



GÖTEBORGS UNIVERSITET
HANDELSHÖGSKOLAN

*En explorativ studie av det svenska
pensionssparandet*

Kandidatuppsats i marknadsföring



Johan Jansson
Andreas Jonasson
Handledare: Jeanette Hauff

Höstterminen 2011

Sammanfattning

Den svenska pensionsreformen innebär att de svenska löntagarna själva får besluta om en del av sitt pensionssparande. Denna liberalisering ställer höga krav på det finansiella beslutsfattandet hos den svenska befolkningen.

Icke parametriska metoder har använts för att analysera data från svenska premiepensionsmyndigheten.

Spridningen i resultat fondsparare mellan är mycket stor, men de som aktivt byter fonder tenderar att ha en högre avkastning än mer passiva sparare. Förenklade beslutsbeteenden så som home bias och $1/n$ bias visade sig ha en liten påverkan på avkastningen trots att de anses som negativ påverkan på avkastningen.

Att använda historisk avkastning som beslutsunderlag i hopp om hög framtida avkastning visade sig vara en god strategi, speciellt under en period med positiv avkastning. Tidigare forskning visar att de som är yngst och äldst i den svenska befolkningen har sämst förutsättningar att fatta rationella finansiella beslut. Resultatet i denna rapport ger inget stöd för denna teori.

Abstract

The Swedish Premium Pension reform made it possible for Swedish workers to influence a part of their future pension. This liberalization put new demands on the Swedish citizens to make qualified financial decisions.

Non-parametric methods were used to analyze data from the Swedish Pensions Agency.

The investors who trade their funds most frequently tends to have a higher return than passive investors. Thus, there are big differences in rate of return even among the active investors. Many biases (for example home bias and $1/n$ bias), that are considered to have a negative impact on rate of return were, found to have a small impact.

To base investment decisions on historically high returns, in order to achieve higher future returns turned out to be a good strategy, especially during periods of positive returns. Previous research indicates that young and old people in Sweden are worse off when making rational financial decisions. Our findings do not support this theory.

Innehåll

Den svenska pensionsreformen	1
Problemställning.....	3
Syfte.....	5
Avgränsningar.....	5
Teoretisk referensram.....	5
Metod och modell	14
Resultat och analys.....	26
Slutsats	63
Slutdiskussion	65
Referenser	66
Bilaga A - Systeminställningar och programkod.....	

Den svenska pensionsreformen

Efter ett riksdagsbeslut 1994 förändrades det svenska pensionssystemet i samband med sekelskiftet. Det har sedan dess bestått av tre delar. Den första delen benämns inkomstpension (16 procent av lönen) och används till att försörja de som är pensionärer idag. Del nummer två motsvarar två och en halv procent av inkomsten och benämns premiepension. Denna del ska förvaltas av inkomsttagare själv för att sedan vara en del av dennes pension när hon går i pension. Den tredje delen är en garanterad framtida pension som ska säkerställa att ingen pensionär blir helt utan ersättning. Denna del bygger på att 18,5 procent av den individuella inkomsten sätts undan till dennes framtida pension. Om dessa 18,5 procent överstiger sju och ett halvt prisbasbelopp kommer det som överstiger sju och ett halvt prisbasbelopp inte tillfalla individen men ändå dras från dennes inkomst (Nordén, 2010).

Beslutet om att byta ut det gamla systemet var långt ifrån enhälligt då Socialdemokraterna motsatte sig reformen medan högerpartierna argumenterade att det nya systemet skulle komma att spela en viktig roll. Utgången blev en kompromiss där staten fortfarande förvaltar en väldigt stor andel av den framtida pensionen. Sverige beskrivs som något av ett föregångsland då många länder i västvärlden kommer att vara tvungna att göra om sina pensionssystem för att kunna försörja en åldrande befolkning (Palme et al., 2005).

Denna uppsats fokuserar på premiepensionen då det är denna del som individen själv kan påverka och fatta beslut omkring.

När det nya pensionssystemet lanserades fick spararna själva välja hur de ville placera sina två och en halv procent. Valmöjligheten bestod av 464 olika fonder (utbudet har ökat markant sedan dess, Palme et al., 2005) och placeringen kunde bestå av mellan en och fem olika fondval. För de som inte gjorde ett aktivt val fanns även ett "default val" bestående av sjunde AP-fonden. Den sjunde AP-fonden var från början tänkt att vara en lågriskfond som främst skulle investera i räntepapper. Denna tanke blev dock nedröstad då det ansågs att många låginkomsttagare skulle välja detta default val, och att avkastningen skulle vara alltför låg. Istället är sjunde AP-fonden en långsiktig lågriskfond innehållandes både aktier och obligationer, där andelen aktier måste ligga mellan 80 och 90 procent av portföljens innehåll och 75 procent av dessa får investeras i utländska aktier (Palme et al., 2005).

Investerarna fick inför pensionsvalet en broschyr med information om fonderna. Informationen bestod bland annat av risk, historisk avkastning och avgifter, där risken definieras som avkastningens standardavvikelse de senaste tre åren. Fonderna delades även upp i tre olika kategorier, 100 procent inhemska investeringar, 50 procent inhemska och 50 procent utländska investeringar eller 100 procent utländska investeringar (Nordén, 2010 och Palme, 2005).

Från och med att en person fyllt 61 kan denne välja att få sin pension utbetald i form av en annuitet från PPM. Individen själv får välja om det ska vara ett fast eller rörligt belopp (Palme et al., 2005).

Ett stort stickprov gjorde när den svenska befolkningen för första gången (år 2000) fick chansen att göra ett aktivt val. Vid denna tidpunkt valde cirka två tredjedelar att aktivt besluta om vilka fonder deras framtida pensionspengar skulle förvaltas i. Den resterande tredjedelen gjorde inte ett aktivt val, varpå deras pensionspengar förvaltades av sjunde AP-fonden (Nordén, 2010).

Som tidigare nämnts var valdeltagandet vid det första valet år 2000 relativt högt, cirka två tredjedelar gjorde då ett aktivt val. Denna siffra har dock sjunkit drastiskt för de årskullar som kommit in på arbetsmarknaden efter år 2000. Redan vid valet ett år senare gjorde endast 17,6 procent ett aktivt val och år 2005 hade den siffran sjunkit ytterligare till 8,0 procent. Enligt Weaver (2003/2004) beror det kraftiga fallet i aktivt val på karaktärsdragen hos arbetarna och beslutssituationen. De som gjorde sina val senare än 2000 var betydligt yngre än medelåldern hos de som valde år 2000, vilket innebär att de hade längre tid till pension. Dessa yngre människor var nya på arbetsmarknaden och hade förhållandevis låga löner. Dessa faktorer tillsammans anses ha gjort valet mindre intressant vid senare årgångar. Dock kan detta endast förklara en lite del av det låga valdeltagandet. Personer i den åldersgruppen som gjorde sitt val år 2000 hade ett valdeltagande på cirka 60 procent. En tredje mycket viktig faktor till det stora fallet som gjort ett aktivt val är spridningseffekten. Eftersom en så stor andel av den svenska befolkningen gjorde sitt första val år 2000 hade andras deltagande och diskussioner påverkan på andra. Det är rimligt att anta att den effekten varit betydligt mindre vid senare val. Dessutom har reklamen och informationen via media varit betydligt mindre omfattande vid senare val (Weaver, 2004).

Cronqvist (2003) uppskattar att de svenska fondbolagen spenderade cirka 65 miljoner dollar på reklam inför valet år 2000, motsvarande siffra år 2003 var 10 miljoner dollar. Det minskade värdet i reklam sammanfaller med det sjunkande antalet sparare som gjort ett aktivt val (Cronqvist et al., 2003).

Inför det första valet användes i stor utsträckning olika tjänster för att assistera det finansiella valet. Användandet av dessa tjänster har minskat dramatiskt sedan dess.

Att alla individer besitter fullständig rationalitet och att hela befolkningen därför gynnas av ett pensionssystem där samtliga individer själva får fatta beslut om sitt sparande, bygger på en enkel nationalekonomisk hypotes som kallas *den effektiva marknadshypotesen* (längre fram i denna uppsats kommer denna att diskuteras mer utförligt). I praktiken är denna hypotes inte fullständigt tillämpbar. Sanningen är att många finansiella beslut fattas utan fullständig rationalitet. Detta kan komma att orsaka stora negativa konsekvenser för den privata ekonomin. Det kan även rubba den finansiella stabiliteten, framför allt i länder med välutvecklade kreditmarknader. Det är dock inte så enkelt som att förse alla samhällets individer med information för att undkomma detta problem, individerna måste också ha en möjlighet och vilja att tillgodogöra sig informationen (Almenberg och Widmark, 2011).

Problemställning

Anders Lundström som grundat tjänsten *Min pension* riktar kritik mot systemet då han anser att informationen är överväldigande för konsumenterna. Detta är information som kommer från bland annat pensionsmyndigheten, fondbolag, förmedlare och valcentraler. De danska forskarna Möller och Nielsen går även de till hård kritik mot det nya svenska pensionssystemet som de hävdar flyttar risken från staten och arbetsgivare till privatpersoner. Vidare kritiserar de systemet för att vara för generellt för att passa alla och främst missgynnar unga och höginkomsttagare.

Även AMF-chefen Camilla Larsson kritiserar det nya pensionssystemet. Hon uttrycker en oro för att systemet är så komplicerat och svårt att sätta sig in att många istället väljer att göra ingenting.

Trots detta ses Sverige som något av ett föregångsland för pensionsreformer och många studier har gjorts angående utfallet av den svenska pensionsreformen (Palme et al., 2005). Flera studier visar på stora skillnader i avkastning mellan individer inom premiepensionssystemet (www.pensionsmyndigheten.se)

I denna komplexa valsituationen blir det viktigt för den individuella spararen att göra bra val i pensionssystemet, då dessa val kommer i förlängningen ha en stor inverkan på den framtida pensionen. Det går inte att anta att alla som behöver göra pensionsval besitter fullständig information eller ha förmåga att dra optimala slutsatser utifrån informationen. För att kunna hantera den här typen av svåra problem använder människor sig av olika förenklade beslutsmekanismer för att trots komplexiteten kunna navigera genom verkligheten på ett effektivt sätt. Dessa förenklade beslutsmekanismer har som karaktäristik att de spar energi och tid genom att de inte ger det optimala svaret, utan ett svar som baserat på tidigare erfarenheter, har fungerat någorlunda väl i många situationer. I en situation som pensionsvalet där vi ställs inför svåra val och där utfallet av valet kan ligga långt fram i framtiden har det visat sig extra svårt att motivera människor att göra väl avvägda val trots att de rationellt kan förstå att valet kan ha stora konsekvenser för dem (Benartzi och Thaler, 2007)

Tidigare forskning pekar på ett flertal faktorer som påverkar förmågan att göra goda finansiella val. I en ny studie av Almenberg och Widmark (2011) presenteras stora brister i den svenska befolkningens räknefärdigheter i allmänhet och finansiella räknefärdigheter i synnerhet. Resultatet av studien visar att det finns stora kunskapsskillnader mellan olika grupper i det svenska samhället. Studien pekar även på ganska stora skillnader i finansiell kunskap mellan olika åldersgrupper.

