet av transportsätt förändras utav en förändring av infrastruktursystem fokuserar på effekterna efter att förändringarna gjorts. Bel (1997) är en studie av hur en förändring i bilresetid påverkar efterfrågan på tågresor i Spanien. Som beroende variabel används förändring i efterfrågan på tågresor (förändring i sålda biljetter 1991 jämfört med 1988). Som oberoende variabler användes tågresetid, bilresetid och förändringar i inhemska flygsträckor. Slutsatsen som dras i studien är att restid med tåg är negativt korrelerat med efterfrågan på tågresor för såväl korta som långa resor, alltså minskar efterfrågan på tåg då restiden ökar. Förhållandet mellan restid med buss och efterfrågan på tågresor var enligt studien positivt speciellt för korta och mellanlånga resor. Förhållandet mellan flygfrekvens och efterfrågan på tåg visade sig vara negativt, alltså ju oftare det går flyg desto mindre blev efterfrågan på tåg.

I en annan studie som det refereras till i NCHRP:s handbok undersökte Kroes et al. (1996) hur reseefterfrågan förändrades utav en utbyggnad av motorvägssystemet i Amsterdam genom att intervjua invånare både före och efter utbyggnaden. Slutsatsen från studien var att förändringar i val av resväg och anländandetid var signifikant större än den hade varit utan utbyggnaden. Däremot kunde inga större förändringar ses i valet av transportsätt, destination eller resefrekvens.

För att summera det som NCHRP skriver i sin handbok så går det inte att hantera sociala effekter som ett enda index utan att de bör studeras var för sig. En god tillgänglighet till arbetsplats, skola, allmänna samhällsfunktioner, underhållning samt vänner och familj leder till förhöjd livskvalitet. Men det finns flera olika typer av tillgänglighetsmått. För den samhälleliga samhörigheten kan ett transportsystem ha både positiv och negativ inverkan. Den kan dels bidra med både fysiska och psykiska barriärer men kan också verka som en sammanbindande länk mellan olika grupper av människor och geografiska områden. Att ha fler transportalternativ kan hjälpa till att uppnå jämlikhet, reservalternativ för de som kör bil, öka transportsystemets effektivitet och förbättra levnadsförhållanden.

Specifikt gällande effekterna av höghastighetsjärnväg har ett flertal studier genomförts internationellt. I Storbritannien, vars första höghastighetsjärnväg togs i full drift 2007, finns en stor mängd forskning som försöker bedöma sociala och socioekonomiska effekter av transportslaget. Preston & Wall (2008) gör just detta för den tidigare nämnda

järnvägen mellan tunneln under engelska kanalen och London som i Storbritannien har namnet High Speed 1 (HS1). De prövar hypotesen om att det enligt teoretiska resonemang baserade på så kallade computable general equilibrium models torde uppstå samhällsekonomiska vinster, främst på grund av ökad konkurrens i mark- och arbetsmarknader samt agglomerationer av ekonomier, som är betydligt större än de som framkommer i konventionella samhällsekonomiska analysen. Deras övergripande resultat är att de samhällsekonomiska nyttorna är oerhört svåra att knyta specifikt till järnvägen och särskilja från utomliggande faktorer. Såväl på nationell som regional och stadsnivå menar författarna att de inte har möjlighet att belägga långsiktiga nyttor av järnvägen eftersom att dessa drunknar i exogena ekonomiska processer. Ett annat grepp för att utvärdera höghastighetsjärnväg presenteras i Martínez Sánchez-Mateos och Givoni (2012) som undersöker HS2 som är tänkt att förlänga HS1 och sträcka sig från London, via Birmingham och vidare norrut mot Manchester, Sheffield och Leeds och senare möjligtvis ända upp till Skottland. Artikelns syfte är att undersöka hur den ökade tillgängligheten från höghastighetsjärnvägen i förhållande till befintliga järnvägsförbindelser fördelar sig inom de påverkade regionerna. Värt att poängtera är att det författarna i denna artikel menar med tillgänglighet är restid mellan orter och skiljer sig därför mot hur begreppet används i forskningen om social exkludering där den snarare betecknar individens möjlighet att i den byggda miljön förflytta sig till de tjänster/aktiviteter som den efterfrågar. I artikeln används en enkel metodik där de beräknar förändring i restider med tåg från olika orter (både sådana som är direkt anslutna till höghastighetsjärnvägen och sådana som är anslutna till den indirekt via befintlig järnväg) i det område där HS2 planeras att bygga. Slutsatserna indikerar att den regionala tillgänglighetsförbättringen är mycket liten och i praktiken främst sker hos de orter som har en väl fungerande anslutning till de större orterna där den nya järnvägen kommer ha station. Dessutom menar författarna att det i Spanien tydligt kunnat ses att orter som är anslutna till det konventionella järnvägsnätet men inte har någon anslutning till höghastighetsjärnvägen har fått sämre tillgänglighet än tidigare eftersom turtäthet och investeringar i underhåll och förbättringar har minskat på den konventionella järnvägen.

Ytterligare ett exempel på en ingång i frågan finns i en artikel som undersöker höghastighetstågens påverkan i Australien. Hensher et al. (2013) tar fram en modell med vilken de studerar vilka fördelar en förbättrad tillgänglighet i form av höghastighetståg ger individer och hushåll. Deras arbete utgår ifrån att ett ökat resande i sig inte bidrar till ökad produktivitet eller sociala nyttor. Istället menar det att det är tillgänglighet och kostnadsbesparingar på resor som skapar ekonomiska och sociala möjligheter och det är dessa som i sin tur genererar ökat resande. Modellen uppskattar värdet på de sociala effekterna både för arbetsrelaterade och icke arbetsrelaterade resor. Detta görs både utifrån den förbättrade restiden och en uppskattning av hur många fler som skulle resa mellan olika städer. Hensher et al. (2013) konstaterar att höghastighetståg ger positiva sociala effekter både för arbetsrelaterade och icke arbetsrelaterade resor men att de största effekterna sker för den senare.

Alltså tyder tidigare forskning av Preston och Wall (2008) att det är ytterst svårt att knyta samhällsekonomiska nyttor till järnvägen och speciellt långsiktiga nyttor då de inte går att urskilja från andra ekonomiska processer. Martínez Sánchez-Mateos & Givoni (2012) konstaterar att tillgänglighetsförbättringar från höghastighetsjärnväg är mycket små och främst sker i orter som har bra anslutning till de större orter där den nya järnvägen har station. Hensher et al. (2013) menar dock att i den mån som höghastighetståg leder till ökad tillgänglighet så genereras också sociala effekter då han antyder att en ökning av tillgänglighet och kostnadsbesparingar leder till positiva sociala effekter.

För att sammanfatta den tidigare forskningen som diskuterats i det här kapitlet så pekar den på att ökad tillgänglighet leder till förbättrade sociala förhållanden. Men det finns vissa forskare som menar att en utbyggnad av höghastighetståg för med sig mycket små tillgänglighetsförbättringar och ibland en minskning, medan andra hävdar att det faktiskt leder till en signifikant tillgänglighetsförbättring. I den mån positiva sociala effekter skapas av höghastighetståg tillfaller den inte alla eftersom vissa sociala grupper eller geografiska områden kan missgynnas. De faktorer som tidigare forskning menar ingår i social nytta är exempelvis hälsa, socialt stöd, känsla av samhörighet, socialt nätverk, tilltro till institutioner, jämlikhet och allmän livskvalitet.

3. Metod

I detta kapitel presenteras de kvantitativa metoder som används i studien. Först ges en introduktion till det ekonometriska angreppssättet metoden är baserad på. Därefter presenteras en enkel kausal hypotetisk modell för transportinfrastrukturens påverkan på sociala förhållanden som sedan formuleras matematiskt utifrån ekonometriska modeller. De metodproblem som är särskilt viktiga diskuteras sedan i ett antal avsnitt. Kapitlet avslutas med en kortfattad beskrivning av den programvara som använts samt en sammanfattning av arbetsgången i studien.

Först och främst bör det poängteras att det enligt författarnas vetenskap vid tiden för genomförandet av detta examensarbete inte finns någon etablerad metodik för att undersöka statistiska samband mellan transportinfrastruktur och dess sociala effekter. Därför har inspiration till möjliga tillvägagångssätt hämtas från besläktade områden från transportforskningen och främst från de områden som rör olika former av ekonometriska metoder för bedömning av samhällsekonomisk påverkan. Ekonometri kan sammanfattas som tillämpandet av statistiska metoder på ekonomisk data för att empiriskt undersöka ekonomiska samband. Baltagi (2011) beskriver det som att de tre grundläggande beståndsdelarna i ekonometri utgörs av ekonomisk teori, ekonomiska data och statistiska metoder. Förutom att empiriskt testa giltigheten i teoretiska ekonomiska samband är ett viktigt fält inom ekonometri att bestämma kvantitativa skattningar av olika former av elasticitetsparametrar som kan vara av intresse vid politiskt beslutsfattande. Som exempel på denna typ av ansats använder Baltagi (2011) frågeställningarna:

- Vad blir effekten av att höja tobaksskatten med x procent?
 - Hur mycket minskar rökandet?
 - Hur stora skatteintäkter genererar det?

I de tidigare nämnda svenska modellerna för lokalisering- och tillväxteffekter till följd av infrastrukturförändringar, Dynlok och Samlok, används liknande ekonometriska elasticiteter mellan tillgänglighet och ekonomiska effekter för att driva modellerna. I denna studie är det sambanden mellan transportinfrastruktur och sociala förhållanden som ämnas att undersökas och en ekonometrisk ansats till detta är väl lämpad av flera orsaker. Ekonometriska metoder är anpassade för liknande typ av data som kan tänkas användas för att studera sociala förhållanden. Det ekonometriska angreppssättet ger också en möjlighet till att kvantitativt beskriva samband mellan tillgänglighet och sociala förhållanden vilket den tidigare forskningen på området inte behandlat i någon större omfattning.

3.1 Tillgänglighet som mått på transportinfrastrukturens funktion

Ett grundläggande problem i den metodologiska ansatsen till detta arbete är att finna ett användbart mått för transportinfrastruktur som är förenligt med arbetes syfte. Börjesson och Eliasson (2015) poängterar att ökad tillgänglighet i regel är den största effekten av en transportinfrastrukturutbyggnad. Tillgänglighet är dessutom ett begrepp som är centralt i det transportpolitiska funktionsmålet. van Wee (2013, s.207) ger en grov och generell tolkning av tillgänglighetsbegreppet som att ”transportsystemet ska tillåta människor att resa och delta i aktiviteter samt företag att transportera gods mellan olika platser”⁶. Begreppet är dock, trots den centrala roll det har i transportplanering, ofta otydligt definierat och tolkas olika i olika kontexter.

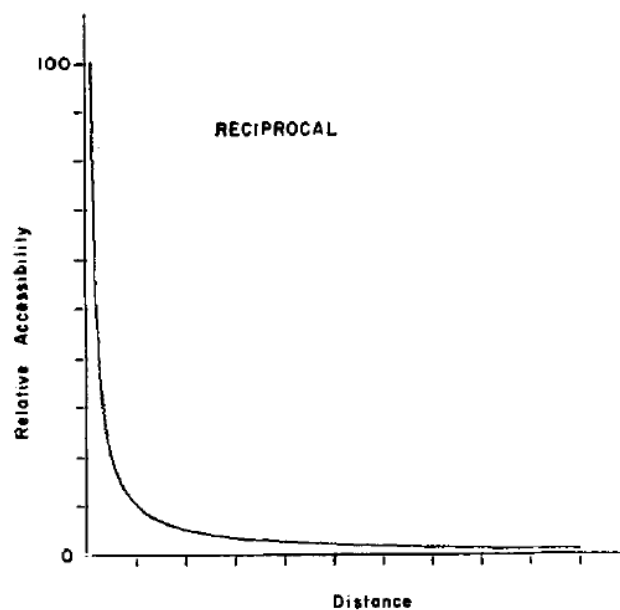
Litman (2011) menar att transportsystemsplanering för närvarande går igenom ett paradigmskifte där det tidigare dominerande analysbegreppet mobilitet ersätts av tillgänglighet. Den centrala skillnaden beskrivs som ett skifte från en analys där färdmedlet, i regel bilen, har stått i centrum till en utvärdering med människor i centrum. Därigenom förändras utgångspunkten för såväl frågeställningar som potentiella lösningar. Författaren framhåller framförallt att begreppet möjliggör större flexibilitet för planeringslösningar eftersom det är transportsystemets påverkan på människor snarare än transportsystemets kvantitativa funktion som utvärderas.

Ett tidigt försök till en matematisk formulering av tillgänglighet är publicerad i en artikel av Hansen (1959). Utgångspunkten var att finna en fungerande matematisk definition för att använda i analys av fysisk planering och bosättningsmönster. Definitionen av tillgänglighet i artikeln är ”potential för möjligheter till interaktion”⁷. Hansens använder sig av ett så kallat potentialbaserat tillgänglighetsmått som för en given geografiskt avgränsad zon förklaras som summan av de möjliga interaktioner som finns i övriga zoner viktat med en linjär avståndsfunktion för respektive zon. Här är det alltså det fysiska avståndet mellan två zoner avgörande i tillgänglighetsmättet.

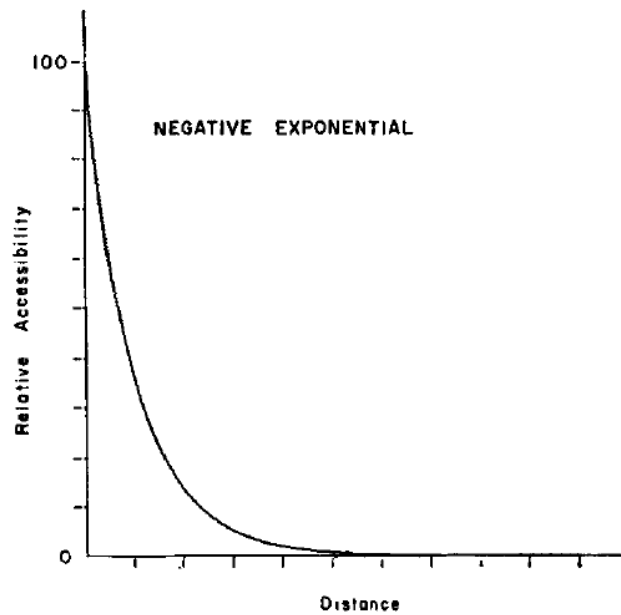
⁶ Översatt från engelska av författarna

⁷ Översatt från engelska av författarna

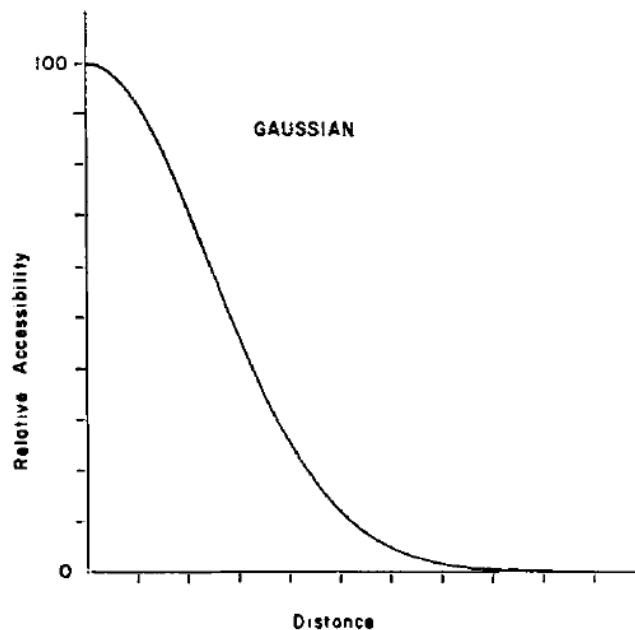
Tillgänglighet som metodologiskt koncept för att kvantifiera interaktionsmöjligheterna i en specifik geografiskt avgränsad zon vidareutvecklades sedan till att använda något mer komplexa potentialbaserade mått. Ingram (1971) beskriver konceptet med icke-linjära viktningsfunktioner. I den just nämnda artikeln föreslås tre olika matematiska funktioner som författaren anser vara användbara som viktningsfunktion både vid kortare och längre avstånd från den undersökta zonen; hyperbolfunktion (se figur 6), negativ exponentialfunktion (se figur 7) samt normalfunktion (se figur 8). Ingram (1971) menar att denna typ av funktioner mycket väl kan beskriva ett flertal mekanismer inom den byggda miljön som har inverkan på interaktionsmöjligheterna och därmed tillgängligheten mellan olika zoner.



Figur 6 - Hyperbol funktion som beskriver relativ tillgänglighet som funktion av avståndet. Ingram (1971)



Figur 7 - Negativ exponentialfunktion som beskriver relativ tillgänglighet som funktion av avståndet. Ingram (1971)



Figur 8 - Normalfunktion som beskriver relativ tillgänglighet som funktion av avståndet. Ingram (1971)

Ingram (1971) använder sig precis som Hansen av fysiskt avstånd för att beskriva tillgängligheten. Därigenom är det ett mått som snarare beskriver hur människor rör sig i den byggda miljön än hur tillgängligheten beror av transportsystemets funktion. För att bättre beskriva tillgänglighet utifrån transportsystemet används ofta så kallad generaliserad kostnad eller synonymt, reseuppostring, som viktfunktion. Koenig (1980) diskuterar ett antal förbättringar som då gjorts i tillgänglighetsbegreppets förankring i

teoretiska ramar för beteendevetenskap och slår fast att det är väl användbart i transportplanering om de ingående parametrarna används på rätt sätt.

För studier där transportinfrastrukturens sociala dimension är central, vilket den är i detta examensarbete, menar van Wee (2013) att just potentialbaserade tillgänglighetsmått är väl lämpade. Dessa kan förstås som att tillgängligheten för en geografiskt avgränsad zon är summan av alla möjliga målpunkter i övriga zoner viktade med en funktion för reseuppofferingen som straffar målpunkter som på grund av transportsystemets utformning är svårnådda eller avlägsna. Därigenom inkluderas både lokaliseringseffekter och transportsystemets utformning tillsammans med antaganden om människors preferenser och vanor kring resor.

En relativt stor del av tidigare forskning på området för sociala effekter av transportinfrastruktur menar att tillgängligheten på flera sätt påverkar förutsättningarna för människors sociala liv (se exempelvis Jones och Lucas, 2013 och NCHRP, 2001). Forskningsgenomgången visar att synen på relationen mellan tillgänglighet och sociala konsekvenser varierar mellan olika forskningsingångar och författare. Något förenklat kan de olika förhållningssätten förstås som att antingen anses förbättrad tillgänglighet i sig vara en social nytta eller så ses förbättringen som något som på sikt kan omsättas av medborgarna till aktiviteter som ger sociala nyttor. Detta arbete utgår ifrån det senare förhållningssättet, alltså att tillgänglighet är en form av potential för positiva effekter på de sociala förhållandena. Utifrån detta perspektiv kan det antas att den förändrade potentialen för sociala förhållanden på grund av förändrad tillgänglighet tar sig mätbara uttryck i statistik som utgörs av indikatorer för sociala förhållanden, givet att potentialen realiserar i tillräckligt stor utsträckning. Dessutom bör detta också innebära att det kan finnas skillnader i sociala förhållanden mellan geografiska entiteter som kan förklaras utifrån skillnader i tillgänglighet. Det är alltså intressant att studera både hur förändringar i tillgänglighet över tid har för samband med sociala förhållanden men även hur nivån av tillgänglighet mellan olika entiteter förhåller sig till sociala förhållanden. Därigenom torde potentialbaserad tillgänglighet vara ett användbart mått för den typ av kvantitativ undersökning i kontexten för sociala nyttor som detta examensarbete behandlar.

I den kvantitativa undersökningen i detta examensarbete definieras tillgängligheten för en geografiskt avgränsad zon som nyttan av att nå alla potentiella målpunkter i andra zoner viktad med en funktion av reseuppofferingen. Denna typ av tillgänglighetsmått som kan beskrivas som ett potentialmått där inflytandet av avlägsna målpunkter eller målpunkter som är svåra att nå minskar på grund av en hög reseuppoffering van Wee (2013). Detta, eller snarlika mått, är väl använda i olika typer av geografisk transportplanering. Bland annat i flera av de svenska lokaliserings- och tillväxtmodeller.

Det mått på tillgänglighet som används i denna studie är detsamma som används i modellen Samlok. Den data över tillgänglighet som används har tillhandahållits av WSP. Den följande beskrivningen av tillgänglighetsmättet utgår ifrån WSP analys och strategi (2015) om inget annat anges.

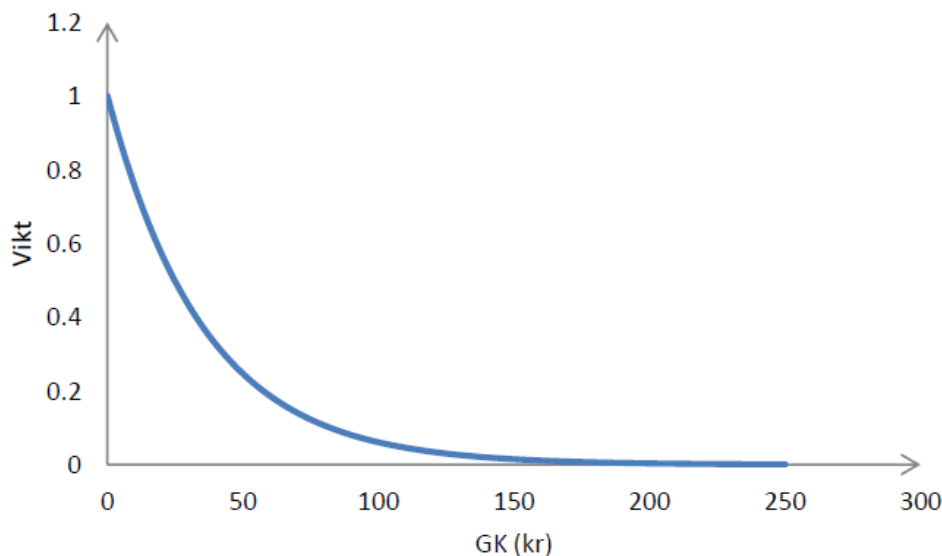
Reseuppofringen består av tre komponenter vilka är restid (inklusive bytes- och anslutningstid), monetär kostnad samt marginalkostnad. Därigenom påverkar faktorer som turtäthet, anslutningstrafik och orters koncentration eller utspridning tillgängligheten. Reseuppofringen kan ses som ett kvantifierat mått på den upplevda reseuppofringen mellan två respunkter. I regel kan en slutpunkt nås med hjälp av flera olika färdmedel. Eftersom de olika färdmedlen innebär olika reseuppofringar behöver dessa vägas samman för att ge ett aggregerat mått på tillgängligheten mellan två respunkter. Sammanvägningen av reseuppofring för alla färdmedel kallas generaliserad kostnad (GK). Detta görs genom att beräkna ett medelvärde av de transportslagsspecifika GK viktat med hur verkliga resflöden mellan målpunkterna är fördelade mellan färdmedlen. I ekvation (3) finns en formel för de huvudsakliga komponenterna vid beräkning av GK mellan startpunkt i och målpunkt j. Den första termen, αt , är uppofringen till följd av resans totala tidsåtgång, där α är resenärens känslighet för tidsåtgång, och t är resans tidsåtgång. Liknande beskriver βc reseuppofringen till följd av monetära kostnader, där β är resenärens känslighet för kostnaden och c kostnaden. γ_m står för marginalkostnaden och utgörs av faktorer så som säkerhet, komfort och liknande (van Wee et al., 2013). Parametrarna α , β och γ är skattade utifrån verkliga reseflöden där t och c är antingen kända eller uppskattade med trafiksimuleringsmodeller. Därigenom kan tillgängligheten med hjälp av simuleringar av framtida resflöden även prognosticeras för planerade infrastrukturåtgärder. För de olika färdmedel är de linjära termerna uppspjäskade i flera termer som består av mer specifika tider och kostnader som uppkommer vid resor med det specifika färdmedlet.

$$GK_{ij} = \alpha t_{ij} + \beta c_{ij} + \gamma_m \quad (3)$$

Tillgänglighetsmättet beskrivs i ekvation (4). A_i är tillgängligheten i kommun i, X_j är antalet arbetsplatser i destinationskommun j. GK_{ij} är den generaliserade kostnaden för resa mellan kommun i och j. β är en skattad parameter som beskriver känsligheten för reskostnaden.

$$A_i = \sum_j X_j * \exp(-\beta GK_{ij}) \quad (4)$$

I figur 10 visas ett exempel på viktningen, alltså känsligheten för GK, som används för att beräkna tillgängligheten.



Figur 9. Exempel på skattad priskänslighet i Samlok-modellen. Alltså hur tillgängligheten skattas beroende på den generaliserade kostnaden för en resa.

3.2 En enkel modell för samband mellan transportinfrastruktur och sociala förhållanden

I figur 6 finns en schematisk bild över den modell som ligger till grund för metoden. Arbetet utgår alltså ifrån en modell som antar att ett politiskt beslut ligger till grund för en förändring av transportinfrastruktur som sedan förändrar tillgängligheten som i sin tur får effekter på sociala förhållanden.



Figur 10. Modellantagande som ligger till grund för metoden. Ett politiskt beslut ligger till grund för en förändring i transportinfrastruktur som antas förändra tillgänglighet som i sin tur får effekter på sociala förhållanden.

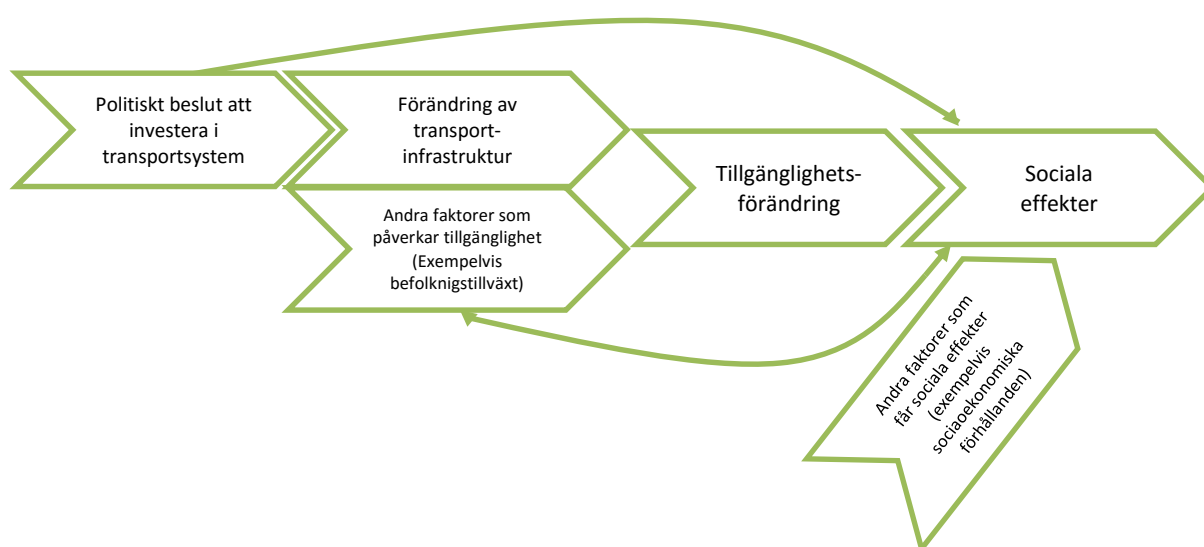
Hur ska då dessa sociala effekter mätas? Utifrån forskningsgenomgången identifierades ett antal variabler som i kvalitativ forskning anses vara relaterade till tillgänglighet. Till dessa har sedan indikatorer knutits utifrån vad som finns tillgängligt i officiell statistik. En utförlig genomgång av datan finns i kapitel 4 men det krävs viss grundläggande vetenskap om datan för att vidare kunna diskutera metoden. En viktig aspekt är att all data som används i arbetet är geografiskt data med upplösning på kommunnivå insamlat på årsbasis vilket får ett antal konsekvenser som diskuteras såväl i detta kapitel som i kapitel 4.

För att formulera en ekonometrisk modell så kan vi utgå ifrån idén om att de sociala förhållandena matematiskt kan förklaras som en multivariat funktion på formen enligt ekvation (5).

$$y = f(x_1, x_2, x_3, \dots, x_n) + \varepsilon \quad (5)$$

De faktiska orsakssambanden som ger upphov till det sociala förhållandet y är att betrakta som ytterst komplexa och beroende av många och svårbedömda faktorer. Som utgångspunkt reduceras ansatsen till en enkel matematisk funktion för att beskriva ett socialt förhållande y i kommun i som en funktion av dess tillgänglighet enligt ekvation (6).

$$y_i = f(x_{\text{tillgänglighet}_i}) + \varepsilon_i \quad (6)$$



Figur 11. Modell som beskriver de samband mellan olika orsaker som påverkar varandra.

