



UPPSALA
UNIVERSITET

UPTEC-ST513024

Examensarbete 30 hp
Juni 2013

Ny metod för mätning av vägkvalitet

En statistisk studie över samband mellan
vägkvalitet och socioekonomisk utveckling

Magnus Pargéus



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

New method for measuring road quality- A statistical study of the relationship between road quality and socioeconomic development

Magnus Pargéus

A new tool with implications for road maintenance has the potential to yield statistical data for road quality in a large number of regions. This thesis investigates the possibility to develop a method, from this data, for evaluating the effects of road quality on socioeconomic development.

The thesis is essentially divided into two parts. Part one looks to qualitatively establish the link between socioeconomic development and road quality and the second part focuses on how to develop a method for evaluating the effects of road quality on socioeconomic development. A method was devised using categorical data and a multinomial regression model. This method provides a comprehensive overview of the complex relationship between a variety of different variables. To illustrate this method a study was conducted to show the relationship between the proportion of poor roads and prosperity in different regions.

The results of the study revealed a strong correlation between the proportion of poor roads and prosperity. This study also illustrates how methods can be developed to raise awareness and increase knowledge over the effects of road quality on socioeconomic development. This information could be useful when evaluating the potential effects of different road projects. It could also make it easier to secure financing for a project, by making the effects of the project easier to predict.

Handledare: Lars Forslöf
Ämnesgranskare: Jesper Rydén
Examinator: Elisabet Andresdottir
ISSN: 1650-8319, UPTec-STS13024

Populärvetenskaplig beskrivning

En fungerande infrastruktur gör det möjligt för människor i ett modernt samhälle att förenas socialt, ekonomiskt och kulturellt. En av de viktigaste samhällsfunktionerna är transportsystem, vilka möjliggör förflyttningen av människor, varor och tjänster. Då större delen av världens transporter sker via vägar spelar väginfrastrukturen en viktig roll för hur bra dessa transportsystem fungerar.

Vägkvalitet har en stark påverkan på hur funktionell och effektiv en väginfrastruktur är. Dåliga vägar leder bland annat till längre restider och större reparationskostnader för de fordon som trafikerar vägarna. Detta påverkar i sin tur alla de människor, företag och institutioner som på ett eller annat sätt använder sig av transportsystemen, vilket betyder att vägkvalitet påverkar en stor del av samhället och dess utveckling.

Att vägkvalitetens inverkan på samhällsutveckling skiljer sig åt beroende på regioners olika förutsättningar är en väletablerad sanning inom forskningsamfundet. Trots detta är det inte ordentligt utrett hur dessa skillnader ser ut och vad de beror på. Detta kan till viss del förklaras med att den information som finns tillgänglig om vägförhållanden i olika regioner till stor del är bristfällig.

Ett nytt verktyg för vägunderhåll kan på sikt ge information om den övergripande vägstandarden i ett stort antal olika regioner. Syftet med denna rapport har varit att visa på att informationen från detta verktyg kan användas för att utveckla en metod som visar hur vägkvalitet inverkar på olika delar av samhällsutvecklingen.

För att kunna utveckla denna metod skapades en databas upp som sammanförde data över samhällsekonomiska variabler med data över vägkvalitet. Utifrån detta utvecklades en statistisk modell som genom att dela in data i kategorier ger en god överblick över samband som normalt är svåra att se. För att visa hur modellen fungerar studerades sambandet mellan dåliga vägar och välstånd i olika regioner.

Resultatet från studien visar på ett tydligt samband mellan andelen dåliga vägar och nivån av välstånd, men också på möjligheterna att utveckla metoder som kan ge större förståelse och ökad medvetenhet över de effekter vägkvalitet har på samhällsutveckling. Detta skulle till exempel vara till hjälp när man utvärderar potentialen hos olika väginfrastrukturprojekt. Det kan också göra det lättare att få finansiering till projekten, då de potentiella effekterna blir lättare att se.

Innehållsförteckning

1. INLEDNING.....	4
1.1 BAKGRUND	4
1.2 SYFTE.....	5
1.3 RAPPORTSTRUKTUR	5
2. VÄGKVALITET.....	6
2.1 VÄGINFRASTRUKTUR.....	6
2.2 UNDERHÅLL AV VÄGAR	6
3. SOCIOEKONOMISKA EFFEKTER	8
3.1 UTVECKLINGSLÄNDER.....	8
3.1.1 <i>Städer</i>	8
3.1.2 <i>Rurala områden</i>	8
3.1.3 <i>Regionala skillnader</i>	9
3.1.4 <i>Fattighedsreduktion</i>	9
3.1.6 <i>Institutionell svaghet</i>	10
3.2 INDUSTRIALISERADE LÄNDER.....	10
3.3 KINA- ETT LAND MED STORA SKILLNADER.....	11
3.5 SAMMANFATTNING	11
4. MÄTNING AV VÄGSTANDARD.....	13
4.1 YTOJÄMNHETER	13
4.2 MÄTMETODER	13
4.2.1 <i>IRI</i>	13
4.2.2 <i>Brister hos IRI</i>	13
4.2.3 <i>Panel ratings</i>	14
4.4 BEHOVSBILD.....	14
4.5 NY MÄTMETOD- ROADROID	15
4.5.1 <i>Historik</i>	15
4.5.2 <i>Teknisk beskrivning</i>	15
4.5.3 <i>Möjligheter med tekniken</i>	16
4.6 SPATIAL DATA MINING.....	16
5. KATEGORISKA DATA OCH MULTINOMINELLA LOGITMODELLER	18
5.1 KATEGORISKA VARIABLER.....	18
5.2 GENERALISERADE LINJÄRA MODELLER.....	18
5.3 LOGITMODELLER.....	18
5.4 <i>Multikategoriska logitmodeller</i>	19
5.5 <i>Betingade logitmodeller</i>	20
5.6 <i>Multinominella logitmodeller</i>	20
5.6.1 <i>Begränsningar</i>	21
5.6.2 <i>Diagnostik</i>	21
6. MODELL	22
6.1 VAL AV TILLÄMPNING.....	22
6.1.2 VAL AV VARIABLER.....	23
6.2 BEGRÄNSNINGAR HOS DATA	23
7. RESULTAT	25
7.1 PRESENTATION AV RESULTAT	25
7.2 TOLKNING AV RESULTAT	26

7.3 PARAMETERSKATTNING	26
7.4 EN MODELL - FLERA MÖJLIGHETER.....	27
8. SLUTSATSER.....	29
9. DISKUSSION	30
10. LITTERATURLISTA	32
10.1 BÖCKER	32
10.2 RAPPORTER OCH VETENSKAPLIGA ARTIKLAR	32
10.3 ELEKTRONISKA KÄLLOR.....	33
11. APPENDIX	34
11.1 R-KOD FÖR DEN FÖRESLAGNA MODELLEN	34
11.2 UTDRAK UR DEN DATABAS SOM ANVÄNDES I MODELLERINGEN.....	35

I. Inledning

I.1 Bakgrund

Ett modernt samhälle är uppbyggt kring ett fungerande och effektivt transportsystem. Transportsystem fungerar som en viktig kanal för förflyttningen av varor, tjänster och människor. Detta innebär att transportsystem är sammanlänkade med ett stort antal samhällsfunktioner och utgör grunden för ekonomisk aktivitet. (Kayondo-Ndandiko, 2012, Umoren et al, 2009)

Vägtransport står för 90 % av passagerartrafiken och 80 % av godstransporterna i världen, därför utgör väginfrastrukturen en fundamental roll i utvecklingen av fungerande transportsystem. (Internationella Väg- och Transportfacket, 2000) Kvaliteten hos infrastrukturen är en viktig aspekt i utvecklandet av ett effektivt transportsystem. Vägkvalitet påverkar främst restider och reparationskostnader för de fordon som trafikerar vägarna, vilket i sin tur påverkar alla de aktörer som på ett eller annat sätt har en koppling till transportsystemen. (Blyde, 2010)

Det finns en allmän konsensus inom forskningsamfundet om att transportsektorn har en avgörande inverkan på en regions utveckling, men inte hur denna inverkar under skilda ekonomiska och geografiska förutsättningar. Detta beror på att det ofta är svårt att utreda eller att det saknas underlag för att kvantifiera dessa i direkta och indirekta effekter. (Umoren et al, 2009)

En litteraturgenomgång rörande vägkvalitetens inverkan på socioekonomisk utveckling visar att regioners skilda förutsättningar påverkar hur vägkvaliteten inverkar på socioekonomisk utveckling. Detta innebär att effekterna från väginfrastrukturprojekt varierar beroende på regioners specifika behovsbild. (Umoren et al, 2009, Fan och Chan-Kang, 2005, Gwilliam, 2002)

De metoder som idag används för att mäta vägkvalitet är dyra och komplexa, eller är inte lämpliga för jämförelser över tid och mellan olika platser. Detta innebär att det underlag som är tillgängligt för beslutsfattare oftast är bristfälligt och inte ger en adekvat bild över den aktuella vägstandarden. Det finns därför ett behov att skapa en ökad förståelse för vägkvalitet och dess effekter under olika förutsättningar. (Kayondo-Ndandiko, 2012, Kumar 2010, Sayers och Karamihas, 1998)

Roadroid är ett nystartat svenskt företag som utvecklat ett kostnadseffektivt system för mätning av vägkvalitet. Systemet är tänkt att fungera som hjälpmedel för effektivt underhållsarbete av vägar, men kan på sikt även ge ett betydande statistiskt underlag över vägförhållanden i ett stort antal regioner (Forslöf, 2013)

1.2 Syfte

Syftet med denna rapport är att visa på hur samhällsekonomisk data kan integreras med data från Roadroids system för mätning av vägstandard. Detta utgör sedan grunden till en metod som kan ge samhällsekonomiskt stöd för bedömningar och prioriteringar rörande underhåll och investeringar i ett givet vägnät.

1.3 Rapportstruktur

Denna rapportens första del är av en samhällsvetenskaplig karaktär och syftar till att ge läsaren en övergripande bild av relationen mellan vägkvalitet och socioekonomiska faktorer, samt på vilket sätt denna påverkas av olika regionala förutsättningar.

Forskning inom detta område fokuserar främst på situationen i utvecklingsländer. Infrastrukturen i dessa länder är ofta underutvecklad och förbättringar är av stor socioekonomisk betydelse. I industrialiserade länder är infrastrukturen mer utvecklad och invånarna åtnjuter redan de mest grundläggande effekterna som effektiva transportsystem medför. Forskning i dessa länder handlar främst om restider och om att hantera köbildningar. Eftersom denna rapport ämnar studera de socioekonomiska effekterna av vägkvalitet ligger fokus mer på utvecklingsländer än industrialiserade länder. Detta då förbättringar av vägkvalitet har en större socioekonomisk effekt i dessa länder.

Rapportens efterföljande del beskriver de metoder som idag används för att mäta vägstandard, de begränsningar som förekommer och den behovsbild som finns hos användarna. Med bakgrund i detta beskrivs ett nytt system för mätning av vägstandard, samt de möjligheter detta medför att skapa en databas som integrerar data över vägkvalitet med samhällsekonomisk data.