Ett tydligt samband finns mellan de som har god avkastning på sitt PPM-sparande och de som frekvent byter ut sina fonder (Dahlquist et al. 2011). Bevisningen Dahlquist lägger fram visar att det är statistiskt säkerställt att frekvensen har en stor betydelse för vilken avkastning individen får på sitt PPM-sparande.

Olika förenklade beslutsmekanismer, så kallade bias, spelar en stor roll för avkastningen på finansmarknaderna. Exempel på bias är:

- **1/n bias** diskuteras av Thaler och Benartzi (2007) som en förenklad beslutsmekanism för att göra en enkel riskspridning. I korthet innebär 1/n bias att en person som ställs inför ett komplext val där han omedelbart behöver välja upp till n stycken alternativ tenderar att välja det maximala antalet tillåtna alternativ (n) i lika stora delar (1/n). I premiepensionssystemet innebär detta att fondsparare tenderar att välja ut fem fonder och placera 20 % av sitt kapital i respektive fond.
- **Home bias** är ett uttryck för preferensen att välja det välkända för att minimera riskerna hos obekanta alternativ, trots att de obekanta alternativen kan vara bättre. Inom premiepensionssystemet visar Nordén med flera, hur fondsparare tenderar att välja svenska fonder trots att de kan få en bättre avkastning och diversifiering om de använder sig av hela den tillgängliga marknaden (Nordén, 2010).
- **Historisk avkastning** kan ses som en förenklad prognosmetod. Om någonting har varit bra förut så är det rimligt att tro att det även kommer vara bra i framtiden. Antagandet om historisk avkastning strider mot den effektiva marknadshypotesen och Nordén berättar att personer med en övertro på sin egna finansiella förmåga tenderar att välja investeringar baserat på historisk avkastning. (Nordén, 2010).

Eftersom frekventa byten, skillnader i finansiell kunskap och användandet av förenklade beslutsmekanismer potentiellt kan påverka den pensionen för många svenskar på ett märkbart sätt så blir det intressant att utreda konsekvenserna av dessa faktorer.

Syfte

Att undersöka effekten på avkastningen hos olika förenklade beslutsmekanismer i premiepensionssystemet. De beteenden vi studerar är:

- 1/n bias – Hur påverkas resultatet av att spararen sprider sina risker genom att dela upp sin portfölj i fem lika stora delar.
- Home bias – Hur påverkas resultatet av att spararen placerat i svenska fonder.
- Aktivitet – Hur påverkas resultatet av att spararen aktivt byter sin fondportfölj.
- Historisk avkastning – Hur påverkas resultatet av att spararen använt historisk avkastning vid valet av fonder.

Vi studerar även skillnaderna i resultat mellan olika ålderskohorter, med utgångspunkt från skillnader i finansiell kunskap.

Avgränsningar

Forskningen visar att även på att faktorer som genetik, overconfidence och financial illiteracy har en stor inverkan på vår finansiella beslutsfattning men dessa kan inte undersökas i vår data. Författarna uppmanar andra forskare att undersöka dessa områden närmre.

Existensen av olika beslutsbeteenden kommer ej att visas i denna rapport, detta anser vi visat av tidigare författare.

Teoretisk referensram

En genomgång av de teoretiska verktyg och koncept som används i rapporten. Fokus ligger på studie av olika bias, aktivitet i sitt sparande, finansiell kunskap, differentiering och den effektiva marknadshypotesen.

Differentiering

Butler och Domian (1993) presenterar i *Financial services review* effekten av differentiering över tiden. De jämför avkastningen på amerikanska statskuldsväxlar, statliga obligationer, företagsobligationer samt stora och små aktier. I artikeln presenterar de resultatet från olika studier på avkastningen och standardavvikelsen på de olika värdepappren. I enlighet med många andra studier visar det sig att aktier, i synnerhet små aktier, har högst volatilitet men även högst avkastning. Gemensamt för studierna är att slutsatsen av dem är att, den högre volatiliteten till trots, är det väldigt liten chans att obligationer genererar en högre avkastning över tiden. Över ett tidsspänn på 30 år är det bara lönsamt i fyra procent av fallen att använda sig av en diversifierad portfölj bestående av obligationer och aktier. Över en kortare tidsintervall har värdepapprets volatilitet en större betydelse, där presterar de båda aktieslagen sämre än de tre stabilare värdepappren sämre i mer än fem procent av fallen. I mer än 75 procent av fallen presterade dock aktierna bättre än övriga, stora aktier presterade något bättre än små aktier. Skillnaden mellan små och stora aktier vid den 25:e percentilen var dock mycket liten och majoriteten av de som valde att spara i små aktier hade bäst avkastning över tio-årsperioden. I samma artikel refererar författarna till en studie

Leibowitz och Langetieg (1989) som har jämfört den långsiktiga avkastningen för portföljer med sju procents riskpremium med portföljer med fyra procents riskpremium mellan åren 1926 och 1989. Slutsatsen av den studien var att portföljer med ett lägre riskpremium presterade betydligt sämre över en lång tidsperiod än de med ett högre riskpremium.

Palmes et al. (2005) rekommenderar, i likhet med Butler och Domian, att en investerare bör sänka riskexponeringen i sitt sparande när tiden till pension minskar. Dock visar de tidigare studie att det motsatt förhållande råder. Äldre investerare tar större risk än yngre investerare.

Bali, Demirtas, Levy och Wolf har utfört en utförlig studie av investerares ålder och deras risktagande. De slår fast att yngre investerare bör ha en högre riskexponering än äldre investerare. De nämner också det faktum att den amerikanska motsvarigheten till det som i Sverige benämns generationsfonder (life cycle funds) är det snabbast växande sparformen i USA. Sparandet i dessa fonder har fördubblats mellan åren 2000 och 2009. Det finns dock forskning som inte går i linje med att riskexponeringen skall öka med tidshorisonten. Merton och Samuelson (ref i Bali et al., 2009) hävdar att oavsett tidshorisont kommer obligationer alltid vara en bättre investering än aktier. Deras modell bygger på att investeringens tidshorisont inte har någon betydelse. Dessa två sätt att fatta beslut skiljer sig åt på grund av att de som förespråkar aktier på lång sikt tar hänsyn till tidshorisonten medan de som förespråkar obligationer inte tar hänsyn till tidshorisonten. Genom att ta med tidshorisonten som en del i beslutet har Bali et al. (2009) har kommit fram till att vid en investeringshorisont över 24 månader är aktier ett bättre alternativ. Om horisonten ökar till 60 månader är endast portföljer med minst 60 procent aktier effektiva.

Vidare drar de slutsatsen att om en portfölj endast ska bestå av 100 procent aktier eller 100 procent obligationer är det endast portföljen med 100 procent aktier som är effektiv på lång sikt. Istället anser de att svårigheten snarare ligger i att bestämma hur fördelningen mellan aktier och obligationer ska vara i portföljen. Även i denna fråga kommer Bali et al. (2009) till slutsatsen att det är en direkt funktion av tidshorisonten. Vid en lång tidshorisont bör andelen aktier vara dominant för att sedan minska allt eftersom tidshorisonten blir kortare. Även vid kortare tidshorisonter är aktieportföljer ett mer effektivt val. Avslutningsvis konstaterar de att genom att applicera samtliga nyttofunktioner kan ingen garanti för att aktier är ett bättre val göras, dock sträcker de sig så långt som att genom att tillämpa alla ekonomiska beslutsregler kommer aktier leda till större nytta än obligationer.

1/n bias

Många studier av svenskarnas beteende vid valet av premiepensionsfonder har gjorts sedan reformen genomfördes. Majoriteten av de två tredjedelar som gjorde ett aktivt val, valde att främst investera i aktiefonder. Det mest populära alternativet var att välja maximalt tillåtna antal fonder (fem stycken). Det var även populärt att välja fonder av olika kategorier och fonder som förvaltades av olika förvaltare. Detta går alltså mot det faktum som bevisas i många andra studier att investerare har en för låg grad av diversifiering. Detta anser Hedesström, Svedsäter och Gärling (2007) bero på att rekommendationer och utförlig information om de olika fonderna bifogas i broschyren som delas ut i samband med premiepensionsvalet. Detta skiljer sig från andra former av sparande då en investerare ofta får tips från någon närstående eller en expert angående en aktie.

Psykologiska tester har visat dock att detta kan ha en helt annan förklaring än att spararna får information i broschyren. Vid tillfällen då personer får ett visst antal valmöjligheter vid ett speciellt tillfälle tenderar de att välja fler alternativ än vad de skulle gjort om de fick dessa valmöjligheter vid flera tillfällen. Detta psykologiska fenomen är känt som 1/n heuristik, eller 1/n bias. Detta gäller vid

olika typ av beslut inte enbart finansiella beslut. Såväl Hedeström et al., som Benartzi och Thaler beskriver detta som en naiv diversifiering snarare än som en positiv diversifiering. En trolig förklaring till detta fenomen är när människor presenteras för flera olika val vid ett och samma tillfälle söker de variation på grund av osäkerhet och riskundvikande. Vid ett finansiellt beslutsfattande är denna typ simultana val optimala endast om investerarna har full förståelse för hur valen korrelerar med varandra. Som exempel på detta tar Hedeström et al. (2007) upp det faktum att många av de svenska investerarna väljer att kombinera generationsfonder med aktiefonder. En mix som gör det mycket svårt att kontrollera kovariansen mellan valen, då fondförvaltarna inte har någon information om hur de andra fonderna investerar och vice versa (Hedeström, Svedsäter och Gärling, 2007). För att minimera risk genom diversifiering är det viktigt att kovariansen mellan investeringarna inte är stark. Starkast kovarians finns mellan investeringar på samma marknad eller samma bransch (Hedeström et al., 2007).

Tester har gjorts för att kontrollera allmänhetens kunskap om kovarians. Resultatet av dessa tester ger stöd för Shanks (ref i Hedeström et al., 2007) slutsats, att många investerare saknar kunskap om kovariansens påverkan i finansiella beslutssituationer. Även andra studier har gjorts för att testa kunskapen om värdepappersportföljer och dess kovarians. Vid en studie med tre grupper psykologistudenter fattade samtliga grupper sina beslut utan att hänsyn till portföljernas kovarians, så kallad *covariation neglect*. Samma studie genomfördes senare med ekonomistudenter, dessa tog viss hänsyn till kovariansen. Psykologistudenterna tenderade att diversifiera genom 1/n heuristik. Detta var även fallet då uppgiften i testet ändrades till att endast syfta till riskminimering. Även de som valde fonder från olika marknader i syfte att riskminimera sitt sparande gjorde det helt utan att ta hänsyn till kovariansen mellan fonderna. De mest risk averse individerna valde flest fonder men inte heller de tog hänsyn till kovariansen mellan fonderna, de tenderade dessutom att inte diversifiera mellan olika branscher, endast mellan olika marknader. Vid val av bransch premierades fonder som inriktade sig på branscher som investerarna bedömde som minst riskfyllda de närmsta fem åren. De som valde generiska fonder (som riktar sig mot många olika industrier) var istället risksökarna. Hedeströms et al. slutsats är att investerare bevisligen saknar kunskap om varför och hur en portfölj bör diversifieras.