Problem uppstår eftersom att det tillgänglighetsmått som i det här arbetet används för att beskriva transportinfrastrukturen påverkas av andra faktorer än endast förändrad transportinfrastruktur. Ytterligare problem uppkommer då de sociala förhållandena i mycket hög grad även påverkas av andra faktorer än tillgänglighet. Därigenom kommer residualerna i feltermen, ε , innehålla en betydande mängd information som inte är rent brus eftersom potentiella andra orsaker i praktiken ingår i modellens felterm ε . Det är därför troligt att en modellansats på denna form innehåller såväl slumpmässiga som systematiska felkällor. I figur 7 presenteras en utvecklad sambandsmodell som beskriver samband mellan utelämnade orsaker som kan ge upphov till systematiska fel. Pilen mellan det politiska beslutet och sociala effekter visar att ett politiskt beslut att investera i infrastruktur kan få sociala effekter då det i sig kan ge framtidstro hos invånarna i den kommun där det sker. Det framgår att tillgänglighetsmättet påverkas av andra faktorer än endast transportinfrastruktur, exempelvis befolkningstillväxt och kommunstorlek, och att dessa faktorer i sig kan ha en direkt påverkan på sociala

förhållanden. Det framgår även att sociala förhållanden påverkas av andra faktorer än tillgänglighet, exempelvis socioekonomiska förhållanden. Detta är exempel på simultanitet samt utelämnade variabler vilket kan ge upphov till endogenitetsproblem. Endogenitet diskuteras mer i avsnitt 3.7.2. Problemet med utelämnade variabler kan mildras genom att inkludera andra förklarande variabler som kan antas ha ett samband med de sociala förhållandena i modellen. Däremot är problemet med att tillgänglighetsmåttet beror på andra faktorer än enbart transportinfrastruktur svarare att handskas med. Detta gör att det utifrån tillgänglighetsmåttet inte går att säga vad som beror på transportinfrastruktur och vad som beror på andra faktorer. Därför kommer det vara problematiskt att säga något om hur ett beslut om att investera i transportinfrastruktur kommer påverka de sociala förhållandena. Men ett första steg kan ändå vara att studera hur tillgänglighetsmåttet korrelerar med olika sociala förhållanden.

I det här arbetet har ett antal variabler som indikerar de socioekonomiska förhållandena inkluderats (dessa är beskrivna i kapitel 4) i ett försök att fånga in andra faktorer än tillgänglighet som kan tänkas påverka sociala förhållanden. Modellen kan nu matematiskt skrivas enligt ekvation (7).

$$y_i = f(x_{tillgänglighet_i}, x_{socioekonomi_i}) + \varepsilon_i \quad (7)$$

Andra utelämnade variabler som kan vara intressant att inkludera är generella trender som sker över tid för alla Sveriges kommuner. Det är rimligt att anta att det finns generella tidsberoende trender av förändring som exempelvis kan bero på såväl nationella som internationella policybeslut eller miljömässiga faktorer. Dessa måste justeras för i modellen för att inte störa skattningar av parametrar. Dessa faktorer inkluderas en genom tidsberoende term i den generella modellen för sociala förhållanden som är gemensam för alla kommuner som beskriver generella trender på en lägre geografisk upplösning än kommunnivå. Modellen beskriven i ekvation (8) kan nu tolkas som att det sociala förhållandet y i kommun i vid tidpunkt t är en linjär funktion av kommunens tillgänglighet och socioekonomiska förhållanden denna tidpunkt samt en generell nivå för alla kommuner denna tidpunkt, som beskrivs av termen S_t .

$$y_{it} = f(x_{tillgänglighet_{it}}, x_{socioekonomi_{it}} + S_t) + \varepsilon_{it} \quad (8)$$

En annan utelämnad variabel som kan inkluderas är faktorer som är specifika för de olika kommunerna men som inte förändras över tid. Dessa kan förstås som kommunernas grundläggande förutsättningar i relation till det analyserade sociala förhållandet. Då Sveriges kommuner inte kan anses vara en homogen population i detta avseende måste denna heterogenitet tas i beaktande i modellen för att den ska kunna anses giltig. Därmed introduceras termen Z_i i modellen som motsvarar kommunspecifika, tidsinvarianta, förutsättningar.

$$y_{it} = f(x_{tillgänglighet_{it}}, x_{socioekonomi_{it}} + S_t + Z_i) + \varepsilon_{it} \quad (9)$$

Termerna S_t och Z_i kallas fixa effekter eftersom de är tids- respektive kommunspecifika och konstanta över kommuner respektive tid.

3.3 Linjär regression

Den ekonometriska metod som detta arbetes kvantitativa del är baserat på bygger i grunden på linjär regression. Enligt Alm och Britton (2008) är regressionsmodeller det grundläggande statistiska verktyget för att studera hur observationer beror på någon annan storhet. Detta kan även sägas vara grunden för de ekonometriska metoderna. Det handlar om att modellera ett samband mellan en eller flera förklarande variabler och en responsvariabel. Alm och Britton skriver att den enklaste och kanske också den vanligaste regressionsmodellen är enkel linjär regression vilken matematiskt ser ut enligt

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (10)$$

Denna formel beskriver det linjära sambandet mellan den förklarande variabeln, x , och responsvariabeln, y . ε kallas för residualen, vilken beskriver skillnaden, felen, mellan våra observationer och den skattade modellen. Parametrarna α och β kallas intercept respektive lutningskoefficient. Lutningskoefficienten talar om hur mycket y förändras av en förändring i x . Termen ε_i är modellens residualterm som innehåller avvikelserna mellan uppmätta och, med modell, skattade värden. Ett viktigt antagande i all linjär regression är att för att modellen ska vara giltig måste residualerna vara normalfördelade med medelvärde noll. För få modellen i ekvation (9) på formen för linjär regression skrivs den om enligt ekvation (11):

$$Y_{it} = \alpha + \beta_1 x_{tillgänglighet_{it}} + \beta_2 x_{Socioekonomi_{1it}} + \beta_{\dots} x_{Socioekonomi_{\dots it}} + S_t + Z_i + \varepsilon_{it} \quad (11)$$

Parametern β_1 är alltså den som främst är av intresse att finna en trovärdig skattning för. Parametrarna skattas med hjälp av minstakvadratskattningarna enligt ekvation (12)

$$Q(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 \quad (12)$$

Genom att minimera ekvation (12) fås skattningar av de α och β som ger minst skillnad mellan den skattade modellen och de faktiska observationerna, alltså de minsta residualerna. Den vanligaste estimatören för β kan skrivas på matrisform enligt ekvation (13).

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (13)$$

Vidare skriver Alm och Britton att en bra modell som är väl anpassad till observationerna har små residualer. För att mäta hur bra en regressionsmodell är används förklaringsgrad som är ett mått av storleken på residualkvadratsumman, summan av alla residualer i kvadrat. Förklaringsgraden kan anta ett värde mellan 0 och

1, där en modell anses bättre ju närmre 1 förklaringsgraden är. En förklaringsgrad på 1 innebär att den skattade linjen förklarar all varians mellan variablerna och att det råder ett exakt linjärt samband emellan dessa. Därigenom kan också värden på relationen mellan variablerna som ligger utanför observationerna troligen predikteras med hög precision.

För att avgöra om det finns ett signifikant linjärt samband mellan en oberoende och en beroende variabel görs ett T-test på regressionskoefficienterna (koefficienterna framför de oberoende variablerna) med följande hypotesformulering (Alm & Britton, 2008):

$$H_0: \beta = 0$$

$$H_1: \beta \neq 0$$

Om tester visar att regressionskoefficienten för en oberoende variabel är skild från noll så finns det ett signifikant samband mellan den oberoende och den beroende variabeln. För att detta ska ske måste nollhypotesen förkastas, vilket den gör om testets p-värde är mindre än den önskade signifikansnivån. Inom forskning väljs ofta en signifikansnivå på 0.1, 0.05 eller 0.01 (Alm & Britton, 2008).

3.3.1 Pearsons korrelationskoefficient

Pearsons korrelationskoefficient är ett mått på det linjära samvarierandet mellan två storheter. I rapporten benämns denna även som enbart korrelationskoefficient. Koefficienten kan anta värden mellan -1 och 1 där -1 tyder på fullständig negativ linjär korrelation och 1 på fullständigt positiv korrelation. Värden nära noll tyder på svag linjär korrelation. I ekvation (14) visas uttrycket för korrelationskoefficienten mellan två slumpvariabler A och B som båda har N observationer (Alm & Britton 2008)

$$p(A, B) = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right) \quad (14)$$

Där μ_A och σ_A är medelvärde respektive standardavvikelsen för A och μ_B, σ_B är desamma för B.

3.3.2 Logaritmisk transformering

Benoit (2011) skriver att transformering till logaritmiska variabler är ett vanligt sätt att hantera olinjära samband mellan de oberoende och beroende variablerna och för att transformera en skev fördelning till en mer approximativ normalfördelning. Han skriver att en linjär regressionsmodell transformeras med logaritmer i fyra olika kombinationer. Dels det enkla linjära regressionsfallet som beskrevs i ekvation (10) men också linjär-log modell, log-linjär modell och log-log modell. Följande är ekvationerna för de olika kombinationerna och hur de bör tolkas:

Linjär-log modell:

$$y_i = \alpha + \beta \log_{10}(X_i) + \varepsilon_i \quad (15)$$

I detta fall tolkas regressionskoefficienten β inte som lutningen av en regressionslinje som i det linjära fallet. Istället innebär denna transformering att storleken på påverkan på den beroende variabeln minskar med högre värden på den oberoende variabeln. En förändring av X_i med $P\%$ motsvarar en ökning i den beroende variabeln enligt ekvation (16).

$$\Delta y_i = \beta \log_{10}\left(1 + \frac{P}{100}\right) \quad (16)$$

Log-linjär modell:

$$\log_{10}(Y_i) = \alpha + \beta x_i + \varepsilon_i \quad (17)$$

Tolkningen av koefficienten β i en log-linjär modell är att en ökning på en enhet i X kommer generera en ökning i $\log_{10}(Y)$ på β enheter. En förändring i X på en enhet ger en förändring i Y på 10^β .

Log-log modell:

$$\log_{10}(Y_i) = \alpha + \beta \log_{10}(x_i) + \varepsilon_i \quad (18)$$

En förändring i en procent i X ger en förändring på β procent i Y vilket alltså motsvarar elasticiteten.

3.4 Paneldataanalys

Det finns ett flertal begränsningar med att använda vanlig linjär regression för den typ av undersökning som detta arbete innefattar. Ett avgörande problem är svårigheten att kombinera analyser av tidsserie- och tvärsnittsdata. Det data som detta arbete har tillgång till är utav paneldatastruktur vilket betyder just att observationer har gjorts för samma entiteter över flera olika år. Alltså är datan av både tidserie- och tvärsnittskaraktär och observationerna kan därför utvärderas över både tids- och tvärsnittsdimensionen. Enligt Baltagi (2005) är data av denna struktur mycket vanlig i ekonometriska analyser. Paneldata innehåller mer information och möjlighet till mer precisa skattningar tack vare fler observationspunkter än enbart tvärsnitts- eller tidsseriedata. Paneldata fångar även dynamiska beroenden som i praktiken inte är möjliga att observera eller kvantifiera i endimensionell utvärdering. En annan viktig fördel med att använda statistiska paneldatametoder är att de kan hantera och justera för heterogeniteter i datan såväl mellan olika entiteter som mellan olika tidsperioder. Som tidigare nämnt är avsaknaden av dessa justeringar en möjlig orsak till endogenitet som stör skattningarna av parametrarna i modellen. Ett exempel på hur paneldata kan se ut finns i tabell 1.

Tabell 1. Exempel på hur ett dataset med paneldatastruktur ser ut.

Individ, i	År, t	Y	X
1	2003	$Y_{1,2003}$	$X_{1,2003}$
1	2004	$Y_{1,2004}$	$X_{1,2004}$
1	t	$Y_{1,t}$	$X_{1,t}$
2	2003	$Y_{2,2003}$	$X_{2,2003}$
2	2004	$Y_{2,2004}$	$X_{2,2004}$
2	t	$Y_{2,t}$	$X_{2,t}$
i	2003	$Y_{i,2003}$	$X_{i,2003}$
i	2004	$Y_{i,2004}$	$X_{i,2004}$
i	t	$Y_{i,t}$	$X_{i,t}$

Paneldata kan vara av både balanserad och obalanserad karaktär. Balanserad paneldata har observationer för alla individer för alla årtal medan obalanserad paneldata saknar observationer för någon eller några individer för något eller några årtal. Det finns metoder för att ta hänsyn till och hantera obalanserad data i paneldataregression.

Enligt Baltagi (2005) så tillåter paneldata större möjligheter att studera dynamiska förändringar. Han skriver att tvärsnittsdata som till synes tycks vara stabil kan innehålla många olika förändringsprocesser när även tidsdimensionen tas i beaktande. Med paneldata går det att studera hur och vilka individer som påverkas och förändras under en utvecklingsprocess. Vidare skriver Baltagi (2005) skriver att en av de största fördelarna med att köra regression på paneldata är att det går att kontrollera för individuell heterogenitet. Han menar att till exempel individer, företag, regioner och länder är heterogena och att det inte alltid är lätt att observera alla variabler som leder till denna heterogenitet. Som tidigare diskuterats i rapporten är det så att när en ordinär regression körs på paneldata riskerar den att ge en felaktig skattning av koefficienterna på grund av att det kan förekomma utelämnade variabler som förklarar heterogeniteten hos den beroende variabeln. I en paneldataregression går det däremot att använda verktyg för att kontrollera för dessa utelämnade variabler utan att observera dem.

En mycket enkel paneldatamodell där icke observerad heterogenitet inte justeras för ges på matrisform av ekvation (19).

$$y_{it} = \alpha + X_{it}\beta + \varepsilon_{it} \quad i = 1, \dots, N; t = 1, \dots, T \quad (19)$$

Där i står för alla olika observerade individer och t för alla årtal. Alltså står indexeringen i för tvärsnittsdimensionen och indexeringen t för tidsseriedimensionen. ε_{it} står för den resterande varians som likt den vanliga feltermen i en regression idealt är vitt brus. X_{it} är en matris med en kolonn av längd $N \times T$ för varje oberoende variabel. I regel är paneldatamodeller multivariata och innehåller flera oberoende variabler X , vilket är fallet även i denna studie, med respektive skattad β parameter som utgör elementen i vektorn β .

3.4.1 Between effects regression/ tvärsnittssregression

En mycket enkel metod för att estimeras utifrån variationen mellan individerna (kommunerna) är att använda en transformation av minstakvadratskattningen i ekvation(12). Genom att skala bort tidsdimensionen i datan genom att ansätta $\tilde{y} = \bar{y}$ och $\tilde{X} = \bar{X}$ där $\bar{y}$ och $\bar{X}$ är tidsmedelvärden kan parametrarna skattas utifrån vanlig minstakvadratskattning enligt ekvation (20). Detta kallas *between effects regression* (Álvarez et al. 2013).

$$\hat{\beta}_{be} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \tilde{y} \quad (20)$$

Eftersom tidsdimensionen försvinner i och med tidsmedelvärdestransformationen är between effects regression mycket lik en vanlig tvärsnittssregression. Det går alltså inte att med hjälp av denna estimator säga något om det tidsberoende förhållandet mellan X och y -variablerna utan det är enbart skillnader i nivåer för en datapunkt för respektive individ (kommun) som används för skattningen.

3.5 Kommunspezifika effekter

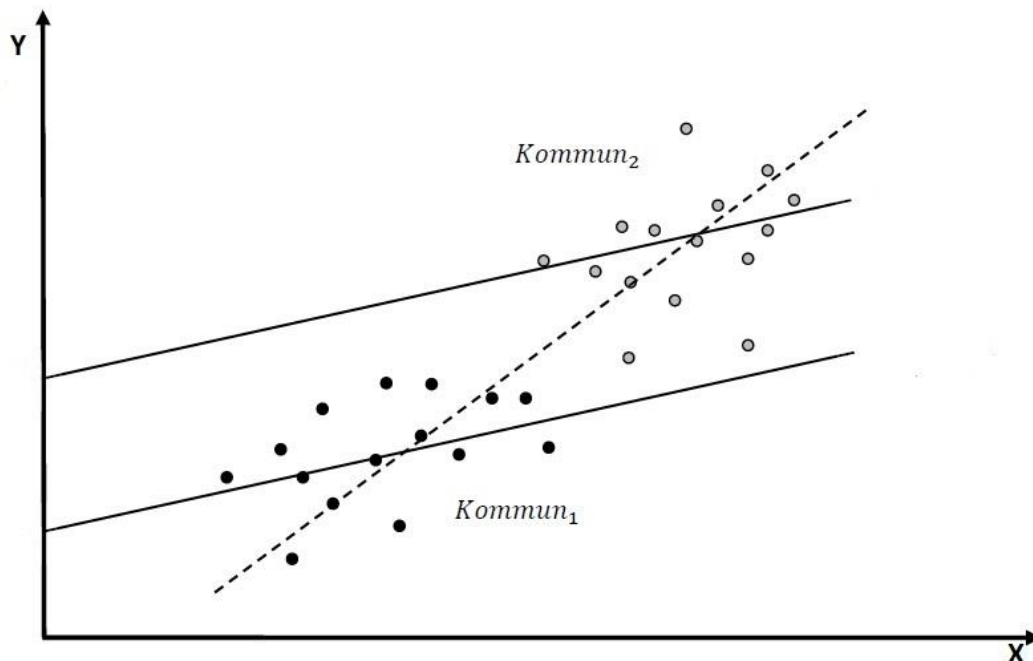
För att förbättra modellen och skattningen av parametrar införs så kallade kommunspezifika fixa effekter. Som utgångspunkt kan en modell för paneldata regression som ser ut enligt följande användas:

$$y_{it} = \alpha + X_{it}\beta + Z_i + \varepsilon_{it} \quad i = 1, \dots, N; t = 1, \dots, T \quad (21)$$

Z_i beskriver oobserverade tidsinvarianta skillnaderna mellan individerna. I övrigt är beteckningarna desamma som de som beskrivs i samband med ekvation (19).

Emanuelsson och Petersson (2011) menar att inkluderingen av de kommunspezifika fixa effekterna Z_i innebär att skattningarna av β parametrarna förbättras eftersom att modellen därigenom justerar för grupperingar i datan till följd av heterogenitet mellan kommunerna. I figur 8 visas en grafisk illustration av hur införandet av kommunspezifika fixa effekter förbättrar skattningen av β vilken ju avgör linjernas lutning. Om inga fixa effekter inkluderas kommer skattningen av parametrarna att behandla datan som observationer från en homogen population, med gemensamt intercept och lutning, och därmed skatta den streckade kurvan. Om istället fixa effekter inkluderas, vilket kan tolkas som att varje gruppering av observationer, alltså varje kommun, har ett eget intercept så blir skattningen av β parametrarna betydligt bättre. Figur 8 kan användas som exempel för att förstå detta bättre. Figuren visar ett paneldataset bestående av två olika kommuner observerade för femton olika tidpunkter. Båda grupperingarna av observationer har en liknande positiv linjär samvarians mellan beroende och oberoende variabel. För hela stickprovet däremot har det linjära förhållandet en större positiv riktning vilket alltså innebär att skattning av en

regressionslinje för hela stickprovet⁸⁸ ger en falsk, överskattad, lutningskoefficient. Om istället varje gruppering av individer tilldelas en fix konstant, som kan tolkas som ett gruppsspecifikt intercept, justeras för heterogenitet bland individerna och skattningen av β blir bättre.



Figur 12. Illustration av regressionsskattningar med kommunspecifika fixa effekter (heldragna linjer) och utan kommunspecifika effekter (streckad linje). Egen bearbetning baserad på Emanuelsson och Petersson (2011).

3.5.1 Regression med fixa effekter

För att skatta β parametrarna i ekvation (19) utifrån den variation som finns inom individerna över tid används *fixed effects regression*. Alltså utnyttjas förändringar i variablerna över tid för att uppskatta de oberoende variabelernas inverkan på den beroende variabeln. Estimeringen av parametrarna görs därför med den information som finns i tidsserievariansen. Denna metod används när individspecifika effekter introduceras i modellen enligt ekvation (21). Denna estimeringsmetod är framförallt användbar när $COV(X_{it}, Z_i) \neq 0$, alltså när det är nödvändigt att justera för icke observerad individheterogenitet (Álvarez et al. 2013). Metoden är baserad på att tilldela en dummyvariabel till respektive individ vars minstakvadratskattade parameter representerar de individspecifika effekterna. Dock innebär detta en förlust av N frihetsgrader samt att är beräkningsmässigt tungt. Detta kan undvikas genom att använda sig av Frisch-Waugh-Lovell teoremet, noggrant förklarat i Baltagi (2005,

⁸⁸ Vilket i detta sammanhang kallas för poolad regressionsanalys

ss.11-14) vilket innebär att parametrarna kan skattas genom att använda estimatoren i ekvation (22), där $\tilde{y} = y - \bar{y}$ och $\tilde{X} = X - \bar{X}$.

$$\hat{\beta}_{fe} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \tilde{y} \quad (22)$$

Parametrarna skattas utifrån avvikelsen från medelvärde över tid. Därmed behöver inte några dummyvariabler beräknas direkt utan de fixa effekterna kan istället lösas ut. Detta kan förklaras genom att skriva om ekvation (19) till formen nedan enligt ekvation (23)

$$y_{it} = X_{it}\beta + u_i + \varepsilon_{it} \quad (23)$$

där $u_i = \alpha + Z_i$ och alltså innehåller både den individuella effekten för kommun i och det gemensamma interceptet. Estimatoren i ekvation (20) eliminerar u_i och därigenom kan vi använda ekvation (24) för att lösa ut de individuella effekterna.

$$\hat{u} = \bar{y} - \bar{X}\hat{\beta}_{fe} \quad (24)$$

En annan viktig egenskap med Frisch-Waugh-Love teoremet är att det innebär att $\sum_{i=0}^N Z_i = 0$. Därigenom kan modellens gemensamma intercept α , respektive de individspecifika effekterna Z_i skattas enligt ekvationerna (25) och (26).

$$\alpha = \bar{u}_i \quad (25)$$

$$Z_i = u_i - \bar{u}_i \quad (26)$$

Alltså kan ekvation (19) fås genom ekvation (27) nedan.

$$y_{it} = \bar{u}_i + X_{it}\beta + (u_i - \bar{u}_i) + \varepsilon_{it} = \alpha + X_{it}\beta + Z_i + \varepsilon_{it} \quad (27)$$

3.5.2 Regression med slumpmässiga effekter

I vissa fall, i regel när observationerna kommer från ett stickprov som är litet relativt en stor population, kan det finnas tidsinvarianta skillnader mellan individerna som är helt slumpmässiga. Denna estimatormetod kallas *random effects regression*. Därigenom behövs inte en parameter skattas för respektive individ utan skillnaderna kan istället inkluderas i feltermen. Detta görs genom att ansätta modellen enligt ekvation (21).

$$y_{it} = \alpha + X_{it}\beta + u_{it} \quad (28)$$

$$u_{it} = Z_i + \varepsilon_{it} \quad (29)$$

Enligt Álvarez (2013) är en stor fördel med denna estimator är att den inte innebär en lika stor förlust av frihetsgrader som tidigare nämnda estimatorer. Dock ställer denna modellformulering relativt starka krav på oberoende mellan parametrarna då följande antaganden måste kunna göras, $\text{COV}(X_{it}, Z_i) = 0$, $\text{COV}(\varepsilon_{it}, Z_i) = 0$, $\text{COV}(X_{it}, \varepsilon_{it}) = 0$.

Härledningen för estimatoren för random effects regression kräver förklaring i många steg och därför inte inkluderad här men kan ses i bland andra, Álvarez (2013). En intuitiv förklaring av hur den fungerar är att den består av en sammanvägning av estimationerna för between effects och fixed effects.

För att avgöra om de antaganden som krävs för att ansätta en random effects regression istället för en fixed effects regression är uppfyllda kan ett så kallat Mundlaktest användas (Álvarez, 2013). Testet bygger på att termen $\hat{X}_i\gamma$ inkorporeras i modellen där $\hat{X}_i$ är tidsmedelvärde för respektive individ. Testhypotesen $H_0: \gamma = 0$ innebär att det inte finns signifikanta icke slumpmässiga individuella effekter och att random effects därmed kan ansättas.

$$y_{it} = \alpha + X_{it}\beta + \hat{X}_i\gamma + Z_i + \varepsilon_{it} \quad i = 1, \dots, N; t = 1, \dots, T \quad (30)$$

3.6 Tid- och kommunspecifika fixa effekter

För att justera för både tid- och kommunspecifik heterogenitet ansätts inkluderas både tid- och individspecifika effekter i modellen. Därmed behöver dummyvariabler implementeras som är specifika för dels alla kommuner dels alla år. Regressionsmodellen då dessa justeringar är gjorda ser ut enligt

$$y_{it} = \alpha + X_{it}\beta + Z_i + S_t + \varepsilon_{it}, \quad i = 1, \dots, N; t = 1, \dots, T \quad (31)$$

Precis som med enbart kommunspecifika effekter kan Frisch-Waugh-Love teoremet användas vilket undviker förlust av frihetsgrader och onödigt tunga beräkningar. Ekvationen för regressionerna med både tids- och kommunspecifika effekter inkluderade hämtas från Baltagi (2005) och ser ut enligt

$$y_{it} = X_{it}\beta + u_i + \lambda_t + \varepsilon_{it} \quad (32)$$

Den estimator som används finns noggrant beskriven i Baltagi (2005, s.33). Estimatoren bygger på att skatta parametrarna i modellen utifrån transformationer av beroende och oberoende variabler där de såväl tid- som individspecifika effekterna utesluts. Estimatoren ser ut enligt följande

$$\hat{\beta}_{fe2way} = (X^T Q X)^{-1} X^T Q y \quad (33)$$

Där Q är en operatormatris som ger

$$Qv = \{v_{it} - \bar{v}_i - \bar{v}_t + \bar{v}_{it}\}_{it} \quad (34)$$

Där $v_{it} = u_i + \lambda_t + \varepsilon_{it}$

Där u_i och λ_t är motsvarar det gemensamma interceptet och de kommunspecifika respektive tidspecifika effekterna, och ser ut som

$$u_i = Z_i + \alpha \quad (35)$$

$$\lambda_t = S_t + \alpha \quad (36)$$

Där Z_i och S_t är de kommunspecifika och tidspecifika effekterna och α är modellens intercept. Som tidigare nämnt ansätts i estimatorn villkoret $\sum_{i=0}^N Z_i = \sum_{t=0}^T S_t = 0$ vilket ger

$$\alpha = \bar{u}_i = \bar{\lambda}_t \quad (37)$$

Där

$$\hat{u}_i = \bar{y}_i - \bar{X}_i \hat{\beta}_{fe2way} \quad (38)$$

$$\hat{\lambda}_t = \bar{y}_t - \bar{X}_t \hat{\beta}_{fe2way} \quad (39)$$

För att modellen ska se ut enligt ekvation (32) måste alltså α subtraheras från ekvation (31) enligt följande

$$y_{it} = X_{it}\beta + u_i + \lambda_t + \varepsilon_{it} - \alpha = X_{it}\beta + S_t + \alpha + Z_i + \alpha - \alpha + \varepsilon_{it} = \alpha + X_{it}\beta + Z_i + S_t + \varepsilon_{it} \quad (40)$$

3.7 Metodproblem

När ett ekonometriskt angreppssätt används för att studera och beskriva policyrelaterade effekter uppstår ett antal centrala metodologiska problem. De viktigaste rör kausalitet och endogenitet mellan modellernas beståndsdelar. I detta kapitel diskuteras innebörden av problemen för detta arbete.

