



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Grundforløbspakker og frafald på danske erhvervsskoler

Del 2: mulige effekter af reformen i 2007

Munk, Martin D.; Baklanov, Nikita; Bohn, Lars

Publication date:
2013

Document Version
Tidlig version også kaldet pre-print

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Munk, M. D., Baklanov, N., & Bohn, L. (2013). *Grundforløbspakker og frafald på danske erhvervsskoler: Del 2: mulige effekter af reformen i 2007.*

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- ? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- ? You may not further distribute the material or use it for any profit-making activity or commercial gain
- ? You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Grundforløbspakker og frafald på danske erhvervsskoler
Del 2: mulige effekter af reformen i 2007

Martin D. Munk, Nikita Baklanov og Lars Bohn

CMR, AAU-CPH, 29. august 2013

MANUSKRIPT

Indholdsfortegnelse

Datagrundlag og formål.....	3
Data om grundforløbspakkerne indhentet af CMR fra Erhvervsskolerne	3
Datakilde til identifikation af grundforløbstyper	4
Data fra erhvervsskolernes administrative database, Easy	5
Skoler med mulighed for identifikation.....	7
Fejlkilder og bortfald af data	8
Databortfald fra skolerne	9
Kvalitetskontrol og rettelser	10
Undersøgelse på aggregeret niveau	11
Registerdata fra Danmarks Statistik	12
Undersøgelhedsdesign	12
Sammenligning af WAS og CC-modeller	14
Metode:.....	14
Data	16
Resultater	18
Sammenligning af grundforløbstyper under CC-modellen	23
Data og metode.....	23
Resultater	25
Diskussion	28
Litteratur:.....	30
Appendiks.....	32
Appendiks A. Oversigt over forklarende variable anvendt i regressionsmodeller	32
Appendiks B.1: Deskriptiv statistik for alle baggrundsvariables for hhv. WAS, CC og hele Bygge og Anlæg (BA) populationen, 2008-2010	33
Appendiks B.2: Deskriptiv statistik for alle baggrundsvariables for hhv. WAS, CC og hele Strøm, Styring og IT (SSI) populationen, 2008-2010.....	34
Appendiks C. Clear Cut Opkvalificering 1 vs. Obligatorisk, LPM.....	35
Appendiks D. Clear Cut Opkvalificering 2 vs. Obligatorisk, LPM.....	36
Appendiks E. Deskriptiv statistik for baggrundsvariables, Clear Cut Analyse	37
Appendiks F. Stratificeret CR Cox Model – Frafald og Fuldførelse – Bygge og Anlæg.....	38
Appendiks G. Logit – Praktikvej – Bygge og Anlæg.....	39
Appendiks H. Sammenligning mellem CR Cox og Fine-Gray modellerne – Frafald – Bygge og Anlæg ..	40

Datagrundlag og formål

Formålet med denne del af forskningsprojektet om frafald på erhvervsuddannelserne er, som det også er omtalt i del 1, at undersøge om det er muligt at påvise betydningen af skolernes indsats mod frafald ved hjælp af en sammenlignende analyse af skoler på bygge og anlæg samt et sammenlignende design.

Data om grundforløbspakkerne indhentet af CMR fra Erhvervsskolerne

Udgangspunktet er den analyse af grundforløbspakkerne, som omtales i del 1. Her viste vi, at skolernes indsats for fastholdelse af grundforløbseleverne kan beskrives på tre dimensioner. Dels kan man skelne mellem forskellige typer grundforløb, målrettet mod forskellige elever. Man kan her tale om obligatoriske hold, målrettet elever, som ved hvad de vil og har hvad der skal til. Disse er typisk af en varighed på 20 uger.

Mange skoler har også forskellige typer elitehold. Det kan være hold for elever med ambitioner om at videreudanne sig efter endt erhvervsuddannelse, eller hold for elever, som har andre ambitioner, fx om at være selvstændig eller om en international karriere. Mange steder har man også mulighed for at gennemløbe grundforløbet hurtigere hvis man har erhvervs erfaring relateret til det valgte fag.

De fleste skoler har også et eller flere holdtyper på et langsommere niveau, målrettet mod elever med forskellige vanskeligheder i forhold til gennemførelse af grundforløbet på de normale 20 uger. Hvis skolen (eller afdelingen på skolen) kun har en holdtype målrettet mod elever med få ressourcer så vidt angår skolefærdigheder, varer den typisk 30 uger, men på nogle skoler har man holdtyper af både 30 og 40 ugers varighed. I disse tilfælde er det almindeligt, at de kortere grundforløbspakker målrettes mod elever, hvis hovedvanskelighed er manglende faglige kompetencer, mens de længere pakker er målrettet elever, som har større vanskeligheder, og typisk ikke er studieparate.

På en anden dimension kan man skelne mellem forskellige grundforløbsmodeller. Modellerne udtrykker forskellige måder at realisere den ovenfor omtalte niveaudeling. Her skelnes mellem en clear-cut model, hvor eleverne deles på forskellige, klart afgrænsede hold, således at de obligatoriske elever går på et hold i 20 uger, og pakkeeleverne på et andet af 30 eller 40 ugers varighed. Andre steder lader man alle elever starte samtidig. De elever, som har brug for ekstra tid eller hjælp, eller som har valgt at supplere forløbet med ekstra elementer, fortsætter så i længere tid, indtil de er klar til at overgå til hovedforløbet. Denne model har vi valgt at betegne som vent-og-se modellen.

Endelig kan man tale om en tredje dimension af fastholdelsesindsatsen. Skolerne har også forskellige tiltag på holdene, hvis sigte er at fastholde eleverne på holdet. Således er der hold, hvor eleverne får morgenmad, enten i form af adgang til gratis morgenmad i kantinen eller et egentligt morgenmads arrangement på holdet. Nogle skoler har en lavere klassekvotient på hold, målrettet mod svagere elever, eller har et lærerkorps, som arbejder mere socialpædagogisk end på de obligatoriske hold. Mange steder har man også en øget indsats for at sikre lavt fravær på holdene.

I del 1 beskrives disse forskelle mere detaljeret, ligesom der også redegøres for datagrundlag og analysemetode for den analyse, som har ført til identifikationen af disse grundforløbstyper og –

modeller. Her ses også hvilke hypoteser om de forskellige tiltag, den aktuelle fraværersforskning giver anledning til.

I det følgende præsenteres resultatet af vores undersøgelse af nogle af disse hypoteser. Dels er vi gået efter at undersøge, om det er muligt at spore tegn på effekter på frafaldet af det enkle forhold, at skolesvage elever får mulighed for at gennemføre grundforløbet over længere tid på de to hovedmodeller. Desuden har vi interesseret os for effekter af de forskellige modeller, der har været brugt til at skabe øget fastholdelse blandt skolesvage elever. Som det fremgår både her og i første del, er variationen ganske omfattende, både med hensyn til grundforløbstyper, men især med hensyn til de modeller, skolerne har anvendt til realisering af grundforløbstyperne. Da man tillige kun med forsigtighed kan sammenligne frafald på forskellige uddannelsesretninger, bliver mulighederne relativt begrænsede mht. hvilke grundforløbsmodeller det har været muligt at undersøge.

Vi har derfor valgt at begrænse undersøgelsen af modellernes effekter til en sammenligning af modellerne clear-cut og vent-og-se på indgangene Bygge og Anlæg (BA) og Strøm, styring og IT (SSI) samt en sammenlignende undersøgelse af clear-cut med henholdsvis en kontrolgruppe og to indsatsgrupper.

Datakilde til identifikation af grundforløbstyper

Undersøgelsen forudsætter imidlertid data om grundforløbenes karakter mht. holdtype og model for de enkelte elever, som Danmarks Statistik ikke kan levere. Derimod giver skolernes administrative database, Easy-A, mulighed for at registrere disse forhold. Her registreres blandt andet elevernes samlede uddannelsesforløb, hvilken uddannelse de har valgt og hvilke forskellige hold, de er tilknyttet, eller har været tilknyttet. Det er også muligt at registrere særlige tiltag, hvilket dog i praksis kun sker mht. grundforløbspakkernes tidsmæssige udstrækning.

Ikke alle skoler udnytter muligheden for at registrere grundforløbstype eller model, og databasen rummer heller ikke nogen enkel måde at indhente disse data til statistisk brug, der hvor det registreres. Derfor har det været nødvendigt for os at undersøge mulighederne for at trække de ønskede oplysninger ud af databasen, samtidig med at vi undersøgte hvilke skoler som har brugt databasen på en måde, så den rummer disse data.

Vi indgik en aftale med UNI-C, der administrerer databasen, om levering af data fra Easy-A, når vi havde identificeret hvilke data der kan knytte eleverne til de grundforløb, de har fulgt og hvilke skoler, der registrerer eleverne på en måde, der gør det muligt. Vores undersøgelse viste, at vi kunne forvente at hente de ønskede data ud fra 20 af de 36 tekniske erhvervsskoler, der er i Danmark, hvorfor vi i foråret 2012 fik data fra UNI-C fra disse 20 skoler: EUC Vest, EUC Syd, EUC Lillebælt, EUC Nordvest, EUC Sjælland, Roskilde, Københavns, Silkeborg, Holstebro, Svendborg, Skive og Randers Tekniske Skoler, Syddansk og Nordsjællands Erhvervsskoler, Viden Djurs, Herningsholm, Teknisk Erhvervsskole Center (TEC), CELF, Aalborg TECH og Mercantec.

Gennemgangen af den endelige dataleverance afslørede, at det på to skoler kun var muligt at sammenknytte elever og grundforløb i meget få tilfælde (EUC Syd i Sønderjylland og Mercantec i Viborg), og at der var betydelige vanskeligheder på andre skoler. Andre steder er data i sig selv af

god kvalitet, men antallet af elever er for lavt til, at man kan producere sikre resultater. Af disse årsager er vi endt med at udvælge et begrænset antal skoler til undersøgelsen, men også den quasi-eksperimentelle tankegang tænker i det gode design, ikke alene på antal. Der er tale om Herningsholm, EUC Vest, Syddansk Erhvervsskole i Odense, Roskilde Tekniske Skole og CELF på Lolland-Falster.

Data fra erhvervsskolernes administrative database, Easy

Opgaven med at hente data fra skolernes administrative database, Easy, har nødvendiggjort at vi har oparbejdet et indgående kendskab til databasens indretning. Vi har fra samtaler med forskellige medarbejdere i UNI-C, og undersøgelser af udleveret materiale, fået et vist kendskab til databasens opbygning. I realiteten er der tale om tre, forbundne databaser, Easy-A, Easy-P og Easy-S. Skolernes administrative data lagres i en lokal udgave af databasen Easy, kaldet Easy-A. Denne del af databasen bruger skolerne til den daglige administration af elever og undervisning, og den rummer også de informationer der bruges til at beregne skolernes taxametertilskud.

En del af data fra Easy-A overføres på daglig basis til en central database, Easy-S, som administreres af UNI-C og er beregnet på statistisk brug. Det er herfra, vi henter de data, som kan bruges til at identificere hvilke grundforløbstyper eleverne har fulgt, når de kobles med de informationer vi har indsamlet fra skolerne. Easy-P registrerer forskellige forhold omkring elevernes praktikaftale. Her registreres det bl.a. om eleverne har en praktikaftale og hvornår den i givet fald er indgået – en faktor som givetvis har en vis betydning for fastholdelse.

Data lagres i forskellige tabeller. To af disse, tabellen skoleforløb og tabellen aktiviteter, kan anvendes på en måde, som ville gøre identifikation mulig. Begge disse overføres til Easy-S.

Alle elever skal registreres i tabellen skoleforløb. Eleverne kan ikke registreres på mere end et skoleforløb, da det er denne registrering, der er grundlag for udbetaling af taxameterpenge til skolerne. Hvis skolen opretter en skoleforløbstabel separat for hver grundforløbstype vil det være muligt at identificere hvilken type grundforløb en given elev har fulgt ud fra skoleforløbet. Vores samtaler med skolerne viste, at dette kun er tilfældet på en skole, nemlig EUC Lillebælt i Fredericia. De øvrige steder korresponderer skoleforløbet typisk blot med elevernes start- og sluttidspunkt på skolen.

Skolerne registrerer imidlertid også eleverne på forskellige hold. Disse holdoplysninger lagres i tabellen aktiviteter. Hver tabel har et unikt navn og registrerer samtlige elever på et givent hold. Eleverne kan være på mange forskellige hold, som hver henviser til forskellige forhold. Der kan være oprettet hold med elever, der deltager i forskellige mindre undervisningsaktiviteter (for eksempel delementer i den faglige uddannelse, førstehjælpskurser o.l.) eller elever som har søgt optagelse på skolen, som ikke er begyndt endnu, og der kan være oprettet hold, som registrerer samtlige elever, der følger en bestemt grundforløbspakke på en specifik uddannelse.

Tablet 2.1: Udsnit af tabellen AKTIVITETER med de felter, som kan anvendes til identifikation.

Kolonne navn	Krævet	Datatype
AKTIVITET (holdkoden)	Ja	Tekststreng max 12 enh.
BETEGNELSE	Ja	Tekststreng max 50 enh.
KORT_BETEGNELSE	Ja	Tekststreng max 15 enh.
STARTDATO	Ja	Dato
SLUTDATO	Ja	Dato

Det er imidlertid ikke tilstrækkeligt, at alle elever på et givent grundforløb er registreret på samme hold. Identifikation af grundforløbstype forudsætter naturligt nok også, at holdet efterfølgende kan forbindes til det grundforløb, som eleverne har gået på. Det er ikke en selvfølge. I nogle tilfælde er det en relativt enkel øvelse at knytte et hold til en bestemt type grundforløb, mens det andre gange er vanskeligt eller ikke muligt.

Tabellen aktiviteter består af en række felter, hvoraf nogle rummer mulighed for identifikation af grundforløbstype, afhængigt af hvordan tabellen bruges af det administrative personale på den konkrete skole og afdeling. Disse felter kan ses i tabel 2.

Holdnavnet i feltet aktivitet er typisk en kode – en kortere eller længere række af tal og bogstaver. På en del skoler indeholder denne aktivitetskode informationer om hvilken grundforløbstype holdets elever har gået på. Det er tilfældet på Herningsholm og EUC Vest, CELF, enkelte indgange på Syddansk Erhvervsskole og Silkeborg og Holstebro Tekniske Skoler og til en vis grad grundforløbene på Aalborg TECH.

Mange steder er det imidlertid ikke muligt at knytte forbindelse mellem hold og grundforløbstype ud fra aktivitetskoden, idet det måske blot er et fortløbende nummer uden nogen specifik information om holdet. Der kan også være tale om en kode, som oplyser forhold som for eksempel uddannelsesretning og år uden at have informationer om grundforløbets type.

I nogle af disse tilfælde har vi brugt holdets angivne start og sluttidspunkt. På de skoler og indgange, hvor holdene oprettes på grundforløbets planlagte længde og der ikke er forskellige grundforløbsmodeller med samme tidslængde, er det muligt at slutte fra holdets planlagte længde til hvilken grundforløbstype holdets elever har gået på. Efterfølgende kan identifikationen knyttes til holdets aktivitetskode i en tabel, vi selv opretter og tilføjer til data.

Denne fremgangsmåde er i udgangspunktet mindre sikker end identifikation direkte ved aktivitetskode, da identifikationen kun er indirekte, men vi har søgt at gardere os mod fejl ved at indhente verifikation af vores identifikationer fra de pågældende skoler. Fremgangsmåden er anvendt på forskellige indgange på Holstebro og Silkeborg Tekniske Skoler, på EUC Nordvest (Thisted), Nordsjællands Erhvervsskole i Hillerød og en enkelt afdeling på Københavns Tekniske Skole.

I et andet tilfælde – Strøm, styring og IT på Syddansk Erhvervsskole – har vi baseret identifikationen på feltet betegnelse, hvor skolens administrative personale kan skrive kommentarer om holdene. Her er det naturligvis muligt også at skrive, hvilken grundforløbstype holdets elever følger, hvilket skolen har gjort i dette tilfælde. Det samme gælder et mindre felt, kort betegnelse, hvor man også kan lave noter om holdene. Brugen af disse felter er lige så sikker som identifikation baseret på aktivitetskoden, da begge dele baserer sig på skolens administrative personales indtastninger. Ofte bruges felterne imidlertid ikke systematisk, således at de kun af og til rummer informationer om grundforløbstype på indgangene. I disse tilfælde kan man altså ikke bruge feltet til sikker identifikation af grundforløbstype for alle holdene, men feltet kan anvendes til dobbelttjekke identifikation ud fra holdenes angivne tidslængde eller aktivitetskode, hvilket er gjort hvor det har været muligt.

Endelig viste det sig på enkelte mindre skoler muligt at identificere elevernes grundforløbstyper selv om ingen af ovennævnte metoder er mulige. Det er tilfældet på Viden Djurs i Grenå, Svendborg Teknisk Skole (Mad til Mennesker), på Tradium, Randers Tekniske skole (Strøm, Styring og IT) samt enkelte uddannelsesretninger på Københavns Tekniske Skole og TEC i hovedstaden. I disse tilfælde kunne skolerne levere en fortegnelse over grundforløbstyper knyttet til aktivitetskoden for samtlige hold.

Skoler med mulighed for identifikation

Databasen Easy rummer således potentielt muligheden for at identificere hvilken type grundforløb den enkelte elev har fulgt. Det forudsætter to ting: dels at man kender indholdet af de forskellige grundforløb, dels at databasens skoleforløb eller hold kan knyttes til disse på en af de omtalte måder.

Vores undersøgelse på skolerne viste at det kun på 3 skoler er muligt at identificere samtlige elevers grundforløbstyper: EUC Vest i Esbjerg, EUC Lillebælt i Fredericia og Herningsholm.

Ud over disse skoler kunne 15 andre skoler give os data som muliggør identifikation af elevernes grundforløbstyper på nogle af deres uddannelsesretninger. Det drejer sig om CELF på Lolland-Falster, Syddansk Erhvervsskole i Odense, Randers, Roskilde, Københavns, Skive, Silkeborg og Holstebro Tekniske Skoler, EUC Nordvest i Thisted, EUC Sjælland i Næstved og Køge, TEC i hovedstaden (Frederiksberg, Hvidovre, Ballerup, Gladsaxe og Lyngby), EUC Syd i Sønderborg, Svendborg og Nordsjællands Erhvervsskole (Hillerød afdeling), Videnscenter Djurs (Grenaa), og Aalborg TECH.