Detta följs av ett teoriavsnitt där kategoriska data och multinominella logitmodeller anges som ett sätt att hantera och presentera de ofta komplexa sambanden mellan vägkvalitet och socioekonomisk utveckling. I detta avsnitt förutsätts läsaren vara bekant med statistisk regression, därför beskrivs inte detta i närmare detalj.

Det finns en rad olika statistiska metoder för att hantera kategoriska data. Den multinominella logitmodellen valdes eftersom den bedömdes vara den mest generella modellen, och den är kapabel att hantera ett större antal olika tillämpningar än de alternativa metoderna.

Denna rapport syftar främst till att visa vilka möjligheterna som ett utökat statistiskt underlag över vägstandard medför. Då detta underlag ännu inte är särskilt stort, har begränsade mätningar fått representera hela regioner. Detta innebär att de resultat som presenteras i denna rapport inte är helt tillförlitliga, utan ska snarare ses som en indikation för hur det förhåller sig i verkligheten.

2. Vägkvalitet

2.1 Väginfrastruktur

Det moderna samhället är beroende av ett effektivt fungerande transportsystem. Vita(2008) identifierar i sin rapport för EU-kommissionen ett fungerande transportsystem som en nyckelfaktor för ekonomisk tillväxt och utveckling. Effektiv infrastruktur och tillgänglighet attraherar produktions- och konsumtionscentra, vilket ger en positiv regional effekt. Det spelar därför en fundamental roll i att nå målen för ökad tillväxt och jobbskapande som dikterats i Lissabonobjektiven. Utvecklas inte transportsystemet i takt med övriga samhället är risken stor att detta begränsar den socioekonomiska utvecklingen (Assafou, 2007).

Då större delen av världens transportsystem är vägburna utgör väginfrastrukturen en särskilt viktig del av ett lands infrastruktur. Den fungerar som en viktig kanal för insamlandet och transporten av varor, tjänster, människor och delning av information. Detta innebär att väginfrastrukturen är sammankopplad med ett stort antal viktiga samhällsfunktioner och utgör grunden i ekonomisk aktivitet. Därför påverkar den även den samhällsekonomiska utvecklingen. (Kayondo-Ndandiko, 2012, Umoren et al, 2009) Exempelvis har Världsbanken en dubbelt så hög genomsnittlig ekonomisk avkastning på väginfrastrukturprojekt jämfört med andra infrastrukturprojekt. (Internationella Väg- och Transportfacket, 2000)

Kvaliteten på infrastruktur spelar en avgörande roll för utvecklandet av ett effektivt transportsystem. Stephan(2001) finner i sin fallstudie över ett antal tyska och franska delstater att kvaliteten hos infrastrukturen är en signifikant förklaringsvariabel för skillnaden i regional ekonomisk output. Bristande kvalitet i infrastrukturen anses också vara den huvudsakliga anledningen till Afrikas dåliga handelsfacit. (Lima och Venables, 1999)

2.2 Underhåll av vägar

I sitt arbete kring hur transportkostnader, till följd av dålig vägkvalitet, påverkar exporten från olika regioner i Columbia identifierar Blyde(2010) att förbättringar av vägkvaliteten har två huvudsakliga effekter. För det första ökar den medelhastigheten, och sänker därmed kostnader associerade med tidsåtgång. För det andra sänks kostnader för service och reparationer för fordonen som trafikerar vägen. Dessa faktorer har stor inverkan på vägburna transportsystems kostnader, och har därför en central roll i ett lands samhällsekonomiska utveckling. Detta då vägkvaliteten påverkar alla de aktörer som på ett eller annat sätt har en koppling till transportsystemen (Gwilliam, 2002).

Även om vägprojekt är de infrastrukturprojekt som i regel har den högsta ekonomiska avkastningen får inte alla projekt önskat resultat. Fan och Chan-Kang(2005) noterar att det i dessa fall ofta rör sig om bristande underhåll. Författarna medger att även om vanliga mått som den totala längden och densiteten av vägar är en användbar indikator på väginfrastrukturen i en region, visar dessa iakttagelser på vikten av att inkludera hur olika typer av vägkvalitet inverkar på avkastningen från ett väginfrastrukturprojekt. Detta får

stöd av Gwilliam(2002) som noterar att länder med lågt BNP per capita även i hög utsträckning lider av bristande vägunderhåll. Även Vita betonar i sin rapport vikten av underhållsarbete i utvecklingen och effektiviseringen av väginfrastrukturen. Dåliga vägar är en av de vanligaste anledningarna till stora trafikolyckor. Vägarbeten stör flöden och leder till förseningar, ökad bensinförbrukning och växthusgasutsläpp.

En genomtänkt planering kring underhållsarbetet innebär att arbetet kan bedrivas på ett effektivt sätt. Te Velde(2008) undersökte två scenarion över investeringar i Ugandas väginfrastruktur och dess effekt på den ekonomiska tillväxten. Både scenarion innefattar en ökad årlig investering med \$100 miljoner, dock med skillnaden att scenario 1 har oförändrad effektivitet och scenario 2 innefattar fördubblad effektivitet. Resultaten från studien indikerar på att scenario 2 skulle ge 50 % större positiv effekt på den årliga tillväxten, jämfört med scenario 1.

3. Socioekonomiska effekter

Det finns en allmän konsensus om att transportsektorn har en stor inverkan på en regions socioekonomiska utveckling, men inte hur denna inverkan ser ut under olika regionala förutsättningar. Detta beror på att det många gånger är svårt eller inte finns ett tillräckligt underlag för att kvantifiera dessa i direkta och indirekta effekter. Kunskapen om transportsektorns tillräcklighet har en stor inverkan på genomförandet och utformningen av regionala utvecklingsplaner och är en nödvändighet för socioekonomisk tillväxt och utveckling i ett modernt samhälle. (Umoren et al, 2009)

De socioekonomiska effekterna av förbättrad vägkvalitet skiljer sig beroende på länders olika geografiska och ekonomiska förutsättningar. I industrialiserade länder är väginfrastrukturen mer utbyggd och håller i allmänhet en högre standard. På grund av detta är potentialen för socioekonomiska effekter till följd av förbättrad vägkvalitet mindre för dessa länder då väginfrastrukturen redan möjliggör grundläggande samhällsekonomiska funktioner. I många utvecklingsländer är väginfrastrukturen inte lika utvecklad och många vägar är i så dåligt skick att det knappt är möjligt att trafikera dessa med vanliga fordon. Denna skillnad i förutsättningar gör att olika länder har olika behovsbild. (Gwilliam, 2002)

3.1 Utvecklingsländer

Då väginfrastrukturen i utvecklingsländer i de flesta fall är underutvecklad är potentialen av förbättrad vägkvalitet stor och kan få effekter inom en rad olika områden. Nedan följer en beskrivning av dessa.

3.1.1 Städer

I utvecklingsländer står den urbana sektorn för ca 50 % av BNP, detta då den ekonomiska agglomerationen samlar de flesta ekonomiska aktiviteter kring större städer. Ansvaret för infrastruktur faller ofta på de regionala myndigheterna, vilka i de flesta fall är institutionellt svaga, underbemannade och har begränsade ekonomiska resurser. Detta illustreras av att trafikpoliser ofta är involverade i trafikplanering eftersom inget civilt alternativ finns att tillgå. Detta i kombination med den kraftigt ökade urbaniseringen, samt att transportsektorn är än mer bunden till väginfrastrukturen i städer har lett till att transportsituationen ofta är mycket dålig. Det har inneburit ineffektiva transporter med långa köer och avgaser som följd. Urbaniseringen och brist på landyta har medfört att det inte är möjligt att öka vägkapaciteten i samma utsträckning som trafikvolymen. För att hantera denna typ av problematik är vägkvalitet bara en del i en större övergripande plan för att skapa ett fungerande transportsystem. (Gwilliam, 2002)

3.1.2 Rurala områden

I rurala områden är problematiken till viss del mycket lik den i urbana områden. Brister i vägkvaliteten begränsar transporten av varor, tjänster och människor. Investeringar i rurala vägar har i flera studier visat sig ha en stark inverkan på både jordbruksproduktionen och icke- jordbruksproduktionen. Detta minskar förlusterna av lättfördärliga jordbruksprodukter samt skapar bättre förutsättningar för småföretag, som i längden ökar behovet av arbetskraft i de rurala områdena. (Fan och Chan- Kang, 2005, Umoren et al, 2009)

3.1.3 Regionala skillnader

När bra vägar existerar blir det lättare för staten och privata aktörer att etablera ekonomisk aktivitet och sociala bekvämligheter. Det har visat sig i flera studier att områden med högre vägstandard tenderar att dra mer nytta av denna typ av satsningar än områden med sämre vägkvalitet. (Umoren et al, 2009)

Priset på varor kan variera regionalt på grund av transportkostnader. Minten och Kyle(1999) konstaterar i sin studie att vägkvalitet har stor inverkan på transportkostnader. Resultatet från studien visade att priset på mat var 2-3 gånger högre i områden där större delen av vägarna inte var belagda, jämfört med områden med högre andel belagda vägar. Detta kan förklaras med att belagda vägar i regel är av en betydligt högre kvalitet än icke-belagda vägar.

Under 1980-talet genomfördes ett projekt för att förbättra tillgängligheten till låginkomstområden i brasilianska städer. Många av dessa vägar var inte belagda och under regnsäsongen inte farbara för vanliga fordon, vilket starkt begränsade framkomligheten för bussar och utryckningsfordon. Projektet gick ut på att med en enkel och kostnadseffektiv metod belägga dessa vägar. Studien visade på en stark positiv socioekonomisk effekt, främst på grund av förbättrade busskommunikationer. Projektet har fått en rad efterföljare (Gwilliam, 2002). Dessa studier visar på att kvaliteten hos väginfrastrukturen är en betydande förklarande variabel för lokala skillnader inom en region, men även för skillnader mellan olika områden inom städer.

3.1.4 Fattighedsreduktion

Fattigdom anses till stor del bero på ett socialt utanförskap. I rurala områden utan en tillförlitlig transportinfrastruktur isoleras lokalbefolkningen från utbildning, sociala och ekonomiska möjligheter. Detta leder effektivt till så kallade ”poverty traps”, vilket är en självförstärkande mekanism som gör fattigdomen bestående. (Deichmann et al, 2000).