Home bias

Med home bias menas i detta fall att många investerare har en förkärlek till att investera på den inhemska marknaden. På de fem största aktiemarknaderna i världen är mer än 90 procent av investeringarna inhemska. Detta trots att fördelarna med internationell differentiering anses som odiskutabla. Bailey, Kumar och Ng (ref i Nordén, 2010) i en studie från 2008 konstaterar att en internationellt diversifierad portfölj eliminerar risker bättre än en portfölj bestående av endast inhemska aktier. Dock visar studien att detta endast är sant om den inhemska portföljen har en låg diversifieringsgrad och består av mindre än tolv tillgångar. För en investerare med en inhemsk, väl diversifierad portfölj finns inga ytterligare fördelar i att även diversifiera internationellt. Inte heller avkastningen tycks skilja väl diversifierade portföljer som antingen innehåller internationella investeringar eller endast inhemska investeringar. Nordén diskuterar även att i portföljer där frekvent bytande av tillgångar förekommer, är avkastningen sämre än i portföljer där frekvensen är lägre.

Båda dessa beteenden (att ha en fallenhet för home bias och att frekvent byta ut tillgångarna i portföljen) förklaras av samma psykologiska fallenhet, *competence effect*. Mer kompetenta investerare tenderar att handla med värdepapper mer frekvent och dessutom ha mer internationellt diversifierade portföljer. Det mer frekventa handlandet beskrivs som en jakt på högre avkastning och *the competence effect* är snarare en effekt över en övertro på sin egen förmåga. Precis som Cesarini et al. (2003) konstaterar Nordén av beteende av denna övertro snarare leder till lägre avkastning.

Nordéns studie (2010) visar att de investerare som har en fallenhet för home bias ofta har en betydligt högre riskexponering än de individer som även investerare i utlandet. Det betyder dock inte att deras avkastning är sämre. Samma studie visar även att de investerare som har en fallenhet för home bias är mer långsiktiga i sitt investerande vilket gör att deras avkastning ofta överstiger de investerare som har en övertro på sin kunskap och därmed investerar internationellt men även byter ut sina investeringar ofta. Korrelationen mellan home bias och övertro är mycket svag (Nordén, 2010). Däremot finns det en tydlig korrelation mellan de som arbetar inom tillverkningsindustrin och en fallenhet för home bias, medan anställda inom bank- och försäkringssektorerna i stor utsträckning diversifierar utomlands. Palme et al., menar att det med stor sannolikhet beror på en större finansiell kunskap.

Personer med en hög utbildningsnivå tenderar att diversifiera sina portföljer bättre än personer med lägre utbildningsnivå. Även en hög inkomst och om individen tror att andelen av den totala pensionen som kommer från premiepensionen är stor tycks leda till ett mer diversifierat sparande (Nordén, 2010).

Aktivitet

Det nya svenska pensionssystemet möjliggör, som tidigare nämnts, att löntagare har möjlighet att själva förvalta en del av sin framtida pension och de har möjlighet att byta ut sina fonder med önskad frekvens utan utökade avgifter. Statistik från PPM visar att frekvensen bland fondbyten skiljer sig mycket åt mellan individer. Enligt teorin om en övertro på sin egen förmåga är det främst de som är duktiga fondsparare och de som tror att de är duktiga som byter fonder mest frekvent, och att endast de som är duktiga tjänar på sitt frekventa bytande (Nordén, 2010).

Debatten om individer tjänar på att byta ut sina fonder ofta har spåtts på genom Dahlquist et al's., rapporter (2011). En jämförelse mellan aktiviteten i fonder i pensionssystemet och fonder utanför pensionssystemet visar att de som sparar i fonder utanför systemet byter ut sina fonder mer frekvent. Dessa rapporter skiljer inte på duktiga investerare och investerare som har en övertro på sin förmåga utan konstaterar att de som sparar i pensionsfonder inte tycks reagera på fondernas historiska avkastning. Detta innebär att pensionsspararna riskerar att fastna i fonder som presterar dåligt under en längre tid. Detta hävdar Dahlquist och Martinez (2011) beror långsamma reaktioner och dålig uppmärksamhet hos de svenska pensionspararna. Historisk avkastning definieras i detta sammanhang som en indikator på hur duktiga fondförvaltarna samt vad framtida avkastning kommer att bli. I detta fall är det alltså mer lönsamt att byta fonder ofta baserat på historisk avkastning. Detta antagande baseras på att byten till och från fonder utanför pensionssystemet ofta har en stark koppling till fondernas historiska avkastning och att investerarna i dessa fonder tenderar att jaga högre avkastning. Däremot visar studien att pensionsparare är okänsliga för fondernas historiska avkastning (Dahlquist och Martinez, 2011).

Dahlquist och Martinez (2011) menar att de som sparar i fonder inom pensionssystemet håller i genomsnitt kvar vid sina fonder 20-30 procent längre i underpresterande fonder än fondinvesterar utanför pensionssystemet. Dessutom tenderar pensionsspararna att fortsätta investera i fonder som presterar dåligt, till skillnad från andra investerare som är snabba att dra sig ur underpresterande fonder. Skillnaden i utflöden från underpresterande fonder är stor. Av de fonder som presterar dåligt är utflödet av investeringar sex procent större från investerare utanför pensionssystemet än från investerare inom pensionssystemet. Kassaflödena till och från fonden Robur Contura är ett bra exempel på detta. Robur Contura "Placerar i svenska och utländska "framtidföretag", dvs. företag som satsar på forskning och utveckling, tex data och läkemedel. Fonden är öppen för Premiepensionssparande" (www.avanza.se).

När svenska befolkningen skulle göra sitt första val år 2000 hade fondens värde ökat med 500 procent de senaste fem åren. Den höga avkastningen gjorde att det var svenskarnas mest populära val efter default-fonden. Efter detta första val bestod fondens förvalta kapital till tio procent av PPM-pengar. Mellan åren 2000 och 2003 minskade fondens värde med 70 procent (sjunde AP-fondens värde minskade med 30 procent under samma period). Detta innebar att Robur Contura under denna period tillhörde de sämst presterande fonderna. Under denna negativa period var de privata fondspararna mer uppmärksamma i det avseendet att de drog sig ur fonden i större utsträckning. Pensionsspararna valde att investera i Robur Contura i stor utsträckning även under åren 2000-2003. Detta anser inte Dahlquist och Martinez vara ett aktivt val vid just den perioden utan snarare var det så att de inte valde att ändra sina förvalda fonder, ett tecken på tröghet (inertia). Tio år efter att de svenska löntagarna gjorde sina första investeringar år 2000 är Robur Contura fortfarande en av de populäraste fonderna i PPM. År 2009 bestod fondens totala förvalta kapital till 60 procent av PPM pengar (Dahlquist och Martinez, 2011).

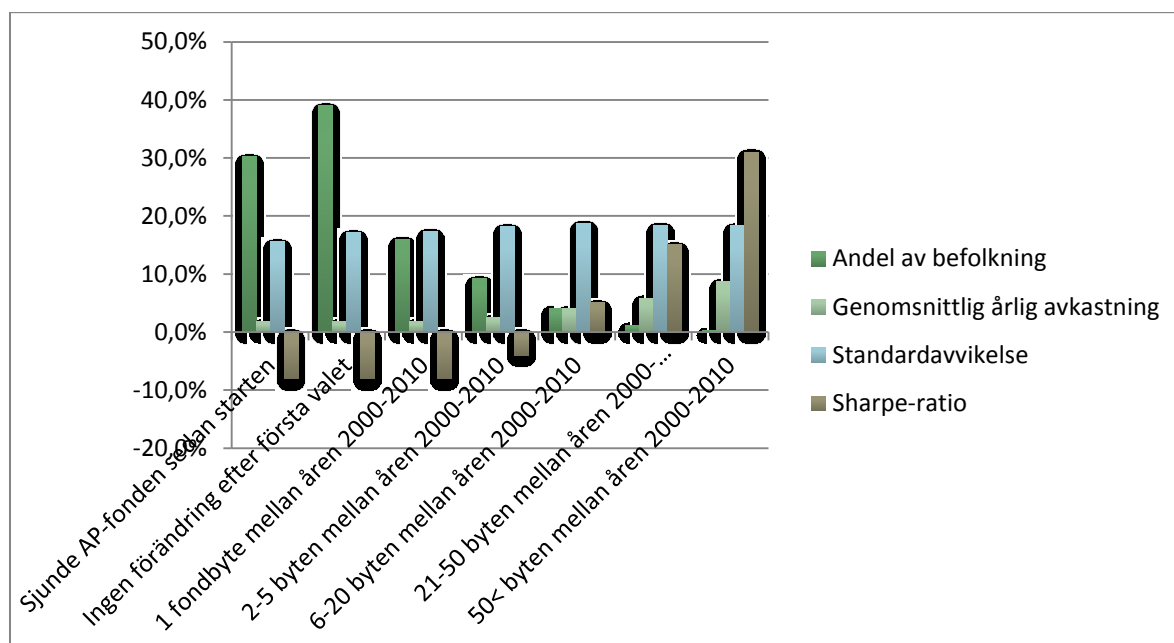
Privata fondsparare reagerar sex till åtta gånger mer på förändrad avkastning än pensionssparare. Denna mindre uppmärksamhet på låg historisk avkastning hävdar Dahlquist och Martinez (2011) leder till en sämre framtida avkastning. 9,18 procent av PPS-pengarna är investerade i någon av de 10 procent sämst presterande fonderna. Motsvarande siffra för privata sparare är 6,88 procent. 10,47 procent av fondspararna inom pensionssystemet har investeringar kontinuerligt går sämre än genomsnittsfonderna. Motsvarande siffra för de privata spararna är 8,50 procent (Dahlquist och Martinez, 2011).

