Ráðstefnurit Netlu – Menntakvika 2011
Menntavísindasvið Háskóla Íslands

Ritrýnd grein birt 31. desember 2011

Helgi Tómasson

Hugleiðing um framvindu náms og brotthvarf í Hagfræðideild Háskóla Íslands

Umræða um brotthvarf er áberandi í umfjöllun um árangur nemenda í skólakerfinu. Í þessari grein er leitast við að draga upp hlutlæga mynda af því sem hugsanlega liggur að baki skráningartölum og framvindu í háskólanámi. Gögnum um einn árgang nýinnritaðra hagfræðinema í Háskóla Íslands var safnað. Það val var viðleitni til að gera hópinn einsleitan. Öll gagnagreining gengur út frá tölfræðilegu líkani og líkanasmíðin er einfaldari fyrir einsleitan hóp. Leitast er við að skýra prófþátttöku með hefðbundnum einvíðum líkönum. Reikna má með að ómældir eiginleikar einstaklingsins ráði miklu um námsframvindu. Til að leiðrétta fyrir ómældum breytum voru margvíð líkön metin með bayesískum aðferðum. Niðurstöður benda til þess að einn ráðandi þáttur skýri próftökuvilja nemenda. Hugsanlega mætti mæla þann þátt með inntökuprófi.

Höfundur er dósent í hagrannsóknum og tölfræði við Hagfræðideild á Félagsvísindasviði Háskóla Íslands.

Reflections on progress and drop-out of students at the Faculty of Economics, University of Iceland

Concepts, such as *drop-out,* is prevalent in the discussion on success rates of students in the school system. This paper is an objective approach to illustrate what lies behind registration numbers and progress of university studies. Data on a cohort of freshly registered students at the Faculty of Economics at the University of Iceland was used. All data analysis is based on a statistical model and modeling is simpler for a homogenous group. Conventional univariate models are used to account for exam participation. Unmeasured variables are a considered to be a likely explanation for much of the study progress. To account for eventual unmeasured heterogeneity some multivariate binary models were estimated using Bayesian methods. The results suggest that there is one dominant factor explaining the motivation to attend exams. This factor can perhaps be measured by means of a test or exam when entering a university program.

The author is associate professor at the Faculty of Economics, School of Social Sciences, University of Iceland.

Inngangur

Markmið þessarar rannsóknar er að greina þætti sem skýrt geta námsframvindu nemenda í háskóla. Á hverju ári innritast fjöldi nemenda og er framvinda þeirra misjöfn. Grunur leikur á að allmargir nemendur séu ekki nægilega vel undirbúnir við upphaf háskólanáms og muni hætta námi. Mikill kostnaður fylgir því að halda úti námsframboði sem aðeins gagnast broti nemenda. Ljóst er að fjárhagslegur ávinningur er af því ef nemendur sem ekki hafa viðeigandi undirbúning geti nýtt tíma sinn skynsamlegar, til dæmis með undirbúningsnámi eða þátttöku á vinnumarkaði. Einnig myndi kennsla og hagnýt atriði eins og notkun á kennslurými verða markvissari ef ástundunarvilji og hæfni nemenda lægi fyrir í upphafi náms. Æskilegt er því að vita hvort og hvernig vitneskja um fyrri feril nemenda áður en þeir innritast í háskólanám hafi þýðingu fyrir námsgengi. Ekki liggur fyrir hvernig nota megi vitneskju um framhaldsskóla, einkunnir og námsbrautir til að spá fyrir um gengi einstaklings í háskóla. Hugsanlegt er að þessi atriði skipti litlu máli og verður þá að nota aðrar aðferðir svo sem inntökupróf í deildir. Til að geta skilgreint og metið slíka þætti er nauðsynlegt að setja fram tölfræðilegt líkan og álykta út frá því um eðli ferlisins.

Þegar talað er um upplýsingar (e. *information*) þarf að mæla þær með einhverjum hætti. Upplýsingar eru háðar líkindadreifingu eða fjölskyldu af líkindadreifingum. Vel þekkt form er til dæmis Shannon-upplýsingar. Shannon-upplýsingar fyrir líkan sem lýst er með þéttifalli (e. *density function*) f eru:

$$I_f = \int log(f(x))f(x)dx.$$

Í bayesískri tölfræði er sett fram fyrirfram-dreifing (e. *prior*) um hið óþekkta, Upplýsingar þeirrar dreifingar eru I_{prior}. Gögnum er safnað og eftir-á-dreifing (e. *posterior*) reiknuð út með reglu Bayes. Það má segja að upplýsingar í gögnum, I_{data}, séu mismunur á upplýsingum úr eftir-á-dreifingu og upplýsingum úr fyrirfram dreifingu:

$$I_{posterior} = I_{prior} + I_{data}.$$

Settar hafa verið fram ýmsar útgáfur af því hvernig skuli setja fram fyrirfram-dreifingu sem lýsir því að engar upplýsingar séu fyrir hendi. Lágmarksupplýsingar eru það að I_{prior} sé lágmarkað sem jafngildir því að upplýsingar úr gögnum I_{data}, séu hámarkaðar. Lausn á því lágmörkunarvandamáli er háð því hvernig upplýsingar eru skilgreindar. Ef notast væri við Fisher-upplýsingar (eins algengt er í tölfræði) er ekki víst að ástandið lágmarksupplýsingar hefði sama stærðfræðilega form og fæst með því að ganga út frá Shannon-upplýsingum. Það er ekki augljóst hvernig lýsa skuli því ástandi að ekkert sé vitað. Um ýmis form á upplýsingahugtakinu og vandann um hvernig lýsa skuli því ástandi að ekkert sé vitað má til dæmis sjá í kennslubókum (Zellner, 1971) og í greinum eins og til dæmis *Noninformative priors do not exist* (Bernardo, 1997). Grundvallaratriðið er að hugtakið upplýsingar er samofið formi þeirrar fjölskyldu af tölfræðilegum líkönum sem unnið er með. Upplýsingar eru ekki gögn heldur ákveðin tenging milli gagna og tölfræðilegs líkans. Þess vegna eru allar ályktanir út frá mælingum háðar vali á líkani. Val á líkani mótast af því vísindalega vandamáli sem greina á hverju sinni. Ákvörðunin um hvort safna skuli gögnum þarf að byggjast á því að kostnaður við gagnaöflun sé veginn á móti ávinningi af auknum upplýsingum. Fyrir tiltekin líkön er gagnaöflun takmörk sett og því ákveðið óvissulögmál fyrir hendi. Í þessari grein er ekki gengið út frá stærðfræðilegri skilgreiningu á upplýsingum heldur skírskotað til þess að gögn þurfa að auka nákvæmni í tilteknu líkani.

