

TRAFIKVERKET

EFFEKTSAMBAND

KONSUMERAD KAPACITET OCH TIDTABELLSTID

2018-02-09



wsp

SWECO 

EFFEKTSAMBAND

Konsumerad kapacitet och tidtabellstid

Trafikverket

KONSULT

WSP Analys & Strategi

121 88 Stockholm-Globen
Besök: Arenavägen 7
Tel: +46 10 7225000
WSP Sverige AB
Org nr: 556057-4880
Styrelsens säte: Stockholm
www.wsp.com

KONTAKTPERSONER

René Braune (Trafikverket)
Håkan Berell (WSP Analys & Strategi)
Sida Jiang (WSP Analys & Strategi)
Anders Lindfeldt (Sweco Norge)

PROJEKT

UPPDRAGSNAMN

Samband tågtidtabell och konsumerad kapacitet - nya effektsamband

UPPDRAGSNUMMER

10238500

FÖRFATTARE

Håkan Berell

DATUM

2018-02-09

ÄNDRINGSDATUM

INNEHÅLL

1	SAMMANFATTANDE SLUTSATSER	5
1.1	DATAFRÅGOR	5
1.2	STATISTISK ANALYS	5
1.3	MODELLESTIMERINGAR OCH SLUTSATSER	6
2	BAKGRUND	8
3	SYFTE	8
4	EFFEKTSAMBAND INOM TRAFIKVERKET	8
4.1	PROGNOSTIDTABELL FÖR JÄRNVÄGEN	9
4.1.1	Trafikering för person- och godstrafiken	9
4.1.2	Tidigare effektsamband	11
4.1.3	Nu gällande effektsamband	11
5	METOD	15
5.1	VAL AV DELSTRÄCKOR ATT STUDERA	15
5.2	KONSUMERAD KAPACITET	15
5.2.1	Linjedelar och dimensionerande sträcka	15
5.2.2	Beräkning av kapacitetsutnyttjande på enkelspår	16
5.2.3	Beräkning av kapacitetsutnyttjande på dubbelspår	17
5.3	HUR TIDTABELLEN KONSTRUERAS	18
5.4	RELATION MELLAN NU GÄLLANDE EFFEKTSAMBAND OCH NYA ESTIMERINGAR	19
5.4.1	Beroende variabel i estimeringarna	19
5.4.2	Omskrivning av gällande effektsamband	20
5.4.3	Estimering	21
5.5	STÄDNING AV DATA	22
6	ESTIMERING AV EFFEKTSAMBAND	25
6.1	ÖVERSIKTLIG PRESENTATION AV EFFEKTSAMBAND	25
6.1.1	Värmlandsbanan (VLB)	25
6.1.2	Norge-/Vänernbanan (NVB)	27
6.1.3	Södra Stambanan (SS)	31
6.2	RESULTAT FRÅN UTÖKAD MODELLSPECIFIKATIONER	33
7	RESULTAT	37
7.1	SLUTSATSER OM DATA	37
7.2	ESTIMERINGSRESULTAT	37
7.3	UTVÄRDERING AV RESULTAT	39
7.4	VIDARE ARBETE MOT EFFEKTSAMBAND	41
7.4.1	Vidare estimeringar	42
7.4.2	Hypotetiska experiment	42
7.4.3	Tidtabellsanalys	43
8	REFERENSER	44

1 SAMMANFATTANDE SLUTSATSER

1.1 DATAFRÅGOR

Projektet har siktat på att estimerar nya effektsamband mellan tidstillägg i tidtabellerna och konsumerad kapacitet med hjälp av data i tidtabells-databasen LUPP. Förhoppningen var att en statistisk ansats som bygger på de verkliga utfallen från tidtabellsprocessen skulle ge mer korrekta och mer tillförlitliga resultat.

Arbetet har emellertid till stor del präglats av svårigheter att förstå datat och med att producera tillförlitliga skattningar. För att studera hur stora tidstillägg som finns i tidtabellen krävs det inte bara uppgifter om den gällande tidtabellen, utan även uppgifter om den tekniska gångtiden för vart och ett av de olika tågen. Den tekniska gångtiden hämtades från Railsys.

I matchningen mellan Lupp och Railsys uppkom vissa problem. Andra problem hängde samman med att avstånden (linjedelarnas längder) i Lupp är bristfälliga, vilket på sina håll gav upphov till tillsynes orimliga hastigheter och tidsmarginaler. Avstånden byttes därför ut mot avstånden i BIS.

Det mest allvarliga problemet uppenbarade sig vid intervjuer med tidtabellskonstruktörer som påpekade att godstrafikens "kommersiella uppehåll" inte är tillförlitligt beskrivna i Lupp. Detta innebär att vi inte kan veta ifall ett långt uppehåll för ett godståg beror på att tåget står och väntar på att få plats på spåren för sin fortsatta färd eller om tåget står stilla för att det skall koppla till eller ifrån vagnar. Därmed kan vi inte veta om ett långt uppehåll är kapacitetsberoende eller inte.

Flera av de problem som uppkommit kunde ha undvikits ifall arbetet skulle ha utgått ifrån databasen Trainplan istället för kombinationen Lupp/Railsys. Enligt uppgift skall de "kommersiella uppehållen" vara bättre beskrivna i Trainplan. Om nya effektsamband skall tas fram med hjälp av en statistisk ansats rekommenderar vi att Trainplan används.

1.2 STATISTISK ANALYS

Några allmänna slutsatser från den statistiska analysen är följande:

- Det finns stora skillnader mellan de olika banorna

Variansen mellan de olika stråken är hög, vilket indikerar att ett nationellt genomsnitt för enkelspår respektive dubbelspår kan vara missvisande om man söker efter det "verkliga" utfallet av en kapacitetsinvestering.

- Det är stora skillnader mellan nu gällande effektsamband och de nya estimaten i denna studie

Detta indikerar att val av metod för att estimerar effektsambanden kan komma att ha en stor effekt på de samhällsekonomiska kalkylerna. Förändrade res- och transporttider är ofta den dominerande kalkylposten vid en kapacitetsinvestering.

- Det har inte varit möjligt att dra slutsatser angående godstrafiken eftersom det finns allt för stora osäkerheter

Det uppkommer stora tidspåslag på många linjedelar där förklaringen skulle kunna vara att tågen gör ett "kommersiellt uppehåll", men där detta inte framgår i Railsys.

1.3 MODELLESTIMERINGAR OCH SLUTSATSER

Regressionsanalys har genomförts med ett antal olika modellspecifikationer. Gemensamt för dessa är att en uppdelning har skett mellan gods- och persontrafik, men där det inte gjorts separeringar mellan typer av persontåg eller typer av godståg.

En första modellspecifikation har haft förändring i "minuter per mil" som beroende variabel och förändring av "konsumerad kapacitet" som enda förklaringsvariabel. Denna modell är enkel att illustrera grafiskt och är lätt att jämföra med de tidspåslag som Trafikverket använder idag. Den främsta nyttan av denna modell är att den är användbar för att studera och förstå datasetet samt hitta felaktigheter. Modellspecifikationen presenteras nedan, med närmare förklaring i kapitel 5.

$$(5.4') \quad \frac{T_{ij}}{D_i} = \alpha K_{ij} + \beta + \varepsilon'_{ij}.$$

Vid estimering av parametrar användes en modellspecifikation där förändring i "minuter" är beroende variabel och där förändring i "konsumerad kapacitet" och "linjedelens längd" är förklaringsvariabler. Denna modell är specificerad för att direkt kunna skatta de parametrar som Trafikverket använder idag. Modellspecifikation presenteras nedan, med närmare förklaring i kapitel 5.

$$(5.4) \quad T_{ij} = \alpha K_{ij} D_i + \beta D_i + \varepsilon_{ij},$$

Så långt har estimeringar gjorts för både gods- och persontrafiken. Eftersom det fanns stora osäkerheter i datamaterialet för godstrafiken (främst vad gäller "kommersiella uppehåll") gjordes fortsatta tester endast för persontrafiken.

För att testa om modellen (5.4) ger tillförlitliga estimat genomfördes regressioner med en fix-effekt modell. Detta innebär förenklat uttryckt att en dummy läggs till för att representera variabler som inte ingår i modellen men som eventuellt borde vara inkluderade. Modellspecifikation presenteras nedan, med närmare förklaring i kapitel 5.

$$(5.6) \quad T_{ij} = \alpha_0 + \alpha K_{ij} D_i + \beta D_i + \delta K_{ij} + \sum_i \gamma_i I_i + \varepsilon_{ij},$$

Där I_i är en indikatorvariabel för linjedel i och parametrarna γ_i är de fixa effekterna för linjedel i .

Ifall estimaten från modell (5.4) är pålitliga kommer koefficienterna inte att förändras i fix-effektmodellen (5.6) och parametern för dummysyn kommer inte att bli signifikant. Dummy-variablerna blir emellertid signifikanta och parameterskattningarna påverkas i hög grad. Detta indikerar att det inte är tillräckligt att bygga en modell för tidstillägg som en funktion av endast axelintercept, konsumerad kapacitet och delsträckornas längder.

Ovanstående slutsats är problematisk. Den kan indikera att statistisk analys och regressionsanalys inte är en framkomlig väg när det gäller att estimera robusta och tillförlitliga parametrar som beskriver sambandet mellan

konsumerad kapacitet och tidspåslag i tidtabellen. Trafikverket är emellertid beroende av att ha tillgång till ett effektsamband av denna typ. Det är en öppen fråga om de effektsamband som används idag är bättre eller sämre än de samband som kan tas fram med hjälp av statistik analys. I rapporten diskuteras ett antal olika sätt att föra analysen framåt. Både sätt att komma vidare med hjälp av alternativa specifikationer av modellen, men även alternativa metoder för att ta fram estimat av den typ som Trafikverket behöver.

Här måste det också övervägas möjligheten att ett dataset som bygger på Trainplan skulle kunna ge ett bättre resultat, vilket dock är långt ifrån säkert.

2 BAKGRUND

Effektsambandet mellan konsumerad kapacitet och tidspåslag i tidtabellen används i samband med samhällsekonomiska analyser av kapacitetspåverkande åtgärder inom järnvägsnätet. Effektsambandet har funnits med under lång tid och har under lång tid varit samma för persontrafiken och godstrafiken – vilket inte är rimligt. I samband med att en samhällsekonomisk studie av Norrbotniabanan togs fram uppmärksammades bristen med att effektsambandet för godstrafiken inte rimligen kan vara lika som för persontrafiken. En FUD-ansökan skickades till Trafikverket där bristen påtalades och det föreslogs att nya effektsamband skulle estimeras med hjälp av data från Trafikverkets databas LUPP.

Inom Trafikverket har det även startats upp ett internt projekt som syftar till att åstadkomma samma sak. En samordning har därför skett mellan det projekt som beskrivs i denna rapport och det Trafikverksinterna projektet. Det Trafikverksinterna projektet avser att estimeras ett effektsamband för enkelspår och dubbelspår för hela Sverige. WSP/Sweco-projektet har därför fått som uttalat mål att vara "metodsonderande" och fokusera på eventuella skillnader mellan olika järnvägsstråk. Detta för att avgöra om det är rimligt att ha ett och samma effektsamband för t.ex. alla enkelspår eller om det finns behov av att ta fram olika samband för olika stråk.

3 SYFTE

Syftet med projektet är att estimeras effektsamband som beskriver hur kapacitetsutnyttjandet (konsumerad kapacitet) påverkar de tidspåslag som behöver göras i tidtabellerna för att trafiken ändå skall vara robust och punktlig. Effektsambanden skall studeras separat för gods- och persontrafik samt mellan enkelspår och dubbelspår.

Eftersom det även finns ett annat projekt inom Trafikverket som avser att göra estimeringar för hela Sverige ligger fokus i denna rapport på att studera metodologiska frågor som kanske är mer tydliga när man studerar enskilda stråk samt att bidra med ett kunskapsunderlag till estimeringsarbetet i Trafikverksprojektet.

4 EFFEKTSAMBAND INOM TRAFIKVERKET

När Trafikverket gör samhällsekonomiska utvärderingar av olika väg- och järnvägsprojekt följer man en strikt och formaliserad process som syftar till att åstadkomma så träffsäkra kalkyler som möjligt och till att säkerställa att kalkylerna är jämförbara (så långt som möjligt). För att uppnå detta har man utvecklat en arbetsprocess där alla kalkyler skall göras med samma förutsättningar, vilket innebär att kalkylerna utgår ifrån en och samma prognos för framtiden, samma kalkylvärden och effektsamband samt att det

finns en godkännandeprocess där man kontrollerar att rätt förutsättningar har använts. Ett av de effektsamband som är av stor vikt i detta sammanhang är sambandet mellan konsumerad kapacitet och tidtabellstid.

När den konsumerade kapaciteten ökar tenderar tidtabellskonstruktörerna att lägga på mer tid i tidtabellerna för att upprätthålla punktligheten. Detta gör i sin tur att res- och transporttiderna blir längre. En järnvägsinvesterings samhällsekonomiska lönsamhet beror till stor del beror av hur många resenärer samt hur många tåg och nettoton gods som trafikerar respektive järnvägssträcka. Hur många resenärer som en järnväg kan attrahera beror i sin tur på vilka restider och hur många tåg per timme som kan erbjudas. Därav följer den samhällsekonomiska betydelsen av effektsambandet mellan konsumerad kapacitet och tidtabellstiden. Ju fler tåg som trafikerar en sträcka i jämförelsealternativet (JA), desto längre restider. Långa restider och stora kapacitetstidspåslag i JA ger stor potential till att skapa tidsvinster i ett utredningsalternativ (UA) där åtgärder vidtagits för att sänka den konsumerade kapaciteten.

Effektsambandet påverkar även godstågens tidtabellstid. I en samhällsekonomisk kalkyl brukar kortare transporttider för godstågen ge en större effekt via reduktioner av tidsberoende operativa transportkostnader än via tidsvinster för själva godset¹. Effektsambandet mellan konsumerad kapacitet och tidtabellstid kommer emellertid att påverka beräkningen av den konsumerade kapaciteten och därmed bedömningen av hur många tåg som kan tillåtas trafikera en viss järnväg i JA.