I genomsnitt har en PPM-sparare som gjort ett aktivt val en gång för att sedan vara helt passiv en årlig avkastning på 1,7 procent. PPM-sparare som är mer aktiva än så har en genomsnittliga årlig avkastning på 2,5 - 8,6 procent. Ett frekvent bytande av fonderna visar på högre avkastning men en oförändrad riskexponering. Även vid användande av dummy variabler för att kontrollera om det verkligen frekvensen som är avgörande, visar tester att så är fallet. De som är mest aktiva visar även tendenser att göra ett bättre första val än de som är passiva. Aktiva investerare lyckas även oftare flytta sina pengar till bättre presterande fonder än de lämnade, medan de som är passiva ofta fastnar i de fonder som presterar allra sämst. En viktig detalj i detta sammanhang är att cirka tio procent av placeringarna i pensionsfonder sker i samråd med experter (motsvarande siffra för fonder i allmänhet är cirka 80 procent). Det finns även tecken på att ett väldigt aktivt förvaltande av pensionsfonderna kan ha en negativ effekt på de passiva spararnas avkastning (Dahlquist, Martinez och Söderlind, 2011). En stor andel aktivt förvalta pensionssparande tvingar förvaltare av icke-pensionsfonder till mer aktivitet vilket gör att de som är uppmärksamma och reagerar på förändringar kan få sämre avkastning.

Under perioden 2000 till 2010 hade 30,2 procent av befolkningen aldrig gjort något aktivt val och hade således haft sin samlade premiepensionspott hos sjunde AP-fonden. 39,0 procent hade gjort ett aktivt val år 2000 och placerat sina premiepensionspengar i en eller flera fonder men därefter inte uppdaterat sitt val. 16,0 procent hade ändrat sitt fondval en gång, 9,2 procent hade ändrat sina fondval mellan två och fem gånger, 4,1 procent mellan sex och tjugo gånger, 1,2 procent mellan 21 och 50 gånger och 0,3 procent mer än 50 gånger.

Vid ett typiskt fondbyte flyttar investeraren cirka 50 procent av sin totala portfölj. Vanligtvis sker bytet inom samma fondtyp (aktiefond, generationsfond och så vidare). Endast i tio procent av fallen sker bytet till fullo till en annan fondtyp. Män är mer aktiva än kvinnor, även högtbildade är överrepresenterade i kategorin aktiva sparare. Generellt sett presterar enskilda sparare sämre än marknadens genomsnitt (Dahlquist et al., 2011).

Den genomsnittliga avkastningen är direkt beroende av bytesfrekvensen. De som byter allra mest frekvent har också den högsta genomsnittliga avkastningen. De som byter minst frekvent har den absolut lägsta genomsnittliga avkastningen. Den enda avvikelser från den korrelationen är att de som aldrig gjort något val presterar bättre än de som gjort ett val år 2000 för att sedan hålla kvar vid det valet.



Som synes i diagrammet ovan sker den ökade genomsnittliga avkastningen utan en högre volatilitet, vilket resulterar i ett bättre Sharpe-ratio (Dahlquist et al., 2011).

Den effektiva marknadshypotesen

Enligt den effektiva marknadshypotesen består en effektiv kapitalmarknad av säkerheter där priserna justeras blixtnabbt så fort ny information tillkommer. Av den anledningen återspeglar det aktuella värdet all information som finns om säkerheten. Tre krav måste uppfyllas för att marknaden ska betecknas som effektiv:

- 1) Ett flertal vinstdrivande och konkurrerande deltagare analyserar och värderar, oberoende av varandra, samtidigt
- 2) Ny information når marknaden slumpvis
- 3) De konkurrerande investerarna strävar efter att justera säkerheternas priser snabbt för att återspegla all tillgänglig information

Eftersom den effektiva marknadshypotesen återspeglar all tillgänglig information och att alla investerare är rationella är den enda skillnaden på priset på säkerheten mellan t och $t+1$ orsakat av händelser som ligger i framtiden och inte kan förutses. Det finns många studier där författarna hävdar sig bevisa att hypotesen håller, men även väldigt många där författarna hävdar sig bevisa det motsatta. Ett flertal nyare rapporter (däribland Jagric et al., 2005) hävdar dock att den effektiva marknadshypotesen inte håller på kort sikt. Hypotesen får dock visst stöd på lång sikt kan den hålla (Jagric et al., 2005). Att den håller på lång sikt men inte på kort kan tyckas oväsentligt då poängen med hypotesen är att priserna justeras blixtnabbt, och därmed endast kan kontrolleras på kort sikt.

Den effektiva marknadshypotesen och prognoser

Som förklarat ovan är det omöjligt enligt den effektiva marknadshypotesen att spekulera i framtida värden. Om framtida priser på värdepapper skulle gå att förutse skulle det leda till obegränsad avkastning. De som försöker använda historisk data för att förutse framtida avkastning misslyckas med största sannolikhet. Den effektiva marknadshypotesen spelar dock en stor roll på den framtida avkastningen och kan inte helt förbises. Detta beror inte nödvändigtvis på att den är till fullo applicerbar utan på att vissa investerare använder historisk data för att prognosticera framtida avkastning och det påverkar priset på säkerheterna.

Sammanfattningsvis slås om den effektiva marknadshypotesen fast; "If the behavior of investors produces efficient markets by their continuous profit seeking, the reverse is that the EMH (Efficient market hypothesis, reds anm) does not rule out predicting many other variables that, although of general interest, are not the basis for a profit making strategy". Om en investerare kan hitta nya mönster och samband mellan en ny variabel och prisförändringar kan denne tillämpa detta som en prognos för framtiden. Om detta ska ske i enlighet med den effektiva marknadshypotesen krävs dock att endast ett fåtal känner till detta sätt att prognosticera. Om inte så är fallet kan inte historisk data användas för att "besegra" marknaden. Om alla investerare har tillgång till samma information kommer det nuvarande priset justeras blixtnabbt för att ligga i relation till det framtida priset (Timmermann och Granger, 2004).

Historisk avkastning och overconfidence

Teorin om overconfidence, eller övertro, utgår från det välkända faktum att alla individer inte är helt rationella. Genom att hålla detta som sant utgår forskningen från att individuella investerare söker maximera sin nytta enligt sina preferenser och åsikter. Detta resulterar i investerare med en övertro på sin förmåga ofta gör mer riskfyllda investeringar och är mer delaktiga i värdepappershandeln än rationella individer. Overconfidence kan liknas vid ett själv-tillkännings bias där ett dåligt resultat beror på otur och ett bra resultat beror på skicklighet (Palomino och Sadrieh, 2011).

Det finns likheter i beteende mellan de som har en övertro på sin egen förmåga och de som är skickliga. Båda dessa personligheter tenderar att byta innehåll i sina portföljer relativt ofta. Det som skiljer dem åt är främst att de som har en övertro snarare än skicklighet väljer fonder främst baserat på en hög historisk avkastning, vilket inte tycks vara en viktig parameter för skickliga investerare (Nordén, 2010). Detta får till viss del stöd Ekholm och Pasternack (2007) även om de hävdar att storleken på transaktioner är ett tydligare tecken på overconfidence. De simplificerar det något och hävdar att alla småsparare är overconfident, medan de stora investerarna (alternativt de finansiella instituten) är skickliga. Denna slutsats baseras på att små aktörer verkar reagera starkare på positiva nyheter än negativa nyheter. Små investerare köper, med deras mått mätt, mycket när en aktiekurs (eller fondkurs) är på väg uppåt, men när det kommer negativa nyheter och kursen vänder nedåt är de inte alls lika uppmärksamma. Detta stödjer alltså Dahlquist et al's. (2011) teorier om att pensionssparare (som definieras som småsparare) har en tendens att fastna i underpresterande fonder. Däremot strider det helt mot Dahlquist et al's. (2011) teorier om att en positiv historisk avkastning tyder på en skicklig förvaltare och därmed är det fördelaktigt att välja fonder baserat på historisk avkastning.

Genetik och kunskap må förklara en hel del av hur vårt finansiella beslutsfattande ser ut. Men även andra personlighetsdrag spelar en stor roll. Självförtroende är ett av dessa drag. En övertro på sin egen finansiella förmåga tenderar att minimera det rationella beslutsfattandet. En övertro pekar, enligt forskning, på två speciella särdrag i det ekonomiska beslutsfattandet. Det ena särdraget är en överdriven tro på sin egen förmåga. Det andra särdraget är att dessa personer grovt överskattar sin egen kunskap, alltså en avsaknad av metakognitiv förmåga (Cesarini, Sandewall och Johannesson, 2003).

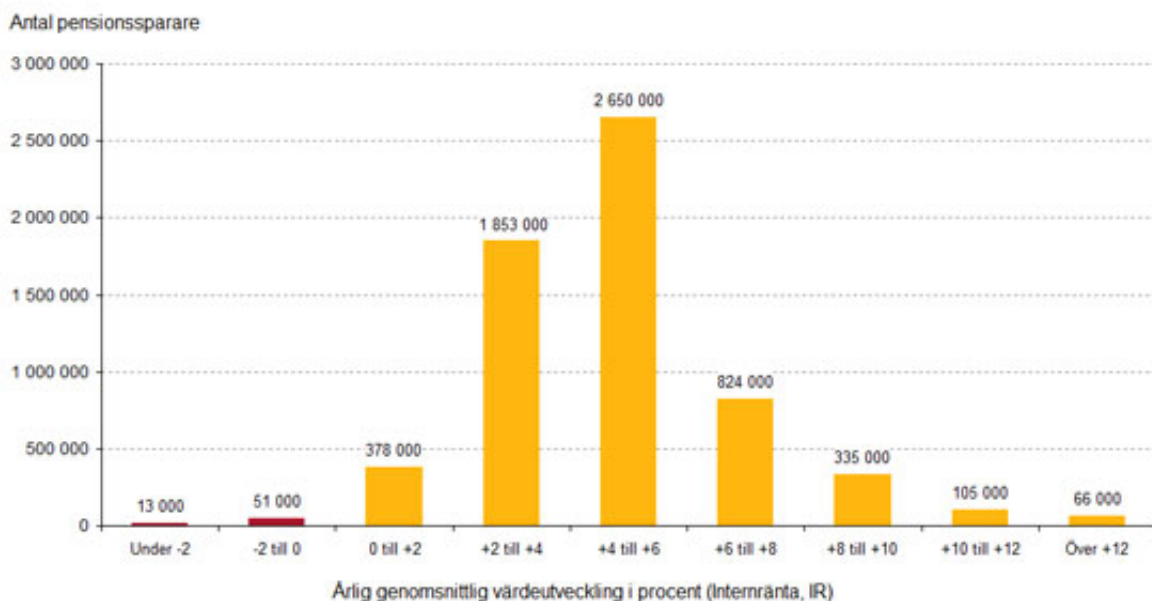
Financial illiteracy

En individs räknefärdighet, i synnerhet inom finansiella områden, har en stor betydelse för hur individen fattar finansiella beslut. Det finns dessutom ett starkt samband mellan en individs räkneförmåga och dess finansiella kunskaper. Denna typ av kunskap, som till stor del bygger på kognitiv förmåga, är sämst hos de yngsta och de äldsta i populationen. Att de äldsta i samhället presterar lågt på denna typ av tester förklaras av att kognitiv förmåga avtar med åldern. Att de yngsta i samhället presterar dåligt är mer förvånande och förklaras av omgivningen, att undervisningen i skolan har mindre fokus på matematik samt andra miljöfaktorer (Almenberg och Widmark, 2011).

