

Hovedoppgave for cand.polit-graden

Svart Arbeid

Endringen i utbredelse fra 1980 til 2001.

Wiljar Gudmund Hansen

November 2002

**Økonomisk institutt
Universitetet i Oslo**

Forord.

Denne hovedoppgaven er skrevet ved Stiftelsen Frischsenteret for samfunnsøkonomisk forskning i Oslo. Prosjektet denne hovedoppgaven rapporterer fra er finansiert av Skattedirektoratet og delfinansiert av Norges Forskningsråd (NFR-prosjekt nr. 145618/510). Prosjektets resultater er i sin helhet presentert i Goldstein m.fl. (2002) *Svart arbeid fra 1980 til 2001*, Frisch-senterets rapportserie 3/2002.

Tone Ognedal og Steinar Strøm har veiledet meg i skrivingen av oppgaven. Jeg ønsker å takke dem for inspirasjon og tålmodighet. I tillegg ønsker jeg å takke Erling Barth ved Institutt for samfunnsforskning for velvillig hjelp til estimeringen og for nyttige kommentarer underveis.

Innholdsfortegnelse

1.	
Innledning.....	1
2.	
Datainnsamling.....	5
3. Metoder for å undersøke omfanget av den svarte økonomien.....	6
4. Beskrivende statistikk.....	10
5. Valgmodellen.....	16
6. Logit-modellen.....	20
7. Estimering.....	23
8. Logit-sannsynligheter for svart arbeid siste 12 måneder uavhengig av bransje.....	27
9. Logit-sannsynligheter for svart arbeid siste 12 måneder avhengig av bransje.....	33
10. Sannsynligheten for svart arbeid, hvis en har muligheten til det.....	41
11. Konklusjon.....	48
12.	
Referanser.....	51

1. Innledning.

I denne hovedoppgaven skal jeg se på endringen i utbredelsen av svart arbeid mellom 1980 og 2001, og hva som kan forklare denne endringen. Med svart arbeid menes arbeid mot betaling som normalt sett skulle vært oppgitt til ligningsmyndighetene, men som ikke blir oppgitt. Det er altså skatteunndragelsen som er ulovlig, ikke arbeidet i seg selv. Dette innebærer at jeg ser bort fra ulovlig inntektsbringende virksomhet som smugling, prostitusjon og narkotikaomsetning. Vi ser bort fra momsunndragelse og får trolig heller ikke med svarte bytter av tjenester. Analysen bygger på data fra spørreundersøkelser i 1980, 1989 og 2001. Disse tre spørreundersøkelsene ble utført med omtrent det samme spørreskjemaet slik at jeg kan sammenligne resultatene mellom årstallene.¹

Basert på data fra de tre spørreundersøkelsene finner vi (Goldstein m.fl. 2002) at utbredelsen av svart arbeid har gått ned: det er færre som svarer at de har jobbet svart siste 12 måneder, færre som svarer at de har jobbet svart noen gang og færre som svarer at de ville arbeidet svart dersom muligheten var tilstede. Jeg skal i denne hovedoppgaven forsøke å forklare hvilke faktorer som ligger bak disse endringene.

Jeg skal forklare endringen i utbredelse av svart arbeid gjennom å analysere hva som påvirker beslutningen om å jobbe svart eller ikke. Valget mellom svart arbeid eller ikke er et moralsk valg som ligner det å stjele, omsette narkotika eller generelt gjøre noe ulovlig. Den moralske kostnaden ved å arbeide svart øker trolig ikke med antall timer en velger å jobbe. Teoretiske analyser av tilbudet av svart arbeid (Allingham og Sandmo 1972) går i liten grad inn på dette valget. De ser på hvor mange timer man vil arbeide svart som et resultat av et kontinuerlig valg. Faktorer som påvirker dette valget er svart og hvit timelønn, skattesatsen, sjansen for å bli tatt og straffeskatten dersom man blir tatt. Jeg skal ikke se på valget av antall svarte timer man eventuelt ønske å arbeide svart, og på den måten ikke se på endringen i omfanget av svart arbeid.

Valget om å arbeide svart eller ikke er et diskret valg hvor en persons moralske kostnad ved svart arbeid, i tillegg til muligheten denne personen har til å arbeide svart, spiller inn på dette

¹ Datafilene med individdata fra 1989 er dessverre tapt, jeg har derfor basert meg på individdata fra 1980 og 2001, samt frekvenstabeller fra 1989.

valget. Personens moralske kostnad og muligheten han har for å arbeide svart varierer trolig med faktorer som kjønn, alder, utdanning, bransje og yrke. Dette er da faktorer som påvirker valget av svart arbeid eller ikke i tillegg til den forventede netto gevinsten ved å arbeide svart. Denne forventede netto gevinsten avhenger av lønn, skattesats, sjansen for å bli tatt og straffeskatt ved oppdagelse.

Analyser av data fra spørreundersøkelsene viser at andelen som arbeider svart har gått ned gjennom perioden. Det samme har andelen som sier at de ville jobbe svart dersom de hadde muligheten. Ved hjelp av logit-analyser har jeg analysert hvilke faktorer som påvirker tilbøyeligheten til å jobbe svart. Jeg finner at det er større sannsynlighet for svart arbeid blant yngre enn eldre, blant menn enn kvinner og blant lavt utdannede enn høyt utdannede. En lav skattemoral medfører høyere sannsynlighet for skatteunndragelse, og dersom personen anser sjansen for å bli tatt som lav er sannsynligheten større for at han unndrar skatt enn om han anser sjansen for å bli tatt som høy.

Det er store variasjoner i frekvensen av svart arbeid mellom ulike bransjer. Omfanget av svart arbeid har falt i samtlige næringer, men fremdeles er det slik at utbredelsen av svart arbeid er særdeles høy i noen bransjer. Analyser gjort på endringen i utbredelse av svart arbeid viser at det ikke er endringen i bransjesammensetningen i samfunnet som er grunnen til fallet i antallet som arbeider svart.

Vi finner at det er flere faktorer som kan bidra til å forklare at utbredelsen av svart arbeid er lavere i 2001 enn hva den var i 1980. For det første er utdanningsnivået i befolkningen blitt betydelig høyere, for det andre har arbeidsstyrken blitt eldre og for det tredje har kvinneandelen i registrert arbeid økt. Logit-analysene viser at dette er kjennetegn som demper tilbøyeligheten for svart arbeid.

En analyse basert på individdata innhentet ved spørreundersøkelser kan være med på å kartlegge hvilke motiver som ligger bak beslutningen om å arbeide svart. Dersom en får større grad av klarhet i motivene bak valget av svart arbeid kan det være enklere å sette inn virkemidler som reduserer det svarte arbeidet og dermed rette opp de feilallokeringene dette medfører, da skatteunndragelse har flere uheldige virkninger utover det at staten kan miste inntekter. For det første kan skatteunndragelse virke vridende på konkurransen i en næring

dersom noen av aktørene i næringen benytter svart arbeidskraft og på den måten har lavere utgifter enn sine konkurrenter. For det andre kan skatteunndragelse føre til en feilallokering av arbeidskraft i samfunnet gjennom at enkelte bransjer tiltrekker seg arbeidskraft ved å tilby gode muligheter for unndragelse. I tillegg til dette kan uregistrert arbeidskraft åpne for brudd på sikkerhetsforskriftene på arbeidsplassen.

Kapittel 2 gir en beskrivelse av datainnsamlingen, utvalgsstørrelser og svarprosenter.

I kapittel 3 går jeg gjennom ulike metoder for å undersøke omfanget av svart arbeid. Jeg diskuterer fordeler og ulemper ved de ulike metodene, og avslutter kapitlet med en diskusjon av problemene rundt utførelsen av en spørreundersøkelse omkring omfanget og utbredelsen av den svarte økonomien.

I kapittel 4 presenteres noen resultater fra spørreundersøkelsene i 1980, 1989 og 2001 i form av frekvenstabeller. Her ser vi at utbredelsen av svart arbeid har falt fra 1980 til 2001 både blant de som har arbeidet svart siste 12 måneder og de som har arbeidet svart noen gang. Jeg viser også at utbredelsen av de som svarer at de ønsker å arbeide svart, hvis de hadde muligheten, har falt fra 1980 til 2001. Det presenteres også tabeller som viser andelen som jobber svart betinget på alder, kjønn, utdanning og noen utvalgte bransjer.

I kapittel 5 viser jeg hovedresultatene Allingham og Sandmos porteføljemodell for skatteunndragelse (1972) gir, og diskuterer hvilke forutsetninger som ligger til grunn for denne modellen. Deretter presenteres den diskrete valgmodellen som dataanalysen bygger på, og drøfter forutsetningen om et diskret valg opp mot antakelsen om kontinuerlige valg som ligger til grunn for porteføljevalgmodellen til Allingham og Sandmo.

I kapittel 6 presenter jeg logitmodellen og dens implikasjoner.

I kapittel 7 presenterer de analysene jeg skal utføre senere i oppgaven og diskuterer hvilke variable som bør være med i slike analyser og hvilke som ikke bør være med.

I kapittel 8 estimerer jeg sannsynligheten for at en person har unndratt inntekt fra beskatning siste 12 måneder for 1980 og 2001. Jeg analyserer endringene i sannsynlighet for svart arbeid mellom de to årstallene gjennom studier av referansepersoner. Deretter dekomponerer

jeg den totale endringen i svart arbeid de siste 12 måneder i variabelendringer og koeffisient endringer.

I kapittel 9 estimerer jeg sannsynligheten for svart arbeid siste 12 måneder når jeg tar hensyn til personenes valg av bransje for 1980 og 2001. Jeg sammenligner resultatene for de to årstallene og dekomponerer den totale endringen i skatteunndragelse i variabelendringer og koeffisientendringer. Avslutningsvis i kapitlet presenterer jeg en enkel metode for å undersøke om det er endringen i bransjesammensetningen i samfunnet som er grunnen til fallet i antall som arbeider svart.

I kapittel 10 estimerer jeg sannsynligheten for at en person ønsker å unndra, hvis det er mulig. Også i denne estimeringen kjører jeg modellen på begge år for å kunne sammenligne resultatene mellom årene. Også i dette kapitlet dekomponerer jeg den totale endringen i utbredelse i variabelendringer og koeffisientendringer. Modellen i kapittel 8 er kjørt på et randomisert utvalg, avslutningsvis i kapitlet viser jeg derfor hvilke endringer i estimatene en får dersom en ikke randomiserer.

Kapittel 11 konkluderer og oppsummerer de viktigste resultatene.

Estimeringen i denne oppgaven er utført i programpakken SAS. Datasettet fra 1980 er tilgjengelig gjennom Norsk Samfunnsvitenskaplig Datatjeneste, NSD. Spørreskjemaet for 2001, samt frekvenstabeller for 2001 er tilgjengelig i Goldstein m.fl.(2002). Frekvenstabeller for 1980 sammen med spørreskjemaet for 1980 er tilgjengelig i Sporastøyl (1982).

2. Datainnsamlingen.

Spørreundersøkelsene vi baserer oss på er foretatt av MMI i henholdsvis september 1980, mai 1989 og september/oktober 2001. I hovedsak er samme spørreskjema blitt benyttet i tre forskjellige år, noe som innebærer at der spørsmålene har samme ordlyd, vil det være mulig å sammenligne resultatene undersøkelsene i mellom for å analysere endringer over tid. For nærmere beskrivelse av datainnsamlingen og utvalgsskjevheter i 1980 henviser jeg til Sporastøyl (1982). Undersøkelsen i 1989 ble utført av MMI etter oppdrag fra SAF og Økonomisk Rapport. Fremgangsmåten i 1989 var tilnærmet identisk med måten 2001 undersøkelsen ble gjennomført på. Dessverre er datafilene fra 1989 tapt, noe som innebærer at resultatene fra denne undersøkelsen kun er å finne i papirkopi. Dette gjør at betingede analyser på 1989 datasettet ikke er mulig. Jeg vil derfor kun bruke resultatene fra 1989 der hvor jeg sammenligner svarprosenten for hele utvalget.

2001-undersøkelsen ble gjennomført i forbindelse med omnibusundersøkelsen til MMI, med personlig intervju i perioden 10.-17. september og 8.-15. oktober. Etter at det personlige intervjuet var ferdig ble samtlige intervjuede personer spurt om de ville delta i en undersøkelse av hvor mye arbeid som ble utført mot betaling som ikke blir oppgitt til ligningsmyndighetene. I denne forespørselen ble det gitt full forsikring om anonymitet. De som sa seg villig til å delta fikk så utlevert et umerket spørreskjema, svarkonvolutt og et informasjonsbrev. Utvalgsstørrelser og svarprosenten for de tre årene er vist i tabellen nedenfor.

Tabell 1. utvalgsstørrelser og svarprosenten.

	1980	1989	2001
Forespurt	1198	1130	1690
Ville delta	80 %	87 %	81 %
Svarprosent	73 %	73 %	58 %

Som det fremgår av tabellen er svarprosentene i de tre årene relativt høye. Svarprosenten er lavere i 2001 enn i 1980. I følge MMI harmonerer nedgangen i svarprosent med tendenser de merker til økende likegyldighet til å delta i slike undersøkelser. Sammensetningen i utvalgene med hensyn til personkjennetegn som alder, kjønn, utdanning, partivalg mm er i godt samsvar med fordelingen av disse kjennetegnene i den norske befolkningen. For en mer utfyllende beskrivelse av utvalgsskjevheter henviser jeg til Goldstein m.fl. (2002).

3. Metoder for å undersøke omfanget av den svarte økonomien.

Det er utført en rekke studier for en rekke land av omfanget av den svarte økonomien. Dette mangfoldet av undersøkelser tar i bruk mange ulike metoder for å finne hvor omfattende skatteunndragelsene er i de ulike landene. I de påfølgende sidene skal jeg komme kort inn på fordelene og ulempene ved noen av disse metodene. Jeg har basert min framstilling på Eide (2000), og henviser til denne rapporten for mer utfyllende drøfting. Det metodiske hovedskillet går mellom de indirekte og de direkte metodene. Mens de indirekte metodene består av å utforske de spor som den svarte økonomien etterlater seg, er de direkte metodene mer opptatt av å studere folks atferd, enten dette er i det virkelige liv eller i konstruerte situasjoner.