3.7.1 Kausalitet

När vi diskuterar begrepp som ”samband” och ”kausala förhållanden” utgår vi ifrån ett tankesystem baserat på att varje observerat empiriskt fenomen föregås av en eller flera orsaker. Hur kausalitet faktiskt är beskaffat har gäckat människor i alla tider och föreställningen om varför en händelse sker har förändrats över tid. Pearl (2010, s. 401-428) avhandlar detta och beskriver hur statistiska beskrivningar av förhållandet mellan variabler, alltså dess korrelation, gett upphov till kontroverser om hur vida det kan användas för att dra slutsatser om kausalitet. Han skriver om hur kausalitet orsakar problem inom olika vetenskapliga discipliner och att ord som ”orsakar” och ”leder till” sällan används i vetenskapliga texter av försiktighetsskäl. Ett exempel hämtat från det ekonometriska fältet som Pearl tar upp för att visa svårigheter med säga något om kausalitet är hur olika läkemedels effekt på återhämtningen från en viss sjukdom kan studeras. Han menar att återhämtningen inte endast beror på vilket läkemedel som

används i behandlingen utan andra faktorer som exempelvis socioekonomiska faktorer, ålder, diet och livsstil också påverkar. Det går då alltså inte att säga om förändringar i återhämtning beror på behandlingen med läkemedlet eller någon av de andra faktorerna. Klart är att statistiska studier baserade på observationer och inte kontrollerade experiment av flera skäl är problematiska att dra kausala slutsatser utifrån. Om vi bortser från det mer filosofiska problemet hur vida det finns verkliga orsakssamband eller om de snarare är en mänsklig tankekonstruktion för att kunna rationalisera tolkningen av omvärlden och istället utgår ifrån att händelser kan förklaras utifrån en eller flera orsaker kan vi använda definitionen av Kenny (1979) som utgångspunkt:

X orsakar Y om följande villkor är uppfyllda.

1. X föregår temporärt Y
2. X måste ha statistisk signifikant korrelation med Y
3. Förhållandet mellan X och Y kan inte förklaras av andra orsaker

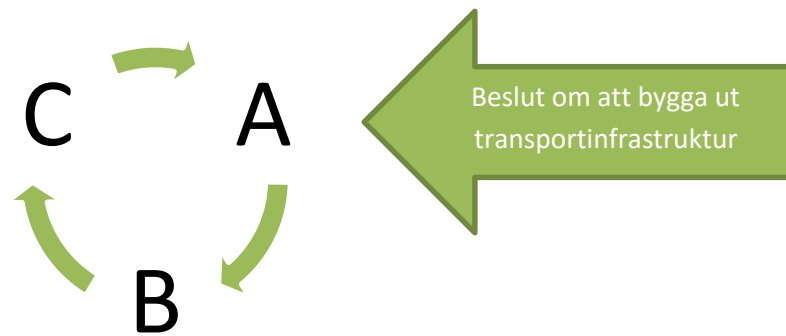
I praktiken är denna stiliserade definition inte tillräcklig för att dra slutsatser om kausalitet i komplexa system som det detta arbete undersöker. Framförallt är villkor nummer tre dels med all sannolikhet inte uppfyllt dels inte möjligt att kontrollera i andra fall än vid ett kontrollerat experiment.

Den absolut tillförlitligaste metoden för att säkerställa att det råder ett kausalt samband mellan variablerna är att genomföra ett kontrollerat experiment där olika värden på x slumpmässigt tilldelas en mängd av stickprovet. Eftersom det inte är möjligt att slumpmässiga tilldela olika kommuner en viss mängd transportinfrastruktur och tillgänglighet i ett experiment uppkommer problem med att analysera orsakssamband. I praktiken är det i regel politiska processer som påverkas av olika former av samhällsekonomiska analyser som ligger till grund för vilka investeringar som görs i transportsystemet (Börjesson et al., 2013).

Problematiken kan illustreras med ett exempel. Om vi utgår ifrån att ett exogent beslut om att bygga ut transportinfrastruktur och får dessa följder

- A) Tillgängligheten förbättras
- B) Tillgängligheten påverkar de sociala förhållandena.
- C) Fler människor flyttar till kommuner med fördelaktiga sociala förhållanden. Fler jobb skapas, samhällen förtätas.

inses att det uppstår en så kallad kausal loop mellan de oberoende och beroende variablerna, se figur 9. Detta gör att det är problematiskt att utifrån enbart eventuellt statistisk korrelation dra några slutsatser om riktningen på kausaliteten mellan tillgänglighet och sociala förhållanden.



Figur 13. En illustration av den kausala loop som visar hur beroende och oberoende variabler kan påverka varandra vilket är en potentiell källa till endogenitet i modellen och bidrar till svårigheter att skatta effekten av infrastrukturutbyggnaden.

När statistiska observationsstudier genomförs, alltså inte kontrollerade experiment, krävs i regel ett gediget forskningsmaterial som hanterar både kvantitativa och kvalitativa aspekter av frågan för att det ska vara rimligt att dra några slutsatser om kausalitet. Enbart en statistisk korrelation mellan beroende och oberoende variabler betyder inte per automatik att värdet på den beroende variabeln på något sätt orsakats av den oberoende variabeln. Några av anledningarna till detta diskuteras vidare i kommande avsnitt.

3.7.2 Endogenitet

Kausala loopar eller omvänd kausalitet är exempel på ett fenomen i samband med regressionsmodeller som kallas endogenitet. Antonakis et al. (2010) menar att detta uppstår när oberoende variabler i modellen inte varierar slumpmässigt över de observerade entiteterna och korrelerar med eventuella orsaker som är utelämnade i modellen. Dessa variabler är då endogena variabler. Om detta inte justeras för blir de skattade parametrarna högst opålitliga och praktiskt taget meningslösa. Motsatsen till endogena variabler är exogena variabler. En exogen variabel är alltså en oberoende variabel som är okorrelerad med modellens felterm. En enkel matematisk definition kan göras utifrån ekvation (41).

$$y = \alpha + \beta x + \varepsilon_i \quad (41)$$

Där x är strikt exogen om villkoret i ekvation (42) är uppfyllt (Antonakis et al., 2010).

$$\text{COV}(x, \varepsilon) = 0 \quad (42)$$

Endogenitetsproblem är viktiga att ta hänsyn till i alla typer av regressionsmodeller som utgår ifrån att det finns ett kausalt förhållande mellan oberoende och beroende variabler. Alltså om modellens syfte är att kunna svara på frågor av formen ”vad blir effekten på y om x förändras med ett visst antal enheter?”. Just denna typ av frågeställning är central för detta arbete men modellformuleringens utformning riskerar att svaret på frågan är

svårt att ge ett exakt svar på grund av endogenitet. Anledningarna till detta diskuteras i resterande del av detta avsnitt. För att en regressionsskattning av den parameter som är av det främsta intresset, β_1 , i modellen i ekvation (31) ska vara fullständigt korrekt och inte påverkad av endogenitetsstörningar gäller följande för kovariansen, ekvation (43) (härledd ifrån Combes et al., 2010 a).

$$\begin{aligned} COV(x_{tillgänglighet_{it}}, x_{Socioekonomi_{1...n,it}} + S_t + Z_i) = \\ \sum_1^n COV(x_{tillgänglighet_{it}}, x_{Socioekonomi_{1...n,it}}) + COV(x_{tillgänglighet_{it}}, S_t) + \\ COV(x_{tillgänglighet_{it}}, Z_i) = 0 \end{aligned} \quad (43)$$

Nollskilda kovarianser ger upphov till störningar i skattningen av parametern för tillgänglighet.

Inom den ekonometriska forskningen för agglomerationseffekter, alltså teorin om att täta samhällen är mer produktiva och har högre lönenivåer än mindre täta samhällen, är endogenitetsproblemet den största potentiella felkällan för skattningar. Även i detta arbete finns det anledning att tro att endogenitet i modellformuleringen riskerar att störa skattningar av parametern för tillgänglighetens samband med den beroende variabeln. En mängd forskning har fokuserat på att uppskatta effekten på produktivitet på grund av ökad täthet. Combes et al. (2010 b) menar att ekonometriska skattningar av en kausal elasticitet mellan täthet och produktivitet eller lönenivå risker att bli felaktiga främst utifrån två grundläggande metodproblem. Det första problemet har att göra med att de mått som används för produktivitet kan vara simultant bestämda med tätheten.

Exempelvis kan produktiva orter attrahera fler arbetare vilket ökar tätheten. En annan aspekt av problemet, som för vissa ur ett ekonometriskt perspektiv är ekvivalent med det tidigare nämnda, är att det finns en i modellen utelämnad variabel som korrelerar med både täthet och produktivitet. Denna variabel menar Combes et al. (2010 a) kan utgöras av att vissa städer har också en attraktionskraft för människor och företag på grund av sitt geografiska läge eller naturliga produktionsfördelar. Dessa platser kan ofta erbjuda en högre lön vilket gör att arbetare väljer att flytta dit, vilket alltså innebär att arbetskraftens lokalisering är endogen. Det rör sig alltså både om en potentiellt motriktad kausalitet, alltså att hög produktivitet ökar tätheten, och ett problem med utelämnade variabler. Combes et al. (2010 b) kallar denna typ av problem för ”endogenous quantity of labour”. I detta arbetes modell finns risk för liknande problem om det finns anledning att misstänka att den fixa effekten som beskriver de utelämnade variablerna som är kommunspecifika har ett samband med tillgängligheten, alltså om $COV(x_{tillgänglighet_{it}}, Z_i) \neq 0$. Den kommunspecifika fixa effekten Z_i är ju olika för de olika beroende variablerna och fångar in den heterogenitet som finns mellan kommunerna som tillgängligheten och de sociala förhållanden inte kan förklara. Kommuner med värden på Z_i som bidrar till att skapa samhällsekonomisk nytta kan ur ett ekonometriskt perspektiv tolkas som kommuner som har en naturlig fördel. Därför finns det anledning att tro att eftersom tillgänglighet, vilket Börjesson et al. (2013) poängterar, påverkas av

politiskt betingade beslut om transportinvesteringar så kan fördelningen av tillgänglighet mellan kommunerna bero på kommunernas grundläggande förutsättningar. Den mest använda metoden för att komma till bukt med denna typ av problem är enligt Combes et al. (2010 a) att använda sig av instrumentvariabler. Något förenklat kan dessa förklaras som en uppsättning variabler som har en stark samvarians med tillgängligheten men som är okorrelerade med de kommunspecifika fixa effekterna.

Det andra problemet benämns i Combes et al. (2010 b) som ”endogenous quality of labour” vilket innebär att produktivare arbetskraft utlokaliseras till tätare områden. Detta kan ha flera anledningar. Produktivitetens fördelarna av hög täthet kan ha större positiva effekter för skickliga arbetare. En annan förklaring kan vara att skicklig arbetskraft dras till täta områden för att de har ett rikare kulturutbud eller liknande. Det kan också finnas stora fördelar för duktiga entreprenörer att verka i en tätbefolkad stad som gör att de söker sig dit. Alltså är det inte bara produktivitet som riskerar att vara simultant bestämt med täthet utan även den lokala arbetskraftens egenskaper och preferenser. I det här arbetet kan det på liknande sätt tänkas att det är personer med bra utbildning och jobb som drar sig större städer med bättre kommunikationer. Dessa människor kan även tänkas inneha ett starkt socialt kapital sedan tidigare. På så sätt är det inte transportinfrastrukturen i sig som leder till att förändra de sociala förhållandena utan snarare en omlokalisering av befolkning betingad till en urbaniseringseffekt. Det går inte heller att säga att en tillgänglighetsförändring förändrar det sociala förhållandet för de som redan bor på en ort då sociala förhållanden lika gärna kan förändras på grund av en omlokalisering av människor. Alltså speglar tillgänglighet egenskaper i kommunerna som kan leda till endogena felkällor i modellen på samma sätt som Combes et al. (2010 b) pratar om ”endogenous quality of labour”.

En åtgärd som Combes et al. (2010 a) menar kan mildra skattningsfelet som endogenitet orsakar är att inkludera observerbara egenskaper för de studerade geografiska områdena, alltså kommuner i denna studie. Som tidigare har diskuterats är, förutom kommunernas tillgänglighet, även indikatorer för socioekonomiska förhållanden inkluderade i modellen vilket motsvarar just denna åtgärd. Detta kan minska modellens endogenitet i viss utsträckning eftersom att det kan minska kovariansen

$COV(x_{tillgänglighet_{it}}, S_t)$ och/eller $COV(x_{tillgänglighet_{it}}, Z_i)$. Dock innebär detta en risk för att de socioekonomiska faktorerna är endogena med den beroende variabeln på så sätt att de både bestäms av gemensamma yttre faktorer.

3.7.3 Heteroskedasticitet

I regel förutsätter regressionsmodeller att residualerna är homoskedastiska alltså att variansen för avvikelserna från modellen är konstant över tid och entiteter. Baltagi (2005) menar att variansen ofta inte är homoskedastisk, alltså heteroskedastisk, för ekonometrisk data och således är det rimligt att misstänka att så även är fallet för det data som detta arbete hanterar. Om residualerna är heteroskedastiska blir skattningen av

parametrarnas standardavvikelse felaktig och i regel underskattad. Detta kan få som konsekvens att när exempelvis t-test genomförs för modellens parametrar kan de felaktigt visa att icke signifikanta parametrar har statistisk signifikans. Heteroskedasticitet kan lyckligtvis hanteras och den metod som används i detta arbete finns förklarad i avsnitt 3.7.3 och 3.7.4.

3.7.4 Robusta standardavvikelser

När det finns misstanke om heteroskedasticitet i residualerna kan detta kompenseras för i fixed- och random effects regression modellerna genom att beräkna robusta standardavvikelser. Álvarez et al. (2013) beskriver en metod för detta som utgår ifrån att beräkna klustrade standardavvikelser där de olika entiteterna, kommunerna, utgör klustren. Denna metod är inkorporerad i programvaran för between effects och fixa effekter när endast kommunspecifika effekter justeras för.

3.7.5 Kolinjäritet

Enligt Weisberg (2013) är en datamatrix med oberoende variabler, X , kolinjär om det finns en vektor med konstanter, a , sådan att $Xa \approx 0$. Om $Xa = 0$ är minst en av de oberoende variablerna en linjär kombination av de andra. Bingham och Fry (2010) skriver att detta är ett vanligt förekommande problem inom multipel regression då det finns fler möjligheter för några av de oberoende variablerna att vara åtminstone approximativt linjärt beroende ju fler av dem som ingår i modellen. De menar att genom att identifiera och justera för multikolinjäritet i en modell går det undvika numerisk instabilitet och samtidigt förminska modellens dimension och därmed beräkningskomplexitet.

Ett vanligt första steg för att identifiera eventuella samband mellan oberoende variabler är genom att studera korrelationsmatrisen vilken visar styrkan på parvis linjära samband mellan variablerna (Belsley et al., 1980). Ett annat vanligt verktyg för att diagnostisera kolinjäritet är variance inflence factor, (VIF). VIF för en oberoende variabel, i , fås av:

$$VIF_i = \frac{1}{1-R_i^2} \quad (44)$$

där $R_i^2 \approx 1$, om variansen för den oberoende variabeln, i , i hög grad förklaras genom en linjär regression av de andra oberoende variablerna, vilket i sin tur ger ett högt VIF. Ett högt VIF-värde tyder på att det finns kolinjäritet mellan de oberoende variablerna. En vanlig gräns är att acceptera VIF-värden under 10 och den används även i detta arbete för att utvärdera graden av kolinjäritet (Belsley et al., 1980).

3.7.6 Spatial korrelation

När geografisk data används i paneldatanalyser uppstår ofta korrelation i tvärsnittsdimensionen mellan entiteterna på grund av deras rumsliga närhet. Detta eftersom att den typ av fenomen som är av intresse att undersöka i denna undersökning.

Om spatiala korrelationen existerar i datan och inte hanteras i modellen kommer detta att ta sig i uttryck som autokorrelation i residualerna. Därmed kan inte en sådan regressionsmodell uppfylla kravet om normalfördelade, alltså slumpmässiga, residualer (Baltagi 2008). Flöhr (2014) har genomfört en studie på hur spatial korrelation påverkar regressionsmodeller för arbetslösheten i Sveriges kommuner. Resultaten tyder på att det finns en rumslig dimension i datan som de oberoende variablerna inte kan förklara. Regressionsmodeller som inkluderade det rumsliga förhållandet mellan kommunerna presterade bättre än de som inte innefattade en spatial dimension. De variabler som ingår i denna studie är i många avseenden av samma karaktär som de Flöhr (2014) undersöker. Det är självklart troligt att även sociala förhållanden har en rumslig fördelning som inte är strikt begränsad av kommungränser och därmed är spatialt korrelerade.

3.8 Programvara

All statistisk modellering, analys och utvärdering har genomförts med hjälp av programvaran MATLAB vilken författarna har tillgång till via studentlicenser från Uppsala Universitet. Främst har två samlingar verktyg utvecklade för MATLAB använts. Dels har ”Statistics and Machine Learning Toolbox” som innehåller en samling metoder för statistisk inferens och analys använts för grundläggande statistiska tester och hantering av data. Denna samling metoder innehåller dock inte några verktyg för ekonometrisk analys av paneldata. För detta ändamål användes istället en samling MATLAB metoder speciellt utvecklade för ekonometriska analyser av paneldata skapade av ekonomer vid Universidad Autónoma de Madrid. Denna innehåller de metoder som använts i studiens estimeringsmodeller för tvärsnittsregressioner samt för fixa och slumpmässiga kommunspecifika effekter. Kring dessa estimeringsmodeller finns ett antal relevanta statistiska tester samt verktyg för att estimerar robusta standardavvikelser vid heteroskedastiska residualer. Metoderna är dokumenterade i Álvarez et. al (2013) och metoderna har i detta arbete körts enligt de instruktioner som finns i dokumentationen.

För att estimerar regressionsmodellerna med både tid- och kommunspecifika effekter programmerades ett script i Matlab enligt formler från Baltagi (2005). Då detta programmerades på egen hand prioriterades metoderna för parameterestimering och signifikanstester vilket innebär att verktyg för att testa och hantera heteroskedasticitet och autokorrelation i residualerna inte har implementerats.

Rådatasammanställning samt enklare sortering och kategorisering av densamma har gjorts i Microsoft Excel.

3.9 Arbetsgång

Tillsammans med handledare, Anna Modin som är utredningssekreterare på Sverigeförhandlingen, bestämdes att vår uppgift skulle vara att studera sociala nyttor i samband med transportsystemutbyggnad. När idén sedan diskuterades med

ämnesgranskare, Anders Karlström, kom vi fram till att ett rimligt omfång på arbetet skulle vara att utifrån ett mått på infrastruktur kartlägga samband med sociala förhållanden. Sedan genomfördes en litteraturgenomgång där tidigare forskning studerades för att kartlägga vad som gjorts tidigare inom detta forskningsfält samt vilka sociala faktorer som kan ge en indikation på social nytta. Insamlingen av data för de sociala faktorerna genomfördes sedan baserat på tidigare forskning, Sverigeförhandlingens definition på social nytta samt vilken statistik som fanns tillgänglig. En modellstruktur ansattes genom att studera etablerade samhällsekonomiska analysmodeller och statistiska metoder för att undersöka sambanden valdes ut och studerades. De metoder som har använts har tillkommit efter arbetets gång allt efter som vi upptäckte vad som behövdes för våra ändamål. Därmed har metodutveckling och framställning av resultat skett parallellt.

4. Data

I detta kapitel presenteras den data som används i undersökningen och hur det insamlats. Dels presenteras var datan kommer ifrån och vilken struktur den har men också vilka problem som uppstått och justeringar som gjorts.

4.1 Geografisk indelning på kommunnivå

All statistik som används i detta arbete är geografiskt uppdelad på kommunnivå. Att det är data med just denna geografiska struktur som används är givetvis inte självklart och det medför ett antal problem och begränsningar som diskuteras mer utförligt senare i detta avsnitt. Det är framförallt två faktorer som utgör de huvudsakliga argumenten för att använda denna typ av data. För det första är kommuner de aktörer som utgör den absoluta majoriteten av förhandlingsparter i Sverigeförhandlingen. Det är alltså av stor vikt i förhandlingsprocessen att kunna använda sig av nyttobedömningar och resultat som är lättillgängliga och tolkningsbara på kommunnivå för att enkelt kunna förhandla. I Sverigeförhandlingen är alltså data och resultat på kommunnivå naturligt att hantera. Den andra faktorn är praktiska skäl. Utifrån arbetets omfattning och prioriteringar samt breda infallsvinkel på vad som utgör social nytta insågs att statistik behövdes för ett flertal olika möjliga aspekter av social nytta som kan tänkas ha ett samband med transportinfrastruktur. På grund av det finns relativt stora mängder lättillgänglig data för ett flertal sociala förhållanden på just kommunnivå utgjorde detta en praktisk och förenklande omständighet i arbetet.

Som tidigare nämnt finns ett antal problematiska omständigheter att ta i beaktande när denna typ av data används. Det kanske viktigaste är att många grundläggande faktorer varierar enormt kommuner emellan. En överblick av skillnaderna ges av Sveriges Kommuner och Landstings kommungruppsuppdelning som klassificerar kommuner utifrån ett antal demografiska och ekonomiska faktorer. Uppdelningen, som gjordes 2011, ser ut enligt följande⁹ (SKL, 2011)

1. **Storstäder** (3 kommuner). Kommuner med en folkmängd som överstiger 200 000 invånare.
2. **Förortskommuner tillorstäder** (38 kommuner). Kommuner där mer än 50 procent av nattbefolkningen pendlar till arbetet i någon annan kommun. Det vanligaste utpendlingsmålet ska vara någon av storstäderna.
3. **Större städer** (31 kommuner). Kommuner med 50 000-200 000 invånare samt en tätortsgrad överstigande 70 procent.

⁹ Sedan denna har Uppsala kommun nått över 200.000 invånare och skulle därför idag klassificeras som "storstad" istället för "större stad".

4. **Förortskommuner till större städer** (22 kommuner). Kommuner där mer än 50 procent av nattbefolkningen pendlar till arbetet i en annan kommun. Det vanligaste utpendlingsmålet ska vara någon av de större städerna i grupp 3.
5. **Pendlingskommuner** (51 kommuner). Kommuner där mer än 40 procent av nattbefolkningen pendlar till en annan kommun.
6. **Turism- och besöksnäringsskommuner** (20 kommuner). Kommuner där antalet gästnätter på hotell, vandrarhem och campingar överstiger 21 nätter per invånare eller där antalet fritidshus överstiger 0,20 fritidshus per invånare.
7. **Varuproducerande kommuner** (54 kommuner). Kommuner där 34 procent eller mer av nattbefolkningen mellan 16 och 64 år är sysselsatta inom tillverkning och utvinning, energi och miljö samt byggverksamhet enligt Svensk Näringsgrensindelning (SNI2007).
8. **Glesbygdskommuner** (20 kommuner). Kommuner med en tätortsgrad understigande 70 procent och mindre än åtta invånare per kvadratkilometer.
9. **Kommuner i tätbefolkad region** (35 kommuner). Kommuner med mer än 300 000 personer inom en radie på 112,5 kilometer.
10. **Kommuner i glesbefolkad region** (16 kommuner). Kommuner med mindre än 300 000 personer inom en radie på 112,5 km.

Dessa variationer mellan kommunerna gör att det uppstår svårigheter med att jämföra olika kommuner med varandra. Eftersom två kommuner kan ha olika stor yta, invånarantal och geografiskt fördelaktiga lägen i Sverige är det ibland svårt att i vissa lägen göra en rättvis jämförelse mellan kommuner. Med andra ord finns det starka anledningar att misstänka att datan på grund av den geografiska uppdelningen är heterogent mellan kommunerna, alltså att de studerade egenskaperna inte går att se som observationer från en slumpvariabel med gemensam fördelning. Dessa problem har diskuterats i kapitel 3. Då detta arbete studerar samband mellan tillgängligheten i olika kommuner och olika sociala förhållanden kan dessa samband eventuellt bero på faktorer som just kommuners storlek och läge i stället för dess sammanlagda tillgänglighet. Ett mer högupplöst data, på stadsdels- eller individnivå, hade till viss del kunna rådgöra om detta problem.

4.2 Oberoende variabler

I dessa avsnitt presenteras den data som utgör de oberoende variablerna i de statistiska analyserna.

4.2.1 Tillgänglighet

Som diskuterats i avsnitt 3.1 är detta examensarbete baserat på ett potentialbaserat tillgänglighetsmått som beskriver tillgängligheten till arbetsplatser. Är det då motiverat att använda detta som variabel för att försöka förklara de sociala förhållandena i olika kommuner där ju även tillgängligheten till andra typer av målpunkter kan vara central? Enligt Svante Berglund¹⁰, transportanalytiker på WSP och doktor inom transport och lokaliseringsanalys, är reseflödena för icke arbetsrelaterade resor i mycket hög grad korrelerade med de arbetsrelaterade resorna. Därigenom torde ett mått på tillgängligheten till arbetsplatser även vara ett bra mått för den övergripande tillgängligheten.

Det tillgänglighetsmått som används i detta arbete är geografiskt uppdelat på kommunnivå. Varje kommuns tillgänglighet kan förstås som den, av GK viktade, tillgängligheten till samtliga övriga kommuner. Viktigt att nämna är att tillgänglighetsmåtten i grunden är baserat på Trafikverkets modell Sampers som beräknar personresor i Sverige. Denna är skattad baserat på verkligt resande mellan så kallade SAMS-områden som är drygt 10 000 geografiskt uppdelade områden som täcker hela Sveriges bebodda yta (Sehlin, 2012 och Trafa, 2011). För att få tillgänglighetsmåtten på kommunnivå har resultaten från Sampersområdena aggregerats upp. Detta arbete är genomfört av WSP som också har kvalitetssäkrat datan för att kompensera för de problem uppaggregering av geografisk data medför.

Tillgänglighetsdatan som författarna fått tillgång till från WSP har två problematiska egenskaper. Det första problemet är att vid analys av datan upptäcktes en trend med sjunkande tillgänglighetsmått över tid för näst intill samtliga Sveriges kommuner vilket inte verkar troligt. Enligt Berglund¹¹ kan detta bero på att olika skattningar använts eller att priskänsligheten, alltså parameter β i ekvation (45), inte är justerade för inflation eller att tillgänglighetsmåtten för de olika tidpunkterna är baserade på olika trafiksimuleringar. För att hantera detta medelvärdesjusterades tillgänglighetsmåtten för 1997 och 2010 så att de fick samma medelvärde som tillgänglighetsmåtten 1985. På detta sätt fångas alla skillnader mellan kommunerna ett visst år in, samtidigt som den generella trenden mellan åren blir av en mer trolig natur än en sjunkande trend. Nackdelen är givetvis att det inte går att säga något om den generella förändringen av tillgänglighet över tid. Tillgänglighetsdatan som använts speglar alltså förhållandet mellan kommuner både i tvärsnitt och över tid korrekt men den genomsnittliga utvecklingen för alla kommuner av författarna har justerats till att vara konstant vilket behöver hållas i åtanke vid fortsatt läsning av rapporten.

Den andra problematiska omständigheten var att tillgänglighetsmåtten endas erhöles för tre tidpunkter (tre år, 1985, 1997 och 2010) medan det är fördelaktigt att ha värden för alla år för att utföra regressioner av tidsseriekaraktär. För att fylla de luckor som fanns

¹⁰ Svante Berglund, transportanalytiker på WSP och doktor inom transport och lokaliseringsanalys. Intervju 2015-10-21 samt efterföljande mailkorrespondens.

¹¹ Ibid.

mellan åren 1985, 1997 och 2010 har en linjärapproximation gjorts för att skatta värdena för de mellanliggande åren. Detta bedöms fånga in åtminstone de grundläggande förändringar av tillgänglighet som sker över då tillgängligheten aggregerad på kommunnivå tenderar att förändras långsamt. Givetvis innebär denna justering att det finns risk för att förändringar av tillgängligheten mellan de tidpunkter som mätvärden finns för inte kan fångas in. Även detta är viktigt att ha i åtanke genom resterande del av rapporten.

4.2.2 Socioekonomiska faktorer

Det finns mycket forskning som tyder på att olika socioekonomiska faktorer har stor inverkan på människors sociala välmående. Till exempel skriver Adler et al. (1994) att det finns starka samband mellan socioekonomiska faktorer och hälsa. I en annan studie argumenterar Leighley och Nagler (1992) för det finns samband mellan socioekonomisk status och valdeltagande. De rapporterade om starka samband mellan bland annat inkomstnivå och valdeltagande.

Då det utan tvekan finns andra faktorer än endast tillgängligheten som förklarar sociala förhållanden hos en grupp människor är det relevant att försöka fånga in dessa i regressionsmodellen. Som tidigare nämndes i metodkapitlet implementeras därför ett antal socioekonomiska faktorer som oberoende variabler tillsammans med tillgänglighet. De socioekonomiska faktorerna som anses intressanta är medelinkomst, medelålder, befolkningstäthet, biltäthet och andel invånare med utländsk bakgrund. En beskrivning av dessa variabler och vilken källa de kommer från ges i tabell 2. Värt att nämna är att medianinkomsten inte är justerad för inflation eller reallöneutveckling och innebär att datan har en stigande trend över tid för hela riket.

Tabell 2. Sammanställning av de socioekonomiska faktorer som anses intressanta att ha med i en modell.