Personer med god matematisk och finansiell förmåga tenderar att ta större finansiella risker. Dessa personer är dessutom mer delaktiga på aktiemarknaden. Denna korrelation står fast även vid jämförelser med dummy variabler som inkomst, utbildning och kön. Däremot finns tydliga samband mellan dessa variabler och räknefärdighet. Män i medelåldern med hög inkomst och lång utbildning presterade bäst på Almenberg och Widmarks tester. Ett liknande samband fanns för finansiell förmåga.

Finansiella beslutsfattande påverkas av många olika faktorer. Den frekvent recenserade professorn Henrik Cronqvist presenterar i sin studie *nature or nurture: What determines investor behavior* (2010) hur stor del av vårt finansiella beteende beror på genetik och hur mycket som beror på omgivning. Studien baseras på agerandet hos 37 504 svenska tvillingar i olika finansiella situationer där tre huvudsakliga mått belyses; delaktighet på aktiemarknaden, hur stor andel av sparandet som sker i form av aktier samt portföljernas volatilitet.

Cronqvist et. al. (2010) hävdar sig bevisa att det korsrelaterade beteendet mellan delaktighet i aktiemarknaden och hur stor andel av det totala sparandet som sker i form av aktier bestäms till cirka en tredjedel av genetik. Denna genetiska faktor påverkas inte av omgivningen och avtar inte heller med åldern, detta bevisas genom att beteendet inte skiljer sig mellan tvillingar som har frekvent och infrekvent kontakt. Vidare skriver de att omgivningen även den har en stor påverkan, men att denna påverkan sker utöver denna tredjedel och snarare är förklaringen till varför familjemedlemmar agerar olika i finansiella situationer. Omgivningens påverkan är dessutom kraftigt avtagande med tiden. Resultaten av denna studie har kontrollerats mot andra faktorer som anses ha påverkan på vårt finansiella beteende, till exempel utbildningsnivå och förmögenhet, med resultatet att den genetiska påverkan inte förändras. Även valet av volatilitet i portföljen anses bevisas till cirka en tredjedel av genetiska faktorer (Cronqvist et al., 2010). Detta är en av de faktorer som förklarar stora skillnader inom grupper som, bortsett från genetik, anses vara homogena.

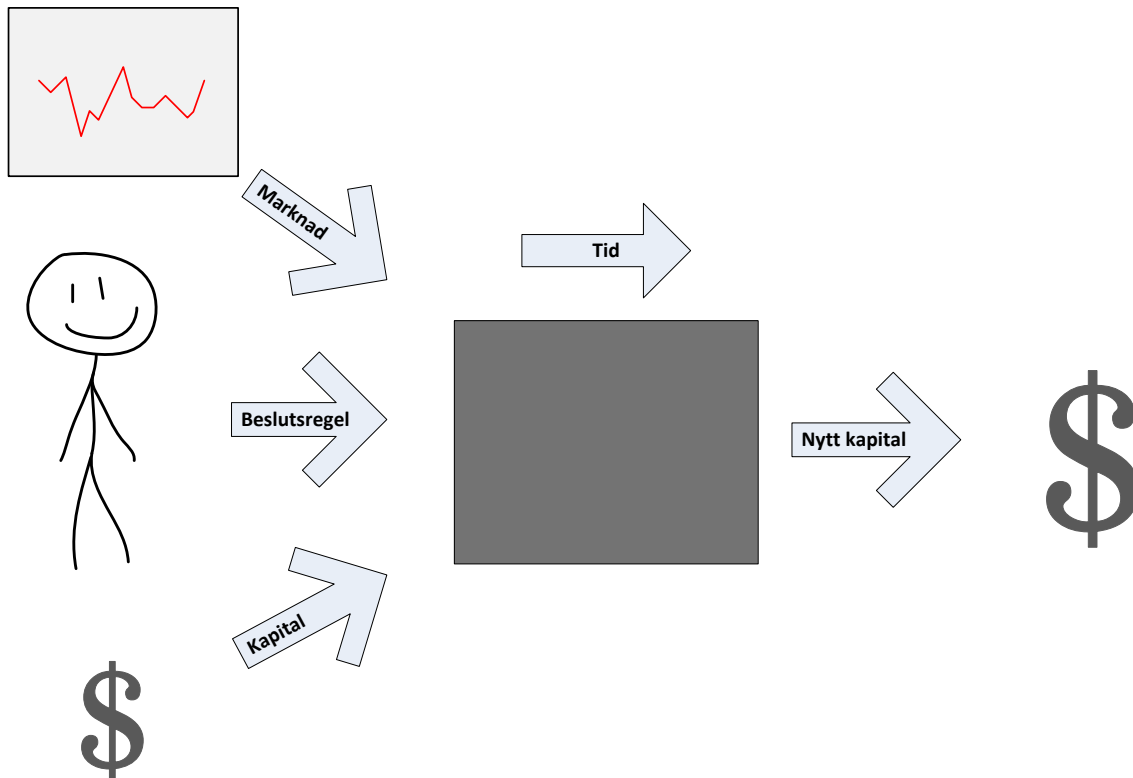


www.pensionsmyndigheten.se

Metod och modell

Vår modell bygger på att spararnas resultat förklaras helt av:

- Hur marknaden i stort utvecklas
- De beslutsregler som påverkar fondvalet
- Hur mycket pengar som investerats



Nytt kapital är resultatet av att låta kapitalet (p) vara investerat under perioden T med räntan r .

$$\text{Nytt kapital} = pr^T$$

Eftersom enskilda avkastningen hos investeringar varierar över tiden så skrivs uttrycket om för att spegla detta.

$$pr^T = p \prod_{t=1}^T r_t = p(r_{t=1} * r_{t=2} * r_{t=3} \dots * r_{t=T})$$

Antagande: Vår investering av kapital är förhållandevis litet och påverkar inte marknaden på något mätbart sätt. Kapitalets storlek påverkar alltså inte räntan och kan förkortas bort och räntan används som mått på hur väl investeringen lyckats.

$$\prod_{t=1}^T r_t$$

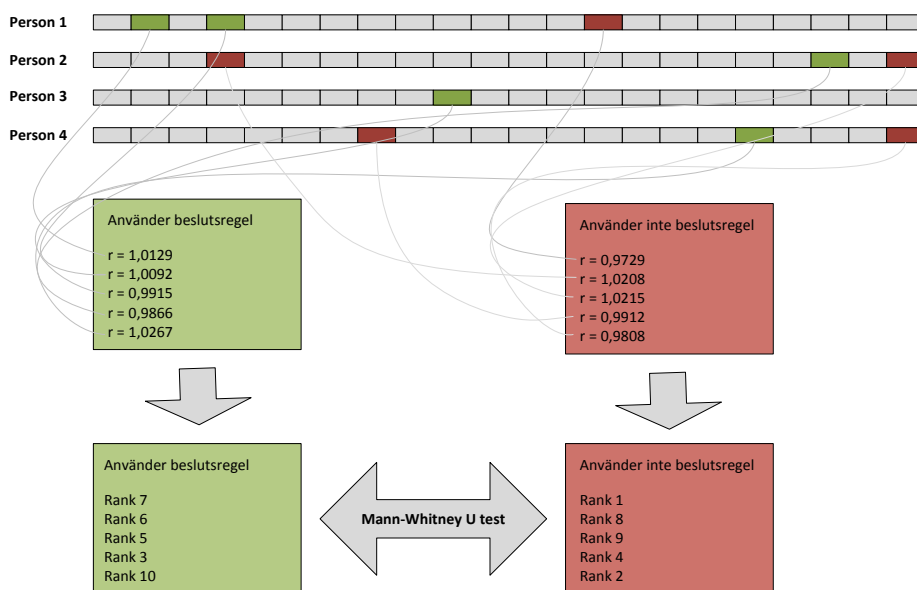
Antagande: Räntan på en hos en viss investering beror på den beslutsregel som används och på marknadens slumpmässiga rörelse. Om beslutsregeln har någon effekt så ger den upphov till en högre eller lägre ränta än marknaden i stort. För att se effekterna av en viss beslutsregel så jämförs avkastningen mellan personer som har använt en viss beslutsregel och de som inte har använt sig av den.

Eftersom individuella investeringar har olika tidshorisont så går de inte direkt att jämföra med varandra. För att få jämförbar data delas investeringen upp i veckor, där veckoräntan mellan olika investeringar jämförs med varandra. Exempel:

$$\text{Investering 1 mellan vecka 1 och 30: } \prod_{t=1}^{30} r_t = r_1 * r_2 * r_3 \dots r_{10} * r_{11} * r_{12} \dots * r_{30}$$

$$\text{Investering 2 mellan vecka 10 och 12: } \prod_{t=10}^{12} r_t = r_{10} * r_{11} * r_{12}$$

Veckorna 10, 11 och 12 innehåller här räntor som kan användas för att se effekten av beslutsregeln. Då totala avkastningen på en investering inte är beroende av ordningen mellan veckoräntorna på grund av kommutativitet; $x * y = y * x$ och associativitet; $x(y * z) = z(x * y)$ så är veckoräntorna för våra syften lika mycket värda oavsett var de befinner sig i tiden - de är alltså jämförbara med varandra.



Test av beslutsregel

För att testa effekten av en beslutsregel tas ett slumpmässigt stickprov av veckor ur vår datamängd. Stickprovets storlek är vanligen 10000 - 100000 datapunkter vilket motsvarar någon procent av den totala datamängden på ca. 7 miljoner punkter. Eftersom punkterna är glest spridda inom stickprovet bör de vara statistisk oberoende av varandra. *Resonemanget är alltså att en beslutsregel har effekt om dess veckoavkastningar i genomsnitt avviker från andra investeringars veckoavkastningar.*

Problem med modellen

Genomgång av de brister i modellen som har identifierats.

Icke linjära samband

Resultatet av en investering är som sagt resultatet av multiplikationen av de ingående veckoräntorna. För att illustrera de icke linjära sambanden så antar vi två likvärdiga investeringar under två veckor:

	Avkastning vecka 1	Avkastning vecka 2
Investering A	2 %	2 %
Investering B	1 %	3 %

Väntevärdet för båda investeringarna är 2 %, medan resultatet av de två investeringarna skiljer sig åt:

Resultat investering A: $1,02 * 1,02 = 1,0404$

Resultat investering B: $1,01 * 1,03 = 1,0403$

Investering A har alltså ett bättre resultat än investering B. I vår modell så bedöms dock dessa investeringar som likvärdiga trots att de egentligen inte är det. Denna avvikelse blir större ju större ränteskillnaden är och ju längre tidshorisonten är, vilket ger ett systematiskt fel i modellen där portföljer med en hög volatilitet blir lite högre värderade relativt fonder med låg volatilitet. Vi har ändå valt att gå vidare med denna modell eftersom felet är förhållandevis små vid våra räntenivåer och att alternativet att göra parvisa jämförelser är tekniskt komplicerat.