Effektsambandet mellan konsumerad kapacitet och tidtabellstid används främst när en prognostidtabell för hela Sveriges järnvägsnät skall tas fram. Det är denna tidtabell som ligger till grund för den basprognos som beskriver hur resandet förväntas se ut år 2040 i ett jämförelsealternativ där infrastrukturen är lika med dagens vägar och järnvägar plus de åtgärder som redan har beslutats. Utifrån denna JA-prognos gör man samhällsekonomiska beräkningar där man studerar de effekter som uppkommer i UA.

4.1 PROGNOTIDTABELL FÖR JÄRNVÄGEN

Riktlinjerna för hur man går tillväga vid framtagandet av en prognostidtabell för järnvägen finns i Trafikverkets dokument "Beräkningshandledning. Trafik- och transportprognoser, bilaga 1". Här följer en destillerad version som syftar till att beskriva de aspekter av processen som påverkar *hur* vi bör gå tillväga för att estimerar nya effektsamband. Målet är att de nya effektsambanden skall passa in i Trafikverkets existerande arbetsprocess med prognostidtabeller.

4.1.1 Trafikering för person- och godstrafiken

För persontrafiken tas det fram en trafikering som beskriver utbudet och där detta presenteras i "linjer". En "linje" är ett numrerat tågupplägg med ett visst antal turer per dygn, där alla tåg inom samma linje har samma uppehållsbild och samma tekniska gångtid. En ort kan vara servad av mer än en tåglinje. Orten Sundbyberg servas t.ex. av flera olika pendeltågslinjer (där skillnaden

¹ I den operativa kostnaden ligger löner och kapitalbindning i lok och vagnar. I tidsvinsten för själva godset ligger kostnaden för kapitalbindning i det som transporteras (varuvärdet).

mellan linjerna är hur långt pendeltågen går) samt flera regionalstågs- och fjärrtågslinjer.

Den tekniska gångtiden för persontågen är lika med restiden för persontåget (av viss fordonstyp) där tåget kör från start till mål samt startar och stannar vid de platser där linjen gör stopp för resandeutbyte. Detta betyder att de tidspåslag som beaktar graden av kapacitetsutnyttjande skall avse både de möten och/eller förbigångar som sker samt de påslag för "slack" som görs i tidtabellen för att förbättra punktligheten.

Godstågen är inte kategoriserade i "linjer" utan hanteras var för sig och redovisas som "enkelturer per dygn". **Den tekniska gångtiden för ett godståg** beräknas som en funktion av dragkraft och tågvikt där tågen framförs från start till mål med påslag för retardation och acceleration inlagt för det förväntade antalet tågmöten längs sträckan. Detta betyder att de tidspåslag som beaktar graden av kapacitetsutnyttjande skall inkludera den tid tågen står stilla vid möten och/eller förbigångar samt tidstillägg för "slack" som skall förbättra punktligheten.

När ett godståg har ett ärende till en viss station, t.ex. att det skall ske ett lokförarbyte eller att fem vagnar skall kopplas till, så uppkommer det en tidsåtgång som *inte* ligger i den tekniska gångtiden. Denna tidsåtgång är dock inte kapacitetsberoende och är konstant oavsett om det berörda tåget är ensamt på banan eller ej. Det blir viktigt att separera denna tidsåtgång från de tidspåslag som *är* kapacitetsberoende.

De **tågkategorier** som används är följande:

- Snabbtåg
- Övriga persontåg
- Lokaltåg
- Godståg

Vid beräkning av kapacitetstillägg och tidtabellstillägg används emellertid följande **tågkategorier**:

- Intercity
- Snabbtåg
- Snabbtåg stambanor
- Regionaltåg
- Natttåg
- Dieseltåg
- Pendeltåg storstäder
- Höghastighetståg
- Snabba regionaltåg
- Godståg

För var och en av dessa 10 tågkategorier finns det sedan tidigare ett effektsamband mellan konsumerad kapacitet och tidtabellstid. Det är dessa effektsamband som nu skall estimeras på nytt.

4.1.2 Tidigare effektsamband

De effektsamband som använts tidigare togs fram med hjälp av simuleringar av tågtrafik under åren 1992 och 1996. De ursprungliga effektsambanden har dock kommit att justeras i olika steg. En tidig version av effektsambandet såg ut enligt nedan.

Kapacitetstillägg min/mil

Kapacitets- utnyttjande	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %	100 %
Lågprio. persontåg	0,4	0,5	0,6	0,7	0,8	0,9	1,1	1,3	1,5	1,7
Högprio. persontåg	0,1	0,2	0,3	0,4	0,5	0,6	0,8	1	1,2	1,4

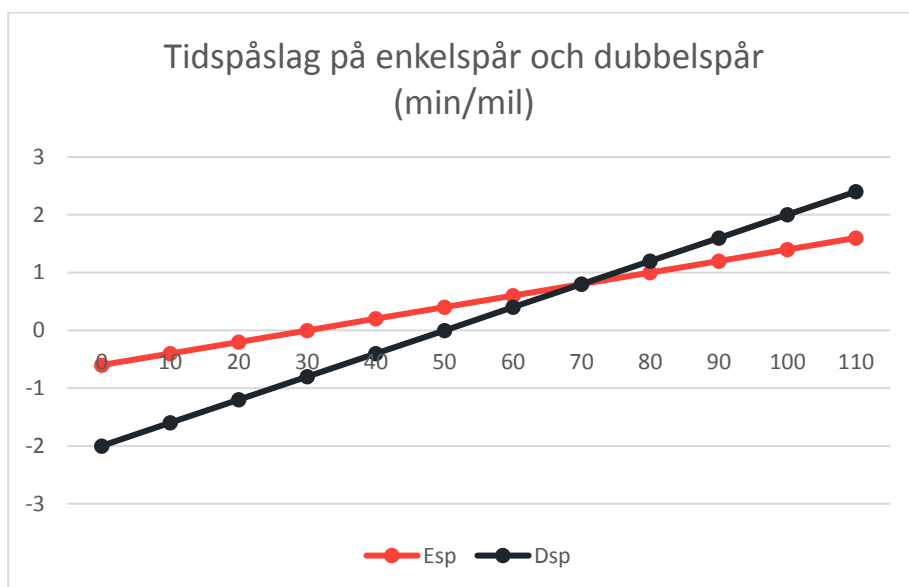
Tabell 1. Effektsamband mellan konsumerad kapacitet och tidtabellspåslag för persontrafiken före år 2005

Det ovanstående effektsambandet sågs över av Lennart Lennefors (år 2005) och därmed uppkom följande linjära samband:

*Enkelspår: Tidstillägg i minuter per mil = Konsumerad kapacitet i % * 2 – 0,6*

*Dubbelspår: Tidstillägg i minuter per mil = Konsumerad kapacitet i % * 4 - 2*

Effektsambanden togs ursprungligen fram för att beskriva effekterna för persontrafiken, men i brist på annat kom formlerna även att användas för godstrafiken. Se figuren nedan.



Figur 1. Tidspåslag i min per mil vid enkelspår respektive dubbelspår, samma påslag för gods- och persontrafik

Negativa tidstillägg görs inte och tidspåslag för högt kapacitetsutnyttjande börjar således tillämpas vid 30 % konsumerad kapacitet på enkelspår och vid 50 % konsumerad kapacitet på dubbelspår.

4.1.3 Nu gällande effektsamband

Problemet med att gods- och persontrafik inte rimligen kan representeras med samma linjära linje ledde år 2015 fram till att projektet Norrbotniabanan

utvecklade ett nytt samband för godstrafik på enkelspår. Den nya formeln togs fram av Göran Hörnell (WSP) med hjälp av metoden "tidtabellsanalys"². Detta samband har därefter inkluderats i Trafikverkets riktlinjer för hur prognostidtabeller bör utformas. Fler justeringar har gjorts av effektsambanden och bland annat har det tagits fram effektsamband för höghastighetståg.

I den instruktion som används för framtagande av prognostidtabell har tidtabellspåslagen delats upp i två delar:

$$\text{Kapacitetstillägg} = (\text{konsumerad_kapacitet} * \alpha_{mn} - \beta_m) * \text{avstånd}$$

$$\text{Tidtabellstillägg} = \gamma_j * \text{avstånd}$$

Där m betecknar enkelspår eller dubbelspår, och n betecknar tågtyp. Variablerna *Kapacitetstillägg* och *Tidtabellstillägg* mäts i minuter, *avstånd* mäts i mil och *konsumerad_kapacitet* är en andel mellan 0 och 1. Det totala tillägget är:

$$\text{Totaltillägg} = \text{Kapacitetstillägg} + \text{Tidtabellstillägg}$$

Kapacitetstillägget rymmer tidspåslag för tågmöten/förbigångar, slack av olika slag men även de avrundningar uppåt till närmaste hel minut som sker vid konstruktionen av tidtabellen (vilket i sig ger upphov till slack).

Tidtabellstillägget motsvarar ett förartillägg som beaktar att det i praktiken inte är möjligt för lokföraren att köra optimalt. Detta tillägg varierar mellan 0,1 till 0,3 min per mil och skiljer sig åt mellan tågtyperna.

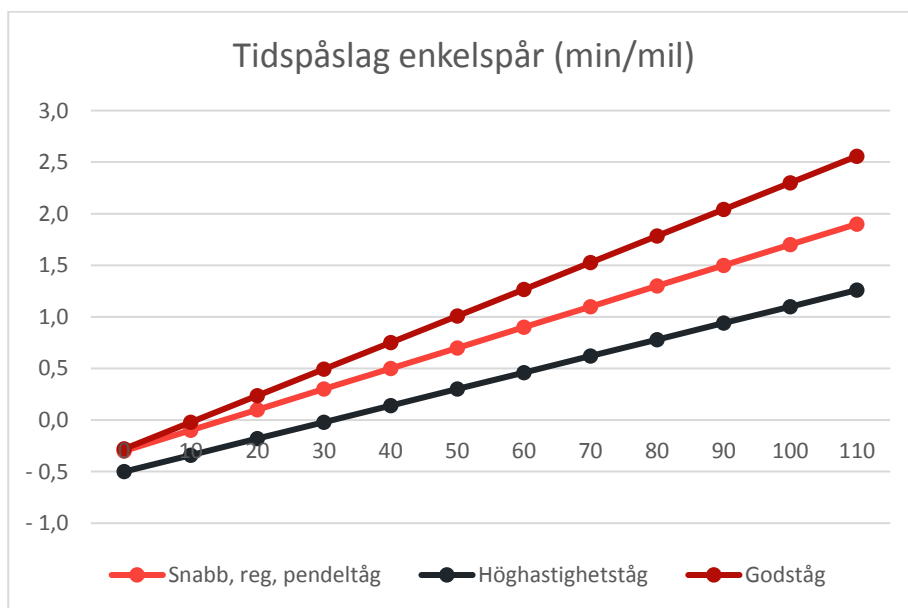
α och β är parametrar som tillsammans ger en linjär approximation av hur de kapacitetsrelaterade tidtabellspåslagen påverkas av graden av kapacitetsutnyttjande. De är anpassade så att alla tidspåslag, vare sig de uppkommer på linjen eller vid noder, fördelas över sträckan och redovisas i enheten "minuter". Utöver parametrarna α , β och γ finns också ett specifikt nodtillägg på 5 min/nod för nattåg och 15 sek/nod för pendeltåg i storstäder. Dessa parametrar i nu gällande effektsamband sammanfattas i Tabell 2 nedan.

² Tidtabellsanalys är här ett samlingsnamn på olika sätt att utifrån dagens tidtabell (en grafisk tidtabell) studerar kapacitetsrelaterade aspekter.

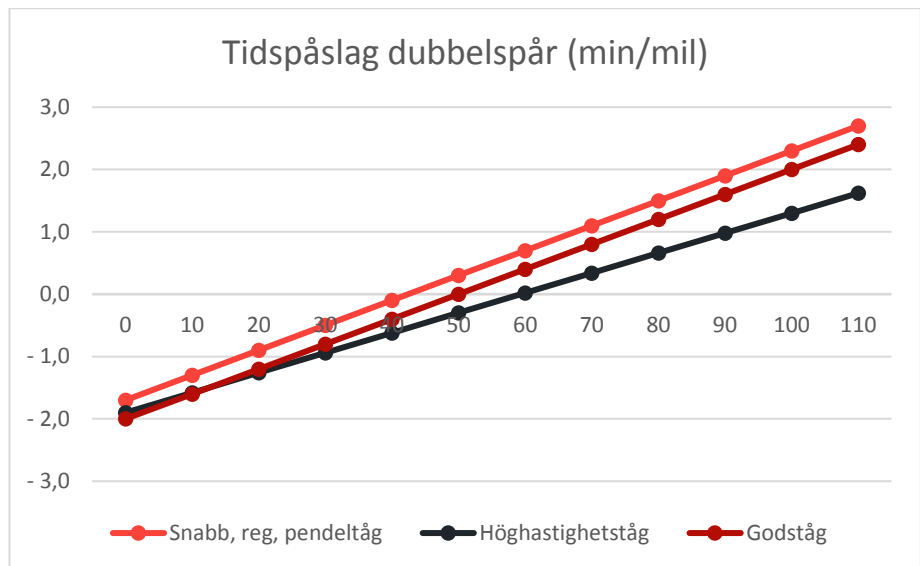
	Tidtabells- tillägg	Dubbelspår		Enkelspår		Nodtillägg
	γ	α	β	α	β	
<i>Intercity</i>	0,3	4,3	2,0	2,15	0,6	
<i>Snabbtåg</i>	0,3	4,0	2,0	2,0	0,6	
<i>Snabbtåg stambanor</i>	0,1	3,2	2,0	1,6	0,6	
<i>Regionaltåg</i>	0,3	4,0	2,0	2,0	0,6	
<i>Natttåg</i>	0,3	4,0	2,0	2,0	0,6	5 min
<i>Dieseltåg</i>	0,3	4,0	2,0	2,0	0,6	
<i>Pendeltåg storstäder</i>	0,3	4,0	2,0	2,0	0,6	15 sek
<i>Höghastighetståg</i>	0,1	3,2	2,0	1,6	0,6	
<i>Snabba regionaltåg</i>	0,1	4,0	2,0	2,0	0,6	
<i>Godståg</i>		4,0	2,0	2,58	0,28	

Tabell 2. Nu gällande effektsamband mellan konsumerad kapacitet och tidspåslag i tidtabellerna

De samband som presenteras i tabellen ovan kan beskrivas grafiskt för enkelspår och dubbelspår enligt nedan. Notera att negativa tidspåslag inte används och att kapacitetspåslag börjar göras först där effektsambandslinjerna skär 0-linjen.