Yderligere to skoler kunne levere data om elevernes konkrete grundforløb, men kun for enkelte hold. EUC Syd i Sønderborg kunne således give os data på eleverne på et særligt fælles 60 ugers ungdomsforløb, beregnet på elever med brug for tid til at blive mere modne, og Mercantec i Viborg havde identifikationsdata for elever på særlige hold for utilpassede og frafaldstruede elever og i Viborg også et enkelt elitehold, men ellers ikke for hverken obligatoriske elever eller pakke elever på deres indgange.

Det er således potentielt muligt at undersøge effekterne af grundforløbsmodellerne på en række af disse 20 skoler. Derfor har vi foretaget en række interview med den eller de uddannelsesansvarlige for disse skoler og/eller de respektive afdelinger på skolerne, samt med de administrative medarbejdere, som havde indsigt i elevregistreringen – enten ved besøg eller telefonisk.

De fleste interview blev gennemført centralt, ofte med deltagelse af personale fra flere afdelinger, og som regel med viden fra hele skolen repræsenteret. Der har været 9 sådanne møder på skolerne: EUC Vest, EUC Lillebælt, EUC Syd, EUC Nordvest, Syddansk Erhvervsskole (Odense), CELF (Nykøbing Falster), Skive Tekniske Skole, Herningsholm, og Aalborg TECH. I 3 tilfælde mødtes vi separat med en eller flere afdelinger. Det er tilfældet på Roskilde og Randers Tekniske Skoler, og TEC i hovedstaden, hvor vi gennemførte interview på tre afdelinger: Frederiksberg, Hvidovre og Ballerup. Telefonsamtalerne er for stort set alle de syv kontaktede skolers vedkommende (Holstebro, Silkeborg, Svendborg og Københavns Tekniske Skoler, samt EUC Sjælland, Nordsjællands Erhvervsskole og Mercantec (Viborg), foregået med skolernes forskellige afdelinger hver for sig.

Blot i et tilfælde er en telefonisk dataindsamling foregået ved en samtale på centralt niveau: Viden Djurs i Grenaa, som er en relativt lille skole.

Fejlkilder og bortfald af data

Som nævnt indledningsvis indgik CMR en aftale med UNI-C om udtræk og leverance af Easy data med henblik på identifikation af de grundforløbspakker, eleverne i undersøgelsen har gået på. Det har imidlertid vist sig at være en mere kompleks opgave end nogen ved projektets påbegyndelse havde forestillet sig.

I august 2010 blev vi stillet et samarbejde i udsigt, som skulle føre til en udbygning af databasen, så den ville blive udbygget med nye data til statistisk brug, samtidig med at vi fik data om elevernes grundforløbspakker, praktikpladssituation, om frafaldsårsager og fravær fra undervisningen, om de fag eleverne har fulgt samt oplysninger om hvilke lærere de enkelte elever har været undervist af. Imidlertid har det vist sig at være forbundet med endda meget store vanskeligheder at hente de ønskede data ud af databasen.

Det første levering kom i maj 2011, men viste sig kun at have de aktivitetskoder, som knytter eleverne til grundforløbsmodellerne for en enkelt skole, EUC Syd. Her måtte vi samtidig konstatere at skolens brug af databasen Easy-A, som leverer data til Easy-S, kun tillod identifikation af ganske få hold, nemlig den såkaldte ungdomsklasse.

Tabellen indeholdt dog oplysninger om samtlige elevers skoleforløb. I praktisk betød det, at data alene kunne anvendes til identifikation af grundforløbstype og model på en skole, nemlig EUC Lillebælt, hvor det som ovenfor anført ikke er tabellen aktiviteter med aktivitetskoden, men tabellen skoleforløb, der kan anvendes til identifikation.

I november samme år modtog vi så en ny leverance, en version 2, som havde aktivitetskoder for næsten samtlige elever på de 20 skoler, vores undersøgelse havde udpeget. Ud af de i alt 91.783 elever i data var der aktivitetskoder for 86.976 elever. Der manglede således aktivitetskoder for 4807 elever.

Dette var ikke nødvendigvis ødelæggende for vores undersøgelse, da en stor del af disse elever givetvis var tilknyttet afdelinger, som ikke kan levere data om elevernes tilknytning til grundforløbsmodel og type, således at det reelle antal manglende elever ville være væsentligt lavere.

Det viste sig imidlertid, at der stadig manglede ganske mange aktivitetskoder, selv om vi nu havde aktivitetskoder på næsten alle elever. Som nævnt er det kun nogle af elevernes aktivitetskoder, der knytter eleverne til en grundforløbstype og model – andre henviser til forhold, der er uinteressante i denne sammenhæng. Da en stor del af de aktivitetskoder, som rummer informationer om elevernes grundforløbstype og model, ikke var med heller i denne leverance, var data reelt ubrugelige.

Den tredje leverance, version 3, leveret ultimo januar 2012, rettede op på dette problem. Nu var alle elever i datasættet forsynet med de for os relevante aktivitetskoder. Imidlertid viste et kontroltjek med data leveret direkte fra erhvervsskolerne Herningsholm og CELF, at der manglede elever svarende til ca. 10 % af eleverne på de tekniske grundforløb. Årsagen til dette er, at UNI-C ikke medregner elever som mindre end 7 måneder forud for start på et grundforløb har haft en uddannelsesstart på en anden erhvervsuddannelse. Derfor var elever, som mindre end 7 måneder

forud for en debut på et teknisk grundforløb havde påbegyndt en uddannelse på en anden erhvervsskole eller på en merkantil eller SOSU uddannelse på samme skole, ikke med i data.

Følgelig fik vi udvirket, at UNI-C leverede endnu et datasæt, en version 4, som vi modtog 20. marts 2012, nu med et bruttoudtræk af samtlige elever, som indenfor vores undersøgelsesperiode har været indskrevet på en teknisk uddannelse på de skoler, vi har valgt at undersøge. Dette dataudtræk rummer ikke samtlige oplysninger om eleverne, men muliggjorde identifikation af de manglende 10 % elever på de skoler, som er med i vores undersøgelse, og når skolerne har data på det, også de grundforløbsmodeller, de har fulgt.

Fra UNI-C har vi modtaget et bruttodatasæt med data på i alt 135.173 elever på de 20 skoler, der er med i vores undersøgelse. En del af disse elever går imidlertid på indgange, som ikke har data på hvilke grundforløb elevernes hold er knyttet til, eller hvor data af tekniske eller registreringsmæssige årsager ikke er brugbare.

Når disse data sorteres fra, kan data identificere grundforløbstyper for 43.136 elever. Imidlertid er der ikke oplysninger om grundforløbsmodeller for 7.321 af disse elever. Der kan derfor kun identificeres grundforløbsmodel for 35.815 elever. Dvs. vi ved på hvilket niveau 43.136 af eleverne på de danske erhvervsskoler har gennemført deres grundforløb og hvilken art niveauet har haft – obligatorisk, elite for iværksættere, studieforberedende eller internationalt orienterede elever eller udvidet eller opkvalificerende pakke. Vi har imidlertid kun data for hvilken model (clear cut, vent-og-se, etc.), der er brugt til at etablere opdelingen mellem de forskellige niveauer for 35.815 af eleverne.

Total antal elever i data	135.173
Elever kun med oplysninger om grundforløbstype	43.136
Elever med oplysninger om grundforløbstype og grundforløbsmodel	35.815
Elever med oplysninger om grundforløbstype og model på de 5-6 udvalgte skoler	18.574

Disse elever kan potentielt blive genstand for vores undersøgelse. Imidlertid er datamængderne for de forskellige uddannelsesretninger med samme holdkategorier og modeller i nogle tilfælde så små, at man må forvente vanskeligheder med at producere signifikante og troværdige resultater. I andre tilfælde er der usikkerhed om

datakvaliteten fordi skolerne ikke anvender Easy på en stringent måde, eller fordi vi har afdækket irregularet, hvis kilde det ikke umiddelbart har været muligt at afdække.

Databortfald fra skolerne

Alle de steder, hvor der er mulighed for fejl og hvor det giver mening, har vi bedt skolerne om at verificere vores holdidentifikationer, dvs. de steder, hvor det ikke er skolerne selv, som har lavet identifikationen for os. Skolernes administrative personale er blevet præsenteret for en liste med identificerede holdnavne, som de efterfølgende har tjekket, rettet og sagt god for.

Derefter er identifikationerne kontrolleret overfor andre identifikationsmuligheder de steder hvor det har været muligt. Som beskrevet i forrige afsnit rummer tabellen ”aktiviteter” foruden holdnavn

også to felter, ”kort betegnelse” og ”betegnelse”, som kan anvendes til at identificere holdtype, i de tilfælde hvor det administrative personale har noteret noget om holdtype her. Feltet er brugt til kontrol i de tilfælde, hvor vi i udgangspunktet har anvendt andre identifikationsmetoder – for eksempel holdnavnet eller holdenes tidsmæssige udstrækning.

Herefter er der lavet datakørsler for at se, om eleverne fordeler sig på de forskellige grundforløbstyper i overensstemmelse med de oplysninger om fordelingen som skolerne har givet.

Det har ført til bortfald i to tilfælde. På uddannelsesretningen Bil, fly og andre transportmidler på Tech College Aalborg og på Bygge-anlæg fra Silkeborg Tekniske Skole er uoverensstemmelserne mellem tidslængde og skolens notater i disse felter af en sådan karakter at vi har valgt at kassere data.

På EUC Nordvest i Thisted har vi identifikation af grundforløbshold for to uddannelsesretninger: Bygge-anlæg og Mad til Mennesker. Bygge-anlæg er imidlertid alene identificeret ud fra holdenes planlagte længde af os og ikke efterfølgende verificeret af skolen. Data viser, at kun 3 elever har fulgt en opkvalificerende pakke i perioden fra 2008 til 2010, mens 286 elever har fulgt et obligatorisk grundforløb i samme periode. Da dette ikke stemmer med de oplysninger skolen har givet i vores interview, har vi konkluderet, at der må være en fejl i identifikationen, og data herfra er derfor kasseret.

Vi havde oprindeligt forventet at kunne identificere den særlige 60 ugers grundforløbspakke, ungdomsklassen, på EUC Syd, ud fra start og sluttidspunktet i tabellen ”aktiviteter”, men gennemgangen af Easy data viste, at det ikke er muligt. Til gengæld kunne vi konstatere, at flere af holdene var navngivet i feltet ”betegnelse”. Imidlertid har vi blot kunnet knytte 30 elever til ungdomsklassen via de identificerede hold, hvilket på ingen måde passer med skolens oplysninger. Det er sikkert, at disse 30 elever har gået på ungdomsklassen, men det er også sikkert, at der har været flere elever på holdet. Endvidere er antallet så lavt, at det vil være vanskeligt at producere signifikante resultater. Det er således tvivlsomt om disse data kan bruges til noget, hvorfor vi har valgt at se bort fra dem i vores undersøgelse.

Kvalitetskontrol og rettelser

Der er også eksempler på uregelmæssigheder, som ikke er ødelæggende, da de alene berører ubetydelige datamængder eller skaber usikkerhed i et begrænset omfang. Felterne ”betegnelse” og ”kort betegnelse” er som nævnt ovenfor også brugt til kontrol af de identifikationer af link mellem grundforløbstype og hold, som er foretaget ud fra holdnavn eller holdenes start og sluttidspunkt. I fire tilfælde er feltet udfyldt i alle holdenes tabeller: Strøm, styring og IT på Syddansk Erhvervsskole, Bil, Fly og andre transportmidler på Silkeborg Tekniske Skole, EUC Vest og TEC i Hvidovre. I Odense er holdene identificeret ud fra oplysningerne i feltet. I de øvrige tre tilfælde viser kontrol med feltets oplysninger 100 % match med identifikation baseret på holdnavn eller holdets tidsudstrækning.

Mange steder er feltet imidlertid brugt usystematisk – dvs. det er tilfældigt om grundforløbstype er noteret i feltet eller ikke. I de tilfælde hvor feltet er brugt, har vi også her anvendt det til kontrol, og kontaktet skolerne i de tilfælde hvor der ikke er match.

Flere steder er enkelte hold blevet omdefineret som følge af denne kontrol. Det er sket i flere tilfælde på Herningsholm uddannelsesretningerne. Således er seks hold ud af i alt 191 på uddannelsesretningerne Bygge-Anlæg, Bil, Fly og andre Transportmidler og Produktion og Udvikling omdefineret fra obligatorisk til turbo-mix, da holdet viste sig at have været brugt til at registrerer to elever, som fulgte et obligatorisk hold i kortere tid end de øvrige, da de havde erhvervs erfaring indenfor området. To hold ud af 104 på uddannelsesretningen Strøm, styring og IT på Herningsholm er blevet omklassificeret fra obligatorisk til opkvalificerende som følge af noter i feltet.

På EUC Vest i Esbjerg har der også været mindre justeringer som følge af kontrol med feltet "betegnelse". Således er et hold på Bygge- anlæg med kun en elev som oprindeligt var klassificeret som opkvalificerende, blevet omklassificeret som følge af oplysninger i feltet – eleven ventede på at komme i skolepraktik og var ikke på et opkvalificerende hold. Der er i alt 1348 hold på Bygge- anlæg i vores data. På Krop og stil og Mad til mennesker, hvor der til sammen er 1234 hold, er et mindre antal hold, fejlagtigt klassificeret som obligatorisk, omklassificeret fordi det i feltet betegnelse var oplyst, at der var tale om et hold for mesterlære-elever. Yderligere seks ældre hold fra før 2008 på Mad til Mennesker blev omklassificeret fra obligatorisk til opkvalificerende type, fordi der i feltet var anført, at der var tale om 40 ugers obligatoriske hold. Skolen oplyser, at man tidligere har betegnet en type opkvalificerende hold på denne måde.

Andre steder har feltet bekræftet vores identifikationer i alle de tilfælde, hvor der har været noteret oplysninger om grundforløbstype i feltet. Det gælder således på Strøm, Styring og IT på Nordsjællands Erhvervsskole og Randers Tekniske Skole foruden CELF på Lolland-Falster. Atter andre steder har kontrol ikke været mulig, da feltet ikke rummer information om grundforløbstype. Det er tilfældet på samtlige hold på EUC Nordvest, EUC Lillebælt, EUC Sjælland, Holstebro, Skive, Roskilde og Københavns Tekniske Skoler, TEC, Svendborg og Nordsjællands Erhvervsskoler, Viden Djurs, foruden på Krop og stil og Produktion og udvikling på EUC Vest, Strøm, styring og IT på TECH Aalborg, og Bil, fly og andre transportmidler på Syddansk Erhvervsskole. I disse tilfælde er der således ikke mulighed for tilsvarende fejlidentifikationer som de ovenfor beskrevne.

De datakontroller vi har foretaget efter holdidentifikationen, er efterfølgende sammenlignet med det kendskab, vi har til elevernes fordeling på skolernes og indgangenes forskellige grundforløbstyper. I de tilfælde der ikke er nævnt her, svarer vores fund til de forventninger om elevernes fordeling på de forskellige grundforløbstyper.

Undersøgelse på aggregeret niveau

Der er i alt 36 erhvervsskoler i Danmark. Da vi har data fra alle afdelinger på 3 skoler og en eller flere afdelinger fra andre 17 skoler har vi således ikke data fra 16 skoler, og fra en række afdelinger på yderligere 17 skoler. Det er naturligvis utilfredsstillende at skulle udelade en så stor del af erhvervsskolerne fra vores undersøgelse – enten helt eller delvist. Derfor har vi også indsamlet data om grundforløbenes indretning og organisering fra de øvrige afdelinger på de skoler hvor vi har data, foruden yderligere 9 skoler, således at vi i alt har været i kontakt med 29 af landets 37 erhvervsskoler. I de fleste tilfælde har kontakten til disse skoler været telefonisk og med

pædagogisk personale på skolernes forskellige afdelinger, men i enkelte tilfælde også med centralt placeret personale, for eksempel studievejledere med kendskab til hele skolen.

De erhvervsskoler, som ikke har været kontaktet, er alle specialiserede skoler som for eksempel slagteriskolen i Roskilde, Jordbrugets uddannelsescenter i Århus og Hotel og restaurantskolen i København.

Kontakten til de øvrige skoler er sket for at kunne beskrive hvilken organisatorisk og indholdsmæssig indretning af grundforløbene også disse skoler betjener sig af, således at der er mulighed for at lave en undersøgelse af effekter på et mere aggregeret niveau. Selv om vi ikke kan identificere hvilken specifik grundforløbstype de forskellige elever har fulgt, kan vi med data fra Danmarks Statistik se på hvilken skole og uddannelsesretning eleverne har taget deres uddannelse. Kendes skolernes indretning og organisering af grundforløbene er det muligt at undersøge effekter på skolens samlede frafald ved at se på det samlede frafald på skolen og dens uddannelsesretninger sammenholdt med skolens indretning af grundforløbene og kontrolleret for elevernes sociale og familiemæssige baggrund.

Registerdata fra Danmarks Statistik

Det er velkendt, at frafaldsrisiko er tæt knyttet til en række sociale og økonomiske faktorer, som fx forældres uddannelse, indtægt, arbejdsmarkedstilknytning og kriminalitet (se fx Nielsen 2011; Jensen, Larsen 2011; Rumberger 2000). Derfor har vi indhentet en række data fra Danmarks Statistik til kontrol for faktorer, som kan tænkes at have betydning for elevernes frafald. De væsentligste registre her er Fertilitetsdatabasen (FTDB) og befolkningsregisteret (FAIN) som gør det muligt at identificere hvem elevernes forældre er, samt IDAP, INDH, INDK, uddannelsesregistre UDDA, UDFK og KOET, der rummer oplysninger om forældrenes indkomst, uddannelse og tilknytning til arbejdsmarkedet. En samlet opgørelse over registerdata fra Danmarks Statistik kan ses i appendiks A. En første delanalyse af registerdata viste, hvilken betydning længden af uddannelsespause har for frafaldet på erhvervsskoler, og resultatet viser, at jo mindre pause og jo større pres af elever ind i systemet desto større frafald (Munk & Park, 2013). En anden analyse viste, at unge fra selvstændige kernefamilier med faglærte forældre har en øget chance for at opnå en læreplads inden grundforløbets begyndelse, og at det er mere sandsynligt, at det giver pote på vestjyske skoler ud af de skoler vi kunne analysere (Munk & Baklanov, 2013); derudover viste analysen, at elever med en praktikplads fra starten i højere grad gennemfører deres uddannelse. Dette resultat førte frem til en afgrænsning med fokus på skolevejselever i model- og type analyserne (Munk & Baklanov, 2013).