Gögn um námsferil hvers nemanda eru skráð við innritun í Háskóla Íslands. Ef í slíkum gögnum felast einhverjar upplýsingar þarf að meta og túlka viðeigandi tölfræðilegt líkan. Engin gagnagreining er fær án tölfræðilegs líkans. Oft er mikilvægt atriði í greiningu gagna um einstaklinga að hver einstaklingur sé rekjanlegur. Þannig er hægt að ná einstaklingshreinsuðu mati á þýðingu ýmissa skýristærða. Breytileiki einstaklinga getur gert áhrif skýristærða ógreinilegri. Rekjanleiki einstaklings þýðir að ekki þarf að gera sérstakt tölfræðilegt líkan um þýðingu einstakra einkenna einstaklings, til dæmis hæðar, þyngdar, kyns og illmælanlegra (ómældra) stærða eins og hæfni. Aðferðafræðin þarf að byggja á því sem á ensku er kallað „repeated-measures“ (einnig „longitudinal-data“ eða „panel-data“). Hugmyndin að baki einstaklingshreinsuðu mati er að gera vinnu við tölfræðilega líkanasmíði hnitmiðaðri. Ef eiginleikar einstaklinga eru ekki rétt formaðir í tölfræðilegu líkani er hugsanlegt að það skekki ályktanir á máta sem ekki leiðréttist með vaxandi gagnamagni.

Í þessari grein er leitast við að kortleggja hversu gagnlegar framhaldsskólaupplýsingar geta verið til að skýra út framvindu nemenda í hagfræði við Háskóla Íslands. Aflað var gagna um framvindu nemenda í Hagfræðideild Háskóla Íslands og nokkur einföld tölfræðilíkön metin og túlkuð. Líta ber á tölulegar niðurstöður í þessari grein sem forkönnun. Markmiðið er að fá gróft mat á upplýsingagildi framhaldsskólagagna fyrir líkön sem skýra árangur í háskóla.

Ýmsar hvatir liggja að baki vali á háskólagrein. Sumir einstaklingar velja af hugsjón og aðrir velja grein með sama hugarfari og fólk kaupir kort í líkamsræktarstöð eftir áramót. Það er vel þekkt að margir sem kaupa kort í líkamsræktarstöð á nýbyrjuðu ári komast lítið áfram í líkamsrækt. Það er því eðlilegt að gera greinarmun á þeim nemendum sem ekki taka nein próf og þeim sem þreyta í það minnsta eitt próf. Greiningarvandinn er því að minnsta kosti tvíþættur. Athuga þarf hvort mælanlegir eiginleikar spái virkni í námi og síðan hvort þeir spái fyrir um framvindu í námi.

Gert er ráð fyrir að einstaklingar sem hefja háskólanám hafi stundað nám í framhaldsskóla. Þeir koma úr ýmsum framhaldsskólum, og innan hvers skóla eru ýmsar brautir, svo sem náttúrufræðibraut, félagsfræðibraut o.s.frv. Í mörgum skólum eru gefnar einkunnir sem mælikvarði á hæfni og ástundun nemanda. Þau próf eru ekki samræmd, einkunnakvarði ekki samræmdur og því ekki ástæða til að ætla að einkunnakvarði milli skóla sé sambærilegur. Jafnvel þó prófin væru samræmd væri heldur ekki augljóst að einkunnir milli skóla væru sambærilegar, m.a. vegna þess að sumir skólar myndu ef til vill einbeita sér að því að kenna nemendum að taka samræmda prófið. Eftir að framhaldsskóla lýkur er stefna háskólanemans samsett úr að minnsta kosti tveim skrefum; ákvörðuninni um að innrita sig og síðan ákvörðun um ástundun. Innritun er forsenda ástundunar og því eðlilegt að mæla ástundun skilyrt á innritun. Fyrsta skref ástundunar er að þreyta eitt próf. Í prófum eru nemendur kvarðaðir og ákveðinn þröskuldur til að hafa staðist prófið skilgreindur. Kvörðunin er því margþætt, í fyrsta lagi að fara í próf, í öðru lagi að standast prófið og í þriðja lagi er einhvers konar kvörðun á getu nemandans lýst með einkunn, sem á Íslandi er víða á bilinu 0–10.

Skipulag greinarinnar er eftirfarandi. Í næsta kafla er rakið mikilvægi þess að taka samtímis tillit til allra þátta sem áhrif geta haft á myndun gagna. Í þriðja kafla er sagt frá þáttum sem hugsanlega skipta máli í námsframvindu. Þá er lýst gagnasafni úr Hagfræðideild Háskóla Íslands. Því næst er sýnd úrvinnsla byggð á einvíðum líkönum og úrvinnsla byggð á margvíðum líkönum í kafla þar á eftir. Í lokakafla eru hugleiðingar um niðurstöður og frekari greiningar.

Tölfræðileg líkön

Nauðsyn þess að taka tillit til allra þátta samtímis

Allar ályktanir byggðar á gagnagreiningu þurfa að styðjast við tölfræðilegt líkan. Í slíku líkani þarf að taka tillit til allra mikilvægra þátta samtímis, ekki eingöngu þeirra þátta sem rannsaka á. Við val á tölfræðilegu líkani er nauðsynlegt að hafa í huga eðli hins undirliggjandi ferlis og hvernig hugsanlegar ályktanir tengist markmiðum starfseminnar.