Indirekte metoder.

De viktigste indirekte metodene tar utgangspunkt i:

- (i) hvor mye penger de illegale transaksjonene legger beslag på,
- (ii) hvor stor differanse det er mellom husholdningenes inntekter og registrerte eller beregnede utgifter,
- (iii) og hvor stor differanse det er mellom faktisk og registrert deltagelse i arbeidsmarkedet.

(i) Pengebruk ved ulovlige transaksjoner – transaksjonsmetoden.²

Denne metoden har som forutsetning at alle svarte transaksjoner gjøres opp i kontanter. I denne metoden måles den svarte økonomien til å være differansen mellom den faktiske pengemengden i landet og den delen av pengemengden som er anslått å skulle betjene de legale transaksjonene. Resultatene av disse undersøkelsene er ganske usikre av flere grunner. For det første baserer de seg på at det i et basisår ikke brukes penger til illegale transaksjoner. For det andre kan det faktum at pengenes omløpshastighet er forskjellig på forskjellige områder bidra til usikkerhet rundt resultatene av disse studiene. For det tredje virker også denne metoden usikker da det kan være mange grunner til at privat sektor ønsker å sitte med kontanter. Mange eldre har fremdeles overraskende mange kontanter uten at

disse skal brukes på kjøp av illegale tjenester. Undergrunnsøkonomien i Norge er, ved bruk av denne metoden, beregnet til å ha økt fra 1,5 % av BNP i 1960 til 18 % av BNP i 1995³. Denne sterke økningen i den svarte økonomien i Norge er det vanskelig å finne annen informasjon som kan bekrefte.

(ii) Differansen mellom utgifter og inntekter.

Enkelte studier tar utgangspunkt i at det i nasjonalregnskapet bør være likhet mellom inntekter og utgifter. Dersom det ikke er likhet mellom inntektene og utgiftene kan denne differansen betraktes som et anslag på den svarte økonomien. Det er så vidt meg bekjent ikke utført noen slik beregning for den norske undergrunnsøkonomien. Også denne metoden er det knyttet stor usikkerhet til. Dersom alle komponentene på utgiftssiden i nasjonalbudsjettet var uten målefeil ville dette være en god metode for estimering av størrelsen på den svarte økonomien. Nasjonalbudsjettet er dessverre ikke uten målefeil.

(iii) Differansen mellom faktisk og registrert deltagelse i arbeidsmarkedet.

En nedgang i arbeidsdeltakelsen i den offisielle økonomien kan brukes som en indikator på økt aktivitet i den svarte økonomien, under forutsetning av at arbeidsstokken er konstant. Slike beregninger er gjort for Italia og USA. Svakheten ved denne metoden er at det kan ligge helt andre grunner bak et fall i yrkesdeltakelsen enn en økning i den svarte økonomien. En kanskje viktigere innvending er at personer både kan arbeide i den offisielle økonomien og i den svarte. Dette er grunner til at estimater utregnet etter denne metoden er svake indikatorer på størrelsen på skyggeøkonomien.

Direkte metoder

Dette er som oftest mikroøkonomiske tilnærminger som omfatter:

- (i) metoder basert på observasjoner av skatteyteres atferd (kontroll av selvangivelser, bokrevisjon),
- (ii) eksperimenter der forskjellige grupper av skatteyttere utsettes for forskjellig behandling,
- (iii) laboratorie-eksperimenter med fiktiv utfylling av forenklete selvangivelser, og
- (iv) spørreundersøkelser om kjennskap til og deltagelse i svart økonomi.

² Denne metoden er utviklet av Feige i flere artikler, bl.a. Feige, Edgar L (1989)

(i) Undersøkelser basert på TCMP.

En rekke studier av årsaker til skatteunndragelse baserer seg på informasjon hentet fra det amerikanske *Taxpayer Compliance Measurement Program (TCMP)*. Disse dataene viser hva personene i utvalget har oppgitt på hver post i selvangivelsen, samt hva en skatterevisor har beregnet det riktige beløpet til å være. Differansen mellom oppgitte beløp og beregnede beløp er et mål på skatteunndragelse. En svakhet ved TCMP-dataene er revisorenes anslag på individenes inntekt. Disse er usikre da skattereviseoren kun har tilgang på to typer av systematiske tilleggsopplysninger: forskuddstrekk og informasjon om renter og utbytte fra debitorer. TCMP-dataene er bedre egnet til å estimere skatteunndragelsens avhengighet av variable som inntekt, region, alder og sivilstatus enn hva de er egnet som et verktøy til å anslå størrelsen på den svarte økonomien.

(ii) Eksperimenter der grupper av skatteyttere blir utsatt for forskjellig behandling.

Det er utført få slike kontrollerte eksperimenter på skattebetaleres atferd. Grunnen til dette kan være at slike eksperimenter er vanskelig å gjennomføre uten at det går på bekostning av personvernet. Eksperimenter av denne typen er heller ikke egnet til å måle størrelsen på skatteunndragelsene, men kan brukes til å undersøke atferden til skatteytterne. Et slikt eksperiment ble utført i Minnesota i USA i 1995. I dette eksperimentet ble 1724 skatteyttere kontaktet og gitt opplysning om at deres selvangivelse ville bli meget grundig kontrollert av skattemyndighetene, og dersom det ble funnet uregelmessigheter ville også foregående års selvangivelser bli kontrollert. Blumenthal et al. (1998) sammenlignet selvangivelsene det året hvor det var gitt opplysninger om kontroll med foregående års selvangivelse for personene i eksperimentet.

(iii) Laboratorie-eksperimenter.

Mange forskere har forsøkt å hente informasjon fra konstruerte laboratorie-eksperimenter. Slike spill foregår som regel over flere perioder hvor deltakerne i hver periode mottar en tenkt inntekt, som de skal betale skatt av under en gitt sannsynlighet for kontroll av skatteinnbetalingene og de må betale straffeskatt dersom de blir oppdaget i skattefusk. Når spillet er over får som regel deltakerne beholde en sum som gjerne er proporsjonal med hva de satt igjen med når spillet var ferdig. Det forskerne ønsker å oppnå med slike eksperimenter er å se hvordan endringer i ulike politikkvariable påvirker unndragelsen.

³ Schneider, Friedrich og Enste, Dominik H (2000) tabell 3 s. 81.

Innvendingene mot slike eksperimenter går mye på laboratoriets unaturlige element hvor det er vanskelig å designe spill slik at moral, følelser og sosialt ansvar kommer inn som komponenter.

(iv) Spørreundersøkelser.

Spørreundersøkelser om skatteunndragelse er gjennomført i en rekke land, bl.a. Norge⁴, Danmark⁵ og Sverige⁶. Fordelen ved slike undersøkelser ligger i den detaljerte informasjonen en kan oppnå omkring strukturen i den svarte økonomien. Men resultater fra slike undersøkelser er sårbare overfor den måte spørsmålene er formulert på, og, som ved alle spørreundersøkelser, avhenger resultatene i stor grad av respondentenes villighet til å utlevere til dels sensitiv informasjon. Et annet problem er at det ofte er vanskelig å sammenligne resultater fra spørreundersøkelser over landegrensler da det er store forskjeller i hvordan undersøkelsene er lagt opp. Spørreundersøkelser er generelt lite egnet til å beregne presise estimater på unndratt beløp, og unndragelser som andel av BNP. Men selv om det er lett å bomme på nivået på unndragelsene kan spørreundersøkelser gi pålitelig informasjon om endringer i utbredelsen av den svarte økonomien. Spørreundersøkelser har størst verdi dersom svarene fra respondentene kobles opp mot registerdata. Dette er gunstig både for å få tilleggsinformasjon på de som har svart, men også for å sjekke at de opplysninger respondentene har gitt om inntekt og skatt er riktige. En spørreundersøkelse om svart arbeid krever anonymitet for å få mest mulig korrekte svar, dette umuliggjør en slik registeroppkobling. En har derfor ikke muligheten til å sjekke om de privatøkonomiske opplysningene som innhentes er riktige.

For å minimere respondentenes uvillighet til å svare ble det gitt full forsikring om anonymitet ved utføringen av vår spørreundersøkelse. De som sa seg villig til å delta fikk utlevert et umerket spørreskjema, informasjonsbrev og en umerket svarkonvolutt. Denne framgangsmåten var den samme i alle tre årene undersøkelsen er gjennomført. En annen forholdsregel for å få korrekte svar var at det for alle tre år ble benyttet et privat institutt for innhenting av informasjon. Grunnen til dette var at de som sa seg villig til å svare ikke skulle ha en oppfatning om at det var staten som sto bak undersøkelsen, noe det kunne ha vært en fare for dersom vi hadde benyttet SSB eller Skattedirektoratet.

⁴ Isachsen og Strøm (1980)

⁵ Mogenstein, Kvist, Körmendi og Pedersen (1995). Her estimerte Mogenstein et al. at svart arbeid i Danmark utgjorde 2,7 % av BNP i 1989, 4,2 % av BNP i 1991 og 3,1 % av BNP i 1994.

⁶ Vogel (1974).

Det kan tenkes at personer med dårlige norskkunnskaper vegrer seg mot å svare på undersøkelser av denne typen. Dersom dette er tilfelle, og det er slik at utbredelsen av svart arbeid er større i grupper som diskrimineres i det registrerte arbeidsmarkedet, kan det hende vi bommer på våre anslag. Det at spørreskjemaet kun er besvart av norske borgere gjør at vi heller ikke har mulighet til å sjekke om andelen av utenlandsk arbeidskraft som arbeider svart i Norge har endret seg. Da vi lever i en åpnere verden i 2001 enn hva tilfelle var i 1980 kan det tenke seg at tilbudet av svart arbeidskraft fra utlandet er høyere i 2001.

4. Beskrivende statistikk.

Hva sier så datamaterialet fra vår spørreundersøkelse om utbredelsen av svart arbeid i Norge? I spørreskjemaene har vi stilt en rekke ulike spørsmål som gir oss flere ulike mål på utbredelsen og omfanget av svart arbeid: Samlede svarte arbeidsinntekter, samlede svarte inntekter, andelen som har arbeidet svart siste året, andelen som noen gang har jobbet svart og andelen som ville jobbet svart dersom de fikk muligheten til det. En grunn til å ha så mange ulike mål på omfang og utbredelse er at vi vil redusere noe av usikkerheten som skyldes ulike tolkninger av enkeltspørsmål når vi spør om det samme på litt forskjellige måter. En annen grunn er at vi er interessert i ulike sider ved omfanget og utbredelsen av svart arbeid for å kunne forstå mer av det svarte arbeidsmarkedet. Jeg skal primært konsentrere meg om utbredelsen målt ved følgende tre spørsmål:

1. *Har du noen gang hatt inntekt som du ikke har oppgitt til ligningsmyndighetene?*
2. *Hvis du hadde mulighet for å få inntekt uten å måtte oppgi den til ligningsmyndighetene, ville du ta imot en slik inntekt?*
3. *Har du i løpet av siste 12 mnd utført et arbeid hvor inntektene av dette ikke er eller vil bli oppgitt til skattemyndighetene?*

Tabellene nedenfor viser svarfordelingene på disse spørsmålene.

Tabell 2: Har du noen gang hatt inntekt som du ikke har oppgitt til ligningsmyndighetene ? Prosent av hele utvalget.

1980	1989	2001
------	------	------

Ja	39	44	33
Nei	60	53	68
Ubesvart	1	3	0
Sum	100	100	100

Fra tabell 2 ser vi at andelen som noen gang har unndratt inntekt fra beskatning gikk opp fra 39 prosent i 1980 til 44 prosent i 1989, for deretter å falle til 32 prosent i 2001, det vil si til et lavere nivå enn i 1980. Selv om reduksjonen fra 1980 til 2001 ikke er like dramatisk som fra 1989 til 2001, er den signifikant forskjellig fra null (signifikansnivå minst 95 %, dvs. feilsannsynlighet lavere enn 5 %).

Tabell 3: Hvis du hadde muligheten for å få inntekt uten å måtte oppgi denne til ligningsmyndighetene, ville du ta i mot en slik inntekt? Prosent av hele utvalget

	1980		1989	2001	
	Prosent	Antall	Prosent	Prosent	Antall
Ja	64	562	59	44	432
Nei	18	158	20	26	255
Vet ikke	Ikke svaralternativ		19	26	255
Ubesvart	18	157	1	3	31
Sum	100	877	100	100	973

Fra tabell 3 ser en at det er en fallende trend fra 1980 til 2001 i antallet som svarer at de ville unndratt inntekt dersom de hadde muligheten til det.

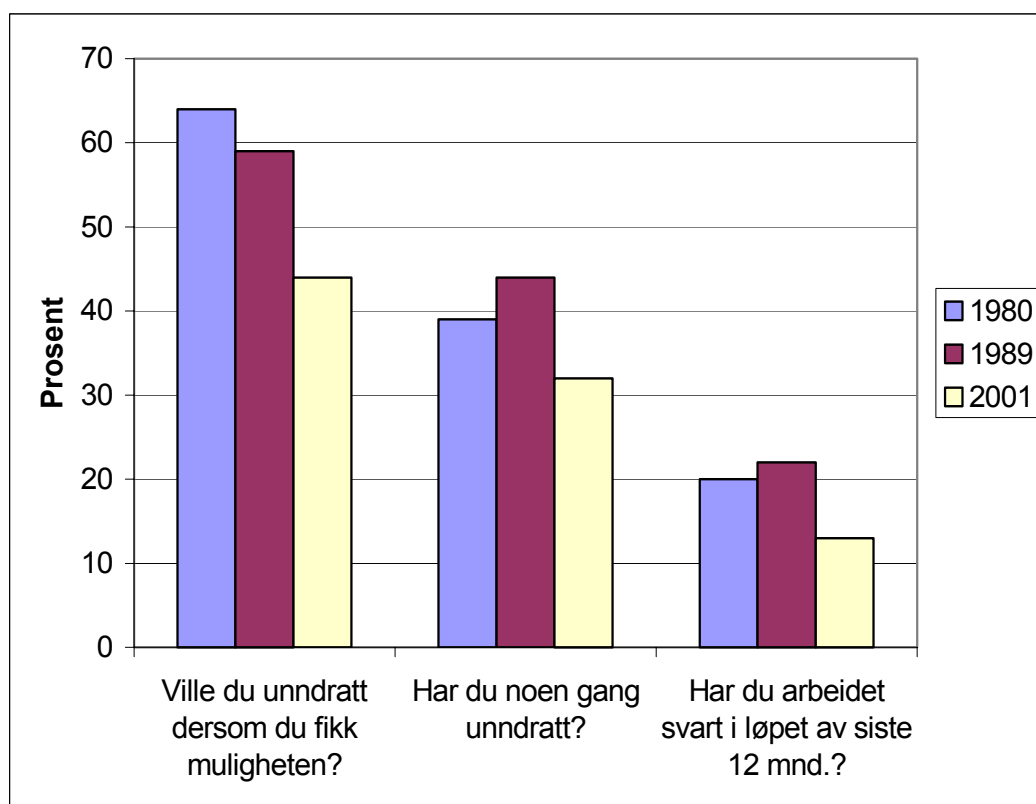
Tabell 4: Har du i løpet av siste 12 måneder utført et arbeid hvor inntekten av dette ikke er blitt eller vil bli oppgitt til skattemyndighetene? Prosent av hele utvalget.