Figur 2. Tidspåslag per mil på enkelspår för några typer av tåg, som funktion av konsumerad kapacitet i %.



Figur 3. Tidspåslag per mil på dubbelspår för några typer av tåg, som funktion av konsumerad kapacitet i %

5 METOD

5.1 VAL AV DELSTRÄCKOR ATT STUDERA

Projektet skall studera effektsambanden för både gods- och persontåg samt även studera hur dessa skiljer sig åt på sträckor med enkelspår och dubbelspår. För att säkerställa kvalitén i skattningarna har en avgränsning gjorts till tre stråk och valet har fallit på Norge/Vänerbanan, Värmlandsbanan och Södra stambanan. De egenskaper som eftersträvades var sträckor med:

- Både enkelspår och dubbelspår
- Varierande trafikbelastning, antingen över dygnet eller längs sträckan
- En blandning av tågtyper
- Inte alltför "styv" tidtabell för persontrafiken

Om persontrafiken följer en styv tidtabell kan man förvänta sig att tidspåslagen i minuter per mil ter sig oberoende av trafikbelastningen alternativt att det är trafikbelastningen i högtrafiktiden som bestämmer vilka tidspåslag som skall gälla under hela trafikdygnet. Eftersom det finns sträckor där styv tidtabell tillämpas kan det finnas anledning att ta särskild hänsyn till detta vid estimeringen av effektsambandet mellan tidspåslag och konsumerad kapacitet. När Trafikverket gör prognostidtabeller är man dock inte intresserad av att veta hur restiderna kommer att variera över dygnet. Därför ligger fokus i denna rapport på att beskriva hur möjliga restider beror av trafikbelastningen på dygnsnivå, vilket gör det rimligt att avgränsa bort det specialfall som styva tidtabeller innebär. Det är *möjlig* så att restiderna för tåg som går enligt styv tidtabell ändå har fått sina tidtabell fastställd utifrån den kapacitetsbelastning som gäller på banan.

5.2 KONSUMERAD KAPACITET

För att beräkna utnyttjad kapacitet används Trafikverkets modell för beräkning av linjekapacitet³. Metoden är en enkel modell där en befintlig tidtabell används till att beräkna den totala tiden som olika linjedelar är belagda av tåg. Beläggningstiden divideras därefter med analysperiodens längd för att erhålla hur stor del av perioden som linjedelen är belagd. Generellt sett anses linjesträckningar som har en beläggning på mindre än 60 % vara i balans. I intervallet 61–80 % kan det bli problem med att systemet är störningskänsligt. Är kapacitetsutnyttjandet över 80 % är systemet mycket störningskänsligt.

5.2.1 Linjedelar och dimensionerande sträcka

För att göra kapacitetsberäkningarna måste sträckningarna som analyseras delas upp i flera linjedelar. Enligt Trafikverkets modell skall indelningen i linjedelar göras utifrån trafikintensitet, trafikblandning och infrastrukturens beskaffenhet. En linjedel skal vara homogen med avseende på ovanstående parametrar och en ny linjedel skal skapas på platser där det sker en förändring.

³ Trafikverkets modell för beräkning av linjekapacitet, PM, Trafikverket 2016

En linjedel består normalt av flera stationssträckor. På enkelspår beräknas utnyttjad kapacitet för varje stationssträcka. Den för linjedelen dimensionerande sträckan blir den stationssträcka som har högst kapacitetsutnyttjande. På dubbelspår görs beräkningarna inte nödvändigtvis för varje stationssträcka. Utnyttjad kapacitet beräknas istället mellan stationer där förbigångar frekvent förekommer. Dimensionerande sträcka blir då den sträcka, mellan två stationer med frekventa förbigångar, som har högst belastning.

För både enkel- och dubbelspår är kapacitetsutnyttjandet på hela linjedelen densamma som kapacitetsutnyttjandet på den dimensionerande sträckan.

5.2.2 Beräkning av kapacitetsutnyttjande på enkelspår

På enkelspår beräknas tiden som varje stationssträcka är belagd (T_{bel}) enligt formeln nedanför. Den totala beläggningstiden beror på och ökar med fler tåg, längre körtider, om stationerna inte har samtidig infart och om sträckan inte har fjärrblockering. I detta projekt hämtas gångtiderna från Lupp. Information om stationerna har samtidig infart eller ej hämtas från BIS och alla banor har fjärrblockering. Tidstillägg på grund av möten hämtas från Trafikverkets modell för beräkning av linjekapacitet, PM (Trafikverket 2016) men finns även redovisade i tabellen nedanför. För att bestämma vilken kategori som tågen tillhör i tabellen nedanför, används gångtiderna från Lupp tillsammans med stationsavstånden från BIS till att beräkna tågens hastighet.

$$T_{bel} = \sum_{k=1}^{k=n} (T_{gång} + T_{möte} + T_{inf} + T_{fjb})_k$$

där:

- n : antal tåg som under tidsperiod för beräkning trafikerar spåret
- k : tågets ordningstal
- T_{inf} : tidstillägg vid ej samtidig infart till stationen, 2 minuter
- T_{fjb} : tidstillägg vid ej fjärrblockering på banan, 1 minuter
- $T_{gång}$: tågets gångtid på stationssträckan stationssträcka
- $T_{möte}$: tidstillägg på grund av tågmöte

Tågtyp	Hastighet [km/h]	$T_{möte}$ [minuter]
Snabbtåg	200	5
I/R-tåg	160 – 180	4
Lokaltåg	130 – 140	3
Godståg	70 – 160	5
Malmtåg	50 – 70	5

Tabell 3. Tidstillägg för tågmöten på enkelspår ($T_{möte}$)

Kapacitetsutnyttjandet per dygn ($K_{u,dygn}$) och i maxtimmen ($K_{u,max2h}$) beräknas enligt formlerna nedanför. Tidpunkten när max 2 timmen inträffar beräknas genom att beräkna kapacitetsutnyttjandet för samtliga 2 timmarsperioder under dygnet. Tidssteget mellan två på varandra följande två timmarsperioder är 5 minuter.

$$K_{u,dygn} = T_{bel} / (24 - T_{kval})$$

$$K_{u,max2h} = T_{bel}/(2 - T_{ban})$$

där:

T_{kval} : 6 timmar, kvalitetstid för förebyggande underhåll och felavhjälpning

T_{ban} : 0 timmar, tid för förebyggande underhåll, felavhjälpning och större banarbeten

5.2.3 Beräkning av kapacitetsutnyttjande på dubbelspår

Den totala belägningstiden beräknas enligt formeln nedanför. $T_{tåg}$ är det tidsutrymme som ett tåg tar upp. T_{konf} är ett tidstillägg som tillkommer när trafiken är heterogen, det vill säga att det finns tåg som kör med olika hastighet på sträckan. T_{kors} är ett tillägg som tillkommer för varje tåg som har en korsande tågväg på sträckan. Beräkningarna görs separat för båda riktningarna och den riktning som har högst kapacitetsutnyttjande blir dimensionerande för sträckan.

$$T_{bel} = \sum_{k=1}^{k=n} (T_{tåg} + T_{konf})_k + \sum_{j=1}^{j=m} (T_{kors})$$

$$T_{konf} = |T_k - T_{medel}|$$

där:

n : antal tåg som under tidsperiod för beräkning trafikerar spåret

m : antal tåg som kör korsande tågvägar under tidsperiod för beräkning

k, j : tågets ordningstal

$T_{tåg}$: utrymme i tidtabellen för tåg

T_{konf} : tidstillägg om tåg följs av ett tåg med annorlunda gångtid

T_{kors} : tidstillägg vid korsande tågväg per enskilt tåg, 4 minuter

T_k : gångtiden för tåg k

T_{medel} : genomsnittlig gångtid för alla tåg på sträckan

$T_{tåg}$ ges av tabellen nedanför och är beroende typ av tåg och tågets hastighet. Enligt "Trafikverkets modell för beräkning av linjekapacitet" (Trafikverket 2016) skall dock riktlinjerna i "Riktlinjer täthet mellan tåg – Planeringsförutsättningar"⁴ användas ifall finns definierade för aktuell sträcka. Riktlinjer finns definierade för Södra stambanan, Västra stambanan och den dubbelspåriga delen av Norge-/Vänerbanan och redovisas för aktuella delsträckor i tabellen nedanför. Efter att T_{bel} beräknats, beräknas kapacitetsutnyttjandet på samma sätt som för enkelspår.

⁴ Tågplan 2017, Trafikverket 2016

Tågtyp	Hastighet [km/h]	$T_{\text{tåg}}$ [minuter]
Persontåg	0 – 130	3
Persontåg	130 – 160	4
Persontåg	>160	5
Godståg	70 – 160	5
Malmtåg	50 – 70	5

Tabell 4. Utrymme i tidtabellen $T_{\text{tåg}}$ dubbelspår.

Sträcka	Minsta avstånd mellan tåg på linjen [minuter]
Järna-Flen	5
Flen-Katrineholm	4
Katrineholm-Åby	5
Åby-Lund	4
Öxnered-Trollhättan	5
Trollhättan-Göteborg	4

Tabell 5. Minsta avstånd mellan tåg på linjen för aktuella dubbelspåriga sträckor.

Beläggningstiden för korsande tågvägar kan ha betydande påverkan på den totala beläggningstiden på flera delsträckor. För att fastställa hur många gånger, och när på dygnet, som tåg korsar ned- respektive uppspår, görs speciella beräkningar på platser där korsande tågvägar kan förekomma. Exempel på platser med korsande tågvägar är platser där flera banor går ihop eller korsar varandra och stationer där många tåg vänder på sidospår.

5.3 HUR TIDTABELLEN KONSTRUERAS

När en tidtabellskonstruktör lägger en tidtabell så används systemet Trainplan. Från Trainplan överförs tidtabellerna till såväl systemet Railsys, där simuleringar av kapaciteten kan göras, och till systemet LUPP, som är ett uppföljningssystem där Trafikverket sammanställer trafik- och anläggningsinformation från andra källsystem (t.ex. den tidigare tågföringsdatabasen TFÖR, Bessy, Ofelia). I LUPP kan man således se både hur det var tänkt att tågen skulle framföras och hur det sen blev.

En tidtabell förändras över tiden fram till dess den träder i kraft. Detta innebär att uppgifterna om antal tåg för viss dag och hur de skall framföras beror av vid vilken tidpunkt tidtabellen avläses. I förevarande rapport har tåguppgifter hämtats från både LUPP och Railsys.

Från LUPP hämtas uppgifter om den planerade tågtrafiken (körtider enligt tidtabellen). För att kunna utläsa hur stora tidspåslag som lagts på – utöver den minsta möjliga körtiden – så behöver man även veta den tekniska gångtiden för varje tåg. Den tekniska gångtiden finns inte i LUPP men är däremot tillgänglig i såväl Trainplan som i Railsys. Det bästa hade varit att

använda sig av de tidtabeller som ligger i Trainplan. Denna databas var emellertid inte tillgänglig, varvid uppgifter om teknisk gångtid för gods- och persontrafiken hämtades från en befintlig och tillgänglig version av Railsys.

Detta har i sin tur gett upphov till två problem:

Antalet tåg för en viss dag är inte identiskt i LUPP och Railsys: Uppgifterna för trafiken är inte identiska i de båda databaserna, vilket förklaras av att information har överförts från Trainplan till LUPP och Railsys vid olika tidpunkter. En del tåg som varit inlagda i tidtabellen har tagits bort och andra tåg har tillkommit i ett sent skede. För att bevara så många observationer som möjligt har tekniska gångtider för liknande tåg använts (för persontågen är den tekniska gångtiden en funktion av tågtyp och för godstågen är gångtiden en funktion av vikten).

Railsys har inte tillgång till fullständiga uppgifter om godstrafikens tekniska gångtider: Under arbetets gång har det visat sig (i samband med intervjuer med tidtabellskonstruktörer) att den tekniska gångtiden i Railsys inte är fullständig vad gäller godstågens ärende under transportens gång. Det vill säga, det går inte att veta om ett längre uppehåll vid viss station beror på att tåget väntar på att få plats på spåret för att fortsätta sin färd (och väntetiden därmed är kapacitetsrelevant) eller om tåget skall koppla till eller ifrån vagnar vid denna plats. Detta är ett potentiellt stort problem! Fullständiga uppgifter finns i Railsys för persontrafiken. För att få fullständiga uppgifter om både gods- och persontrafiken borde databasen Trainplan användas.

5.4 RELATION MELLAN NU GÄLLANDE EFFEKTSAMBAND OCH NYA ESTIMERINGAR

5.4.1 Beroende variabel i estimeringarna

Regressionerna syftar till att studera hur tidtabellskonstruktören förändrar restiden för tåg på en viss linjedel när kapaciteten för linjedelen förändras. Regressionen skulle kunna göras direkt på tidtabellstiden $T_{tt,ik}$ för ett tåg k på linjedel i . En del av tidtabellstiden utgörs av den rena tekniska gångtiden $T_{gång,ik}$ för tåget och som är närapå konstant när kapaciteten förändras. Gångtiden är dock specifik för varje linjedel, det finns därför en poäng att istället för att använda $T_{tt,ik}$ som beroende variabel i regressionerna använda differensen mellan tidtabellstiden och gångtiden som beroende variabel, det vill säga

$$T_{ik} = T_{tt,ik} - T_{gång,ik},$$

Vi kallar denna tidsdifferens för tidstillägg. Fördelen med att använda tidstillägg som beroende variabel i regressionerna är att specifika faktorer som styr gångtiden och som varierar mellan linjedelarna inte behöver tas med i regressionen. Detta är ett sätt att homogenisera data. Den slutliga beroende variabeln som har använts i regressionerna är T_{ij} som är genomsnittet för T_{ik} över ett visst dygn (vardag) j .

Separata modeller estimeras för var och en av de tre stråken Södra Stambanan, Norge/Vänerbanan och Värmlandsbanan uppdelade på person- och godståg. Det finns behov av ytterligare differentiering efter typ av tåg, till exempel snabbtåg, intercity och regionaltåg samt godståg efter vikt, men detta har inte gjorts i detta projekt. I de estimeringar som beskrivs nedan

utgör en observation en kombination av linjedel (se avsnitt 5.2.1) och dygn (vardagar).