Socioekonomisk faktor	Källa	Beskrivning
Medianinkomst	SCB	Medianen av förvärvsinkomst, tkr
Medelålder	SCB	Befolkningens medelålder
Befolkningstäthet	SCB	Invånare per kvadratkilometer
Biltäthet	Trafikanalys	Antal bilar per 1000 invånare
Andel med utländsk bakgrund	SCB	Antal personer med utländsk bakgrund delat med totala antalet invånare

4.2.3 Kolinjäritet

Som tidigare diskuterats i metodkapitlet finns risk att de oberoende variablerna är alltför linjärt korrelerade. Detta kan ge felaktiga parameterskattningar och därmed bidra till fel i modellen. För att undersöka kolinjäriteten i de oberoende variablerna har variance inflation factor (VIF) och Pearsons korrelationskoefficient använts. Dessa har beräknats utifrån ett medelvärde över femton år för respektive kommun och variabel. Se tabell 3 respektive tabell 4. VIF för tillgänglighet är högre än gränsvärdet på 10. Eftersom att tillgängligheten är central för undersökningens syfte är det inte ett alternativ att utelämna den i regressionerna. För att minska VIF för tillgänglighet kan istället de variabler som har hög korrelation med tillgängligheten uteslutas.

Tabell 3. De oberoende variablernas variance inflation factor (VIF). Tillgänglighet har en VIF högre än gränsvärdet på 10.

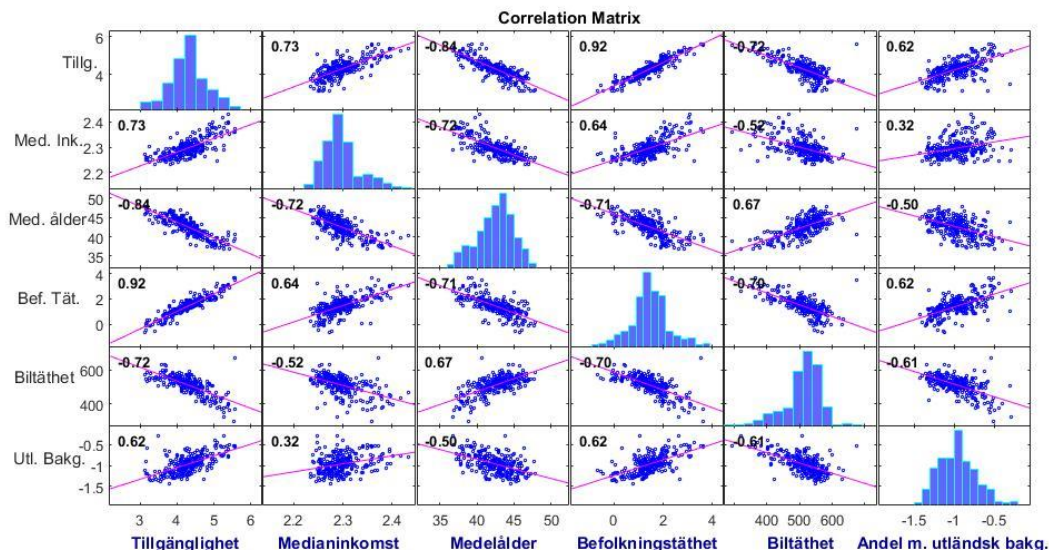
Variabel	VIF
<i>Tillgänglighet</i>	13.2830
<i>Medianinkomst</i>	2.4945
<i>Medelålder</i>	4.2776
<i>Befolkningstäthet</i>	8.0704
<i>Biltäthet</i>	2.4545
<i>Andel med utländsk bakgrund</i>	1.9174

Som ses i tabell 4 är korrelationen mellan tillgänglighet och medelålder respektive befolkningstäthet de två starkaste. Dessa två variabler utesluts därför ur modellen för att minska multikolinjäriteten i de oberoende variablerna. Dessutom tyder en så pass hög linjär korrelation att de till stor del beskriver samma variationer mellan kommunerna som tillgänglighet vilket medför att det är relativt liten mängd information som går till spillo. En intressant anmärkning att göra kring detta är att tillgänglighet alltså till viss del kan tjäna som ett mått på en kommuns befolkningstäthet och vice versa.

Tabell 4. Korrelationskoefficienterna för de oberoende variablerna i modellen visar hur pass korrelerade de är med varandra. Både medelålder och befolkningstäthet korrelerar till hög grad med tillgänglighet.

Korrelationskoeff.	Tillg.	Med. Inkomst	Med. Ålder	Bef. Täthet	Biltäthet
<i>Tillgänglighet</i>	1.00				
<i>Medianinkomst</i>	0.7257	1.00			
<i>Medelålder</i>	-0.8355	-0.7245	1.00		
<i>Befolkningstäthet</i>	0.9250	0.6437	-0.7086	1.00	
<i>Biltäthet</i>	-0.7206	-0.5217	0.6738	-0.7013	1.00
<i>Andel med utländsk bakgrund</i>	0.6162	0.3158	-0.5024	0.6185	-0.6084

I figur 11 finns spridningsdiagram över variablernas förhållanden för att grafiskt illustrera korrelationerna. Diagrammen innehåller också en skattad enkel regressionslinje som illustrerar korrelationens riktning. Diagonalelementen i figuren visar ett histogram över respektive variabel och kan därmed ge en grundläggande uppfattning av variablernas fördelningar. Notera att vissa av variablerna är transformerade vilket beskrivs i avsnitt 4.2.4.



Figur 14. Spridningsdiagram för de oberoende variablerna för att ge en grafisk överblick av de oberoende variablernas fördelning samt regressioner dessemellan. I diagonalen ses fördelningen av observationerna. I resterande celler går det att se korrelationskoefficienten samt en regression mellan variabeln på X-axeln och variabeln på Y-axeln.

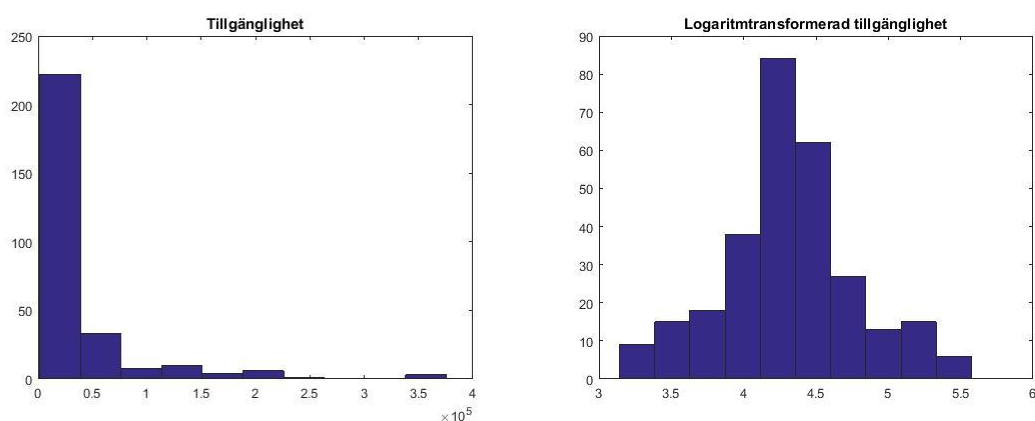
Efter att medelålder och befolkningstäthet utesluts är VIF för samtliga variabler betydligt lägre än tidigare och befinner sig med marginal under gränsvärdet 10, se tabell 5. Detsamma gäller för korrelationskoefficienterna dem emellan. Uppsättningen av oberoende variabler bestående av tillgänglighet, medianinkomst, biltäthet samt andel i befolkningen med utländsk bakgrund är därmed tillräckligt icke kollinjär för att inte leda till felaktiga parameterkattningar. På grund av det här skälet kommer rapporten fortsättningsvis använda denna uppsättning av oberoende variabler.

Tabell 3. De oberoende variablernas variance inflation factor då befolkningstätheten och medelåldern tagits bort.

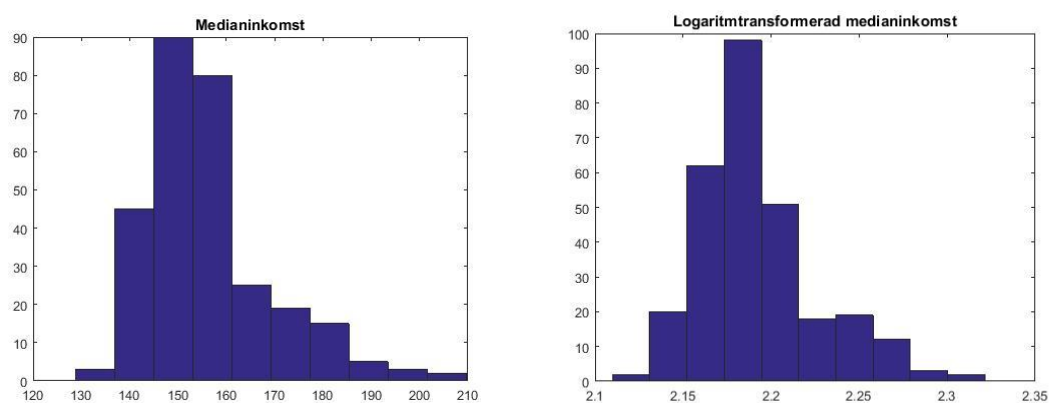
Variabel	VIF
Tillgänglighet	3.8169
Medianinkomst	2.2577
Biltäthet	2.3006
Andel med utländsk bakgrund	1.8944

4.2.4 Observationernas fördelning

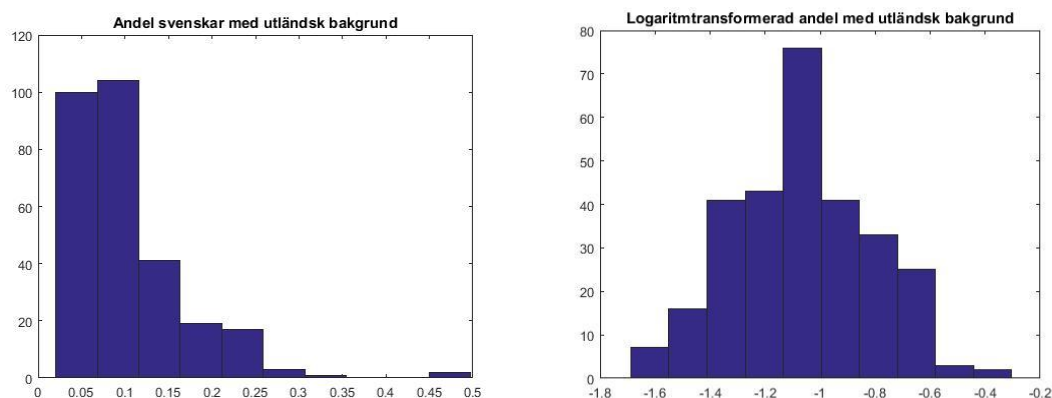
En del av de oberoende variablerna som används i rapportens regressioner har en fördelning av observationerna som är något så när approximativt lognormal. För att få denna skeva fördelning att bli något mer normalfördelad görs en logaritmtransformering med 10 logaritmen. De oberoende variabler som har logaritmtransformerats i den här rapporten är tillgänglighet, medelinkomst och andel med utländskbakgrund. I figurerna 12-14 nedan visas histogram för dessa observerade variabler. Histogrammen till vänster visar fördelningen av de observerade värdena utan någon transformation och de till höger visar fördelningen efter att variablerna har logaritmtransformerats.



Figur 15. Histogram för tillgängligheten i Sveriges kommuner där tillgängligheten är angiven i linjär skala till vänster och logaritmtransformerad till höger.



Figur 16. Histogram för medianinkomsten i Sveriges kommuner där medianinkomsten är angiven i linjär skala till vänster och logaritmtransformerad till höger.



Figur 17. Histogram för andelen med utländsk bakgrund i Sveriges kommuner där andelen med eftergymnasial utbildning är angiven i linjär skala till vänster och logaritmtransformerad till höger.

Logaritmering av variabler kan användas vid situationer när förhållandet mellan ej transformerade oberoende och beroende variabler är icke linjärt men ändå ansätta en linjär regressionsmodell (Benoit, 2011). När en oberoende variabel logaritmeras förändras tolkningen av förhållandet mellan den och den beroende variabeln. Se avsnitt 3.3.2 för en förklaring av detta.

4.3 Beroende variabler – Sociala förhållanden

Urvalet för vilka sociala faktorer som studeras i den här rapporten gjordes genom att studera tidigare forskning på social inverkan av transportsystem samt Sverigeförhandlingens definition av sociala nyttor. Tidigare forskning pekar hälsa, socialt stöd, känsla av samhörighet, socialt nätverk, tilltro till institutioner, jämlikhet och allmän livskvalitet som viktiga indikatorer. Sverigeförhandlingen framhåller främst socialt kapital som en social nytta som de menar kan mätas i form av tillit, trygghet, delaktighet i val och föreningslivsaktivitet. De nämner dock att det kan vara svårt att få tag i geografiskt högupplöst data för dessa indikatorer.

Utifrån den tidigare forskningen och Sverigeförhandlingens definition av social nytta söktes statistiska databaser som SCB, och SKLs databas KOLADA igenom för att undersöka vilka tänkbara indikatorer för sociala förhållanden det fanns tillförlitlig data på för alla Sveriges kommuner över en längre tidperiod. Att datan finns för alla kommuner över en lång tidsperiod har varit ett krav i datainsamlingen och ett antal variabler som skulle kunna vara av intresse har fått väljas bort på grund av detta. Varje utvald variabel tjänar alltså som en beroende variabel i analyserna. I tabell 6 finns en sammanfattning av de mätbara indikatorer som anses kunna ge en uppfattning om socialt välstånd och social nytta som den här studien har valt att undersöka. I tabellen framgår vilka indikatorerna är, vilka sociala effekter de indikerar och vilken myndighet datan kommer ifrån. Genom att studera dessa faktorer hoppas studien fånga in de

sociala effekter som kan kopplas till social nytta. I appendix A återfinns myndigheternas egna beskrivning av de olika variablerna.

Tabell 4. Sammanställning av de beroende variabler som undersöks och antas vara indikatorer på social nytta.

Indikator	Social effekt	Källa	Beskrivning
Förväntad livslängd, män	Hälsa	SCB	Beräknad återstående medellivslängd i antal år vid födseln.
Förväntad livslängd, kvinnor	Hälsa	SCB	Beräknad återstående medellivslängd i antal år vid födseln.
Deltagande i senaste riksdagsval	Tillit till institutioner	SCB	Antal angivna röster i senaste riksdagsvalet dividerat med antal röstberättigade, multiplicerat med 100.
Sjukpenningtalet, dagar/registrerad försäkrad	Hälsa	Försäkringskassan	Antal utbetalda dagar med sjukpenning registrerad försäkrad i åldrarna 16-64 år.
Lönegap, medianlön kvinnor – medianlön män anställda av kommunen, kr	Jämställdhet	SCB	Lönegap median kvinnor – median män anställda av kommunen, kronor (i månaden).
Kvinnor som tar ut tillfällig föräldrapenning (VAB), andel av antal dagar (%)	Jämställdhet	Försäkringskassan	Andel tillfällig föräldrapenning som kvinnor tar ut. Avser: Dagar i samband med vård av barn.
Barn 1-12 år inskrivna i förskola och skolbarnsomsorg, andel (%)	Socialt nätverk	SCB Skolverket	Antal inskrivna 1-12 år i förskola och skolbarnsomsorg dividerat med antal barn 1-12 år.
Bidragshushåll ekonomiskt bistånd, antal/1000 inv	Social exkludering/Känsla av samhörighet	SCB och Socialstyrelsen	Antal hushåll med ekonomiskt bistånd, dividerat med antal invånare totalt dividerat med 1000.
Deltagartillfällen i idrottsföreningar, antal/inv 7-20 år	Socialt nätverk/Hälsa	Riksidrottsförbundet	Antal deltagartillfällen i åldern 7-20 år i LOK-stödsberättigade idrottsföreningar under året, dividerat med antal invånare 7-20 år den 31/12.

Förvärvsarbetande invånare 20-64 år, andel (%)	Socialt nätverk/Känsla av samhörighet	SCB (RAMS)	Antal förvärvsarbetande i åldern 20-64 år dividerat med antal invånare i åldern 20-64 år.
Gymnasiefrekvens, (%)	Utbildning/Socialt nätverk	SCB Skolverket	Antal barn 16-18 inskrivna i gymnasieskolan i kommunen dividerat med antal barn 16-18 år i kommunen.
Kvinnor som tar ut föräldrapenning, andel av antal dagar (%)	Jämställdhet	Försäkringskassan	Andel av föräldrapenning (sett till nettodagar) som kvinnor tar ut.
Andel av invånare 25-64 år med eftergymnasial utbildning (%)	Utbildning/Socialt nätverk	SCB	Invånare med eftergymnasial utbildning 25-64 år, andel (%). Eftergymnasial avser: eftergymnasial utbildning kortare än 3 år, längre än 3 år samt forskarutbildning.

4.3.1 Förväntad livslängd

Jones & Lucas (2013) diskuterar att transportsystemet på lång sikt kan påverka hälsan hos de som interagerar med systemet på flera olika sätt. Däremot menas att dessa kopplingar kan vara svåra att observera eller dra några kausala slutsatser utifrån. I de transportpolitiska målen framhålls att utformningen av transportsystemet ska bidra till en ökad hälsa (Näringsdepartementet 2009). Kopplingen mellan den av transportsystemet genererade tillgängligheten och hälsa är däremot inte lika framträdande i litteraturen. Förväntad livslängd för män respektive kvinnor används i arbetet som en indikator för hälsa. Givetvis finns det, som för alla indikatorer på komplexa och subjektiva fenomen, ett antal svagheter med dessa. Den kanske tydligaste är att ett långt liv inte per automatik synonymt med ett hälsosamt liv och vice versa. Det finns dessutom över tid finns en, för samtliga kommuner, generell ökande trend som beror på yttre faktorer som exempelvis allmänna förbättringar i sjukvård, livskvalité och liknande som måste tas i beaktande (Folkhälsomyndigheten 2014). Däremot torde skillnader mellan kommuner kunna samvariera på skillnader i tillgänglighet.

Ökad förväntad livslängd anses i arbetet vara en social nytta.

4.3.2 Deltagande i senaste riksdagsval

Valdeltagande tjänar som en indikator för det som kallas vertikalt socialt kapital. Alltså människors tillit till samhällets institutioner. I 2013 års Stockholmsförhandling (2013) som undersöker utbyggnaden av tunnelbanan i Stockholms sociala konsekvenser så omnämns valdeltagande som en väl fungerande indikator på detta. Högt valdeltagande

tyder på högt vertikalt socialt kapital vilket är nyttskapande då det bidrar till ett välfungerande demokratiskt samhälle. I samma rapport diskuteras att det är fördelaktigt att inkludera fler indikatorer baserade på undersökningar om medborgares trygghet och tillit för att få ett mer robust mått på socialt kapital. Tyvärr har inte denna typ av data funnits tillgängligt på kommunnivå och över lång tid och har därmed inte kunnat användas i arbetet.

Ökat valdeltagande anses i arbetet vara en social nytta.

4.3.3 Sjukpenningtalet

Sjukpenningtalet definieras som det genomsnittliga antalet dagar med sjukpenning som tas ut av de arbetsföra invånarna i kommunen och används som indikator på hälsa och kompletterar därmed den förväntade livslängden. Folkhälsomyndigheten (2014) menar att det finns starka samband mellan hälsa och social position även om orsakssambanden i många avseenden är dubbelriktade. God hälsa kan alltså både bidra till människors möjlighet att stärka sin sociala position men även gäller att en stark social position kan generera god hälsa. Därigenom torde sjukpenningtalet fånga in en bild av den generella hälsan hos den arbetande delen av befolkningen och därmed även säga något om de sociala förhållandena i kommunen.

Minskat sjukpenningtal anses i arbetet vara en social nytta.

4.3.4 Lönegap mellan kvinnor och män

Jämlikhet mellan könen är inkluderat i transportpolitikens så kallade funktionsmål där det formuleras som att ” Transportsystemet ska vara jämställt, dvs. likvärdigt svara mot kvinnors respektive mäns transportbehov” (Näringsdepartementet 2009). Lönegapet mellan kvinnor och män är en indikator som inte direkt kan kopplas till kvinnors respektive mäns transportbehov. Snarare ger det en uppskattning av den ekonomiska jämställdheten i kommunerna vilket i sig är ett intressant socialt förhållande att undersöka. Däremot kan det finnas indirekta kopplingar mellan tillgänglighet och indikatorn som kan ge en fingervisning om huruvida transportsystemet ger jämlika förutsättningar. Sandow & Westin (2006) visar på ett antal intressanta skillnader i pendlingsmönster mellan könen män och kvinnor. Kvinnor pendlar mer sällan, i regel kortare sträckor och är mer positivt inställda till att använda kollektivtrafik än män. Därigenom är det rimligt att tänka sig att tillgängligheten kan inverka på könen möjlighet till matchning på arbetsmarknaden. Lönegapet mellan kvinnor används som en indikator för jämställdhet på arbetsmarknaden vilket även görs på samma sätt i Försäkringskassan (2015) . Viktigt att poängtera är dock att det data som används i arbetet enbart rör anställa i kommunal verksamhet och därför bara speglar en liten del av kommuninvånarnas arbetssituation.

4.3.5 Vård av Barn (VAB) och Föräldrapenning

I en studie från Försäkringskassan (2015) hävdas att en ojämn fördelning av föräldrapenning korrelerar med en ojämn fördelning av hushållsarbete. I studien används just fördelningen av uttag av VAB samt föräldrapenning mellan föräldrarna som indikator för jämställdhet i hemmet. På samma sätt används dessa data i detta arbete. De förutsättningar transportinfrastrukturen skapar i form av tillgänglighet kan tänkas påverka villkoren för vem av föräldrarna som tar ut föräldrapenning eller vårdar sjukt barn.

VAB och föräldrapenning som går mot en kvot på 0.5, alltså att män och kvinnor tar ut lika många dagar, anses i arbetet vara en social nytta.

4.3.6 Gymnasiefrekvens

Enligt Skolverket (2014) är gymnasieutbildning centralt för framtida möjligheter till etablering på arbetsmarknaden och därmed en stark social position. Det framhålls att det som främst påverkar etableringen är att ha slutfört årskurs tre. I detta arbete däremot används frekvensen ungdomar som påbörjar gymnasiet som indikator istället. Detta eftersom att skolgången i sig, oavsett avslutade studier eller inte, kan bidra till att bygga ett socialt nätverk och därmed stärka det horisontella sociala kapitalet och därmed utgöra en nytta.

Ökad gymnasiefrekvens anses i arbetet vara en social nytta.

4.3.7 Förvärsarbetande

Fors och Brülde (2012) skriver att arbete ger identitet och social status och därför innebär arbetslöshet ofta en knäck för självförtroendet och självbilden. I samma artikel presenteras belegg för att förvärsarbetande visar på signifikant högre resultat för både livstillfredsställelse och subjektivt välbefinnande än arbetslösa i Sverige 2010 och 2011. Jönsson (2003) argumenterar för att arbetslöshet bidrar till att skapa negativa självkänslor och befäster en negativ social position.

Ökad andel förvärsarbetande i idrottsföreningar anses i arbetet vara en social nytta.

4.3.8 Deltagartillfällen i idrottsföreningar

Den här variabeln beskriver antalet deltagartillfällen i en idrottsförening per person för kommuninvånare i åldern 7-20 år. I en studie gjord av Riksidrottsförbundet (FoU 2005) visade det sig att ungdomar som var aktiva inom en idrottsförening hade en mer positiv inställning till livet än ungdomar i övrigt. Studien konstaterar att ungdomar som idrottar upplever en minskad stress. Idrottsföreningar är även en mötesplats för ungdomar där de kan skapa ett socialt kontaktnät och det finns också resultat som tyder på positiva effekter ur ett integrationssyfte (Riksidrottsförbundet, 2010). Forskning tyder alltså på flertalet positiva sociala effekter som följd av idrottande. Ett hinder för idrottsutförande

kan vara transporten till och från platsen där idrotten utövas. Om då en ökad tillgänglighet till dessa platser kan leda till ökat idrottande får det se som en positiv effekt. Till skillnad från de övriga variablerna som fanns tillgängliga för åren mellan 2000 och 2014 så fanns antalet deltagartillfällen i idrott endast tillgänglig för åren mellan 2002 och 2014.

Ökat antal deltagartillfällen i idrottsföreningar anses i arbetet vara en social nytta.

4.3.9 Bidragshushåll

Som diskuterats i kapitlet om tidigare forskning är social exkludering ett av de flitigaste använda begreppen i forskning kring sociala konsekvenser av transportsystemet. Arbetslöshet, framförallt långtidsarbetslöshet, är ofta en central beståndsdel i begreppet. Transportsystemets utformning och förutsättningar de genererar för människors möjlighet till arbete, och företagets möjlighet till arbetskraft, tycks enligt denna forskning kunna få stor inverkan. Antalet hushåll med ekonomiskt bidrag bör därför kunna vara en indikator på ekonomisk utsatthet i kommunen som kan bero på långtidsarbetslöshet eller andra faktorer. Socialstyrelsen (2013) menar att ekonomiskt biståndstagande har flera fördelar som indikator på fattigdom och ekonomisk utsatthet eftersom det tydligt kopplar till inkomstnivåer snarare än andra mått baserade på förmögenhet eller liknande. Märk väl att bidrag i sig ju mer än något annat är ett verktyg för att lindra ekonomisk utsatthet och därför i sig inte utgör någon social onytta. Ur ett samhällsekonomiskt perspektiv däremot är det givetvis mer fördelaktigt att färre människor behöver ekonomiskt bidrag från samhället för att undvika ekonomisk utsatthet.

Viktigt att poängtera är att andel förvärvsarbetande inte nödvändigtvis behöver innebära komplementet till andelen arbetslösa. Detta eftersom studenter och liknande inte är inkluderat. Dock är det rimligt att tro att en hög grad förvärvsarbetande i hög grad samvarierar med låg arbetslöshet.

Minskat antal bidragshushåll anses i arbetet vara en social nytta.

4.3.10 Barn inskrivna i förskola och skolbarnsomsorg

Denna beroende variabel är något mer komplex att förhålla sig till än de övriga. Att många barn går i förskola behöver inte i sig vara en social nytta. Det kan dock indikera att båda föräldrarna är sysselsatta och därmed utgöra ett tecken på jämställdhet, ekonomiskt välstånd och socialt deltagande i samhället. Angående barnens utveckling pekar forskningen på att barn som går i förskola utvecklar sin sociala förmåga och får ett försprång i inlärningsförmågan jämfört med barn som inte gör det (Barnomsorg och skolkommittén 1997). Å andra sidan kan det tänkas att i mycket välbärgade familjer behöver inte båda föräldrarna förvärvsarbete vilket innebär att en av föräldrarna kan ta hand om barnen i hemmet och att barnet inte går förskola. I datan ingår även

skolbarnsomsorg för barn upp till tolv år, alltså fritids. Även för barn i lite äldre åldrar kan umgänget med andra barn främja deras sociala utveckling.

Ökad förskolefrekvens anses i arbetet, med viss försiktighet, som en social nytta.

4.3.11 Eftergymnasial utbildning.

Enligt Folkhälsomyndigheten (2015) finns det starka samband mellan utbildningsnivå och såväl fysisk som psykisk hälsa. Även sambandet mellan högre utbildning och arbetsmöjligheter väl underbyggt. Dessutom kan genomförandet av en eftergymnasial utbildning i sig tänkas stärka det sociala kapitalet och sociala nätverk. Indikatorn som används för detta är helt sonika andel av befolkningen i kommunen med någon form av eftergymnasial utbildning. På samma sätt som den förväntade livslängden är även eftergymnasial utbildning något som har en generellt ökande trend över tid som troligen till största del är exogen från transportsystemet och tillgänglighetens utveckling över tid och som därför bör justeras för.

Ökad andel av befolkningen med eftergymnasial utbildning anses i arbetet vara en social nytta.