	1980	1989	2001
Ja	20	22	13
Nei	76	75	87
Ubesvart	4	3	0

Sum	100	100	100
-----	-----	-----	-----

Vi ser av tabell 4 at andelen som svarer at de har unndratt inntekt fra beskatning siste 12 måneder går litt opp fra 1980 til 1989, fra 20 prosent til 22 prosent, for så å falle til 13 prosent i 2001, det vil si til et lavere nivå enn i 1980. Reduksjonen fra 1980 til 2001 er signifikant.

Figur 1. Utviklingen i utbredelse av svart arbeid.



I vårt datamateriale finner vi altså at utbredelsen av svart arbeid i Norge har falt fra 1980 til 2001 uavhengig av hvilket mål en bruker. I Goldstein et al (2002) er det gjort beregninger av samlede svarte arbeidsinntekter og samlede svarte inntekter.

For en nærmere statistisk drøfting av hvorvidt nedgangen i utbredelse av svart arbeid etter alle tre mål er signifikant henviser jeg til Goldstein m.fl.(2002)

Hvem er det så som jobber svart, og er det noen grupper som peker seg ut som mer aktive i det svarte arbeidsmarkedet enn andre? En første tilnærming til dette er å se på hvordan den gjennomsnittlige andelen som unndrar varierer med noen ulike kjennetegn.

Tabell 5. Prosentandelen som har jobbet svart siste 12 mnd. Kvinner og menn.

	1980	1989	2001
Kvinner	11	16	8
Menn	28	32	18

Tabell 5 viser at svart arbeid er mer utbredt blant menn enn hva det er blant kvinner i alle tre årene. Tabellen viser også at den fallende trenden fra 1980 til 2001, med en topp i 1989, gjelder uavhengig av kjønn.

Tabell 6. Prosentandelen som har jobbet svart siste 12 mnd. etter alder.

Alder	Unndratt 1980	Unndratt 2001
30år eller lavere	32	21
30 – 49år	19	17
50 år eller eldre	12	6

Fra tabell 6 ser vi at andelen som arbeider svart avtar med alder både i 2001 og i 1980. Andelen som unndrar har gått ned blant den eldste og yngste gruppen, mens et er liten endring for de mellom 30 og 49 år.

Dersom jeg grupperer etter alder og kjønn finner jeg at menn unndrar mer enn kvinner i alle aldersgrupper både i 1980 og i 2001.

Tabell 7. Prosentandel som har jobbet svart siste 12 mnd. etter alder og kjønn.

Alder	Kvinne 80	Kvinne 01	Mann 80	Mann 01
Under 30	12	11	47	33
30 – 49	13	9	24	28
Over 50	7	6	15	7

Fra tabell 7 ser vi at det har skjedd små endringer i andelen som unndrar i de ulike aldersgruppene. For menn har andelen som unndrar gått ned i den eldste og den yngste

aldersgruppen, mens det kun er en liten endring for de mellom 30 og 49 år. Vi ser også at for menn er andelen som unndrar betydelig høyere for den yngste gruppen sammenlignet med den eldste. Dette fenomenet er ikke like sterkt for kvinnene. Fra tabellen kommer det også fram at det er blant mennene den største endringen i andelen som unndrar mellom 1980 og 2001 har funnet sted.

Hva så med utdanningsnivå, er det forskjeller i andelen som unndrar etter antall år utdanning?

Tabell 8. Prosentandelen som har jobbet svart siste 12 mnd, etter utdanningsnivå.

Utdanningsnivå	1980	2001
Folkeskolenivå	14	6
Ungdomsskole	26	12
Gymnasnivå	27	20
Universitetsnivå	15	10

Vi ser av tabell 8 at det ikke er noen klar tendens til at den gjennomsnittlige andelen som unndrar faller med utdanningsnivå. Fra tabellen er det også verdt å merke seg at det har vært en svak nedgang i andelen med universitetsutdannelse som unndrar inntekt fra beskatning, selv om det er langt flere som har en slik utdanning i 2001 enn det var i 1980, og selv om universitetsutdannede er en langt mindre homogen gruppe nå enn i 1980. I tabellen ser vi at det har vært en markant nedgang i andelen som unndrar blant de som har ungdomsskole som sin høyeste utdanning. Noe av grunnen til dette kan være at det er langt færre som avslutter skolegangen etter ungdomsskolen nå enn hva det var før, noe som innebærer at mange av de som tilhører denne gruppen i dag er eldre mennesker, og vi har sett at eldre unndrar langt mindre enn yngre. Dersom jeg krysser alder med utdanningsnivå støttes denne hypotesen da 56 % av de som svarte at ungdomsskolen var deres høyeste utdanning i 2001 er over 50 år mot 20 % i 1980.

For å finne om det er forskjell i utbredelse av svart arbeid mellom ulike bransjer har jeg krysset bransjetilhørigheten med svaret på spørsmålet om personen har arbeidet svart siste 12 måneder. Tabell 9 viser andelen som svarer at de har arbeidet svart i noen utvalgte bransjer.

Tabell 9. Prosentandelen som har jobbet svart siste 12 mnd i noen utvalgte bransjer.

Bransje	1980	2001
Primærnæring	20	18
Bygg og anlegg	65	48
Industri	26	23
Varehandel	18	10
Transport	36	13
Helse og sosial	7	9
Byråkrati / forvaltning ⁷	13	4
Annen tjenesteyting	22	21
Undervisning	15	11
Totalt	23	16

Vi ser av tabell 9 at det er store forskjeller mellom bransjer når det gjelder andelen som jobbet svart sist år. Prosenten som jobber svart har imidlertid falt i samtlige bransjer. I bygg og anlegg har den falt fra 65 % i 1980 til 48 % i 2001. Det har altså vært en relativt stor forbedring i bransjen, men fremdeles arbeider nesten halvparten av de spurte svart. I transportbransjen falt prosenten fra 36 i 1980 til 13 i 2001. Det relativt dramatiske fallet kan skyldes at denne kategorien også inneholder de som jobber innen post og telekommunikasjon, en sektor hvor det har vært store endringer de siste 20 år. Denne utviklingen kan ha bidratt til at det er vanskeligere å arbeide svart i denne bransjen nå enn tidligere.

Vi har sett fra disse tabellene at utbredelsen av svart arbeid har falt fra 1980 til 2001 etter alle mål. Vi har også sett at det er flere menn enn kvinner, samt flere yngre enn eldre som jobber svart. Tabell 9 viste at det er store variasjoner i utbredelsen av skatteunndragelse mellom ulike bransjer. Jeg har altså identifisert en del kjennetegn ved personene som ser ut til å påvirke deres valg av svart arbeid eller ikke. Slike tabeller kan derimot ikke si oss hvilken betydning de enkelte kjennetegnene isolert sett har på dette valget. Årsaken til dette ligger i at når en ser på gjennomsnittlige andeler som unndrar etter ulike kjennetegn tar en

⁷ Denne kategorien inneholder de som har krysset av for alternativene bank, advokat, politi/forsvar og offentlig administrasjon.

ikke hensyn til at andre kjennetegn kan variere. Gjennom slike frekvenstabeller får en heller ikke tatt hensyn til at noen kjennetegn som påvirker en persons valg er observerbare, mens andre slike kjennetegn ikke er det. Av disse grunnene har jeg spesifisert en valgmodell som jeg vil presentere i neste kapittel.

5. Valgmodellen⁸

Porteføljevalgmodellen for skatteunndragelse i Allingham og Sandmo (1972)⁹ er et pionerarbeid innen modellering av det svarte arbeidsmarkedet. Denne modellen tar utgangspunkt i at svart arbeid er et resultat av et kontinuerlig valg over antall svarte timer en person ønsker å arbeide. Denne personen kan sette sammen en portefølje av hvite og svarte arbeidstimer fra alt hvitt og ingenting svart til alt svart og ingenting hvit. I modellen til Allingham og Sandmo velger altså skatteyteren hvor stor del av sin inntekt han vil oppgi til ligningsmyndighetene. Gevinsten ved å velge å oppgi en inntekt som er lavere enn den faktiske inntekten avhenger av hvorvidt personen blir tatt i skatteunndragelse eller ikke. Dersom han blir avslørt kommer han dårligere ut enn ved å velge å oppgi all sin inntekt til ligningsmyndighetene, hvis han ikke blir avslørt kommer han bedre ut. I følge Allingham og Sandmo velger skatteyteren den inntekt som maksimerer hans forventede nytte.

$$5.1) \quad E[U] = (1 - P)U(I - \theta X) + PU(I - \theta X - \pi(I - X)),$$

hvor θ er skattesatsen, X er oppgitt inntekt, I er faktisk inntekt og $P\pi$ er forventet straff. Første ledd i denne funksjonen er sannsynligheten for å ikke bli tatt i skatteunndragelse multiplisert med nytten av unndratt inntekt. Andre ledd på høyre side av likhetstegnet er sannsynligheten for å bli tatt for skatteunndragelse multiplisert med nytten av unndratt inntekt når denne er fratrukket straffeskatten som blir pålagt ved avsløring. Fra maksimeringsproblemet i 5.1 ser en at personen vil ønske å ha en høy andel av svarte timer i sin portefølje dersom sjansen for å bli tatt er liten, eller satsen på straffeskatten er lav. Differensiering av relasjon 5.1 med hensyn på oppgitt inntekt, når oppgitt inntekt er lik faktisk inntekt, gir at skatteyteren vil oppgi all sin inntekt dersom forventet straff er høyere

⁸ Modelleringen av veien fra nytteteori til estimering av sannsynligheter er i hovedsak basert på Train (1993)

enn skattesatsen. Dersom $\theta < P\pi$ får vi bestemt inntekten som oppgis som en funksjon av skattesatsen, sannsynligheten for oppdagelse og størrelsen på straffeskatten.¹⁰

Siden svart arbeid er en ulovlig/kriminell handling er det rimelig å tro at det å jobbe svart innebærer en moralsk kostnad som er uavhengig av hvor mye man jobber svart. Det vil si at en slik kostnad ikke nødvendigvis øker med antall svarte timer en person arbeider, men av valget svart arbeid eller ikke. Det er altså i realiteten to valg som avgjør mengden av svart arbeid:

1. Jobbe svart eller ikke.
2. Hvor mange timer svart arbeid.

Mens Allingham og Sandmo modellerer valget på trinn 2 ønsker jeg å modellere valget på trinn 1. Dette innebærer at det å arbeide svart er et resultat av et diskret valg hvor den involverte har to alternativer å velge mellom, skal han arbeide svart eller ikke. Dersom beslutningstakeren velger å jobbe svart er dette valget, i følge økonomisk teori, et resultat av at den maksimerte nytten ved å jobbe svart er større enn den maksimerte nytten ved å avstå fra slikt arbeid. Denne nytten varierer både med karakteristika ved de ulike personene og ved handlingen som personene velger. Den moralske kostnaden forbundet med svart arbeid inngår som individ karakteristika i denne avveiningen. Jeg skal nå sette opp en modell som viser veien fra nytteteorien til estimering av valgsannsynligheter.

Dersom vi betegner valget å jobbe svart med verdien 1 og valget å ikke jobbe svart med verdien 0, ser vi at nytten ved å arbeide svart er lik:

$$5.2) \quad U_n(Y_n=1 | X_n)$$

⁹ Framstillingen av Allingham og Sandmos modell er i hovedsak basert på Eide (2000). For en grundigere gjennomgang finnes hele modellen i *Journal of Public Finance* 1 (1972), eller i Atkinson, A.B. (red): *Modern Public Finance vol 1* (1991).

¹⁰ Modellen til Allingham og Sandmo er den første som anvender nobelprisvinner Gary Beckers artikkel fra 1968 "Crime and Punishment: An Economic Analysis" sammen med Arrows (1974) teori om adferd under usikkerhet, på skatteunndragelse. I følge Beckers teori vil kriminaliteten øke dersom gevinsten ved lovbrudd øker eller prisen for å begå lovbrudd avtar. Resultatene fra Allingham og Sandmo er i samsvar med dette.

Y_n betegner utfallet til responsvariabelen for person n , hvor Y kan ta verdien 0,1. X -vektoren er de sett av karakteristika som spiller inn på valget, og fotskriften n betegner personen, $n \in [1, N]$. Tilsvarende er nytten ved å ikke jobbe svart lik:

$$5.3) \quad U_n(Y_n=0 | X_n)$$

Denne X -vektoren kan splittes i to.

Z_{in} = vektor av alle relevante karakteristika ved handlingen som person n velger, her inngår økonomiske variable som svart og hvit lønn, marginals katt, straffeskatt og sjansen for å bli tatt¹¹. Z_{in} inneholder altså de variablene som spilte en avgjørende rolle i Allingham og Sandmos porteføljevalgmodell, men jeg oppfatter disse økonomiske variablene som endogene i estimeringen. Dette kommer jeg nærmere tilbake til senere i oppgaven.