Fattiga hushåll lägger normalt 10- 30 % av sin totala inkomst på transport vilket innebär att många inte har råd att ta sig till de områden där den potentiella anställningen finns. Det är heller inte ovanligt att en arbetare spenderar tre till fem timmar om dagen på pendling till och från arbetet. Det går till viss del att fokusera på väginfrastruktursatsningar som gynnar de fattiga. Projekt som fokuserar på att förbättra vägkvalitet på rutten som skulle underlätta masstransporter och öka tillgängligheten till fattigare områden är ett sådant exempel. (Gwilliam, 2002)

3.1.5 Trafiksäkerhet

En halv miljon människor dör och 15 miljoner skadas bara i urbana trafikolyckor i utvecklingsländer varje år. Totalt sker 85 % av världens trafikrelaterade olyckor i dessa länder, och det är den nionde vanligaste dödsorsaken globalt. Det har i flera studier visats att faktorer som vägkvalitet och förändringar av lutningar vid så kallade black spots (extra olycksdrabbade platser) kan vara mycket effektiva åtgärder för att minska antalet allvarliga olyckor. Det är trots detta svårt att övertyga regeringar och andra beslutsfattande organ om att prioritera trafiksäkerhet. Det finns en tro om att den

mänskliga faktorns inblandning i olyckor gör att det inte går att genomföra kostnadseffektiva åtgärder för att förhindra olyckor. För att ändra detta synsätt krävs än tydligare metoder som visar på hur bland annat vägkvalitet inverkar på trafiksäkerheten. (Gwilliam, 2002)

3.1.6 Institutionell svaghet

Den institutionella svaghet som nästan alla länder med bristande väginfrastruktur lider av beror, förutom begränsade ekonomiska resurser, på bristande samarbete mellan beslutsfattande organ, bristande teknisk kunskap samt dålig funktionell och organisationell koordination. Uganda är ett exempel på ett sådant land. Där har vägnätet utvecklats utan någon egentlig struktur och underhållsarbetet har skett sporadiskt utan någon egentlig övergripande plan. Detta har starkt bidragit till att de flesta väginfrastrukturprojekt som genomförts de senaste åren inte har fått önskade resultat. (Kayondo-Ndandiko, 2012)

Faktorer som bidrar till institutionell svaghet är begränsade ekonomiska resurser och ombytliga politiska klimat. Detta innebär att satsningar på infrastruktur kan utebli om de styrande i ett land inte kan se den politiska nyttan med investeringarna. För att trygga kontinuerliga investeringar är det viktigt att visa på vilka effekter väginfrastrukturprojekt faktiskt kan få. Dessa fortlöpande investeringar är en förutsättning för att bygga upp ett långsiktigt fungerande vägnät. (Booth och Golooba-Muebi, 2009)

3.2 Industrialiserade länder

Då infrastrukturen och de institutionella strukturerna är betydligt mer utvecklade i industrialiserade länder kommer effekterna av ytterligare förbättringar av väginfrastrukturen inte vara lika stora som i många utvecklingsländer, då man redan åtnjuter de mest grundläggande fördelar som en fungerande väginfrastruktur ger. Detta betyder dock inte att det inte finns någon socioekonomisk potential i ytterligare effektiviseringar av väginfrastrukturen. Stephan(2001) påvisade att kvaliteten hos infrastrukturen var en betydande faktor i att förklara skillnaden i regional produktivitet i Frankrike och Tyskland, vilka båda har en väl utvecklad infrastruktur.

Internationella Väg- och Transportfacket(2000) betonar att eftersom vägtransport står för 90 % av världens passagerartrafik och 80 % av godstransporterna, får även mindre störningar i väginfrastrukturen stor effekt. Till exempel leder trafikstockningar till att bränsleförbrukningen ökar med en faktor 3, vilket i sin tur ökar utsläppen av växthusgaser. Inom EU uppgick kostnaderna, år 2000, till följd av störningar i väginfrastrukturen till 0,5 % av BNP, vilket motsvarar cirka €40 miljarder.

I takt med att världen blir allt mer global har konkurrensen mellan regioner ökat.

Vita(2008) anser i sin rapport att Europas konkurrenskraft till stor del beror på förmågan att hantera den predikterade 69-procentiga ökningen av transporterat gods. Det finns därför ett stort behov av att öka kapaciteten och effektiviteten på befintliga vägar. En viktig faktor i detta är att snabbt förbättra vägarna i Östeuropa så att de når samma standard som övriga Europa. Även om utbyggnader av den befintliga väginfrastrukturen är nödvändiga för att öka kapaciteten skulle en förbättring av vägstandarden leda till

mindre allvarliga olyckor och färre omfattande vägarbeten, vilket förutom att minska antalet döda i trafiken också skulle bidra till att minska köer och trafikstockningar.

3.3 Kina- ett land med stora skillnader

Även om det finns en stor skillnad mellan industrialiserade- och utvecklingsländers behov, kan de egenskaper som karaktäriserar dessa existera inom samma land. Kinas ekonomiska framväxt de senaste årtiondena har till stor del berott på stora investeringar i infrastrukturprojekt som stora motorvägar mellan viktiga städer. Dessa satsningar har dock varit fokuserade på städerna i kustregionen, vilket lett till att det idag är en mycket stor skillnad mellan landets olika provinser. De rika kustregionerna har till stor del en samhälls- och infrastrukturutveckling i nivå med de industrialiserade länderna, medan de fattiga provinserna i nordväst har mer gemensamt med situationen i många utvecklingsländer. (Deichmann et al, 2000)

Fan och Chan-Kang visar på att sentida stora högkvalitativa vägprojekt i urbana regioner har sett en minskad ekonomisk avkastning. Enklare projekt som att belägga rurala vägar har idag en högre ekonomisk avkastning, och hjälper fler människor ur fattigdom. Studien visar även att infrastrukturens satsningar fått olika effekt i olika delar av landet. De östra och centrala delarna har fått högst ekonomisk avkastning, medan bidraget till att minska fattigdomen varit störst i västra Kina. Detta menar författarna implicerar behovet av att formulera olika regionala policys, där de lokala beslutsfattarna får göra avvägningar och prioriteringar efter vilka aspekter de anser viktigast.

3.5 Sammanfattning

Förbättrad vägkvalitet kan inte ensamt lösa den problematik som diskuterats i detta avsnitt, det kan däremot skapa förutsättningar för andra ekonomiska och samhällsnyttiga funktioner att göra detta. För att skapa en effektiv väginfrastruktur, och därmed effektiva transportsystem, krävs stora satsningar efter en väl genomarbetad plan. Beskrivningen av hur olika typer av regioner påverkas av vägkvalitet är generell i sin natur för att tydliggöra hur regionernas olika förutsättningar påverkar vägkvalitetens inverkan på den socioekonomiska utvecklingen.

De skillnader som beskrivs visar att regioner har olika behov, vilket innebär att effekterna från väginfrastrukturprojekt varierar beroende på de specifika regionernas behovsbild. I Europa skulle till exempel bygget av en ny motorväg, med mycket hög vägstandard, mellan två viktiga städer vara den satsning som skulle ge bäst resultat utifrån de rådande förutsättningarna. I fallet med de fattiga områdena i Brasilien visade det sig att enklare belagda vägar hade en stark positiv effekt på socioekonomisk utveckling.

Denna typ av exempel visar på behovet av regionala policys rörande väginfrastrukturens satsningar. En ökad förståelse för väginfrastrukturens inverkan och effekter på olika regioner skulle underlätta arbetet kring att planera, prioritera och göra avvägningar mellan olika typer av väginfrastrukturprojekt. Det skulle även kunna avhjälpa problem med institutionell svaghet, genom att göra det enklare att förutse

effekter från satsningar på väginfrastruktur. Det skulle då bli lättare att säkra investeringar, då den politiska nyttan med investeringarna skulle bli tydligare.

4. Mätning av vägstandard

4.1 Ytojämnheter

Vägkvalitet mäts genom att studera ytojämnheter i en vägyta. Ytojämnheter ger upphov till vibrationer i fordon som färdas över vägytan. Dessa vibrationer har en negativ inverkan på körupplevelsen och kan få effekter som ökat fordonsslitage och drivmedelsförbrukning. Kraftiga vibrationer ökar risken för trafikolyckor och förlänger restider. Ytojämnheter mäts genom att kvantifiera den vertikala avvikelserna hos en verklig yta från dess perfekta form. När avvikelserna är små visar detta på att ytan är slät, medan större avvikelser indikerar en grov yta. (Ihs et al, 2006, Sayers och Karamihas, 1998)

En grov yta ökar slitaget på vägytan, då friktionen blir högre och fordonens fjädrande massa sätts i gungning och leder till en längre dynamisk resonans, som ger en längre belastning ner i vägbanan. (Sayers och Karamihas, 1998, Granlund 2009)

Ojämnheter kan även visa på andra brister i vägbanan, till exempel sprickbildning. Ett förebyggande underhållsarbete leder till minskade kostnader för reparationer av vägskador, då mindre reparationer inte kräver lika stora resurser som större skador. (Kayondo-Ndandiko, 2012)

4.2 Mätmetoder

Det finns ett antal olika metoder för att mäta vägkvalitet. Idag är International Roughness Index (IRI) den vanligaste metoden, även om Panel Ratings också används i viss utsträckning. Nedan följer en kortare beskrivning av dessa metoder.

4.2.1 IRI

Ojämnheter i en vägyta mäts inte i en specifik punkt utan som en funktion av ojämnheter över längd. Den teknik som främst används idag är IRI och uttrycks som ett standardfordons ackumulerade stötdämparrörelse dividerat med sträckan som färdats under mätperioden, uttryckt på formen mm/m eller m/km. IRI betecknar även det värde som metoden genererar. Metoden har en öppen mätskala mellan 0-20mm/m, där 0 indikerar en helt slät vägyta och värden över 14 indikerar en väg som inte är framkomlig med standardfordon. Insamlandet av vägprofildata görs med hjälp av en avancerad typ mätutrustning. Denna mätutrustning registrerar våglängder från 0,5 till 100m, detta sägs täcka in de våglängder som främst påverkar ett fordon's körning över en vägyta. Vägprofildata körs sedan genom en mjukvarubaserad kvartsfordonsmodell som simulerar en körning i 80 km/h. Eftersom mätningarna sker under samma förhållanden kan mätningar utföras med olika mätutrustning på skilda tider och platser och ändå vara jämförbara (Sayers och Karamihas, 1998)

4.2.2 Brister hos IRI

Mätutrustningen som används för att mäta vägprofildata är mycket dyr och involverar en komplex kalibreringsprocess som kräver hög kompetens och resurser. Det finns en fördröjning på ett par månader mellan när mätningen genomförs och när resultatet från denna är tillgänglig för beslutsfattare, eftersom vägprofildata måste köras genom kvartsfordonsmodellen. (Kumar, 2010, Sayers och Karamihas, 1998).

Komplexiteten och kostnaden hos mätprocessen innebär att mätningar normalt inte sker på kontinuerlig basis, på grund av resursbegränsningar. Det dröjer därför på många platser flera år mellan mätningarna. Dessa brister innebär att den information beslutsfattarna besitter ofta är daterad och inte beskriver den aktuella vägkvaliteten. (Booth och Golooba-Muebi, 2009)

4.2.3 Panel ratings

En annan metod som används för att mäta vägkvalitet är Panel Ratings. Dessa utförs vanligtvis genom att låta ett antal personer köra en sträcka och sedan fråga dem om hur de upplevde vägkvaliteten. Denna metod är direkt och kostnadseffektiv, men är högst subjektiv, då den är starkt korrelerad till individernas normala standard, åsikter och humör. Även utfrågningens utförande kan påverka resultatet, till exempel vilka instruktioner som ges och hur frågorna formuleras. Metoden ger en indikation av allmänhetens åsikter rörande vägkvalitet, men är inte lämplig för en övergripande eller jämförande analys av vägkvalitet. (Sayers och Karamihas, 1998)

4.4 Behovsbild

Begränsningarna hos IRI-metoden innebär att den inte ger ett bra underlag för beslutfattande rörande underhåll av vägar, då informationen ofta är daterad. Kostnaden och komplexiteten innebär att många länder använder enklare och mer kostnadseffektiva metoder för majoriteten av mätningarna av vägkvalitet. I Columbia använder man sig av panel ratings, som sedan ungefärligt översätts till IRI. I bland annat Indien används en primitivare form av responssystem, en så kallad Fifth wheel bump integrator¹, till majoriteten av mätningarna. Dessa metoder är fullgoda alternativ inom begränsade tillämpningsmetoder, men är inte jämförbara över tid och plats. (Sayers och Karamihas, 1998, Kumar, 2010, Blyde, 2010)

Flera studier, bland annat Booth och Golooba-Muebi(2009), har identifierat förebyggande underhållsarbete som en nyckelfaktor till att förbättra den övergripande vägstandarden i en region. Ett hinder för detta är att de inrapporteringssystem som finns, särskilt i de länder med bristande vägunderhåll, inte ger en bra bild av den aktuella vägstandarden. Detta leder till att underhållsarbete blir ineffektivt. Behovet av adekvata, enklare och framför allt mer kostnadseffektiva verktyg för att underlätta underhållsarbetet är därför stort.