Att använda beroende variabeln tidstillägg homogeniserar data vilket leder till enklare specifikationer av regressionerna och säkrare parameterskattningar. Som visas längre fram i rapporten finns ändå problem med kvarvarande specifika effekter för linjedelar i modellen. Ett alternativ för att åtgärda det problemet är att estimerar regressionerna på differensen

$$\Delta T_{ij} = T_{ij} - T_{ij_0},$$

som oberoende variabel, där j_0 är ett visst godtyckligt valt referensdygn. Detta tar bort alla additiva specifika effekter för linjedelar från både data och modellspecifikationen. Det ger ytterligare en homogenisering av problemet. Detta har inte applicerats på regressionerna som redovisas i denna rapport, men det kan övervägas i eventuellt kommande regressioner.

5.4.2 Omskrivning av gällande effektsamband

Det nu gällande effektsamband som används för framtagande av prognostidtabell gavs i avsnitt 4.1.3 av följande två ekvationer:

$$(5.1) \quad \text{Kapacitetstillägg} = (\text{konsumerad_kapacitet} * \alpha_m - \beta_m) * \text{avstånd}$$

$$(5.2) \quad \text{Tidtabellstillägg} = \gamma_n * \text{avstånd}$$

Där m betecknar enkelspår eller dubbelspår, och n betecknar tågtyp. Utöver dessa två ekvationer fanns även ett nodstillägg. Ekvationerna (5.1) och (5.2) kan skrivas i en gemensam ekvation tillsammans med nodstillägget, vilket ger

$$(5.3) \quad \text{Totaltillägg} = (\text{konsumerad_kapacitet} * \alpha_m + (\gamma_n - \beta_m)) * \text{avstånd}$$

I de gällande effektsambandet finns även Nodstillägget på 15 sekunder/nod för pendeltåg kring storstäder, vilket är tämligen marginellt. Vid estimering har de 5 procent långsammaste tågen exkluderats. Denna utslutning täcker in de övriga tåg där detta nodstillägg är betydande, till exempel nattåg på Södra stambanan. Därmed har vi valt att bortse från nodstillägget och tar inte med det i någon ekvation. I alla ekvationer mäts avstånd i mil och kapacitetstillägg eller tidtabellstillägg i minuter. Om vi som i föregående avsnitt använder beteckningen T_{ij} för det genomsnittliga totala tidstillägget per linjedel i och vardagsdygn j , och låter konsumerad kapacitet betecknas med KK_{ij} och sträcklängd för linjedel (avstånd) med D_i , och multiplicerar in avstånd i parentesen i ekvation (5.3) så kan den ekvationen skrivas som

$$(5.4) \quad T_{ij} = \alpha KK_{ij} D_i + (\gamma - \beta) D_i$$

Denna ekvation blir utgångspunkten för estimering. För att förenkla uttrycken har index m och n tagits bort i ekvation (5.4), separata modeller kommer att estimeras för alla kombinationer av enkel-/dubbelspår och person-/godståg. Detta ger en viss skillnad mot nu gällande effektsamband i det att parametern $\gamma - \beta$ får ett unikt värde för alla kombinationer av m och n istället för den något mer restriktiva differensen $\gamma_n - \beta_m$.

5.4.3 Estimering

Generellt vid estimeringen så kan vi inte skilja på γ och β i differensen $\gamma - \beta$ i ekvation (5.4). Istället har en parameter β estimerats som ger den egentliga differensen $\gamma - \beta$. Så vid estimeringarna används regressionsekvationen

$$(5.4) \quad T_{ij} = \alpha K K_{ij} D_i + \beta D_i + \varepsilon_{ij},$$

som en utgångspunkt, där ε_{ij} är ett normalfördelat slumpfel. En variant av ekvationen där båda leden i (5.4) delas med sträcklängden D_i , är

$$(5.4') \quad \frac{T_{ij}}{D_i} = \alpha K K_{ij} + \beta + \varepsilon'_{ij}.$$

Denna ekvation ger samma estimerade parametrar som (5.4) men inte nödvändigtvis samma R^2 -värde. Ekvation (5.4') har en fördel gentemot (5.4) redovisningsmässigt genom att den endast innehåller en oberoende variabel. Därmed kan data och regressionslinje samtidigt visas grafiskt i ett diagram, vilket utnyttjas i avsnitt 6. Som utgångspunkt för vidare specifikation av modellen ses dock modellen i (5.4) som den naturliga varianten.

Modellen (5.4) har fördelen att den är enkel och att effektsambanden som ges av de två parametrarna α och β går att uttrycka som tidstillägg i minuter per mil vilket är enkelt att hantera i samhällsekonomiska kalkyler, till exempel kan man undvika att definiera linjedelar enligt avsnitt 5.2. Denna modellformulering har dock ett antal problem. Ett av dessa är att specifikationen i (5.4) saknar intercept. Det innebär till exempel att om konsumerad kapacitet är noll så måste alla tidstillägg vara distansberoende om (5.4) ska ge korrekt estimerade parametrar. Ett fall där detta inte är uppfyllt är vid korsande tågvägar. Eftersom de korsar den studerade linjedelen tar de upp mycket lite kapacitet samtidigt som de orsaker tidstillägg. Därför bör specifikationen i (5.4) utökas med ett intercept. Generellt är det också bra att estimeras en modell med intercept även om det inte kommer att användas vid tillämpning av modellen, annars kan rena tillfälligheter i specifika data ge felaktiga värden för parametrarna av intresse. Ett intercept införs i den utökade specifikationen

$$(5.5) \quad T_{ij} = \alpha_0 + \alpha K K_{ij} D_i + \beta D_i + \delta K K_{ij} + \varepsilon_{ij}.$$

I den specifikationen har även en huvudeffekt δ för konsumerad kapacitet inkluderats. Om parametern δ är statistiskt signifikant så blir α i modell (5.4) inte korrekt estimerad. Den fångar i så fall upp nivåskillnader i konsumerad kapacitet mellan linjedelarna. Därför är specifikationen i (5.5) mer robust än (5.4). Det är värt att notera att detta innebär ett mer komplicerat effektsamband att använda. Bland annat blir det mer kritiskt att korrekt definiera linjedelar enligt Trafikverkets metod för att beräkna konsumerad kapacitet (se avsnitt 5.2).

Sträcklängd D_i i ekvation (5.5) varierar enbart mellan olika linjedelar i , den är den dessutom den enda variabeln som beskriver olika linjedelar. Det är en kritisk egenskap i modellen given av (5.5). Om det finns ytterligare variabler som beskriver linjedelarna och som påverkar tidstilläggen T_{ij} , så kan det faktum att de inte ingår i modellen påverka estimeringen av parametrar på samma sätt som uteslutande av parametern δ påverkar modell (5.5).

Förekomsten av eventuella sakande variabler som beskriver linjedelar kommer att testas genom att inkludera specifika additiva effekter för linjedelar, även kallad fixa effekter. Det innebär att modell (5.5) bör utökas enligt

$$(5.6) \quad T_{ij} = \alpha_0 + \alpha K_{ij} D_i + \beta D_i + \delta K_{ij} + \sum_i \gamma_i I_i + \varepsilon_{ij},$$

Där I_i är en indikatorvariabel för linjedel i och parametrarna γ_i är de fixa effekterna för linjedel i . De fixa effekterna motsvarar att varje linjedel i modell (5.6) har sitt eget specifika intercept.

5.5 STÄDNING AV DATA

Steg 1: De första estimeringarna visade på väldigt långa tidspåslag per mil på Norge/Vänerbanan och Värmlandsbanan. En närmare undersökning visade att detta hängde samman med att en hög andel av tågen hade en genomsnittlig hastighet på över 200 km/h – vilket i sin tur visade sig bero på felaktiga avståndsangivelser i LUPP. När delsträckornas längder korrigerats med data från BIS blev det bättre, men det finns fortfarande observationer med genomsnittliga hastigheter över 200 km/h. Detta gäller alla de tre banor som vi studerat. Vi har tagit bort hastigheter över 200 km/h (4 % av observationerna) och även tagit bort negativa hastigheter (ca 2 % av observationerna).

Steg 2: Eftersom observationerna även beror av indata från Railsys var det nödvändigt att "felsöka" i Railsys. En av de egenheter som dök upp handlar om "timing-points". Railsys är av betydelse eftersom det är från detta system som den tekniska gångtiden hämtas. Emellertid finns det inte en teknisk gångtid noterad för varje station/tidtabellspunkt (se den kolumn 4-5 i tabellen nedan).

Delsträcka	Länk	Tid i LUPP	TGT i Railsys	Tidspåslag
1	a-b	12	NA	-
1	b-c	8	NA	-
1	c-d	10	22	8

Tabell 6. Beskrivning av hur data ser ut beträffande teknisk gångtid (TGT) och hur det används.

När det saknas data för Teknisk gångtid (TGT) för viss station antas det att tidspåslaget skall fördelas över de föregående delsträckorna. I exemplet ovan skall 22 min fördelas över länkarna a-d. Men att det står NA i en kolumn kan dock även innebära att det saknas data och vi har kunnat se exempel där 320 sekunder skall fördelas över 7 delsträckor, vilket ger en extremt kort TGT och ett stort slack. Problemet kan identifieras via en beräkning av hastigheterna i Railsys. Vi har tagit bort alla observationer där genomsnittlig hastighet är högre än 200 km/h.

Problemet fanns inte på Värmlandsbanan, så där tappade vi 0 % av observationerna. Längs Norge-/Vänerbanan tappade vi -28 % av observationerna och längs Södra stambanan hela 44 % av observationerna från de korta delsträckorna Norrköping-Åby och Linköping-Norrköping. Observationer från övriga delsträckor längs SSB används.

Steg 3: En plats där det fortfarande uppkom problem var Charlottenberg, där det kunde identifieras att det fanns en positiv körtid från Cg till Cg. Detta hängde samman med Charlottenberg i LUPP är en (1) plats men att Charlottenberg i verkligheten och i Railsys är två olika platser (Charlottenberg och Charlottenberg gränsen). Avståndet mellan Cg och Cggr är förhållandevis långt. Det finns fler platser där det i LUPP bara finns 1 station men där verkligheten ser annorlunda ut (t.ex. Laxå). Problemet har hanterats genom att observationer för delsträckor som Cg-Cg och Lå-Lå har tagits bort ifall det i LUPP finns en positiv körtid "inom" stationen.

Detta påverkar antal observationer i marginell utsträckning (ca 1 %).

Steg 4: När ovanstående korrigeringar gjorts har vi ändå funnit det motiverat att ta bort de observationer som har de 5 % högsta och 5 % lägsta hastigheterna baserat på LUPP-beräkningarna.

Här tappas 10 % av observationerna och detta gäller både för godståg och persontåg.

Steg 5: Ett av de problem som vi fortfarande har är att vi inte vet säkert vad godstågen gör när de har långa uppehållstider. Väntar tågen på att få plats på spåret för att kunna fortsätta sin färd, eller skall de koppla till eller ifrån vagnar? Detta avgör ifall tiden är kapacitetsrelaterad eller ej. I Tidtabellsboken kan vi utläsa vad persontrafiken har för ärende, men vår intervju med Pär Stemark (som konstruerar tidtabeller på Södra stambanan) visar att Tidtabellsboken inte alltid fångar in godstrafikens "kommersiella uppehåll". På Södra stambanan visar det sig att ovanstående korrigeringar och rensningar inte räcker till. Där testar vi att även lägga till ett femte kriterium, där vi tar bort godstågsobservationer där tidspåslagen överstiger 20 min per mil.

Här förloras ca 14 % av godstågsobservationerna. Detta steg genomförts även för Värmländsbanans godståg för de långa tidspåslagen som är inlagda vid norsk gränsen. Det testades även att bort de godstågsobservationer där tidspåslagen överstiger 10 min per mil, men vi konstaterade att det finns risk att långa tidspåslag som är kapacitetsrelevanta blir exkluderade, därför höjdes spärren för exkludering till 20 min per mil istället. Detta värde bör diskuteras tillsammans med tidtabellskonstruktören från fall till fall.

Enligt uppgift skall Trainplan innehålla bättre uppgifter om godstrafikens kommersiella uppehåll.

Nedan ges några sammanfattande deskriptiva uppgifter för datamaterialet efter städning av data.

Spår- och tågtyp		Mean			
		tidstillägg (min)	avstånd (mil)	tpm (min/mil)	kapacitet (dygn-nivå)
Enkel	RST	3,9	4,2	0,8	0,6
	GT	16,1		3,5	
Dubbel	RST	4,5	3,6	1,6	0,6
	GT	8,6		1,9	

Tabell 7. Genomsnittlig statistik för tidspåslag uppdelat på spår- och tågtyp (RST: persontrafik, GT: godståg och Tpm: Tidspåslag per mil)

Spår- och tågtyp		Standard avvikelse			
		tidstillägg (min)	avstånd (mil)	tpm (min/mil)	kapacitet (dygn)
Enkel	RST	3,2	2,3	0,7	0,2
	GT	28,1		3,3	
Dubbel	RST	5,4	0,7	2,3	0,2
	GT	12,7		3,1	

Tabell 8. Standardavvikelse för tidspåslag uppdelat på spår- och tågtyp (RST: persontrafik, GT: godståg och Tpm: Tidspåslag per mil)

6 ESTIMERING AV EFFEKTSAMBAND

I det första avsnittet nedan presenteras effektsamband i grafisk form. För att med tillförlitlighet kunna skatta dessa samband måste, som togs upp i avsnitt 5.4.3, en mer utvecklad modellspecifikation användas än ursprungsmodellen (5.4) som kan ses som en direktöversättning av nu gällande effektsamband till en modellspecifikation. Estimeringsresultat som motiverar den utökade modellspecifikationen presenteras i avsnitt 6.2 nedan.

6.1 ÖVERSIKTLIG PRESENTATION AV EFFEKTSAMBAND

Nedan redovisas sambanden grafiskt mellan konsumerad kapacitet och tidstillägg (tidtabellstillägg) per mil. Tidstillägg per mil är en form för den beroende variabeln som följer ursprungsmodellen (5.4) från avsnitt 5.4.3. Den modellen har visat sig ge icke tillförlitliga parameterskattningar. Däremot ger denna transformation av den beroende variabeln en möjlighet att presentera data i ett tvådimensionellt diagram som ändå inbegriper de tre variablerna tidstillägg (T), konsumerad kapacitet (KK) och sträcklängd för linjedel (D).