Undersøgelingsdesign

Datakvaliteten gør det ikke muligt at lave en udtømmende undersøgelse af samtlige de modeller, analysen har identificeret på skolerne, i langt mindre grad at tage højde for samtlige de varianter der er med hensyn til tiltag. Som nævnt tidligere har vi valgt at koncentrere undersøgelsen om tre forhold:

- i) virkninger af at sætte skolesvage elever på grundforløb af længere varighed end de typisk 20 uger et obligatorisk forløb varer,

- ii) virkninger af modeller som skaber gode betingelser for en klar holdafgrænsning, hvilket formodes at styrke elevens oplevelse af socialt tilhør,
- iii) virkninger af modeller med ringe grad af niveaudeling, hvilket formodes at styrke elevernes mulighed for positiv peer-learning og svække muligheden for negativ peer-learning.

Vi har som sagt desuden valgt i første omgang at koncentrere undersøgelsen om skoler med gode data og med eksempler på udbredte modeller, som repræsenterer grundtræk af væsentlig interesse for fastholdelse. Disse modeller er clear-cut modellen, som dels med 25.445 elever er den mest udbredte model i vores datamateriale, dels stærkest prioriterer hensynet til stabilitet og afgrænsning på bekostning af hensynet til positive effekter af at have stærke og svage elever sammen.

På den baggrund er det naturligt også at fokusere på vent-og-se modellen, som med sine 10.154 er den næstmest udbredte model i vores materiale. Modellen er muligvis relativt mere udbredt på samtlige tekniske erhvervsskoler, da mange af de mindre, og fravalgte erhvervsskoler arrangerer en del af deres grundforløb på denne måde. Modellen repræsenterer også en modpol til clear-cut modellen, da den er den model, som i mest udpræget grad samler stærke og svage elever på bekostning af muligheden for stabil afgrænsning af holdene.

I analyserne anvendes vent-og-se modellen og clear-cut-modellen. Vi undersøger om der er forskelle mellem obligatoriske forløb og opkvalificeringsforløb, og eventuelle forskelle mellem vent- og-se modellen og clear-cut-modellen med hensyn til frafaldsrisiko.

Vi arbejder med følgende hypoteser:

H1: Længere grundforløb kan styrke fastholdelsen af eller i længere tid.

H1b: Modeller helt uden opdeling af elever vil ikke føre til et optimalt resultat. Modeller med en afbalanceret fordeling af svagere og stærke elever vil derimod mindske frafald og øge fuldførelse.

H1c: WAS modellen (sen opdeling) vil i større omfang styrke fastholdelsen af elever over tid i sammenligning med CC modellen (tidlig opdeling).

H2: Opkvalificerende type 1 under CC modellen reducerer frafaldet sammenlignet med det obligatoriske grundforløb.

H2b: Elever på opkvalificerende type 2 under CC har en forhøjet frafaldssandsynlighed sammenlignet med elever på obligatorisk forløb.

Vi foretager fire analyser her, dels analyser hvor vent-og-se modellen med clear-cut-modellen for en indgang sammenlignes med frafald efter et år som udfald og med overlevelsestiden indtil begivenheden frafald som udfald (uddannelseslængden indtil frafald), dels en analyse hvor vi opererer med 29 hold udelukkende baseret på CC-modellen, 15 for opkvalificeringstype 1, og 14 for opkvalificeringstype 2.

Sammenligning af WAS og CC-modeller

I modelanalyserne undersøgte vi om de tekniske skolers forskellige måder at organisere grundforløb påvirker frafaldet. Registreringspraksis på skolerne indebærer, at analysen omfatter en væsentlig delmængde af skoler på bygge og anlæg (samt visse andre indgange), men ikke alle. I den første delanalyse sammenlignes to hovedmodeller – den såkaldte vent-og-se (WAS) model med sen opdeling, og clear-cut (CC)-model med tidlig opdeling – for at undersøge, om tidslængden spiller en rolle. Vores mål er at sammenligne de to alternative organiseringsmodeller henblik på frafald og tidspunkt for frafald.

Metode:

Medens den typiske skolefrafaldsforskning primært er baseret på binære valganalyser af de observerede udfald på et givet tidspunkt efter skolestart (f.eks. logit), kigger kun få på tidspunktet og udviklinger af frafald over tid (Willet & Singer, 1991; DesJardins et al., 1999; DesJardins et al., 2002). Vi implementerer et alternativ til den binære valganalyse: varighedsanalyse. Varighedsanalyser beskæftiger sig med varigheden indtil et givent udfald indtræffer (overlevelsestiden), således er den specifikke tid for en begivenhed det centrale element¹. Vi undersøger således overlevelsestiden, indtil begivenheden frafald indtræffer, i stedet for kun at undersøge frafaldet på et givet tidspunkt (fx efter et år). Et centralt begreb indenfor varighedsanalyse er hazard-funktionen, der angiver den øjeblikkelige sandsynlighed for forekomst af en given begivenhed på tidspunkt t , givet at begivenheden endnu ikke er indtruffet (Cameron & Trivedi, 2005). Den simpleste og mest udbredte varighedsmodel er Cox Proportional Hazard (PH) modellen (Cox, 1972), formelt:

$$\lambda(t|\mathbf{x}) = \lambda_0(t)\phi(\mathbf{x}, \beta)$$

hvor λ_0 er baseline hazarden som funktion af t ; mens ϕ er en funktion af en vektor af kovariater, \mathbf{x} , og normalt antages $\phi(\mathbf{x}, \beta) = \exp(\mathbf{x}'\beta)$. Den største fordel ved denne model, er at den i modsætning til alternativerne ikke antager en bestemt funktionel form for baseline hazard-funktionen. Da Cox PH modellen er en semi-parametrisk model, kombinerer den en nonparametrisk estimering af hazard-funktionene og en parametrisk estimering af koefficienterne for kovariaterne. I Cox modellen antages derimod at de forklarende variable har en proportional effekt på reference-hazarden, dvs. ændringer i disse variable kan fortolkes at have en multiplikativ effekt på reference-hazarden (Cameron & Trivedi, 2005). Dermed muliggør Cox model estimering af de forklarende variables effekt på hazarden for frafald og estimering af den kontrollerede hazard-funktion. Vi kan bruge denne model til at teste hvorvidt WAS elever, som falder fra, gør det senere eller tidligere end CC.

En udvidelse af varighedsanalyse er competing-risk (CR) analyse, som tillader at flere konkurrerende udfald kan forekomme. Competing-risk muliggør dermed modeller således at varigheden til frafald modelleredes under antagelsen af, at elever kan færdiggøre eller falde fra i samme periode. Fordelen ved competing-risk modellen er dermed, at vi kan bruge alle tilgængelige

¹ Lignende metoder er implementeret i f.eks. Feng & Sass (2012) og Plank & DeLuca (2008).

observationer, medens ulempen er, at vi bliver nødt til at foretage nogle strenge antagelser. En af den simpleste CR model, foreslået af Kalbfleisch & Prentice ([1980]2002), er en udvidelse af Cox varighedsmodellen, med den udfaldsspecifikke hazard funktion:

$$\lambda_j(t|\mathbf{x}) = \lambda_{0j}(t)\exp(\mathbf{x}'\beta_j)$$

hvor j angiver det specifikke udfald: i vort tilfælde $j = (\text{frafald, fuldførelse})$. Denne model kræver en antagelse om at udfaldene er uafhængigt af hinanden (sammenligneligt med IIA antagelsen i multinomial logit modellen). Antagelsen betyder, at begge udfald vil ske, men at vi kun har tiden for den første. Dette kan til en vis grad antages i vores tilfælde, dvs. hvis eleven ikke falder fra vil han/hun eventuelt fuldføre uddannelsen. Det virker yderligere fornuftigt at antage at hazarden for frafald og fuldførelse ikke vil følge den samme udvikling over tid, da disse udfald er påvirket af væsentligt forskellige overvejelser og tidsmæssige begrænsninger. Dermed kan vi antage en højere risiko for frafald i begyndelsen af en uddannelse, for eksempel på grund af uddannelsesom- og fejlvalg; medens fuldførelse vil først være sandsynlig at forekomme senere i forløbet, efter at vise krav er indfriet. Da der ikke er noget helt egnet statistisk test for antagelsen, estimerer vi under de samme modelspecifikationer den mere generelle og robuste Fine-Gray CR model (Fine & Gray, 1999). Da estimererne for denne model er i overensstemmelse med Cox CR modellen, taler det for den sidstes anvendelighed. Vi foretrækker Cox CR modellen til hovedanalysen, pga. af dennes fleksibilitet og bredt udvalg af udvidelser. I modsætning til Fine-Gray modellen, der kun betragter den aggregerede effekt af en forklarende variabel, muliggør Cox CR modellen en opdeling af den direkte effekt af skolemodel på frafald og den indirekte effekt på frafald fra ændringer i fuldførelsesvarigheden (Dignam *et al*, 2012; Wolbers *et al*, 2009).

Cox CR modellen er begrænset i den måde de forklarende variable behandles. Der antages såkaldte proportionale hazarder, dvs. de forklarende variable antages at påvirke frafaldsrisikoen med en konstant faktor i forhold til baseline hazarden over hele perioden. Dermed tillader den simple Cox CR model ikke, at en forklarende variabel kan påvirke frafaldet anderledes på forskellige tidspunkter. Der findes dog modeludvidelser for Cox, der kan bruges til at løse dette problem: f.eks. den stratificerede Cox CR model, hvor baseline hazarden kan variere frit for forskellige værdier af en given variabel (den stratifikationsvariablen), medens alle andre koefficienter er konstante. Da denne model ikke giver en koefficient for stratifikationsvariablen, bruger vi modellen til at lave en graf over reference-hazarden for hver strata. Den stratificerede Cox CR model kan udtrykkes som:

$$\lambda_{jk}(t|\mathbf{x}) = \lambda_{0jk}(t)\exp(\mathbf{x}'\beta_j)$$

Hvor k indikerer strata. Dermed kan vi se at i denne model kan hazard ratioerne variere over begivenhed som før, mens de er konstante over begge værdier af den stratificerede variabel (dvs. over modeltype – WAS/CC). Baseline hazarden varierer derimod både over begivenhed og strata. Vi antager dermed, at de forklarende variable har forskellige virkninger på hhv. frafald og fuldførelse men er konstante over de to skolemodeltyper. Vi bruger således den stratificerede Cox CR model for at estimere og sammenligne de udfalds-specifikke hazard funktioner for hhv. WAS og CC elever, medens vi kontrollerer for baggrundsvariable.

For at bedre visualisere den aggregerede modeeffekt på frafald, medens vi kontrollerer kovariater, kan vi beregne den kumulative incidents (CI) kurve for frafald på baggrund af den stratificerede Cox CR model. CI kan beregnes som:

$$I_1(t) = \int_0^t \lambda_1(s) \exp\left(-\int_0^s \lambda_1(u) + \lambda_2(u) du\right) ds, \text{ hvor } j = (1,2)$$

som måler den aktuelle risiko for at begivenhed $j=1$ indtræffer inden tidspunktet t : $I_1(t) = P(T \leq t \text{ and } k = 1)$ (Kalbfleisch & Prentice, 2002; StataCorp, 2009). Et særligt aspekt ved CI funktionen er, at den afhænger både af den udfalds-specifikke hazard for den interesserede begivenhed men også for den konkurrerende begivenhed. Dermed ved at estimere CI kurven for frafald, kan vi observere den aggregerede skolemodeeffekt: dvs. den direkte effekt fra påvirkning af hazarden for frafald og den indirekte fra påvirkning af fundførelse.

Data

Analysedata er opbygget fra registerdata fra Danmarks Statistik (DST). Datasættet indeholder både respons- og forklarende variable, hvor den sidst nævnte inkluderer demografiske karakteristika, familiebaggrund og tidligere uddannelseshistorik. Da skoleanalysen viste, at modeltypen (f.eks. WAS) typisk er konstant over forskellige hold indenfor den samme indgang på en skole eller afdeling for en given årgang, har vi kunnet kombinere resultaterne fra skoleanalysen og DSTs komprimerede elevregister (KOET) register for 2011. Dermed kan vi knytte modeltype til skole og elev. Denne tilgang muliggjorde inklusion af flest mulige skoler til analysen. Dog giver KOET ikke mulighed for at skelne mellem forskellige retninger indenfor samme indgang (f.eks. tømrere og VVS), og derfor har vi ikke kunnet bruge tre skoler, hvor WAS blot var brugt på dele af BA. Vi kunne derimod godt skelne mellem forskellige afdelinger af samme hovedskole (f.eks. KTS Glostrup og Herlev), da institutionsnumre indeholder kommunekoderne. Således har vi identificeret 4 skoler, der bruger WAS modellen. Til sammenligning har vi brugt skoler, der primært anvender CC modellen på BA. På nogle skoler er det dog ikke muligt at skelne mellem CC obligatoriske hold og nogle opkvalificeringshold, der bruger f.eks. præ-pakke modellen. Således kan kontrolgruppen ikke betragtes som en ren CC model.

Tabel 1. Oversigt over skoler brugt til analysen

Model	Skole	2008	2009	2010	Afdeling
Wait-and-See	6 Randers TS	x	x	x	Alle
	9 Syddansk Erhvervsskole	x	x	x	Alle
	12 Silkeborg	x	x		Alle
	15 KTS	x	x	x	Glostrup og Herlev
Clear-Cut	1 EUC Vest	x	x	x	Alle
	3 EUC Nordvest	x	x	x	Thisted
	8 Herningsholm	x	x	x	Alle
	10 CELF	x	x	x	Nykøbing og Nakskov
	13 Holstebro TS	x	x	x	Alle
	14 TECH Ålborg	x	x	x	Alle

Tabel 1 giver en oversigt af de skoler og afdelinger, der bruger enten WAS eller CC for the periode 2008 to 2010. Den endelige analysepopulation omfatter fire WAS skoler og seks CC skoler.

Tabel 2. Oversigt over identificerede elever som andel af den samlede BA-tilgang pr. begyndelsesår

	Kun skolevejs elever				Alle elever			
	2008	2009	2010	Samlet	2008	2009	2010	Samlet
CC	17.0	17.1	18.2	17.4	19.1	18.1	19.0	18.8
WAS	16.6	17.1	14.4	16.0	16.6	16.7	14.2	15.9
CC & WAS	11.5	11.0	11.1	5.9	10.7	10.4	10.6	10.6
Andre modeller	32.5	33.4	33.8	38.5	30.1	32.1	32.4	31.5
Ikke undersøgt	22.4	21.4	22.5	22.1	23.5	22.7	23.8	23.3
Total	7004	6878	6784	20666	9070	8022	7855	24947

Tabel 2 viser fordelingen af alle nye elever på Bygge og Anlægs indgangen for perioden 2008-2010 over modeltype. Det kan således ses, at 78% af eleverne gik på en skole, som var med i vores undersøgelse; af denne gruppe var 60% organiseret efter enten WAS eller CC modellerne. Da 14% gik på skoler, hvor det ikke var muligt at skelne mellem WAS og CC elever, bruger vi 45% af den samlede population af nye elever for perioden til analyserne.

En præliminær undersøgelse af elevpopulationen viste store forskelle mellem skolevejs- og praktikvejs elever – både i forhold til baggrundskarakteristika og frafaldsadfærd. Det tyder på at praktikvejs elever er udsat for en helt anderledes uddannelsesoplevelse og -proces med anderledes incitament og forventninger, der samlet bidrager til en særdeles anderledes frafaldsudvikling for denne gruppe. Ligeledes afhænger sandsynligheden for at påbegynde på praktiksvejsuddannelse af elevens baggrundsvARIABLE (Appendiks G): unge fra kernefamilier med selvstændige eller faglærte forældre har en øget chance for at opnå en læreplads inden grundforløbets begyndelse og dermed ende på en praktikvejsuddannelse. Det er således også uklart hvordan og i hvilken grad praktikvejs elever bliver påvirket af skolens modelvalg, da de tilbringer en del af deres uddannelse på udenfor skolerne og bliver dermed mindre påvirket af skolens organisering af grundforløb. Derfor har vi valgt at udelukke fokusere skolevejs elever i vores analyser.

WAS elever er identificerede gennem en dummyvariabel (1=WAS; 0=CC). Derudover inkluderede vi en række demografi- og familiebaggrundsvARIABLE (målt to år inden uddannelsesstart); samt kontrolvariable for startsår, sommerstart (start før august) og skole; gennemsnitseksamens karakterer fra grundskoler i dansk og matematik, samt dummy variable for manglende værdier.

En ulempe ved data er, at vi ikke er i stand til at adskille model- og den (indgangsspecifikke) skoleeffekt på frafald, da modellen er konstant indenfor hele indgangen på en skole. Dermed er koefficienten for WAS variabelen et udtryk for den samlede model- og skoleeffekt. Der er flere potentielle metoder, der kan bruges for at tackle dette identifikationsproblem: f.eks. fixed-effekt (Cameron & Trivedi, 2005) og såkaldte placebo regressioner (f.eks. Olsson, 2009). Fixed-effekt regression kan potentielt bruges til at estimere en separat skolemodeffekt ved at sammenligne elever før og efter skolereformen i 2007. Denne metode antager, at skoleeffekten er konstant over tid og reformen. Desværre kan vi ikke bruge denne metode, da denne kræver, at vi skulle have kendskab til organisationsmodellen på skolerne inden 2007 reformen. Placebo regressioner kan isolere en skoleeffekt ved tilføje en eller flere yderligere indgange for hver skole med konstant organisationsmodel, dvs. vi ønsker at tilføje en ”placebo” indgang, der varierer over skole men ikke

over model. Disse yderligere indgange ville muliggøre estimering af en reference skoleeffekt, der er adskilt fra modeffekten. For at denne metode skulle være succesfuld, kræver det, at skoleeffekten er konstant over indgange og at forskelle mellem indgange (indgangseffekten) er konstante over skoler. Den deskriptive statistik af frafald på Bygge og Anlæg samt to mulige "placebo" indgange – Strøm, Styring og IT (SSI) og Bil, fly og transport (BFT) – viser imidlertid, at skoleeffekten varierer usystematisk over indgange, og at indgangsforskelle varierer usystematisk over skoler. Dermed ville metoden føre til fejlestimerede skole- og indgangseffekter og medfører bias i estimatet for modeffekten. Dette støttes af en række prøve-regressioner, hvor WAS estimatet var meget ustabil i signifikans, størrelse og retning, afhængigt af hvilke skoler og indgange som inkluderes.

Givet utilgængeligheden af en succesfuld identifikationsmetode, forslår vi at bruge den kombinerede model- og den indgangsspecifikke skoleeffekt for at afdække den overordnede forskel i frafaldsniveau og -tidspunkt mellem CC og WAS skoler. Således skal de følgende estimater fortolkes med omhu. Den kombinerede effekt har dog også en naturlig samlet fortolkning, idet at modeffekten kan betragtes som en del af den mere generelle indgangsspecifikke skoleeffekt. Ved at bruge adskillige skoler for hver model kan vi estimere en gennemsnitseffekt. Da vi yderligere inkluderer skole dummy variable, opfanger vi noget af skoleforskellene, mens den fælles modeffekt over skoler, bliver delvis renset.