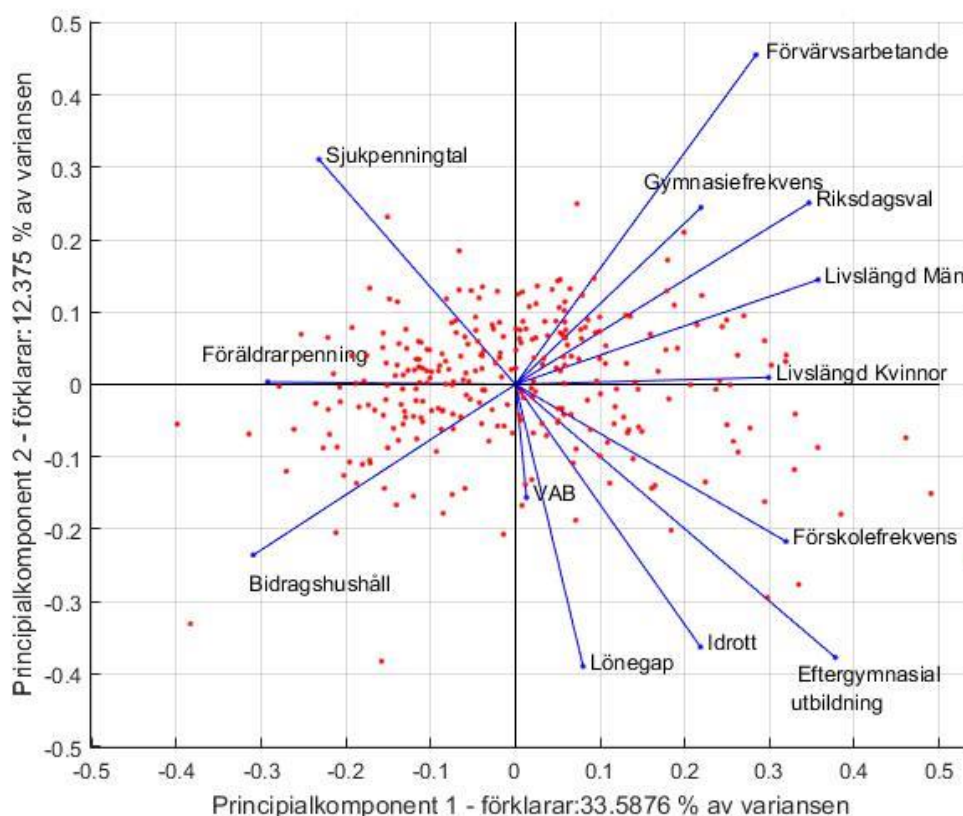
4.3.12 Relevanta sociala effekter som inte ingår i undersökningen

Av de sociala effekter som nämns i tidigare forskning och av Sverigeförhandlingen fångas de flesta in av de variabler som finns i tabell 6. Dock saknas en bra indikator för trygghet och genomsnittlig nöjdhet. Data för indikatorer för både trygghet och nöjdhet återfanns i KOLADA i form av medborgarundersökningar men ansågs inte vara tillämpbara då de inte sträckte sig tillräckligt långt tillbaka i tiden samt att det förelåg en oregelbundenhet i vilka år det fanns tillgängligt. Långtidsarbetslöshet var också något som den tidigare forskningen pekar på som en påverkande faktor till social nytta men som det inte fanns tillräckligt bra statistik för. Datan för långtidsarbetslöshet fanns tillgänglig från 2006 och framåt vilket anses vara ett för kort historiskt tidsperspektiv. Istället undersöks andelen förvärvsarbetande i varje kommun då det kan fungera som en liknande indikator som långtidsarbetslösheten.

4.3.13 Förhållanden mellan de beroende variablerna

Denna undersökning omfattar alltså 13 beroende variabler. Eftersom variablerna är tänkta att belysa olika aspekter av de sociala förhållandena i kommunen som kan tänkas påverkas av tillgänglighet och i förlängningen transportinfrastruktur är det relevant att säkerställa att de inte är alltför korrelerade och beskriver ”samma” underliggande fenomen. Om så är fallet är det helt enkelt onödigt att inkludera flera indikatorer på samma fenomen i analysen. Ett etablerat verktyg för att undersöka strukturer i flerdimensionellt data är så kallad principalkomponentsanalys (PCA). En enkel förklaring ges av Sehlens (2014) som beskriver PCA som en statistisk metod för att undersöka graden av redundans av information i olika variabler i ett dataset. Metoden

bygger på att transformera riktningen på datasetets baser så att så stor del som möjligt av variansen kan förklaras av så få dimensioner som möjligt. Om variansen i ett mångdimensionellt dataset till stor del kan förklaras av bara ett fåtal roterade dimensioner, kallade principalkomponenter, tyder det på att datasetet har hög redundans av information och därmed kan eventuellt vissa variabler vara överflödiga. I figur 15 finns en grafisk tolkning av resultaten från PCA på de beroende variablerna. Figuren är en tvådimensionell plot där axlarna spänns upp utav de två principalkomponenter som förklarar störst andel av den totala variansen i datasetet, alltså den optimala tvådimensionella komprimeringen av datan. De röda prickarna representerar hur de olika observationerna, alltså kommunerna, placerar sig i detta plan och de blåa linjerna visar i vilken riktning de olika variablernas varians förhåller sig till de två principalkomponenterna. Utan att gå in allt för detaljerat så kan spridningen av variablerna i figuren tolkas som att det inte finns några uppenbart starka korrelationer mellan variablerna. Dessutom visar texten vid axlarna i figuren att de två dominerande principalkomponenterna tillsammans utgör cirka 46 procent av datasetets totala varians, vilket är relativt lågt. Även detta tyder på att det inte finns någon extremt dominant latent faktor som förklarar variansen mellan variablerna.



Figur 18. Principalkomponentsanalys som visar förhållandet mellan de beroende variablerna.

Även Pearsons korrelationskoefficient kan användas för detta ändamål. Samtliga korrelationer mellan de beroende variablerna finns sammanställda i tabell 7. Det största absolutbeloppet bland korrelationerna är 0.65 och majoriteten lägre än 0.5 vilket tyder på att de olika variablerna innehåller relativt stor andel unik information och därav är relevanta att inkludera i undersökningen.

Tabell 5. Korrelationsmatris för de beroende variablerna som beskriver hur de korrelerar med varandra.

	Livsl. M	Livsl. K	Val	Sjuk	Lönegap	VAB	För.- Penn.	Gymn.- Frekv.	Arbete	Idrott	Bidrag	F. skola
<i>Livslängd Män</i>	1.00											
<i>Livslängd Kvinnor</i>	0.55	1.00										
<i>Riksdags- val</i>	0.60	0.37	1.00									
<i>Sjuk- penningtal</i>	-0.34	-0.40	-0.22	1.00								
<i>Lönegap</i>	-0.04	-0.08	0.01	-0.02	1.00							
<i>VAB</i>	-0.03	0.06	-0.05	-0.13	0.01	1.00						
<i>Föräldra- penning</i>	-0.29	-0.16	-0.40	0.08	-0.34	0.23	1.00					
<i>Gymnasie- frekvens</i>	0.28	0.22	0.33	-0.11	-0.07	0.08	-0.20	1.00				
<i>Förvärvs- arbete</i>	0.41	0.26	0.57	-0.03	0.01	-0.06	-0.46	0.37	1.00			
<i>Idrott</i>	0.20	0.18	0.18	-0.29	0.13	0.02	-0.26	0.23	0.05	1.00		
<i>Bidrags- hushåll</i>	-0.44	-0.34	-0.40	0.21	0.01	-0.10	0.36	-0.29	-0.52	-0.12	1.00	
<i>Förskola</i>	0.42	0.33	0.36	-0.24	0.18	0.11	-0.36	0.20	0.23	0.34	-0.35	1.00
<i>E.-gymn.- utb.</i>	0.48	0.47	0.43	-0.54	0.34	0.05	-0.51	0.18	0.17	0.49	-0.37	0.62

4.4 Bortfall av observationer

I tillgänglighetsdatan som inhämtats från WSP saknas data för tre kommuner, Heby, Knivsta och Nykvarn. Vilket troligtvis beror på att dessa kommuner tillkommit på senare år. Därför har dessa kommuner uteslutits ur samtliga variabler och är därmed utelämnade helt och hållet i undersökningen. Därmed kan datasettet ses som ett stickprov bestående av 287 av Sveriges 290 kommuner. Stickprovet utgör alltså knappt 99 procent av den totala populationen och därmed torde bortfallet av de tre kommunerna inte få någon betydande inverkan på de sammanvägda resultaten.

Datasettet som används i det här arbetet är av paneldataform vilket innebär att varje kommun har en observation för alla de åren som ingår i undersökningen. I de statistiska databaser som har sökts igenom har det för många av variablerna inte gått att få tag på data för alla kommunerna längre tillbaka i tiden. För att ändå få med så många beroende variabler som möjligt att undersöka valdes att ta med observationer från år 2000 till 2014.

4.5 Datasetets korrelationer

För att få en överblick över hur oberoende och beroende variabler korrelerar kan Pearsons korrelationskoefficient användas. Viktigt att ha i åtanke är att denna inte nödvändigtvis speglar något verkligt samband och framförallt inte säger någonting om eventuella kausala beroenden utan bara huruvida datamaterialet samvarierar. I tabell 8 ses korrelationskoefficienterna för alla beroende variabler mot de fyra oberoende variablerna. Det finns ett antal intressanta drag hos korrelationerna att poängtera. Tillgänglighet och medianinkomst har samma tecken, alltså positiva eller negativa koefficienter, för alla de beroende variablerna. Medianinkomst är den oberoende variabel som generellt tycks ha starkast korrelationer med de beroende variablerna. Bilinnehav och utländsk bakgrund har för de flesta beroende variabler låga korrelationer och för vissa beroende variabler skilda riktningar mot korrelationerna för övriga oberoende variabler vilket kan tyda på att de speglar en annan dynamik än tillgänglighet och medianinkomst. Korrelationerna är generellt sett relativt låga, bortsett från ett fåtal korrelationer med medianinkomsten, vilket tyder på att det inte finns några helt uppenbara givna linjära samband mellan de beroende och oberoende variablerna var för sig.

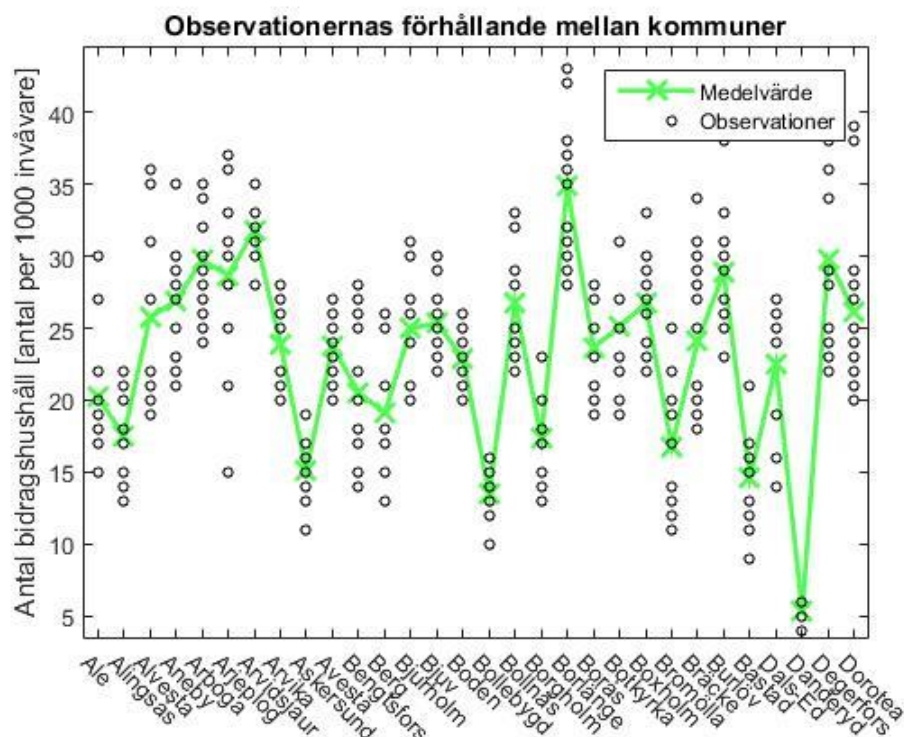
Tabell 6. Resultat för Pearsons korrelationstest mellan beroende och oberoende variabler.

Varabler	X_{Tillgänglighet}	X_{MedianI}	X_{Bil}	X_{UtländskB}
<i>Livslängd män</i>	0.3778	0.6812	-0.0202	0.1981
<i>Livslängd kvinnor</i>	0.3361	0.5986	-0.0510	0.2253
<i>Valdeltagande</i>	0.3575	0.6585	-0.0825	0.0638
<i>Sjukpenning</i>	-0.2304	-0.6952	-0.0440	0.3995
<i>Lönegap</i>	0.0153	0.3241	-0.0487	0.0785

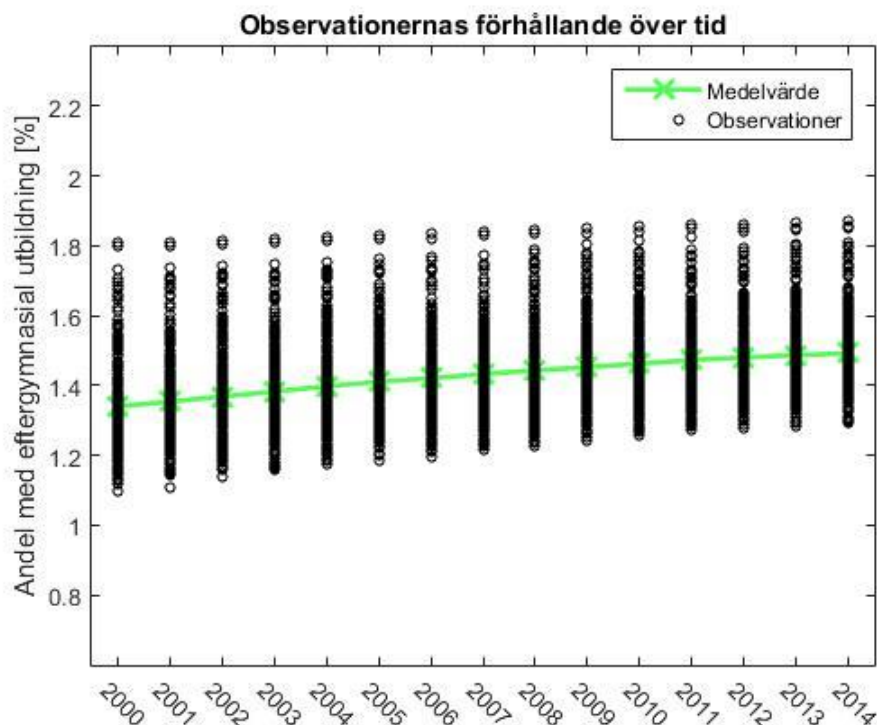
VAB	0.1190	0.1074	-0.0346	0.2127
Föräldrарpenning	-0.0878	-0.7744	-0.0932	-0.1403
Gymnaisefrekvens	-0.1223	-0.0113	0.1290	-0.3507
Förvärvsarbetande	0.1564	0.4207	0.1032	-0.2459
Idrott	0.4227	0.1306	-0.2772	0.1227
Bidragshushåll	-0.1267	-0.2044	-0.1087	0.1376
Förskolefrekvens	0.4289	0.8086	-0.1235	0.3730
Eftergymnaisal utbildning	0.6034	0.6598	-0.4172	0.3114

4.6 Observationernas heterogenitet

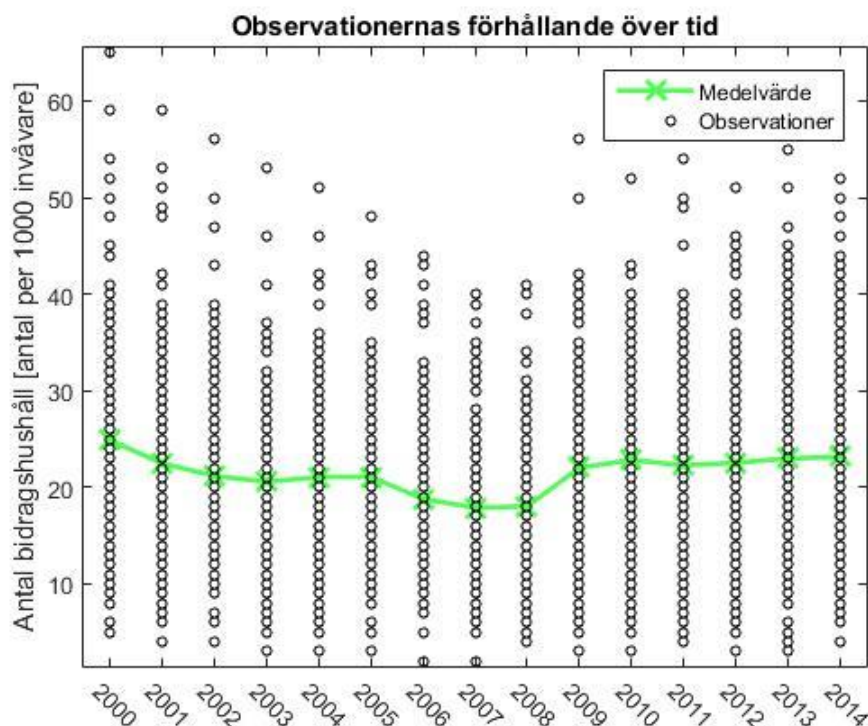
I avsnitt 3.7 diskuterades en del problem som kan uppstå vid paneldata regression. Ett av problemen var att det kan förekomma heterogenitet mellan de beroende variablerna både mellan kommuner och över tid. Detta kan leda till att det finns variationer i datan som inte fångas in av regressionsmodellen. Om så är fallet bör detta justeras för i regressionsmodellen. För att hantera detta så gjordes en kontroll av heterogenitet i de beroende observationerna i den här undersökningen. Med hjälp av ett statistiskt poolabilitetstest som gjordes upptäcktes att det förekommer heterogenitet i alla de beroende variablerna. Det går därmed inte att se alla observationer som att de kommer från samma fördelning varken över tvärsnitt eller tid. Observationerna är alltså heterogena både mellan kommuner och mellan år. Alltså behövs hänsyn tas till kommuninvarianta och tidsinvarianta observationer i paneldataregressionerna. En illustration av exempel på heterogenitet kan ses i figurerna 16-18. Dessa grafer finns för alla de beroende variablerna i appendix B. I figur 16 går det att se att det mellan kommuner finns stora skillnader för antalet bidragshushåll. Dessa skillnader i tvärsnitt är gemensamt för alla de beroende variablerna. Mellan åren finns det däremot olika struktur på heterogeniteten där vissa följer en linjär trend, se figur 17, medan andra varierar mellan åren utan tydligt mönster, se figur 18.



Figur 19. Grafen illustrerar den heterogenitet som existerar mellan olika kommuner när det gäller antalet bidragshushåll. I grafen visas endast antalet bidragshushåll bland de 30 första kommunerna i bokstavsordning.



Figur 20. Grafen illustrerar den heterogenitet som existerar över tid när det gäller andelen med eftergymnasial utbildning. Grafen tyder på att det finns en linjärt uppåtgående trend över tid.



Figur 21 Grafen illustrerar den heterogenitet som existerar över tid när det gäller antalet bidragshushåll. Grafen tyder på att antalet bidragshushåll varierar mellan åren men att det inte finns ett tydligt mönster.

4.7 Datastruktur

Tabell 9 visar strukturen på det dataset som använts i det här arbetet. Bland annat framgår vilka beteckningar som kommer användas för de olika variablerna och vilken enhet de är angivna i. Får att ge en uppfattning om vilka värden de olika variablerna antar visas även deras medelvärde och standardavvikelse.

Tabell 7. Strukturen på den data som använts i det här arbetet.

Variabelnamn	Typ	Beteckning	Enhet	Medelvärde	Standardavvikelse
Förväntad livslängd, män	Beroende	$Y_{\text{LivsL M}}$	År	78.5	1.47
Förväntad livslängd, kvinnor	Beroende	$Y_{\text{LivsL K}}$	År	82.7	1.12
Deltagande i senaste riksdagsval	Beroende	Y_{Val}	%	82.1	3.5
Sjukpenningtal	Beroende	Y_{Sjuk}	Dagar/ invånare	12	5.22
Lönegap mellan kvinnor och män	Beroende	$Y_{\text{Lönegap}}$	Kronor	-1380	875

VAB	Beroende	Y_{VAB}	%	63.6	3.34
Föräldrapenning	Beroende	$Y_{FöräldrarP}$	%	79.1	3.50
Gymnasiefrekvens	Beroende	$Y_{Gymnasie}$	%	93	1.98
Förvärsarbetande	Beroende	Y_{Arbete}	%	77.8	3.98
Idrott	Beroende	Y_{Idrott}	Tillfällen/ invånare	28.1	8.59
Bidragshushåll	Beroende	Y_{Bidrag}	Antal/1000 invånare	21.2	7.94
Förskolefrekvens	Beroende	$Y_{Föskola}$	%	61.6	7.49
Eftergymnasial utbildning	Beroende	$Y_{Utbildning}$	%	1.4	0.13
Tillgänglighet	Oberoende	$\text{Log}(X_{\text{Tillgäng.}})$	-	4.3	0.48
Medianinkomst	Oberoende	$\text{Log}(X_{\text{MedianInk}})$	1000 Kronor	2.3	0.054
Biltäthet	Oberoende	X_{Bil}	Bilar/1000 invånare	514	57.8
Andel med utländsk bakgrund	Oberoende	$\text{Log}(X_{\text{UtländskB}})$	Andel av total befolkning	-1	0.23

5. Resultat

5.1 Mått på transportinfrastruktur

En deluppgift i arbetet var att identifiera ett lämpligt mått på transportinfrastruktur för att kunna genomföra kvantitativa studier av sambanden mellan transportsystem och social nytta. Utifrån tidigare forskning om social nytta samt en undersökning av etablerade modeller för modellering av andra nyttor i samband med transportinfrastruktur framkom att tillgänglighet bör vara ett lämpligt mått för denna uppgift. Begreppet är väl etablerat i såväl forsknings- som trafikplaneringskretsar även om vissa skillnader i hur det definieras finns i olika användningsområden. I detta fall tycks ett potentialbaserat tillgänglighetsmått vara lämpligt utifrån dels argumentationen i van Wee (2013) dels eftersom att det används i ett liknande sammanhang i exempelvis Samlok-modellen.

Däremot är det svårt att utifrån tillgänglighetsmättet dra några slutsatser om hur transportsystem påverkar sociala förhållanden då måttet består av andra faktorer än endast tillgång till transportsystem. För att kunna särskilja transportsystemsåtgärder från övriga egenskaper av tillgänglighet behövs vetskap om transportsystemets funktionsförändring. Exempel på detta är tillgänglighetsförändring på grund av förändrad generaliserad kostnad (GK) vilket används i Samlok men som inte funnits tillgängligt för detta arbete.

5.2 Kvantifiering av sociala nyttor

För att möjliggöra en statistisk undersökning av social nytta behövdes en lämplig kvantifiering av begreppet göras. Utifrån tidigare forskning samt undersökning av sociala data kan det fastslås att social nytta bör studeras utifrån sina olika beståndsdelar. Inget generellt index bör skapas utan i stället ska de olika komponenter som utgör social nytta studeras var för sig. De komponenter som identifierades som beståndsdelar av social nytta är hälsa, socialt stöd, känsla av samhörighet, socialt nätverk, tilltro till institutioner, jämlikhet och allmän livskvalitet.

5.3 Tvärsnittsestimering – Between effects

Som diskuterats i metodavsnittet är en estimering utifrån between effects ett sätt att skatta modellparametrar utifrån ett tidsmedelvärde på de beroende och oberoende variablerna. Därmed utesluts tidsdimensionen helt ur skattningen. Regressionsmodellen som använts ser ut enligt följande för respektive socialt förhållande y:

$$y_{it} = \alpha + \beta_1 \log_{10}(x_{tillgänglighet_{it}}) + \beta_2 \log_{10}(x_{MedianI_{it}}) + \beta_3(x_{Bil_{it}}) + \beta_4 X_{UtländskB_{it}} + \varepsilon_{it} \quad (45)$$

Där parametrarna skattas med följande estimator

$$\hat{\beta}_{be} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \tilde{y} \quad (46)$$

Där $\tilde{y} = \bar{y}$ och $\tilde{X} = \bar{X}$.

Modellen beskriver alltså sambandet mellan kommunernas genomsnittliga värde för de observerade åren på de beroende och oberoende variablerna.

I tabell 10 presenteras regressionsresultaten för den förväntade livslängden för kvinnor och män, deltagande i riksdagsval, sjukpenningtalet samt lönegap mellan kvinnor och män. För förväntad livslängd för både kvinnor och män är samtliga oberoende variabler signifikanta. Tillgängligheten är positivt korrelerad med den förväntade livslängden med en signifikansnivå på 99 procent för både män och kvinnor. Detta tyder på att det existerar ett samband mellan ökad tillgänglighet och ökad förväntad livslängd.

För deltagande i riksdagsval fås signifikans för alla parametrar utom biltäthet. Tillgängligheten tycks här vara positivt korrelerad med deltagande i riksdagsval vilket betyder att det finns ett samband mellan ökad tillgänglighet och ökat deltagande i riksdagsval. Detta kan tolkas som att samhällets tillit till statliga institutioner är högre där tillgängligheten är hög.

Även för sjukpenningtalet är tillgängligheten statistiskt signifikant på en 99-procentig nivå. Modellen ger en negativ korrelation mellan och antalet dagar med sjukpenning. Färre dagar med sjukpenning tas alltså ut där tillgängligheten är hög vilket tyder på en positiv social nytta.

För lönegapet mellan kvinnor och män används som beroende variabel fås svag förklaringsgrad. Korrelationskoefficienten för tillgänglighet är statistiskt signifikant. Då den är negativ är sambandet mellan tillgänglighet och lönegapet mellan kvinnor och män negativt. Eftersom statistiken för löneskillnaden mellan kvinnor och män anges som kvinnors medianlön minus mäns medianlön innebär det att ökad tillgänglighet har ett samband med lägre lön för kvinnor i förhållande till män. Sambandet tyder alltså på social onytta då jämställdheten mellan kvinnor och män minskar med ökad tillgänglighet. Dock bör det poängteras att förklaringsgraden för denna modell endast är 0.10 vilket får anses som en väldigt svag modell.

*Tabell 8. Resultaten av between effects regression. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ***=99%, **=95% och *90%).*

	Y_{LivsL M}	Y_{LivsL K}	Y_{val}	Y_{Sjuk}	Y_{Lönegap}
<i>Log(X_{Tillgäng.})</i>	1.7*** (0.20)	0.83*** (0.16)	2.0*** (0.43)	-1.9*** (0.42)	-580*** (160)
<i>Log(X_{MedianInk})</i>	6.3*** (2.1)	6.1*** (1.6)	39*** (4.4)	-11** (4.3)	5900*** (1600)
<i>X_{Bil}</i>	0.0037*** (0.0014)	0.0018* (0.0010)	-0.0037 (0.0030)	-0.0060** (0.0029)	-4.1*** (1.1)
<i>Log(X_{UtländskB})</i>	-1.8***	-0.72***	-6.50***	-2.4***	-190

	(0.31)	(0.24)	(0.66)	(0.64)	(240)
<i>Konstant (alfa)</i>	53*** (4.4)	63*** (3.4)	-21** (9.4)	46*** (9.0)	-11000 *** (3400)
<i>Modellens förklaringsgrad</i>	0.43	0.33	0.57	0.37	0.10

Tabell 11 presenterar resultaten för regressionerna när andelen kvinnor som tar ut VAB respektive föräldrapenning, gymnasiefrekvens samt förvärvsarbete. För andelen dagar som kvinnor tar ut i VAB fås inget signifikant samband med tillgängligheten och dessutom har modellen en låg förklaringsgrad, $r^2 = 0.09$. Varken då gymnasiefrekvensen eller andel dagar med föräldrapenning som kvinnor tar ut används som beroende variabel fås ett signifikant samband med tillgängligheten. Resultaten för regressionen med andel förvärvsarbete som beroende variabel ger signifikans för tillgängligheten på en 90-procentig signifikansnivå då den justeras för BE. Modellen har en förklaringsgrad på 0.54. Resultatet pekar på att högre tillgänglighet sammanfaller med högre andel förvärvsarbete vilket är en positiv social nytta.