Diagrammen visar framför allt spridning och struktur för data. Även om modell (5.4') inte ger ett slutgiltigt effektsamband så har regressionslinjen från den modellen lagts in i diagrammen. Den visar främst hur en regressionslinje kan anpassas till data utifrån de förutsättningar som gäller för modell (5.4'). Många av diagrammen nedan visar en tydlig tendens till klustring av data. Klustren motsvarar linjedelar som är grupper av observationer i data. Dessa kluster utgör en motivation till varför en modellspecifikation innehållande fixa effekter är nödvändig, se avsnitt 7.3.

6.1.1 Värmlandsbanan (VLB)

Den studerade sträckan går från Laxå vid Västra stambanan till Charlottenberg vid norska gränsen. Järnvägen är enkelspårig och trafikeras av både gods- och persontrafik. Vid analyserna har sträckan delats upp i fyra linjedelar där tågantalet är lika inom linjedelarna men olika mellan linjedelarna.

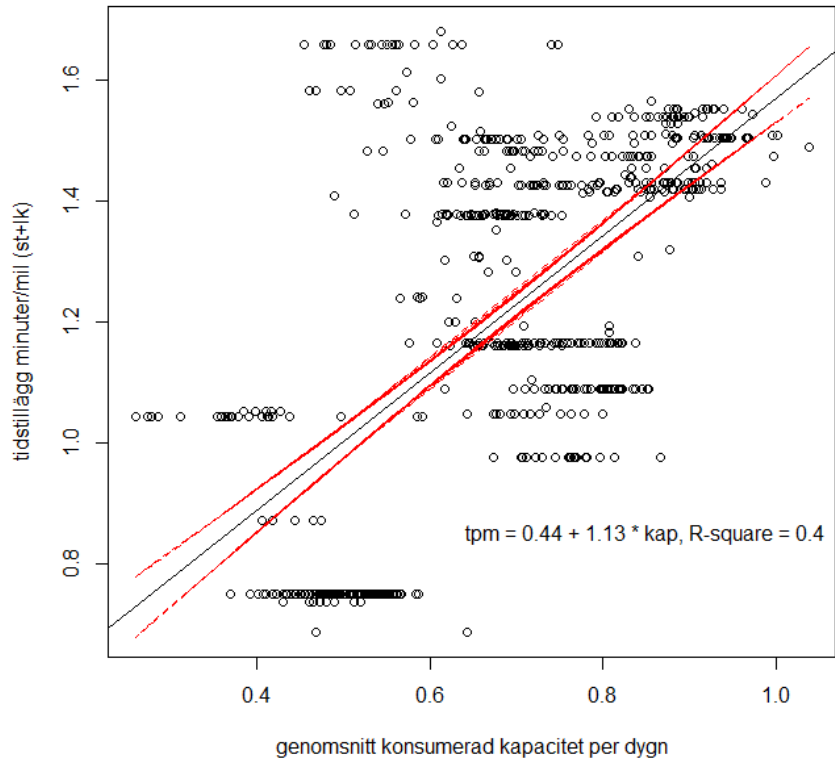
Tabellen nedan visar hur stor andel (%) av den totala tidtabellstiden som utgörs av "tid för möten/förbigångar och slack på VLB", uppdelat på enkelspår (Esp) och dubbelspår (Dsp).

VLB	Esp	Dsp	Esp+Dsp
Persontåg	22,3%	Gäller ej	
Godståg	39,8%		

Tabell 9. Tidspåslagens andel av den totala tidtabellstiden uppdelat på enkelspår, dubbelspår samt persontåg och godståg

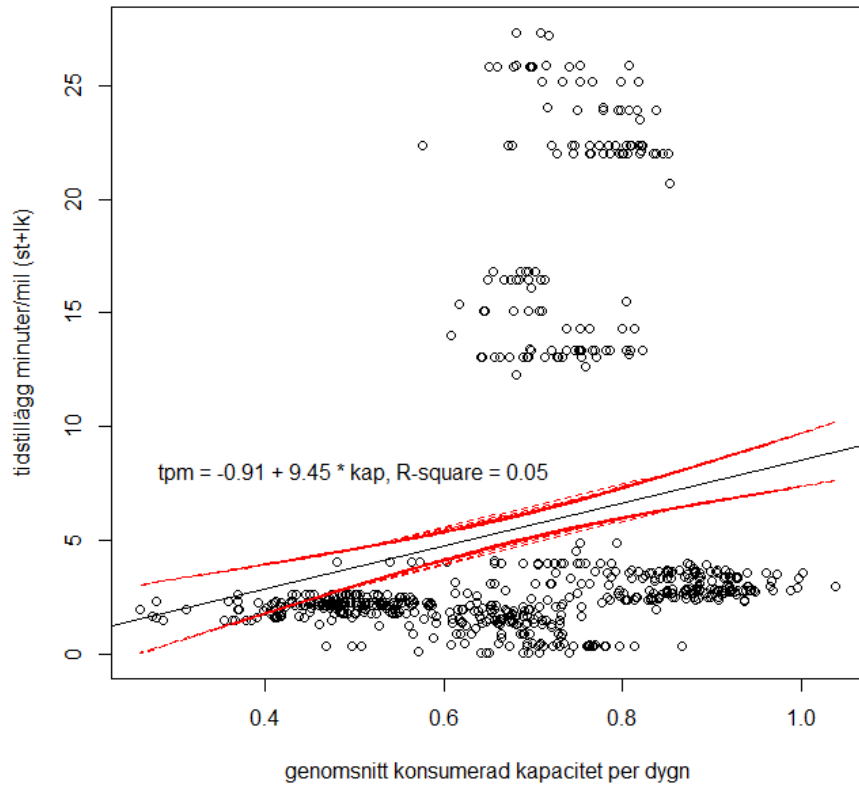
I figurerna nedan gäller följande tmp = "tidstillägg per mil". Kap = "konsumerad kapacitet" och (St+lk) betyder "station och länk".

Persontåg



Figur 4. Effektsamband mellan tidspåslag (tpm) i min/mil och konsumerad kapacitet (kap = KK) för persontrafiken på Värmlandsbanan år 2016

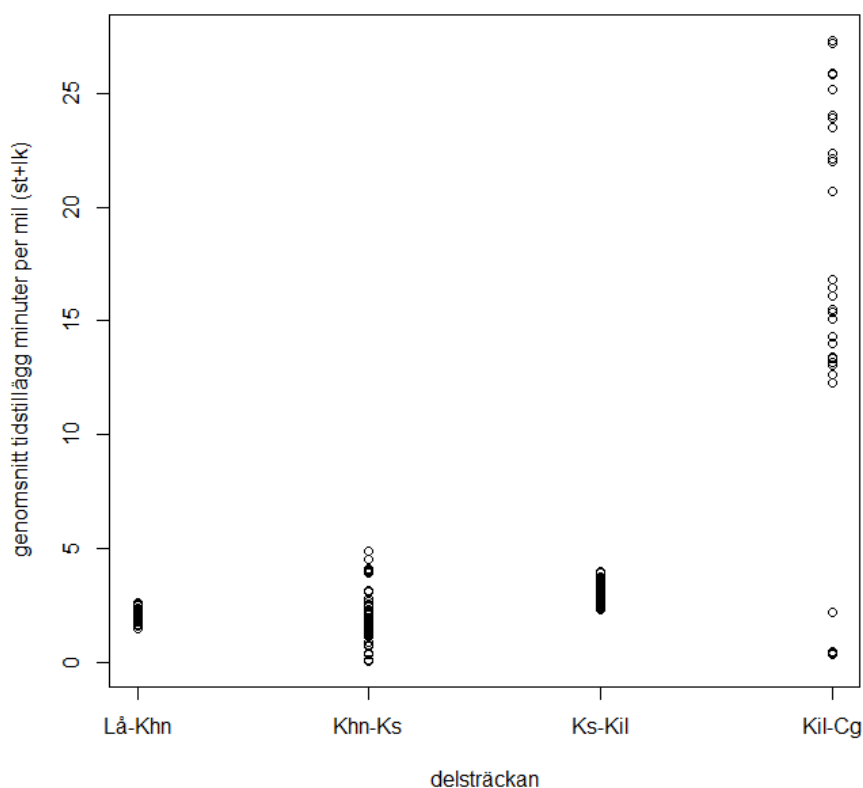
Godståg



Figur 5. Effektsamband mellan tidspåslag i min/mil och konsumerad kapaciteten för godstrafiken på Värmlandsbanan år 2016

Man kan notera att det för godstrafiken finns en hel del observationer med mycket positiva tidspåslag. Dessa observationer härrör från sträckan närmast Charlottenberg (Cg), vilket framgår i nästa figur som plottar sambandet mellan tidspåslag i min/mil mot de fyra delsträckor som studeras.

På sträckan Kil till Charlottenberg (Kil-Cg) finns det många tidspåslag som är över 10 minuter/mil som tidstillägg. Delvis väntar godståg längre vid gränsen, delvis finns här även ett signaturproblem som innebär att Charlottenberg i LUPP endast är en (1) station. I verkligheten finns det två stationer som ligger förhållandevis långt ifrån varandra (Charlottenberg (Cg) och Charlottenberg gränsen (Cggr)) där Charlottenberg gränsen i LUPP räknas som en understation till Charlottenberg. Som noterades i kapitlet om "Städning av data" rensades alla observationer med positiva tidtabellsangivelser från Cg-Cg, från Lå-Lå osv. Men vi har fortfarande problemet med att vi har tidsangivelser från näst sista stationen till Cg och att vi inte vet om denna tidsangivelse avser tiden till Cg eller till Cggr.



Figur 6. Tidspåslagen i min/mil uppdelat på delsträcka (med stora tillägg vid Charlottenberg)

Ovanstående problem infinner sig inte endast vid Cg/Cggr utan det finns "understationer" även vid andra platser. Laxå är ett annat exempel där det finns "understationer". Ett sätt att hantera problemet skulle vara att helt enkelt bortse från de observationer som kommer från sträckan Kil-Cg. Det kan dock vara onödigt att utesluta alla delsträckor där det finns "understationer". Vi upptäcker inte några problem för sträckan Laxå-Kristinehamn.

6.1.2 Norge-/Vänernbanan (NVB)

Längs Norge-/Vänernbanan förekommer det både sträckor med dubbelspår (två delsträckor närmast Göteborg) och enkelspår (upp till Kil). Det finns anledning att förväntas sig att sambanden skiljer sig åt mellan enkel- och dubbelspår och därför redovisas resultaten separat.

Dubbelspåret in mot Göteborg har en hög trafikbelastning med godståg och persontåg med olika uppehållsbild. Eftersom det bara finns två spår måste tågupplägg med olika genomsnittlig hastighet samsas på samma spår. Stora tidspåslag läggs in i tidtabellen för att hantera en hög sannolikhet för att snabbare tåg får ligga bakom långsammare tåg.

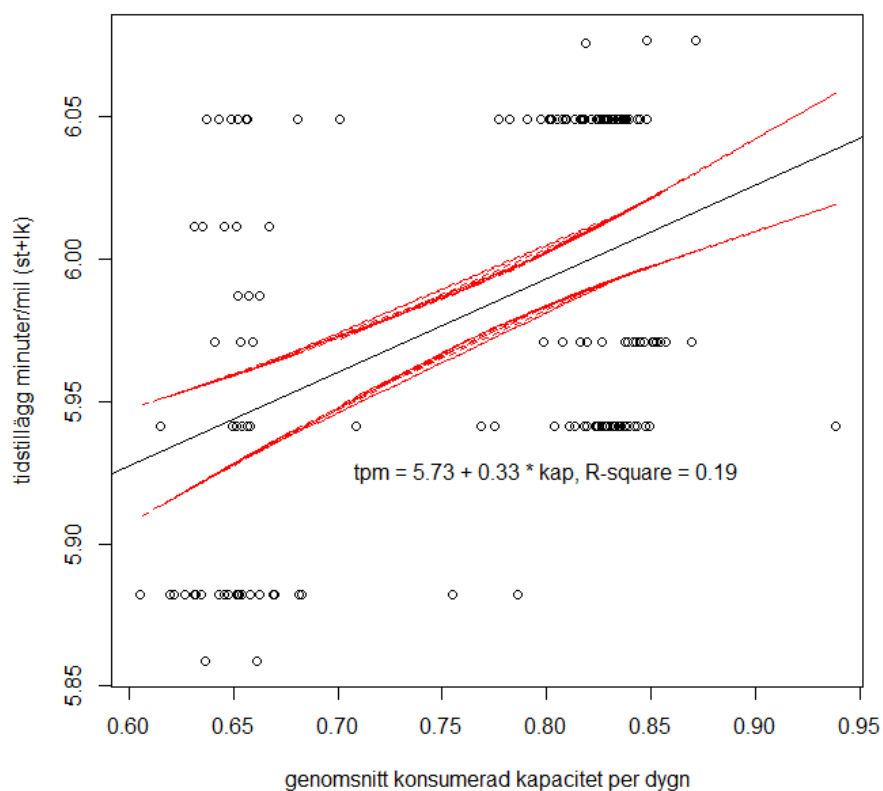
Tabellen nedan visar hur stor andel (%) av den totala tidtabellstiden som utgörs av "tid för möten/förbigångar och slack på NVB", uppdelat på enkelspår (Esp) och dubbelspår (Dsp).