Resultater

Vi estimerer to empiriske modeller som beskrevet under modelafsnittet: en simpel Cox CR model med WAS som en forklarende dummy variabel og under antagelse af proportional hazard for modeffekten; og en stratificeret Cox CR model, hvor reference hazarden estimeres separat for hver skoleorganisationsmodel, medens de andre forklarende variable antages have konstant effekt på tværs af organisationsmodel.

Tabel 3 viser estimaterne fra fire simple Cox CR regressioner for hver begivenhed (frafald og fuldførelse), hvor vi har kontrolleret for forskellige baggrundsvariable. Disse regressioner er udført for at fastlægge den generelle retning og størrelse af effekten på frafaldsniveauet fra WAS og de andre forklarende variable, samt at observere ændringerne i den estimerede WAS koefficient, da vi kontrollerer for yderligere baggrundsvariable. Estimaterne præsenteres som hazard-ratioer, som fortolkes som den relative afvigelse fra baseline hazarden, dvs. en hazard ratio på 0.90 svarer ca. til et fald på 10% i hazard for dette udfald. I de først tre regressioner er WAS estimatet konstant med en hazard ratio af 0.87, dvs. WAS elever har en 13% lavere frafaldsrisiko sammenlignet med CC elever. Tilsvarende for de samme regressioner men for fuldførelse har WAS en hazard ratio på 1.25. Dermed ser det ud til, at på WAS skoler øges fuldførelsesraten og frafaldsrisikoen mindskes. Når skoledummyer inkluderes (model IV) øges forskellen i frafaldsrisikoen til 17%, hazard ratioen for fuldførelse falder og mister signifikans. Regressionerne indikerer også større større variation i frafald mellem WAS skoler: KTS virker til at være den bedste til at fastholde sine elever mens Randers klarer sig betydeligt ringere.

Table 3: Cox CR model – Varighed indtil hhv. frafald og fuldførelse (Bygge og Anlæg)

	Frafald				Fuldførelse			
	(I)	(II)	(III)	(IV)	(I)	(II)	(III)	(IV)
WAS	0.870*** (0.00)	0.874*** (0.00)	0.868*** (0.00)	0.833* (0.02)	1.250*** (0.00)	1.243*** (0.00)	1.250*** (0.00)	1.097 (0.24)
<i>Startår (ref: 2008)</i>								
2009	0.816*** (0.00)	0.813*** (0.00)	0.811*** (0.00)	0.803*** (0.00)	0.835*** (0.00)	0.823*** (0.00)	0.821*** (0.00)	0.817*** (0.00)
2010	0.947 (0.24)	0.950 (0.28)	0.954 (0.32)	0.940 (0.20)	0.831*** (0.00)	0.808*** (0.00)	0.796*** (0.00)	0.789*** (0.00)
Sommer	1.412*** (0.00)	1.404*** (0.00)	1.401*** (0.00)	1.400*** (0.00)	1.026 (0.56)	1.037 (0.40)	1.045 (0.32)	1.012 (0.78)
Kvinde	1.366*** (0.00)	1.321*** (0.00)	1.318*** (0.00)	1.311*** (0.00)	1.849*** (0.00)	1.963*** (0.00)	1.983*** (0.00)	1.940*** (0.00)
Kernefamilie (15-årsalderen)	0.764*** (0.00)	0.804*** (0.00)	0.820*** (0.00)	0.820*** (0.00)	1.355*** (0.00)	1.248*** (0.00)	1.192*** (0.00)	1.185*** (0.00)
<i>Ethnicitet (ref: Dansk)</i>								
Indvander	1.224** (0.01)	1.113 (0.21)	1.062 (0.47)	1.048 (0.58)	0.653*** (0.00)	0.819* (0.05)	0.914 (0.36)	0.864 (0.13)
Efterkommere	1.420*** (0.00)	1.281* (0.02)	1.232* (0.04)	1.270* (0.02)	0.694*** (0.00)	0.860 (0.20)	0.958 (0.71)	0.944 (0.62)
Allder	1.104*** (0.00)	1.094*** (0.00)	1.064* (0.02)	1.064* (0.02)	0.996 (0.84)	1.009 (0.66)	1.052* (0.02)	1.043 (0.06)
Allder ²	0.998** (0.00)	0.999** (0.01)	0.999* (0.04)	0.999* (0.04)	1.001 (0.23)	1.000 (0.28)	1.000 (0.89)	1.000 (0.84)
<i>Farens uddannelse (ref: grundskole)</i>								
Erhvervsfaglig	-	0.892* (0.01)	0.901* (0.02)	0.896* (0.01)	-	0.992 (0.84)	0.968 (0.42)	0.968 (0.43)
Gymnasial	-	0.927 (0.59)	0.953 (0.73)	0.958 (0.76)	-	0.945 (0.68)	0.886 (0.35)	0.869 (0.27)
Kort/mellemlang videregående	-	0.877 (0.09)	0.886 (0.12)	0.879 (0.09)	-	1.075 (0.27)	1.022 (0.74)	1.007 (0.91)
Lang videregående	-	0.886 (0.44)	0.893 (0.47)	0.908 (0.54)	-	0.848 (0.20)	0.825 (0.11)	0.855 (0.18)
<i>Morens uddannelse (ref: grundskole)</i>								
Erhvervsfaglig	-	0.908* (0.03)	0.921 (0.07)	0.924 (0.08)	-	1.156*** (0.00)	1.121** (0.01)	1.115** (0.01)
Gymnasial	-	0.958 (0.70)	0.978 (0.84)	0.962 (0.73)	-	1.161 (0.13)	1.111 (0.27)	1.086 (0.38)
Kort/mellemlang videregående	-	0.930 (0.35)	0.951 (0.51)	0.959 (0.59)	-	1.124 (0.07)	1.073 (0.26)	1.069 (0.29)
Lang videregående	-	1.098 (0.53)	1.135 (0.39)	1.144 (0.36)	-	0.850 (0.32)	0.781 (0.12)	0.774 (0.10)
<i>Farens beskæftigelse (ref: lønmodtager lavt niveau)</i>								
Selvstændig	-	1.021 (0.79)	1.026 (0.74)	1.026 (0.74)	-	1.116 (0.07)	1.098 (0.12)	1.103 (0.11)
Lønmodtager, mellem/højt niveau	-	1.131 (0.08)	1.139 (0.07)	1.130 (0.09)	-	0.956 (0.46)	0.957 (0.45)	0.954 (0.43)
Lønmodtager, andet	-	1.064 (0.30)	1.065 (0.29)	1.066 (0.28)	-	0.908 (0.09)	0.903 (0.07)	0.892 (0.05)
Ledig	-	1.025 (0.82)	1.025 (0.81)	1.024 (0.82)	-	0.834 (0.08)	0.829 (0.08)	0.805* (0.05)
Udenfor arbejdsmarkedet	-	1.135* (0.04)	1.128* (0.05)	1.133* (0.04)	-	0.964 (0.56)	0.978 (0.72)	0.977 (0.71)
<i>Morens beskæftigelse (ref: lønmodtager lavt niveau)</i>								
Selvstændig	-	0.776* (0.04)	0.768* (0.04)	0.777* (0.05)	-	1.125 (0.18)	1.132 (0.16)	1.133 (0.16)
Lønmodtager, mellem/højt niveau	-	0.945 (0.43)	0.963 (0.60)	0.974 (0.71)	-	1.042 (0.47)	0.995 (0.93)	1.022 (0.71)
Lønmodtager, andet	-	1.018 (0.79)	1.025 (0.72)	1.036 (0.60)	-	1.017 (0.79)	0.990 (0.87)	1.011 (0.85)
Ledig	-	1.111 (0.19)	1.091 (0.28)	1.081 (0.34)	-	0.967 (0.67)	1.006 (0.93)	0.985 (0.86)
Udenfor arbejdsmarkedet	-	1.189** (0.00)	1.179** (0.00)	1.183** (0.00)	-	0.876** (0.02)	0.881* (0.02)	0.893* (0.04)
Familieindkomst (gennemsnit)	-	0.999 (0.94)	1.002 (0.91)	1.002 (0.91)	-	1.053*** (0.00)	1.049*** (0.00)	1.043*** (0.00)
Eksamenstype, 9.kl.	-	-	0.945*** (0.00)	0.942*** (0.00)	-	-	1.130*** (0.00)	1.123*** (0.00)
<i>Skole (ref: 15. KTS (WAS))</i>								
Randers TS (WAS)	-	-	-	1.499*** (0.00)	-	-	-	1.355** (0.00)
Syddansk Erhvervsskole (WAS)	-	-	-	1.007 (0.93)	-	-	-	0.913 (0.21)
Silkeborg TS (WAS)	-	-	-	1.205 (0.17)	-	-	-	0.987 (0.91)
EUC Vest (CC)	-	-	-	1.121 (0.10)	-	-	-	0.892 (0.08)
EUC Nordvest (CC)	-	-	-	1.080 (0.53)	-	-	-	1.119 (0.29)
Hermingsholm Erhvervsskole (CC)	-	-	-	1.001 (0.99)	-	-	-	0.919 (0.26)
CELF (CC)	-	-	-	0.983 (0.81)	-	-	-	0.372*** (0.00)
Uddannelsescenter Holstebro (CC)	-	-	-	1.101 (0.32)	-	-	-	1.168 (0.09)
Observationer	6843	6843	6843	6843	6843	6843	6843	6843
Pseudo R-squared	0.006	0.007	0.008	0.009	0.007	0.009	0.012	0.016
AIC	47519.614	47504.213	47470.748	47454.188	57312.585	57245.567	57059.099	56846.241
BIC	47594.755	47743.297	47723.495	47761.582	57387.726	57484.651	57311.846	57153.635

Ann.: Hazard Ratioet; p-værdier i parentes. Statistisk signifikante estimater er markeret: *, 5%; **, 1%; ***, 0.1%. Kategorier for manglende værdier er inkluderet. Efron method for ties (Cleves et al., 2010: s.151)

De demografiske variable virker til at have en betragtelig påvirkning af hazard ratioerne: kvindelige elever har højere hazard for både frafald og fuldførelse, hvilket gør den samlede effekt utydelig. Elever med indvandrebaggrund har højere risiko for at falde fra; denne effekt er stærkest for andengenerations-indvandrere med en hazard ratio på 1.2. Elever som kommer fra en kernefamilie har signifikant lavere hazard for frafald, på 18%. Alder har en stigende positiv effekt på frafald.

Elever med forældre med erhvervsuddannelse er mindre tilbøjelige for at falde fra sammenlignet med andre forældreuddannelsesgrupper; resultatet ser ud til at holde for begge forældre. Elever med forældre udenfor arbejdsmarkedet har større hazard ratio for at falde fra (ca. 13% for faderen og ca. 18% for moderen), medens de med en selvstændig mor har betydeligt lavere risiko sammenlignet med børn af en ufaglært ansat mor. Forældres indkomst påvirker ikke hazarden for frafald, men virker til at øge hazarden for fuldførelse. Endeligt har grundskolekarakterer en negativ effekt på frafald og en positiv effekt på fuldførelse.

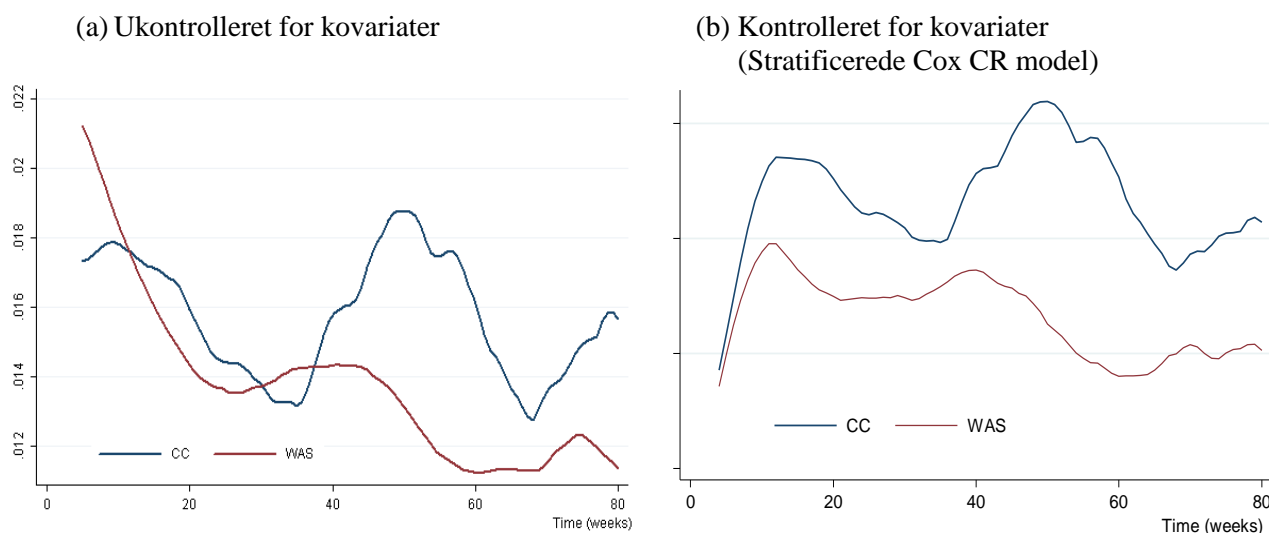
Som nævnt før kan den simple Cox CR model kun fortælle os, at WAS overordnet bidrager til at reducere hazarden for frafald. Imidlertid når vi estimerer den udglattede hazard funktion ved brug af en nonparametrisk metode (figur 1 (a)) uden at kontrollere for baggrundsvariable, kan vi se at hazarden for WAS udvikler sig anderledes over tid i forhold til CC. For at kontrollere for denne differens estimerer vi en stratificeret Cox CR model, hvor baseline hazarder tillades at variere mellem WAS og CC, mens de andre variable estimeres på tilsvarende måde som i tabel 3. De estimerede hazard ratioer for de fire stratificerede Cox CR regressioner fremlægges i Appendiks F. De estimerede hazard ratioerne er næsten identiske med Tabel 3, med undtagelsen af WAS, der ikke længere har et estimat.

Ved at bruge model IV i Appendiks F som beregningsmodel (funktion) kan vi estimere og sammenligne reference hazarden for både WAS og CC, mens vi kontrollerer for baggrundsvariable (Cleves *et al*, 2010; Royston, 2011). De estimerede reference hazard kurver er rapporterede i figur 1 (b). For at estimere de udglattede hazard funktioner har vi brugt den alternative Epanchikov kernel (epan2 i Stata). Vi har afprøvet flere forskellige kerneller inklusiv den Gaussianske og den ordinære Epan(chikov) kernel (Cleves *et al*, 2010), medens den alternative Epanchikov kernel var den bedste til at fange variansen i vore data. Især omkring enderne af funktionen var denne kernel overlegen, da den bruger en separat kernel omkring endepunkterne for at tage højde for venstrecensurering af beregningsværdier tæt på $t=0^2$. Dette tillader os mere nøjagtigt at estimere hazarden for frafald kort efter indskrivning. Vi bruger en båndbredde på $h=9.0$, der specificerer vinduet af data omkring hvert punkt der skal bruges til udglatningen. Vi valgte denne værdi baseret på en stribe forsøg for at finde den mest retvisende afbildning af data. Vi har også forsøgt med en optimeringsmetode, hvor den båndbredde der minimerer variansen vælges; men denne metode blev forkastet, da den giver overglattede kurver.

² Det samme gælder højrecensurering tæt på max. antal uger, dog er det i vort tilfælde ikke relevant, da vi kun kigger på uddannelseslængde op til 80 uger.

Figur 1 viser udviklingen i den udfalds-specifikke hazard for frafald over de første 80 uger efter uddannelsesstart. Når vi betragter den kombinerede hazard for begge skolemodeller, fås at hazardkurven starter ret højt omkring 10 uger og falder efterfølgende over de næste 25 uger (indtil knap 40 uger); efter 40 uger vokser hazarden igen over de næste 10-20 uger; hvorefter den igen falder indtil uge 70 efterfulgt af en mindre stigning. Dette indikerer, at risikoen for frafald er relativt høj i begyndelsen af uddannelsen, men falder i takt med at mere tid investeres i uddannelsen; derudover virker der til at være en skarp (men aftagende) stigning i risikoen efter hvert af de to skæringspunktet (normeret uddannelseslængder af hhv. 40 og 60 uger³). Dette tyder på, at risikoen for frafald er højest i starten af uddannelsen, hvor frafaldet er karakteriseret ved uopfyldte forventninger for uddannelsesvalget eller dårlig motivation og forudsætninger; og omkring 40 uger, når den normerede uddannelseslængde udløber.

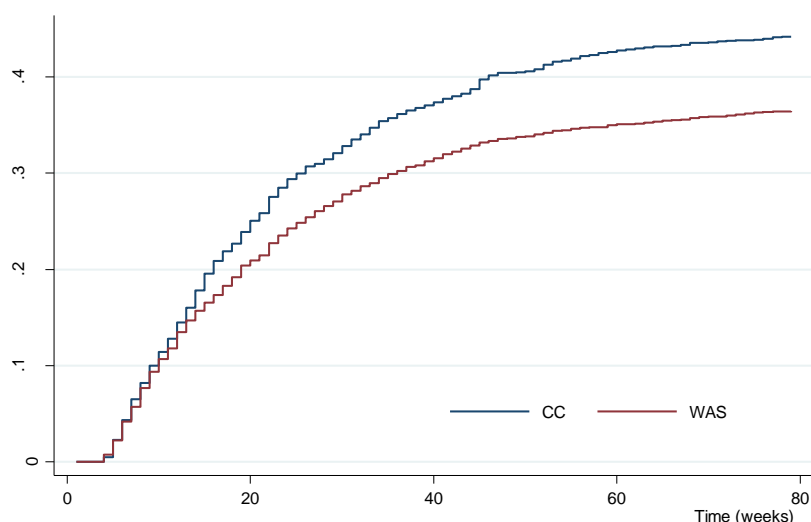
Del (b) af Figur 1 viser, at den estimerede hazard funktion for WAS ligger under den for CC i hele perioden. Overordnet viser figuren, at udviklingen i hazarden for WAS er mere jævn end for CC og har færre udsving i hazarden. Differencen mellem de to hazard funktioner topper to gange: en mindre top mellem 15 og 20 uger; og en større top mellem 50-60 uger. Dermed tyder det på, at den ellers relativt høje frafaldsrisiko ved uddannelsesstart er lavere på WAS skoler i forhold til CC skoler. Endvidere indikerer figuren, at en stigning i risikoen mellem 40-50 uger er betydeligt stejlere for CC elever end for WAS, der i stedet har en mindre stigning omkring 40 uger efterfulgt af et fald umiddelbart efter (dvs. 10 uger inden faldet hos CC elever). Dette virker plausibelt, da skolerne under WAS-modellen får mulighed for at forlænge elever med en større frafaldsrisiko op til uddannelsesslutningen (20 eller 40 uger); medens CC studerende er tvunget til at fortsætte i det samme forløbsspor og tage eksamen.