Kumar(2010) menar att sådana verktyg skulle vara till hjälp i framtagandet av standarder för vägunderhåll och ligga till grund för ekonomiska analyser. Även Kayondo-Ndandiko(2012) pekar på vikten av bättre verktyg till stöd för underhållsarbete. Hennes arbete visar på hur mobila lösningar kan utnyttja geografiska informationssystem (GIS), vilket är en metod som används för att presentera geografispecifik data, för att skapa ett verktyg som kan ge en mer uppdaterad bild över den aktuella vägkvaliteten. Detta beräknas kunna sänka kostnaderna för datainsamling och bidra till ett mer

¹ Ett hjul monterat på en släpkärra kopplat till standardfordon via dragkrok. För mer information se Kumar, 2010.

kostnadseffektivt underhållsarbete. Eftersom arbetet med väginfrastruktur ofta sker med begränsade resurser skulle ökad effektivitet kring underhållsarbetet ge en ökad ekonomisk möjlighet att adressera andra brister relaterade till väginfrastrukturen.

4.5 Ny mätmetod- Roadroid

Roadroid är ett svenskt företag som utvecklat ett system för att mäta vägkvalitet. Systemet använder sig av den inbyggda accelerometer som finns i moderna smartphones för att mäta vibrationer i väglaget. Eftersom smartphones även är utrustade med GPS går data från mätningar att positionsbestämma. Dessa data matchas med digitala geografiska kartor och vägnät med GIS. Detta möjliggör en överblick över vägnät från en global nivå ända ner till en specifik punkt på ett vägavsnitt. (Forslöf, 2012)

4.5.1 Historik

Företaget grundar sig på ett forskningsprojekt som drevs i samarbete med Trafikverket och KTH mellan 2004 och 2006. En prototyp utvecklades där en accelerometer monterades på ett fordonets bakaxel, som kopplades till en laptop inne i fordonet, med extern GPS-sändare och GSM-modem. Resultaten från projektet visade på en stor potential, det fanns dock vissa tekniska begränsningar som mätutrustningens utsatta läge, prestanda och priset på mättekniken. 2010 hade utvecklingen av mobila plattformar kommit såpass långt att alla komponenter från det initiala systemet fanns tillgängliga i en enda mobil enhet, vilket föranledde att projektet reviderades och återupptogs. Grunderna för signalanalys och hastighetens inverkan på mätningen var områden där de flesta frågetecknen rätats ut redan under det initiala projektet. Under 2011 utvecklade företaget en applikation till det öppna mobila operativsystemet Android som möjliggjorde samma typ av mätningar som den ursprungliga prototypen. Applikationen innebar att kostnaderna för att utföra mätningarna minskade, samtidigt som användandet och underhållet av utrustningen blivit betydligt enklare. (Forslöf, 2012, Hildebrand, 2012)

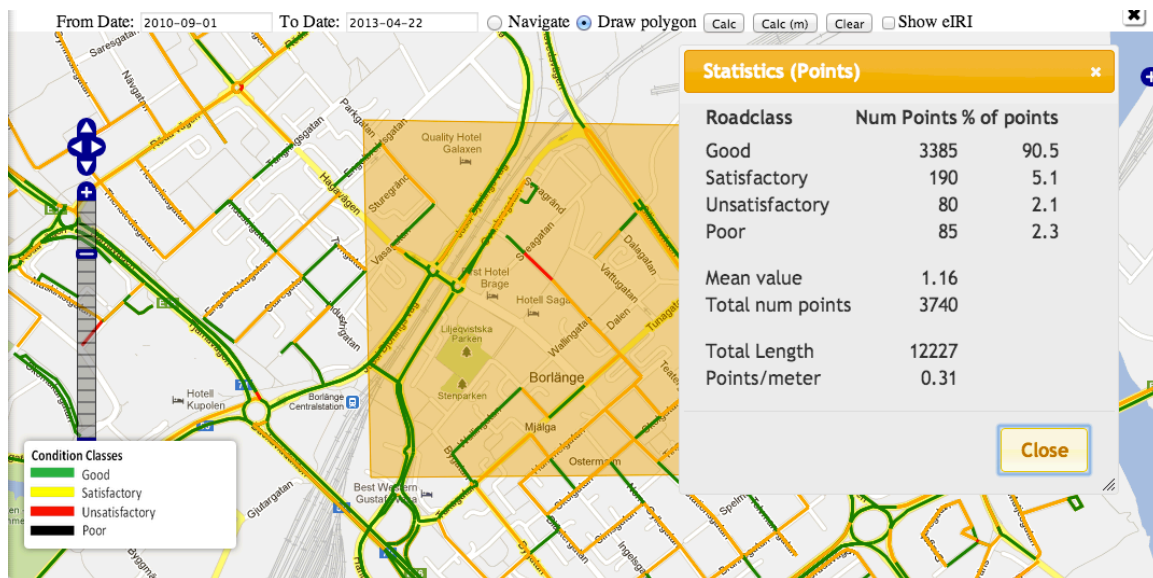
4.5.2 Teknisk beskrivning

En smartphone, utrustad med Roadroids applikation, monteras i ett bilfäste i fordonet. Applikationen registrerar de signaler som telefonens inbyggda accelerometer (en form av lägesgivare) samplar med en frekvens på 100Hz, detta summeras en gång per sekund till en mätpunkt med en estimerad korrelation till IRI. Applikationen använder sig av GPS, så att all data tilldelas en egen koordinat. Data kategoriseras grovt inom fyra olika kategorier², för att representera olika typer av vägkvalitet. Dessa kategorier är färgkodade, där grön betyder god, gul tillfredsställande, röd ej tillfredsställande och svart undermålig vägkvalitet. Den senare kategorin är den som har störst negativ inverkan på fordonen som färdas på vägen. Applikationen skickar sedan data till Roadroids webserver där data översätts till punkter i Google Maps. Eftersom enbart punkter på kartor skulle bli svåröverskådligt används Nationella Vägdatabasen i Sverige, och Open Street View globalt, så att en genomsnittlig standard för ett vägavsnitt kan visas. (Forslöf, 2012, Hildebrand, 2012)

² Kategoriseringen är ursprungligen gjord på engelska som Good, Satisfactory, Unsatisfactory och Poor. Dessa har sedan fritt översatts till svenska

4.5.3 Möjligheter med tekniken

Vägar där mätningar ofta genomförts har många fler mätpunkter än sträckor där mätningar enbart skett ett fåtal gånger. För att kunna jämföra olika sträckor på ett objektivt sätt presenteras resultatet därför även som procentandelar över de olika kategorierna. För att sammanställa vägkvaliteten för ett specifikt område markeras detta område på kartan med hjälp av ett polygonverktyg, på så vis går det att se sammanställningar av vägkvalitet på olika nivåer.



Figur 4.1 Ett område i centrala Borlänge markeras med polygonverktyget i Roadroids användargränssnitt. I rutan till höger visas statistik över vägkvaliteten för det markerade området.

I figur 4.1 illustreras hur Roadroids användargränssnitt ser ut, samt hur statistiken för det valda området presenteras. Detta verktyg ger en kontinuerlig överblick över kvaliteten hos väginfrastrukturen, vilket kan ge ett underlag som leder till effektivare underhållsarbete och i förlängningen förbättrad vägkvalitet. Det kan på sikt även ge ett övergripande statistiskt underlag rörande vägförhållanden i ett stort antal regioner och länder.

4.6 Spatial Data Mining

Data mining är en samling av metoder som går ut på att, utifrån stora mängder data, hitta mönster som sedan kan ge användaren viktig information som kan utnyttjas i dennes verksamhet. Spatial data mining är en gren inom data mining som fokuserar på att finna mönster i data med avseende på geografi. Den geografiska analysen utifrån GIS är inte särskilt utvecklad men har en stor potential att utvecklas till ett viktigt verktyg för beslutsstöd. (Shekhar, Zhang, Huang, 2005)

Det statistiska underlaget som Roadroid kan ge över regional vägkvalitet skulle kunna integreras med socioekonomisk data. Detta skulle göra det möjligt att utnyttja spatial data mining för att se effekterna av vägkvalitet på socioekonomisk utveckling, med avseende på olika länders geografiska och ekonomiska förutsättningar. Polygonverktyget skulle

kunna utvecklas till att innefatta förutbestämda regioner, så att data för vägkvalitet är tillgänglig i samma dimension som övrig data.

När man studerar komplexa samband, med många variabler inblandade, ger kontinuerliga variabler ofta svåröverskådliga resultat. Detta problem kan avhjälpas genom att gruppera de kontinuerliga data efter givna kategorier, där varje kategori anses spegla vissa karaktärsdrag. (Agresti, 1996)

Tabell 1. Data över vägkvalitet för olika länder från Roadroids mätningar integrerad med data över BNP/ per capita för de olika regionerna.

Country	Good (%)	Satisfactory (%)	Unsatisfactory (%)	Poor (%)	Region	GDP/capita ³	Income group ⁴
Sweden	91	5	2	2	Europe	57134	High
Spain	70	11	7	12	Europe	31820	High
South Africa	51	20	14	7	Africa	8090	Middle
Ethiopia	28	19	15	38	Africa	367	Low
India	60	16	10	14	Asia	1528	Low

I tabell 1 visas hur data för vägkvalitet över olika regioner kan integreras i en databas med publik socioekonomisk data för samma region. Kategoriseringen av inkomstgrupper, i kolumnen längst till höger, baseras på Världsbankens klassificeringssystem. Det finns egentligen två kategorier för medelinkomst, dessa är dock sammanslagna till en i detta exempel.

Denna databas skulle även kunna utvecklas till att innefatta flera nivåer. Sverige i tabell 1 skulle kunna bestå av en underkategori där vägstandarden för landets kommuner innefattas. Dessa nivåer skulle möjliggöra analys från global ner till lokal nivå med ett och samma verktyg, under förutsättning att socioekonomiska data finns tillgänglig för alla nivåer.

Även om den generella vägstandarden i ett land som Sverige är betydligt mycket högre än i säg Kambodja kan det finnas stora individuella avvikelser. Till exempel kan ett område i Gävle kommun ha samma vägstandard som Kambodja, eller så kan ett område i Kambodja ha en högre vägstandard än vad Sverige generellt har. Det är därför viktigt att kunna undersöka förhållanden på global så väl som på lokal nivå.