*Tabell 9. Resultaten av between effects regression. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parentes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ***=99%, **=95% och *90%).*

	Y_{VAB}	Y_{FöräldraP}	Y_{Gymnasie}	Y_{Arbete}
<i>Log(X_{Tillgäng.})</i>	-0.59 (0.60)	0.26 (0.33)	-0.049 (0.29)	1.1* (0.61)
<i>Log(X_{MedianInk})</i>	15** (6.1)	-25*** (3.4)	12*** (3.0)	69*** (6.2)
<i>X_{Bil}</i>	0.0067 (0.0041)	0.010*** (0.0023)	0.0026 (0.0020)	0.015*** (0.0041)
<i>Log(X_{UtländskB})</i>	4.0*** (0.92)	4.2*** (0.51)	-3.1*** (0.45)	-8.1*** (0.94)
<i>Konstant (alfa)</i>	32** (13)	140*** (7.2)	62*** (6.3)	-100*** (13)
<i>Modellens förklaringsgrad</i>	0.09	0.4	0.28	0.54

Resultaten för regressionerna där idrottsdeltagande, bidragstagande, förskolefrekvens och eftergymnasial utbildning har använts som beroende variabler presenteras i tabell 12. Idrottsdeltagande ger ett statistiskt signifikant samband med tillgänglighet och en förklaringsgrad för modellen på 0.22. Alltså sammanfaller högt idrottsdeltagande med hög tillgänglighet vilket är en positiv social nytta då det kan leda till förhöjd livskvalitet och känsla av samhörighet. För andelen bidragstagande får tillgängligheten ett

signifikant samband på en 90-procentig signifikansnivå. Regressionskoefficienten är negativ vilket tyder på att det finns ett samband mellan ökad tillgänglighet och ett minskat bidragstagande. Detta får ses som en positiv social nytta då det tyder på minskad ekonomisk utsatthet. Regressionsresultaten för förskolefrekvens tyder på signifikanta samband mellan tillgänglighet och andelen barn inskrivna i förskola och skolbarnsomsorg. Modellens förklaringsgrad är 0.56. Högre tillgänglighet sammanfaller alltså med högre andel barn inskrivna i förskola och skolbarnsomsorg vilket tolkas som en positiv social nytta i det här arbetet. För andelen med eftergymnasial utbildning fås en hög förklaringsgrad på 0.58 och en signifikant korrelation med tillgängligheten. Den positiva koefficienten tyder på att det existerar ett samband mellan ökad tillgänglighet och ökad andel med eftergymnasial utbildning. Detta ses som en positiv social nytta då det tycks leda till bättre psykisk hälsa och ökat socialt kapital.

*Tabell 10. Resultaten av between effects regression. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ****=99%, **=95% och *90%).*

	YIdrott	YBidrag	YFörskola	YUtbildning
<i>Log(X_{Tillgäng.})</i>	6.7*** (1.7)	-2.7* (1.4)	5.8*** (0.84)	6.0*** (1.5)
<i>Log(X_{MedianInk})</i>	48*** (17)	-110*** (14)	44*** (8.6)	95*** (15)
<i>X_{Bil}</i>	0.014 (0.011)	-0.060*** (0.0010)	-0.0081 (0.0058)	-0.045*** (0.0010)
<i>Log(X_{UtländskB})</i>	-1.9 (2.6)	4.2* (2.2)	-4.2*** (1.3)	-9.0*** (2.2)
<i>Konstant (alfa)</i>	-120*** (36)	310*** (30)	-65*** (18)	-200*** (31)
<i>Modellens förklaringsgrad</i>	0.24	0.33	0.56	0.56

Sammanfattningsvis så får regressionsmodellerna som använder tvärsnittsestimering signifikans för tillgängligheten som oberoende variabel för:

- Förväntad livslängd för kvinnor och män
- Deltagande i senaste riksdagsval
- Sjukpenningtalet
- Lönegapet mellan kvinnor och män
- Andel förvärvsarbetande
- Antal deltagartillfällen i idrottsföreningar per person
- Andel bidragshushåll
- Andel barn inskrivna i förskola och skolbarnsomsorg
- Andel med eftergymnasial utbildning

Av dessa så fås bäst modellanpassning när deltagande i riksdagsval, andel barn inskrivna i förskole och skolbarnsomsorg eller andel med eftergymnasial utbildning används som oberoende variabel.

För samtliga ovanstående variabler, förutom lönegapet mellan kvinnor och män, är korrelationen mellan tillgänglighet och den beroende variabeln att betrakta som en social nytta utifrån tolkningen av vad som utgör en social nytta för respektive beroende variabel i kapitel 4.

För lönegapet mellan kvinnor och män är korrelationen dock negativ vilket indikerar att hög tillgänglighet samvarierar med stora löneskillnader mellan kvinnor och män, där kvinnor tjänar mindre. Detta är utifrån tolkningen i kapitel 4 att betrakta som en social onytta.

T-test av residualerna tyder på att för samtliga modeller estimerade med between effects är residualernas väntevärde inte signifikant nollskilda. Alltså kan de antas ha väntevärde noll.

5.4 Estimering med kommunspezifika effekter

Vid genomförande av Mundlak-test för att avgöra om de kommunspezifika effekterna kan antas vara slumpmässiga visade det sig att för samtliga beroende variabler kan förekomsten av slumpmässiga kommunspezifika effekter förkastas. Alltså behövs fixa effekter användas i estimeringen. Därför presenteras inga resultat för estimeringar med slumpmässiga effekter.

Presentera vilka samband som fås av estimeringen.

I detta avsnitt återfinns resultaten från körningar utifrån när fixed effects estimator används. Regressionsmodellen som används ser ut enligt följande

$$y_{it} = \alpha + \beta_1 \log_{10}(x_{tillgänglighet_{it}}) + \beta_2 \log_{10}(x_{MedianI_{it}}) + \beta_3 (x_{Bil_{it}}) + \beta_4 \log_{10}(x_{UtländskB_{it}}) + Z_i + \varepsilon_{it} \quad (47)$$

Och skattningar på parametrarna fås med estimatorn:

$$\hat{\beta}_{fe} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \tilde{y} \quad (48)$$

där $\tilde{y} = y - \bar{y}$ och $\tilde{X} = X - \bar{X}$.

Termen Z_i innehåller alla de kommunspezifika effekterna vilka talar om hur varje kommun förhåller sig till riksgenomsnittet. β -parametrarna motsvarar då de genomsnittliga förändringskoefficienterna då kommunspezifika effekter har bortjusterats. Alltså tar modellen hänsyn till skillnader mellan de olika kommunerna.

Resultaten för förväntad livslängd för både män och kvinnor visar det enbart medianinkomsten och andel med utländsk bakgrund som är signifikanta oberoende variabler för modellen. Tillgängligheten kan alltså inte sägas ha någon signifikant korrelation med denna estimeringsmodell.

När det gäller deltagande i riksdagsval så är det endast tillgängligheten som inte är signifikant av de oberoende variablerna. Alltså kan inget samband mellan tillgänglighet och deltagande i riksdagsval när det justeras för skillnader mellan kommuner. Inte heller vid regressionen för antalet dagar med sjukpenning per invånare är tillgängligheten signifikant. För lönegapet mellan kvinnor och män används som beroende variabel fås svag förklaringsgrad och ingen signifikans fås för tillgängligheten.

*Tabell 11. Resultaten av regressioner justerade för kommunspecifika effekter Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ***=99%, **=95% och *90%).*

	Y_{LivsL M}	Y_{LivsL K}	Y_{Val}	Y_{Sjuk}	Y_{Lönegap}
<i>Log(X_{Tillgäng.})</i>	0.21 (0.88)	-0.63 (1.1)	-1.9 (1.38)	-2.2 (3.5)	280 (610)
<i>Log(X_{MedianInk})</i>	16*** (0.78)	9.0*** (0.90)	18*** (0.80)	-52*** (2.6)	4900*** (530)
<i>X_{Bil}</i>	0.00038 (0.0017)	-0.00067 (0.0015)	0.013*** (0.0018)	-0.016*** (0.0054)	-0.89 (1.0)
<i>Log(X_{UtländskB})</i>	1.1** (0.56)	1.7** (0.72)	6.9*** (0.55)	-17*** (2.1)	420 (380)
<i>Konstant (alpha)</i>	42*** (2.0)	67*** (2.0)	50*** (4.5)	130*** (17)	-13000*** (3000)
<i>Modellens förklaringsgrad</i>	0.67	0.46	0.56	0.64	0.25

I tabell 14 presenteras resultaten då andelen dagar med VAB respektive föräldrapenning som kvinnor tar ut, gymnasiefrekvens och andel förvärvsarbetande har använts som beroende variabler. När andelen dagar som kvinnor tar ut VAB sätts som beroende variabel fås en svag förklaringsgrad för modellen. Däremot så är tillgängligheten signifikant och då positivt korrelerad vilket tyder på att ökad tillgänglighet sammanfaller med ökad andel av VAB-dagar som kvinnor tar ut. Modellen där andelen av dagar antal dagar med föräldrapenning som kvinnor tar ut ger ingen signifikans för tillgängligheten. För gymnasiefrekvensen är tillgängligheten signifikant men då modellens förklaringsgrad endast är 0.068 är inte dessa siffror speciellt tillförlitliga. Tillgänglighetens negativa koefficient tyder dock på att gymnasiefrekvensen minskar med ökad tillgänglighet. Mot andelen förvärvsarbetande är tillgängligheten inte signifikant.

Tabell 12. Resultaten av regressioner justerade för kommunspecifika effekter. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ***=99%, **=95% och *90%).

	Y _{VAB}	Y _{FöräldraP}	Y _{Gymnasie}	Y _{Arbete}
$\text{Log}(X_{\text{Tillgäng.}})$	6.8*** (2.6)	0.12 (2.2)	-3.0** (1.4)	1.5 (2.4)
$\text{Log}(X_{\text{MedianInk}})$	-12*** (2.1)	-78*** (2.0)	9.7*** (1.5)	17*** (2.5)
X_{Bil}	0.028*** (0.0054)	0.021*** (0.0056)	-0.012*** (0.0032)	0.015** (0.0060)
$\text{Log}(X_{\text{UtländskB}})$	5.7*** (1.5)	3.1** (1.5)	-6.6*** (1.2)	-5.3** (2.1)
Konstant (alpha)	54*** (13)	250*** (9.0)	83*** (7.7)	18 (12)
Modellens förklaringsgrad	0.06	0.81	0.068	0.19

Tabell 15 visar resultaten för regressioner med Idrottsdeltagande får ett signifikant samband med tillgängligheten på en 99-procentig signifikansnivå och modellens förklaringsgrad är 0.24. Tillgänglighet tenderar alltså att vara hög där antalet deltagartillfällen i en idrottsförening är högt. Tillgängligheten tycks därför vara kopplad till positiv social nytta i detta avseende. Regressionsmodellen ger inget signifikant samband mellan tillgänglighet och andelen bidragstagare i kommunerna. Modellen har en förklaringsgrad på 0.33. Regressionsresultaten för förskolefrekvens tyder på signifikanta samband mellan tillgänglighet och andelen barn inskrivna i förskola och skolbarnsomsorg. Fler barn är alltså inskrivna i förskola och skolbarnsomsorg där tillgängligheten är hög. Detta ses med försiktighet, som angett i avsnitt 4.3.10 som en positiv nytta. Modellens förklaringsgrad är 0.78 vilket får anses som en bra modell. Regressionen visar ingen signifikans mellan tillgänglighet och andelen med eftergymnasial utbildning. Däremot har modellen en hög förklaringsgrad på 0.92.

Tabell 13. Resultaten av regressioner justerade för kommunspecifika effekter. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ***=99%, **=95% och *90%).

	Y _{Idrott}	Y _{Bidrag}	Y _{Förskola}	Y _{Utbildning}
$\text{Log}(X_{\text{Tillgäng.}})$	18*** (6.6)	-5.0 (5.6)	11*** (4.2)	-0.095 (1.1)
$\text{Log}(X_{\text{MedianInk}})$	-39***	-43***	110***	61***

	(6.0)	(5.3)	(4.2)	(1.7)
X_{Bil}	0.020** (0.0085)	0.020* (0.010)	0.018** (0.0077)	-0.012*** (0.0036)
$Log(X_{UtländskB})$	-4.0 (4.4)	28*** (4.2)	1.6 (2.7)	0.098 (0.84)
<i>Konstant (alpha)</i>	28 (30)	160*** (30)	-250*** (23)	-110*** (6.3)
<i>Modellens förklaringsgrad</i>	0.22	0.098	0.78	0.94

De beroende som ger signifikans för tillgänglighet som beroende variabel när regressionsmodellen justerar för kommunspecifika effekter är:

- Andelen dagar som kvinnor tar ut VAB
- Andel ungdomar inskrivna i gymnasieskola
- Deltagartillfällen i idrottsföreningar
- Andel barn inskrivna i förskola och skolbarnsomsorg

Av dessa så är det endast modellen för andelen barn inskrivna i förskola och skolbarnsomsorg som får en bra förklaringsgrad.

Gällande deltagartillfällen i idrottsföreningar och andelen barn inskrivna i förskola och skolbarnsomsorg är korrelation mellan tillgänglighet och den beroende variabeln att betrakta som en social nytta utifrån tolkningen av vad som utgör en social nytta för respektive beroende variabel i kapitel 4.

För andelen dagar som kvinnor tar ut VAB och andel ungdomar inskrivna i gymnasieskola är korrelationen av en sådan form att hög tillgänglighet samvarierar med det som utgör social onytta för dessa indikatorer utifrån tolkningen i kapitel 4.

T-test av residualerna tyder på att för samtliga modeller estimerade med kommunspecifika fixa effekter är residualernas väntevärde ej signifikant nollskilda. Alltså kan de antas ha väntevärde noll.

5.5 Estimering med kommunspecifika och tidsspecifika effekter - Fixed effects

Vid genomförande av Mundlak-test för att avgöra om de kommunspecifika effekterna kan antas vara slumpmässiga visade det sig att för samtliga beroende variabler kan förekomsten av slumpmässiga kommunspecifika effekter förkastas. Alltså behövs fixa effekter användas i estimeringen. Därför presenteras inga resultat för estimeringar med slumpmässiga effekter.

De resultat som presenteras nedan är baserade på följande modell

$$y_{it} = \alpha + \beta_1 \log_{10}(x_{tillgänglighet_{it}}) + \beta_2 \log_{10}(x_{MedianI_{it}}) + \beta_3 (x_{Bil_{it}})$$

$$+\beta_4 \log_{10}(X_{UtländskB_{it}}) + Z_i + S_t + \varepsilon_{it} \quad (49)$$

I tabell 16 presenteras regressionsresultaten för modellerna där förväntad livslängd för män och kvinnor, valdeltagande, sjukpenningtalet samt lönegapet mellan kvinnor och män har använts som beroende variabel. Modellen visade inget signifikant samband för tillgängligheten med varken valdeltagande eller förväntad livslängd på både män och kvinnor. Då sjukpenningtalet sätts som beroende variabel fås ett samband med tillgängligheten på en 99-procentig signifikansnivå. Förklaringsgraden för modellen är 0.91 vilket tyder på en väl anpassad modell mot verkliga värden. Den negativa koefficienten för tillgänglighet tyder på att sjukpenningtalet sjunker med ökad tillgänglighet vilket tolkas som en positiv social nytta. Även för lönegapet mellan kvinnor och män fås ett samband med tillgängligheten men på en 95-procentig signifikansnivå. Modellens förklaringsgrad på 0.77 får också ses som hög. Då tillgänglighetskoefficienten är positiv betyder det att kvinnors lön tenderar att vara högre i förhållande till män där tillgängligheten också är hög. Detta ses som positivt för jämställdheten i de kommuner där kvinnor redan tjänar mindre än män men kan vara negativt om kvinnor redan tjänar mer än männen.

*Tabell 14. Resultaten av regressioner justerade för både tid- och kommunspecifika effekter. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ***=99%, **=95% och *90%).*

	Y _{LivsL M}	Y _{LivsL K}	Y _{Val}	Y _{Sjuk}	Y _{Lönegap}
<i>Log(X_{Tillgäng.})</i>	0.44 (0.36)	-0.42 (0.36)	0.28 (0.44)	-3.0*** (0.99)	570** (250)
<i>Log(X_{MedianInk})</i>	5.9*** (1.4)	-2.3 (1.4)	1.9 (1.7)	38*** (3.8)	-910 (940)
<i>X_{Bil}</i>	-0.00090 (0.00070)	-0.0019*** (0.00077)	0.0030*** (0.00091)	-0.016*** (0.0021)	-17*** (0.52)
<i>Log(X_{UtländskB})</i>	0.085 (0.24)	0.72*** (0.25)	0.090*** (0.0021)	-11*** (0.68)	-670*** (170)
<i>Konstant (alfa)</i>	63*** (3.5)	91*** (3.5)	74*** (4.3)	-63*** (9.7)	-1500 (2400)
<i>Modellens förklaringsgrad</i>	0.85	0.72	0.96	0.91	0.77

Resultaten för regressionerna då andelen dagar som kvinnor tar ut av VAB respektive föräldrapenning, andel inskrivna på gymnasium och andelen förvärvsarbetande har använts som beroende variabel presenteras i tabell 17. Andelen dagar som kvinnor tar ut VAB ger en positiv signifikant korrelation med tillgängligheten. Alltså samvarierar högre andel dagar som kvinnor tar ut VAB med hög tillgänglighet. Modellens

förklaringsgrad är 0.63 vilket är en relativt bra modellanpassning men sämre än de flesta andra regressionerna gjorda med både tid- och kommunspecifika effekter. Den positiva korrelationen kan tyda på en social onytta ur ett jämställdhetsperspektiv eftersom det innebär samvariation mellan hög tillgänglighet och hög andel av VAB som tas ut av kvinnor. Även andelen inskrivna på gymnasiet får signifikant tillgänglighetskoefficient. Här är dock koefficienten negativ vilket betyder att andelen inskrivna på gymnasium minskar med ökad tillgänglighet vilket tyder på samvariation för negativ social nytta med ökad tillgänglighet i detta avseende. Men modellen har en förklaringsgrad på endast 0.54 vilket inte är speciellt högt. Varken andelen dagar som kvinnor tar ut föräldrapenning eller andelen förvärvsarbetande får signifikans för tillgängligheten då de används som beroende variabel.

*Tabell 15. Resultaten av regressioner justerade för både tid- och kommunspecifika effekter. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ***=99%, **=95% och *90%).*

	Y _{VAB}	Y _{FöräldraP}	Y _{Gymnasie}	Y _{Arbete}
<i>Log(X_{Tillgäng.})</i>	6.9*** (1.2)	0.76 (0.92)	-2.7*** (0.82)	0.25 (0.75)
<i>Log(X_{MedianInk})</i>	-25*** (4.7)	-50*** (3.5)	14*** (3.1)	90 (2.7)
<i>X_{Bil}</i>	0.031*** (0.0026)	0.021*** (0.0020)	-0.013*** (0.0017)	0.021 (0.0015)
<i>Log(X_{UtländskB})</i>	5.1*** (0.86)	2.0*** (0.63)	7.1*** (0.56)	1.6 (0.50)
<i>Konstant (alfa)</i>	81*** (12)	180*** (9.0)	71*** (8.0)	-140*** (7.0)
<i>Modellens förklaringsgrad</i>	0.63	0.87	0.54	0.93

I tabell 18 presenteras resultaten för antal deltagartillfällen i idrottsföreningar, andel barn inskrivna i förskola och skolbarnsomsorg samt andel med eftergymnasial utbildning. Med idrottsdeltagande som beroende variabel fås positiv korrelation på en 99-procentig signifikansnivå och en förklaringsgrad på 0.88 för modellen. Detta är en positiv social nytta då ökad tillgänglighet samvarierar med ökat idrottsdeltagande som anses förbättra olika sociala aspekter. Även andelen barn inskrivna i förskola och skolbarnsomsorg har positiv korrelation på 99-procents signifikansnivå. Tillgängligheten tenderar att vara hög där även andelen barn inskrivna i förskola och skolbarnsomsorg är hög. Förklaringsgraden på 0.87 får ses som en bra modellanpassning. Då förskola och skolbarnsomsorg med försiktighet kan antas leda till förbättrade sociala förhållanden kan transportsystem kopplas till en positiv social nytta i

detta avseende. Andelen med eftergymnasial utbildning gav inget signifikant samband med tillgängligheten. Paneldatan för andelen bidragshushåll var obalanserad vilket gjorde att vi inte kunde köra denna regressionsmodell på den datan.

*Tabell 16. Resultaten av regressioner justerade för både tid- och kommunspecifika effekter. Siffrorna som redovisas är betaparametrarna för de olika oberoende variablerna och inom parantes står standardavvikelsen för de signifikanta betaparametrarna. (Stjärnorna visar resultatens signifikansnivå där ****=99%, **=95% och *90%).*

	Y Idrott	Y Bidrag	Y Förskola	Y Utbildning
$Log(X_{Tillgäng.})$	17*** (2.4)	Obalanserad panel. Ej möjlig att estimeras.	11*** (1.7)	0.32 (0.41)
$Log(X_{MedianInk})$	-7.8 (7.9)		52*** (6.6)	40*** (1.6)
X_{Bil}	0.021*** (0.0043)		0.020*** (0.0036)	-0.016*** (0.00087)
$Log(X_{UtländskB})$	-0.054 (1.5)		1.4 (1.2)	-1.9*** (0.28)
Konstant (alfa)	-37** (21)		-120*** (17)	-58*** (4.1)
Modellens förklaringsgrad	0.88		0.87	0.99

De beroende variabler som tillgängligheten har signifikans för då modellen justerar för både tid- och kommunspecifika effekter är:

- Sjukpenningtalet
- Lönegap mellan kvinnor och män
- Andel dagar som kvinnor tar ut VAB
- Andel ungdomar inskrivna i gymnasieskola
- Antal deltagartillfällen i idrottsföreningar
- Andel barn inskrivna i förskola och skolbarnsomsorg

Modellerna får högst förklaringsgrad då sjukpenningtalet, idrottsdeltagande eller förskola och skolbarnsomsorg används som beroende variabel.

Angående tolkningen om korrelationen mellan tillgänglighet och de sociala indikatorerna utgör sociala nyttor eller onyttor går det att utlösa följande. För sjukpenningtalet, lönegap mellan kvinnor och män, antal deltagartillfällen i idrottsföreningar samt andel barn inskrivna i förskola och skolbarnsomsorg är korrelationen av en sådan form att den är att betrakta som en social nytta utifrån tolkningen i kapitel 4.

Däremot gäller för andel dagar som kvinnor tar ut av VAB, andel ungdomar inskrivna i gymnasieskola att korrelationen är att tolka som social onytta.

deltagartillfällen i idrottsföreningar och andelen barn inskrivna i förskola och skolbarnomsorg är korrelation mellan tillgänglighet och den beroende variabeln att betrakta som en social nytta utifrån tolkningen av vad som utgör en social nytta för respektive beroende variabel i kapitel 4.

För andelen dagar som kvinnor tar ut VAB och andel ungdomar inskrivna i gymnasieskola är korrelationen av en sådan form att hög tillgänglighet samvarierar med det som utgör social onytta för dessa indikatorer utifrån tolkningen i kapitel 4.

T-test av residualerna tyder på att för samtliga modeller estimerade med tid- och kommunspecifika fixa effekter är residualernas väntevärde ej signifikant nollskilda. Alltså kan de antas ha väntevärde noll.

5.6 Sammanfattande resultat av paneldataanalys

För samtliga beroende variabler, förutom andel dagar som kvinnor tar ut av föräldrapenning, finns signifikanta statistiska samband med tillgänglighet för minst en typ av estimeringsmetod. Att statistiska samband existerar mellan beroende och oberoende variablerna är dock i sig inget bevis för att de är praktiskt lämpliga att använda eller att ett kausalt orsakssamband råder dem emellan.

I tabell 19 sammanfattas för vilka av de beroende variablerna som tillgängligheten är en signifikant oberoende variabel för de olika modellerna som använts i arbetet.

Livslängden för både män och kvinnor, deltagande i riksdagsval, förvärvsarbetande, antalet bidragshushåll, samt eftergymnasial utbildning har signifikant skattad parameter för tillgänglighet enbart för between effects estimeringen. Detta kan tolkas som att tillgängligheten är signifikant för att förklara skillnader mellan kommunerna medan den inte kan förklara förändringar av dessa beroende variabler över tid. Det motsatta gäller för andelen kvinnor som tar ut VAB samt gymnasiefrekvensen där den skattade parametern för tillgänglighet enbart är signifikant när estimeringen görs med fixa effekter och inte signifikant för between effects. Sjukpenningtalet, lönegap mellan kvinnor och män, idrottsdeltagande och förskolefrekvens har signifikanta skattningar för tillgänglighet vid såväl between effects som vid justering för kommun- och/eller kommun- och tidspecifika fixa effekter. Detta tyder på att såväl skillnader mellan kommunerna i tillgänglighet och skillnader i tillgänglighet över tid har ett samband med dessa beroende variabler.

Andelen kvinnor som tar ut föräldrapenning har ingen signifikant skattning för parametern för tillgänglighet och därmed har, med dessa modeller, inget statistiskt samband dem emellan kunnat observeras.

Tabell 17. Sammanfattning av signifikansnivåerna för tillgängligheten som oberoende variabel när de olika beroende variablerna ansätts i de tre olika modellerna. Stjärnorna visar signifikansnivåerna, där ***=99%, **=95% och *90%).(+) eller (-)indikerar om korrelationen har en sådan form att den är att betrakta som en social nytta respektive onytta utifrån tolkningarna i kapitel 4.

	Between effects	Justering för kommunspezifika effekter	Justering för tid- och kommunspezifika effekter
$Y_{LivsL\ M}$	*** (+)		
$Y_{LivsL\ K}$	*** (+)		
Y_{Val}	*** (+)		
Y_{Sjuk}	*** (+)		*** (+)
$Y_{Lönegap}$	*** (-)		** (+)
Y_{VAB}		*** (-)	*** (-)
$Y_{FöräldrarP}$			
$Y_{Gymnasie}$		** (-)	*** (-)
Y_{Arbete}	* (+)		
Y_{Idrott}	*** (+)	*** (+)	*** (+)
Y_{Bidrag}	* (+)	-	-
$Y_{Föskola}$	*** (+)	*** (+)	*** (+)
$Y_{Utbildning}$	*** (+)		

5.7 Osäkerhet och felkällor

5.7.1 Tillgänglighetsdata

En stor brist med det här arbetet är att det tillgänglighetsmått som används som oberoende variabel endast erhöles för tre olika år; 1985, 1997 och 2010. Dessutom var inte dessa tillgänglighetsmått inte justerade för konsumentprisutveckling vilket gör att måtten för de olika åren inte är direkt jämförbara. Detta justerades istället manuellt genom att multiplicera tillgänglighetsmåtten för 1997 och 2010 med en faktor så att medelvärdet för båda åren blev samma som för 1985. För åren emellan de år som tillgänglighetsmåtten erhöles för har en linjär förändring antagits. Båda dessa justeringar och antaganden borde göra att tidsserieinformationen för tillgänglighetsmåttet inte fullt ut beskriver den verkliga förändringen och utelämnar information. Eftersom paneldataregressionen som justeras för fixed effects använder sig av tidsserieinformationen borde den utelämnade tidsserieinformationen påverka tillgänglighetens inverkan i sambandsanalysen. Detta tyder även resultaten på då tillgängligheten är signifikant för fler av de beroende variablerna då between effects används än när regressionerna justeras med fixed effects. Eftersom tillgänglighetsdatans genomsnitt på riksnivå är statiskt över tid, på grund av medelvärdesjusteringen som gjordes på grund av bristerna i tillgänglighetsdatan, så blir möjligheterna till identifikation av effekter av tillgänglighet med hjälp av de estimatorer som används för metoderna med fixa effekter begränsad. Detta då dessa bygger på att utnyttja

tidsseriedimensionen av datan vilket ju dels har medelvärdesjusterats, dels har linjärapproximerats, vilket givetvis är potentiella källor till störningar av skattningar. Medelvärdesjusteringen innebär dock inte att varje kommuns tillgänglighetsförändring övertid är konstant utan enbart att det generella riksmedelvärdet är konstant. Resultaten för skattningar med between effects estimering däremot påverkas mindre av detta eftersom att förhållandet i tillgänglighet mellan kommunerna inte påverkas.

Undersökningen av korrelationen mellan de beroende variablerna, som presenterades i kapitel 4 så framgick att korrelationen mellan tillgänglighet och befolkningstäthet var mycket hög. Detta tyder på att de samband som i det här arbetet tillskrivs tillgängligheten kan till hög grad också bero på befolkningstäthet. Givetvis beror sambandet mellan dessa två variabler mycket på att befolkningstäthet är en underliggande faktor av tillgänglighetsmättet.

5.7.2 Endogenitet

De endogenitetsproblem som diskuterades i metodkapitlet är viktiga att poängtera vid tolkning av resultaten. För de olika beroende variablerna kan olika former av endogenitetsproblem förekomma och störa skattningarna.