Figur 1: Den udglattede hazard funktion for WAS og CC (a) ukontrolleret and (b) kontrolleret for baggrundvariable

³ For 60 uger virker der til at være mindre forsinkelse, som sandsynligvis opstår på grund af medregnede ferieuger.

Når vi sammenligner de estimerede hazarder kontrolleret for baggrund i del (b) og den ”rå” nonparametriske hazard i (a), ser vi overensstemmelse i toppene. Den første halvdel af hazard kurven for WAS virker til at være lavere, når vi kontrollerer for baggrund: dette tyder på, at WAS eleverne har ringere baggrundsforudsætninger i forhold til CC elever (dette understøttes af den deskriptive statistik i Appendiks B.2). I begge figurer adskiller perioden mellem 50 og 60 uger sig som den med den største forskel i frafaldsrisikoen mellem de to skolemodeller.



Figur 2: Den kumulative incidens (Stratificeret Cox CR model)

Figur 2 viser den kumulative incidens for frafald for WAS og CC. Igen bruger vi den stratificerede Cox CR model (Appendix F), hvor de udfalds-specifikke hazarder for frafald og fuldførelse estimeres på baggrund af grundmodellen (hhv. for frafald og fuldførelse). Det ses at den aggregerede frafaldsrate for de to modeller udvikler sig overordnet ens: stiger kontinuert for hele perioden men med aftagende kraft. Frafallsraterne for de to modeller begynder dog at skilles kort efter start med en voksende forskel. Udviklingen i CI er i overensstemmelse med Figur 1: den første større afvigelse sker mellem 10 og 20 uger og den anden fra mellem 40 og 50 uger. Omkring 20 uger er rateforskellen mellem de to modeller omkring 3 pct.point medens den cirka er fordoblet efter 60 uger. Generelt viser figuren, at elever på WAS modellen, har en lavere risiko for frafald i forhold til CC elever. Denne forskel er gældende over hele perioden og akkumuleres til risikofald på ca. 6-7 efter 60 uger.

Varighedsanalyserne viser samlet, at elever i gennemsnit falder fra senere, og har et lavere frafald på skoler, hvor WAS-modellen anvendes, sammenlignet med skoler hvor CC-modellen anvendes. Det betyder, at senopdeling i gennemsnit giver et lavere frafald end tidlig opdeling (jf. Gamoran, 2010). Vi fandt, at frafallsforskellen mellem de to modeller over tid især topper omkring 20 uger og efter 40 uger, og at frafaldet for elever på WAS-modellen udvikler sig mere jævnt sammenlignet med elever på CC-modellen. Elever på vent-og-se forløbet erfarer formentlig et mere dynamisk og måske fleksibelt forløb, da de undervejs får mulighed at komme på forlængede hold, som i højere grad er afstemt med forudsætninger og kunnen, simpelthen fordi informationen om eleverne, som skolen og lærerne råder over efter 15 ugers forløb, er større. Skolerne får desuden med WAS-

modellen mulighed for at forlænge elever med en større frafaldsrisiko efter 18-20 uger, og efterfølgende for at blande eleverne mere hensigtsmæssigt, end det er tilfældet på CC-modellerne, hvor elever i vid udstrækning under de forlængede forløb fortsætter på de samme hold.

En alternativ forklaring på, at elever under WAS-modellen gennemsnitligt har et lavere frafald, er, at ressourcetsvage elever faktisk får gavn af først at gå sammen med andre elever, som afslutter deres forløb efter 20 uger, og naturligvis også med nogle elever, som falder fra før de 20 uger. Forlængelsen af WAS-forløbet er måske alt andet lige bedre end at anvende end CC-modellen, idet eleverne ikke opdeles de første 18-20 uger og derved formentlig får gavn af at møde stærke elever (positiv peer-læring), dog under forudsætning af en balanceret blanding af stærke og mindre stærke elever. Omvendt kan det ikke udelukkes, at WAS-hold kan indebære en negativ peer-effekt for ressourcetsvage elever i forhold ressourcestærke elever, fordi den reelle grund til mindre risiko for frafald hænger sammen med, at de ressourcetsvage elever får mindst to forsøg til at gennemføre uddannelsen, og at de altså (kun) gennemfører pga. muligheden for ekstra forsøg. Endelig kan en mulig højere lærerforventning på WAS-modellerne alt andet lige bidrage til at holde frafaldet nede.

Sammenligning af grundforløbstyper under CC-modellen

I denne del kigger vi nærmere på grundforløbstyper indenfor Clear-Cut modellen for at se hvorvidt der er en forskel.

I udgangspunktet undersøges om de 15 + 14 hold, som vi får ved at dele op i forhold til 5 skoler, indgange og år, fører til et reduceret frafald. Dels har vi foretaget separate analyser delt op på skoler, indgang og år, dels har vi sammenlagt de 15 hold under opkvalificeringstype 1 for at få tilstrækkeligt med observationer til at kunne udsige noget mere generelt om tiltagene, hvilket vi også har gjort for opkvalificeringstype 2 med 14 hold. Vi har opdelt begge analyser efter om eleverne fulgte et opkvalificerende forløb 1 eller et opkvalificerende forløb 2. Designet følger således en tankegang, hvor vi forsøger at undersøge om det er muligt at påvise betydningen af skolernes indsats mod frafald.

Vort primære design for undersøgelsen af de 15 og de 14 holdstudier følger en logik hvor vi forsøger at forbedre analyserne gennem brug af interne kontrolgrupper se fx Shadish et al. (2002: 122). Den interne kontrolgruppe vil bestå af elever fra samme skole og indgang, så at sige fra samme population, som startede på et obligatorisk hold, så de er sammenlignelige. Ifølge Shadish et al. (2002) skal kontroleleverne være "plausible similar to" eleverne i indsats gruppen. I og med at vi tager højde for centrale baggrundskarakteristika vurderer vi, at eleverne på de obligatoriske hold og på opkvalificerende type1 holdene er relativt ens.

Data og metode

Vi tager udgangspunkt i alle elever på skoler og indgange, se tabel 4, der var påbegyndt et grundforløb i august måned i et givet år (2007, 2008, 2009, 2010) og fulgte dem 43 uger frem. I denne analyse har vi betinget på, at eleven ikke tidligere havde påbegyndt uddannelsesforløb indenfor erhvervsuddannelserne, dvs. undersøgelsespopulationen bestod af førstegangsstartere., og ved hjælp af indsamlede holdkoder fra skolerne.

Ved at kæde oplysninger om holdkoder fra skoleundersøgelsen præsenteret i del 1 sammen med data fra UNI-C har vi identificeret og tilknyttet elever til de forskellige grundforløbstyper og modeller Skolernes administrative database, EASY-A, som UNI-C omsætter til EASY-S data, giver mulighed for at registrere elevernes samlede uddannelsesforløb, hvilken uddannelse de har valgt og hvilke forskellige hold, de er tilknyttet, eller har været tilknyttet. Det er dog ikke alle erhvervsskoler, som har en registreringspraksis, der gør det muligt at kæde hold med bestemte pakker, modeller og elever sammen. På den baggrund er vi endt op med at anvende 5 skoler ud af 20 mulige omfattende EUC Vest, Roskilde, Herningsholm, Syddansk og CELF (i alt 1805 elever på forskellige indgange).

De fleste af clear cut eleverne er kun tilknyttet en grundforløbstype i undersøgelses-perioden, men nogle elever har været tilknyttet flere typer i samme periode. For at sikre, at en elev ikke optræder i både kontrol- og indsatsgruppen (eller på andre grundforløbsmodeller i samme periode) har vi antaget, at elever der har været tilknyttet et opkvalificeringshold i mere end 60 % af tiden kan betragtes som opkvalificeringselever. Det indebærer bl.a. at elever som har brugt lige lang tid på obligatorisk og opkvalificeringsholdet blev kategoriseret som opkvalificeringselever. Eksempel: Hvis elev gik 15 uger i alt, og i 14 af de 15 uger såvel foregik på et obligatorisk forløb som på et opkvalificeringsforløb så placeres eleven i opkvalificeringsforløbet, omvendt hvis en elev i alt gik 29 uger, hvoraf de 15 uger var obligatorisk forløb og de 14 uger var opkvalificeringsforløb, så placeres eleven på obligatorisk forløb. Kravet om de 60 % er en antagelse, som kan diskuteres. Som det også fremgår af første del, viste det sig under dataindsamlingen, at der var en del støj i de aktivitetskoder vi skulle bruge til at identificere grundforløbstyperne. For eksempel ser ud til, at vi i nogle tilfælde har fået 'for mange' obligatoriske hold, dvs. hold som ikke er stamhold, men hold der går på tværs af grundforløbstyperne (og som derfor klart ikke kan bruges som identifikation).

Ved hjælp af oplysninger fra Danmarks Statistiks komprimerede elevregister (KOET2011) har vi defineret en variabel, der angiver elevens status på holdets efter 43 uger. Hvis en elev har gennemført grundforløbet indenfor hold-perioden er status = "Gennemført", hvis en elev har afbrudt uddannelse i hold-perioden er status = "Afbrudt" og hvis en elev hverken har gennemført eller afbrudt uddannelsen i hold-perioden er status = "I gang". Disse er så

Ud fra status-variabel har defineret to udfalds-variable til de statistiske analyser. Den ene, *completed*, har værdien 1 hvis eleven har gennemført uddannelsen indenfor undersøgelsesperioden, dvs. hvis status = "Gennemført", og nul ellers. Den anden, *dropout*, har værdien 1 hvis eleven er faldet fra i undersøgelsesperioden, dvs. hvis status = "Afbrudt", og nul ellers.

Som forklarende variable i de statistiske analyser af opkvalificeringstype 1 og opkvalificeringstype 2 under clear-cut modellen har vi brugt registervariable fra Danmarks Statistik. En oversigt over definitioner og datakilde for samtlige anvendte baggrundsvariable ses i appendiks A. I forbindelse med analyserne af de enkelte hold var det ikke muligt at inkludere alle variablene på grund af den begrænsede stikprøvestørrelse.

Tabel 4. Oversigt over skolerne udvalgt til analyse af clear cut modellen

		Startår	Opkvalificerings- type 1	Opkvalificerings- type 2
EUC Vest	Bygge og anlæg	2008		13
		2009		10
		2010	11	
	Fra jord til bord	2008	11	20
		2009	20	11
	Håndværk og teknik	2008	10	
	Teknologi og kommunikation	2008		23
		2009		14
Roskilde TS	Mekanik, transport og logistik	2008	45	
		2009	44	12
		2010	13	11
Herningsholm	Bygge og anlæg	2008	20	
		2009	17	
		2010	11	
	Teknologi og kommunikation	2008	18	
		2009	23	
		2010	15	
Syddansk Erhvervsskole	Mekanik, transport og logistik	2008	16	
CELF	Bygge og anlæg	2008	34	
	Fra jord til bord	2008		13
		2008		14
	Mekanik, transport og logistik	2009		55
		2010		47
	Teknologi og kommunikation	2009		18
2010			13	
I alt			308	274

For at undersøge om der er forskel på elever på Clear Cut modellen på obligatoriske hold og opkvalificeringstype-hold har vi estimeret en lineær sandsynlighedsmodel for gennemførelse og afbrud indenfor for den givne tidsperiode. Modellen estimerer sandsynligheden for, at en elev falder fra eller ej som funktion af de inkluderede kontrolvariable (se Appendiks A). Parameterestimererne viser hvorvidt en enheds stigning i kontrol-variablen øger eller reducerer sandsynligheden for at falde fra, når de andre kontrol-variable holdes konstante. Der er forbehold og de vigtigste nævnes her. For det første kan der være endogenitets- og selektionsproblemer. For det andet kan treatmentvariablen kan være korreleret med uobserverede elev-karakteristika, som f.eks. motivation for at gennemføre uddannelse.

Resultater

Vi ser nu nærmere på frafald og gennemførelse for de udvalgte skoler i undersøgelsen af clear-cut-modellen, og specielt om der er forskel på eleverne på det obligatoriske forløb og eleverne på opkvalificeringsforløbene type 1 og type 2 med hensyn til frafald og gennemførelse på første forløb på grundforløbet. Vi har også undersøgt alle forløb, og det gav i store træk samme hovedresultat.

Vi har opstillet en samlet model for alle sammenligninger, henholdsvis 15 for opkvalificerende type 1 og 14 for opkvalificerende type 2 og udregnet et estimat for den gennemsnitlige forskel i frafaldsraten og gennemsnitraten for i hver indsatsgruppe sammenlignet med kontrolgruppen. Appendiks C og D viser estimater fra en model hvor vi delt op stikprøven op efter opkvalificeringstype. Estimaterne er udregnet i en lineær sandsynlighedsmodel, hvor vi har kontrolleret for skole-, indgang- og årseffekter samt baggrundsvARIABLENE, som angivet i appendiks A.

Det fremgår af tabel 5, at eleverne i indsatsgrupperne har et større frafald end i kontrolgruppen. Analysen viser, at elever som gik på type1 og type 2 sammenlignet med normal sporet (obligatorisk) har en øget frafaldsrisiko. Især i forhold til type 1 elever som på observerede karakteristika ligner normalsporelever (kontrolgruppen) er der grund til at notere sig, at en opdeling af elever ikke fremmer fastholdelsen, hvilket dog ikke betyder, at der kan være elever på udvalgte indgange og skoler som haft gavn af at gå på type 1 forløb eller sågar type 2. Frafallsraten i opkvalificeringstype 1 holdene er i gennemsnit 0.10 højere end frafallsraten på de obligatoriske hold, og for opkvalificeringstype 2 hold er den i gennemsnit er 0.17 højere. Ser vi på forskelle i gennemførselsrater klarer eleverne på opkvalificeringsholdene sig også relativt dårligt. Opkvalificeringstype 1 hold har en gennemførselsrate, der 0.15 lavere end leverne i kontrolgruppen, og for opkvalificeringstype 2 er gennemførselsraten 0.30 lavere.

Man skal være påpasselig med at fortolke estimaterne fra den lineære sandsynlighedsmodel, da de i sig selv ikke siger noget om hvilken kausal effekt opkvalificeringstyperne har på frafald og gennemførsel. Der kan være alternative forklaringer. Det kan tænkes, at de to opkvalificeringshold 1 og 2 faktisk reducerer frafallssandsynligheden i forhold til den sandsynlighed, der ville have været gældende, hvis eleven i stedet for havde fulgt et obligatorisk hold.

Ideelt set vil vi gerne vurdere om den ekstra fastholdelse er med til at formindske frafaldet. Problemet er, at vi ikke kan se frafaldet for eleverne på indsatsgrupperne, hvis de ikke havde gået på opkvalificeringshold (med andre ord observerer vi ikke det kontrafaktiske udfald), så der kan være flere alternative forklaringer. At eleverne på opkvalificeringstype 1 og 2 klarer sig relativt dårligt i forhold kontrol-gruppen, kan skyldes forskelle før de startede på uddannelsen, men vi ved i hvert fald, at elever på obligatorisk type og opkvalificerende type 1 er ret lig hinanden, målt på baggrundskarakteristika, som vi kontrollerer for. Vi kan dog ikke helt udelukke, at indsatsen har haft en vis positiv effekt. Elever på opkvalificerende type 1 og især type 2 under clear-cut modellen har øget risiko for frafald i sammenligning med elever på de obligatoriske typer, hvilket kan skyldes, at elever på type 1 havde lidt ringere faglige forudsætninger, medens elever på type 2 foruden har en lidt svagere baggrund såvel fagligt som socialt. Ovenstående analyse siger noget om interaktionen mellem eleverne og de forskellige grundforløbstyper, mere end selve effekten af grundforløbstyperne.

Tabel 5: Lineær sandsynlighedsmodel for afbrud og gennemførelse. Elever på Clear Cut modellen

Variabel	Afbrud		Gennemførelse	
	Parameterestimat (std. fejl)			
<i>Grundforløbspakke (ref: obligatorisk)</i>				
Opkvalificering 1	0.10 ***	(0.03)	-0.15 ***	(0.03)
Opkvalificering 2	0.17 ***	(0.03)	-0.30 ***	(0.03)
<i>Skole (ref: EUC Vest)</i>				
CELF	-0.13 ***	(0.03)	-0.14 ***	(0.03)
Roskilde TS	-0.12 **	(0.05)	0.26 ***	(0.05)
Herningsholm	-0.08 **	(0.04)	0.00	(0.03)
Syddansk Erhvervsskole	-0.10 **	(0.05)	0.10 **	(0.05)
<i>Indgang (ref: Bygge og anlæg)</i>				
Fra jord til borg	-0.16 ***	(0.04)	0.10 ***	(0.04)
Håndværk og teknik	0.07	(0.05)	-0.02	(0.05)
Mekanik, transport og logistik	-0.08 **	(0.04)	0.04	(0.04)
Teknologi og kommunikation	0.05	(0.03)	-0.07 **	(0.03)
<i>Startår (ref: 2008)</i>				
2009	0.02	(0.02)	-0.06 **	(0.02)
2010	0.13 ***	(0.03)	0.01	(0.03)
Kvinde	0.06	(0.04)	-0.03	(0.04)
Alder	0.04 **	(0.02)	0.00	(0.02)
Alder ²	0.00 **	(0.00)	0.00	(0.00)
<i>Ethnicitet (ref: Dansk)</i>				
Indvandrere	0.04	(0.09)	-0.03	(0.09)
Efterkommere	-0.06	(0.07)	0.05	(0.07)
Kernefamilie	-0.01	(0.03)	0.05 *	(0.03)
Farens alder	0.00	(0.00)	0.00	(0.00)
Morens alder	0.00	(0.00)	0.00	(0.00)
<i>Farens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	-0.07 ***	(0.02)	0.08 ***	(0.02)
Gymnasial	0.02	(0.07)	-0.06	(0.07)
Kort/mellemlang videregående	0.02	(0.05)	-0.04	(0.04)
Lang videregående	-0.05	(0.09)	0.10	(0.09)
<i>Morens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	-0.01	(0.03)	0.04 *	(0.02)
Gymnasial	-0.08	(0.05)	0.03	(0.05)
Kort/mellemlang videregående	-0.02	(0.04)	0.10 **	(0.04)
Lang videregående	0.14	(0.12)	-0.21 *	(0.12)
Familieindkomst (gennemsnit)	-0.01	(0.01)	0.00	(0.01)
<i>Farens beskæftigelse (ref: lønmodtager mellem/højt)</i>				
Lønmodtager, lavt niveau	0.02	(0.04)	-0.05	(0.04)
Lønmodtager, andet	0.02	(0.04)	0.00	(0.04)
Self-beskliftiget	0.09 *	(0.05)	-0.06	(0.05)
Ledig	-0.04	(0.09)	0.04	(0.09)
Udenfor arbejdsmarkedet	0.08	(0.05)	-0.06	(0.05)
<i>Morens beskæftigelse (ref: lønmodtager mellem/højt)</i>				
Lønmodtager, lavt niveau	0.04	(0.03)	0.00	(0.03)
Lønmodtager, andet	0.02	(0.04)	0.00	(0.04)
Self-beskliftiget	-0.11 *	(0.06)	0.04	(0.06)
Ledig	0.13 *	(0.08)	-0.03	(0.08)
Udenfor arbejdsmarkedet	0.07	(0.04)	-0.02	(0.04)
Eksamenskarakter i skriftlig matematik, 9.kl.	-0.03 ***	(0.00)	0.04 ***	(0.00)
Eksamenskarakter i skriftlig dansk, 9.kl.	0.00	(0.00)	0.01 *	(0.00)
Konstantled	0.22	(0.24)	-0.04	(0.24)
	R ² = 0.13		R ² = 0.24	
	N=1805			

Anm.: Statistisk signifikante estimater er indikeret: *: 10%, **: 5% and ***: 1%. Kategorier for manglende værdier er inkluderet. Modellen er estimeret ved OLS.