³ FN, 2013. *National accounts main aggregates database*. Tillgänglig online:<http://unstats.un.org/unsd/snaama/selbasicFast.asp>(2013-05-15)

⁴ Världsbanken, 2013. *How we classify countries*. Tillgänglig online: <http://data.worldbank.org/about/country-classifications>(2013-05-15)

5. Kategoriska data och multinominella logitmodeller

Metodiken för att hantera kontinuerlig data utifrån kategorier är vanlig inom samhällsvetenskapliga och biomedicinska studier, men har på senare år även blivit vanligare inom mer kvantitativa områden. Det finns en rad olika metoder för att hantera kategoriska data, i detta arbete används multinominella logitmodeller. Denna metod kan till exempel användas för att se hur valet av högskoleutbildning påverkas av resultat på högskoleprovet och vilket samhällsskikt personen i fråga tillhör.

5.1 Kategoriska variabler

En kategorisk variabel definieras som att skalan man mäter består av kategorier, detta skulle exempelvis kunna vara politisk orientering eller religiös tillhörighet. Statistiska regressionsmodeller undersöker hur en responsvariabel influeras av en eller flera förklarande variabler. Respons- och de förklarande variablerna kan vara såväl kategoriska som kontinuerliga, i många tillämpningar brukar dock en kategorisk variabel fungera som responsvariabel. Kategoriska variabler kallas även kvalitativa för att skilja dem från variabler med numeriska värden, som kallas kvantitativa (Agresti, 1996).

Det finns två typer av kategoriska variabler: ordinala och nominella. Ordinala kategorier har given ordning, till exempel lång, medel och kort. För nominella kategorier har ordningen av kategoriseringen ingen betydelse, eftersom det för dessa modeller inte påverkar resultatet. Ordinala variabler kan därför användas i nominella modeller, men inte tvärtom då ordinala modeller använder sig av ordningen i sina beräkningsprocesser. (Agresti, 1996)

5.2 Generaliserade linjära modeller

De kategoriska modellerna som används i detta arbete är en form av generaliserade linjära modeller (GLM). GLM består av en slumpmässig komponent som identifierar responsvariabeln Y , och antar en sannolikhetsfördelning efter denna. En systematisk variabel anger de förklarande variablerna och använder dessa som prediktorer. En så kallad länk beskriver det funktionella sambandet mellan den systematiska komponenten och det förväntade värdet hos den slumpmässiga komponenten. GLM ger en funktion för det förväntade värdet till förklaringsvariablerna genom en prediktionsekvation på linjär form (Agresti, 1996):

$$\begin{aligned}\mu &= E(Y) \\ g(\mu) &= \alpha + \beta_1 x_1 + \dots + \beta_n x_n\end{aligned}\tag{5.1}$$

där n betecknar antalet förklarande variabler.

5.3 Logitmodeller

Den traditionella linjära regressionsmodellen har en strukturell begränsning eftersom sannolikheter ligger mellan 0 och 1, medan linjära funktioner kan anta värden över hela reella axeln. Om länkfunktionen ser ut som $g(\mu) = \log(\mu/(1 - \mu))$ modelleras logaritmen av ett odds $\mu/(1 - \mu)$, en sådan modell kallas logitmodell. Sannolikheter

begränsas mellan (0,1), medan en logitmodell kan anta alla reella värden. Denna typ av modell medger ett betydligt större tillämpningsområde än den traditionella linjära regressionsmodellen, därför använder sig GLM-modeller ofta av logaritmiska odds istället för sannolikhet. För en regressionsmodell med en förklarande variabel x kan oddset formuleras som

$$\frac{\pi(x)}{1-\pi(x)} = \exp(\alpha + \beta x) = e^{\alpha}(e^{\beta})^x \quad (5.2)$$

För varje ökning av x ökar oddset med multipliciteten e^{β} . När $\beta=0$ sker ingen förändring vid ökning av x . Denna typ av tolkning är dock inte naturligt för de flesta och de logaritmiska oddsen omvandlas ofta till sannolikheter eller relativ risk när resultaten från modelleringen presenteras. (Agresti, 1996)

Multinominella modeller använder ofta relativ risk istället för odds. Inom enkel logistisk regression anses kategorierna vara ordnade då denna bara antar värdena 1 eller 0, och då inte är oberoende. Inom multinominella modeller antas kategorierna vara oberoende och inte ordnade. Relativ risk visar på risken eller sannolikheten att hamna i en given kategori jämfört med de andra kategorierna och kan i sin enklaste form beskrivas som $\frac{\pi_1}{\pi_2}$. (Hilbe, 2009)

Det analoga sambandet mellan odds och relativ risk kan skrivas som:

$$Odds\ ratio = \frac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)} = Relativ\ risk * (\frac{1-\pi_2}{1-\pi_1}) \quad (5.3)$$

Detta samband är användbart då det för vissa data inte går att beräkna relativ risk, men det är möjligt att beräkna oddskvoten och använda dessa för att approximera den relativa risken. (Agresti, 1996)

5.4 Multikategoriska logitmodeller

Den enklaste logitmodellen är den binomiala, vilket ofta benämns som logistisk regression. För att beskriva kategoriernas påverkan i modellen används dummyvariabler. Betrakta till exempel en modell med två förklarande variabler X och Z . Dessa är dummyvariabler och fungerar som binära prediktorer och indikerar kategorier för de förklarande variablerna. 0 och 1 representerar två skilda alternativ, att tillhöra gruppen som dummyvariabeln representerar eller inte. Denna modell kan skrivas som

$$logit(\pi) = \alpha + \beta_1 x + \beta_2 z \quad (5.4)$$

När antalet responskategorier överstiger 2 används en generalisering av den logistiska regressionsmodellen som kallas multikategoriska logitmodeller. För varje kombination av nivåer för de förklarande variablerna antas responsvariabelns kategorier vara multinominellt fördelade. Detta innebär att modellen kan referera till alla par av kategorier samtidigt, samt beskriva oddsen för varje respons i en kategori istället för de andra. I varje individuellt fall behandlas en kategori som responsvariabel, medan de andra variablerna ses som förklarande. För nominella responsvariabler paras varje

responskategori med en baselinekategori, som är godtycklig. När den sista kategorin J är baseline ser modellen ut som

$$\log\left(\frac{\pi_j}{\pi_J}\right) = \alpha_j + \beta_j x, \text{ för } j=1, \dots, J-1 \quad (5.5)$$

Modellen består av J-1 logitekvationer med separata parametrar för varje. Detta betyder att effekterna varierar till följd av vilken kategori som paras med baseline. För godtyckligt par kategorier a och b gäller:

$$\log\left(\frac{\pi_a}{\pi_b}\right) = \log\left(\frac{\pi_a/\pi_J}{\pi_b/\pi_J}\right) = \log\left(\frac{\pi_a}{\pi_J}\right) - \log\left(\frac{\pi_b}{\pi_J}\right) = (\alpha_a - \alpha_b) + (\beta_a - \beta_b)x \quad (5.6)$$

Alternativt kan (5.6) uttryckas direkt i form av respons sannolikheter:

$$\pi_j = \frac{\exp(\alpha_j + \beta_j x)}{\sum_{h=1}^{J-1} \exp(\alpha_h + \beta_h x)}, \text{ för } j=1 \dots J-1 \quad (5.7)$$

5.5 Betingade logitmodeller

I många fall tar en förklarande variabel olika värden för olika responskategorier. Hur mycket en person är beredd att betala för en ny bil varierar till exempel mycket beroende på vilket bilmärke och modell det handlar om. I multikategoriska modeller är de förklarande variablerna lika över alla responskategorier. Betingade logitmodeller⁵ är ett sätt att hantera detta. Med responskategori j, antalet förklarande variabler p och valmängd C_i , blir modellen för sannolikheten att välja val j:

$$\pi_j(X_i) = \frac{\exp(\beta_j X_{ij})}{\sum_{h \in C_i} \exp(\beta_h X_{ih})} \quad (5.8)$$

För ett godtyckligt par kategorialternativ a och b, kan modellen uttryckas som

$$\log[\pi_a(X_i)/\pi_b(X_i)] = \beta(X_{ia} - X_{ib}), \quad (5.9)$$

Oddset att välja a framför b är inte beroende av de övriga kategorierna eller värdena hos dess förklarande variabler. Denna egenskap kallas *oberoende av irrelevanta alternativ*. (Agresti, 2002)

5.6 Multinominella logitmodeller

Agresti(2002) sid 366, visar att Modell 5.7 är ett specialfall av 5.8 och kan genom omskrivningar skrivas som

$$\pi_j(X_i) = \frac{\exp(\alpha_j + \beta_j X_i)}{\exp(\alpha_1 + \beta_1 X_i) + \dots + \exp(\alpha_J + \beta_J X_i)} = \frac{\exp(\beta_j X_{ij})}{\exp(\beta_j X_{i1}) + \dots + \exp(\beta_j X_{iJ})} \quad (5.10)$$

⁵ Denna modell uttrycks inte på samma sätt som (5.7) vilket beror på att Agresti(2002) inte uttrycker sig på samma sätt som Agresti(1996)

Detta innebär att modell 5.8 kan hantera både kvalitativa och kvantitativa variabler, samt att de förklarande variablerna kan anta olika värden för olika responskategorier. Denna typ av modell kallas multinominell logitmodell och är till följd av nyss nämnda egenskaper mycket generell och kan användas inom ett stort antal tillämpningar.

5.6.1 Begränsningar

Multinominella logitmodeller har ett mycket brett tillämpningsområde, det finns dock vissa begränsningar som bör beaktas i arbetet med dessa modeller. Multinominella logitmodeller använder sig av maximum likelihood-skattning och de många parametrarna leder ofta till ekvationssystem. Detta innebär att dessa modeller är i behov av ett större antal datapunkter än binära och ordinala modeller. Tillämpningar där tillräcklig datamängd inte är tillgänglig kan påverka resultatets tillförlitlighet (Agresti, 2002, Hilbe, 2009).

Antagandet om oberoende av irrelevanta variabler är inte realistiskt i vissa tillämpningar. McFadden(1974) illustrerar detta genom valet av transport mellan en bil och en röd buss. Man antar att 80 % väljer att färdas med bil med ett odds på 4.0. Om man ändrar transportalternativen till att innefatta bil, röd buss och blå buss. Enligt 5.9 borde oddset att välja bil över röd buss fortfarande vara 4.0, men intuitivt förväntar vi oss att oddset bör vara 8.0 (10 % väljer varje bussalternativ). Tillämpningar bör därför vara begränsade till situationer där alternativ rimligen kan antas vara entydiga och oberoende ur beslutsfattarens synvinkel. Det finns andra modelltyper som kan minska effekten av antagandet av oberoende av irrelevanta alternativ. Exempel på sådana är probit- och nested-logitmodeller. Dessa modeller har dock andra begränsningar och behandlas inte vidare i denna rapport.