Í allri tölfræðivinnu er nauðsynlegt að leiðrétta fyrir truflandi þáttum. Í kennslubókum, eins og til dæmis Poirier (1995), er þetta stundum sýnt með því að láta nemendur sýna það að sé mikilvægri skýribreytu í aðhvarfslíkani (e. *regression model*) sleppt þá muni það leiða til bjagaðs mats á þýðingu annarra skýristærða. Ef sanna líkanið er

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon,$$

og breytunni x_2 sleppt, það er metið líkan er

$$y = \alpha_0 + \alpha_1 x_1 + \epsilon,$$

fær nemandinn það verkefni, að meta hversu villandi mynd matið á α_1 gefur af sannri þýðingu breytunnar x_1, það er β_1. Þessi gerð villu er útbreidd í daglegri gagnavinnu. Í greiningu á námsferlum er einstaklingurinn mikilvæg breyta. Því er mikilvægt að hægt sé að sigta einstaklingseinkenni frá. Í kennslubók í hagrannsóknum eftir Greene (2003) er sagt frá frægri greiningu Hausman og Taylor (1981). Þeir bera saman nokkur líkön þar sem m.a. má álykta um launamun hvítra og svartra í Bandaríkjunum. Ef ekki er leiðrétt fyrir ómældum einstaklingsáhrifum virðist launamunur hvítra og svartra vera um 8% en þegar leiðrétt er fyrir einstaklingsáhrifum virðist munurinn vera um 2%. Aðferðafræði Hausman og Taylor byggist á því að nota margar mælingar á sama einstaklingnum og sigta frá ómælda einstaklingsþætti, svo sem hæfni. Líkanasmiðurinn þarf því ekki að beita hugviti sínu til að giska á hvernig eiga að forma hæfniþátt líkansins, né heldur að safna gögnum sem mæla hæfni. Einfaldasta formið á aðferðafræði af þessari gerð er paraður samanburður, t.d., sami einstaklingur skoðaður fyrir og eftir meðferð. Í allri hlutlægri vinnu er nauðsynlegt að nota viðmið. Í tölfræðilegri líkanasmíði þarf yfirleitt að vinna með nokkr-ar gerðir líkana til að geta tjáð sig í þá veru að eitt sé betra en annað.

Atriði sem skipta máli í námsframvindu

Mældar skýristærðir við skráningu í háskóla eru fæðingarár, kyn, innritunarár í háskóla, framhaldsskóli, lokaeinkunn, námsbraut og útskriftarár úr framhaldsskóla. Í háskólanámi safnast síðan mælingar á útkomum í einstökum námskeiðum. Útkoman er tvívíð, þ.e. hvort einstaklingur náði prófi og hvað viðkomandi fékk í einkunn.

Það að ná prófi er í eðli sínu 0/1-breyta og því er vandamálið það að meta líkur á að við-komandi mæti í próf. Ef gefið er að viðkomandi fari í próf fæst einhver einkunn sem hugs-anlega skýrist að einhverju leyti af mælanlegum eiginleikum einstaklingsins. Ætla má að ómældir eiginleikar einstaklings fylgi honum á háskólaferlinum.

$$\text{árangur} = f(\text{framhaldsskólaupplýsingar, kyn, aldur, o.s.frv}) + \text{hæfni}$$

Hér táknar f eitthvert fall af mælanlegum eiginleikum. *Repeated-measures*-eiginleikann má nota til að leiðrétta fyrir ómældum breytum (eins og hæfni). Það að framhaldsskólar eru ekki eins, námsbrautir ólíkar og notkun á einkunnakvarða mismunandi, eru atriði sem taka á tillit til í líkanasmíðinni. Þetta eru í eðli sínu margvíðar aðferðir, sem bjóða upp á flóknari ályktanir en einvíðar. Aðferðum er lýst í mörgum bókum, eins og t.d. Pinheiro og Bates (2000), Wooldridge (2002), Cameron og Trivedi (2005), Rossi, Allenby og McCulloch (2005).

Öflun gagna

Af tæknilegum ástæðum var ákveðið að athuga aðeins nýskráningar í BS-nám haustið 2008 en þá voru nýskráðir 88 einstaklingar. Á fyrsta ári voru skráðir um það bil 50 til við-bótar í BA-nám og endurinnritaðir með forsögu í Háskóla Íslands. Sú ákvörðun að velja 88 nýnema sem skráðir voru í BS-nám var því viðleitni til að skoða tiltölulega einsleitan hóp. Ef skoða ætti aðra hópa, t.d. þá sem skráðir eru í BA-nám og þá sem koma úr öðr-

Tafla 1 – Meðaltal og staðalfrávik lokaeinkunnar úr framhaldsskóla.			
	Ýmsir skólar	**Verzlunarskóli Íslands**	**Menntaskólinn í Reykjavík**
Meðaleinkunn úr framhaldsskóla	7,16	7,17	6,83
Staðalfrávik	1,49	0,72	0,98

um deildum Háskóla Íslands kallaði það á flóknari tölfræðilega líkanasmíði. Upplýsingar um framhaldsskóla, námsbraut og einkunn hjá þessum 88 lágu fyrir. Einkunnir úr námskeiðunum *Stærðfræði I, Tölfræði I, Þjóðhagfræði I, Rekstrarhagfræði I* og *Haglýsingu* á fyrsta misseri liggja fyrir eftir eitt próf og sjúkrapróf. Gagna var aflað hjá skrifstofu Viðskipta- og Hagfræðideildar Háskóla Íslands. Gögnin voru dreifð í mörgum Excel-skjölum og ljóst að mikil hætta er á villum þegar slíkum skrám er steypt saman. Nemendur komu af 10 tegundum námsbrauta úr 21 framhaldsskóla. Af þessum 88 nemendum stóðust 12 fimm fyrstu prófin og 8 fjögur próf af fimm. Alls voru 51 sem ekki stóðust neitt próf. Flestir komu úr Verzlunarskóla Íslands (27), næst flestir úr Menntaskólanum í Reykjavík (17) og færri (alls 44) úr öðrum skólum. Í *Töflu 1* er sýnt meðaltal og staðalfrávik einkunna úr framhaldsskóla. Alls komu 30 nemendur af náttúrfræðibraut, 23 af hagfræðibraut, 8 af félagsfræðibraut og 7 af viðskiptabraut Verzlunarskóla Íslands.

Tafla 2 – Fylgni milli einkunna í 5 BS-hagfræðigreinum og einkunnar úr framhaldsskóla.						
	Stl	**Töl**	**Þjl**	**Rel**	**Hag**	**Frh**
Stl	1,00	0,98	0,95	0,94	0,99	0,25
Töl	0,98	1,00	0,95	0,95	0,98	0,18
Þjl	0,95	0,95	1,00	0,95	0,94	0,11
Rel	0,94	0,95	0,95	1,00	0,98	0,14
Hag	0,99	0,98	0,94	0,98	1,00	0,21
Frh	0,25	0,18	0,11	0,14	0,21	1,00

Fylgnifylki einkunnar úr framhaldsskóla og einkunnar í 5 fyrstu greinum BS-hagfræði er sýnt í *Töflu 2.* Hin mikla fylgni milli háskólagreinanna innbyrðis skýrist af því að þeir sem féllu, féllu yfirleitt í mörgum greinum. Greinarnar bera nöfnin, *Stærðfræði I*=stl, *Tölfræði I*=töl, *Þjóðhagfræði I*=þjl, *Rekstrarhagfræði I*=rel og *Haglýsing*=hag. Frh stendur fyrir lokaeinkunn úr framhaldsskóla.