Vi finder således, at forlængede grundforløb i højere grad er med til at fastholde elever, når vi sammenligner WAS- og CC-modellerne, dels lige efter 20 uger, og især efter 30-40 uger. Endelig finder vi, at elever på normalsporet i gennemsnit har en lavere frafaldssandsynlighed end elever på opkvalificeringstype 1 og opkvalificeringstype 2 under CC-modellen.

Forskellen i frafaldsrater mellem grundforløbstyper kan variere for forskellige elevtyper. For eksempel kan forskellen være relativ stor for elever med et lavt karaktergennemsnit (GPA). Tabel 6 viser gennemsnittet af de individuelle beregnede frafaldssandsynligheder baseret på den lineære sandsynlighedsmodel i tabel 5 over grundforløbstype og karakterplacering. Gruppen af elever med lavere gennemsnit er defineret som elever med grundskolekarakterer i matematik og dansk under 4; gruppen med højt gennemsnit er defineret som elever med karakterer over 7.

Tabel 6: Beregnet frafaldssandsynlighed over grundforløbstype og karaktergennemsnit

	Obligatorisk			Opkvalificeringstype 1			Opkvalificeringstype 2		
	<i>Estimat (laveste CL, 5%) (øverste CL, 95%)</i>								
Alle elever	0,25	(0,23)	(0,28)	0,36	(0,31)	(0,41)	0,44	(0,39)	(0,50)
Lav GPA	0,34	(0,30)	(0,37)	0,41	(0,36)	(0,47)	0,48	(0,42)	(0,54)
Høj GPA	0,19	(0,16)	(0,23)	0,28	(0,22)	(0,34)	0,33	(0,26)	(0,41)

Note: Sandsynligheder er beregnet på baggrund af den lineære sandsynlighedsmodel.

Konfidensintervallerne er estimeret ved bootstrapping med 1000 replikationer.

Vi kan således se, at elever på obligatorisk hold med en lav karaktergennemsnit fra grundskolen har en beregnet frafaldsrisiko på 34 pct. i gennemsnit sammenlignet med 19 pct. for elever med høj gennemsnit. Elevgruppen med et højt karaktergennemsnit under opkvalificeringstype 1 har sammenlignet med elevgruppen med et lavt karaktergennemsnit på normalsporet en lavere frafaldsrisiko på 28 pct. mod 34 pct. Det betyder, at opkvalificeringen måske særligt kan hjælpe elever med lidt bedre forudsætninger.

Diskussion

De to analyser viser, at der er forskelle mellem skoler, indgange og år. Analyserne er blevet testet på forskellig vis og kørt med forskellige modelantagelser. Når sammenligninger foretages, som vi har gjort det, melder der altid en række spørgsmål i retning af om det er plausibelt at gøre som vi gjort, om sammenligningerne ikke står over for en række relevante validitetstrusler. For at optimere analyserne gjorde vi meget ud af kontroller i såvel designet som i den statistiske model.. En relevant indvending er at forskellige indgange har forskellig sværhedsgrad, men her er pointen, at vi blot sammenligner indenfor indgangen. En anden fejlkilde kunne være at der er forskellige typer af elever på forskellige skoler, hvilket betyder at vi kun sammenligner elever inden for den enkelte skole. Vi forsøgte ligeledes i analyserne af de enkelte hold at sammenligne elever inden for samme begyndelsesår, idet forholdene på skolerne kan ændres over tid. Det er dog ikke sikkert, at disse forskelle er så store givet at alle årene vedrører år efter reformen. I forhold til det Shadish et al. (2002, side 20) kalder for trusler mod construct validity, kan det diskuteres om vores typer og WAS- og CC- modeller dækker over det samme på alle skoler og indgange (eller om de dækker over det, vi siger de dækker over, vores teoretiske konstruktioner). Vi ved jo fra første del, at det

faktisk er tilfældet for en del skoler, men at der også er skoler og indgange, som har mere blandede typer og modeller. Dem har vi så vidt muligt ikke medtaget i analyserne. Bias på elevniveau er forsøgt minimeret ved hjælp af baggrundsvARIABLE i de statistiske modeller. Analyserne kan være truet af, at estimaterne bliver drevet af nogle få ekstreme observationer og på den baggrund er en række modelkontroller foretaget.

Alt i alt har vi i de sammenlignende kvantitative undersøgelser bidraget med nye analyser på langt mere detaljerede niveauer, end man normalt ser i en dansk kontekst og for den sags skyld internationalt, og vi har, med brug af modeller med sen opdeling (late tracking), bidraget med nye statistiske analyser, som viser, at elevfrafaldet er mindre på nogle skoler. Imidlertid kunne vi ønske os, at skolerne fremadrettet får tilsagt en mere ensartet registreringspraksis, som vil muliggøre en kausal analyse af alle erhvervsskoler og deres tiltag og interventioner. Det kræver dog, at Ministeriet for Børn og Undervisning sammen med UNI-C og forskere går mere aktivt og forpligtende ind i et samarbejde med skolerne omkring systematisk registrering af data.

Litteratur:

BEK nr. 1518 af 13.12.2007: Bekendtgørelse om erhvervsuddannelser, UVM, offentliggjort på hjemmesiden "Retsinformation":

<https://www.retsinformation.dk/Forms/R0710.aspx?id=114118>, set 4/1 2012.

Cameron, A. C., & Trivedi, P. K. (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.

Cleves, M., Gutierrez, R., Gould, W., & Marchenko, Y. (2010): *An Introduction to Survival Analysis Using Stata*. Stata Press.

Cox, D. R. (1972): Regression Models and Life Tables. *Journal of the Royal Statistical Society, Series B (Methodological)* 34, 187-220.

DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999): An event history model of student departure. *Economics of Education Review* 18, 375-390.

DesJardins, S. L., McCall, B. P., Ahlburg, D. A., & Moye, M. J. (2002): Adding a Timing Light to the "Tool Box". *Research in Higher Education* 43, 83-114.

Dignam, J. J., Zhang, Q., & Kocherginsky, M. (2012): The Use and Interpretation of Competing Risks Regression Models. *Clinical Cancer Research* 18, 2301-2308.

EVA (2009): *Grundforløbspakker på erhvervsuddannelserne*. Rapport udgivet af Det danske Evalueringsinstitut (EVA).

Feng, L., & Sass, T. (2012): *Competing Risk Analysis of Dropout and Educational Attainment for Students with Disabilities*. Andrew Young School of Policy Studies: Working Paper 2012-2-3.

Fine, J. P., & Gray, R. J. (1999): A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association* 94, 496-509.

Gamoran, A. (1992): The Variable Effects of High School Tracking. *American Sociological Review* 57, 812-828. Reprinted in *Advances in Educational Research* 1, 104-120

Gamoran, A. (2010): Tracking and Inequality: New Directions for Research and Practice. Pp. 213-228, i M. Apple; S. J. Ball; L. A. Gandin (Eds.): *The Routledge International Handbook of the Sociology of Education*. London: Routledge.

Kalbfleisch, J. D., & Prentice, R. L. (2002): *The Statistical Analysis of Failure Time Data*. John Wiley & Sons.

Langbein, L. & Felbinger, C. L. (2006): *Public Program Evaluation. A Statistical Guide*. M. E. Sharpe, Armonk, New York

- Larsen, B. Ø. & Jensen, T. Pilegaard (2010): *Fastholdelse af elever på de danske erhvervsskoler*. Rapport. AKF
- Munk, M. D. & Park, D.-Y. (2013): Explaining the Rise in Danish Vocational Education System Dropouts: The Effect of the Gap Year. Centre for Mobility Research, Aalborg, University.
- Munk, M. D. & Baklanov, N. (2013): Vocational school dropout and the effect of late versus early tracking and prolongation of basic courses, Centre for Mobility Research, Aalborg University.
- Nielsen, K. (2011): Hvad siger international forskning om frafald? Et forsøg på et overblik, i Jørgensen, C. Helms (red): *Frafald i erhvervsuddannelserne*. Roskilde Universitetsforlag, Roskilde.
- Olsson, M. (2009): Employment protection and sickness absence. *Labour Economics* 16, 208-214.
- Plank, S. B., DeLuca, S., & Estacion, A. (2008): High School Dropout and the Role of Career and Technical Education: A Survival Analysis of Surviving High School. *Sociology of Education* 81, 345-370.
- Royston, P. (2011): *Estimating a smooth baseline hazard function for the Cox model*. Research report No. 314: University College London.
- Rumberger, R. & Thomas, S. L. (2000): The distribution of Dropout and Turnover Rates Among Urban and Suburban High Schools. *Sociology of Education* 73, 39-67.
- Shadish, W. R.; Cook, T. D.; Campbell, D. T. (2002): *Experimental and Quasi-experimental Designs for Generalized Causal Interference*. Houghton Mifflin Company, Boston.
- StataCorp. (2009): *Stata survival analysis epidemiological tables. Reference manual. Release 11*. Texas: Stata Press.
- DesJardins, S. L., McCall, B. P., Ahlburg, D. A., & Moye, M. J. (2002): Adding a Timing Light to the "Tool Box". *Research in Higher Education* 43, 83-114.
- Tanggaard, L., Jørgensen, C. Helms, Koudahl, P., & Nielsen, K. (2012): Preventing dropout in vocational education – students' experiences of streaming. Work in Progress.
- Willet, J. B., & Singer, J. D. (1991): From Whether to When: New Methods for Studying Student Dropout and Teacher Attrition. *Review of Educational Research* 61, 407-450.
- Wolbers, M., Koller, MT., Witteman, JC., & Steyerberg, EW. (2009): Prognostic models with competing risks: methods and application to coronary risk prediction. *Epidemiology* 20, 555–561.

Appendiks

Appendiks A. Oversigt over forklarende variable anvendt i regressionsmodeller

Variabel	Definition	DST register	WAS	CC
Kvinde	1= Kvinde og 0=Mand	Befolkningsregistret (BEF)	x	x
Alder	Alder i år	Befolkningsregistret (BEF)	x	x
Kernefamilie	1=Kernefamilie ved 15 årsalder og 0 ellers	Befolkningsregistret (BEF)	x	x
Karaktergennemsnit	Gennemsnit af skriftlige eksamens karakter i matematik og dansk i 9. klasse	Folkeskolekarakterer (UDFK)	x	x
Matematik og dansk	Gennemsnitskarakter fra skriftlig eksamen i 9. klasse i hhv. matematik og dansk	Folkeskolekarakterer (UDFK)		x
Ikke-vestlig	1=Indvandrere/Efterkommere og 0 ellers	Indvandrere og efterkommere (IEPE)	-	x
Etnicitet	0=Danish; 1=Indvandrere; 2=Efterkommere	Indvandrere og efterkommere (IEPE)	x	x
Familieindkomst	Sum af familieindkomst (100.000 DKK). Indkomster lavere end 1% fraktilen er omkodet til 1% fraktilen. Tilsvarende omkodning til indkomster over 99% fraktilen	Indkomstdata (INDH)	x	x
Farens og morens højeste fuldførte uddannelse	0=Grundskole 1=Erhvervsfaglig 2=Gymnasial 3=Kort/mellemlang videregående 4=Lang videregående	Uddannelsesregistret (UDDA)	x	x
Farens/ morens social klassifikation	1= Selfbeskæftigede (medarbejdende ægtefælle) 2= Lønmodtager, grundniveau (<i>Reference</i>) 3= Lønmodtager, mellem og højt niveau 4= Lønmodtager, andet 5= Ledig 6= Udenfor arbejdsmarkedet	Indkomstdata (INDH)	x	x
Sommer	Uddannelsestart 0=Mellem januar og juli 1= Mellem august og december	Komprimeret Elevregister (KOET)	x	-

Anm: Kategorier for manglende værdier er også inkluderet i analyserne.

Appendiks B.1: Deskriptiv statistik for alle baggrundsvariables for hhv. WAS, CC og hele Bygge og Anlæg (BA) populationen, 2008-2010

	WAS		CC		Samlet BA tilgang			
	Middel- værdi	Standard- afvigelse	Middel- værdi	Standard- afvigelse	Middel- værdi	Standard- afvigelse	Min.	Maks.
WAS	1,00	0,00	-	-	-	-	-	-
<i>Startår (ref: 2008)</i>	0,35	0,48	0,33	0,47	0,34	0,47	0	1
2009	0,35	0,48	0,33	0,47	0,33	0,47	0	1
2010	0,29	0,46	0,34	0,47	0,33	0,47	0	1
Sommer	0,35	0,48	0,28	0,45	0,35	0,48	0	1
Kvinde	0,17	0,38	0,13	0,34	0,15	0,35	0	1
Familietype (15-årsalderen; ref.: <i>Andet</i>)	0,32	0,47	0,30	0,46	0,32	0,47	0	1
Kernefamilie	0,65	0,48	0,68	0,47	0,65	0,48	0	1
Ukendt	0,03	0,17	0,02	0,14	0,03	0,16	0	1
<i>Ethnicitet (ref: Dansk)</i>	0,87	0,33	0,94	0,24	0,89	0,32	0	1
Indvandrere	0,07	0,26	0,04	0,20	0,06	0,24	0	1
Efterkommere	0,05	0,22	0,02	0,13	0,05	0,22	0	1
Alder	15,81	4,81	15,32	4,44	15,65	4,76	11	60
<i>Farens uddannelse (ref: grundskole)</i>	0,31	0,46	0,34	0,47	0,31	0,46	0	1
Erhvervsfaglig	0,41	0,49	0,44	0,50	0,42	0,49	0	1
Gymnasial	0,02	0,13	0,01	0,12	0,02	0,15	0	1
Kort/mellemlang videregående	0,12	0,32	0,09	0,28	0,10	0,30	0	1
Lang videregående	0,01	0,12	0,02	0,13	0,03	0,16	0	1
Ukendt	0,13	0,34	0,10	0,30	0,12	0,33	0	1
<i>Morens uddannelse (ref: grundskole)</i>	0,38	0,49	0,37	0,48	0,35	0,48	0	1
Erhvervsfaglig	0,38	0,49	0,40	0,49	0,39	0,49	0	1
Gymnasial	0,03	0,17	0,03	0,16	0,04	0,18	0	1
Kort/mellemlang videregående	0,13	0,33	0,14	0,35	0,15	0,36	0	1
Lang videregående	0,02	0,12	0,01	0,10	0,02	0,13	0	1
Ukendt	0,06	0,24	0,05	0,21	0,06	0,24	0	1
<i>Farens beskæftigelse (ref: lønmodtager lavt niveau)</i>	0,07	0,26	0,08	0,28	0,08	0,27	0	1
Selvstændig	0,13	0,34	0,11	0,31	0,13	0,34	0	1
Lønmodtager, mellem/højt niveau	0,37	0,48	0,43	0,50	0,38	0,48	0	1
Lønmodtager, andet	0,13	0,33	0,14	0,34	0,13	0,34	0	1
Ledig	0,04	0,19	0,03	0,18	0,04	0,19	0	1
Udenfor arbejdsmarkedet	0,17	0,38	0,14	0,34	0,16	0,37	0	1
Ukendt	0,09	0,29	0,07	0,26	0,08	0,28	0	1
<i>Morens beskæftigelse (ref: lønmodtager lavt niveau)</i>	0,03	0,17	0,03	0,18	0,03	0,18	0	1
Selvstændig	0,15	0,36	0,17	0,37	0,18	0,39	0	1
Lønmodtager, mellem/højt niveau	0,40	0,49	0,43	0,49	0,40	0,49	0	1
Lønmodtager, andet	0,09	0,29	0,09	0,29	0,09	0,29	0	1
Ledig	0,06	0,24	0,05	0,23	0,06	0,24	0	1
Udenfor arbejdsmarkedet	0,23	0,42	0,21	0,40	0,20	0,40	0	1
Ukendt	0,03	0,17	0,02	0,15	0,03	0,17	0	1
Familieindkomst (gennemsnit)	5,39	2,23	5,53	2,13	5,43	2,23	0,00	12,97
Ukendt familieindkomst	0,01	0,11	0,01	0,09	0,01	0,10	0	1
<i>Skole (ref.: 15. KTS (WAS))</i>	0,18	0,39	-	-	-	-	-	-
Randers TS (WAS)	0,16	0,37	-	-	-	-	-	-
Syddansk Erhvervsskole (WAS)	0,61	0,49	-	-	-	-	-	-
Silkeborg TS (WAS)	0,05	0,21	-	-	-	-	-	-
EUC Vest (CC)	-	-	0,22	0,41	-	-	-	-
EUC Nordvest (CC)	-	-	0,06	0,24	-	-	-	-
Hermingsholm Erhvervsskole (CC)	-	-	0,12	0,33	-	-	-	-
CELF (CC)	-	-	0,15	0,36	-	-	-	-
Uddannelsescenter Holstebro (CC)	-	-	0,08	0,28	-	-	-	-
Tech College Aalborg (CC)"	-	-	0,36	0,48	-	-	-	-
Eksamenskarakter, 9.kl.	2,86	2,58	3,18	2,56	3,03	2,57	-1,13	12,00
Manglende eksamenskarakter	0,31	0,46	0,25	0,43	0,28	0,45	0	1
Obs.	3.286		3.560		20.405			