Q_n = vektor av alle relevante karakteristika ved personen, her inngår variable som kjønn, alder og utdanning.

Dersom jeg lar fotskriften i betegne utfallet av valget, hvor $i = 0,1$, kan nytten skrives som:

$$5.4) \quad U_{in} = U(Z_{in}, Q_n)$$

Forutsatt nyttemaksimering velger da beslutningstakeren det alternativet som gir ham størst nytte. Svart arbeid blir utfallet kun dersom $U_{1n}^* > U_{0n}^*$, hvor $*$ indikerer at det er de maksimerte verdiene vi ser på.

Dersom størrelsen på nytten ved de ulike alternativene skal beregnes, må en kjenne alle komponentene i X -vektoren, det er ikke mulig her. En kan derfor tenke seg at vektorene Z_{in} og Q_n består av både observerbare og uobserverbare ledd. Jeg velger å betegne den observerbare delen av Z_{in} med W_{in} , og den observerbare delen av Q_n med L_n . Dersom vi dekomponerer relasjon 5.4 i to underfunksjoner, den observerbare og den uobserverbare, får vi følgende sammenheng:

$$5.5) \quad U_{in} = U(Z_{in}, Q_n) = V(W_{in}, L_n, \beta) + \epsilon_{in}$$

¹¹ Da formuleringen "sjansen for å bli tatt" er brukt i spørreskjemaet vil jeg gjennomgående holde meg til denne, selv om den mer presise formuleringen ville vært "risikoen for å bli tatt".

β symboliserer den ukjente parametervektoren som jeg skal estimere. Således består V funksjonen bare av observerbare faktorer. De uobserverbare faktorene i personens valg er symbolisert ved e_{in} vektoren. En ser at e_{in} vektoren er differansen mellom den sanne nytten og den observerbare delen av nytten.

Jeg definerer $P(S_{1n})$ som suksessannsynligheten, det vil si sannsynligheten for at individ n velger å arbeide svart, slik at:

$$\begin{aligned} 5.6) \quad P(S_{1n}) &= \text{Prob}(U_{1n} > U_{0n}) \\ &= \text{Prob}(V_{1n} + e_{1n} > V_{0n} + e_{0n}) \\ &= \text{Prob}(e_{0n} - e_{1n} < V_{1n} - V_{0n}) \end{aligned}$$

Ved å se nærmere på høyresiden av siste ledd i relasjon 5.6 ser vi at differansen mellom de to stokastiske variablene e_{0n} og e_{1n} selv er en stokastisk variabel som varierer over beslutningstakere med samme nyttekomponenter. Vi ser da at sannsynligheten for at personen velger utfall 1 er lik sannsynligheten for at den stokastiske differansen er mindre enn den observerbare differansen, noe som er identisk med en kumulativ fordelingsfunksjon.

Dersom en kjenner fordelingen til det stokastiske bidraget i relasjon 5.6 kan en beregne sannsynligheten for suksessutfallet. Diskrete valgmodeller spesifiserer en fordeling til de ukjente nyttekomponentene og utleder funksjoner for suksessannsynligheten. Ulike fordelinger gir opphav til ulike valgmodeller.

6. Logit-modellen¹²

Jeg ønsker å modellere sannsynligheten for at en person velger å arbeide svart. For å kunne gjøre dette trenger jeg en sannsynlighetsmodell som har to spesielle kjennetegn. (Gujarati 1995)

1. Når X_n øker, øker $P_n = E(Y = 1 \mid X)$, men overskrider ikke intervallet $[0,1]$.

¹² I dette kapitlet baserer jeg meg på Green (2000), Gujarati (1995), samt Goldstein m.fl. (2002).

2. Forholdet mellom P_n og X_n er ikke-lineært.

Blant de modellene som tilfredsstill disse kriteriene finner en Logit-modellen og Probit-modellen. Disse to modellene skiller seg fra hverandre ved at de tar utgangspunkt i forskjellige fordelinger på det stokastiske bidraget i relasjon 5.6 Ved å forutsette at den stokastiske komponenten er normalfordelt får vi en Probit-modell. Dersom det derimot forutsettes at e_{0n} og e_{1n} er ekstremverdifordelt, er differansen $e_{0n} - e_{1n}$ logistisk fordelt og vi har en Logit-modell.

I følge Green (2000) er det vanskelig å begrunne valget av den ene modellen framfor den andre da disse gir veldig likt resultat. Begge tetthetsfunksjonene er symmetriske rundt null, men tetthetsfunksjonen til standardnormalfordelingen har tettere konsentrasjon rundt null enn hva den logistiske fordelingen har.

Jeg har valgt å estimere valgsannsynligheten ved en Logit-modell. Denne Logit-modellen er spesifisert på følgende måte.

$$6.1) \quad Prob(Y=1/X) = P(S_n) = \frac{e^{X_i \beta}}{1 + e^{X_i \beta}}$$

Hvor X er en vektor av personspezifiske variable og β er den ukjente parametervektoren.

Estimering av logitmodellen baserer seg på "maximum likelihood" (ML) metoden. Ved å betrakte hver observasjon som en uavhengig trekning fra en Bernoulli fordeling, med suksessannsynlighet $P(S_n)$, blir likelihoodfunksjonen, for et utvalg på N observasjoner, gitt ved:

$$6.2) \quad L = \prod_{n=1}^N [p(S_n)]^{Y_n} [1 - p(S_n)]^{1 - Y_n}$$

Ved å ta logaritmen til likelihoodfunksjonen finner en fram til log-likelihoodfunksjonen.

$$6.3) \quad \text{Log } L = \sum_{n=1}^N [Y_n \text{Log } P(S_n) + (1 - Y_n) \text{Log}[1 - P(S_n)]]$$

ML-estimatorene framkommer ved å maksimere log-likelihoodfunksjonen med hensyn på de ukjente parameterene. Betinget på at forklaringsvariablene er lineært uavhengige av hverandre vil førsteordensbetingelsene være tilstrekkelige betingelser for å kunne løse maksimeringsproblemet, da log-likelihoodfunksjonen er strengt konkav.

$$6.4) \quad \frac{\delta \text{Log} L}{\delta \beta} = \sum_{n=1}^N (Y_n - P(S_n)) x_n$$

I min modell er parametervektoren β estimert på data fra henholdsvis 1980 og 2001. Dersom jeg lar N_s være antall personer som har sagt at de jobber svart og lar N_h være de som sier at de ikke har gjort det vil log-likelihoodfunksjonen ta følgende form.

$$6.5) \quad L = \sum_{n=1}^{N_s} \ln P(S_n) + \sum_{n=1}^{N_h} \ln(1 - P(S_n))$$

De maksimale verdiene av L og $\text{Log} L$ får vi ved å sette inn verdier for de β -er som maksimerer L . Vi kaller dem L_{\max} og $\text{Log} L_{\max}$.

En alternativ verdi på $\text{Log} L$ kan vi få ved å anta at de forklaringsvariable vi har trukket inn, ikke betyr noe i forklaringen av den svarte jobbsannsynligheten og at personene velger å jobbe svart rent tilfeldig, dvs kaste krone/mynt. I så fall vil den korresponderende simultane sannsynligheten være lik $L_0 = (0,5)^N$, hvor $N = N_s + N_h$. Logaritmen til denne simultane sannsynligheten kaller vi $\text{Log} L_0$.

To mål på hvor godt vår modell forklarer data i forhold til en modell hvor personene kaster kron/mynt om de skal jobbe svart eller ikke, er:

$$6.6) \quad G = \frac{[L_{\max}]^{1/N}}{[L_0]^{1/N}}$$

$$6.7) \quad \rho^2 = 1 - \frac{\text{Log} L_{\max}}{\text{Log} L_0}$$

Det første målet er forholdet mellom det geometriske gjennomsnittet av valgsannsynlighetene i vår modell når de ukjente parametrene, β , er bestemt slik at L er maksimal, og det geometriske gjennomsnittet av valgsannsynligheten når individet velger

helt tilfeldig. Dersom G er større enn 1, forklarer vår modell data bedre enn rent tilfeldige valg. Nedre verdi for G er 1.

Det andre målet kalles ofte McFadden's rho og dette målet varierer mellom 0 og 1, jo nærmere 1 det er, desto bedre forklarer vår modell data i forhold til en modell hvor individene velger rent tilfeldig.

Det er følgende sammenheng mellom G og ρ^2 :

$$6.8) \quad \rho^2 = \frac{N \ln G}{-L_0}$$

Tolkning av koeffisienter i den binære logit-modellen.

I vanlig lineær regresjon tolkes regresjonskoeffisientene som endringen i den avhengige variabelen når vi endrer den aktuelle uavhengige variabelen med en enhet. Når det gjelder den binære logit-modellen er tolkningen av regresjonskoeffisientene noe annerledes. For å illustrere dette kan det være nyttig å fremstille modellen på følgende måte;

$$6.9) \quad \text{Log} \left[\frac{P(Y_n = 1)}{P(Y_n = 0)} \right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K$$

Her ser vi at koeffisienten i logit-modellen kan tolkes som en endring i forholdet mellom sannsynligheten for suksess og sannsynligheten for ikke-suksess som følge av en enhets endring i den aktuelle uavhengige variabelen. Dette betegnes ofte som en endring i log-oddsen som følge av en enhets endring i den aktuelle koeffisienten.

Testing av lineære restriksjoner.

Jeg vil benytte meg av en Wald test for å teste lineære restriksjoner knyttet til de ukjente parametrene. For et tenkt tilfelle med nullhypotesen gitt ved $H_0 : \beta_i = 0$, blir Wald-testobservatoren, W , gitt ved:

$$6.10) \quad W = \frac{\hat{\beta}_i}{\text{Est.Asy.Var}(\hat{\beta}_i)} \sim \text{Asy.}\chi^2(1)$$

Nevneren i uttrykket for testobservatoren er den estimerte asymptotiske variansen til ML-estimatoren. Wald-observatoren er asymptotisk Chi-kvadratfordelt med en frihetsgrad under H_0 , (Green 2000). Generelt er antall frihetsgrader lik antall restriksjoner som testes simultant.

7. Estimering.

I de følgende kapitlene skal jeg estimere sannsynligheten for at en person har arbeidet svart siste 12 måneder gjennom to modeller. I den første estimeringen ser jeg på sannsynligheten for svart arbeid uavhengig av bransje, mens jeg tar hensyn til personens valg av bransje i modell nummer to. Disse to første modellene tar utgangspunkt i følgende spørsmål i spørreskjemaet:

”Har du i løpet av siste 12 mnd utført et arbeid hvor inntektene av dette ikke er eller vil bli oppgitt til skattemyndighetene?”

Deretter vil jeg estimere sannsynligheten for at en person ønsker å unndra inntekt fra beskatning, hvis det er mulig. I denne modellen tar jeg ikke hensyn til personens valg av bransje. Spørsmålet som ligger til grunn for denne estimeringen er:

”Hvis du hadde mulighet for å få inntekt uten å måtte oppgi den til ligningsmyndighetene, ville du ta imot en slik inntekt?”

Jeg vil estimere disse modellene på data både fra 1980 og 2001 for å kunne sammenligne svarene de gir over tid.

Alle disse tre modellene er estimert med de samme basisvariablene.

- Kjønn
- Alder
- Utdanning
- Skattemoral

- Sjansen for å bli tatt

Nedenfor følger en kort diskusjon av hvorfor jeg har valgt å inkludere disse variablene i analysen, mens jeg har utelatt andre.

Kjønn, alder og utdanning.

I kapitlet med beskrivende statistikk så vi at utbredelsen av svart arbeid varierte med kjennetegn som kjønn, alder og utdanning. Det er naturlig å ha disse individkjennetegnene med i analysen for å kunne se hvilken betydning det enkelte kjennetegn har isolert sett. Dette er variable som kan fange opp noe av forskjellene i muligheten til svart arbeid. Det er ikke nødvendigvis slik at en kvinne arbeider mindre svart enn en mann fordi hun er kvinne, men fordi hun har mindre mulighet til svart arbeid enn hva en mann har. Personkjennetegn som kjønn, alder og utdanning kan si noe om muligheten for svart arbeid over disse kjennetegnene.

Skattemoral.

I spørreskjemaet blir respondenten bedt om å ta stilling til fire påstander om skatt og skatteunndragelse.

1. *"Skattetrykket gjør at folk ikke er interessert i å jobbe overtid"*
2. *"Dagens skattenivå er nødvendig for å ha en velferdsstat som år"*
3. *"Det er forståelig at folk gjerne utfører arbeid mot skattefri betaling"*
4. *"Skatteunndragelse bør straffes hardt"*

For hver av disse påstandene er det fire svaralternativer: "Enig", "uenig", "hverken enig eller uenig" og "vet ikke". Det spørsmålet som er mest egnet til å måle folks skattemoral er nr.3 av de som er listet opp over. Svaret på påstanden: *"Det er forståelig at folk gjerne utfører arbeid mot skattefri betaling"*, kan oppfattes som en pekepinn på den moralske kostnaden ved å arbeide svart. Jeg har valgt å kalle denne variabelen for *"norm"*. Denne normvariabelen kan oppfattes som endogen i analysen, med dette mener jeg at en persons svar på spørsmål om holdninger til skatteunndragelse kan være farget av hvorvidt denne personen arbeider svart eller ikke. Dersom det er slik at det valget du gjør, av svart arbeid eller ikke, påvirker din skattemoral så er normvariabelen endogen i analysen. Dersom det er slik at din skattemoral påvirker det valget du gjør er ikke denne variabelen endogen. I utgangspunktet vil jeg behandle denne variabelen som eksogen, vil komme nærmere inn på dette i selve estimeringen.

Sjansen for å bli tatt.

Spørreskjemaet gir også opplysninger om respondentens subjektive oppfatning av hvor stor sjansen er for å bli tatt dersom en unnlater å oppgi en inntekt til ligningsmyndighetene.