Kapitalets inverkan på resultatet

Då det investerade kapitalet ökar efter hand som en person bygger upp sitt pensionsgrundande kapital, så spelar veckoresultatet för individen större roll ju senare i tiden de är, eftersom räntan då får verka på ett större kapital. Syftet med undersökningen är att jämföra resultatet av olika beslutsregler oavsett när i tiden de sker; det är så att säga bättre att vara bättre än genomsnittet, oavsett om det sker i början eller slutet av investeringens horisonten.

Mycket små skillnader som mäts

Det är troligt att spridningen i individuella resultat är mycket stor jämfört med de effekter vi kommer mäta. Även konjunktur och enskilda fluktuationer kommer spela en stor roll. Risken för typ II fel, dvs. att det finns en skillnad som vi inte kan upptäcka är relativt hög, trots den stora mängden data. Inom den studerade tidsperioden ryms både en hög- och en lågkonjunktur. Det är dock svårt att säga om resultaten från perioden kan generaliseras.

Validitet

Validiteten är hur nära sanningen undersökningen kommit. Vid en jämförelse mellan resultatet hos de som uppvisar ett visst beteende och de som inte uppvisar beteendet så är skillnaderna ofta mycket små. Detta gör att känsligheten för systematiska fel är hög. Även om de statistiska testen görs med mycket stora urval så slår de systematiska felen igenom ändå eftersom de inte tar ut varandra på samma sätt som slumpmässiga variationer. Med bakgrund av detta är det främst två problem som påverkar validiteten i vår undersökning negativt:

Informationen om nedlagda fonder är ofullständig. Med tanke på att ca 4 % av veckoresultaten saknas på grund av att avkastningsdata för fonden saknas så är tappet ganska omfattande. Det är förmodligen rimligt att anta att flera av fonderna lades ner på grund av att de hade en sämre utveckling än genomsnittet och att detta drar troligen ner resultatet för de som gjort aktiva val.

Vår modell är förmodligen en allt för enkel modell som inte hanterar icke-linjära samband som avkastning över tid. Räntan ligger generellt sett nära 1 så avvikelserna blir ganska små. Konsekvenserna av att använda icke-parametriska metoder är svåra att förutsäga, men troligen minskar de effekten av extrema värden och minskar då även det icke-linjära felet. Felet borde främst drabba fonder med låg varians som undervärderas relativt andra fonder. Troligtvis påverkar detta mest äldre sparare eftersom de tenderar välja mer räntefonder med en lägre varians.

Vi bedömer validiteten som bristfällig.

Reliabilitet

Reliabiliteten är förmågan hos undersökningen att reproducera konsistenta resultat mellan flera försök. Indata till undersökningen kommer från väl etablerade källor som har mycket god kompetens i statistiska metoder (Pensionsmyndigheten och SCB). Stickprovet med individer speglar därför sannolikt populationen väl. Övrig grunddata, förutom avkastning för nedlagda fonder, bedöms som exakt. SQL-script för att genomföra databehandling innan körningarna i SPSS kan ha buggar och fel. Ett omfattande arbete har lagts på att testköra scripten och verifiera resultatet. De statistiska test som gjorts i SPSS är på stora stickprov. Vi har inte märkt skillnader i resultat mellan olika stickprov. Icke parametriska metoder har används vid analysen eftersom de ger ett robust resultat även då data ej är normalfördelat.

Vi bedömer reliabiliteten som relativt god.

Statistisk metod för att se effekten av en beslutsregel

För att se effekten av en beslutsregel A så används Mann–Whitneys U-test med hypoteserna:

$$H_0: \text{Strategin påverkar inte avkastningen, } P(R|A > R|\bar{A}) = 0,5$$

$$H_1: \text{Strategin påverkar avkastningen, } P(R|A > R|\bar{A}) \neq 0,5$$

Ett dubbelsidigt test används eftersom vi inte vet om en beslutsregel är bra eller dålig jämfört med övriga investeringar. Vi använder en signifikans på 5 % ($\alpha = 0,025$) för att avgöra när nollhypotesen kan förkastas.

Mann–Whitneys U-test har följande grundantaganden:

- **Slumpmässigt stickprov.** Materialet vi använder är ett stickprov från premiepensionsmyndigheten. De har personal som jobbar med statistik som försett oss med stickprovet så vi utgår från att stickprovet är representativt. Vi använder oss av hela materialet och tror inte att det skett några systematiska snedvridningar under databehandlingen.
- **Att olika värden hos den beroende variabeln kan rangordnas i storlek.** Veckoräntan är en variabel med nio decimalers upplösning vilket gör att även små skillnader kan mätas.
- **Den underliggande fördelningen behöver ej vara normalfördelad.**
- **Att antalet observationer inom respektive grupp är större än 10.** Vi har betydlig fler observationer inom varje grupp.

(Newbolt et al, 2007)

Statistisk metod för att mäta korrelation mellan olika variabler

För att mäta korrelationen mellan olika variabler används Spearmans rank-korrelation med hypoteserna:

H_0 : Ingen korrelation mellan X och Y , $r_s < r_{s,\alpha/2}$

H_1 : Korrelation mellan X och Y , $r_s > r_{s,\alpha/2}$

Där r_s är den observerade Spearman rankingen och $r_{s,\alpha/2}$ är tabellvärdet för en Spearman ranking vid signifikansnivån α .

Ett dubbelsidigt test används eftersom två variabler kan vara endera positivt eller negativt korrelerade med varandra. Vi använder en signifikans på 5 % ($\alpha = 0,025$) för att avgöra när nollhypotesen kan förkastas.

Spermans rank-korrelation har följande grundantaganden:

- Att olika värden hos variablerna kan rangordnas i storlek.
- Den underliggande fördelningen hos variablerna behöver ej vara normalfördelad.
- Att förhållandet mellan de två variablerna är monotont, dvs. att de stiger eller sjunker tillsammans, alternativt när den ena variabeln stiger så sjunker den andra.

(Newbolt et al, 2007)

Studerad tidsperiod

Den studerade tidsperioden i den undersökningen är mellan 2001-12-27 och 2010-12-30. Vi kommer även kontrollera olika om olika samband beter sig olika vid en period av positiv respektive negativ utveckling på fondmarknaden för att försöka se om det finns skillnader i utfallet.

Inflationsjustering

Veckoräntan är inflationsjusterad för att spridningen skall vara jämförbar under olika tidsperioder. Som mått på inflationen har konsumentprisindex på månadsnivå från SCB används (SCB, <http://www.ssd.scb.se>). Eftersom KPI informationen har varit på månads basis och våra fondavkastningar varit på veckobasis så är approximerat KPI med linjär interpolation:

När veckan ligger helt inom en månad:

$$KPI_{vecka} = KPI_{månad}^{\frac{1}{4}}$$

När veckan delas mellan två månader:

$$KPI_{interpolerat} = (KPI_{månad 1} * dagar_{månad 1} + KPI_{månad 2} * dagar_{månad 2}) / 7$$

$$KPI_{vecka} = KPI_{interpolerat}^{\frac{1}{4}}$$

Riskfri ränta

Den riskfria räntan är ej dragen från avkastningen. Den riskfria räntan används ofta i modeller som ett uttryck för alternativkostnaden hos en riskfri placering. Investeringarna i premiepensionssystemet är låsta och spararen har ingen möjlighet att placera dem i en i någon riskfri placering.

Förvaltningskostnader

Fondernas avkastning är minskad med deras förvaltningskostnad så att avkastningen mellan dem är jämförbar.

Algoritm för beräkning av 1/n bias

I diskussionen om 1/n bias berättar Thaler och Benartzi (2007) om hur människor som ges ett komplext val förenklar valet genom att välja ut så många alternativ (n) de kan och fördela dessa val i lika stora andelar (1/n) för att sprida risken. Eftersom fondsparare i premiepensionssystemet kan fördela sina pengar i upp till fem fonder så definieras 1/n bias som att ha valt fem olika fonder i lika stora delar (5 st. á 20 %):

$$\text{Har } 1/n \text{ bias: } a_1 = a_2 = a_3 = a_4 = a_5 = 20$$

$$\text{Har ej } 1/n \text{ bias: Övriga fördelningar hos } a_1 \dots a_5$$

Där $a_1 \dots a_5$ är fondandelen i procent för fondvalen 1 till 5. (Citering Nudge/Cronqvist?)

Algoritm för beräkning av home bias

Home bias mäts vanligen för aktieportföljer. Då anges home bias som den andel i individens aktieportfölj som har företag lokaliserade i hemlandet, dividerat med samma kvot hos marknaden i stort (Mavruk, 2010). Fonder innehåller i sig många olika slags värdepapper och att försöka beräkna hur stor andel av dessa värdepapper som kommer från Sverige är för tidskrävande för detta arbete. Vi har istället valt att använda oss av premiepensionsmyndighetens kategorisering av fonder. Om en fond är kategoriserad som svensk så anser vi att den är 100 % lokaliserad i Sverige. Andelen svenska fonder i pensionsspararens portfölj divideras sedan med den genomsnittliga andelen den veckan:

$$\text{Personlig andel svenskt} = pas = \frac{1}{100} \sum_{i=1}^5 S(\text{fond}_i) * \text{andel}_i_procent_i$$

$$S(\text{fond}) = \begin{cases} 1, & \text{Då fonden är kategoriserad som svensk} \\ 0, & \text{Då fonden ej är kategoriserad som svensk} \end{cases}$$

$$\text{Personens home bias} = \frac{pas}{\frac{1}{N} \sum_{i=1}^N pas_i}$$

Där N är antalet personer i stickprovet den veckan.

Det är värt att notera är att Premiesparfonden är kategoriserad som Svensk. Detta bör dock inte snedvrta jämförelserna vid övergången till AP7-Såfa (det nuvarande default alternativet) eftersom home bias är en kvot.

Algoritm för beräkning av aktivitet vid fondval

Vårt mått på aktivitet bygger på hur lång tid sedan fondspararen senast bytte fonder. Vi definierar en aktiv sparare som en person som har bytt fond de senaste 50 veckorna och som då inte har bytt till något av statens alternativ (eftersom vi inte kan skilja mellan personer som placeras i statens fonder och de som aktivt väljer statens fonder). Även Dahlquist et al. (2011) använder ett byte per år som mått för aktivitet.

Den verkliga aktiviteten hos en fondsparare är svår att mäta eftersom han kanske aktivt bevakar sina fonder utan att ändra fördelningen. Denna svaghet i vårt aktivitetsmått borde ge en lägre förklaringsgrad än den verkliga aktiviteten. Vi antar att det är relativt osannolikt att samma uppsättning fonder är optimala vid olika tillfällen och att den aktive spararen även byter fonder frekvent. I synnerhet då han inte behöver betala någon individuell transaktionskostnad. Vårt mått bör vara en god approximation för den verkliga aktiviteten.