Appendiks B.2: Deskriptiv statistik for alle baggrundsvariables for hhv. WAS, CC og hele Strøm, Styling og IT (SSI) populationen, 2008-2010

	WAS		CC		Samlet SSI tilgang			
	Middel-værdi	Standard-afvigelse	Middel-værdi	Standard-afvigelse	Middel-værdi	Standard-afvigelse	Min.	Maks.
WAS	1,00	0,00	0,00	0,00	0,10	0,30	0	1
<i>Startår (ref: 2008)</i>								
2009	0,38	0,49	0,38	0,48	0,38	0,48	0	1
2010	0,34	0,47	0,35	0,48	0,35	0,48	0	1
Sommer	0,20	0,40	0,24	0,43	0,24	0,43	0	1
Kvinde	0,02	0,15	0,05	0,22	0,05	0,21	0	1
<i>Familietype (15-årsalderen; ref.: Ande</i>								
Kernefamilie	0,67	0,47	0,72	0,45	0,71	0,45	0	1
Ukendt	0,04	0,19	0,04	0,21	0,04	0,20	0	1
<i>Ethnicitet (ref: Dansk)</i>								
Indvandrere	0,05	0,22	0,06	0,24	0,06	0,24	0	1
Efterkommere	0,03	0,18	0,02	0,15	0,02	0,15	0	1
Alder	15,78	4,96	16,60	5,71	16,52	5,65	11	52
<i>Farens uddannelse (ref: grundskole)</i>								
Erhvervsfaglig	0,44	0,50	0,43	0,50	0,43	0,50	0	1
Gymnasial	0,03	0,16	0,03	0,17	0,03	0,17	0	1
Kort/mellemlang videregående	0,12	0,33	0,13	0,34	0,13	0,34	0	1
Lang videregående	0,01	0,12	0,02	0,14	0,02	0,14	0	1
Ukendt	0,09	0,29	0,11	0,31	0,11	0,31	0	1
<i>Morens uddannelse (ref: grundskole)</i>								
Erhvervsfaglig	0,31	0,46	0,32	0,47	0,32	0,47	0	1
Gymnasial	0,45	0,50	0,40	0,49	0,41	0,49	0	1
Kort/mellemlang videregående	0,05	0,22	0,04	0,21	0,04	0,21	0	1
Lang videregående	0,13	0,34	0,17	0,37	0,16	0,37	0	1
Ukendt	0,01	0,09	0,01	0,12	0,01	0,12	0	1
Ukendt	0,05	0,22	0,05	0,22	0,05	0,22	0	1
<i>Farens beskæftigelse (ref:</i>								
<i>lønmodtager lavt niveau)</i>								
Selvstændig	0,06	0,24	0,08	0,27	0,08	0,27	0	1
Lønmodtager, mellem/højt niveau	0,18	0,38	0,15	0,35	0,15	0,36	0	1
Lønmodtager, andet	0,41	0,49	0,39	0,49	0,39	0,49	0	1
Ledig	0,13	0,34	0,14	0,35	0,14	0,35	0	1
Udenfor arbejdsmarkedet	0,03	0,17	0,03	0,18	0,03	0,18	0	1
Ukendt	0,15	0,36	0,13	0,33	0,13	0,34	0	1
Ukendt	0,04	0,20	0,08	0,27	0,08	0,27	0	1
<i>Morens beskæftigelse (ref:</i>								
<i>lønmodtager lavt niveau)</i>								
Selvstændig	0,03	0,16	0,03	0,17	0,03	0,17	0	1
Lønmodtager, mellem/højt niveau	0,19	0,40	0,19	0,39	0,19	0,39	0	1
Lønmodtager, andet	0,44	0,50	0,41	0,49	0,41	0,49	0	1
Ledig	0,09	0,28	0,10	0,29	0,09	0,29	0	1
Udenfor arbejdsmarkedet	0,05	0,22	0,06	0,23	0,06	0,23	0	1
Ukendt	0,18	0,38	0,19	0,39	0,19	0,39	0	1
Ukendt	0,02	0,15	0,03	0,18	0,03	0,18	0	1
Familieindkomst (gennemsnit)	5,82	2,14	5,54	2,28	5,56	2,26	0,00	12,97
Ukendt familieindkomst	0,01	0,07	0,01	0,12	0,01	0,11	0	1
<i>Skole</i>								
Randers TS (CC)	0,00	0,00	0,07	0,25	0,06	0,24	0,00	1,00
Uddannelsescenter Holstebro (CC)	0,00	0,00	0,07	0,25	0,06	0,24	0,00	1,00
EUC Sjælland (WAS)	0,89	0,32	0,00	0,00	0,09	0,28	0,00	1,00
Syddansk Erhvervsskole (CC)	0,00	0,00	0,40	0,49	0,36	0,48	0,00	1,00
EUC Vest (CC)	0,00	0,00	0,08	0,27	0,07	0,25	0,00	1,00
Herningsholm Erhvervsskole (CC)	0,00	0,00	0,08	0,27	0,07	0,26	0,00	1,00
Djursland (WAS)	0,11	0,32	0,00	0,00	0,01	0,11	0,00	1,00
Skive (CC)	0,00	0,00	0,03	0,18	0,03	0,17	0,00	1,00
Tech College Aalborg (CC)	0,00	0,00	0,27	0,44	0,24	0,43	0,00	1,00
Eksamenskarakter, 9.kl.	3,45	2,75	3,74	2,93	3,71	2,91	-0,17	12,00
Manglende eksamenskarakter	0,27	0,44	0,28	0,45	0,28	0,45	0	1
Obs.	367		3349		3716			

Appendiks C. Clear Cut Opkvalificering 1 vs. Obligatorisk, LPM

Variabel	Afbud		Gennemførelse	
	Parameterestimat (std. fejl)			
<i>Grundforløbspakke (ref: obligatorisk)</i>				
Opkvalificering 1	0.13 **	(0.03)	-0.15 ***	(0.03)
<i>Skole (ref: EUC Vest)</i>				
CELF	-0.09 **	(0.03)	-0.15 ***	(0.04)
Roskilde TS	-0.24 **	(0.06)	0.26 ***	(0.06)
Herningsholm	-0.11 **	(0.04)	0.00	(0.04)
Syddansk Erhvervsskole	-0.16 **	(0.05)	0.10 *	(0.05)
<i>Indgang (ref: Bygge og anlæg)</i>				
Fra jord til borg	-0.22 **	(0.04)	0.13 ***	(0.04)
Håndværk og teknik	0.07	(0.05)	0.00	(0.06)
Mekanik, transport og logistik	-0.01	(0.04)	0.06	(0.04)
Teknologi og kommunikation	0.09 **	(0.03)	-0.09 **	(0.04)
<i>Startår (ref: 2008)</i>				
2009	0.01	(0.03)	-0.06 **	(0.03)
2010	0.17 **	(0.03)	0.00	(0.03)
Kvinde	0.07	(0.04)	-0.03	(0.05)
Alder	0.04 **	(0.02)	0.01	(0.02)
Alder ²	0.00 **	(0.00)	0.00	(0.00)
<i>Ethnicitet (ref: Dansk)</i>				
Indvandrere	0.03	(0.10)	-0.04	(0.10)
Efterkommere	-0.12	(0.08)	0.04	(0.09)
Kernefamilie	-0.01	(0.03)	0.05 *	(0.03)
Farens alder	0.00	(0.00)	0.00	(0.00)
Morens alder	0.00	(0.00)	0.00	(0.00)
<i>Farens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	-0.08 **	(0.03)	0.08 ***	(0.03)
Gymnasial	0.05	(0.08)	-0.05	(0.08)
Kort/mellemlang videregående	0.01	(0.05)	-0.05	(0.05)
Lang videregående	-0.05	(0.09)	0.09	(0.09)
<i>Morens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	-0.02	(0.03)	0.05 *	(0.03)
Gymnasial	-0.06	(0.06)	0.04	(0.06)
Kort/mellemlang videregående	0.02	(0.04)	0.11 **	(0.05)
Lang videregående	0.13	(0.13)	-0.19	(0.14)
Familieindkomst (gennemsnit)	0.00	(0.01)	0.00	(0.01)
<i>Farens beskæftigelse (ref: lønmodtager mellem/højt)</i>				
Lønmodtager, lavt niveau	0.02	(0.04)	0.00	(0.04)
Lønmodtager, andet	0.02	(0.04)	-0.05	(0.04)
Self-beskftiget	0.06	(0.05)	-0.05	(0.05)
Ledig	-0.01	(0.11)	0.05	(0.11)
Udenfor arbejdsmarkedet	0.09 *	(0.05)	-0.07	(0.05)
<i>Morens beskæftigelse (ref: lønmodtager mellem/højt)</i>				
Lønmodtager, lavt niveau	0.04	(0.04)	-0.01	(0.04)
Lønmodtager, andet	0.03	(0.04)	-0.01	(0.04)
Self-beskftiget	-0.09	(0.06)	0.03	(0.07)
Ledig	0.14	(0.09)	-0.04	(0.09)
Udenfor arbejdsmarkedet	0.08 *	(0.04)	-0.03	(0.05)
Eksamenskarakter i skriftlig matematik, 9.kl.	-0.03 **	(0.00)	0.04 ***	(0.00)
Eksamenskarakter i skriftlig dansk, 9.kl.	0.00	(0.01)	0.01	(0.01)
Konstantled	0.22	(0.25)	-0.07	(0.27)
		R ² = 0.14	R ² = 0.19	
N=1531				

Anm.: Statistisk signifikante estimater er indikeret: *: 10%, **: 5% and ***: 1%. Kategorier for manglende værdier er inkluderet. Modellen er estimeret ved OLS.

Appendiks D. Clear Cut Opkvalificering 2 vs. Obligatorisk, LPM

Variabel	Afbrud		Gennemførelse	
	Parameterestimat (std. fejl)			
<i>Grundforløbspakke (ref: obligatorisk)</i>				
Opkvalificering 1	0.15 ***	(0.03)	-0.29 ***	(0.03)
<i>Skole (ref: EUC Vest)</i>				
CELF	-0.12 ***	(0.03)	-0.20 ***	(0.03)
Roskilde TS	-0.06	(0.06)	0.21 ***	(0.06)
Hemingsholm	-0.12 ***	(0.04)	0.05	(0.04)
Syddansk Erhvervsskole	-0.10 *	(0.05)	0.04	(0.05)
<i>Indgang (ref: Bygge og anlæg)</i>				
Fra jord til borg	-0.13 ***	(0.04)	0.14 ***	(0.04)
Håndværk og teknik	0.09	(0.06)	-0.03	(0.06)
Mekanik, transport og logistik	-0.07 *	(0.04)	0.07 *	(0.04)
Teknologi og kommunikation	0.07 *	(0.04)	-0.02	(0.03)
<i>Startår (ref: 2008)</i>				
2009	0.01	(0.03)	0.00	(0.03)
2010	0.09 ***	(0.03)	0.07 **	(0.03)
Kvinde	0.06	(0.04)	-0.02	(0.04)
Alder	0.04 **	(0.02)	0.00	(0.02)
Alder ²	0.00 **	(0.00)	0.00	(0.00)
<i>Ethnicitet (ref: Dansk)</i>				
Indvandrere	0.04	(0.10)	-0.01	(0.09)
Efterkommere	0.00	(0.08)	0.05	(0.08)
Kernefamilie	-0.03	(0.03)	0.04	(0.03)
Farens alder	-0.01 **	(0.00)	0.00	(0.00)
Morens alder	0.00	(0.00)	0.00	(0.00)
<i>Farens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	-0.07 **	(0.03)	0.10 ***	(0.03)
Gymnasial	0.06	(0.08)	-0.12	(0.08)
Kort/mellemlang videregående	-0.01	(0.05)	-0.03	(0.05)
Lang videregående	-0.06	(0.10)	0.11	(0.10)
<i>Morens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	-0.03	(0.03)	0.05 *	(0.03)
Gymnasial	-0.09	(0.06)	0.04	(0.06)
Kort/mellemlang videregående	-0.02	(0.04)	0.08 *	(0.04)
Lang videregående	0.02	(0.14)	-0.14	(0.13)
Familieindkomst (gennemsnit)	-0.01	(0.01)	0.00	(0.01)
<i>Farens beskæftigelse (ref: lønmodtager mellem/højt)</i>				
Lønmodtager, lavt niveau	0.01	(0.04)	-0.04	(0.04)
Lønmodtager, andet	0.01	(0.04)	0.00	(0.04)
Self-beskliftiget	0.06	(0.05)	-0.02	(0.05)
Ledig	-0.08	(0.10)	0.12	(0.10)
Udenfor arbejdsmarkedet	0.08	(0.05)	-0.05	(0.05)
<i>Morens beskæftigelse (ref: lønmodtager mellem/højt)</i>				
Lønmodtager, lavt niveau	0.03	(0.04)	-0.02	(0.04)
Lønmodtager, andet	0.04	(0.04)	-0.02	(0.04)
Self-beskliftiget	-0.14 **	(0.07)	0.05	(0.07)
Ledig	0.12	(0.09)	-0.04	(0.09)
Udenfor arbejdsmarkedet	0.03	(0.05)	-0.03	(0.04)
Eksamenskarakter i skriftlig matematik, 9.kl.	-0.03 ***	(0.00)	0.04 ***	(0.00)
Eksamenskarakter i skriftlig dansk, 9.kl.	0.01	(0.01)	0.01	(0.01)
Konstantled	0.22	(0.25)	0.02	(0.25)
	R ² = 0.13		R ² = 0.28	
	N=1497			

Anm.: Statistisk signifikante estimater er indikeret: *: 10%, **: 5% and ***: 1%. Kategorier for manglende værdier er inkluderet. Modellen er estimeret ved OLS.

Appendiks E. Deskriptiv statistik for baggrundsvariable, Clear Cut Analyse

	Obligatorisk (N=1223)				Opkvalificering 1 (N=308)				Opkvalificering 2 (N=274)			
	Gnst.	Std	Min	Max	Gnst.	Std	Min	Max	Gnst.	Std	Min	Max
<i>Skole</i>												
EUC Vest	0,32	0,47	0	1	0,18	0,39	0	1	0,32	0,47	0	1
CELF	0,25	0,43	0	1	0,1	0,31	0	1	0,57	0,5	0	1
Roskilde TS	0,06	0,24	0	1	0,28	0,45	0	1	0,09	0,29	0	1
Herningsholm	0,24	0,42	0	1	0,34	0,47	0	1	0	0,06	0	1
Syddansk Erhvervsskole	0,14	0,35	0	1	0,09	0,29	0	1	0,02	0,13	0	1
<i>Indgang</i>												
Bygge og anlæg	0,34	0,48	0	1	0,27	0,45	0	1	0,1	0,3	0	1
Fra jord til bord	0,15	0,36	0	1	0,1	0,31	0	1	0,17	0,38	0	1
Håndværk og teknik	0,05	0,22	0	1	0,05	0,22	0	1	0,02	0,15	0	1
Mekanik, transport og logistik	0,25	0,44	0	1	0,38	0,49	0	1	0,46	0,5	0	1
Teknologi og kommunikation	0,2	0,4	0	1	0,19	0,4	0	1	0,25	0,43	0	1
<i>Startår</i>												
2008	0,39	0,49	0	1	0,48	0,5	0	1	0,34	0,48	0	1
2009	0,36	0,48	0	1	0,32	0,47	0	1	0,41	0,49	0	1
2010	0,25	0,43	0	1	0,2	0,4	0	1	0,25	0,43	0	1
<i>Kvinde</i>												
Alder	16,42	3	14	47	15,94	1,37	14	28	15,91	1,29	14	28
Alder2	278,76	157,94	196	2209	255,91	51,41	196	784	254,76	48,34	196	784
Indvandrere	0,01	0,12	0	1	0,02	0,13	0	1	0,03	0,16	0	1
Efterkommere	0,02	0,14	0	1	0,03	0,18	0	1	0,05	0,21	0	1
Kernefamilie	0,79	0,41	0	1	0,78	0,41	0	1	0,7	0,46	0	1
Fars alder	44,26	6,12	30	84	42,95	5,54	31	61	43,1	5,98	31	67
Mors alder	41,2	5,27	29	79	40,64	4,79	30	57	40,42	4,71	29	57
<i>Farens uddannelse</i>												
Grundskole	0,33	0,47	0	1	0,34	0,47	0	1	0,36	0,48	0	1
Erhvervsfaglig	0,51	0,5	0	1	0,52	0,5	0	1	0,53	0,5	0	1
Gymnasial	0,02	0,14	0	1	0,03	0,17	0	1	0,01	0,12	0	1
Kort/mellemlang videregående	0,09	0,29	0	1	0,07	0,26	0	1	0,07	0,25	0	1
Lang videregående	0,02	0,14	0	1	0,02	0,15	0	1	0,01	0,09	0	1
Ukendt	0,02	0,15	0	1	0,02	0,13	0	1	0,02	0,15	0	1
<i>Morens uddannelse</i>												
Grundskole	0,3	0,46	0	1	0,36	0,48	0	1	0,39	0,49	0	1
Erhvervsfaglig	0,49	0,5	0	1	0,45	0,5	0	1	0,4	0,49	0	1
Gymnasial	0,04	0,21	0	1	0,05	0,21	0	1	0,03	0,18	0	1
Kort/mellemlang videregående	0,15	0,36	0	1	0,12	0,33	0	1	0,15	0,36	0	1
Lang videregående	0,01	0,09	0	1	0,01	0,1	0	1	0,01	0,09	0	1
Ukendt	0,01	0,12	0	1	0,01	0,11	0	1	0,01	0,1	0	1
Familieindkomst (gennemsnit)	3,14	1,91	-3,43	45,55	3,09	1,14	0,95	10,35	2,82	1	-3,58	10,35
Ukendt familieindkomst	0	0	0	0	0	0	0	0	0	0	0	0
<i>Farens beskæftigelse</i>												
Lønmodtager lavt niveau	0,39	0,49	0	1	0,36	0,48	0	1	0,42	0,49	0	1
Lønmodtager, mellem/højt niveau	0,15	0,35	0	1	0,12	0,33	0	1	0,12	0,33	0	1
Lønmodtager, andet	0,24	0,43	0	1	0,27	0,44	0	1	0,23	0,42	0	1
Selvstændig	0,1	0,31	0	1	0,14	0,34	0	1	0,07	0,25	0	1
Ledig	0,01	0,1	0	1	0,02	0,14	0	1	0,03	0,17	0	1
Udenfor arbejdsmarkedet	0,11	0,31	0	1	0,1	0,3	0	1	0,13	0,33	0	1
<i>Morens beskæftigelse</i>												
Lønmodtager lavt niveau	0,41	0,49	0	1	0,37	0,48	0	1	0,39	0,49	0	1
Lønmodtager, mellem/højt niveau	0,2	0,4	0	1	0,17	0,38	0	1	0,15	0,36	0	1
Lønmodtager, andet	0,16	0,37	0	1	0,19	0,39	0	1	0,17	0,37	0	1
Selvstændig	0,04	0,19	0	1	0,06	0,23	0	1	0,03	0,18	0	1
Ledig	0,02	0,14	0	1	0,02	0,15	0	1	0,02	0,15	0	1
Udenfor arbejdsmarkedet	0,17	0,37	0	1	0,19	0,39	0	1	0,24	0,43	0	1
Eksamenskarakter, Matematik 9.kl.	4,8	3,18	-1,5	12	4,32	3,06	-1,5	12	3,58	3,11	-1,5	12
Manglende karakter, Matematik	0,11	0,31	0	1	0,12	0,33	0	1	0,19	0,4	0	1
Eksamenskarakter, Dansk 9.kl.	3,62	2,64	-3	12	3,51	2,57	-3	10	3,05	2,63	0	12
Manglende karakter, Dansk	0,12	0,32	0	1	0,14	0,35	0	1	0,23	0,42	0	1