Denne subjektive oppfatningen har jeg tatt med som et argument i modellene. Spørsmålet denne variabelen bygger på er formulert på følgende måte:

”Hvor stor tror du sjansen er for å bli tatt av ligningsmyndighetene dersom du ikke oppgir beløpet”

Ordet: *”beløpet”* i dette spørsmålet henviser til formuleringen i spørsmålet foran, og viser til en inntekt uten spesifisert størrelse. Svaralternativene på spørsmålet om sjansen for å bli tatt for skatteunndragelse er:

- Blir helt sikkert tatt (1)
- Blir ganske sikkert tatt (0,75)
- Blir kanskje tatt (0,5)
- Blir ganske sikkert ikke tatt (0,25)
- Blir helt sikkert ikke tatt. (0)

Tallene i parentes figurerer ikke i spørreskjemaet, men er vår numeriske tolkning av svaralternativene. Det er ikke sikkert de ulike respondentene har samme tolkning som oss. Det er heller ikke sikkert alle respondentene oppfatter spørsmålet på samme måte, noen kan tolke formuleringen til å gjelde for én periode, mens andre kan tolke det til å gjelde for mange perioder. Svarene vi får kan variere mellom å gjelde sannsynligheten for å bli tatt dersom det unnlates å oppgi en inntekt til ligningsmyndighetene én gang, og sannsynligheten for å bli tatt i unndragelse dersom man unndrar inntekt over en lengre periode. Da dette spørsmålet er formulert på samme måte i 1980 og 2001 antar jeg at dersom respondentene har hatt ulike oppfatninger om hvilken tidshorisont spørsmålet dreier seg om så er dette likt fordelt i de to årene, og da vi er mest opptatt av endringer i utbredelsen, trenger det ikke bety så mye. De fleste artikler og modeller som omhandler skatteunndragelse, og kriminaløkonomi generelt, har som en del av sin konklusjon at omfanget av den kriminelle atferden avtar dersom prisen en betaler for handlingen øker. Jeg har inkludert den subjektive sannsynligheten for å bli tatt i skatteunndragelse i modellen for å undersøke om dette stemmer for våre data.

Bransje

Jeg har modellert sannsynligheten for svart arbeid siste 12 måneder både avhengig og uavhengig av bransje. Grunnen til at jeg har inkludert bransjen personen arbeider innen i modellen er at jeg ønsker å finne om bransjen i seg selv spiller en rolle i det svarte valget, eller, for eksempel, om det er slik at den høye andelen som arbeider svart i bygg og anleggsbransjen skyldes at det arbeider mange unge menn der, ikke selve bransjen.

Jeg har ikke inkludert økonomiske variable som marginalsatt og inntekt i estimeringen av sannsynlighetene. Grunnen til dette er at jeg oppfatter disse variablene som endogene i modellen. Dette innebærer at størrelsen på en persons oppgitte inntekt ikke er uavhengig av hvorvidt han jobber svart eller ikke. Ex ante er det grunn til å tro at det er en korrelasjon mellom lav oppgitt inntekt og svart arbeid. Da marginalsatten avhenger av oppgitt inntekt, er heller ikke denne uavhengig av om personen arbeider svart eller ikke. Dersom en ikke tar hensyn til dette problemet i estimeringen, kan en lett finne en falsk negativ sammenheng mellom marginalsatt og svart arbeid. Disse økonomiske variablene er med som elementer i den maksimerte nyttefunksjonen i valgmodellen, men ikke i estimeringen. Jeg sitter da igjen med ligningene på redusert form. En forklaring på dette er at i en generell likevektsmodell for hele økonomien bestemmes faktorprisene og sysselsettingen i likevekt. Det eneste som skifter tilpasningen til folk i modellen er da eksogene endringer. Jeg ønsker å modellere personers sannsynlighet for svart arbeid gitt disse eksogene størrelsene som kjønn, alder, utdanning, skattemoral og sjansen for oppdagelse.

8. Logit-sannsynligheter for svart arbeid siste 12 måneder uavhengig av bransje.

I denne estimeringen har jeg fjernet alle over 60 år fra utvalget. Grunnen til dette er at jeg i analysen primært er interessert i de yrkesaktive. Det er også noen manglende observasjoner av kjennetegn for enkelte personer, disse har da falt ut av analysen. Svarfordelingen til utvalget jeg da sitter igjen med på spørsmålet om de har arbeidet svart siste 12 måneder er gjengitt i tabellen nedenfor.

Tabell 10. "Har du i løpet av siste 12 mnd utført et arbeid hvor inntektene av dette ikke er eller vil bli oppgitt til skattemyndighetene?"

	1980	2001
Ja	23	16
Nei	77	84
Sum	100	100
Antall svar på dette spm.	709	720

I tabell 10 kan en se at den gjennomsnittlige sannsynligheten for å ha arbeidet svart siste 12 måneder i utvalget var 0,23 i 1980 og 0,16 i 2001.

Jeg har tidligere definert sannsynligheten for å arbeide svart som følgende Logit-sannsynlighet:

$$8.1) \quad \text{Prob}(Y=1 | X) = P(S_n) = \frac{e^{x_i\beta}}{1 + e^{x_i\beta}}$$

Her er x en vektor av observerte variable som antas å påvirke sannsynligheten for at et individ jobber svart. I denne første tilnærmingen til sannsynligheten for å arbeide svart uavhengig av bransjetilknytning består denne vektoren av følgende variable. (tilhørende koeffisienter er angitt i parentes)

- Konstantledd (β_0),
- Kjønn = 1 dersom personen er mann, 0 ellers (β_k),
- Alder, målt i antall år, (β_a),
- Utdanning = 1 dersom personen har universitetsutdannelse, 0 ellers (β_u),
- Norm = 1 dersom personen er enig i følgende påstand: "Det er forståelig at folk gjerne utfører arbeid mot skattefri betaling", 0 ellers, (β_n),
- Sjanse = 1 hvis en svarer at sjansen for å bli tatt for skatteunndragelse er ≥ 0.75 ¹³, 0 ellers (β_s).

Resultatet av estimeringen er vist i tabellen nedenfor.

¹³ Dette tilsvarer svaralternativene "helt sikkert tatt" og "ganske sikkert tatt" jf. diskusjonen av denne variabelen tidligere.

Tabell 11. Estimeringsresultater¹⁴. Sannsynligheten for å ha jobbet svart de siste 12 måneder.

VARIABLE	1980		2001	
	Estimerer	Standardfeil	Estimerer	Standardfeil
Konstant	-0,8184	0,4061	-2.1105	0,4605
Kjønn	1.2546	0,2110	1.2170	0,2342
Alder	-0,0411	0,00908	-0,0259	0,00940
Utdanning	-0,9194	0,2624	-0,5883	0,2416
Norm	0,5449	0,2625	1.3145	0,7352
Sjansse	-1.2910	0,6224	-1.7225	0,2908
LogL	-334.66		-267.69	
G	1.247		1.379	
ρ^2	0,32		0,46	
N	709		720	

Fra tabell 11 kan en se at vi forklarer data noe bedre ved vår modell i 2001 enn i 1980. Bedømt ut fra G-målet, ser vi at modellen for 2001 forklarer dataene 37.9 % bedre enn en modell hvor valgene forutsettes å skje rent tilfeldig. Koeffisientene er alle skarpt bestemt. Normvariabelen har ikke bare økt i betydning fra 1980 til 2001, den blir også klart skarpere bestemt i 2001 enn i 1980.

Fra estimeringsresultatene kan en også se at koeffisientene er forbausende like dersom en tar hensyn til usikkerheten i anslagene.

En metode for å illustrere virkningen av de forskjellige koeffisientenes størrelse og retning på sannsynligheten for å arbeide svart, er å ta utgangspunkt i en referanseperson. Man tilegner så denne personen kjennetegn etter hvilke variable en har i analysen. Deretter endrer jeg ett og ett kjennetegn for å se hvordan dette virker inn på den svarte valgsannsynligheten. Jeg skal presentere to slike referansepersoner. Første referanseperson er en kvinne på 40 år med høy utdanning, som synes det er forståelig at folk gjerne utfører arbeid mot skattefri betaling og som anser at sjansen for å bli tatt for skatteunndragelse er lavere enn 75 %. Jeg

¹⁴ Alle estimater er signifikant forskjellige fra 0.

kan ved hjelp av estimatene gitt i tabell 11 beregne en sannsynlighet for referansepersonen arbeidet svart siste 12 måneder for de to aktuelle årstallene.

Tabell 12. Sannsynligheten for svart arbeid siste 12 måneder for 1980 og 2001.

	Referanse- person	Kjønn	Alder	Utdan.	Norm	Sjans
Kjennetegn						
Kjønn	Kvinne	Mann	✓	✓	✓	✓
Alder	40	✓	30	✓	✓	✓
Utdanning	Høy	✓	✓	Lav	✓	✓
Norm	Svart	✓	✓	✓	Hvit	✓
Sjansen	Liten	✓	✓	✓	✓	Stor
Valgsannsynlighet						
Svart valgsannsynlighet						
i prosent for 1980	6	17	8	13	3	2
Svart valgsannsynlighet						
i prosent for 2001	8	23	10	14	2	2

I tabellen indikerer ✓ at vi holder det aktuelle kjennetegnet uendret i forhold til referansen.

Som ventet ser vi at den svarte valgsannsynligheten stiger i begge år dersom vi endrer kjønn på referansepersonen fra kvinne til mann. Det er også en lavere sannsynlighet for at personen har arbeidet svart siste 12 måneder dersom hun har en oppfatning om at det er stor sannsynlighet for å bli tatt for skatteunndragelse. Fra tabellen ser en også at norm variabelen har en betydelig dempende effekt i begge år. Den svarte valgsannsynligheten er høyere jo lavere alder personen har og den er lavere for personer med universitetsutdannelse enn for de uten. Hvis en sammenligner sannsynligheten for å arbeide svart mellom de to årstallene ser en at de er relativt like over kjennetegnene.

Motstykket til den høyt utdannede middelaldrende kvinnen vist i tabell 12 kan man si er en lavt utdannet ung mann. Dersom jeg konstruerer en mannlig referanseperson på 25 år med lav utdannelse, som synes det er forståelig at folk gjerne utfører arbeid mot skattefri betaling

og som antar at sannsynligheten for å bli tatt for skatteunndragelse er lavere enn 75 %, får en et innblikk i en helt annen verden enn hva den middelaldrende kvinnen representerer.

Tabell 13. Sannsynligheten for svart arbeid i 1980 og 2001.

	Referanse- person	Kjønn	Alder	Utdan.	Norm	Sjansse
Kjennetegn						
Kjønn	Mann	Kvinne	✓	✓	✓	✓
Alder	25	✓	40	✓	✓	✓
Utdanning	Lav	✓	✓	Høy	✓	✓
Norm	Svart	✓	✓	✓	Hvit	✓
Sjansen	Liten	✓	✓	✓	✓	Stor
Valgsannsynlighet						
Svart valgsannsynlighet						
i prosent for 1980	49	21	34	28	36	21
Svart valgsannsynlighet						
i prosent for 2001	44	19	35	31	18	12

I tabellen indikerer ✓ at vi holder det aktuelle kjennetegnet uendret i forhold til referansen.

Fra tabell 13 ser vi at sannsynligheten for at den unge mannen skal ha arbeidet svart siste 12 måneder er ganske lik for 1980 og 2001. I 1980 var den svarte valgsannsynligheten i prosent ca 49 for at den mannlige referansepersonen valgte å arbeide svart, mot 44 i 2001, tall som er langt mer dramatiske enn de presentert i tabell 12. Både sjansen for å bli tatt og normen har også her en betydelig dempende virkning på tilbøyeligheten til å jobbe svart, og en ser også at denne dempende virkningen er sterkere i 2001 enn i 1980.

Av tabell 12 og 13 ser vi at sannsynligheten for å ha arbeidet svart siste 12 måneder er høyere i 2001 enn hva den er for 1980. Grunnen til dette er at referansepersonene har blitt tilegnet en svart norm. Fra tabell 11 ser vi at normvariabelen har en langt mer dempende effekt på den svarte sannsynligheten i 2001 enn hva den har for 1980. Dette innebærer at når jeg gir referansepersonene en svart norm så slår dette langt sterkere ut i 2001 enn hva det gjør i 1980. Dette bidrar dermed til at sannsynligheten for å ha arbeidet svart siste 12 måneder for et gitt sett av kjennetegn for en del tilfeller er høyere i 2001 enn hva de er i

1980, selv om vi har sett fra tabell 10 at den gjennomsnittlige sannsynligheten for svart arbeid er lavere i 2001 enn i 1980.

Det må understrekes at de svarte valgsannsynlighetene i tabell 12 og 13 er estimater og derigjennom heftet med stor usikkerhet¹⁵.

Dersom jeg kjører modellen beskrevet i dette kapitlet, men denne gangen uten å inkludere normvariabelen, har dette tilnærmet ingen effekt på estimatene til de andre variablene i analysen. Ut i fra dette kan jeg ikke konkludere med at normvariabelen er eksogen i analysen, men det viser at normvariabelen har en selvstendig effekt på sannsynligheten for svart arbeid.

Dekomponering av endringen i den svarte jobsannsynligheten fra 1980 til 2001.

Jeg ønsker nå å se nærmere på hvorvidt endringen i sannsynligheten for å arbeide svart fra 1980 til 2001 skyldes endringer i folks adferd eller endringer i sammensetningen av de individuelle kjennetegnene i befolkningen. For å se på hvordan endringer i adferd og endret sammensetning med hensyn til kjennetegn i befolkningen har påvirket sannsynligheten for å jobbe svart, har jeg dekomponert endringen i sannsynligheten for å ha arbeidet svart siste 12 måneder fra 1980 til 2001 i variabelendringer (kjennetegn) og koeffisientendringer (adferd). I følge tabell 10 var sannsynligheten for å ha arbeidet svart siste 12 måneder i gjennomsnitt over alle i logit-utvalget 0,16 i 2001 og 0,23 i 1980. Den gjennomsnittlige sannsynligheten for å jobbe svart har altså gått ned med 7 prosentpoeng. Dekomponeringen som jeg benytter, gir et anslag på hvor stor del av reduksjonen som skyldes endrete kjennetegn i befolkningen og hvor stor del som skyldes endring i adferd.