5.6.2 Diagnostik

Inom enkel logistisk regression finns det många metoder tillgängliga för att diagnostiskt utvärdera en modell för att finna outliers eller datapunkter med stor inverkan. För multinominella logitmodeller är det dock inte lika enkelt. Det är dock möjligt att dela upp modellen i separata logitmodeller och köra vanlig diagnostik på dessa.(Agresti, 2002, Hilbe, 2009)

För att testa om en enkel logistikmodell ger en bra anpassning till data används normalt hypotesprövning med Pearson χ^2 eller likelihood-ratio G^2 .

$$\chi^2 = \sum \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}, \quad G^2 = 2 \sum y_i \log \left(\frac{y_i}{\hat{\mu}_i} \right) \quad (5.11)$$

där y_i symboliserar de observerade värdena och $\hat{\mu}_i$ de anpassade värdena. Höga värden på dessa test visar på en dålig anpassning hos modellen. När anpassningen är dålig beskriver residualer och andra diagnostiska mått den påverkan individuella observationer har på modellens anpassning och belyser orsakerna till modellens otillräcklighet. (Agresti, 1996)

6. Modell

I den målspecifikation som utarbetades i samarbete med Roadroid bestämdes att den föreslagna modellen skulle:

- Utifrån givet geografiskt område matcha data över vägstandarden med samhällsekonomisk data för samma område.
- Kunna ge svar på hur vägkvaliteten förhåller sig mot givna förhållanden och ge indikationer på vilka effekter som kan uppnås med en förbättrad vägstandard samt vad/var man bör prioritera.
- Vara applicerbar både på lokal och global nivå.

Som ett första steg för att uppnå dessa mål skapades en databas där vägkvalitet och socioekonomisk data integreras enligt det presentationssätt som illustreras i tabell 1.

En multinominell logitmodell utarbetades i statistikprogrammet R, med utgångspunkt i den exempelkod som Hilbe, 2009, sid 407-409, presenterar för multinominella logitmodeller. Modellen har sedan utvecklats och anpassats för denna specifika tillämpning. R är ett program med öppen källkod vilket innebär att det finns en rad olika paket som kan utföra samma funktioner. I denna tillämpning valdes funktionen Multinom, från paketet nnet (CRAN, 2013). Detta då denna funktion inte kräver att data omformas och därför är smidigare att arbeta med än alternativa funktioner. Koden för den färdiga modellen är tillgänglig i appendix 1.

6.1 Val av tillämpning

Roadroids system har testats i ett 20-tal länder världen över. För att visa på modellens skalbarhet, det vill säga att den går att tillämpa på lokal så väl som global nivå, samt att utnyttja att mätningar gjorts i flera länder har ett globalt exempel valts för att illustrera den föreslagna modellen. Då denna rapport fokuserar på sambandet mellan vägkvalitet och socioekonomisk utveckling har ett exempel tagits fram för att belysa just detta samband. Eftersom undermåliga vägar är den vägkvalitetskategorisering som har störst negativ effekt på transportsystemens kostnader visar den föreslagna modellen hur ett lands välstånd påverkas av andelen undermåliga vägar, samt vilken region landet tillhör.

Tabell 2. De variabler ur databasen som används i den föreslagna modellen.

Country	Poor (%)	Percentage of poor roads	Region	HDI(0-1) ⁶
Norway	3	Low	Europe	0,955
Czech Rep.	10	Medium	Europe	0,873
Germany	1	Low	Europe	0,920
England	11	Medium	Europe	0,875
South Africa	7	Low	Africa	0,629
Ethiopia	38	High	Africa	0,369
India	14	Medium	Asia	0,554

⁶ UNDP, 2013. Human Development Report 2013. Tillgänglig online:
<http://hdr.undp.org/en/media/HDR2013%20Summary%20Swedish.pdf>(2013-05-05)

Sri Lanka	30	High	Asia	0,715
Malaysia	2	Low	Asia	0,769

I tabell 2 visas de variabler från databasen som används i den valda tillämpningen. Den fullständiga databasen inkluderar data för bland annat export, andelen döda i trafiken och BNP/capita. Databasen skapades i Excel, då den begränsade datamängden inte motiverade användningen av mer avancerade programvara för databashantering.

Som databasen är uppbyggd går det enkelt att tillföra ytterligare variabler som anses relevanta, så länge dessa data är i samma dimension som övrig data. Detta görs genom att utöka databasen med en extra kolumn. I den föreslagna modellen väljer användaren vilken data från databasen som modellen ska inkludera, utifrån vad denne anser vara relevant i den specifika tillämpningen. Data i denna tillämpning anses vara entydiga och oberoende och således uppfyller antagandet om oberoende av irrelevanta variabler.

6.1.2 Val av variabler

Som mått på välbstånd har Human Development Index (HDI) valts som indikator för detta. HDI är en sammanvägning av BNP, förväntad livslängd och utbildningsnivå. Detta mått används av bland andra FN och kan anses vara en standard när det kommer till att mäta välbstånd globalt. Precis som andra typer av globala mått har även HDI sina begränsningar, detta beror på att begreppet välbstånd är komplext och innefattar ett stort antal variabler. Det finns i dagsläget ingen användbar modell som kan hantera alla dessa på ett överskådligt sätt. Kritiken mot HDI handlar främst om att det i viss utsträckning går att manipulera uppskattningarna av de olika variablerna utifrån beställarens intressen, (Segerfeldt, Wallen, 2003)

Valet att använda Europa, Asien och Afrika som regioner är att det är de regioner som på förhand ansågs ha störst inbördes skillnad, samtidigt som de är de regioner där flest mätningar av vägkvalitet genomförts.

Data över andelen undermåliga vägar har delats in i tre kategorier, där low står för en andel undermåliga vägar på 0-10%, medium för 10-20% och high för en andel över 20 %, detta har gjorts för att tydliggöra vägkvalitetens påverkan som prediktor för välbstånd. Denna kategorisering har gjorts godtyckligt då inget vedertaget klassificeringssystem för detta existerar.

6.2 Begränsningar hos data

Roadroid är ett nytt företag och de mätningar som utförts har ofta varit begränsade i sin omfattning i många länder. Syftet i dessa fall har främst varit att ge lokala aktörer en möjlighet att testa systemet. För att genomföra detta exempel har den data som funnits tillgänglig fått representera en övergripande standard för hela landet. Av programmeringstekniska skäl har procentandelarna för de olika nivåerna av vägkvalitet avrundats till närmaste heltal.

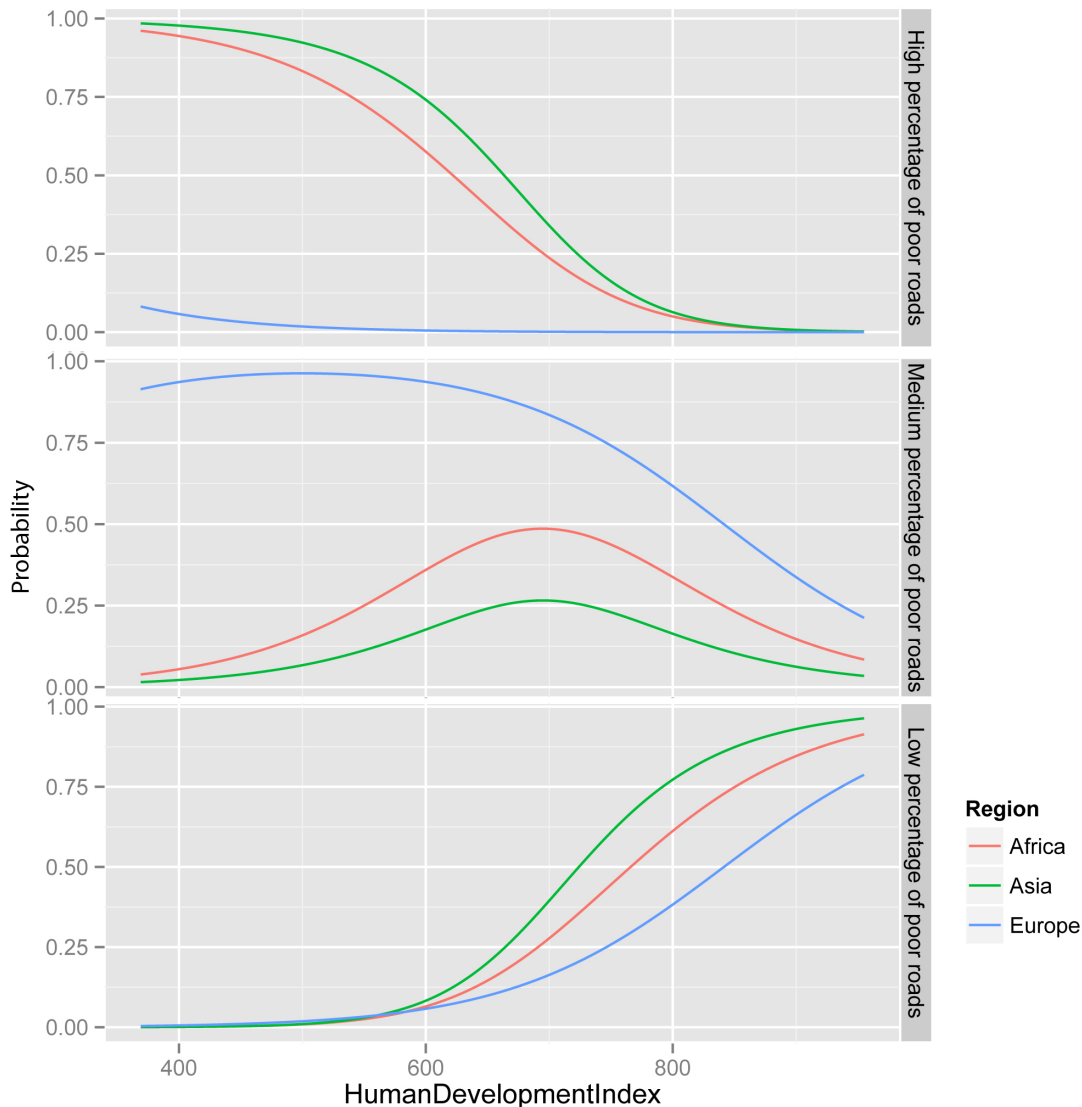
Som beskrivs i avsnitt 5.6.1 kräver multinominella logitmodeller ett relativt stort antal datapunkter. För att hantera detta har databasen utvidgats till att innefatta länder där

mätningar ännu inte genomförts. Med utgångspunkt i databasens socioekonomiska variabler har vägstandarderna i dessa länder antagits vara likvärdig med länder, där mätningar genomförts, som uppvisar liknande värden i övriga data. Dessa begränsningar påverkar tillförlitligheten hos resultatet från detta exempel. Det viktiga är dock inte att göra en statistiskt säkerställd analys, utan att visa på möjligheterna som finns med kategorisk data och multinominella logitmodeller.

7. Resultat

I detta avsnitt presenteras resultatet från tillämpningen av den multinominella modell som beskrevs i avsnitt 6. Resultatet diskuteras utifrån både en grafisk representation och en parameterskattning. Avsnittet syftar till att visa på hur denna typ av modell ska tolkas och hur den kan användas.

7.1 Presentation av resultat



Figur 7.1. Graf som visar sambandet mellan andelen undermåliga vägar och välstånd för länder i olika regioner.

Modellen i figur 7.1 visar sannolikheten för att uppnå ett visst värde på HDI- skalan för olika regioner, givet tre olika kategorier för andelen undermåliga vägar. Regionerna representeras av färger och varje delgraf visar sannolikheten för ett land i respektive region att ha ett visst HDI värde, givet en viss kategori av undermåliga vägar. I varje

punkt på x-axeln är den totala sannolikheten för varje färg i de olika kategorierna lika med 1. HDI mäts normalt mellan 0-1, med tre decimaler. I detta fall visas den skala 1-1000, då det blir mer överskådligt. Då inget land har ett HDI-värde under 300 har modellen skalats från 300-1000.