Ef aðeins eru skoðaðir þeir 12 sem náðu öllum prófum i fyrstu tilraun þá verður fylgnitaflan eins og sýnt er í *Töflu 3.* Þetta grófa mat (mælingar aðeins 12) bendir til að einkunn úr

Tafla 3 – Fylgni einkunna í fyrstu 5 greinum BS-náms í hagfræði og framhaldsskólaeinkunnar.						
	Stl	**Töl**	**Þjl**	**Rel**	**Hag**	**Frh**
Stl	1,00	0,65	-0,14	0,43	0,35	0,15
Töl	0,65	1,00	-0,25	0,18	0,16	0,26
Þjl	-0,14	-0,25	1,00	0,68	0,16	0,01
Rel	0,43	0,18	0,68	1,00	0,43	0,29
Hag	0,35	0,16	0,16	0,43	1,00	0,49
Frh	0,15	0,26	0,01	0,29	0,49	1,00

framhaldsskóla geti ekki verið mjög afgerandi skýristærð hvað varðar einkunnir í upphafi háskólaferils. Ef sönn fylgni milli einkunnar úr framhaldsskóla og einkunna í háskólaprófum væri mikil (t.d. meiri en 0,9) myndi slíkt sjást þó að mælingar væru fáar. Í töflunni sést lítil fylgni milli einkunna í háskóla og einkunnar í framhaldsskóla.

Einvíð greining: Nokkur einföld líkön (0/1-líkön)

Í fyrirliggjandi gögnum er mest áberandi hve margir taka engin próf. Því var ákveðið að einblína á atriði sem ákvarða próftökuviljann. Fjöldi þeirra sem fer í próf er það lítill að ályktanir um hvaða þættir skýra árangur í prófi verða ekki nákvæmar. Eftir að einstaklingur hefur ákveðið að skrá sig í námskeið kemur að því að taka ákvörðun um að fara í próf. Ákvörðunin að fara í fyrsta prófið, *Stærðfræði I,* má reyna að skýra með binomial-líkani, þ.e. gengið er út frá því að líkur á að gefinn einstaklingur fari í prófið megi skýra út frá því hvaða framhaldsskóla hann var í og hvað hann fékk í einkunn þar. Hér táknar $y_1 = 1$ það að viðkomandi hafi farið í prófið og $y_1 = 0$ að hann hafi ekki farið í það. Tvenns konar útfærsla af líkaninu var prófuð, a) „linear-probability" (LP),

$$E(y_1) = P(y_1 = 1) = \boldsymbol{X\beta}$$

og b) probit,

$$E(y_1) = P(y_1 = 1) = \Phi(\boldsymbol{X\beta})$$

þar sem Φ er dreififall staðlaðrar normalhendingar. Kostir þess að meta LP-líkanið (a) eru að túlkun $\boldsymbol{\beta}$ á þýðingu skýribreytnanna $\boldsymbol{X}$ er auðveld. Galli við LP-líkanið er að við mat á því með venjulegu tölfræðiforriti er hugsanlegt að einstaklingi sé úthlutað neikvæðum líkum eða líkum stærri en 1 á að fara í prófið. Mat á LP-líkani þannig að $P(y_1 = 1)$ sé þvingað til að vera á bilinu [0,1] er tæknilega erfiðara og gerir túlkun á þýðingu skýribreytnanna erfiðari. Í þessari greiningu var það hliðarskilyrði þvingað fyrir gefin gögn en ekki er víst að þvingunin haldi fyrir öll möguleg gildi á $\boldsymbol{X}$. Kostir probit-líkans (b) eru að $P(y_1 = 1)$ þvingast sjálfkrafa á bilið [0,1]. Í *Töflum 4* og *5* er sýnt mat á LP- og probit-líkönum fyrir próftöku í *Stærðfræði I.* Í *Töflum 6* og *7* er sýnt mat þegar einkunn úr framhaldsskóla hefur verið bætt við. Eðlilegt er að gera ráð fyrir möguleikanum á villum sem varað er við í kennslubókum í tölfræði eins og t.d. að skýribreytur vanti. Að baki þessum töflum liggja 80 mælingar, þar sem einkunn úr framhaldsskóla vantaði fyrir 8 einstaklinga.

Það hafa farið í stærðfræðiprófið hefur afgerandi þýðingu fyrir líkur á að standast tölfræðiprófið. Að gefnum upplýsingum um hvort viðkomandi hefur mætt í stærðfræðipróf, hafa upplýsingar um framhaldsskóla og einkunn lítið að segja um líkur á að fara próf í *Tölfræði I.* Þetta má lesa úr *Töflum 8* til *13* þar sem sýndar eru nokkrar útfærslur af því hvernig einkunn, framhaldsskóli og það hvort viðkomandi hafi farið í próf í *Stærðfræði I* hafa áhrif á líkur á að fara í próf í *Tölfræði I.* Tengsl einkunnar og framhaldsskóla gera að verkum að Verzlunarskóli Íslands fær neikvætt formerki í *Töflum 12* og *13.* Skýringin er að einkunnir í Verzlunarskólanum eru að meðaltali hærri en einkunnir úr Menntaskólanum í Reykjavík og einkunnabreyta fer því að leika hlutverk Verzlunarskóla-breytunnar. Ljóst er hins vegar á staðalvillu í mati að nákvæmnin er ekki mikil (vegna fárra mælinga). Svipuð greining á prófþátttöku í *Þjóðhagfræði I, Rekstrarhagfræði I* og *Haglýsingu* bendir til að þegar útkoma úr einu prófi er gefin sé lítil skýrigeta í upplýsingum um framhaldsskóla og einkunn. Þess má geta að skoðun á Bayes-þætti gefur ekki tilefni þess að hægt sé að greina á milli frammistöðu LP-líkans og probit-líkans. Þetta eru í eðli sínu einvíð líkön sem ekki taka tillit til þess að einstaklingar eru ólíkir. Þessa tegund líkana og útkomur í *Töflum 8* til *13* má því gagnrýna vegna þess að þau taka ekki tillit til ómældra mikilvægra breytna, eins og t.d. hæfni. Einhver margvíð greining í anda Hausman og Taylor (1981), sem tekur á ómældum einstaklingseinkennum, er því æskileg. Þessar töflur byggja á hefðbundnum tölfræðiaðferðum og framsetning á niðurstöðum er úr hefðbundnu tölfræðiforriti.