NVB	Esp	Dsp	Esp+Dsp
Persontåg	3,4%	50,7%	19,1%
Godståg	24,2%	9,7%	17,3%

Tabell 10. Tidspåslagens andel av den totala tidtabellstiden uppdelat på enkelspår, dubbelspår samt persontåg och godståg

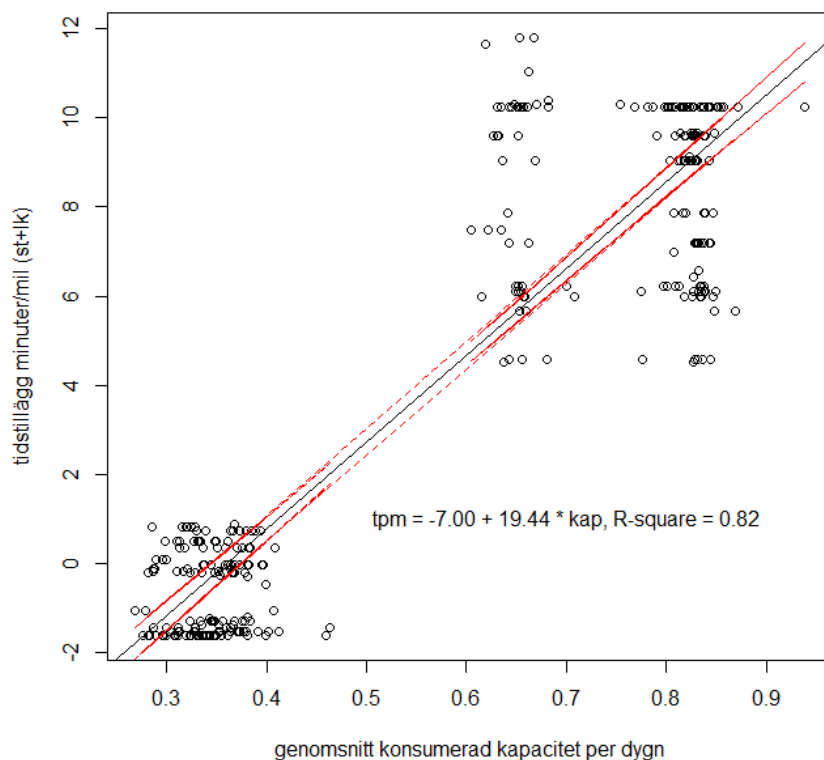
På dubbelspår delen har persontågen mycket stora tidspåslag i tidtabellen (50,7 %), vilket indikerar att persontågen har anpassat sig till godstrafikens hastigheter.

Dubbelspår- persontåg



Figur 7. Effektsamband mellan tidspåslag i min/mil och konsumerad kapacitet för persontåg på dubbelspår delen av NVB år 2016

Dubbelspår - godståg



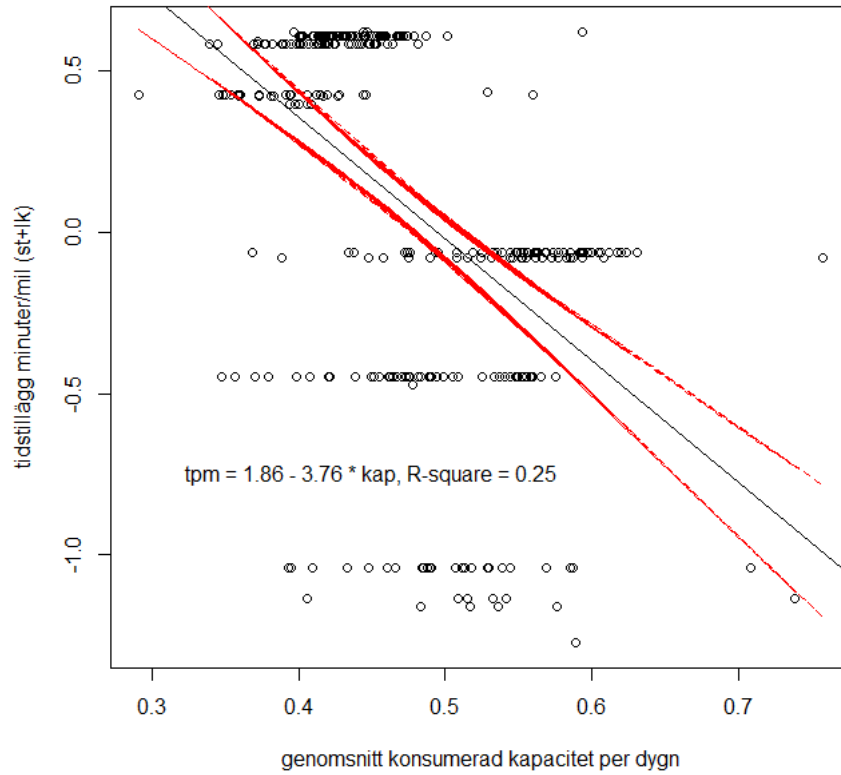
Figur 8 Effektsamband mellan tidspåslag i min/mil och konsumerad kapacitet för godståg på dubbelspårsgodståg på NVB år 2016

Det finns vissa sträckor där planerad tid är kortare än teknisk körtid. Det kan bero på att tidsangivelse för godstrafiken är sämre än för resandetågen i Lupp-databasen. Negativa tillägg erhålls även för godstrafiken på södra stambanan, vilket redovisas i 6.1.3 nedan.

Enkelspår

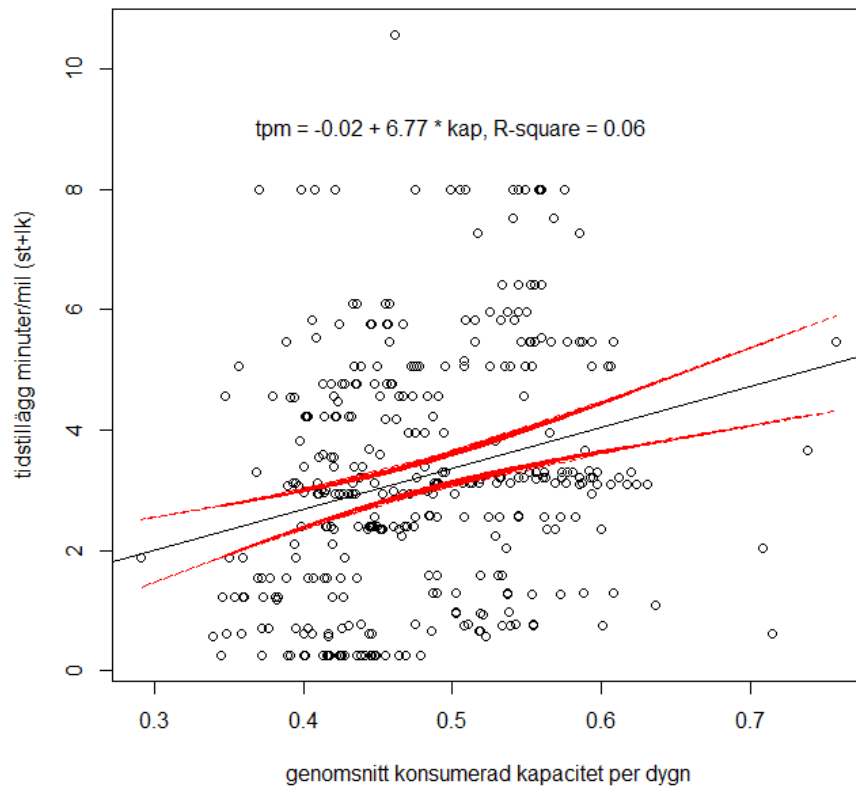
För enkelspårsgodståg på NVB är det svårt att finna ett tydligt effektsamband. I figurerna nedan kan man se att det är mycket svårt att bestämma hur man lämpligen drar en linje genom observationerna. För persontrafiken lutar linjen åt fel håll. För godstrafiken erhålls förvisso en linje med "rätt" lutning, men det är inte alls uppenbart hur den linjen bör dras.

Persontrafik



Figur 9. Effektsamband mellan tidspåslag i min/mil och konsumerad kapacitet för persontåg på enkelspårsdelen av NVB år 2016

Godståg



Figur 10. Effektsamband mellan tidspåslag i min/mil och konsumerad kapacitet för godståg på enkelspårsdelen av NVB år 2016

Merparten av observationerna ligger i ett snävt intervall för konsumerad kapacitet (40-60 %). Spridningen är också betydande och i princip slår den ut klusterstrukturen i data.

6.1.3 Södra Stambanan (SS)

Södra stambanan är en dubbelspårig järnväg med ett högt kapacitetsutnyttjande under delar av dygnet och alla typer av tåg samsas på spåren. I data är den indelad i 5 linjedelar.

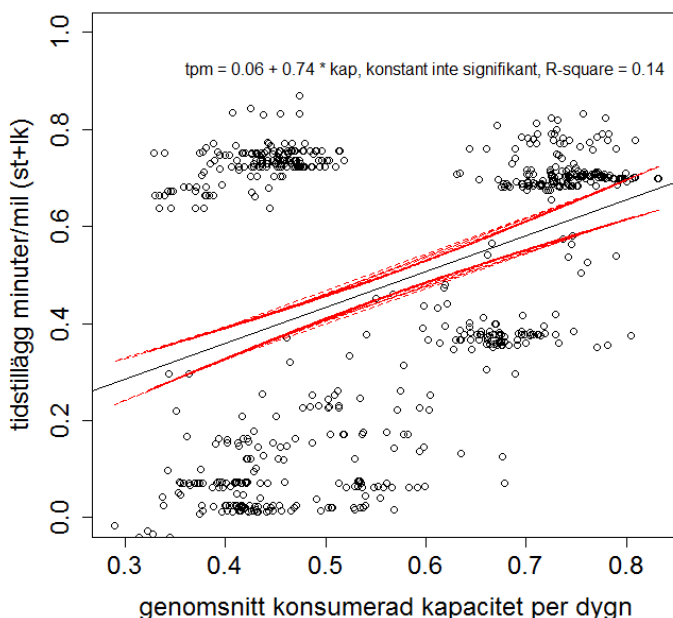
En intervju har gjorts med Pär Stemark, som under lång tid har gjort tidtabellerna för Södra stambanan söder om Nässjö. Vid denna intervju framgick det att persontrafiken till rätt stor del följer en "styv tidtabell", även om avsteg från denna princip framtvings för att inte binda upp allt för mycket kapacitet. Pär Stemark förväntar sig att det för godstrafiken bör finnas ett positivt samband mellan konsumerad kapacitet och tilläggen i tidtabellen, men att detta bör vara betydligt svagare för persontrafiken.

Tabellen nedan visar hur stor andel (%) av den totala tidtabellstiden som utgörs av "tid för möten/förbigångar och slack på SS", uppdelat på enkelspår (Esp) och dubbelspår (Dsp).

SS	Esp	Dsp	Esp+Dsp
Persontåg	Gäller ej	11,1%	Gäller ej
Godståg		28,3%	

Tabell 11. Tidspåslagens andel av den totala tidtabellstiden uppdelat på enkelspår, dubbelspår samt persontåg och godståg

Persontåg



Figur 11. Effektsamband mellan tidspåslag (tpm) i min/mil och konsumerad kapacitet (kap = KK) för persontrafiken på Södra Stambanan år 2016

Klustringstendensen är tämligen tydlig i figuren. Lutningen på linjen är också tämligen markant, vilket motsäger Pärs förväntningar för persontågstrafiken. Men det är också tämligen tydligt att regressionslinjen får sin lutning främst

från placeringen av klustren. Inom klustren, framför allt de tre övre, verkar lutningen vara tämligen marginell. Så inom klustren (linjedelarna) överensstämmer den visuellt skattade lutningen för en tänkt regressionslinje tämligen väl med Pärs uppfattning att det bör vara ett svagt samband för persontrafiken mellan konsumerad kapacitet och tidstillägg.

Godståg

För godstrafiken finns det många observationer med höga tidstillägg per mil. Detta gäller även efter att datarensning gjorts utifrån de fyra stegen som beskrivs i avsnitt 5.5 ("Städning av data") ovan. Möjliga felkällor kan vara att vi inte har kontroll på godstrafikens "kommersiella uppehåll", dvs att vi inte vet när godstågen stannar för att koppla till och/eller koppla ifrån vagnar. Här kan även finns tidsåtgång som är kopplade till banarbeten. I figuren nedan redovisas resultatet givet att observationer där godstågen får tidstillägg som är längre än 20 min per mil har tagits bort.

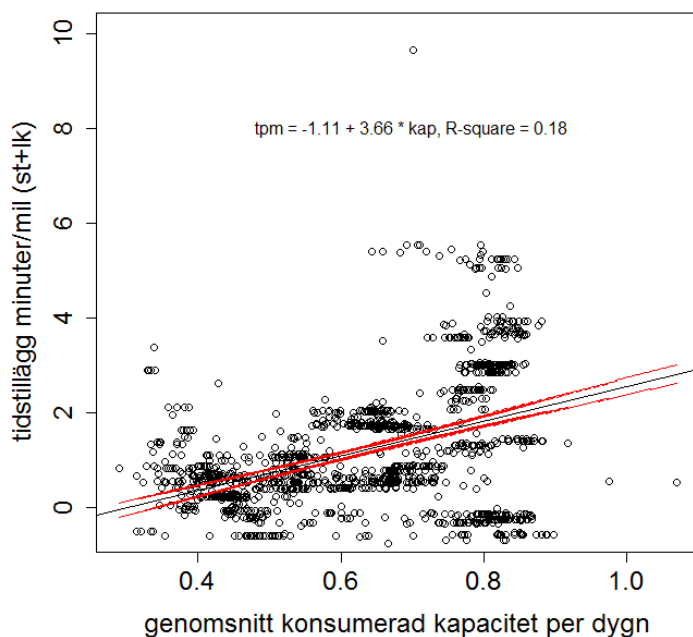
Det finns även en del observationer med negativa tidspåslag per mil. Detta kan härröra från matchningsproblem mellan tidtabellstider från LUPP och teknisk gångtid från Railsys. Problemet består i att placeringen av "timing points" mellan vilka teknisk gångtid kan beräknas, sträcker sig över flera länkar i LUPP. Matchning av data från de två källorna ger då upphov till saknade värden för teknisk gångtid. Placeringen av saknade värden behöver inte följa lätt tolkade principer för hur tidstillägg ska beräknas utifrån tidtabellstid från LUPP och teknisk gångtid från Railsys.

Tabellen nedan illustrerar problemet.

Linjedel	Länk	Tid i LUPP	Tek. gångtid i Railsys	Tidstillägg
1	a-b	12	NA	-
1	b-c	8	NA	-
1	c-d	10	22	8
1	d-e	15	20	-5
2	e-f	5	NA	-

Tabell 12. Illustration av hur systemet med "timing points" och saknade värden kan påverka estimeringarna

Saknade värden är markerade som NA. De två första NA är tolkade som att tekniska gångtid på 22 minuter gäller sammanlagt över länkarna a-b, b-c och c-d. Detta ger ett tidstillägg på $12 + 8 + 10 - 22 = 8$ minuter. NA i den sista raden tolkas som att den tekniska gångtiden på 20 minuter gäller för länken d-e enbart och ger upphov till ett negativt tillägg på $15 - 20 = 5$ minuter. Den korrekta hanteringen här skulle då vara att den tekniska gångtiden på 20 minuter gäller för de kombinerade länkarna d-e och e-f vilka skulle ha tidstillägg $15 + 5 - 20 = 0$. Om det tillägget hade blivit positivt skulle det också behövts fördelas mellan linjedel 1 och 2.