Ett tydligt exempel är eftergymnasial utbildning som i modellen ansätts vara den beroende variabeln. I detta fall kan flera källor till endogena störningar vara aktuella. Först och främst kan det vara så att kommuner med universitet eller andra institutioner för högre utbildning har välutbildad befolkning samt ofta har goda kommunikationer och har större orter och har därmed hög tillgänglighet. Det är också rimligt att anta att det kan vara av politiskt intresse att förbättra tillgängligheten i dessa kommuner just för att folk ska kunna nå utbildningsmöjligheterna. Dessutom är det rimligt att människor som är välutbildade söker sig till kommuner med hög tillgänglighet till arbetsmarknaden. Alltså är det troligt att modellen för eftergymnasial utbildning lider av motsvarande problem som Combes et al. (2010 b) kallar för "endogenous quantity of labour" och "endogenous quality of labour". En annan form av endogenitetsproblem som kan tänkas förekomma för eftergymnasial utbildning är att det inte behöver vara ett enkelriktat samband mellan medianinkomst, som används som oberoende variabel, och andelen med eftergymnasial utbildning. Snarare är det högst troligt att kommuner med en hög andel med eftergymnasial utbildning har höga inkomstsnivåer just på grund av sin välutbildade population. Här finns alltså en stor risk att en kausal loop mellan oberoende och beroende variabler uppstår vilket leder till felaktiga skattningar.

En annan beroende variabel där den skattade parametern för tillgänglighet uppvisar stark signifikans är deltagandetillfällen i för ungdomar i idrottsföreningar. Korrelationen är positiv vilket innebär att hög tillgänglighet samvarierar med många deltagandetillfällen. I detta fall finns det inte någon uppenbar risk för motriktad kausalitet, alltså att högt idrottsdeltagande ökar tillgängligheten. Däremot kan det finnas en i modellen utelämnad variabel som förklarar idrottsdeltagandet som stör skattningen.

Liknande endogenitetsproblem uppstår även för de övriga modellerna och därför är det inte utifrån det data och de kvantitativa metoder detta arbete använt möjligt att fullt besvara frågan om exakt vilka sociala effekter förbättrad transportinfrastruktur leder till. På grund av endogeniteten i modellerna kan de statistiska sambanden inte användas för att förutspå förändringen av den beroende variabeln med en given förändring av tillgängligheten (eller någon annan av de oberoende variablerna) med övriga oberoende variabler fixa. Detta eftersom att det inte går att säkerställa ett giltigt kausalt orsakssamband mellan oberoende och beroende variabler. Endogenitetsproblemen innebär sammanfattningsvis att resultaten visar formen för eventuell samvarians mellan tillgänglighet och de sociala förhållandena men ingenting om transportsystemets inverkan eller vad som orsakar vad.

5.7.3 Generaliserade modeller

Den regressionsmodell som ansätts i det här arbetet är utav en generell karaktär. Samma modell har använts för att undersöka alla de olika beroende variablerna trots att ett rimligt antagande kan vara att de olika beroende variablerna beror på olika oberoende variabler. För att få lämpligare modeller för de beroende variablerna kan det därför behövas läggas till eller tas bort oberoende variabler. Syftet med detta arbete är dock inte att utveckla en modell utan snarare att utreda vilka av de beroende variablerna kan vara intressant att utveckla en modell för.

5.7.4 Hantering av stora datastrukturer

Regressionerna i det här arbetet är baserade på en stor datamängd. Den totala datan som sammanställts består av 13 beroende variabler och sex oberoende variabler, alla dessa variabler består i sig av observationer för 287 kommuner över 15 år. All data utom för tillgänglighet har insamlats från statistiska onlinedatabaser och sedan importerats till Excell och därefter sorterats för att få på det på paneldatastruktur. Då det finns många potentiella felkällor i denna process har en försiktighet vidtagits vid utförandet men detta ger ingen fullständig garanti för att enstaka fel kan förekomma i datamaterialet.

5.7.5 Offentlig statistik

Då datan är insamlad från offentliga databaser är det svårt att få insikt i statistikens tillförlitlighet. Då det handlar om offentliga organisationer som tillhandahåller datan är ändå utgångspunkten att den är tillförlitlig.

6. Diskussion

I det här kapitlet förs en diskussion om hur de resultat som framkommit kan relateras till tidigare forskning och vilken betydelse de kan ha i praktiken. Till sist framkommer även förslag på förbättringar och intressanta uppslag för framtida forskning inom området.

6.1 Relation till tidigare forskning

Först och främst bör det nämnas att det enligt författarnas vetskap i skrivande stund aldrig tidigare gjorts en liknande studie om samband mellan tillgänglighet och sociala faktorer. Därför är detta arbete främst en ansats för att studera hur sambanden mellan tillgänglighet och sociala förhållanden ser ut och hur starka de är. Resultaten bör därmed främst ligga till grund för vidare forskning inom området.

Resultaten av detta examensarbete ger en fingervisning om att den generella diskussion som förs i kvalitativ forskning om transportinfrastrukturens sociala konsekvenser i vissa avseenden är möjliga att identifiera med statistiska metoder på paneldata. Hur de signifikanta parametrar för tillgänglighet som skattats fram i detta arbete kopplar till de orsakssamband den kvalitativa forskningen på området framhåller är dock svårt att förhålla sig till. De studier som gjorts på tillgänglighetens påverkan på ekonomisk tillväxt framhåller vikten av att vara mycket försiktig med att dra slutsatser kring sambandens kausalitet (se Börjesson et al., 2013). Detsamma gäller för denna studie. De metoder som använts kan ge, och ger för vissa beroende variabler, en fingervisning om att variationen i tillgänglighet mellan kommuner samvarierar med de beroende variablerna och att det tycks finnas ett samband mellan dem. Däremot innebär detta inte på något sätt att det, utifrån dessa resultat, går att dra slutsatsen att tillgängligheten orsaker förändringen i den beroende variabeln. För att kunna använda tillgänglighetens påverkan sociala förhållanden på liknande sätt som i modellerna för agglomerationsekonomier behövs vidare forskning för att koppla samman de kvalitativa och kvantitativa förståelsen av tillgänglighetens påverkan.

En viktig poäng från den tidigare forskningen är att fördelningen av tillgänglighetsförbättringen mellan olika grupper är central för att förstå transportinfrastrukturens sociala konsekvenser. Exempelvis argumenterar Frei et al. (2009) för att människor med hög mobilitet har ett starkare socialt nätverk och mer stabilt socialt kapital. Liknande menar Lucas (2012) att låginkomsttagare upplever sämre tillgänglighet och har skilda resmönster från övriga socioekonomiska grupper. Denna fördelningsdimension är givetvis intressant och fångas inte direkt upp i den kvantitativa delen av detta arbete. Detta på grund av den uppenbara anledningen att data på kommunnivå är alltför aggregerad för att på något meningsfullt sätt tydligt fånga skillnader mellan socioekonomiska grupper.

Vissa internationella studier (se exempelvis Martínez Sánchez-Mateos & Givoni (2012)) menar att tillgänglighetsförbättringar av höghastighetsjärnväg är mycket små och mer eller mindre försumbara. Om detta hålls för sant är det ändå givetvis rimligt att anta höghastighetsjärnväg ändå kommer få inverkan på människors liv i en social kontext.

Det kan mycket väl vara så att det finns alternativa mått till tillgänglighet som fångar in andra effekter än möjligheten att resa till fler målpunkter som kan påverka de sociala förhållandena.

6.2 Uppslag för vidare studier

De statistiska metoder baserade på en ekonometrisk ansats applicerade på data av panelstruktur tycks vara lämpade för att kvantitativt undersöka kopplingar mellan transportinfrastruktur, tillgänglighet och sociala förhållanden. För att resultaten undersökningar med denna struktur på allvar ska kunna få praktisk relevans vid exempelvis förhandlingar om medfinansiering av stora infrastrukturutbyggnader likt Sverigeförhandlingen är författarnas uppfattning att det krävs betydligt mer geografiskt högupplöst data för såväl beroende som oberoende variabler, förslagsvis på SAMS-nivå. Dessutom skulle eventuellt ett tillgänglighetsmått inkluderandes fler typer av målpunkter än enbart arbetsplatser vara bättre lämpat för denna typ av studie. Förhoppningsvis kan denna studie tjäna som material för att ta ut riktningen för Sverigeförhandlingens och andra berörda aktörers framtida arbete med att utveckla en metodik för att kvantifiera transportinfrastrukturens sociala nyttor.

Vad gäller förbättringsåtgärder för kommande studier är det först och främst datamaterialet som studeras som kan förbättras. Dels bör mer kontinuerlig tillgänglighetsdata studeras och dels bör datan som studeras vara på en mer högupplöst nivå. Som tidigare nämnt i avsnitt 5.7.1 är en brist i detta arbete den tillgänglighetsdata som använts. Den har endast erhållits för tre tidpunkter och dessutom hade den till en början en generell sjunkande trend. Den linjärapproximation som gjorts mellan dessa tre tidpunkter i det här arbetet gör att förändringar som skett mellan dessa år inte till fullo fångas in. När det gäller upplösningen på data som använts har det här arbetet använt data som är på kommunnivå. Att istället använda data på en mer högupplöst nivå, exempelvis stadsdelsnivå eller individnivå, skulle med stor sannolikhet kunna ge en tydligare bild av sambanden mellan tillgänglighet och sociala förhållanden. Detta beror till stor del på att både tillgänglighet i allmänhet och tillgång till transportsystem i synnerhet samt olika sociala faktorer varierar stort inom en kommun. Exempelvis är fördelningseffekter svåra att fånga in på kommunnivå. Som tidigare forskning antydde så tillfaller för- och nackdelar från transportsystem olika sociala grupper och geografiska områden. Hur detta tillfaller kan bero på hur pass välfungerande den lokala stadstrafiken är för olika stadsdelar, hur priskänsliga en social grupp är och om till exempel en järnvägsräls rent av försämrar framkomlighet på stadsdelsnivå.

Vissa av de beroende variablerna kan vara olämpliga mått på social nytta. Även om de beroende variablerna som använts i det här arbetet är utvalda för att de antas vara indikatorer på sociala förhållanden kan de i vissa fall finnas en bristfällig koppling dessemellan. Ofta finns det endast en indirekt koppling mellan de indikatorer som använts och de sociala förhållandena de ska beskriva och ibland kan kopplingen vara svagare än så. Det skulle vara önskvärt att hitta statistik och mått som är mer direkta indikatorer på sociala förhållanden. En annan viktig aspekt rörande de mått som använts som indikatorer för sociala förhållanden är att de i sin natur är objektiva då de är sammanställningar av befolkningsstatistik i Sveriges kommuner. När det är sociala

faktorer kan det dock vara mer relevant att kolla på självskattade faktorer som samlats in exempelvis via enkätundersökningar. Dessa skulle kunna ge en mer subjektiv bild av de sociala förhållandena och på ett mer direkt sätt säga hur invånarna mår och upplever sin tillvaro. Data av detta slag verkar dock vara svår att få tag på då det kräver att dessa medborgarundersökningar har gjorts för alla kommuner över en längre tid. I det här arbetet kunde inte data från medborgarundersökningar användas då det varken fanns tillgängligt för alla kommuner eller för alla år. Däremot så testades under arbetets gång att köra en linjära regressioner för ett antal självskattade sociala förhållanden vid ett år mot tillgänglighet vilket resulterade i höga korrelationsnivåer. Att istället använda sig av indikatorer på sociala förhållanden av denna karaktär är troligtvis mer lämpligt då de dels speglar de sociala förhållandena bättre, dels verkar ha en starkare koppling till tillgänglighet. Det skulle också innebära att det går att täcka in flera komponenter av social nytta, t.ex. trygghet, som inte täcks av de indikatorer som används i den här rapporten.

En viktig aspekt som inte tas upp i detta arbete men som är central för att bedöma sociala nyttor är vilka faktiska värden dessa nyttor faktiskt bidrar med. I en förhandlingssituation likt Sverigeförhandlingen är det rimligen viktigt att tydligt kunna säga vilken aktör nyttan tillfaller och värdet av den. Detta bör därför studeras vidare.

Ett naturligt nästa steg för att utveckla den modellstruktur med fixa effekter som använts i detta arbete är att i modellen inkludera spatiala variabler. Det är troligt att det i de beroende variablerna finns en geografiskt betingad samvarians och därmed skulle modellerna prestera bättre och generera bättre skattningar för övriga parametrar om spatiala variabler inkluderas. Ett konkret exempel på hur detta skulle kunna implementeras finns i metodiken som presenteras i Flöhr (2014).

Metoden som använts i det här arbetet tycks vara lämplig för att kunna kartlägga de grundläggande statistiska sambanden och beräkna fram elasticitet mellan tillgänglighet och sociala nyttor. För att det ska vara möjligt att modellera hur transportinfrastruktur påverkar social nytta måste framtida forskning utreda kausaliteten mellan tillgängligheten och social nytta. Lämpliga instrumentvariabler bör införas i en modell för att justera för den endogenitet som existerar mellan beroende och oberoende variabler.

Angående kausaliteten mellan tillgänglighet och sociala förhållanden går det inte att säga någonting om den utifrån det här arbetet. Detta beror på problem med endogenitet bland de beroende och de oberoende variablerna. För att hantera detta bör instrumentvariabler införas som justera för denna endogenitet. Då både tid och kunskap om lämpliga instrumentvariabler saknades av författarna har inga instrumentvariabler införts i det här arbetet. Går inte att säga något om kausalitet, krävs mer avancerade regressionsmetoder med instrumentvariabler.

Ett av resultaten i det här arbetet var att social nytta bör studeras genom de olika beståndsdelar var för sig eftersom begreppet innehåller många olika aspekter. En viktig deluppgift för att kunna kvantifiera de sociala nyttor som uppstår från en transportsystemutbyggnad är då att vikta de olika beståndsdelarna för att kunna jämföra de totala nyttorna ur ett samhällsekonomiskt perspektiv mellan olika kommuner.

7. Slutsatser

Arbetet fastslår att det utifrån enbart tillgänglighet är svårt säga något om hur transportinfrastruktur påverkar sociala förhållanden eftersom att måttet mäter även andra faktorer än endast transportsystem. Dock verkar potentialbaserad tillgänglighet ha flera fördelaktiga egenskaper för att göra kvantitativa bedömningar av transportsystem. För att kunna särskilja transportsystemsåtgärder från övriga egenskaper av tillgänglighet behövs vetenskap om transportsystemets funktionsförändring. Exempel på detta är tillgänglighetsförändring på grund av förändrad generaliserad kostnad (GK).

Sociala effekter av transportinfrastruktur som kan utgöra sociala nyttor bör inte vägas samman till ett mått genom någon form av indexering eller liknande utan istället bör dess beståndsdelar studeras var för sig. De beståndsdelar som identifierats som lämpliga för att kvantifiera social nytta är hälsa, socialt stöd, känsla av samhörighet, socialt nätverk, tilltro till institutioner, jämlikhet och allmän livskvalitet. Dessa är inte nödvändigtvis heltäckande för begreppet social nytta utan snarare aspekter som diskuterats i tidigare forskning.

Utifrån de paneldataregressioner som utförts kan följande slutsatser dras. Vid tvärsnittsestimering är tillgängligheten signifikant för förväntad livslängd för kvinnor och män, deltagande i senaste riksdagsval, sjukpenningtalet, lönegapet mellan kvinnor och män, andel förvärvsarbetande, antal deltagartillfällen i idrottsföreningar per person, andel bidragshushåll, andel barn inskrivna i förskola och skolbarnsomsorg och andel med eftergymnasial utbildning. För alla dessa utom lönegapet mellan kvinnor och män är korrelationen i sådan riktning att den tyder på att tillgängligheten korrelerar med positiv social nytta.

Vid estimering med justering för kommunspecifika fixa effekter är tillgängligheten signifikant för andelen dagar som kvinnor tar ut VAB, andel ungdomar inskrivna i gymnasieskola, antal deltagartillfällen i idrottsföreningar och andel barn inskrivna i förskola och skolbarnsomsorg. Av dessa tyder korrelationen på positiv social nytta för idrott och förskola och skolbarnsomsorg men negativ för VAB och gymnasiefrekvens

Estimeringen med justering för både kommun- och tidspecifika effekter ger signifikans för tillgängligheten för sjukpenningtalet, lönegap mellan kvinnor och män, andel dagar som kvinnor tar ut VAB, andel ungdomar inskrivna i gymnasieskola, antal deltagartillfällen i idrottsföreningar och andel barn inskrivna i förskola och skolbarnsomsorg. Även här tyder korrelationerna på negativ social nytta för VAB och andel ungdomar inskrivna i gymnasieskola. De andra visar dock på en korrelation som tyder på att tillgängligheten korrelerar med en positiv social nytta.

Sammanfattningsvis tyder resultaten från regressionsanalyserna att tillgänglighet korrelerar med negativ social nytta för andelen dagar som kvinnor tar ut VAB, andel ungdomar inskrivna i gymnasieskola. Lönegapet mellan kvinnor och män visar på både positiv och negativ korrelation med tillgängligheten. Positiv signifikant korrelation med

tillgänglighet fås för förväntad livslängd för kvinnor och män, deltagande i senaste riksdagsval, sjukpenningtalet, andel förvärvsarbete, antal deltagartillfällen i idrottsföreningar per person, andel bidragshushåll, andel barn inskrivna i förkola och skolbarnsomsorg och andel med eftergymnasial utbildning.

På grund av endogenitet kan de skattade statistiska sambanden inte användas för att uppskatta effekten av en transportinfrastrukturinducerad tillgänglighetsförbättring. Alltså kan inga kausala samband mellan transportinfrastruktur och social nytta bevisas. Detta eftersom utifrån detta arbete inte går att säkerställa ett giltigt kausalt orsakssamband mellan oberoende och beroende variabler. Däremot går det utifrån resultaten inte att förkasta ansatsen om att en utbyggnad av transportsystem leder till positiva sociala nyttor. Flera indikatorer för sociala nyttor har signifikant samvarians med tillgänglighet. Dock behövs ytterligare studier av dessa statistiska samband för att kunna utröna hur det kausala förhållandet mellan transportinfrastruktur och sociala förhållanden är.

Referenser

- 2013 års Stockholmsförhandling (2013), Tunnelbana som investering i socialt kapital, 2013 års Stockholmsförhandling: Stockholm.
- 2013 års Stockholmsförhandling (2014), Överrenskommelse om finansiering och medfinansiering av utbyggnad av tunnelbanan samt ökad bostadsbebyggelse i Stockholms län enligt 2013 års Stockholmsförhandling, 2013 års Stockholmsförhandling: Stockholm.
- Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994), "Socioeconomic status and health: the challenge of the gradient", *American psychologist*, vol. 49, nr 1.
- Alm, S. E., Britton, T. (2008), *Stokastik: sannolikhetsteori och statistikteori med tillämpningar*, 1:a uppl., Liber: Stockholm.
- Álvarez, I. C., Barbero, J., Zofío, J. L. (2013), *A Panel Data Toolbox for MATLAB*, Working Papers in Economic Theory 2013/5, Department of Economics Universidad Autónoma de Madrid: Madrid.
- Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R. (2010), "On Making Causal Claims: A Review And Recommendations", *The Leadership Quarterly*, vol. 21, nr. 6, ss. 1086-1120.
- Baltagi, B. H. (2005), *Econometric analysis of panel data*, 3. Uppl., Wiley: Chichester.
- Barnomsorg och skolakommittén (1997), *Att erövra omvärlden – Förslag till läroplan för förskolan* (SOU 1997:157), Utbildningsdepartementet: Stockholm.
- Bel, G. (1997), "Change in Travel Time Across Modes and Its Impact on the Demand for Inter-Urban Rail Travel" *Transportation Research*, vol. 33, nr. 1, ss. 43–52.
- Belsley, D. A., Kuh, E., Welsch, R. E. (1980), *Regression Diagnostics*, John Wiley & Sons: Hoboken, NJ.
- Benoit, K. (2011), *Linear Regression Models with Logarithmic Transformations*, Methodology Institute London School of Economics: London.
- Börjesson, M., Eliasson, J., Isacson, G. (2013), *Infrastrukturens påverkan på ekonomisk tillväxt*, Specialstudie nr. 37, Konjunkturinstitutet: Stockholm.
- Börjesson, M., Eliasson, J., (2015) *Kostnadseffektivitet i valet av infrastrukturinvesteringar*, Finanspolitiska rådet: Stockholm.
- Center for Transportation and the environment North Carolina State University (2008), *Improved methods for assessing social, cultural, and economic effects of transport projects*, NCHRP Project 08-36, Task 66.
- Combes, P., Duranton G., Gobillon, L. (2010 a), "The Identification Of Agglomeration Economies", *Journal of Economic Geography*, vol. 11, nr. 2, ss. 253-266.

- Combes, P., Duranton G., Gobillon, L., Roux, S. (2010 b), "Estimating Agglomeration Economies with History, Geography, and worker Effects", I Glaser, E.L., *Agglomeration Economics*, National Bureau of Economic Research: Cambridge, ss. 253-266.
- Cuba, L., David M. Hummon, D. M. (1993), "A Place to Call Home: Identification with Dwelling, Community, and Region" *The Sociological Quarterly*, vol. 34, nr. 1, ss. 111–131.
- Currie, G., Delbosc, A., (2010), "Modelling the social and psychological impacts of transport disadvantage", *Transportation*, vol. 18, ss. 31–41.
- Currie, G., Richardson, T., Smyth, P., Vella-Broderick, D., Hine, J., Lucas, K., Stanley, J., Morris, J., Kinnear, R., Stanley, J. (2009), "Investigating links between transport disadvantage, social exclusion and well-being in Melbourne - Preliminary Results", *Transport Policy Special Issue International Perspectives on Transport and Social Exclusion*, vol.16, nr.3, ss. 90–96.
- Department for transport, (2013), *Valuing the social impacts of public transport - Final report*, Department for transport: London.
- Emanuelsson, R., Petersson, L., (2011), *The Demand for Swedish Higher Education – an Econometrical Panel Data Analysis*, Linköpings universitet Filosofiska fakulteten: Linköping.
- Finansdepartementet (2009), *Redovisning av kommunal medfinansiering*, (Regeringens proposition 2008/09:228), Stockholm: Regeringskansliet.
- Flöhr, A. (2014), *Rumslig autokorrelation i kommunal arbetslöshet*, Ekonomihögskolan, Lunds universitet: Lund.
- Folkhälsomyndigheten (2014), *Folkhälsan i Sverige: årsrapport 2014*, Stockholm: Folkhälsomyndigheten.
- Frei, A., Axhausen, K. W., Ohnmacht, T., (2009), "Mobilities and Social Network Geography: Size and Spatial Dispersion—the Zurich Case Study" i: Ohnmacht, T., Maksim, H., Bergman, M. M. (Red.), *Mobilities and Inequalities*, Ashgate: Surrey.
- Fors, F., Brülde, B., (2012). *Varför arbetslösa mår sämre*. I Weibull, L. et al. (red.) *I framtidens skugga: fyrtiotvå kapitel om politik medier och samhälle: SOM-undersökningen 2011*.ss. 247-259.
- Försäkringskassan (2015), *Jämställdhet och sjukfrånvaro: Förstagångsföräldrar och risken för sjukfrånvaro vid olika jämställdhetssituationer och effekter på sjukfrånvaron av reformer inom föräldraförsäkringen* (Socialförsäkringsrapport 2015:3), Försäkringskassan Analys och prognos: Stockholm.
- Geurs, K. T., Boon, W., & Van Wee, B. (2009), "Social impacts of transport: literature review and the state of the practice of transport appraisal in the Netherlands and the United Kingdom", *Transport reviews*, vol. 29, nr. 1, ss. 69-90.

- Grieco, M. (2006), "Gender, transport and social empowerment: Investigating the consequences of the interaction between gender and constrained mobility", *Naturbilder und Lebensgrundlagen-Konstruktionen von Geschlecht*, vol. 53.
- Gutiérrez, J., Monzón, A., Piñero, J. M. (1998), "Accessibility, Network Efficiency, and Transport Infrastructure Planning", *Environment and Planning*, vol. 30, ss. 1337–1350.
- Handy, S. L., Niemeier, D. A. (1997), "Measuring Accessibility: An Exploration of Issues and Alternatives" *Environment and Planning*, vol. 29, ss. 1175 –1194.
- Hansen, G. H., (1959), "How Accessibility Shapes Land Use", *Journal of the American Institute of Planners*, vol. 25:2, ss. 73-76.
- Ingram, D. R., (1971), "The concept of accessibility: A search for an operational form", *Regional Studies*, vol. 5:2, ss. 101-107.
- Jönsson, L. R., (2003), *Arbetslöshet ekonomi och skam – om att vara arbetslös i dagens Sverige*, School of Social Work: Lund.
- Koenig, J. G., (1980), "Indicators of Urban Accessibility: Theory and Appliation", *Transportation*, vol. 9, ss. 145-172.
- Kenny, D. A. (1979), *Correlation and causality*, Wiley-Interscience: New York.
- Kroes, E., Daly, A., Gunn, H., van der Hoorn, T. (1996), "The Opening of the Amsterdam Ring Road: A Case Study on Short-Term Effects of Removing a Bottleneck", *Transportation*, vol. 23, nr. 1, ss. 71–82.
- Leighley, J. E., Nagler, J. (1992), "Socioeconomic Class Bias in Turnout, 1964–1988: The Voters Remain the Same", *American Political Science Review*, vol. 86, ss. 725-736.
- Litman, T. (2011), *Evaluating accessibility for transportation planning*, Victoria Transport Policy Institute: Victoria, BC.
- Lucas, K. (2012), "Transport and social exclusion: Where are we now?", *Transport Policy*, vol. 20, ss.105-113.
- Martínez Sánchez-Mateos, H., Givoni, M. (2012), "The accessibility impact of a new High-Speed Rail line in the UK – a preliminary analysis of winners and losers", *Journal of Transport Geography*, vol. 25, ss.105-114.
- Medfinansieringsutredningen (2011), *Medfinansiering av transportinfrastruktur (SOU:2011:49): Ett nytt system för den långsikta planeringen av transportinfrastruktur samt riktlinjer och processer för medfinansiering*, Regeringskansliet: Stockholm.
- Mild, J. (2009), *Järnvägar, Tillgänglig online*:
<http://www.bgf.nu/historia/6/jarnvagar.html#1> (2015-10-21)

- National Cooperative Highway Research Program (NCHRP) (2001 a), Guidebook for Assessing the Social and Economic Effects of Transportation Projects, National Academy Press: Washington D.C.
- National Cooperative Highway Research Program (NCHRP) (2001 b), Assessing the Social and Economic Effects of Transportation Projects, National Academy Press: Washington D.C.
- Nohrstedt, L. (2014), "Brett stöd för höghastighetståg", Ny Teknik, 3 juli, Tillgänglig online: http://www.nyteknik.se/nyheter/fordon_motor/jarnvag/article3835762.ece (2015-10-21).
- Näringsdepartementet (2009), Mål för framtidens resor och transporter (Regeringens proposition 2008/09:93), Regeringskansliet: Stockholm.
- Näringsdepartementet (2013), Utbyggnad av tunnelbanan och ökad bostadsbebyggelse i Stockholm län (Kommittédirektiv 2013:22), Regeringskansliet: Stockholm.
- Näringsdepartementet (2014 a), Utbyggnad av nya stambanor samt åtgärder för bostäder och ökad tillgänglighet i storstäderna. (Kommittédirektiv 2014:106), Regeringskansliet: Stockholm.
- Näringsdepartementet (2014 b), Tilläggsdirektiv till Utredningen om utbyggnad av nya stambanor om utbyggnad av nya stambanor samt åtgärder för bostäder och ökad tillgänglighet i storstäderna. (Kommittésirektiv 2014:113), Regeringskansliet: Stockholm.
- Pearl, J. (2000), Causality, Cambridge University Press: Cambridge.
- Preston, J. and Wall, G. (2008), "The Ex-ante and Ex-post Economic and Social Impacts of the Introduction of High-speed Trains in South East England", Planning Practice and Research, vol. 23, nr. 3, ss.403-422.
- Riksidrottsförbundet (2010), Idrott och integration, Tillgänglig online: http://www.rf.se/ImageVaultFiles/id_29196/cf_394/Idrott_och_integration.pdf (2016-01-07).
- Riksidrottsförbundet (FoU 2005), Idrottens sociala betydelse – en statistisk undersökning hösten 2004, Tillgänglig online: http://www.rf.se/ImageVaultFiles/id_29192/cf_394/Idrottens_sociala_betydelse.pdf (2016-01-07).
- Salomon, K. (2012), "Tågets historia har inte gått spårlöst förbi". Svenska Dagbladet, 26 maj, Tillgänglig online: <http://www.svd.se/tagets-historia-har-inte-gatt-sparlost-forbi> (2015-10-24).
- Sandberg, M. (2015), Granskning av Sverigeförhandlingens arbete med nyttor för nya stambanor, Sweco: Stockholm.
- Sehlin, J. (2012), Indata till den nationella svenska persontrafikmodellen Sampers för basår 2010 med avseende på socioekonomisk statistik, TRV 2011/87875, Statistiska Centralbyrån (SCB).