7.2 Tolkning av resultat

Från figur 7.1 går det att utläsa ett tydligt samband mellan välstånd och andelen undermåliga vägar i ett land. En hög andel undermåliga vägar motsvarar en mycket liten sannolikhet att uppnå ett HDI-värde över 800. En låg andel undermåliga vägar sammanfaller med en hög sannolikhet med höga HDI-värden.

Ser man till den delgraf som visar på en medelhög andel undermåliga vägar, det vill säga mellan 10-20 % av landets vägar, går det att urskilja hur regioner påverkas olika av bristande vägkvalitet. Ser man till Afrika och Asien är det sannolikt att länder i dessa regioner har ett HDI-värde runt 700, även om sannolikheten för detta är dubbelt så stor i Afrika. För Europa är det i denna kategori med mycket hög sannolikhet en indikator på låga HDI-värden, med en kraftigt avtagande sannolikhet för högre värden.

Ser man till hur stora sannolikheterna är för de olika regionerna ser man att det finns en skillnad för hur stor inverkan vägkvalitet har på välståndet. Dessa skillnader kan bero på ett stort antal olika variabler. Att ett land i Europa med en medelhög andel undermåliga vägar med 25 % sannolikhet kan ha ett HDI-värde på runt 900, medan det i Asien bara har någon procents chans att uppnå detta kan troligen förklaras av hur utbyggd övrig transportinfrastruktur är. Om till exempel tåg och tunnelbana är väl utbyggda så kan detta betyda att vägberoendet generellt inte är lika stort i denna region.

Som nämns i avsnitt 5.6.1 kräver multinominella logitmodeller ett större antal datapunkter för att utföra en tillförlitlig prediktion. Då mätningar i Europa främst har skett i regionens mer utvecklade delar finns det inga länder med hög andel undermåliga vägar, vilket förklarar utseendet av kurvan för Europa i den övre delgrafen. Fullständig data över regionen skulle även inkludera de mindre utvecklade länderna och eventuellt ge ett annat resultat.

7.3 Parameterskattning

Nedan följer en utskrift av parameterskattningen hos modellen vars resultat illustreras i figur 7.1. Detta för att bättre visa på hur den underliggande modellen fungerar

Call:

```
multinom(formula = Partpoor2 ~ Region + HumanDevelopmentIndex, data = land)
```

Coefficients:

	(Intercept)	RegionAsia	RegionEurope	HumanDevelopmentIndex
Low percentage of poor roads	-16.232545	-0.008242197	4.554504	0.02341673
Medium percentage of poor roads	-7.597558	-0.964734139	5.627277	0.01187973

Std. Errors:

	(Intercept)	RegionAsia	RegionEurope	HumanDevelopmentIndex
Low percentage of poor roads	1.012783	1.295931	0.7797361	0.001711880
Medium percentage of poor roads	2.435513	1.070606	0.7782435	0.004036704

Residual Deviance: 62.39987

AIC: 78.39987

Figur 7.2: Utskrift för parameterskattning av modellens output, uttryckt i logaritmiska odds

Under *call* i utskriften visas den formel som modellen använder sig av. För att återkoppla till den teoretiska beskrivningen i avsnitt 5 så motsvarar *Partpoor2* responsvariabel Y . *Region* och *HumanDevelopmentIndex* motsvarar de förklarande variablerna x_1 och x_2 . Parametervärdena för de förklarande variablerna motsvarar β_1 och β_2 och *intercept* betecknar α . Den övre matrisen i figur 7.2 visar koefficienterna för variablerna och den undre medelfelen för de samma.

Siffrorna i utskriften presenteras som logaritmiska odds. I detta fall har en hög andel undermåliga vägar valts som baseline. En ökning med en enhet i *HumanDevelopmentIndex* innebär ett ökat logaritmiskt odds med 0,0234 för att tillhöra en låg andel undermåliga vägar, jämfört med baseline hög andel undermåliga vägar. Ökningen av det logaritmiska oddset är något mindre för den medelhöga andelen undermåliga vägar, där ökningen bara är 0,0119 jämfört med baseline. Avläsningen av regionerna sker på liknande sätt, där regionerna jämförs med Afrika. Det logaritmiska oddset för att ha en låg andel undermåliga vägar ökar med 4,5545 om regionen ändras från Afrika till Europa.

Då logaritmiska odds, som det beskrivs i avsnitt 5.3, inte är särskilt intuitiva går det även att undersöka parameterskattningen genom att se samma utskrift fast som relativ risk, det vill säga risken eller sannolikheten att hamna ett alternativ i en kategori jämfört med baselinekategorin.

	(Intercept)	RegionAsia	RegionEurope	HumanDevelopmentIndex
Low % of poor roads	8.91857e-08	0.9917917	95.05961	1.023693
Medium % of poor roads	5.016750e-04	0.3810845	277.90449	1.011951

Figur 7.3: Utskrift för parameterskattning av modellens output, uttryckt i relativ risk

Utskriften över relativ risk i figur 7.3 avläses på samma sätt som i 7.2. det vill säga att en enhets ökning i *HumanDevelopmentIndex* ger en relativ risk på 1,0237 att tillhöra en låg andel undermåliga vägar jämfört med baseline hög andel undermåliga vägar. Den relativa risken att tillhöra en låg andel undermåliga vägar jämfört med baseline om regionen ändras från Afrika till Europa är 95,0596. Genom att presentera parameterskattningen på detta sätt blir det än mer tydligt att det finns en stor regional skillnad när det kommer till andelen undermåliga vägar, något som också illustreras i figur 7.1.

Residual Deviance och *AIC* (Akaikes informationskriterium), som syns i figur 7.2, är mått som används för att jämföra hur bra olika modeller är och eftersom detta inte görs i denna rapport diskuteras det inte vidare.

7.4 En modell - flera möjligheter

Den föreslagna modellen skulle även kunna användas till att studera hur andelen dåliga vägar påverkas av HDI-värde, genom att kategorisera HDI och använda kontinuerliga

värden för andelen undermåliga vägar. En annan möjlighet är att byta plats på de förklarande kategorierna, där varje delgraf visar en region och varje färg motsvarar en kategori av undermåliga vägar. Den föreslagna modellen ger på detta sätt en stor möjlighet att studera och visualisera kopplingarna mellan ett antal variabler på flera olika sätt beroende på hur användaren väljer att tillämpa modellen.

8. Slutsatser

Genom att applicera metodiken kring spatial data mining på data över socioekonomiska faktorer och vägkvalitet har det i denna rapport visats på en metod för analys för hur vägkvalitet inverkar på socioekonomisk utveckling. För att genomföra detta byggdes en enklare databas upp, som integrerar data för de olika variablerna. Genom att använda kategoriska data och en multinominell logitmodell har en metod som möjliggör analys över sambandet mellan variablerna utvecklats. Utifrån en målspecifikation över modellens tillämpningsområde valdes en tillämpning som visar på samband mellan välstånd och andelen undermåliga vägar ur ett globalt perspektiv.

Modellens resultat visar på en tydlig koppling mellan andelen undermåliga vägar och ett lands välstånd. Även om det statistiska underlaget till viss del är något bristfälligt är resultatet så pass tydligt att det går att dra slutsatsen att det finns en stark korrelation mellan vägkvalitet och välstånd i en region. Välstånd som i mångt och mycket är målet med socioekonomisk utveckling. Det mest intressanta resultatet är dock inte bara att det finns en stark korrelation mellan variablerna, men också att en medelhög andel undermåliga vägar begränsar välståndet inom ett visst intervall.

Den föreslagna modellen har utvecklats utifrån devisen att den ska vara användarvänlig, belägga komplexa samband på ett överskådligt sätt, samt vara applicerbar på flera olika nivåer. På sikt skulle denna modell kunna utvecklas till ett verktyg för att visa på vägkvalitetens inverkan och effekter under olika förutsättningar, vilket skulle bidra till att öka förståelsen för vägkvalitetens inverkan på samhällsutveckling. Detta verktyg skulle göra det lättare att prioritera och göra avvägningar kring vilka projekt som en region bör satsa på. Anses fattigdomsreduktion i rurala områden vara det som ses som viktigast för den regionala utvecklingen kan detta verktyg indikera på vilken typ av projekt som ger bäst önskad effekt för att uppnå detta mål. Detta skulle även underlätta arbetet med regionala policys, både i utformningen och uppföljningen av dessa.

En ökad förståelse för effekterna av väginfrastruktursatsningar skulle även göra det enklare att visa nyttan med investeringarna, vilket i sin tur skulle göra det lättare att motivera på det politiska planet. I länder med skiftande politiskt klimat skulle detta kunna användas för att säkra kontinuerliga investeringar i väginfrastrukturen och bidra till att bygga upp en långsiktigt fungerande infrastruktur.

Idag är det vanligt att privata aktörer är inblandade i väginfrastrukturprojekt och eftersom kostnaden för dessa projekt är mycket stora och många länder har begränsade ekonomiska resurser skulle en ökad inblandning av privata aktörer göra det möjligt att genomföra viktiga projekt som det inte finns utrymme för i statliga och regionala budgetar. Att kunna visa för företagen hur exempelvis deras transportkostnader skulle kunna sänkas genom bygget av en ny väg, med hög vägstandard, skulle göra det enklare att involvera dem i finansieringen av projektet.

9. Diskussion

Förutom vägkvalitet skulle även andra faktorer som påverkar transportsystemens effektivitet kunna analyseras med den föreslagna modellen. Till exempel skulle mått som längd och densitet av väginfrastrukturen, som normalt används som mått på hur utvecklad en regions infrastruktur är, kunna användas i modellen. Det skulle även vara möjligt att göra någon form av analys där dessa jämförs med vägkvalitet, för att närmare studera deras olika inverkan på väginfrastrukturens kvalitet. En sådan analys skulle även kunna ge indikationer på hur faktorerna samverkar, vilket skulle kunna användas för att utarbeta strategier för hur man på bästa sätt uppnår en effektiv väginfrastruktur.

I den tillämpning som analyserats har fokus legat på undermåliga vägar, detta för att göra kopplingarna till de socioekonomiska variablerna tydligare. Inom andra tillämpningar skulle även de andra kategorierna av vägkvalitet kunna användas. Ett sådant exempel skulle exempelvis vara en vidareutveckling av Stephans(2001) studie rörande regional ekonomisk output. I en sådan studie skulle alla fyra kategorier över vägkvalitet kunna inkluderas som förklarande variabler för den regionala ekonomiska outputen.

För att kunna utvecklas till en bra metod för stöd och investeringar rörande underhåll och investeringar i väginfrastruktur så behöver korrelationerna mellan de olika variablerna studeras närmare. Då tillräcklig data inte varit tillgänglig i detta fall har fokus främst legat på att visa hur en sådan metod kan se ut, och inte på en statistiskt säkerställd analys gällande till exempel modellval och modellens giltighet. Då mer data blir tillgänglig ökar möjligheterna till korrelationsanalys, där sambanden mellan vägkvalitet och samhällsekonomiska variabler ytterligare kan kvantifieras.