Tafla 4 - Probit líkan fyrir *Stærðfræði I.*				
	Mat	**Staðalvilla**	**Z-gildi**	**Pr(>|z|)**
(Fasti)	-1,2074	0,2486	-4,86	0,0000
VI	0,5618	0,3600	1,56	0,1186
MR	0,9844	0,3948	2,49	0,0127

Tafla 5 – LP líkan fyrir *Stærðfræði I.*				
	Mat	**Staðalvilla**	**T-gildi**	**Pr(>|t|)**
(Fasti)	0,1136	0,0487	2,33	0,0219
VI	0,1456	0,0987	1,48	0,1436
MR	0,2981	0,1308	2,28	0,0252

Tafla 6 – Probit líkan fyrir *Stærðfræði I* þegar einkunn úr framhaldsskóla hefur verið bætt við.				
	Mat	**Staðalvilla**	**Z-gildi**	**Pr(>|z|)**
(Fasti)	-7,2048	1,9043	-3,78	0,0002
VI	0,7102	0,4193	1,69	0,0903
MR	1,5054	0,4951	3,04	0,0024
Einkunn	0,8008	0,2404	3,33	0,0009

Tafla 7 – LP-líkan fyrir *Stærðfræði I* þegar einkunn úr framhaldsskóla hefur verið bætt við.				
	Mat	**Staðalvilla**	**T-gildi**	**Pr(>|t|)**
(Fasti)	0,0000	0,0000	0,00	1,0000
VI	0,0764	0,0994	0,77	0,4446
MR	0,2172	0,1270	1,71	0,0914
Einkunn	0,0230	0,0083	2,77	0,0071

Tafla 8 – Probit-líkan fyrir *Tölfræði I.*				
	Mat	**Staðalvilla**	**Z-gildi**	**Pr(>|z|)**
(Fasti)	-0,4728	0,1968	-2,40	0,0163
VI	0,1419	0,3151	0,45	0,6524
MR	0,6958	0,3645	1,91	0,0562

Tafla 9 – LP-líkan fyrir *Tölfræði I.*				
	Mat	**Staðalvilla**	**T-gildi**	**Pr(>|t|)**
(Fasti)	0,3182	0,0714	4,45	0,0000
VI	0,0522	0,1185	0,44	0,6608
MR	0,2701	0,1409	1,92	0,0587

Tafla 10– Probit líkan fyrir *Tölfræði I* þegar þátttöku í stærðfræðiprófi hefur verið bætt við.				
	Mat	**Staðalvilla**	**Z-gildi**	Pr(>|z|)
(Fasti)	-0,6930	0,2130	-3,25	0,0011
VI	-0,1377	0,3554	-0,39	0,6984
MR	0,2703	0,4189	0,65	0,5188
Stærðfræði I	1,8939	0,4295	4,41	0,0000

Tafla 11 – LP líkan fyrir *Tölfræði I* þegar einkunn úr framhaldsskóla hefur verið bætt við.				
	Mat	**Staðalvilla**	**T-gildi**	**Pr(>|t|)**
(Fasti)	0,0000	0,0000	0,00	1,0000
VI	-0,0125	0,1218	-0,10	0,9182
MR	0,2344	0,1412	1,66	0,1011
Einkunn	0,0525	0,0111	4,74	0,0000

Tafla 12 – Probit líkan fyrir *Tölfræði I* þegar einkunn úr framhaldsskóla og þátttöku í stærðfræðiprófi hefur verið bætt við.				
	Mat	**Staðalvilla**	**Z-gildi**	**Pr(>|z|)**
(Fasti)	-1,1304	1,1157	-1,01	0,3110
VI	-0,2105	0,3672	-0,57	0,5664
MR	0,2376	0,4369	0,54	0,5866
Stærðfræði I	1,8042	0,4490	4,02	0,0001
Einkunn	0,0741	0,1535	0,48	0,6293

Tafla 13 – LP líkan fyrir *Tölfræði I* þegar einkunn úr framhaldsskóla og þátttöku í stærðfræðiprófi hefur verið bætt við.				
	Mat	**Staðalvilla**	**Z-gildi**	**Pr(>|z|)**
(Fasti)	0,0000	0,0010	0,00	0,9997
VI	-0,0379	0,1040	-0,36	0,7165
MR	0,0772	0,1076	0,72	0,4751
Einkunn	0,0375	0,0104	3,61	0,0006
Stærðfræði I	0,5833	0,1010	5,78	0,0000

Margvíð líkön

Lausleg skoðun á einvíðum líkönum gefur til kynna að hugsanlega vanti mikilvægar skýristærðir. Því gæti verið upplýsandi að meta margvítt líkan þar sem til dæmis prófþátttaka væri skýrð. Þar sem mælingar eru fáar þarf greinandi að setja viðeigandi skorður til að fá túlkanlegar útkomur. Hér var valin sú leið að meta margvítt probit-líkan með bayesískum aðferðum til að skýra prófþátttöku. Lausleg útskýring á tæknilegum atriðum varðandi uppsetningu og bayesískt mat á margvíðu probit-líkani er sýnd í viðauka. Hugmyndin að baki þeirri margvíðu nálgun sem beitt var hér er að nokkrar mælanlegar stærðir skýri prófþátttöku og síðan sé ómælanleg stærð,vilji til próftöku eða próftökuvilji, mismunandi eftir einstaklingum.