- Shlens, J. (2014), A Tutorial on Principal Component Analysis, eprint arXiv:1404.1100, ARXIV, Tillgänglig online: <http://arxiv.org/pdf/1404.1100v1.pdf> (2015-09-14).
- Skolverket (2014), Vad ungdomar gör efter gymnasieskolan – en registerstudie, Skolverket: Stockholm.
- Social Exclusion Unit (2003), Making the Connections: Final Report On Transport and Social Exclusion, Office of the Deputy Prime Minister: London.
- Socialstyrelsen (2013), Ekonomisk utsatthet och välfärd bland barn och deras familjer 1968-2010: Underlagsrapport till Barns och ungas hälsa, vård och omsorg 2013, Socialstyrelsen: Stockholm.
- Sverigeförhandlingen (2015), Metoder och redovisning av nyttoberäkningar i Sverigeförhandlingen, Sverigeförhandlingen: Stockholm.
- Sverigeförhandlingen (2016 a), Förhandlingskonferenser. Tillgänglig online: <http://forhandling.sverigeforhandlingen.se/forhandlingskonferenser/> (2016-01-05).
- Sverigeförhandlingen (2016 b), Delrapport från Sverigeförhandlingen: Höghastighetsjärnvägens finansiering och kommersiella förutsättningar, SOU 2016:3, Tillgänglig online: http://media.sverigeforhandlingen.se/2016/01/SOU-2016_03_webbpdf-160112-2.pdf (2016-01-13).
- Talen, E. (1998), “Visualizing Fairness: Equity Maps for Planners”, Journal of the American Planning Association, vol. 64, nr. 1, ss. 22–38.
- Talen, E., Anselin, L. (1998), “Assessing Spatial Equity: An Evaluation of Measures of Accessibility to Public Playgrounds”, Environment and Planning, vol. 30, ss. 595–613.
- Trafikanalys (2011), Trafikverkets arbete med modeller för samhällsekonomiska analyser- statusrapport 2011 (rapport 2011:7), Trafikanalys: Stockholm.
- Trafikverket (2012), Introduktion till samhällsekonomisk kalkyl, Publikation 2012:220, Tillgänglig online: http://www.trafikverket.se/contentassets/81dcc49542364c39a084c867737d4bf8/pm_2012_01_introduktion_till_samhallsekonomisk_analys.pdf (2015-09-28).
- Trafikverket (2015 a), Utbyggnadsstrategier och förhandlingsunderlag för höghastighetsjärnvägar – Underlag till Sverigeförhandlingen, Publikation 2015:241, Tillgänglig online: http://media.sverigeforhandlingen.se/2015/01/Trafikverket_Utbyggnadsstrategier_for_handlingsunderlag_for_hhj_151204.pdf (2016-01-04).
- Trafikverket (2015 b), Trafikering med nya höghastighetsbanor Stockholm-Göteborg/Malmö, Publikation 2015:4, Tillgänglig online: http://www.trafikverket.se/contentassets/899cd278cd2c4686bc4dde4f2324780a/planer--beslutsunderlag/trafikering_nya_stambanor_150529.pdf (2015-10-15).
- Truelove, M. (1993), “Measurement of Spatial Equity”, Environment and Planning, vol. 11, ss. 19–34.

- Urry, J., (2000), *Sociologies beyond Societies: Mobilities for the Twenty-First Century*, Routledge: New York.
- Utredningen för höghastighetsbanor (2009), *Höghastighetsbanor – ett samhällsbygge för stärkt utveckling och konkurrenskraft* (SOU 2009:74), Regeringskansliet: Stockholm.
- van Wee, B., Annema, J., Banister, D. (2013), *The transport system and transport policy*, Edward Elgar Publishing Limited: Cheltenham.
- Widgery, R. N. (1982), "Satisfaction with the Quality of Urban Life: A Predictive Model", *American Journal of Community Psychology*, vol. 10, Nr. 1, ss. 37–48.
- Wikhall, M. (2014), *Kommunal medfinansiering av statlig infrastruktur och högskoleverksamhet*.(2014:8), Statskontoret: Stockholm. Tillgänglig online: <http://www.statskontoret.se/globalassets/publikationer/2014/201408.pdf> (2015-11-04).
- Winter, K (2015), *Sociala nyttor i Sverigeförhandlingen*, KTH: Stockholm.
- von Sydow, S. (2011), *Miljöaspekter i samhällsekonomiska kalkyler: exemplet Botniabanan*, Nationalekonomiska institutionen, Uppsala universitet: Uppsala.

Appendix A

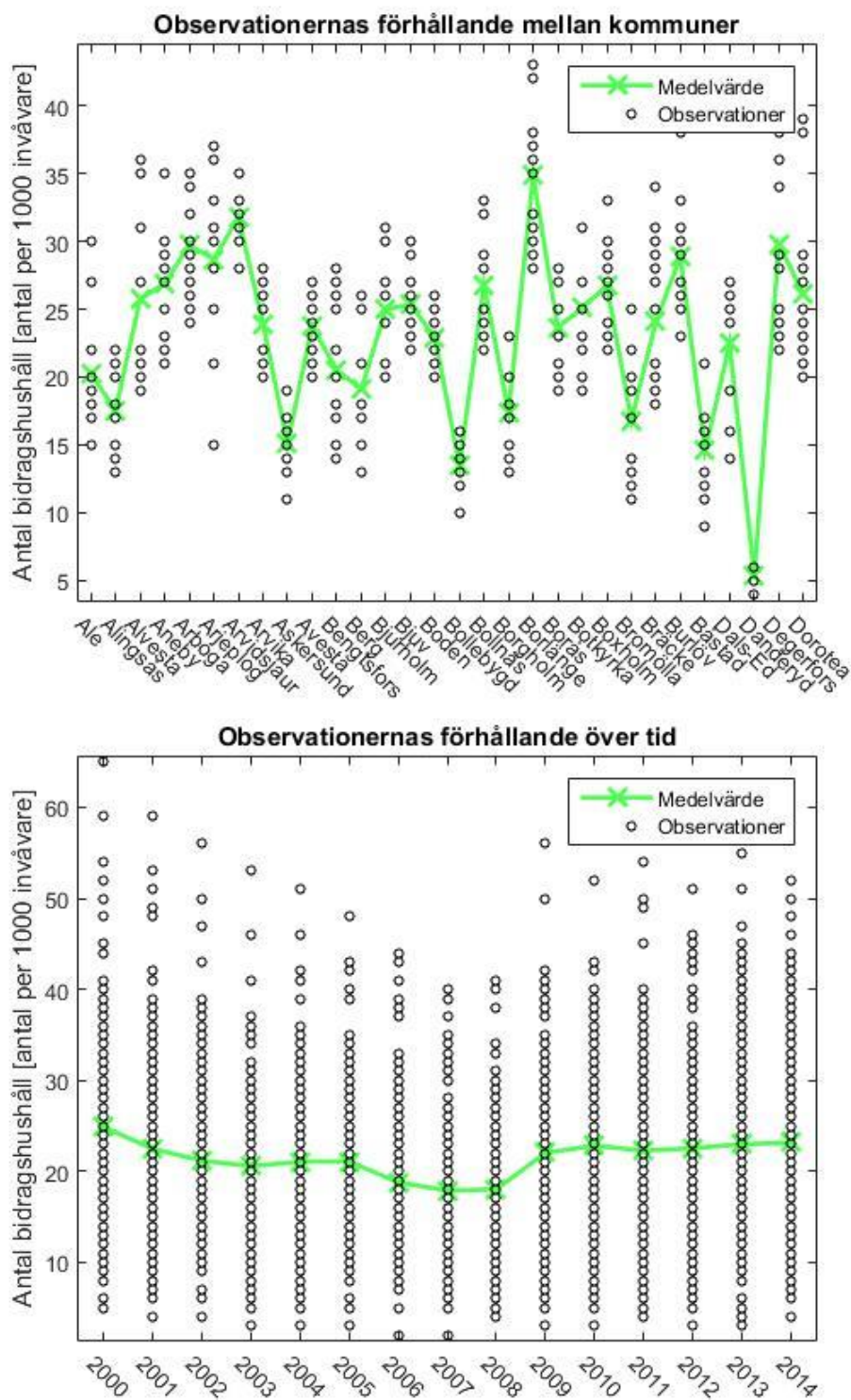
Tabell över de beroende variablerna och dess beskrivning från den myndighet de inhämtats från.

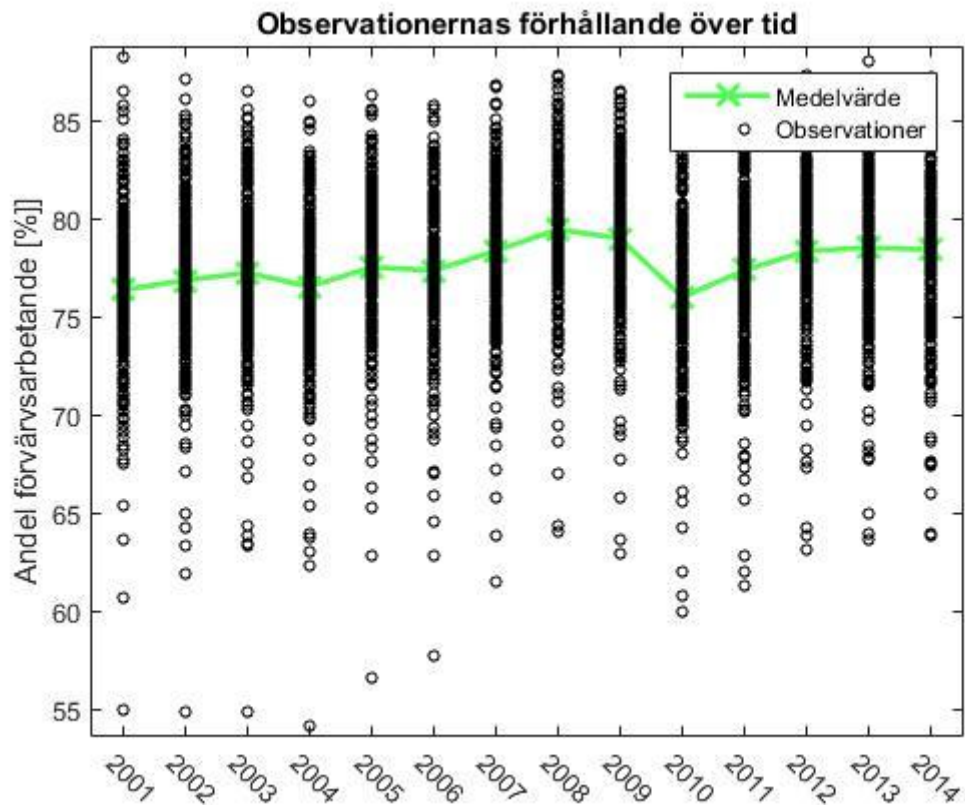
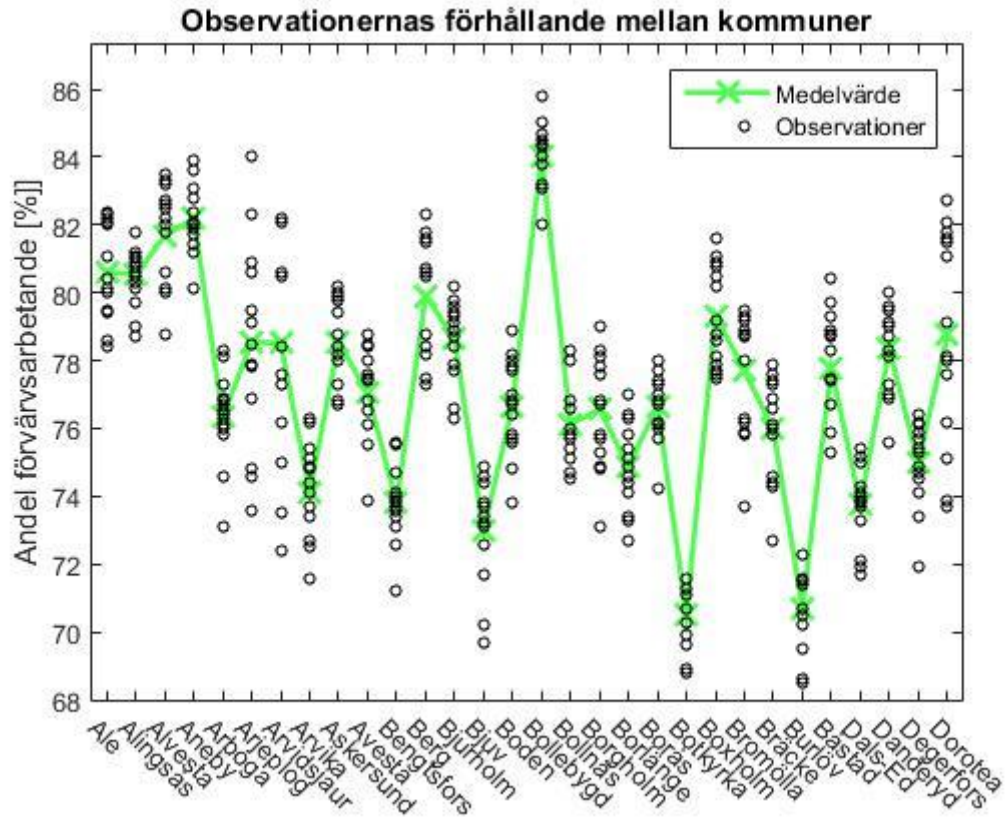
Indikator	Källa	Beskrivning
Medellivslängd, kvinnor	SCB	Beräknad återstående medellivslängd i antal år vid födseln perioden T-4 till år T för kvinnor födda i Sverige, kommun.
Medellivslängd, män	SCB	Beräknad återstående medellivslängd i antal år vid födseln perioden T-4 till år T för män födda i Sverige, kommun.
Deltagande i riksdagsval	SCB	Antal angivna röster i senaste riksdagsvalet (giltiga och ogiltiga) dividerat med antal röstberättigade, multiplicerat med 100.
Sjukpenningtalet, dagar/registrerad försäkrad	Försäkringskassan	Antal utbetalda dagar med sjukpenning och rehabiliteringspenning per registrerad försäkrad i åldrarna 16-64 år exklusive försäkrade med hel sjukersättning eller aktivitetsersättning (före år 2003 hel förtidspension eller helt sjukbidrag). Alla dagar är omräknade till nettodagar.
Lönegap, medianlön kvinnor – medianlön män anställda av kommunen, kr	SCB	Lönegap median kvinnor – median män anställda av kommunen, kronor (i månaden). All lön som redovisas är uppräknad till heltidslön. I heltidslön ingår grundlön plus rörliga tillägg och förmåner. Avser anställda 18-66 år (t.o.m. 2013 avsågs 18-64 år) som är månads- eller timavlönade. Anställda i kommunägda bolag ingår inte.
Kvinnor som tar ut tillfällig föräldrapenning (VAB), andel av antal dagar (%)	Försäkringskassan	Andel tillfällig föräldrapenning (sett till antal nettodagar) som kvinnor tar ut. Avser: Dagar i samband med vård av barn.
Barn 1-12 år inskrivna i förskola och skolbarnsomsorg, andel (%)	SCB Skolverket	Antal inskrivna 1-12 år i förskola och skolbarnsomsorg dividerat med antal barn 1-12 år. Avser samtlig regi.
Bidragshushåll ekonomiskt bistånd, antal/1000 inv	SCB och Socialstyrelsen	Antal hushåll med ekonomiskt bistånd, dividerat med antal invånare totalt 31/12 dividerat med 1000. Avser samtlig regi.

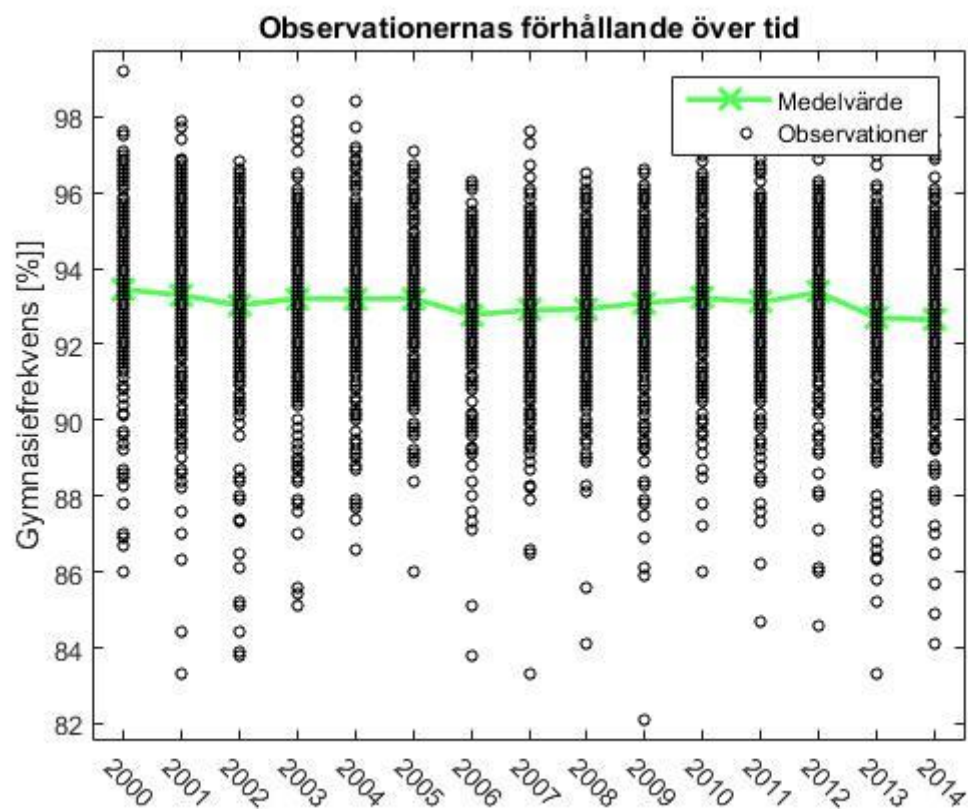
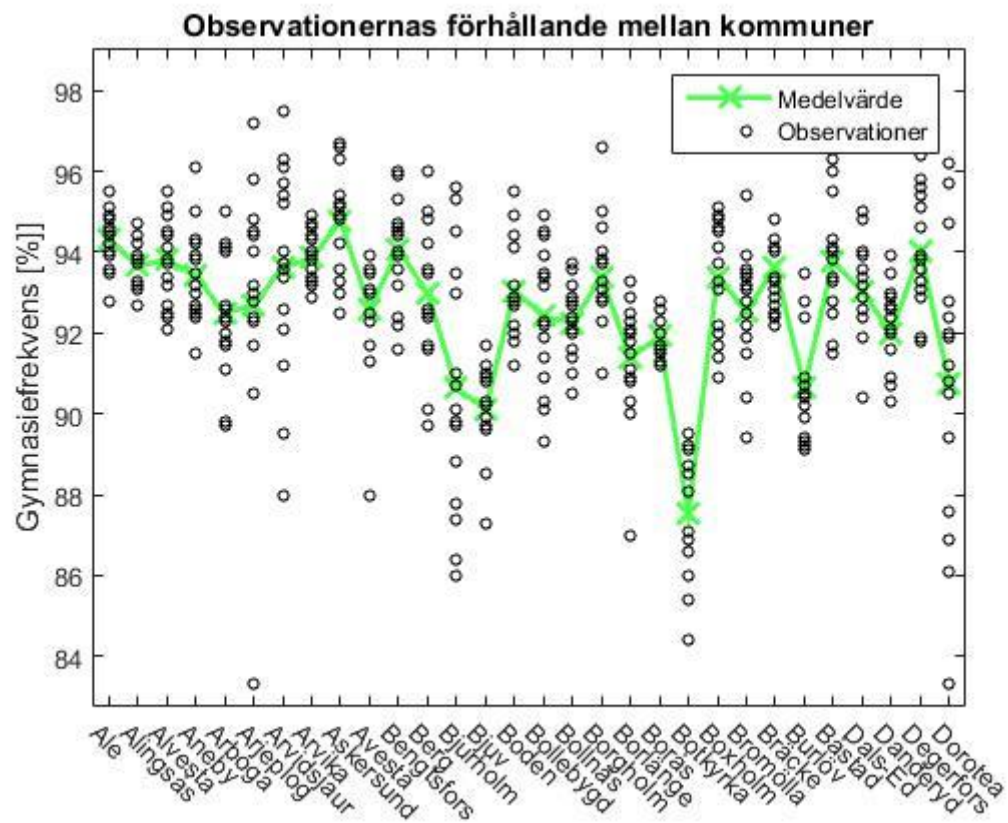
Deltagartillfällen i idrottsföreningar, antal/inv 7-20 år	Riksidrottsförbundet	Antal deltagartillfällen i åldern 7-20 år i LOK-stödsberättigade idrottsföreningar under året, dividerat med antal invånare 7-20 år den 31/12. Deltagartillfällen är summan av antalet närvarande deltagare i av RF godkända sammankomster för vilken föreningen sökt LOK-stöd (Statligt lokalt aktivitetsstöd). LOK-stöd ges till aktiviteter för pojkar och flickor 7-20 år. Stöd ges även till handikappidrottare som är äldre än 20 år, men detta är exkluderat i denna sammanräkning. En sammankomst är en ledarledd gruppaktivitet inriktad mot föreningens idrottsliga verksamhet.
Förvärvsarbetande invånare 20-64 år, andel (%)	SCB (RAMS)	Antal förvärvsarbetande i åldern 20-64 år dividerat med antal invånare i åldern 20-64 år den 31/12. Med förvärvsarbetande avses personer med löneinkomst av anställning under november månad, och personer med inkomst av aktiv näringsverksamhet.
Gymnasiefrekvens, (%)	SCB Skolverket	Antal barn 16-18 inskrivna i gymnasieskolan den 15/10 i kommunen dividerat med antal barn 16-18 år i kommunen den 15/10.
Kvinnor som tar ut föräldrapenning, andel av antal dagar (%)	Försäkringskassan	Andel av föräldrapenning (sett till nettodagar) som kvinnor tar ut.

Appendix B

Diagram som synliggör de beroende variabelnas heterogenitet både mellan kommuner och över tid.







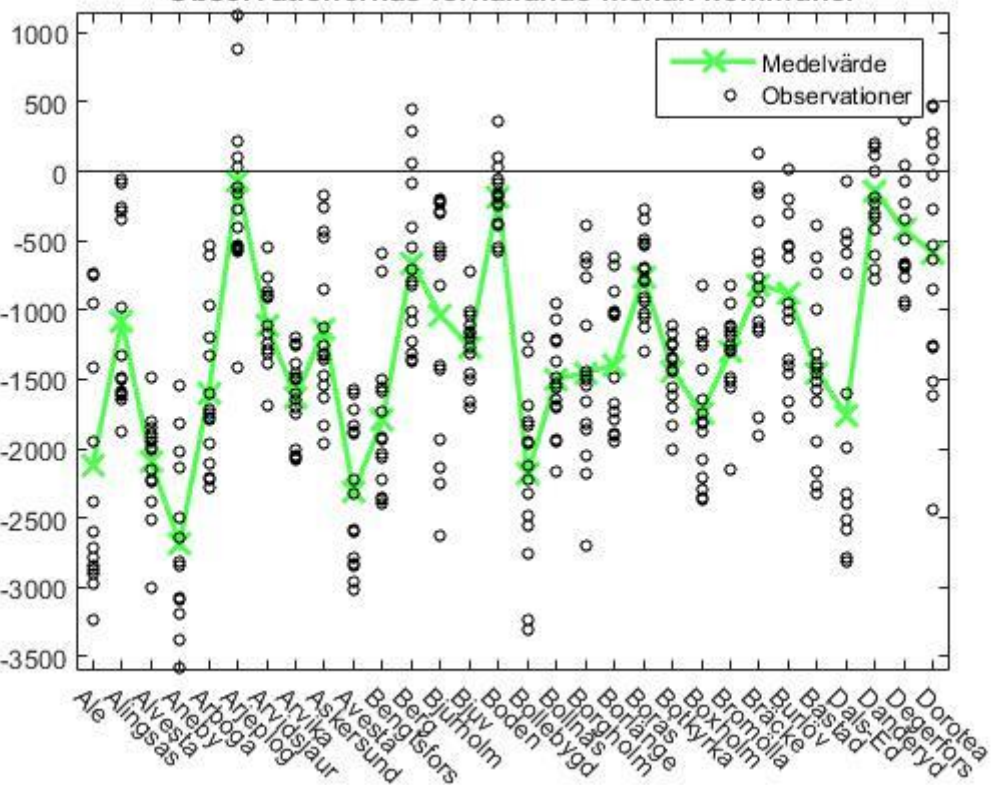
Förväntad livslängd för Kvinnor [år]

Medelvärde
Observationer

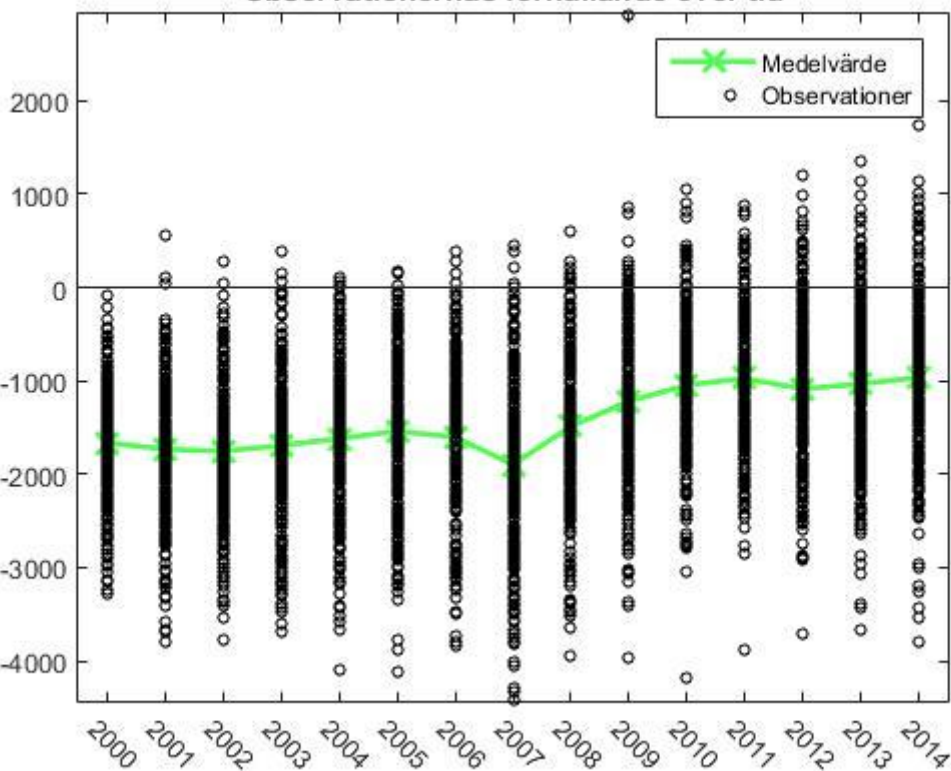
År	Medelvärde (År)	Observationer (År)
2000	81.8	77.0, 78.8, 79.3, 79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2001	82.0	77.8, 78.8, 79.0, 79.2, 79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2002	82.0	77.0, 78.8, 79.0, 79.2, 79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2003	82.1	78.2, 78.8, 79.0, 79.2, 79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2004	82.2	77.5, 78.5, 78.8, 79.0, 79.2, 79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2005	82.3	79.3, 79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2006	82.5	79.0, 79.2, 79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2007	82.7	79.5, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2008	82.8	80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2009	82.9	80.3, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2010	83.0	79.7, 79.8, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2011	83.2	80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2012	83.3	80.3, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2013	83.4	79.5, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0
2014	83.7	79.4, 80.0, 80.2, 80.5, 80.8, 81.0, 81.2, 81.5, 81.8, 82.0, 82.2, 82.5, 82.8, 83.0, 83.2, 83.5, 83.8, 84.0, 84.2, 84.5, 84.8, 85.0, 85.2, 85.5, 85.8, 86.0

Löneskillnad mellan kvinnor och män anställda av kommunen [t]

Observationernas förhållande mellan kommuner



Observationernas förhållande över tid



[illegible]

The figure is a scatter plot with a line graph overlay. The y-axis is labeled 'Sjukpenningtalet [dagar per invånare]' and ranges from 5 to 35. The x-axis shows years from 2000 to 2014. A legend in the top right corner indicates that the green line with 'x' markers represents the 'Medelvärde' (average) and the open circles represent 'Observationer' (observations). The average sick days start at approximately 17 in 2000, rise to a peak of about 21 in 2002-2003, and then generally decline to around 11 by 2014. The observations are clustered around the average line, showing a similar downward trend over the period.