Den föreslagna modellen skulle initialt kunna fungera som en extra funktion i Roadroids nuvarande användargränssnitt. Detta skulle vara en lösning som under utvecklingsfasen av verktyget kan ge användarna en chans att se utvecklingsmöjligheterna samt ge feedback på hur verktyget kan förbättras. I takt med att verktyget utvecklas och databasen växer skulle ett fristående användargränssnitt kunna utvecklas.

I modellen används andelen undermåliga vägar som responsvariabel, vilket förklaras av HDI och region. Intuitivt så kan det tyckas att HDI borde vara responsvariabel och förklaras av de övriga två. Syftet har dock varit att visa på sambandet mellan vägkvalitet och socioekonomisk utveckling på ett överskådligt sätt som är enkelt för användaren att tolka. Som den grafiska presentationen av modellen ser ut ges en mer övergripande bild av sambandet mellan vägkvalitet och socioekonomisk utveckling om andelen undermåliga vägar fungerar som responsvariabel.

För flygtrafik har Flightradar 24 utvecklat ett användargränssnitt där användaren visuellt kan följa flygtrafiken över hela världen utifrån flygplanens positionssändare. Signalen från positionssändaren är del av flygsäkerheten och signalen kan fångas upp alla som har tillgång till en mottagare. Databasen är uppbyggd kring data som fångats upp av privatpersoner med tillgång till dessa sändare. Företagets site har cirka 10 miljoner besökare varje månad och har bidragit till att öka allmänhetens intresse för luftfarten. (Flightradar24, 2013)

Detta koncept skulle kunna användas även i detta fall. En web-domän skulle kunna införskaffas, där ett användargränssnitt med olika typer av data över väginfrastruktur och socioekonomiska faktorer kan studeras på både ett lokalt och ett globalt plan. Detta skulle kunna innefatta många fler sätt att presentera data på förutom det som föreslagits i denna rapport. En väl utvecklad databas ger möjlighet till flera olika metoder att presentera data på. Detta användargränssnitt skulle, om det utvecklats på ett bra sätt, bidra till att öka kunskapen och intresset för väginfrastrukturens betydelse för socioekonomisk utveckling. Precis som för Flightradar24 skulle användare uppmuntras till att bidra med relevant data för att ytterligare bygga ut databasen och öka möjligheterna till nya och spännande analyser.

10. Litteraturlista

10.1 Böcker

- Agresti, Alan. 1996. An introduction to categorical data analysis. Sid 1-2, 71-75,89, 106-107,111-113, 126, 205-208. Wiley-Interscience. Hoboken New Jersey
- Agresti, Alan. 2002. Categorical data analysis, Sid 298-300, 366. Wiley-Interscience. Hoboken New Jersey
- Hilbe, Joseph M. 2009. Logistic regression models. Sid 385-387,396-398, 407-409, Boca Raton, CRC press, Florida
- Shekhar, Shashi, Zhang, Pusheng, Huang Yan. 2005. Data mining and knowledge discovery handbook. Sid 833-851. Springer US, New York

10.2 Rapporter och vetenskapliga artiklar

- Blyde, Juan. 2010. Paving the road to export: the trade impacts of domestic transport costs and road quality. Inter-American Development bank. Washington, DC
- Booth, David, Golooba-Mutebi, Frederick. 2009. Aiding economic growth in Africa- The political economy of roads reform in Uganda. Working paper 307 Overseas Development Institute. London
- Deichmann, U., M. Fay, J. Koo, S. V. Lall. 2000. Economic structure, productivity, and infrastructure quality in southern Mexico. Policy Research Working Paper Series 2900. Världsbanken. Washington, D.C.
- Fan, Shebggen, Chan-Kang, Connie. 2005. Road development, economic growth and poverty reduction in China. Research Report 138, International food policy institute, Washington, DC
- Gwilliam, Ken. 2002. Cities on the move- A world bank urban transport strategy review. Världsbanken, Washington DC
- Ihs, Anita, Andrén, Peter, Mård-Berggren, Selina, Sjögren, Leif. 2006. Utformning av väg och vägyta för säkerhet, komfort och hälsa- En förstudie. Statens väg- och transportforskningsinstitut, Linköping
- Kayondo-Ndandiko, Lydia Mazzi. 2012. Geographical information technologies- Decision support for road maintenance in Uganda. Doktorsavhandling Blekinge tekniska högskola, Karlskrona
- Limao, N, A. J. Venables. 1999. Infrastructure, geographical disadvantage and transport costs. Policy Research Working Paper Series 2257. Världsbanken. Washington, D.C.
- McFadden, Daniel. 1974. Conditional logit analysis of qualitative choice behaviour. Sid 105-142 in Frontiers in Econometrics, Academic press. New York
- Sayers, Michael W, Karamihas, Steven M, 1998. Little book of profiling- Basic information about measuring and interpreting road profiles. University of Michigan. Ann Arbor
- Stephan, Andreas. 2001. Regional Infrastructure Policy and Its Impact on Productivity: A Comparison of Germany and France, Discussion Paper FS IV 01-02, Wissenschaftszentrum, Berlin

- te Velde, D.W. 2008 Economic Policy Scenarios for Addressing Growth Constraints in Uganda'. Background Note 1. DFID-funded Project on Political Incentives and Growth. ODI. London.
- Umoren, V, Ikurekong, E.E, Emmanuel, A, Udida, A.A. 2009. Development of road infrastructure as a tool of transforming Ibiono Ibom local government area. Global journal of social sciences vol 8, nr 2. Nigeria
- Vita, Laura. 2008. Road Infrastructure- The backbone of transport system. EU-kommissionen, Bryssel

10.3 Elektroniska källor

- Assafou, Aka. 2007, The importance of infrastructure investments in the urban agglomerations cities of Africa. Presentation International seminar on sustainable road financing and investment, Arusha, Tanzania. Tillgängligt online: <http://www.roadsfundtz.org/web/pdf/session%204/Aka%20Assafou.pdf>(2013-05-24)
- CRAN. 2013. Package "nnet". Tillgänglig online: <http://cran.r-project.org/web/packages/nnet/nnet.pdf>(2013-05-15)
- Fligtradar24. 2012. How it works. Tillgängligt online: <http://www.fligtradar24.com/how-it-works> (2013-05-23)
- Forslöf, Lars. 2012. Roadroid- Smartphone road quality monitoring.-final paper ITS world congress, Wien. Tillgängligt online: <http://www.roadroid.se/common/References/Roadroid%20Final%20Paper%20ITS%20World%20Congress%20Vienna%202012.pdf>(2013-05-24)
- Granlund, Johan. 2009. Konsekvenser av dålig veggeometri, Vectura, Norge. Tillgängligt online: http://www.vegvesen.no/_attachment/388988/binary/666275(2013-05-24)
- Hildebrand, Emelie. 2012. Smartphones mäter vägarnas kvalitet. Infrastrukturnyheter. Tillgänglig online: <http://www.infrastrukturnyheter.se/2012/05/smartphones-m-ter-v-garnas-kvalitet> (2013-05-04)
- Internationella Väg- och Transportfacket. 2000. Improving road infrastructure and traffic flows- IRU resolution, Bryssel. Tillgängligt online: http://www.iru.org/cms-filesystem-action?file=en_Resolutions_Mobility/00_infrastructure_e.pdf(2013-05-24)
- Kumar, Pradeep, 2010. Current practices for pavement condition monitoring of road infrastructure in India. Central road research institute, New Delhi. Tillgängligt online: http://www.arrb.com.au/admin/file/content13/c6/Pradeep%20Kumar_CRRI_India.pdf(2013-05-24)
- Segerfeldt, Fredrik, Wallen, Fabian. 2003. FN-statistiken felaktig. Artikel Svenska Dagbladet. Tillgängligt online: http://www.svd.se/opinion/brannpunkt/fn-statistiken-felaktig_99863.svd (2013-05-05)

II. Appendix

II.I R-Kod för den föreslagna modellen

```
land<-read.xls("Example2.xls")
names(land)=c("Country","Good","Satisfactory","Unsatisfactory","P
oor","Partpoor","Region","GDPcapita","Incomegroup","HumanDevelopm
entIndex","HDIgroup","Exportpercapita","Highwaycapita","Carsperca
pita","Killed","Deadintrafficismotor") # Läser in data från
databasen.
with(land,table(Partpoor,Region))# skapar en ny tabell med enbart
den önskade datan
land$Partpoor2<-relevel(land$Partpoor,ref="High percentage of
poor roads") # Väljer baselinekategori
test<-multinom(Partpoor2~Region+HumanDevelopmentIndex,data=land)#
körning av funktionen
Summary(test) # summerar test
exp(coef(test)) # Visar test som relativ risk
dcount <- data.frame(Region = rep(c("Asia","Africa", "Europe"),
each = 587), HumanDevelopmentIndex = rep(c(369:955),3))# beräknad
sannolikhet för olika värden av HDI för varje nivå av region
pp.write <- cbind(dcount, predict(test, newdata = dcount, type =
"probs", se = TRUE))# lagrar de beräknade sannolikheterna för
varje värde av HDI och Region
by(pp.write[, 3:5], pp.write$Region, colMeans) # beräknar
medelvärde för sannolikheten för varje nivå av Region
lpp <- melt(pp.write, id.vars = c("Region",
"HumanDevelopmentIndex"), value.name = "probability") # smälter
data set till format long för att anpassa för ggplot
ggplot(lpp, aes(x = HumanDevelopmentIndex, y = value, colour =
Region)) + geom_line() + facet_grid(variable ~ ., scales =
"free") #plot
```

11.2 Utdrag ur den databas som användes i modelleringen

Country	Good	Satisfactory	Unsatisfactory	Poor	Partpoor	Region	HDI	HDIgroup	Highwaycapita	Carspercapita	Deadintraffictomotor
Sverige	91	5	2	2	Low	Europe	916	Very High	24	462	7
Spanien	70	11	7	12	Medium	Europe	885	Very High	17	481	10
Tjeckien	70	13	7	10	Medium	Europe	873	Very High	5	427	20
Tyskland	95	3	1	1	Low	Europe	920	Very High	3	517	7
Storbritannien	63	17	9	11	Medium	Europe	875	Very High	6	457	7
Sydafrika	51	20	14	7	Low	Africa	629	Medium	8	112	209
Etiopien	28	19	15	38	High	Africa	369	Low	1	2	11667
Indien	60	16	10	14	Medium	Asia	554	Medium	3	12	315
Sri Lanka	39	17	14	30	High	Asia	715	High	5	20	221
Förenade Arab	93	5	1	1	Low	Africa	818	Very High	1	313	119
Malaysia	81	11	4	2	Low	Asia	769	High	3	325	37
Kina	80	9	5	6	Low	Asia	699	Medium	1	44	36
Kambodja	28	10	10	46	High	Asia	543	Medium	1	21	3200
Danmark	88	8	2	3	Low	Europe	901	Very High	13	390	6
Thailand	53	25	11	6	Low	Asia	690	Medium	1	67	119
Österrike	76	10	6	8	Low	Europe	895	Very High	25	529	12
Belgien	62	13	8	17	Medium	Europe	897	Very High	14	489	17
Turket	63	16	10	4	Low	Asia	722	High	6	104	97