Appendiks F. Stratificeret CR Cox Model – Frafald og Fuldførelse – Bygge og Anlæg

	Frafald				Fuldførelse			
	(I)	(II)	(III)	(IV)	(I)	(II)	(III)	(IV)
WAS	-	-	-	-	-	-	-	-
Startår (ref: 2008)	-	-	-	-	-	-	-	-
2009	0.816*** (0.04)	0.813*** (0.04)	0.812*** (0.04)	0.803*** (0.04)	0.835*** (0.03)	0.824*** (0.03)	0.822*** (0.03)	0.819*** (0.03)
2010	0.947 (0.04)	0.950 (0.04)	0.954 (0.05)	0.940 (0.04)	0.828*** (0.04)	0.804*** (0.04)	0.792*** (0.03)	0.786*** (0.04)
Sommer	1.419*** (0.06)	1.410*** (0.06)	1.407*** (0.06)	1.405*** (0.06)	1.027 (0.05)	1.039 (0.05)	1.047 (0.05)	1.015 (0.05)
Kvinde	1.365*** (0.08)	1.320*** (0.07)	1.318*** (0.07)	1.309*** (0.07)	1.859*** (0.11)	1.974*** (0.12)	1.992*** (0.12)	1.949*** (0.12)
Kernefamilie (15-årsalderen)	0.764*** (0.03)	0.804*** (0.03)	0.820*** (0.03)	0.821*** (0.03)	1.357*** (0.05)	1.250*** (0.05)	1.194*** (0.05)	1.189*** (0.05)
Ethnicitet (ref: Dansk)	-	-	-	-	-	-	-	-
Indvandrer	1.226** (0.09)	1.114 (0.09)	1.063 (0.09)	1.048 (0.09)	0.653*** (0.05)	0.820* (0.08)	0.914 (0.09)	0.862 (0.08)
Efterkommere	1.428*** (0.13)	1.288** (0.13)	1.237* (0.13)	1.273* (0.13)	0.695*** (0.07)	0.863 (0.10)	0.960 (0.11)	0.945 (0.11)
Alder	1.103*** (0.03)	1.093*** (0.03)	1.063* (0.03)	1.063* (0.03)	0.998 (0.02)	1.011 (0.02)	1.054* (0.02)	1.044 (0.02)
Alder ²	0.998** (0.00)	0.999** (0.00)	0.999** (0.00)	0.999** (0.00)	1.000 (0.00)	1.000 (0.00)	1.000 (0.00)	1.000 (0.00)
Farens uddannelse (ref: grundskole)	-	-	-	-	-	-	-	-
Erhvervsfaglig	-	0.892* (0.04)	0.902* (0.04)	0.897* (0.04)	-	0.993 (0.04)	0.969 (0.04)	0.969 (0.04)
Gymnasial	-	0.937 (0.13)	0.964 (0.13)	0.969 (0.13)	-	0.956 (0.13)	0.896 (0.11)	0.885 (0.11)
Kort/mellemlang videregående	-	0.878 (0.07)	0.887 (0.07)	0.880 (0.07)	-	1.076 (0.07)	1.023 (0.07)	1.009 (0.07)
Lang videregående	-	0.894 (0.14)	0.902 (0.14)	0.919 (0.14)	-	0.850 (0.11)	0.827 (0.10)	0.858 (0.10)
Morens uddannelse (ref: grundskole)	-	-	-	-	-	-	-	-
Erhvervsfaglig	-	0.907* (0.04)	0.920 (0.04)	0.923 (0.04)	-	1.154*** (0.05)	1.119** (0.05)	1.113* (0.05)
Gymnasial	-	0.959 (0.11)	0.979 (0.11)	0.964 (0.11)	-	1.152 (0.11)	1.103 (0.11)	1.080 (0.10)
Kort/mellemlang videregående	-	0.930 (0.07)	0.950 (0.07)	0.958 (0.07)	-	1.116 (0.07)	1.066 (0.07)	1.063 (0.07)
Lang videregående	-	1.108 (0.16)	1.146 (0.17)	1.152 (0.17)	-	0.851 (0.14)	0.783 (0.13)	0.778 (0.12)
Farens beskæftigelse (ref: lønmodtager lavt niveau)	-	-	-	-	-	-	-	-
Selvstændig	-	1.018 (0.08)	1.023 (0.08)	1.022 (0.08)	-	1.113 (0.07)	1.094 (0.07)	1.099 (0.07)
Lønmodtager, mellemhøjt niveau	-	1.132 (0.08)	1.140 (0.08)	1.131 (0.08)	-	0.955 (0.06)	0.954 (0.06)	0.951 (0.06)
Lønmodtager, andet	-	1.065 (0.06)	1.066 (0.06)	1.067 (0.06)	-	0.910 (0.05)	0.905 (0.05)	0.893 (0.05)
Ledig	-	1.029 (0.11)	1.031 (0.11)	1.029 (0.11)	-	0.833 (0.09)	0.826 (0.09)	0.800* (0.09)
Udenfor arbejdsmarkedet	-	1.135* (0.07)	1.128* (0.07)	1.132* (0.07)	-	0.961 (0.06)	0.975 (0.06)	0.972 (0.06)
Morens beskæftigelse (ref: lønmodtager lavt niveau)	-	-	-	-	-	-	-	-
Selvstændig	-	0.780* (0.10)	0.772* (0.10)	0.782 (0.10)	-	1.140 (0.10)	1.148 (0.10)	1.149 (0.10)
Lønmodtager, mellemhøjt niveau	-	0.943 (0.07)	0.960 (0.07)	0.971 (0.07)	-	1.044 (0.06)	0.997 (0.06)	1.025 (0.06)
Lønmodtager, andet	-	1.019 (0.07)	1.026 (0.07)	1.036 (0.07)	-	1.023 (0.06)	0.998 (0.06)	1.019 (0.06)
Ledig	-	1.111 (0.09)	1.091 (0.09)	1.082 (0.09)	-	0.973 (0.08)	1.013 (0.08)	0.990 (0.08)
Udenfor arbejdsmarkedet	-	1.188** (0.06)	1.177** (0.06)	1.181** (0.06)	-	0.873* (0.05)	0.879* (0.05)	0.890* (0.05)
Familieindkomst (gemensnit)	-	0.999 (0.01)	1.002 (0.01)	1.002 (0.01)	-	1.053*** (0.01)	1.049*** (0.01)	1.042*** (0.01)
Eksamenskarakter, 9-kl	-	-	0.944*** (0.01)	0.942*** (0.01)	-	-	1.130*** (0.01)	1.122*** (0.01)
Skole (ref.: 15. KTS (WAS))	-	-	-	-	-	-	-	-
Randers TS (WAS)	-	-	-	1.487*** (0.14)	-	-	-	1.356** (0.13)
Syddansk Erhvervs-skole (WAS)	-	-	-	1.002 (0.08)	-	-	-	0.894 (0.07)
Silkeborg TS (WAS)	-	-	-	1.187 (0.16)	-	-	-	0.959 (0.12)
EUC Vest (CC)	-	-	-	1.118 (0.08)	-	-	-	0.892 (0.06)
EUC Nordvest (CC)	-	-	-	1.078 (0.13)	-	-	-	1.109 (0.12)
Hermingsholm Erhvervs-skole (CC)	-	-	-	0.996 (0.09)	-	-	-	0.909 (0.07)
CELF (CC)	-	-	-	0.970 (0.07)	-	-	-	0.370*** (0.03)
Uddannelsescenter Høstebro (CC)	-	-	-	1.109 (0.11)	-	-	-	1.181 (0.11)
Observationer	6843	6843	6843	6843	6843	6843	6843	6843
Pseudo R-squared	0.007	0.008	0.009	0.010	0.007	0.009	0.012	0.017
AIC	43463.812	43448.767	43415.153	43398.866	52175.753	52108.712	51922.341	51705.802
BIC	43552.122	43681.021	43661.068	43699.429	52244.063	52240.965	52168.256	52006.365

Ann.: Hazard Ratioer, p-værdier i parentes. Statistisk signifikante estimater er indikeret: *, 5%; **, 1%; ***, 0.1%. Kategorier for manglende værdier er inkluderet. Efron method for ties (Clevens et al. 2010: s.15)

Appendiks G. Logit – Praktikvej – Bygge og Anlæg

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
<i>Startår (ref: 2008)</i>							
2009	-	-	0.567*** (0.02)	0.560*** (0.02)	0.556*** (0.02)	0.533*** (0.02)	0.507*** (0.02)
2010	-	-	0.550*** (0.02)	0.500*** (0.02)	0.494*** (0.02)	0.466*** (0.02)	0.442*** (0.02)
Sommer	-	-	1.782*** (0.06)	1.580*** (0.06)	1.590*** (0.06)	1.613*** (0.06)	1.765*** (0.07)
Kvinde	-	-	-	0.204*** (0.02)	0.205*** (0.02)	0.210*** (0.02)	0.216*** (0.02)
Kernefamilie (15-årsalderen)	-	-	-	1.900*** (0.08)	1.766*** (0.08)	1.633*** (0.07)	1.522*** (0.07)
<i>Ethnicitet (ref: Dansk)</i>							
Indvandrere	-	-	-	0.191*** (0.02)	0.245*** (0.03)	0.303*** (0.04)	0.318*** (0.04)
Efterkommere	-	-	-	0.125*** (0.02)	0.153*** (0.03)	0.193*** (0.03)	0.242*** (0.04)
Allder	-	-	-	1.158*** (0.02)	1.171*** (0.02)	1.188*** (0.02)	1.139*** (0.02)
Allder ²	-	-	-	0.998*** (0.00)	0.998*** (0.00)	0.998*** (0.00)	0.999*** (0.00)
<i>Farens uddannelse (ref: grundskole)</i>							
Erhvervsfaglig	-	-	-	-	1.243*** (0.05)	1.150*** (0.05)	1.179*** (0.05)
Gymnasial	-	-	-	-	0.765 (0.11)	0.674** (0.10)	0.691* (0.10)
Kort/mellemlang videregående	-	-	-	-	0.992 (0.07)	0.897 (0.06)	0.894 (0.06)
Lang videregående	-	-	-	-	0.612*** (0.08)	0.479*** (0.07)	0.549*** (0.08)
<i>Morens uddannelse (ref: grundskole)</i>							
Erhvervsfaglig	-	-	-	-	1.225*** (0.05)	1.125** (0.05)	1.157** (0.05)
Gymnasial	-	-	-	-	1.321** (0.13)	1.220 (0.13)	1.277* (0.14)
Kort/mellemlang videregående	-	-	-	-	1.070 (0.06)	0.970 (0.07)	0.987 (0.07)
Lang videregående	-	-	-	-	0.626* (0.12)	0.523** (0.10)	0.578** (0.12)
<i>Farens beskæftigelse (ref: lønmodtager lavt niveau)</i>							
Selvstændig	-	-	-	-	-	1.452*** (0.09)	1.451*** (0.09)
Lønmodtager, mellem/højt niveau	-	-	-	-	-	0.866* (0.05)	0.934 (0.06)
Lønmodtager, andet	-	-	-	-	-	1.008 (0.06)	1.016 (0.06)
Ledig	-	-	-	-	-	0.744* (0.09)	0.803 (0.10)
Udenfor arbejdsmarkedet	-	-	-	-	-	0.859* (0.06)	0.924 (0.06)
<i>Morens beskæftigelse (ref: lønmodtager lavt niveau)</i>							
Selvstændig	-	-	-	-	-	1.178 (0.11)	1.154 (0.11)
Lønmodtager, mellem/højt niveau	-	-	-	-	-	0.976 (0.06)	1.032 (0.06)
Lønmodtager, andet	-	-	-	-	-	1.199** (0.07)	1.190** (0.07)
Ledig	-	-	-	-	-	0.848 (0.08)	0.849 (0.08)
Udenfor arbejdsmarkedet	-	-	-	-	-	0.893 (0.05)	0.907 (0.06)
Familieindkomst (gennemsnit)	-	-	-	-	-	1.099*** (0.01)	1.114*** (0.01)
<i>Skole (ref.15. KTS (WAS))</i>							
KTS	-	-	-	-	-	-	0.231*** (0.02)
TEC	-	-	-	-	-	-	1.187 (0.11)
Erhvervsskolen Nordsjælland	-	-	-	-	-	-	0.587*** (0.05)
Roskilde TS	-	-	-	-	-	-	0.593*** (0.06)
Tradium	-	-	-	-	-	-	1.137 (0.12)
Uddannelsescenter Holstebro	-	-	-	-	-	-	1.489** (0.18)
EUC Sjælland	-	-	-	-	-	-	0.446*** (0.04)
CELF	-	-	-	-	-	-	0.311*** (0.05)
Syddansk Erhvervsskole	-	-	-	-	-	-	0.668*** (0.05)
Svendborg Erhvervsskole	-	-	-	-	-	-	0.237*** (0.04)
EUC Syd	-	-	-	-	-	-	0.692*** (0.07)
EUC Vest	-	-	-	-	-	-	0.901 (0.08)
Herningsholm Erhvervsskole	-	-	-	-	-	-	3.096*** (0.24)
Silkeborg Tekniske Skole	-	-	-	-	-	-	0.492*** (0.10)
Skive Tekniske Skole	-	-	-	-	-	-	1.034 (0.14)
EUC Nordvest	-	-	-	-	-	-	1.853*** (0.22)
Mercantec	-	-	-	-	-	-	1.722*** (0.18)
Tech College Aalborg	-	-	-	-	-	-	0.510*** (0.04)
Eksamenskarakter, 9.kl.	-	-	-	-	-	-	1.060*** (0.01)
Observationer	8640	24947	24631	24631	24631	24612	24612
Pseudo R-squared	0.007	0.026	0.099	0.107	0.118	0.170	0.177
AIC	8578.345	22287.289	20359.343	20214.095	19990.169	18834.721	18694.668
BIC	8592.473	22319.787	20448.572	20384.442	20274.080	19264.604	19140.772

Appendiks H. Sammenligning mellem CR Cox og Fine-Gray modellerne – Frafald – Bygge og Anlæg

	Dropout			
	Cox CRM		Fine-Gray CRM	
WAS	0.833*	(0.02)	0.803**	(0.01)
<i>Startår (ref: 2008)</i>				
2009	0.803***	(0.00)	0.896*	(0.01)
2010	0.940	(0.20)	0.990	(0.82)
Sommer	1.400***	(0.00)	1.376***	(0.00)
Kvinde	1.311***	(0.00)	0.935	(0.24)
Kernefamilie (15-årsalderen)	0.820***	(0.00)	0.779***	(0.00)
<i>Ethnicitet (ref: Dansk)</i>				
Indvandrere	1.048	(0.58)	1.100	(0.24)
Efterkommere	1.270*	(0.02)	1.269*	(0.01)
Alder	1.064*	(0.02)	1.041	(0.15)
Alder ²	0.999*	(0.04)	0.999	(0.06)
<i>Farens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	0.896*	(0.01)	0.919	(0.06)
Gymnasial	0.958	(0.76)	0.964	(0.79)
Kort/mellemlang videregående	0.879	(0.09)	0.882	(0.10)
Lang videregående	0.908	(0.54)	0.954	(0.76)
<i>Morens uddannelse (ref: grundskole)</i>				
Erhvervsfaglig	0.924	(0.08)	0.885**	(0.01)
Gymnasial	0.962	(0.73)	0.925	(0.49)
Kort/mellemlang videregående	0.959	(0.59)	0.931	(0.35)
Lang videregående	1.144	(0.36)	1.242	(0.14)
<i>Farens beskæftigelse (ref: lønmodtager lavt niveau)</i>				
Selvstændig	1.026	(0.74)	0.961	(0.62)
Lønmodtager, mellem/højt niveau	1.130	(0.09)	1.159*	(0.04)
Lønmodtager, andet	1.066	(0.28)	1.084	(0.17)
Ledig	1.024	(0.82)	1.113	(0.26)
Udenfor arbejdsmarkedet	1.133*	(0.04)	1.113	(0.08)
<i>Morens beskæftigelse (ref: lønmodtager lavt niveau)</i>				
Selvstændig	0.777*	(0.05)	0.746*	(0.02)
Lønmodtager, mellem/højt niveau	0.974	(0.71)	0.953	(0.50)
Lønmodtager, andet	1.036	(0.60)	1.030	(0.66)
Ledig	1.081	(0.34)	1.091	(0.27)
Udenfor arbejdsmarkedet	1.183**	(0.00)	1.208***	(0.00)
Familieindkomst (gennemsnit)	1.002	(0.91)	0.986	(0.32)
Eksamenskarakter, 9.kl.	0.942***	(0.00)	0.897***	(0.00)
<i>Skole (ref.: 15. KTS (WAS))</i>				
Randers TS (WAS)	1.499***	(0.00)	1.370***	(0.00)
Syddansk Erhvervsskole (WAS)	1.007	(0.93)	1.068	(0.37)
Silkeborg TS (WAS)	1.205	(0.17)	1.121	(0.42)
EUC Vest (CC)	1.121	(0.10)	1.154*	(0.03)
EUC Nordvest (CC)	1.080	(0.53)	1.036	(0.77)
Herningsholm Erhvervsskole (CC)	1.001	(0.99)	1.022	(0.80)
CELF (CC)	0.983	(0.81)	1.350***	(0.00)
Uddannelsescenter Holstebro (CC)	1.101	(0.32)	0.994	(0.95)
Observationer	6843		6843	

Note: (Sub)Hazard Ratioer; p-værdier i parentes