Figur 12. Effektsamband mellan tidspåslag (tpm) i min/mil och konsumerad kapacitet (kap = KK) för godstrafiken på Södra Stambanan år 2016

6.2 RESULTAT FRÅN UTÖKAD MODELLSPECIFIKATIONER

Resultatet av estimering av modellspecifikationen enligt ekvation (5.4) ovan, ges i tabell 13, för de totalt 8 separata modellerna givna av stråk, tågtyp och spårtyp. Som jämförelse visas i de två högra kolumnerna motsvarande parametrar från de nu gällande effektsambanden.

Modell Specifikation			Parametrar (effektsamband)			Nu gällande	
Stråk	Typ av tåg	Typ av spår	α	β	Övrigt	α	β
VLB	RST	Enkel	0,9	0,5		2,0	-0,3
	GT	Enkel	15,4	-4,4		2,58	-0,28
NVB	RST	Dubbel	0,3	5,7	*	4,0	-1,7
	GT	Dubbel	19,6	-7,3	*	4,0	-2,0
	RST	Enkel	-4,0	2,1	Fel estimering	2,0	-0,3
	GT	Enkel	-3,1	3,4	Fel estimering	2,58	-0,28
SSB	RST	Dubbel	0,8	0,001	$\gamma - \beta$ icke signifikant	4,0	-1,7
	GT	Dubbel	5,1	-1,4		4,0	-2,0

* NVB dubbelspår persontrafik går inte att estimeras på pga för få frihetsgrader.

Tabell 13. Estimeringsresultat för modeller estimerade enligt ekvation (5.4)

Tabell 13 visar en synnerligen stor spridning i erhållna parameter, både avseende tecken och värden. Endast för tre av godstågsmodellerna erhålls

rätt tecken för båda parametrarna α och β . Ett sätt att hantera detta är att aggregera modellerna så att de blir mer av genomsnittsmoeller. Detta görs i avsnitt 7.2 genom att modellerna för de tre stråken slås ihop.

Men spridningen visar också på specifikationsproblem. Därför redovisas här estimering av de olika specifikationerna givna av ekvationerna (5.4) – (5.6) i avsnitt 5.4. Den redovisningen begränsas till persontåg (RST) för Värmlandsbanan (VLB).

Grundmodellen har följande ekvation

$$(5.4) \quad T_{ij} = \alpha K K_{ij} D_i + \beta D_i + \varepsilon_{ij} .$$

Estimeringsresultatet för den modellen ges i tabell 14 nedan. Eftersom modellen inte har något intercept så är R^2 inte nödvändigtvis ett tillförlitligt mått på hur god anpassning till data modell (5.4) har. Notera att parametern för sträcklängd, β , får fel tecken när grundmodellen estimeras på data för VLB RST.

VLB persontåg, modell (5.4)			
Parameter	Värde	Std.av.	P-värde
α	0,85709	0,05609	<2e-16
β	0,51505	0,03746	<2e-16
Frihetsgrader	658		
RSE	1,109		
R^2	0,967		

Tabell 14. Estimeringsresultat för grundmodellen (5.4) för Värmlandsbanans persontåg

Om ekvation (5.4) utökas med en intercept-parameter påverkar inte parametererna för interaktionen mellan konsumerad kapacitet och sträcklängd, α , se tabell 8 nedan. Däremot minskar värdet för β (sträcklängd) med 40 procent vilken är en indikation på att intercept är nödvändigt att inkludera i modellen.

VLB persontåg, modell (5.4) + intercept			
Parameter	Värde	Std.av.	P-värde
(Intercept)	1,2869	0,0897	<2e-16
α	0,8358	0,0490	<2e-16
β	0,3199	0,0354	<2e-16
Frihetsgrader	657		
RSE	0,969		
R^2	0,827		

Tabell 15. Estimeringsresultat för grundmodellen (5.4) med intercept inkluderat, för Värmlandsbanans persontåg

Den ytterligare utökade specifikationen som inkluderar både intercept och huvudterm gavs av ekvation (5.5), alltså

$$(5.5) \quad T_{ij} = \alpha_0 + \alpha K K_{ij} D_i + \beta D_i + \delta K K_{ij} + \varepsilon_{ij} .$$

Resultaten från estimering av denna modell ges i tabell 16 nedan. Med denna specifikation får β korrekt tecken.

VLB persontåg, modell (5.5)			
Parameter	Värde	Std.av.	P-värde
α_0	9,5112	0,6376	<2e-16
α	2,67	0,1477	<2e-16
β	-1,1384	0,1165	<2e-16
δ	-10,1089	0,7776	<2e-16
Frihetsgrader	656		
RSE	0,864		
R ²	0,863		

Tabell 16. Estimeringsresultat för modellen (5.5) för Värmlandsbanans persontåg

Slutligen om modell (5.5) utökas med fixa effekter, γ_i , för de enskilda linjedelarna fås följande specifikation (se avsnitt 5.4)

$$(5.6) \quad T_{ij} = \alpha_0 + \alpha K_{ij} D_i + \beta D_i + \delta K_{ij} + \sum_i \gamma_i I_i + \varepsilon_{ij} ,$$

Tabell 17 nedan visar resultaten av en estimering enligt specifikation (5.6).

VLB persontåg, modell (5.6)			
Parameter	Värde	Std.av.	P-värde
α_0	-10,8106	0,6218	< 2e-16
α	-0,6573	0,1243	1,68e-07
β	2,8708	0,1166	< 2e-16
δ	0,3690	0,6669	0,58
γ_2	6,6033	0,1668	< 2e-16
γ_3	8,9024	0,2835	< 2e-16
Frihetsgrader	654		
RSE	0,4562		
R ²	0,9619		

Tabell 17. Estimeringsresultat för modellen (5.6) för Värmlandsbanans persontåg

Den största skillnaden gentemot modell (5.5) är att tecknen för parametrarna α och β har kastats om, även δ har bytt tecken (men är inte längre statistiskt signifikant). Parametrarna för de fixa effekterna är signifikanta, detta tillsammans med stora förändringar i övriga parametrars värden gör att det är svårt att trovärdigt förBISe de fixa effekterna vid estimering av en modell för effektsamband mellan konsumerad kapacitet och tidtabellstillägg.

Slutsatsen av tester av specifikationerna (5.4) – (5.6) för persontåg på Värmlandsbanan var att ändringar i specifikationerna i samtliga fall gav stora förändringar i estimerade parametrar. Denna brist på robusthet i resultaten kan, utifrån signifikans av fixa effekter för linjedelar, bero på att fler variabler än sträcklängd är nödvändiga för att beskriva linjedelarna. Specifikationen (5.6), som innehåller fixa effekter för linjedelarna, är ytterst svår att tolka avseende de parametrar som är intressant, α , β och δ (se avsnitt 8.3). Den specifikationen tjänar mer till att identifiera problem än att komma med en lösning. Vidare implikationer av detta tas upp i avsnitt 7.3.

Notera också att Värmlandsbanan består av fyra linjedelar. Det kan enbart estimeras fixa effekter (γ_2, γ_3) för två linjedelar, de två övriga fixa effekterna tas upp av interceptet α_0 och δ , se vidare avsnitt 7.3 för mer om detta.

7 RESULTAT

I kapitlet diskuteras resultat som framkommit under projektet och som har relevans för vidare arbete med effektsamband mellan konsumerad kapacitet och tidstillägg i tidtabellen. En stor del av arbetet i projektet har bestått av databearbetning. Erfarenheter från detta moment beskrivs i avsnitt 7.1. Resultat från estimeringar av regressionsmodeller av den typ som motsvarar nu gällande effektsamband tas upp i avsnitt 7.2. Vid estimeringsarbetet framkom problem med så kallade fixa effekter för linjedelarna, Dessa problem kan potentiellt innebära de estimerade parametrar som redovisas i avsnitt 7.2 avviker från deras korrekta värde. I avsnitt 7.3 ges utvärdering av dessa problem tillsammans med slutsatser för hur de kan hanteras i kommande arbete.

7.1 SLUTSATSER OM DATA

Projektet har estimerat de effektsamband som kan utläsas utifrån det tågföringsdata som finns inlagt i databasen LUPP. För att få fram de totala tidstillägg som finns i tidtabellen, d.v.s. tid utöver den tekniska gångtiden, används LUPP i kombination med de tekniska gångtider som finns i Railsys. Att använda sig av LUPP/Railsys har inneburit en hel del problem, som beskrivits närmare i avsnitt 5.5 ("Städning av data"). Det största problemet har varit att Railsys inte innehåller tillräckligt bra information om godstrafikens "kommersiella uppehåll" (d.v.s. var tågen kopplar till/ifrån vagnar etc.) och detta har i sin tur gjort att det är svårt att dra slutsatser om hur effektsambandet mellan konsumerad kapacitet och tidspåslag i tidtabellen ser ut för godstrafiken. Det hade sannolikt varit bättre att använda sig av databasen Trainplan och detta får gälla som en rekommendation. Om man skall estimerar ett effektsamband mellan konsumerad kapacitet och tidspåslag i tidtabellen med hjälp av en statistisk ansats så bör man testa att använda data från Trainplan. Vilka problem som då försvinner respektive uppkommer kan man dock inte veta säkert förrän man testat.

Modellkörningarna visar att konsumerad kapacitet är en signifikant och viktig variabel. Emellertid framgår det också att modeller med enbart konsumerad kapacitet och linjedelens längd som förklarande variabel inte är tillräckliga. Någoting annat bör också vara med. I nästa avsnitt presenteras de modellresultat som framkommit och som bygger på estimeringar enligt modell (5.4) i avsnitt 5.4.3. Resultaten jämförs med de gällande effektsamband som Trafikverket använder idag.

7.2 ESTIMERINGSRESULTAT

Här presenteras en sammanfattning för estimerade effektsamband för de studerade stråken: Värmlandsbanan (VLB), Norge-/Vänernbanan (NVB) och Södra Stambanan (SSB).

Modell Specifikation	Nya Effektsamband	Nu gällande
----------------------	-------------------	-------------

Stråk	Typ av tåg	Typ av spår	α	$\gamma - \beta$	Övrigt	α	$\gamma - \beta$
VLB	RST	Enkel	0,9	0,5		2,0	-0,3
	GT	Enkel	15,4	-4,4		2,58	-0,28
NVB	RST	Dubbel	0,3	5,7	*	4,0	-1,7
	GT	Dubbel	19,6	-7,3	*	4,0	-2,0
	RST	Enkel	-4,0	2,1	Fel estimering	2,0	-0,3
	GT	Enkel	-3,1	3,4	Fel estimering	2,58	-0,28
SSB	RST	Dubbel	0,8	0,001	$\gamma - \beta$ icke signifikant	4,0	-1,7
	GT	Dubbel	5,1	-1,4		4,0	-2,0

* NVB dubbelspår persontrafik går inte att estimeras på pga för få frihetsgrader.

Tabell 18. Jämförelse mellan nya estimerade effektsamband och nu gällande samband

De nya estimerade effektsambanden enligt tabell 18 ovan omfattar kapacitetstillägg, tidtabellstillägg och nodtillägg. Däremot ger de inte samma omfattande uppdelning på typ av resenärtåg som finns för de nu gällande effektsamband. Från tabell 18 är det uppenbart att de estimerade parametrarna i hög grad varierar över de åtta separata modellerna som estimerades. Orsaken till denna spridning i resultat diskuteras i nästa avsnitt. Utöver de åtta redovisade modellerna i tabell 18 ovan så estimerades modeller där stråken slogs ihop men som skiljer på de fyra kombinationerna av tågtyp och enkel- eller dubbelspår. Resultaten från de modell-estimeringarna redovisas i tabell 19 nedan där de nya modellerna kallas för Aggregerade modeller.

Aggregerade modeller			Effektsamband (fr. agg. mod.)			Nu gällande	
Stråk	Typ av tåg	Typ av spår	α	$\gamma - \beta$	Övrigt	α	$\gamma - \beta$
VLB + NVB	RST	Enkel	1,3	0,2	$R^2 = 0,91$	2,0	-0,3
	GT	Enkel	16,1	-5,2	$R^2 = 0,52$	2,58	-0,28
NVB + SSB	RST	Dubbel	5,1	-2,0	$R^2 = 0,42$	4,0	-1,7
	GT	Dubbel	6,6	-2,4	$R^2 = 0,58$	4,0	-2,0

Tabell 19. Parameter för generell modell estimering jämför med nu gällande modell

Alla estimerade parametrar är signifikant skilda från noll och generellt ger de större tidspåslag för en given ökning av den konsumerade kapaciteten än motsvarande nu gällande effektsamband, ett undantag är för persontrafik på enkelspår där tidspåslaget skulle bli mindre. Utöver modellen för persontåg på enkelspår där $\gamma - \beta$ har fel tecken så stämmer också de också överens med nu gällande effektsamband avseende tecken på parametrar. Godståg har högre känslighet för konsumerad kapacitet (högre tidspåslag

vid ökad konsumerad kapacitet), i synnerhet gäller det för godståg på enkelspår.

Tidspåslag Min per mil	Nya effektsamband (fr. agg. modeller)				Nu gällande effektsamband			
	Enkelspår		Dubbelspår		Enkelspår		Dubbelspår	
Kons. kapacitet	RST	GT	RST	GT	RST	GT	RST	GT
40%	0,7	1,2	0	0,2	0,5	0,8	0	0
50%	0,9	2,9	0,2	0,9	0,7	1,0	0,3	0
60%	1,0	4,5	3,1	1,6	0,9	1,3	0,7	0,4
70%	1,1	6,1	3,6	2,2	1,1	1,5	1,1	0,8
80%	1,2	7,7	4,1	2,9	1,3	1,8	1,5	1,2
90%	1,4	9,3	4,6	3,5	1,5	2,0	1,9	1,6

Tabell 20. Jämförelse av tidstillägg (min/mil) mellan nya effektsamband och nu gällande effektsamband

Tabell 20 ovan visar vilken skillnad som uppstår, i tidspåslag per mil, för några olika värden på konsumerad kapacitet om de aggregerade modellerna skall ersätta de nu gällande effektsambanden. Effektsambandet från de aggregerade modellerna ger i princip ingen förändring gentemot nu gällande effektsamband för persontrafik på enkelspår. Däremot ger de betydligt högre tidspåslag än de nu gällande effektsambanden för den övriga trafiken, persontåg på dubbelspår och godståg oavsett typ av spår.