¹⁵ En metode for å få bukt med denne usikkerheten er å regne ut det asymptotiske standardavviket til den svarte valgsannsynligheten for å kunne konstruere et 95 % konfidensintervall rundt disse. Formelen for dette er gitt ved:

$$asy.Std.(P(S)) = \left. \frac{\partial P(S)}{\partial x\beta} \right|_{\beta=\hat{\beta}} \cdot Std.(x\hat{\beta}),$$

hvor $asy. var(x\hat{\beta}) = x' \hat{\epsilon}_{\beta} x$. Hvor x er en vektor med de aktuelle kjennetegnene ved personen, x' er da den transponerte vektoren.

$\hat{\epsilon}_{\beta} = COV(\hat{\beta})$. En framstilling av denne metoden er gitt i Green (2000) s.824.

Fra likning (8.1) ser vi at for et hvert individ kan den svarte jobbsannsynligheten $P(S)$ skrives som en funksjon av $X\beta$. Det betyr at for et individ i , kan jeg skrive endringen i sannsynligheten for å arbeide svart som:

$$\begin{aligned} (2) \quad P(S_{i,2001}) - P(S_{i,1980}) &= P(X_{i,2001}\beta_{2001}) - P(X_{i,1980}\beta_{1980}) \\ &= P(X_{i,2001}\beta_{2001}) - P(X_{i,1980}\beta_{1980}) \pm P(X_{i,2001}\beta_{1980}) \\ &= [P(X_{i,2001}\beta_{1980}) - P(X_{i,1980}\beta_{1980})] + [P(X_{i,2001}\beta_{2001}) - P(X_{i,2001}\beta_{1980})] \end{aligned}$$

På venstre side av uttrykket over finner en den totale endringen i sannsynligheten for å arbeide svart mellom 1980 og 2001. Den første hakeparentesen på høyre side av likhetstegnet viser endringene i sannsynligheten for å arbeide svart for et individ fra 1980 til 2001, dersom atferden uttrykt ved koeffisientvektoren β hadde vært den samme, men kjennetegnet ved individet har endret seg (alder, utdanning, normoppfatning osv). Denne første hakeparentesen viser altså variabelendringen mellom 1980 og 2001. Den andre hakeparentesen holder kjennetegnene ved individet konstant og lik hva de var i 2001, men koeffisienten, dvs atferden, endrer seg. Relasjonen ovenfor viser at den totale endringen i sannsynlighet for å arbeide svart, mellom 1980 og 2001, er lik endringen i individkjennetegn pluss endringen i folks atferd. Dette kan vi gjøre for hvert individ og ta gjennomsnittet over individene. I litteraturen er dekomponeringer av denne typen kjent som en Blinder–Oaxaca dekomponering¹⁶. Resultatet er vist i tabell 26.

Tabell 14. Dekomponering av endringer i sannsynligheten for å jobbe svart.

Total endring	Variabelendring	Koeffisientendring
-0,0687262	-0,0569	-0,01182
100 %	82.8 %	17.2 %

Vi ser altså at størstedelen av endringen i utbredelse av svart arbeid fra 1980 til 2001 skyldes endringer i variablene. Kun 17.2 % av endringen skyldes endring i folks adferd, når vi holder kjennetegnene ved personene konstant og lik verdien i 2001.

9. Logitsannsynligheter når jeg tar hensyn til individets valg av bransje.

¹⁶ For nærmere beskrivelse av denne metoden henviser jeg til Green (2000) s.252.

I dette kapitlet skal jeg ta hensyn til individets bransjetilhørighet i estimeringen av den svarte valgsannsynligheten. Det er tidligere vist at det er store forskjeller mellom bransjer når det gjelder andelen som jobbet svart sist år. Det jeg nå ønsker å undersøke, er hvorvidt et individs valg av bransje påvirker sannsynligheten for at personen jobber svart når vi holder fast de andre kjennetegnene, og om det å innføre variable for bransje endrer noe på viktigheten av de andre variablene. Det kunne for eksempel tenkes at grunnen til at det er så mange som arbeider svart i bygg og anlegg ikke så mye skyldes bransjen i seg selv, men at det er mange unge menn i bransjen. Ved å bringe et individs valg av bransje inn i modellen kan vi både se på effekten av bransjen alene, og effekten på de andre variablene av å innføre bransjevalg i analysen.

Denne utvidede logit-analysen har jeg spesifisert på følgende måte:

Variable:(med tilhørende koeffisienter i parentes)

- Konstantledd (β_0)
- Kjønn = 1 dersom personen er mann, 0 ellers.(β_k)
- Alder, målt i antall år.(β_a)
- Utdanning =1 hvis universitetsutdannelse, 0 ellers.(β_u)
- Norm: Det er forståelig at folk gjerne utfører arbeid mot skattefri betaling.
Norm = 1 dersom enig, 0 ellers.(β_n)
- Sjansse = 1 dersom sjansen for å bli tatt anses som større eller lik 0,75, 0 ellers.(β_s)

Dette er de samme basisvariablene som inngikk i logit-analysen i kapittel 8. Individets bransjevalg har jeg inkludert i estimeringen ved hjelp av dummyvariable. Disse dummyvariablene er spesifisert på følgende måte:

Dummyvariable for bransjer, gitt inntektsgivende arbeid:

- Primærnærings = 1 dersom personen jobber i denne bransjen, 0 ellers. (β_p)
- Industri = 1 dersom personen jobber innen industrien, 0 ellers.(β_i)
- Bygg og anlegg = 1 dersom personen er ansatt i denne næringen, 0 ellers.(β_b)
- Varehandel = 1 dersom personen arbeider innen varehandel, 0 ellers.(β_v)
- Transport = 1 dersom personen jobber innenfor samferdsel, transport, post eller tele,
0 ellers.(β_t)

- Undervisning = 1 dersom personen jobber innen undervisning eller forskning, 0 ellers. (β_{un})
- Helse = 1 dersom personen arbeider innen helse og sosialvesenet, = ellers. (β_h)
- Annen tjenesteyting = 1 dersom personen arbeider innenfor dette feltet, 0 ellers. (β_{tj})
- Byråkrati = 1 dersom personen er ansatt i offentlig administrasjon, i bank, som advokat, i politiet eller i forsvaret, 0 ellers. (β_{by})
- Intet yrke = 1 dersom personen ikke har et spesifikt yrke, inneholder også de av husmødrene i 1980 som hadde inntektsgivende arbeid, 0 ellers. (β_{intet})
- Annet = 1 dersom personen har et annet yrke enn alternativene i spørreskjemaet, i 1980 inneholder denne variabelen også de av studentene som har inntektsgivende arbeid, 0 ellers. (β_{an})
- Missing = 1 dersom personen har unnlatt å svare på dette spørsmålet gitt at han ikke har inntektsgivende arbeid, i 1980 inneholder denne variabelen også de av studentene/skoleelevene og husmødrene som ikke har inntektsgivende arbeid. Missing = 0 dersom individet ikke har disse kjennetegnene. (β_m)

Som referansevariabel til disse bransjedummyene har vi valgt bygg og anlegg.

Estimatene til denne utvidede analysen er vist i tabellen nedenfor.

Tabell 15. Estimer

Variable	1980		2001	
	Estimer	Standardavvik	Estimer	Standardavvik
Konstant	0,1104	0,5001	-0,7057	0,6409
Kjønn	1,2327	0,2412	1,1841	0,2661
Alder	-0,0406	0,00952	-0,0299	0,0100
Utdanning	-0,9340	0,3064	-0,6930	0,2755

Norm	0,5951	0,2710	1,3825	0,2992
Sjanse	-1,3901	0,6433	-1,6482	0,7385
Primærnærings	-0,9675	0,7544	-1,5235	0,7909
Industri	-1,0874	0,3377	-1,1137	0,4822
Varehandel	-1,3948	0,4265	-1,9724	0,5893
Transport	-0,9113	0,4540	-1,7635	0,6432
Undervisning	-0,8829	0,4939	-0,7444	0,6188
Helse	-1,5559	0,5501	-1,3508	0,5699
Annen tjenesteyting	-1,0893	0,4911	-0,8616	0,6199
Byråkrati	-1,6681	0,4361	-2,4466	0,8492
Intet yrke	-0,0480	0,9130	-0,7179	0,8351
Annet yrke	-0,4111	0,4560	-0,8799	0,5160
Missing	-1,0737	0,3875	-1,6662	0,4953
L	-321.9195		-255.728	
G	1.27		1.40	
ρ^2	0,35		0,49	
N	709		720	

Fra tabell 15 ser vi at både alder og utdanning har en svakere effekt på sannsynligheten for svart arbeid i 2001 enn i 1980, mens normvariabelen har en langt sterkere dempende effekt i 2001 enn i 1980. Tro på at det er liten sjanse for å bli tatt i skattefusk virker mer dempende på sannsynligheten for skatteunndragelse i 2001 enn hva det gjorde i 1980.

Fra standardavvikene kan en se at konstantleddene ikke er signifikant forskjellig fra 0 i noen av årene. Som i den tidligere presenterte modellen er også her alle basisvariablene skarpt bestemt. Modellen for 2001 forklarer sannsynligheten for svart arbeid noe bedre enn modellen for 1980, ut i fra G - målet ser vi at modellen for 2001 forklarer sannsynligheten for svart arbeid 40 % bedre enn en modell hvor valgene forutsettes å skje rent tilfeldig, mot 27 % bedre enn rent tilfeldig for modellen for 1980. Også ρ^2 - målet viser at vår modell for 2001 har størst forklaringskraft.

Grunnen til at estimatene for bransjene er negative, er i at jeg har valgt bygg og anlegg som referanse for bransje-dummyvariablene. Da bygg og anlegg er den bransjen som slår sterkest

ut på sannsynligheten for å arbeide svart siste 12 måneder, vil enhver annen bransje ha et negativt estimat i forhold til referansen. Dette innebærer ikke nødvendigvis at å jobbe i en av de andre bransjene virker negativt på sannsynligheten for å arbeide svart, men at det virker negativt på sannsynligheten for å arbeide svart i forhold til om personen arbeidet innen bygg og anlegg. Fra tabellen ser en at standardfeilene knyttet til estimatene til bransjedummyene er store. En må derfor være forsiktig med å si hvilken bransjetilhørighet som gir størst sannsynlighet for å jobbe svart. Tabellen viser imidlertid at arbeid innen sektorer som helse, byråkrati, varehandel og til dels industri gir mindre sjanser for svart arbeid enn i de øvrige sektorene.

Ved testing av lineære restriksjoner blant bransjevariablene i modellen for de to årstallene, det vil si test av:

H_0 : Primær = Industri = annentjeneste = varehandel = transport = undervisning = helse = missing = byråkrati = annet = intet = 0, mot

H_1 : minst en av disse forskjellig fra 0, får jeg følgende resultat.

Tabell 16. Test av lineære restriksjoner.

	Chi-square	DF	Pr > Chisq.
1980	25,2211	11	0,0085
2001	22,1254	11	0,0234

Tabell 16 viser at det laveste signifikansnivået H_0 kan forkastes på er henholdsvis 0,0085 for 1980 og 0,023 for 2001. Jeg kan forkaste H_0 , noe som innebærer at minst en av variablene for bransje har forklaringskraft. Utformingen av testobservatoren er diskutert i kapittel 6. Hva skjer så med de to referansepersonenes svarte valgsannsynlighet når jeg tar hensyn til deres bransjetilhørighet i analysen.

Tabell 17. Referanseperson.

	Referanse- person	Helse	Primær- næring	Vareh.	Uten inntekts- givende arbeid
Kjennetegn					
Kjønn	Kvinne	✓	✓	✓	✓

Alder	40	∨	∨	∨	∨
Utdanning	Høy	∨	Lav	Lav	Lav
Norm	Svart	∨	∨	∨	∨
Sjansen	Liten	∨	∨	∨	∨
Bransje	Undervisning	Helse	Primær- næring	Vareh.	Missing

Valgsannsynlighet

Svart valgsannsynlighet

i prosent for 1980

6 %	3 %	13 %	9 %	12 %
-----	-----	------	-----	------

Svart valgsannsynlighet

i prosent for 2001

12 %	7 %	11 %	8 %	10 %
------	-----	------	-----	------

I tabellen indikerer ∨ at vi holder det aktuelle kjennetegnet uendret i forhold til referansen

Tabell 17 viser at dersom vår kvinne på 40 år, med høy utdanning, som synes det er forståelig at folk unndrar inntekt fra beskatning og som har en oppfatning om at sjansen for å bli tatt for skatteunndragelse er liten, jobber innenfor undervisningssektoren er det 6 % sjanse for at hun arbeidet svart i 1980. Det tilsvarende tallet for 2001 er 12 %. Tabellen viser videre at dersom den samme kvinnen arbeidet innen helsevesnet faller den svarte valgsannsynligheten for begge år. Den svarte valgsannsynligheten for de tre siste kategoriene er regnet ut for den samme kvinnen, men gitt at hun har lav utdanning. Det er verdt å merke seg at det ikke ser ut til å være store forskjeller mellom 1980 og 2001 når det gjelder svart valgsannsynlighet innen de ulike bransjene, sett bort fra undervisningssektoren. Grunnen til at vi for enkelte bransjer får høyere tall for 2001 enn 1980, selv om den gjennomsnittlige sannsynligheten for svart arbeid er lavere i 2001 enn i 1980, er at normvariabelen slår sterkere ut i 2001. Noe som også var tilfelle for logit-analysen uavhengig av bransje.

Tabell 18 viser hvordan den svarte valgsannsynligheten til vår mannlige referanseperson varierer med bransjetilhørighet.

Tabell 18. Referanseperson.