Algoritm för beräkning av placering i fonder som under förra året var bland de 25 procent bästa fonderna

Fondval som endast innehåller de 25 procent högst avkastande fonderna från förra året är definierade som att de använt historisk avkastning som grund för sitt val. Personer som har en eller fler fonder som inte är bland fonderna med högst avkastning räknas som att de inte använt historisk avkastning.

Måttet är ganska trubbigt eftersom det egentliga skälet till fondvalet inte visas.

Algoritm för beräkning av ålderskohort

Ålderskohorter räknas ut genom att beräkna födelseåret och sedan placera personen i den ålderskohort som motsvarar födelseårets decennium.

Personer som fyller år under senare delen av november och december år nio i en ålderskohort kan ha blivit placerade i den yngre kohorten av misstag. Detta beror på att vi haft ålder som indata och inte födelseår, vilket gör att datumet då datat lästes ut ur ppm-systemet spelar roll. En ganska liten andel av populationen påverkas av detta fel.

Underlag

Underlaget till denna undersökning kommer från fyra datafiler som lästs in i en databas och bearbetats. Analysen har mestadels gjorts med SPSS medan kurvor med ackumulerad avkastning har skapats med SQL och Excel.

Datafiler

Fondbyten - Innehåller alla transaktioner för 16900 personer i premiepensionssystemet under perioden 2000-09-28 och 2011-11-22. Personerna är slumpvis utvalda från de ca 6 miljoner svenskar som finns i systemet (Pensionsmyndigheten, <http://www.pensionsmyndigheten.se>). Stickprovets kvalitet bör vara god eftersom det är samma datamängd som myndigheten själva använder. För att undvika de extra krav som ställs vid behandling av personuppgifter så är personnummer ersatta med ett unikt id-nummer på 10 siffror. Den enda bakgrundsvariabel i datat som kan knytas till en person är ålder. Strukturen på fondbyten är:

Person ID	Fond ID	Andel i %	Datum för byte	Ålder
10009988774321	110924	20	2002-02-14	68

Personer som avlidit är ej med i filen.

Avkastning - Innehåller avkastningen för alla fonder som är valbara under 2011. Avkastningen för fonderna är uträknade den ränta fonden haft sedan den föregående veckan. Fondens administrativa avgifter är dragna från veckoräntan. Datafilen kommer från Pensionsmyndigheten och innehåller tillförlitlig data. Dock saknas information för fonder som har lagts ner under perioden. Strukturen på avkastningsfilen är:

Datum	Fond ID	Veckoavkastning
2002-02-14	110924	1.021364000

Precisionen på veckoavkastningen är 6 decimaler. Veckor de fonden inte har handlats så är veckoavkastningen 1,0.

Fullständig avkastning - Innehåller avkastningen hos alla fonder under hela perioden. Även dessa filer kommer från Pensionsmyndigheten. Filerna innehöll dock dubletter i vissa fall. Vi har valt att kasta alla veckor som innehåller dubletter och använda de övriga. Vid en mer rigorös undersökning behövs fullständig avkastningsinformation för alla fonder under perioden. Konsekvensen av den ofullständiga fondinformationen är att vi har 95,38 procent av de teoretiskt möjliga datapunkterna. Det är svårt att förutse effekterna av det saknade datat. Möjligen har nedlagda fonder haft en sämre utveckling än genomsnittet. I så fall leder detta till ett systematiskt fel där de som gjort aktiva val i PPM-systemet fått ett för högt resultat i vår undersökning.

Filerna har samma format som avkastningsfilen för aktiva fonder.

Fondinformation - Innehåller fakta om alla fonder, som namn, värdepapperslag, förvaltare och den kategorisering som vi använder för att avgöra om en fond innehåller svenska värdepapper eller ej. Filen kommer från Pensionsmyndigheten och anses tillförlitlig. Relevanta fält för vår del är:

Fond ID	Fondnamn	Fondkategori
110924	Premiesparfonden	Svenska aktier och räntor

KPI - Innehåller konsumentprisindex på månadsbasis. Informationen kommer från Statistiska Centralbyrån och bedöms som tillförlitlig.

Databehandling

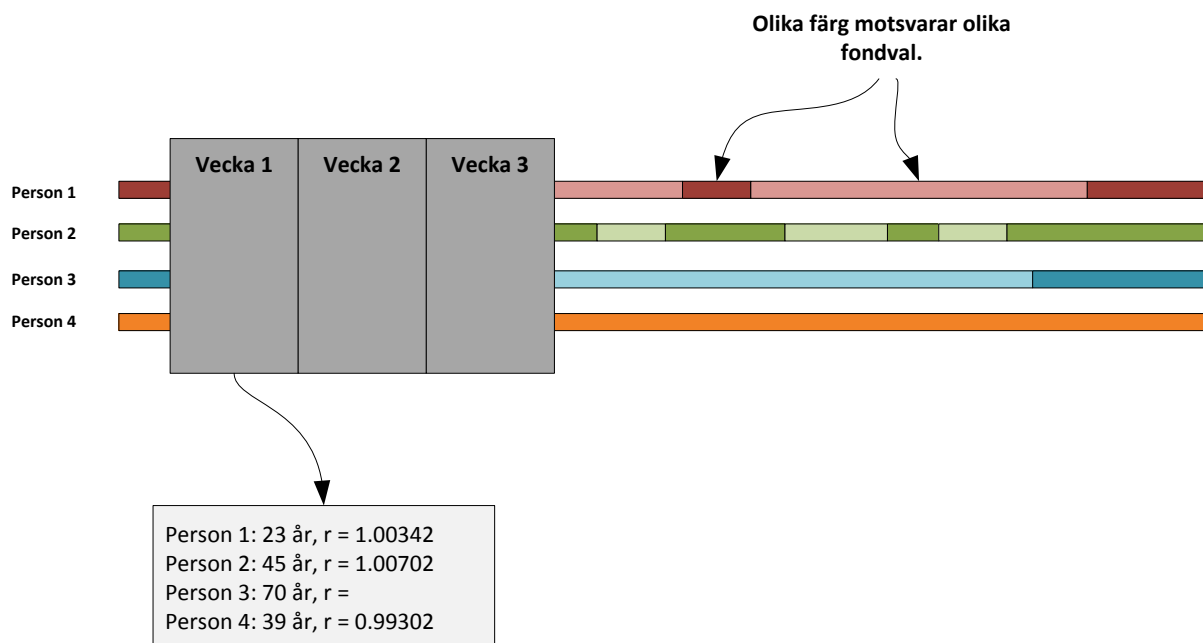
Schematiskt så har målet varit att få bearbeta datat från Pensionsmyndigheten och SCB så att den får följande format:

Person ID Veckonummer Ålder Beskrivande variabel Veckoavkastning

De beskrivande variablerna är en lista av de strategier, bias, ålderskohorter som är beskrivna ovan. Detta är ett format som sedan har varit ganska enkelt att analysera i SPSS. Eftersom datamängden är stor valde vi att använda en databas för databehandlingen. Programkod och inställningar finns i bilaga B.

För att komma till det önskade dataformatet så har vi transformerat transaktionsloggen från Pensionsmyndigheten i två steg:

1. Transaktionsloggen innehåller en rad per händelse, ett val av en portfölj av fonder var alltså uppdelad på flera rader. I första steget görs transaktionsloggen om till en lista av fondval. Fondvalen innehåller hela personens uppsättning av fonder och fördelningarna mellan dem på en rad.
2. För att kunna jämföra avkastningar mellan fonder behövde tidshorisonten på placeringarna normaliseras. Vi valde att göra detta genom att dela upp varje placering i dess komponenter, i form av de hela veckor, som placeringen bestod av.



När detta var gjort kunde vi jämföra resultatet mellan olika strategier på en standardiserad tidshorisont på en vecka.

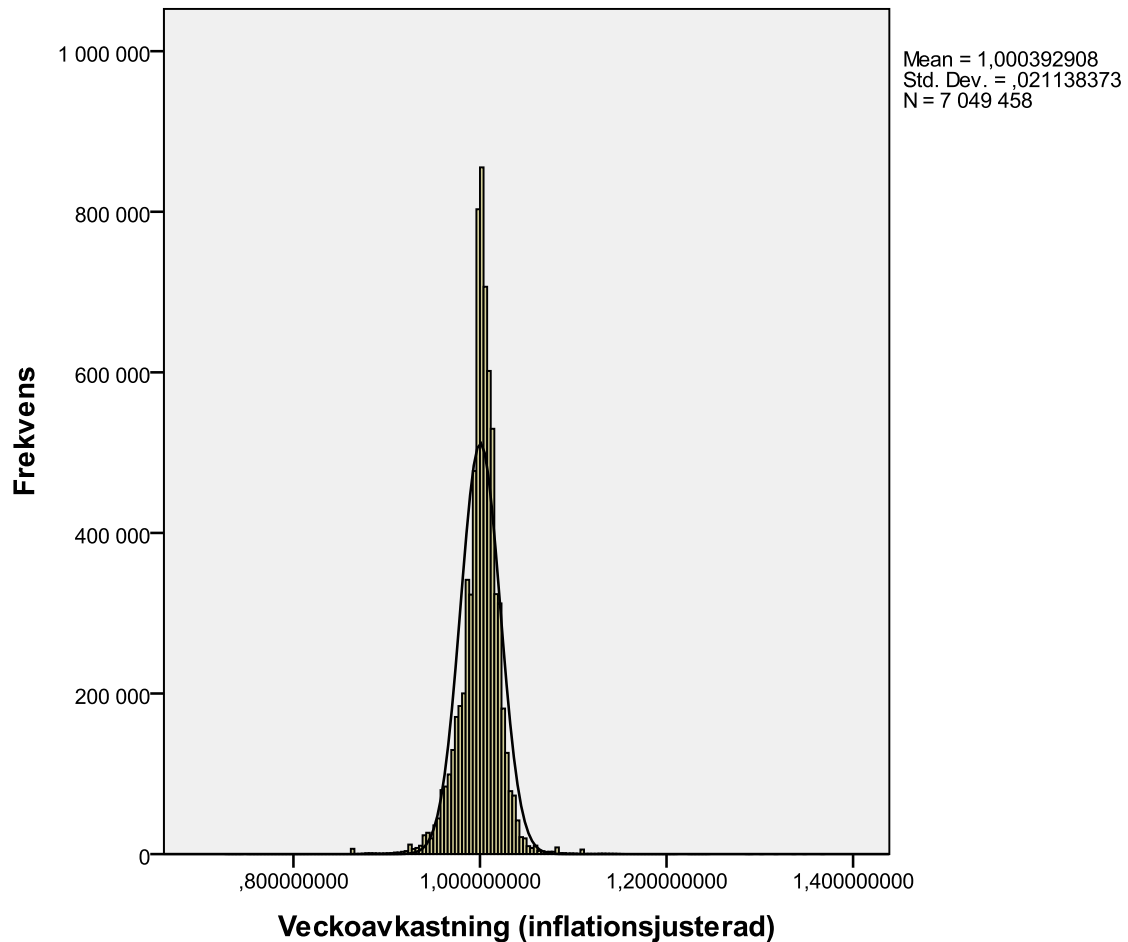
Resultat och analys

För att göra rapporten lättare att läsa presenteras resultat och analys samtidigt.