Tafla 14 – Metið fylgnifylki próftökuvilja.				
Stl	**Töl**	**Þjl**	**Rel**	**Hag**
1,00	0,85	0,87	0,83	0,80
0,85	1,00	0,95	0,95	0,74
0,87	0,95	1,00	0,98	0,89
0,83	0,95	0,98	1,00	0,85
0,80	0,74	0,89	0,85	1,00
1,00	0,85	0,87	0,83	0,80

Tafla 15 – Áhrif skóla á líkur á próftökuvilja.					
Breyta	$\hat{\beta}$	**Staðalfrávik ($\hat{\beta}$)**	**5%**	**50%**	**95%**
(Fasti)	-0,79	0,22	-1,17	-0,78	-0,47
VI	0,13	0,26	-0,27	0,12	0,58
MR	0,31	0,28	-0,12	0,29	0,81

Hluti útkomu er sýndur í *Töflum 14* til *16. Tafla 14* sýnir metið fylgnifylki fyrir próftökuvilja þegar leiðrétt hefur verið fyrir hugsanlegum mun á milli skóla. Í þáttagreiningu (e. *factor analysis*) hafa menn stungið upp á þeirri þumalfingursreglu að fjöldi þátta sé fjöldi eigingilda fylgnifylkisins sem eru stærri en einn. Þáttagreining er ákveðin nálgun í margvíðri tölfræði (e. *multivariate statistics*) sem gengur út á að skoða fylgnifylki með það í huga að raunveruleg vídd fylkisins sé minni en fjöldi breyta. Gengið er út frá því að lítill fjöldi þátta geti skýrt allar breyturnar. Tæknin byggir á því að eigingildi fylkins eru reiknuð. Í því tilfelli að einungis eigingildið væri frábrugðið 0 væri það túlkað þannig að einn þáttur réði útkomum í öllu breytunum. Ein þumalfingursregla til að ákvarða fjölda þátta er að telja fjölda eigingilda fylgnifylkis sem eru stærri en 1. Í fylkinu í *Töflu 14* er stærsta eigingildið 4,46, það eina sem er stærra en 1 og það næsta er 0,27. Heildarsumma eigingilda í fimmvíðu fylgnifylki er 5. Einnig mætti orða þetta þannig að 90% eigingildasummunar séu í stærsta eigingildinu. Samkvæmt þumalfingursreglunni virðist því sem próftökuvilji hvers einstaklings í öllum fögunum sé að mestu einn þáttur. Í *Töflu 15* er sýnt mat á þýðingu skóla. Hér var það hliðarskilyrði sett að skóli hefði sömu áhrif á prófgetu í öllum greinunum fimm. Þetta er eins konar margvíð aðhvarfsgreining þar sem háða breytan er 0/1. Líkanið er metið með því hliðarskilyrði að skóli hafi sömu áhrif í öllum greinum. Ef meiri gögn væru aðgengileg væri ef til vill hægt að álykta um það hvort tiltekinn skóli gæfi sérstaklega góðan undirbúning í einni grein. Samkvæmt þessari töflu er próftökuviljinn mestur hjá þeim sem hafa verið í Menntaskólanum í Reykjavík. Nákvæmnin er eins og vænta mátti ekkert sérstök og stærð stikanna gefur tæplega tilefni til þess að umgangast nemendur úr þeim skóla með öðrum hætti en aðra. *Töflur 15* til *16* sýna bayesískt mat á stikum. Þar er sýnt

Tafla 16 – Áhrif skóla á líkur á próftökuvilja.					
Breyta	$\hat{\beta}$	**s.e($\hat{\beta}$)**	**5%**	**50%**	**95%**
(Fasti)	-1,723	0,693	-2,9400	-1,684	-0,669
VI	1,448	0,783	0,2626	1,398	2,797
MR	1,249	1,191	-0,6580	1,226	3,271
Einkunn	0,141	0,087	0,0095	0,137	0,291
Einkunn x VI	-0,169	0,107	-0,3525	-0,164	-0,005
Einkunn x MR	-0,035	0,171	-0,3110	-0,038	0,257

meðaltal og kvantílar í eftir-á-dreifingu (e. *a posteriori*) stikanna. Túlkunin á *Töflum 15* og *16,* er frábrugðin hefðbundinni tölfræðitúlkun. Í bayesískri tölfræði er litið á óþekkta stika í líkani sem hendingar og óvissu um þá lýst með líkindadreifingu.

Tæknilega er ekkert því til fyrirstöðu að fjölga breytum. Til dæmis mætti bæta við þetta líkan einkunn og jafnvel samspili einkunnar og skóla. Ekki er ástæða til að ætla að einkunnakvarðinn sé samanburðarhæfur milli skóla, hvorki hvað varðar staðsetningu (meðaltal) né skrefastærð (staðalfrávik). Metin voru nokkur líkön af þessari gerð þar sem einkunn hafði verið bætti við. Nákvæmnin í þeim líkönum býður ekki upp á að hægt sé að draga sterkar ályktanir. Dæmi um líkan þar sem einkunn úr framhaldsskóla hefur verið bætt við er sýnt í *Töflu 16.* Frjálsleg túlkun hennar væri t.d. að einstaklingur sem hvorki kemur úr Verzlunarskóla Íslands né Menntaskólanum í Reykjavík og hefur fengið 7 í einkunn hefði líkur

$$\Phi(-1,723+0,143*7)=0,23$$

á að fara í próf. Sams konar einstak-lingur með 8 í einkunn hefði líkur

$$\Phi(-1,723+0,141*8)=0,28$$

á að fara í próf. Einnig sú (órökrétta) túlkun að einstaklingur úr Verzlunarskóla Íslands með einkunn 8 hefði líkur

$$\Phi(-1,723+1448-0,141*8-0,169*8)=0,31$$

á að fara í próf en sá sem væri með 7 í einkunn hefði líkur

$$\Phi(-1,723+1,448+0,141*7-0,169*7)=0,32$$

á að fara í próf.

Lokaorð og frekari rannsóknir

Þau gögn sem liggja til grundvallar þessum hugleiðingum byggjast á einum árgangi nýinnritaðra hagfræðinema. Til að hafa gögnin einsleit var ákveðið að einskorða þau við nýinnritaða BS-nema; BA-nemum og eldri nemendum sleppt. Ljóst er að ef nota á líkön af þeirri gerð sem hér hafa verið kynnt til að álykta nákvæmlega um þýðingu undirbúnings úr framhaldsskóla, svo sem einkunnar og námsbrautar í framhaldsskóla og framhaldsskólans sjálfs, væri til bóta að hafa aðgang að fleiri árgöngum. Ýmsir kynnu að segja að það þyrfti að minnsta kosti 10 sinnum fleiri einstaklinga til að nákvæmni sé viðunandi. Hvað síðan eigi að gera við það nákvæma mat er annað mál. Ef það kæmi til dæmis fram að líkur á að nemandi úr tilteknum framhaldsskóla nái prófi eru 0,2 en 0,3 hjá nemanda úr öðrum skóla, á þá að hafa mismunandi stefnu gagnavart inntöku nemenda úr þessum skólum? Vissulega má segja að gögnin að baki þessari greiningu séu rýr. Þau eru þó ekki gagnslaus. Skorti á upplýsingum verður ekki mætt með stórfelldri gagnasöfnun. Stórfelld gagnasöfnun, til dæmis úr öllum Háskóla Íslands, myndi kalla á flóknari líkanasmíði. Mat á hnitmiðuðu tölfræðilíkani fyrir vel skilgreint vísindalegt vandamál getur dregið stórlega úr óvissu.