Det är oklart i hur stor utsträckning problemen med bristfälliga uppgifter om godstrafikens "kommersiella uppehåll" har påverkat dessa estimeringar.

7.3 UTVÄRDERING AV RESULTAT

Under projektets gång har ett antal varianter av regressionsmodeller för effektsambandet mellan konsumerad kapacitet (KK) och tillägg till tidtabelltid (T) estimerats. Utgångspunkten har varit den underliggande regressionsmodellen för dagens effektsamband. Grundmodellen för att estimerade de gällande effektsambanden har följande specifikation (feltermen har uteslutits här och nedan)

$$(5.4) \quad T_{ij} = \alpha KK_{ij} D_i + \beta D_i.$$

Där D_i betecknar sträcklängden för linjedel i och j ett givet vardagsdygn. Den modellen har fördelen att den är enkel och att effektsambanden som ges av de två parametrarna α och β går att uttrycka som tidstillägg i minuter per mil vilket är enkelt att hantera i samhällsekonomiska kalkyler (till exempel kan man undvika att definiera linjedelar enligt avsnitt 5.2). Denna modellformulering har dock ett antal potentiella problem. Ett av dessa är att om den utökade specifikationen

$$(5.5) \quad T_{ij} = \alpha_0 + \alpha KK_{ij} D_i + \beta D_i + \delta KK_{ij},$$

ger att parametrarna α_0 och/eller δ är statistiskt signifikanta så blir parametrarna i modell (5.4) inte korrekt estimerade. Vid estimeringarna gick det att konstatera att α_0 och δ var signifikanta i flera delmodeller (avsnitt 6.2 visar detta för persontåg på Värmlandsbanan). Där så är fallet bör modell (5.5) användas som grund för effektsamband. Det är värt att notera att detta

innebär ett mer komplicerat effektsamband att använda, bland annat blir det mer kritiskt att korrekt definiera linjedelar enligt Trafikverkets metod för att beräkna konsumerad kapacitet (se avsnitt 5.2).

Vidare gick det att konstatera att både modell (5.4) och (5.5) ovan var känsliga för inkluderande av specifika additiva effekter för linjedelar (se avsnitt 7.2), även kallad fixa effekter. Det innebär att modell (5.5) bör utökas enligt följande

$$(5.6) \quad T_{ij} = \alpha_0 + \alpha K_{ij} D_i + \beta D_i + \delta K_{ij} + \sum_i \gamma_i I_i,$$

Där I_i är en indikatorvariabel för linjedel i och parametrarna γ_i är de fixa effekterna för linjedel i . De fixa effekterna motsvarar att varje linjedel i modell (5.6) har sitt egna specifika intercept. Estimering av modell (5.6) gav att de fixa effekterna var i hög grad signifikanta. Detta gällde för alla separat modeller. Däremot ger specifikation (5.6) svårtolkade parametrar för effektsambanden, se nedan, så detaljerade estimeringsresultat visades endast för persontåg på Värmlandsbanan i avsnitt 6.2.

Att de fixa effekterna var signifikanta innebär (på samma sätt som att utesluta ett signifikant intercept) att parametrarna i modell (5.5) inte blev korrekt estimerade. Det visade sig också att parametrarna α , β och δ i modell (5.5) skilde sig i betydande grad från motsvarande parametrar i modell (5.6). Totalt sett leder detta till att modell (5.5) bör förkastas. Frågan är då om en modell specificerad enligt ekvation (5.6) kan användas som grund för effektsamband?

Det finns generellt sett problem med att tillämpa fix-effekt modeller. I detta fall uppstår det till exempel problem om vi vill använda modell (5.6) som effektsamband för att räkna på en nyinvestering. Det kan bli mycket svårt att avgöra vilka fixa effekter som ska ansättas för de nya linjedelarna. Det kan dock vara möjligt att hitta approximationer som kan vara godtagbara. En sådan approximation skulle i detta fall vara att ersätta termen $\sum_i \gamma_i I_i$ med den genomsnittliga fixa effekten $\frac{1}{l} \sum_i \gamma_i$ där l är antalet linjedelar som ingick i estimeringen. Vi använder då samma genomsnittliga fixa effekt för alla linjedelar, gamla såväl som nya, som vi räknar på.

Det finns ett annat problem med specifikationen i modell (5.6), vilket leder till att de distansberoende effekterna som ges av parametrarna α och β blir svåra att tolka som just *distansberoende* effekter. För att se detta notera att sträcklängden D_i är helt bestämd av linjedelen i . Antalet unika sträcklängder i data är exakt lika många som antalet unika linjedelar i data och det finns ingen variation mellan i och D . Detta gör att parametern δ i modell (5.6) blir helt bestämd av de fixa effekterna. Om vi tänker oss en modell (5.6') som är modell (5.6) med termen βD_i utesluten och där de fixa effekterna betecknas med γ'_i , så kan vi efter att modell (5.6') har estimerats direkt beräkna vad parametrarna β och γ_i skulle få för värde när modell (5.6) estimeras. Eftersom sträcklängden D_i är helt bestämd av linjedelen i kommer vi inte att kunna estimeras alla parametrarna β och γ_i i modell (5.6). Det kan lösas genom att vi sätter $\gamma_{i_0} = 0$ för någon viss linjedel i_0 . Då får vi följande ekvationssystem för parametrarna i modell (5.6)

$$\beta = \frac{\gamma'_{i_0}}{D_{i_0}}$$

$$\gamma_i = \gamma'_{i_0} - \beta D_i \quad (i \neq i_0)$$

Specifikt så är β direkt bestämd av den ”rena” fixa effekten för linjedel i_0 och sträcklängden för samma linjedel. Men linjedelen i_0 hade vi valt helt godtyckligt. Så bara om de fixa effekterna γ'_i står i exakt samma proportion till D_i för alla linjedelar, är β entydigt bestämd. Faktum är att endast om de fixa effekterna enbart består av ett linjärt beroende till sträcklängd, det vill säga att $\gamma'_i = \beta D_i$ för alla linjedelar, så är β entydigt bestämd. I det fallet så förenklas specifikationen för modell (5.6) till att bli samma som för modell (5.5) och de fixa effekterna i modell (5.6) ska då alla estimeras till noll. I det fallet kan vi tolka termen βD_i som att den står för den sträcklängdsberoende delen av ett effektsamband för tidtabellstillägg. Nu vet vi att de fixa effekterna estimerades till att vara signifikant skilda från noll i modell (5.6). Därmed faller också möjligheten att tolka β som en effekt för sträcklängdens samband med tidtabellstillägg. Detta för att vi vet att de fixa effekterna inte enbart består av ett linjärt beroende till sträcklängd.

För interaktionseffekten α mellan sträcklängd D_i och konsumerad kapacitet KK_{ij} kan vi inte enkelt beräkna α utifrån de fixa effekterna, detta på grund av att KK_{ij} varierar både för i och j (dygn). Vi kan däremot konstatera att eftersom de fixa effekterna för linjedelar innehåller ytterligare effekter utöver sträcklängd så gör det faktum att det inte går att separera sträcklängd från linjedelar, att α högst troligt också innehåller effekter från linjedelar utöver sträcklängdsberoende. Därför bör α inte tolkas som ett trovärdigt effektsamband för interaktionseffekten mellan sträcklängd och konsumerad kapacitet. Eftersom huvudeffekten för konsumerad kapacitet δ beror på interaktionseffekten mellan konsumerad kapacitet och sträcklängd, α , så även δ kommer att påverkas av problemet med sträcklängd D_i .

Resultatet av de utökade estimeringarna gav alltså att de enkla modellspecifikationer för det sökta effektsambandet bör förkastas. De kan inte förväntas ge korrekt skattade (eller tolkningsbara) parametrar på grund av förekomst av statistiskt signifikanta fixa effekter för linjedelar. En utökad modellspecifikation som explicit tar hänsyn till de fixa effekterna har estimerats. På grund av att variabeln för sträcklängd enbart varierar med avseende på linjedel, så kommer en fix-effekt modell för linjedelar inte att ge parametrar som lämpar sig för effektsamband (så länge sträcklängd också är med i modellen). Därför tar nästa avsnitt upp en diskussion om möjligheterna för vidare estimeringsarbete som både tar hänsyn till de identifierade problemen men som också kan ge mer fullständiga effektsamband för tidtabellstillägg.

7.4 VIDARE ARBETE MOT EFFEKTSAMBAND

Grundproblemet är att vi har påvisat att det finns signifikanta fixa effekter för linjedelar i de modellspecifikationerna som har testats. Om dessa fixa effekter inte hanteras i modellspecifikationerna så kommer parametrar som är av intresse för att beskriva effektsambandet inte att estimeras korrekt. Problemet försvåras ytterligare a att det finns ett intresse av att inkludera

variabler som inte har någon dygnsvariation och därmed endast varierar med linjedel. Exempel på en sådan variabel är sträcklängden för linjedelar. Dessutom är det av intresse att inkludera interaktionstermer med sådana variabler och variabler som varierar med både linjedel och dygn, till exempel interaktionen mellan sträcklängd och konsumerad kapacitet.

Vi vill föra fram tre sätt att gå vidare för att hantera detta problem. Det första är fortsatta estimeringar med data av samma typ som har använts i denna studie men med utvecklade modellspecifikationer och en viss förändring i hanteringen av data. Men vi vill också föra fram möjligheten att genomföra hypotetiska experiment med tidtabellsläggare. Då finns möjligheten att variera variabler som sträcklängd så att estimeringar kan göras utan att hamna i problemet med fixa effekter för linjedelar. Det tredje sättet att gå vidare på är att ta fram nya effektsamband med hjälp av tidtabellsanalys, dvs med den metod som användes för att ta fram nu gällande effektsamband för godstrafiken på enkelspår. Dessa tre möjligheter beskrivs i avsnitten nedan.

7.4.1 Vidare estimeringar

En sannolik orsak till förekomsten av fixa effekter i de testade modellspecifikationerna är icke-inkluderade förklarande variabler för linjedelarnas effekt på tidtabellstilläggen. Den enda variabel som beskriver linjedelarna i de testade modellerna är sträcklängd. Det finns även andra möjliga förklaringar till förekomsten av fixa effekter till exempel att modellens funktionssamband inte är linjärt. Men vi föreslår att grundhypotesen i fortsatta estimeringar är att fler variabler som beskriver linjedelarnas effekt på tidtabellstillägg bör inkluderas i modellspecifikationerna.

I denna studie har modeller estimerats på data som förutsätter att parametrar för de fixa effekternas additiva huvudeffekter förs in explicit i modellspecifikationen. Detta går att undvika genom att använda så kallade differensdata (se avsnitt 5.4). Då utgår de fixa effekternas huvudeffekter samt huvudeffekter för variabler som enbart beror på linjedel. De enda effekter som återstår är de som berör variabler som varierar med både dygn, samt deras interaktioner med variabler som enbart varierar med linjedel (t.ex. sträcklängd). Om differensdata används så försvinner problemet med att huvudeffekterna för variabler som enbart varierar med linjedel är beroende av specifikationen av de fixa effekterna. Detta bör leda till robustare estimeringsresultat, men har inte testats.

Om estimeringen utförs på en datamängd med betydligt fler linjedelar, till exempel hela järnvägssystemet, finns det risk att estimeringarna blir numeriskt ohanterbara. Detta problem försvinner också om differensdata används.

7.4.2 Hypotetiska experiment

Grunden för problemet med de fixa effekterna är att vi vill basera delar av effektsambandet på variabler som enbart varierar med linjedel. I Hypotetiska experiment med tidtabellsläggare skulle man kunna frånga den begränsningen genom att helt enkelt låta sträcklängden vara variabel för linjedelarna. Detta kan vara det enda rimliga tillvägagångssättet om det visar sig att problemet med de fixa effekterna är i grunden svårt att lösa. Hypotetiska experiment är heller ingen ovanlig metod för att framställa

underlag till samhällsekonomiska kalkyler inom transportområdet, värderingar baserade på SP-studier är exempel på underlag där hypotetiska experiment har använts.

7.4.3 Tidtabellsanalys

Även om det inte är möjligt att ta fram ett perfekt effektsamband så måste Trafikverket ha någonting att använda som underlag till sina samhällsekonomiska kalkyler. Eftersom detta effektsamband påverkar hur stora tidsvinster som olika kapacitetsutbyggnader kan ge upphov till – och tidsvinster ofta är den centrala nyttan i de samhällsekonomiska kalkylerna – så är detta ett potentiellt viktigt effektsamband.

En budgetlösning skulle vara att helt enkelt genomföra ett antal tidtabellsanalyser av den typ som Göran Hörnell gjorde i samband med projekt Norrbotniabanan och som ligger till grund för det effektsamband som idag tillämpas för godstrafik på enkelspår. Detta innebär att man utgår ifrån nu gällande grafisk tidtabell (där det i utgångsläget råder ett visst kapacitetsutnyttjande) och därefter tar bort eller lägger till tåg och studerar vilka res- och transporttider som då blir möjliga i tidtabellen. Det som faller ut vid en analys av denna typ är "ett möjligt utfall" (ett stickprov med 1 observation) ur en större population med möjliga utfall. Om man låter 3 olika personer göra samma sak, för en enkelspårig respektive en dubbelspårig banan, erhålls 3 observationer. Om variansen mellan dessa observationer är låg så har Trafikverket ett samband som är någorlunda tillförlitligt. Om variansen är stor så finns det mer arbetet att göra – men Trafikverket kommer att ha lite bättre information om storleksordningen på effektsambandet.

=====

8 REFERENSER

Trafikverkets modell för beräkning av linjekapacitet, PM, Trafikverket 2016

Riktlinjer täthet mellan tåg – Planeringsförutsättningar, Tågplan 2017,
Trafikverket 2016

INTERVJUER

Bengt Eriksson, tidtabellskonstruktör, 2017-09-29

Pär Stemark, tidtabellskonstruktör, 2017-10-23

WSP Sverige AB

121 88 Stockholm-Globen
Besök: Arenavägen 7

T: +46 10 7225000
Org nr: 556057-4880
Styrelsens säte: Stockholm
wsp.com