Referanse- person	Industri	Høyt utdannet	Vareh.	Uten inntekts- givende arbeid
----------------------	----------	------------------	--------	----------------------------------

		byråkrat			
Kjennetegn					
Kjønn	Mann	∨	∨	∨	∨
Alder	25	∨	∨	∨	∨
Utdanning	Lav	∨	Høy	∨	∨
Norm	Svart	∨	∨	∨	∨
Sjansen	Liten	∨	∨	∨	∨
Bransje	Bygg	Industri	Byråkrati	Vareh.	Missing
Valgsannsynlighet					
Svart valgsannsynlighet					
i prosent for 1980	72	46	16	38	46
Svart valgsannsynlighet					
i prosent for 2001	75	50	12	30	37

I tabellen indikerer ∨ at vi holder det aktuelle kjennetegnet uendret i forhold til referansen.

Vi ser av tabell 18 at dersom referansemannen (25 år gammel, med lav utdanning osv.) jobber i bygge- og anleggsbransjen er sannsynligheten for at han jobber svart i 2001 oppe i 75 %. Dersom han i stedet arbeider i industrien, alt annet likt, faller sannsynligheten for svart arbeid til 50 %. Ved høy utdanning og jobb i byråkratiet synker den svarte valgsannsynligheten til 12 %. Variabelen Missing betyr at personen ikke hadde inntektsgivende arbeid som han skattet av siste 12 måneder. For en slik person er den svarte jobbsannsynligheten 37 %. Som i tabell 17 er det heller ikke her store forskjeller i svart valgsannsynlighet mellom de to årene i de ulike bransjene. Noen av forskjeller er imidlertid blitt forsterket i 2001. For eksempel er forskjellen mellom bygg og anleggsbransjen og varehandelen større når det gjelder sannsynligheten for å arbeide svart i 2001 enn i 1980. Dersom referansepersonen var en høyt utdannet byråkrat faller sannsynligheten for å arbeide svart med 56 prosentpoeng i 1980, mot et fall på 63 prosentpoeng i 2001. Sammenlignet med tabell 17 viser tabell 18 tall på den helt andre enden av skalaen. Det er for eksempel en forskjell i svart valgsannsynlighet på 68 prosentpoeng mellom referansemannen, gitt at han jobber i bygg og anleggsbransjen, og referansekvinnen, gitt at hun arbeider i helsevesenet, for 2001.

De svarte valgsannsynlighetene i tabell 17 og 18 er estimater og er derfor heftet med usikkerhet.

Dekomponering av logit-analysen med bransjevariable.

Jeg dekomponerer endringen i sannsynligheten for å arbeide svart i variabelendringer og koeffisientendringer for å finne hvor mye av endringen i utbredelse av svart arbeid mellom 1980 og 2001 som skyldes endrede kjennetegn i befolkningen og hvor mye som skyldes endring i folks adferd. Metoden jeg bruker her er den samme som vist tidligere.

Tabell 19. Dekomponering

Total endring	Variabel endring	Koeffisient endring
-0,0687271	-0,0574915	-0,0112356
100 %	83.65 %	16.35 %

Tabell 19 viser at når jeg tar hensyn til bransje i logit-analysen kan ca 84 % av endringen i utbredelse kan tilskrives endringer i befolkningens kjennetegn, mens ca 16 % av endringen skyldes endringer i folks adferd. Dette er tilnærmet likt det resultatet vi fant når vi dekomponerte endringen mellom 1980 og 2001 i logit-analysen uavhengig bransjetilknytning. Jeg kan derfor konkludere med at det å innføre bransjer i modellen ikke hadde noen innvirkning på hvor mye av endringen som skyldes endret adferd og hvor mye som skyldes endrede kjennetegn i befolkningen. Dette innebærer at fallet i utbredelsen av svart arbeid mellom 1980 og 2001 ikke kan tilskrives endringer i sysselsettingen mellom bransjer.

En metode for å undersøke hvorvidt det er endringer i bransjesammensetningen i befolkningen som er grunnen til at utbredelsen av svart arbeid har falt fra 1980 til 2001 er å tilegne sysselsettingsandelene til de ulike bransjene i 2001 andelen som unndrog i disse bransjene i 1980. Telleren i relasjon 9.1 er endringen i utbredelse av svart arbeid når vi endrer bransjesammensetning ved å tilegne sysselsettingsandelene i 2001 andelen som unndro i disse bransjene i 1980. Nevneren er endringen i andeler som unndrar mellom de to årene. Størrelsen $\frac{A}{B}$ skal symbolisere denne relative forholdet.

$$9.1) \quad \frac{A}{B} = \frac{\sum_{i=1}^m P_i^{80} \alpha_i^{01} - \sum_{i=1}^m P_i^{80} \alpha_i^{80}}{\sum_{i=1}^m P_i^{01} \alpha_i^{01} - \sum_{i=1}^m P_i^{80} \alpha_i^{80}}$$

Hvor:

P_i^t = andelen som unndrar i bransje i, år t.

α_i^t = sysselsettingsandelen for bransje i, år t.

$i = 1, 2, 3, 4, \dots, m$

$t = 80, 01.$

Dersom vi lar

u_i^t = antall som unndrar i bransje i, år t.

n_i^t = antall sysselsatte i bransje i, år t.

N = samlet antall sysselsatte, finner vi at

$$9.2) \quad \sum_{i=1}^m P_i^t \alpha_i^t = \sum_{i=1}^m \frac{u_i^t}{n_i^t} \times \frac{n_i^t}{N^t} = \frac{\sum_{i=1}^m u_i^t}{N^t} \equiv \lambda^t ; \text{ hvor}$$

λ^t = andelen som unndrar i år t

Uttrykket en da sitter igjen med er:

$$9.3) \quad \frac{A}{B} = \frac{\sum_{i=1}^m \frac{u_i^{80}}{n_i^{80}} \times \frac{n_i^{01}}{N^{01}} - \lambda^{80}}{\lambda^{01} - \lambda^{80}}$$

På våre data finner jeg at $\frac{A}{B} = 0,007193568$

Denne lave størrelsen på A/B innebærer at endringen i bransjesammensetningen mellom 1980 og 2001 har betydd relativt lite for utviklingen i den svarte økonomien mellom årstallene. Dersom A/B hadde vært lik 1 ville endringen i bransjesammensetningen i befolkningen stått for hele endringen i utbredelse av svart arbeid. Dersom A/B var lik 0 ville ikke noe av endringen i utbredelse skyldes endrede bransjesammensetninger. Konklusjonen fra dekomponeringen støttes altså av utregningen vist ovenfor.

10. Sannsynligheten for å jobbe svart, hvis en har mulighet til det

I kapitlet med beskrivende statistikk så vi at mens det bare er ca 13 prosent i 2001-utvalget som har jobbet svart siste år, er det ca 44 prosent som svarer ja på spørsmålet om de ville gjøre det hvis de hadde muligheten. For en del grupper er avviket enda større. En viktig grunn til at folk ikke unndrar er altså at de oppfatter det slik at de ikke har muligheten. I dette kapitlet skal jeg beregne en persons sannsynlighet for svart arbeid hvis han hadde muligheten til det.

I spørreundersøkelsen ble det stilt følgende spørsmål:

”Hvis du har mulighet for å få inntekt uten å måtte oppgi den til ligningsmyndighetene, vil du ta i mot en slik inntekt?”

Hensikten med dette spørsmålet er å kunne si noe om *potensialet* for svart arbeid i befolkningen, og i dette kapitlet skal jeg se på hvilke faktorer som påvirker potensialet. Svarfordelingen i 1980 og 2001 for utvalget i analysen er som vist i tabell 20. Også i denne analysen har jeg fjernet de over 60 år da jeg primært ønsker å analysere de yrkesaktive.

Tabell 20. Prosentandelen som ville jobbe svart hvis en har muligheter til det.

	1980		2001	
	Antall	Prosent	Antall	Prosent
Ja	493	68	343	48
Nei	110	15	173	24
Vet ikke/mangler svar	126	17	204	28

Sum	729	100	720	100
-----	-----	-----	-----	-----

Jeg ønsker å estimere sannsynligheten for at et individ ville ha arbeidet svart, hvis dette individet hadde muligheter til det. Som det fremgår av tabell 20 er det en betydelig andel som har svart ”vet ikke”, eventuelt ikke svarer. Jeg vil kalle disse ”vet ikke”- gruppen. Dette har jeg tatt hensyn til i estimeringen. Jeg har ikke kastet denne gruppen ut av utvalget, men foretatt 100 trekninger, hvor sannsynligheten for at en ”vet ikke” person skal komme i ja gruppen er lik den gjennomsnittlige sannsynligheten for ja, blant dem som har svart ja eller nei. Det betyr at i 1980 er det en sannsynlighet på 0,82 for at en ”vet ikke” person ville ha svart ja og 0,18 for at ”vet ikke” personen ville ha svart nei. De korresponderende sannsynlighetene for ”vet ikke” personer i 2001 er henholdsvis 0,63 for ja og 0,37 for nei. Dette betyr at svarene til ”vet ikke” gruppen er randomisert med 100 trekninger ut fra populasjonsfordelingen blant de som har svart ja eller nei. Dette innebærer at jeg har forutsatt at de som ikke svarer på dette spørsmålet, eller svarer at de ikke vet, har samme holdning til om de ville jobbe svart dersom dette var mulig som de som har svart ja eller nei. Denne randomiseringen er en av mange mulige metoder for å utnytte informasjonen i en ”vet ikke” -gruppe, og trolig heller ikke den beste. Ett alternativ ville ha vært og utnyttet informasjonen i ”missing-data” til å estimere sannsynligheter for svaralternativer ut fra denne informasjonen. Min fremgangsmåte her er noe røffere og resultatene må derfor tolkes med noen forsiktighet.

Modellen er spesifisert på følgende måte:

X er en vektor av observerte variable som antas å påvirke sannsynligheten for at et individ vil jobbe svart, dersom dette er mulig.

Vektoren (tilhørende koeffisienter i parentes) består av følgende variable:

- Konstantledd (β_0),
- Kjønn = 1 dersom personen er mann, 0 ellers (β_k),
- Alder, målt i antall år, (β_a).
- Utdanning = 1 dersom personen har universitetsutdanning, 0 ellers (β_u),
- Norm = 1 dersom personen er enig i følgende påstand: ”Det er forståelig at folk gjerne utfører arbeid mot skattefri betaling”, 0 ellers, (β_n),

- Sjanse =1 hvis en svarer at sjansen for å bli tatt for skatteunndragelse er ≥ 0.75 , 0 ellers (β_s).

Resultatet av estimeringen er gitt i tabellen nedenfor.

Tabell 21. Estimering. Sannsynligheten for å ville jobbe svart dersom det var mulig

Variable	1980		2001	
	Estimater	Standardfeil	Estimater	Standardfeil
Konstant	1,1064	0,2089	0,3161	0,2044
Kjønn	0,8770	0,1059	0,4057	0,0915
Alder	-0,0223	0,0048	-0,0132	0,0043
Utdanning	-0,9775	0,1265	-0,3077	0,1080
Norm	1,8108	0,1048	1,5128	0,0915
Sjanse	-1,3305	0,1543	-0,7625	0,1499
L		-279,008		-404,245
G		1,36		1,14
ρ^2		0,45		0,19
N		729		720

Fra tabell 21 kan en se at data forklares langt dårligere ved modellen i 2001 enn i 1980.

Bedømt ut fra G-målet, ser vi at i 2001 forklarer modellen data bare 14 % bedre enn en modell hvor valgene forutsettes å skje rent tilfeldig. Koeffisientene er skarpt bestemt, med unntak av koeffisienten knyttet til konstantleddet, alder og utdanning i 2001.

For å kunne illustrere virkningen av de forskjellige variablene på sannsynlighetene for ”svart arbeid hvis mulig” i de to årene, har jeg som her som tidligere tatt utgangspunkt i en to referansepersoner for de to aktuelle årstallene. Referansepersonene er de samme som tidligere. Det vil si to personer som synes det er forståelig at folk arbeider mot skattefri betaling og som antar at sjansen for å bli tatt for skatteunndragelse er lavere enn 75 %, hvor den ene personen er en høyt utdannet kvinne på 40 år og den andre en lavt utdannet mann på 25 år. Resultatene av denne illustrasjonen er vist i tabell 22 og 23.

Tabell 22. Referanseperson.

	Referanse- person	Kjønn	Alder	Utdan.	Norm	Sjansse
Kjennetegn						
Kjønn	Kvinne	Mann	✓	✓	✓	✓
Alder	40	✓	30	✓	✓	✓
Utdanning	Høy	✓	✓	Lav	✓	✓
Norm	Svart	✓	✓	✓	Hvit	✓
Sjansen	Liten	✓	✓	✓	✓	Stor
Valgsannsynlighet						
Svart valgsannsynlighet						
i prosent for 1980	74	87	78	88	32	43
Svart valgsannsynlighet						
i prosent for 2001	73	80	75	94	37	56

I tabellen indikerer ✓ at vi holder det aktuelle kjennetegnet uendret i forhold til referansen.

Vi ser fra tabell 22 at sannsynlighetene for å ville jobbe svart, dersom det var mulig, er svært høye. Men at de gjør et betydelig fall dersom personen har hvit norm eller en oppfatning om at det er stor sjansse for å bli tatt for skatteunndragelse. Dette indikerer at det er et stort potensial for svart arbeid i befolkningen.

Hvordan er det så med de svarte valgsannsynlighetene for referansemannen?

Tabell 23. Referanseperson.