Ljóst er af margvíðu greiningunni að einstaklingseiginleikar eru ríkjandi afgerandi. Þessar niðurstöður gefa ekki tilefni til að álykta að þýðing framhaldsskóla sé afgerandi þáttur í væntanlegri framvindu nemanda í háskóla. Eðlilegt er að gera ráð fyrir að framhaldsskólar noti einkunnakvarða á ólíkan hátt. Einn meginþáttur, sem kalla mætti próftökuviljann, er ráðandi fyrir árangur á fyrsta ári í háskóla. Ef gott mat fengist á þennan þátt snemma hausts, til dæmis með inntökuprófi áður en kennsla hefst, mætti hugsanlega gera kennsluna markvissari öllum til hagsbóta. Ályktanir byggðar á einvíðum líkönunum sem sýnd eru í þessari grein eru misvísandi og tal um tölfræðilega marktækni í slíkum líkönum markleysa. Auk þess ber að hafa í huga að tölfræðileg marktækni er mælikvarði á nákvæmni í mati, en hefur ekkert með mikilvægi að gera.

Eðlilegt er að gera ráð fyrir að skólar þróist yfir tíma. Ef álykta ætti út frá mælingum sem safnað væri í mörg ár þyrfti að taka tillit til þess í tölfræðilegu líkani, sem þá yrði flóknara. Líkan sem tæki á þróun yfir tíma gæti t.d. litið svona út:

$$
\begin{aligned}
\boldsymbol{y}'_{it} &= (\Phi(y_{i1t}), \ldots, \Phi(y_{i5t})) \\
y_{ijt} &= \begin{cases} 1 \text{ ef } y^*_{ijt} > 0 \\ 0 \text{ annars} \end{cases} \\
y_{ijt} &= \boldsymbol{X}_i \boldsymbol{B} + \psi_t + \varepsilon_{ijt} \\
\psi_t &= \text{ einhvers konar dýnamískt ferli, AR, MA, etv. fall af } \boldsymbol{X}.
\end{aligned}
$$

Ef meta á marga árganga saman þarf með einhverjum hætti að taka tillit til þróunar í tíma sem hér að ofan er gert með liðnum ψ_t. Af þessari forkönnun er ljóst að einstaklingurinn skiptir afar miklu máli og því nauðsynlegt að hægt sé að rekja gögn til einstakra nemenda. Stundum vekur krafan um að tengja niðurstöður við einstaklinga tortryggni, ýmsir tala um persónuvernd og gera kröfur um leyfi og gagnaeyðingu. Það er brýnt að skilja að einstaklingsrekjanleiki verður ekki lagður að jöfnu við persónunjósnir. Vinna af þeirri gerð sem hér er lýst hefur að markmiði að skilja hvaða þættir skýra námsárangur við Háskóla Íslands. Gögnin eru til innan Háskóla Ísalnds og verða þar áfram. Einungis er spurning um hvort og hvernig eigi að nota þau og ekki síst hvað kostar að gera þau aðgengileg.

Á að safna frekari gögnum? Að fenginni reynslu er vitað að tæknileg vinna verður umfangsmikil. Það er ljóst að óhemju vinna liggur í því raða excel-skrám og öðru illa samræmdu efni saman. Það var mikil vinna fyrir höfund þessara lína að raða saman Excel-skjölum. Án efa urðu til margar gagnavillur í því ferli. Þetta er mikið gæðamál, ekki einungis fyrir Hagfræðideild heldur fyrir Háskóla Íslands í heild og reyndar hvern sem er að bókhaldsgögn séu í röð og reglu. Nauðsynlegt er að viðkomandi aðilar skilgreini markmið og leiðir í slíkri vinnu. Er það þess virði að staðla gögn svo hægt sé að meta líkön í anda þeirra líkana sem hér hefur verið lýst? Það fer eftir því hver gagnsemi vel metins líkans er. Það er ekki augljóst að það sé besta meðferð á fé að safna endalaust meiri gögnum. Þó að óvissa sé mikil þá má minnka hana mikið með hnitmiðuðu líkani og hóflegu gagnamagni. Mikilvægt er að átta sig á hver hugsanlegur ávinningur er af aukinni nákvæmni (upplýsingum) í mati á tölfræðilegum líkönum. Vinna við hönnun hnitmiðaðra tölfræðilíkana er nauðsynleg þannig að hægt sé að forðast villur. Mikið magn gagna mun ekki leiðrétta bjagað mat sem myndast þegar mikilvægum breytum er sleppt.

Þegar þetta er ritað, haustið 2011, höfðu 4 af þeim 88 sem innrituðust í BS-nám í hagfræði haustið 2008 lokið náminu á þremur árum (BS-námið er þriggja ára nám) vorið 2011. Veturinn 2011–2012 stefnir í að 3 til viðbótar muni ljúka því. Af árganginum eru enn 15 í námi í hagfræði og virðist sem hluti þeirra stefni á að ljúka vorið 2012. Nokkrir þessara 15 eru þó aðeins hálfnaðir með námið. Af hinum 66 er það að segja að 27 eru í öðru námi í Háskóla Íslands haustið 2011, 17 hafa horfið úr námi og ekki fengust upplýsingar um 22. Það mætti því segja að þremur árum eftir að 88 nemar hófu BS-nám í hagfræði í Háskóla Íslands sé rúmur helmingur enn skráður í eitthvert nám við skólann.

Mín ályktun af þessari vinnu er að einstaklingsbreytileikinn virðist það mikill að lítil efnisleg rök séu fyrir því að meðhöndla beri nemendur úr tilteknum framhaldsskólum á ólíkan hátt. Einstaklingsbreytileikinn er hinn ráðandi þáttur. Mikilvægt er að ná fljótt mati á ástundunarhæfni einstaklingsins þannig að skólastarfið verið sem markvissast.

Nemendur sem ná prófum eru nokkurn veginn eins í öllum framhaldsskólum. Hlutfall slíkra nemenda er ef til vill aðeins mismunandi og notkun einkunna og samsetning náms breytileg eftir skólum og námsbrautum. Einvíðu líkönin sem lýst er hér geta verið að gefa villandi skilaboð því að í þau vantar hugsanlega mikilvæga skýristærð, t.d. hæfni. Margvíða greiningin bendir til þess að mældu skýristærðirnar, framhaldsskóli, einkunn úr framhaldsskóla, o.s.frv., geti ekki skýrt mikið. Einn meginþáttur virðist vera mest áberandi eiginleikinn. Hugsanlega er hægt að mæla þann þátt með inntökuprófi snemma í háskólanámi og stuðla þannig að betri nýtingu á tíma nemenda og kennara.

Þakkir

Það er mikil vinna að gera mælingar úrvinnsluhæfar. Starfsfólk á skrifstofu Viðskipta- og Hagfræðideildar Háskóla Íslands á þakkir skilið. Sérstaklega ber að þakka Sólveigu Ástvaldsdóttur fyrir að finna færslur í tölvum Háskóla Íslands. Einnig þakka ég ritrýnum fyrir gagnlegar ábendingar.

Heimildir

Bernardo, J. M. (1997). Noninformative priors do not exist. *Journal of Statistical Planning and Inference*, *65*, 159–185.

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

Greene, W. (2003). *Econometric Analysis*. New Jersey: Prentice Hall.

Hausman, J. og Taylor, W. (1981). Panel data and unobservable individual effects. *Econometrica*, *49*, 1377–1398.

Pinheiro, J. og Bates, D. (2000). *Mixed-effects models in S and S-Plus*. New York: Springer.

Poirier, D. J. (1995). *Intermediate statistics and Econometrics: A comparative approach*. Cambridge: Massachusetts Institute of Technology.

Rossi, P.; Allenby, G. og McCulloch, R. (2005). *Bayesian Statistics and Marketing*. Hoboken, NJ: John Wiley og Sons.

Wooldridge, J. (2002). *Econometric Analysis of cross section and panel data*. Cambridge: MIT press.

Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. Hoboken, NJ: John Wiley og Sons.

Helgi Tómasson. (2011).
Hugleiðing um framvindu náms og brotthvarf í Hagfræðideild Háskóla Íslands.
Ráðstefnurit Netlu – Menntakvika 2011. Menntavísindasvið Háskóla Íslands.
Sótt af http://netla.hi.is/menntakvika2011/013.pdf

Viðauki

Mat á margvíðu líkani með bayesískum aðferðum

Metið líkan var á forminu

$$\boldsymbol{y}_i^* = \boldsymbol{X}_i \boldsymbol{B} + \boldsymbol{\varepsilon}_i, \quad \boldsymbol{\varepsilon}_i \sim N(0, \Sigma), \quad (1)$$

$$\boldsymbol{y}_i' = (y_{i1}, y_{i1}, y_{i3}, y_{i4}, y_{i5}),$$

$$y_{ij} = \begin{cases} 1 \text{ ef } y_{ij}^* > 0 \\ 0 \text{ annars} \end{cases}, \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \vdots & \ddots & & \vdots \\ \sigma_{1k} & \cdots & \cdots & \sigma_{kk} \end{bmatrix}.$$

Hér táknar mælda breytan $y_{ij} = 1$ að einstaklingur i hafi farið í próf j. Fylkið $\boldsymbol{X}_i$ táknar fylki af skýribreytum fyrir einstakling i. Ef fyrir hendi eru k skýribreytur hefur fylkið vídd $5x(5k)$. Vægi skýribreytnanna er lýst með fylkinu $\boldsymbol{B}$

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_5 \end{bmatrix}, \quad \boldsymbol{\beta}_j' = (\beta_{1j}, \beta_{2j}, \ldots, \beta_{kj}),$$

og $\boldsymbol{X}_i$ má skrifa sem

$$\boldsymbol{X}_i = \begin{bmatrix} X_i & 0 & 0 & 0 & 0 \\ 0 & X_i & 0 & 0 & 0 \\ 0 & 0 & X_i & 0 & 0 \\ 0 & 0 & 0 & X_i & 0 \\ 0 & 0 & 0 & 0 & X_i \end{bmatrix}.$$

Í margvíðri greiningu kemur fylki skýribreyta X_i, fyrir í hverri jöfnu svo að úr verður stórt fylki skýribreytna $\boldsymbol{X}_i$ fyrir einstakling i. Til að hægt sé að meta þetta líkan þarf að setja hliðarskilyrði á stikana í Σ og $\boldsymbol{B}$. Ljóst er að það að margfalda með fasta í gegnum jöfnu (1) mun ekki breyta líkaninu. Eðlilegt skilyrði er að breyta Σ-fylkinu í fylgnifylki með því að skala líkanið á þann hátt að margfaldað sé í gegnum jöfnu (1) með fylkinu Λ:

$$\Lambda = \begin{bmatrix} 1/\sqrt{\sigma_{11}} & \cdots & 0 \\ 0 & \ddots & \vdots \\ \cdots & \cdots & 1/\sqrt{\sigma_{kk}} \end{bmatrix},$$

Metnir stikar verða því $\Lambda\boldsymbol{B}$ og $\Lambda\Sigma\Lambda = R$ (fylgnifylki). Einnig eru önnur skilyrði hugsanleg.

Bayesísk mat á líkani af þessari gerð var metið fyrir fimmvíða mælingu á prófárangri 88 nýnema í BS-námi í hagfræði. Skýribreytan var einungis framhaldsskóli, (VÍ, MR annað). Fyrirfram-dreifing (e. *a priori dreifing*) stika var skilgreind:

$$\boldsymbol{B} \sim N(\boldsymbol{0}, A^{-1})$$

$$\Sigma^{-1} \sim Wishart(\nu, V)$$

Þar sem ákveðið hefur verið að meta fylgnifylki er eðlilegt að setja V=einingarfylki. Óvissan um Σ var ákveðin mikil með því að velja $\nu = 3$. Fyrirfram nákvæmnin í $\boldsymbol{B}$, Λ var ákveðin lítil, t.d. einingarfylki margfaldað með 0,001. Líkanið var metið með Gibbs-MCMC (Markov-Chain-Monte-Carlo) eins og sýnt er í kennslubók eftir Rossi, Allenby og McCulloch (2005).