	Referanse- person	Kjønn	Alder	Utdan.	Norm	Sjansse
Kjennetegn						
Kjønn	Mann	Kvinne	✓	✓	✓	✓
Alder	25	✓	40	✓	✓	✓
Utdanning	Lav	✓	✓	Høy	✓	✓
Norm	Svart	✓	✓	✓	Hvit	✓
Sjansen	Liten	✓	✓	✓	✓	Stor

Valgsannsynlighet

Svart valgsannsynlighet

i prosent for 1980	96	91	95	91	81	87
--------------------	----	----	----	----	----	----

Svart valgsannsynlighet

i prosent for 2001	87	82	85	83	60	76
--------------------	----	----	----	----	----	----

I tabellen indikerer √ at vi holder det aktuelle kjennetegnet uendret i forhold til referansen

Tabell 23 viser at de svarte valgsannsynlighetene gjennomgående er enda høyere i denne tabellen enn den forrige. Sannsynlighetene for å ville jobbe svart har gått ned fra 1980 til 2001 for referansemannen, og for de andre fem alternativene vi ser på. Det er en slående forskjell mellom sannsynligheten for å jobbe svart og sannsynligheten for å ville gjøre det hvis mulig. Mens kjønn, alder og utdanning hadde en klart dempende effekt på tilbøyeligheten til på å jobbe svart, finner vi ikke så stor effekt på sannsynligheten for å ville jobbe svart. Vi ser videre at det først og fremst er normvariabelen som virker dempende på sannsynligheten for å ville jobbe svart og at denne variabelen har en langt mer dempende effekt på referansemannens svarte valgsannsynlighet i 2001 enn hva den hadde i 1980. Dette fenomenet finner vi ikke hos referansekvinnen. Fra tabell 22 ser en at den dempende effekten av normvariabelen er tilnærmet like stor for begge årene. Grunnen til dette kan ligge i at referansekvinnen er tilegnet en høy utdanning, mens referansemannen er lavt utdannet. Dersom referansemannen tillegges en høy utdanning i tillegg til en hvit norm finner jeg at den svarte valgsannsynligheten i prosent er 61 for 1980 mot 52 for 2001, en differanse som harmonerer med de øvrige i tabell 23.

De svarte valgsannsynlighetene i tabell 22 og 23 er estimater og er derfor heftet med usikkerhet.

Dekomponering av endringen i sannsynligheten for å ville jobbe svart hvis muligheten var tilstede.

Den gjennomsnittlige sannsynligheten for å ville jobb svart var på ca 89 % i 1980 og på ca 73 % i 2001. Dekomponering av dette fallet på 16 prosentpoeng i en variabelendring og koeffisientendring er vist i tabellen nedenfor. Metoden her er den samme som ved tidligere dekomponeringer.

Tabell 24. Dekomponering

Total endring	Variabelendring	Koeffisientendring
-0,161489	-0,0898882	-0,0716008
100 %	55,7 %	44,3 %

Vi ser fra tabell 24 at koeffisientendringen, dvs. atferdsendringen, nå forklarer nær halvparten av reduksjonen i andelen som ville jobbe svart hvis mulig.

Estimater dersom jeg ikke randomiserer

Modellen vist i dette kapitlet er kjørt på et randomisert utvalg hvor jeg tilegnet de som ikke svarte på spørsmålet og de som svarte vet-ikke den gjennomsnittlige sannsynligheten for ja blant de som svarte ja eller nei.. Dersom jeg kjører samme modell, men denne gangen på bare de som svarer ja eller nei i tabell 20 finner jeg, som naturlig er, at modellens forklaringskraft øker for begge år. Koeffisientene blir i tillegg skarpere bestemt, men viser samme trend som i tabell 21.

Tabell 25 viser estimatene når jeg kjører modellen på dette mindre utvalget.

Tabell 25. Estimater når jeg ikke randomiserer.

Variable	1980		2001	
	Estimater	Standardfeil	Estimater	Standardfeil
Konstant	0,7046	0,5052	-0,0143	0,4322
Kjønn	1,2624	0,2717	0,4245	0,2217
Alder	-0,0272	0,0119	-0,0167	0,00942
Utdanning	-1,4390	0,3209	-0,3753	0,2281
Norm	2,5649	0,2755	2,3063	0,2238
Sjånse	-1,9144	0,4143	-1,3616	0,3418
LogL	-199,08		-258,14	

G	1,44	1,23
N	603	528

Dersom en sammenligner tabell 25 med tabell 21 ser en at normvariabelen for begge år har en langt sterkere effekt i tabell 25 enn hva den har i tabell 21. Sammenlignet med estimatene fra modellen kjørt på det randomiserte utvalget er effekten av kjønn sterkere i 1980, mens effekten av sjansevariabelen er sterkere i 2001. I tillegg ser det ut til at utdanning har en mindre dempende effekt i 1980 sammenlignet med estimatet fra tabell 21.

Det er vanskelig å trekke klare konklusjoner fra denne sammenligningen. En ting som kan antydes er at det virker som om effekten av normvariabelen er undervurdert i modellen kjørt på det randomiserte utvalget. Resultatene fra analysen av det begrensede utvalget, presentert i tabell 25, er trolig heller ikke helt korrekt da jeg utelater de som svarer vet-ikke/ubesvart. Det kan dermed tenkes at estimatet på normens effekt på sannsynligheten for svart arbeid dersom dette er mulig er overvurdert i tabell 25. Den sanne verdien på skattemoralens effekt vil da trolig ligge et sted i mellom estimatene fra tabell 21 og 25.

Problemet med analysen av ”svart arbeid hvis mulig”, er hvorvidt man kan behandle ubesvartgruppen i 1980 på samme måte som vet-ikke gruppen fra 2001. Fra tabell 20 så vi at vet-ikke ikke var noe svaralternativ i 1980. Dette kan ha bidratt til at flere har følt seg tvunget til å svare ja eller nei på dette spørsmålet. Dersom dette er tilfelle kan jeg ikke i like stor grad som i logit-analysene av svart arbeid sist år sammenligne resultatene fra analysen gjort på 1980 utvalget med analysen gjort på utvalget fra 2001.

11. Konklusjon

I denne hovedoppgaven har jeg forsøkt å forklare fallet i utbredelsen av svart arbeid mellom 1980 og 2001 ved bruk av logit-analyser. Resultatene fra disse analysene viser at sannsynligheten for svart arbeid er større blant menn enn blant kvinner. Yngre har større tilbøyelighet til skatteunndragelse enn eldre, og lavt utdannede har større sannsynlighet for å unndra inntekt enn hva høyt utdannede har. Videre er det slik at en lav skattemoral, samt en oppfatning om at det er liten sjanse for å bli tatt i skatteunndragelse medfører høyere sannsynlighet for svart arbeid. Logit-analysene viser også at den dempende effekten av en høy skattemoral er sterkere i 2001 enn hva den var i 1980. Det at en høy subjektiv

oppfatning om hvor stor sjansen er for å bli tatt for unndragelse virker dempende på sannsynligheten for svart arbeid er i samsvar med resultatene fra de fleste modellene som omhandler skatteunndragelse, herunder porteføljevalgmodellen til Allingham og Sandmo. Ved å sammenligne estimatene for de to årene og gjennom bruk av referansepersoner vises det at skattemoralen har blitt viktigere i forklaringen av utbredelsen av svart arbeid. Den bedre skattemoralen viser seg også ved at det er færre som synes at svart arbeid er forståelig i 2001 enn hva det var i 1980. Det er også færre som tror at svart arbeid er noe som kommer til å øke i omfang i årene som kommer, og færre som tror at skatteunndragelse er noe som kommer til å godtas mer og mer.

Når jeg tar hensyn til bransjevalg i logit-analysen ser vi at bransjer har en selvstendig effekt på sannsynligheten for svart arbeid, men at endringen i utbredelse ikke kan tilskrives endret bransjesammensetning i befolkningen. Dette ser vi både gjennom å sammenligne dekomponeringen i analysen som er avhengig av bransje med dekomponeringen i analysen som er uavhengig av bransje, og gjennom å tilegne sysselsettingsandelene til de ulike bransjene i 2001 andelen som unndrog i disse bransjene i 1980.

Det er flere faktorer som kan bidra til å forklare fallet i antallet som arbeider svart. For det første er det langt flere som har universitetsutdannelse i 2001 enn hva det var i 1980. Fra logit-analysene har vi sett at høy utdannelse virker dempende på sannsynligheten for unndragelse. Selv om denne dempende effekten er lavere i 2001 enn hva den var i 1980 så bidrar trolig det at flere har høy utdannelse nå enn tidligere til at utbredelsen av svart arbeid er lavere. For det andre har arbeidsstyrken blitt eldre. Dette bidrar også til at færre arbeider svart nå enn tidligere. Vi har sett at alder har en dempende effekt på sannsynligheten for svart arbeid, det at arbeidsstyrken har blitt eldre har da ført til at færre arbeider svart. For det tredje har kvinneandelen i det registrerte arbeidsmarkedet økt. Et av hovedresultatene fra undersøkelsene gjort i 1980 var at det var en lavere andel kvinner enn menn som jobbet svart, men de som gjorde det, jobbet i gjennomsnitt flere timer enn mennene¹⁷. Goldstein m.fl.(2002) viser at antall svarte timer blant kvinner har falt betraktelig. Dette har sannsynligvis sammenheng med at kvinneandelen i det registrerte yrkeslivet har økt siden 1980. For det fjerde har trolig en utvikling mot et mer profesjonalisert næringsliv bidratt til at færre arbeider svart nå enn tidligere. Vi har sett at endringen i unndragelse ikke kan

tilskrives en endret bransjesammensetning, men noe av grunnen ligger kanskje i endringer i selve bransjene. Det norske næringslivet har blitt betydelig mer profesjonalisert over de siste 20 år, det er færre små lokale bedrifter nå enn tidligere, samt at enkelte bransjer har hatt en intern opprydding. Denne endringen kan slå ut i enkelte variable i vår modellering. Det holdnings skapende arbeidet som er gjort i enkelte høyrisikobrancher kan for eksempel slå ut i en bedret skattemoral.

En kraftig reduksjon i marginalskatten for store grupper har trolig bidratt til å redusere gevinsten ved skatteunndragelse. Den maksimerte nytten av hvitt arbeid er høyere jo lavere marginalskatten er, for noen kan dette medfører at nytten av hvitt arbeid blir større enn nytten av svart arbeid. Reduksjonen i marginalskatt over perioden er et aspekt jeg ikke får med i estimeringen da økonomiske størrelser som svart og hvit inntekt, og skattesatser er endogene i modellen. Men det er imidlertid klart at endringen i skattesystemet har hatt en innvirkning på utbredelsen av svart arbeid i befolkningen. Dette er et tema som vi ikke har gått inn på i denne første delen av prosjektet om det svarte arbeidsmarkedet, men som det jobbes videre med i videreførelsen av prosjektet. Skattereduksjonen kan også ha bidratt til en bedring av skattemoralen, siden vi måler denne med om folk synes det er forståelig eller ikke at det unndras skatt.

De fleste av faktorene som har bidratt til at andelen som unndrar er lavere i 2001 enn hva den var i 1980 slår ut i at muligheten til skatteunndragelse har blitt mindre. Høy utdanning gir stillinger og posisjoner i samfunnet hvor muligheten for å unndra inntekt er mindre, det at yrkesdeltakelsen blant kvinner har økt gjør at færre har muligheten til å arbeide svart nå enn tidligere, og en strukturendring i det norske næringslivet har trolig også ført til at færre har muligheten til å arbeide svart.

Selv om det har vært en reduksjon i utbredelsen av svart arbeid er det fremdeles slik at en stor andel av utvalget ville unndratt dersom de hadde muligheten til det. Potensialet for skatteunndragelse er fremdeles ganske stort i samfunnet. Sammenligning av andelen som unndro i 1980, 1989 og 2001 viser at utbredelsen hadde en topp i 1989 før det falt ned til et nivå lavere enn i 1980. En av grunnene til at andelen som unndro i 1989 var høyere enn før og etter kan være at slutten av 80-tallet var en tid preget av økende arbeidsledighet og

¹⁷ Isachsen og Strøm (1981) s.91-93.

gjeldskriser. Dersom arbeidsledigheten på nytt skulle komme til å stige kan det hende en større del av det potensielle svarte arbeidet blir realisert. Hvor stor økning i utbredelse dette kan medføre er avhengig av hvilke yrkesgrupper arbeidsledigheten rammer i tillegg til antallet ledige. Men dersom arbeidsledigheten vokser i tiden som kommer er det grunn til å tro at utbredelsen av svart arbeid også vil vokse.

Referanser.

Allingham, M. and A. Sandmo (1972), *Income Tax Evasion: A Theoretical Analysis*, Journal of Public Economics, 1 (3/4), s.323-38.

Arrow, K (1974): *Essays in the Theory of Risk-Bearing*. Amsterdam, North-Holland.

Blumenthal, M, C. Christian and J. Slemrod (1998): *The Determinants of Income Tax Compliance: Evidence from a Controlled Experiment in Minnesota*. National Bureau of Economic Research, Working paper 6575.

Eide, E. (2000), *Oversikt over litteratur om svart arbeid og skatteunndragelse*, rapport 6/2000, Frischsenteret.

-
- Feige, E. L. (1989): *The underground economies: Tax evasion and information distortion*. Cambridge, Cambridge University Press, s. 13-56.
- Goldstein, H., W.G.Hansen, T.Ognedal og S.Strøm (2002): *Svart arbeid fra 1980 til 2001*, rapport 3/2002, Frischsenteret.
- Green, W. H. (2000) *Econometric analysis*, Prentice-Hall International Inc. (Fourth edition)
- Gujarati, D. N. (1995) *Basic Econometrics*, McGraw-Hill Book Co. (Third edition).
- Isachsen, A. J. and S. Strøm (1980): *The Hidden Economy: The Labour Market and Tax Evasion*. Scandinavian Journal of Economics, 82, s. 304-311
- Isachsen, A. J. and S. Strøm (1981), *Skattefritt, svart sektor i vekst*, Universitetsforlaget.
- Mogensen, G., H. Kvist, E. Körmendi and S. Pedersen (1995): *The Shadow Economy in Denmark 1994: Measurement and Results*. Study No. 3, Rockwool Foundation Research Unit, Copenhagen.
- Schneider, F. og D. H. Enste (2000): *Shadow Economics: Size, Causes, and Consequences*, Journal of economic literature, XXXVIII (March), s. 77-114.
- Sporastøyl, J.O. (1982), *Svart arbeid – en utvalgsundersøkelse*, hovedoppgave til cand.oecon.graden, Sosialøkonomisk Institutt, Universitetet i Oslo.
- Train, K. (1993) *Qualitative Choice Analysis Theory, Econometrics and an Application to Automobile Demand*, MIT Press, Cambridge, Massachusetts, London, England, s. 3-53.
- Vogel, J (1974): *Taxation and Public Opinion in Sweden: An Interpretation of Recent Survey Data*. National Tax Journal 27, 4, s.499-513.