Fördelningen av veckoränta för sparnas fondval

För att kunna avgöra vilka statistiska metoder som var lämpliga behövde datat testas för att se om antagandet om normalfördelning håller.

Histogram över veckoavkastningen hos fondspararna i vårt stickprov



Fördelningen hos avkastningen har följande egenskaper:

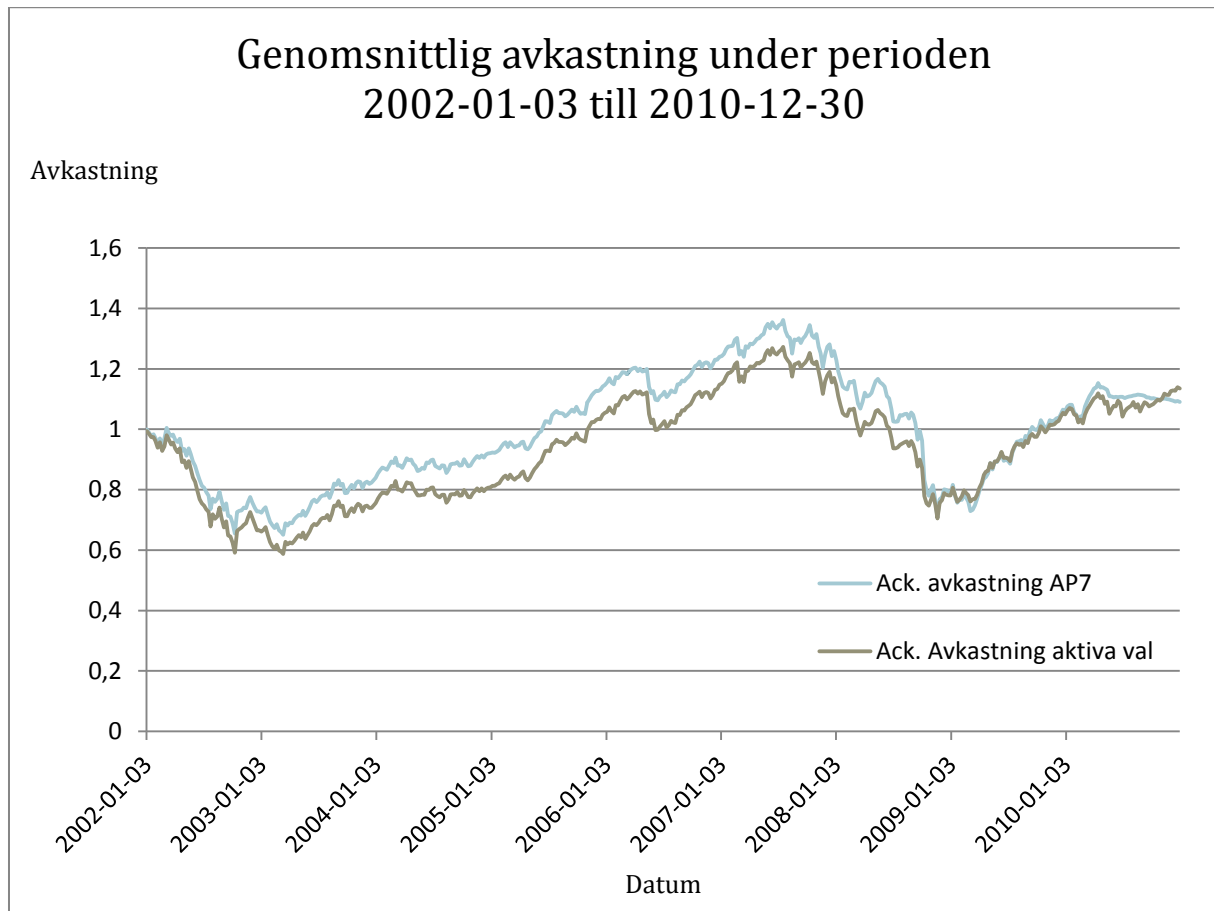
- Avkastningen är asymmetriskt fördelad kring medelvärdet.
- Värden något över genomsnittet är kraftigt överrepresenterade.
- Extrema värden är överrepresenterade.

Tests of Normality

	Kolmogorov-Smirnov ^a		
	Statistic	df	Sig.
Veckoavkastning (inflationsjusterad)	,099	7049458	,000

a. Lilliefors Significance Correction

Signifikansen i Kolmogorov-Smirnov-testet är mindre än 0,05, detta betyder att antagandet om att fördelningen är normalfördelad kan förkastas. Eftersom avkastningen ej är normalfördelad så används icke-parametriska metoder.



Grafen ovan visar den genomsnittliga ackumulerade avkastningen (uttaget per vecka) under perioden 2002-01-03 till 2010-12-30. Sjunde AP-fonden presterade något bättre än de aktivt valda fonderna i början av perioden men under senare delen av perioden är förhållandet det motsatta.

Den ackumulerade avkastningen för de som inte gjort ett aktivt val var under denna nio-årsperiod var 9,03 procent och för de som gjort ett aktivt val 13,5 procent. Vilket ger en genomsnittlig årlig avkastning på:

$\sqrt[9]{1,0903} \Rightarrow 1,0\%$ för de som inte gjort ett aktivt val och $\sqrt[9]{1,135} \Rightarrow 1,4\%$ för de som gjort ett aktivt val. Viktigt att nämna i detta sammanhang är att det finns väldigt stora skillnader i genomsnittlig avkastning bland de som gjort ett aktivt val.

Utifrån denna graf har vi använt följande perioder för positiv och negativ utveckling på fondmarknaden:

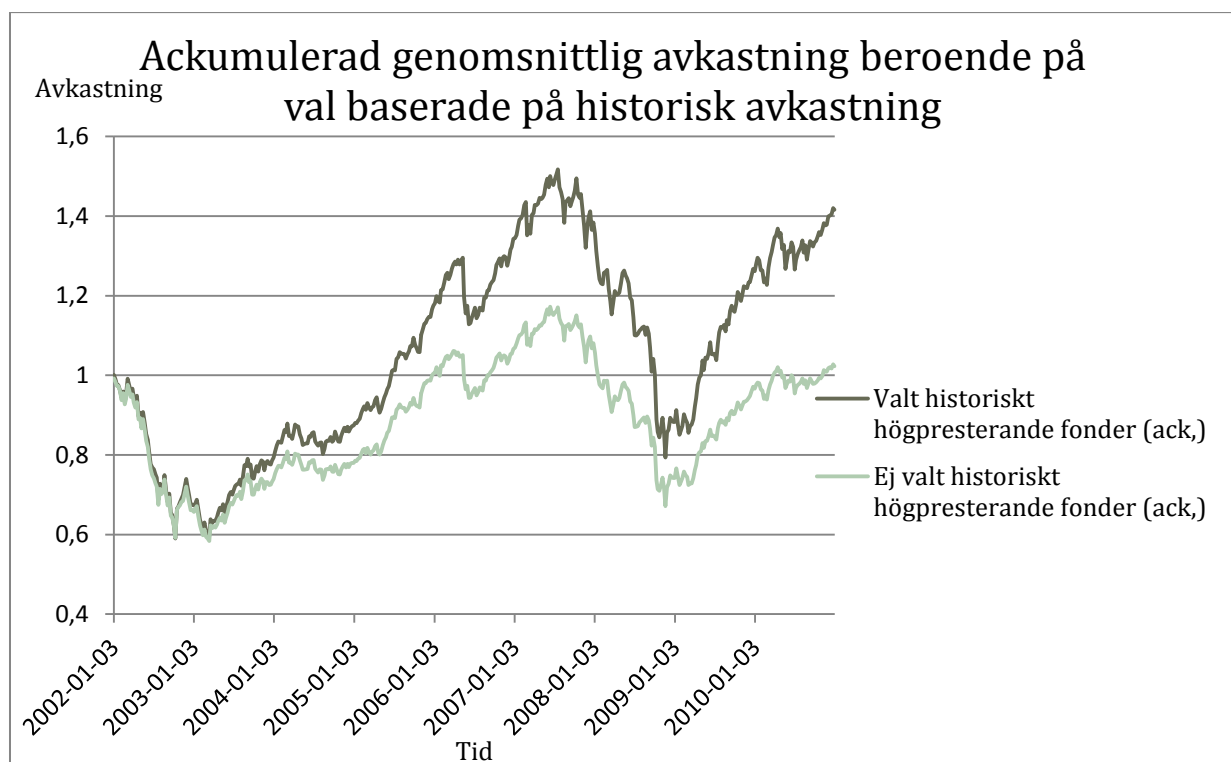
- Period med negativ avkastning: 2007-10-18 - 2009-03-05
- Period med positiv avkastning: 2009-03-12 - 2010-12-30

Avkastning hos personer som baserat sitt fondval på historisk avkastning

Ett intressant psykologiskt bias är övertro på sin egen förmåga (overconfidence). Nordén (2010) menar att det som skiljer duktiga fondsparare från de som tror att de är duktiga är att i jakten på hög avkastning väljer de som tror att de är duktiga fonder baserat på hög historisk avkastning medan de som är verkligen duktiga väljer fonder baserat på andra parametrar. Enligt den effektiva marknadshypotesen är det inte möjligt att med fördel prognosticera framtida avkastning med hjälp av historisk avkastning.

Dahlquist et al. (2011) menar att en hög historisk avkastning visar på att fondförvaltaren är skicklig och att det därför är klokt att välja fonder som har haft hög historisk avkastning.

Även i detta fall visar vårt test på stora skillnader beroende på konjunkturen. Under hela tidsperioden (som innehåller perioder med både positiv och negativ avkastning) har de som valt fonder baserat på historisk avkastning en högre avkastning än de som valt fonder baserat på andra aspekter.



De som valt fonder baserat på historisk avkastning hade under perioden 2002 till 2010 en ackumulerad avkastning på 42 procent vilket innebär en årlig avkastning motsvarande $\sqrt[3]{1,42} \Rightarrow 4\%$. För de som inte valt fonder baserat på historisk avkastning var den genomsnittliga ackumulerade avkastningen 2 procent vilket motsvarar en genomsnittliga årlig avkastning på $\sqrt[3]{1,02} \Rightarrow 0\%$.

Dessa siffror är inte statistiskt säkerställda, endast det inbördes förhållandet mellan dem enligt testet nedan.

Ranks

Historisk avkastning		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Har ej (enbart) de 25% bästa fonderna från föregående år	56697	34824,82	1,97E9
	Har endast de 25% bästa fonderna från föregående år i sin portfölj	13485	36212,74	4,88E8
	Total	70182		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	3,672E8
Wilcoxon W	1,974E9
Z	-7,150
Asymp. Sig. (2-tailed)	,000

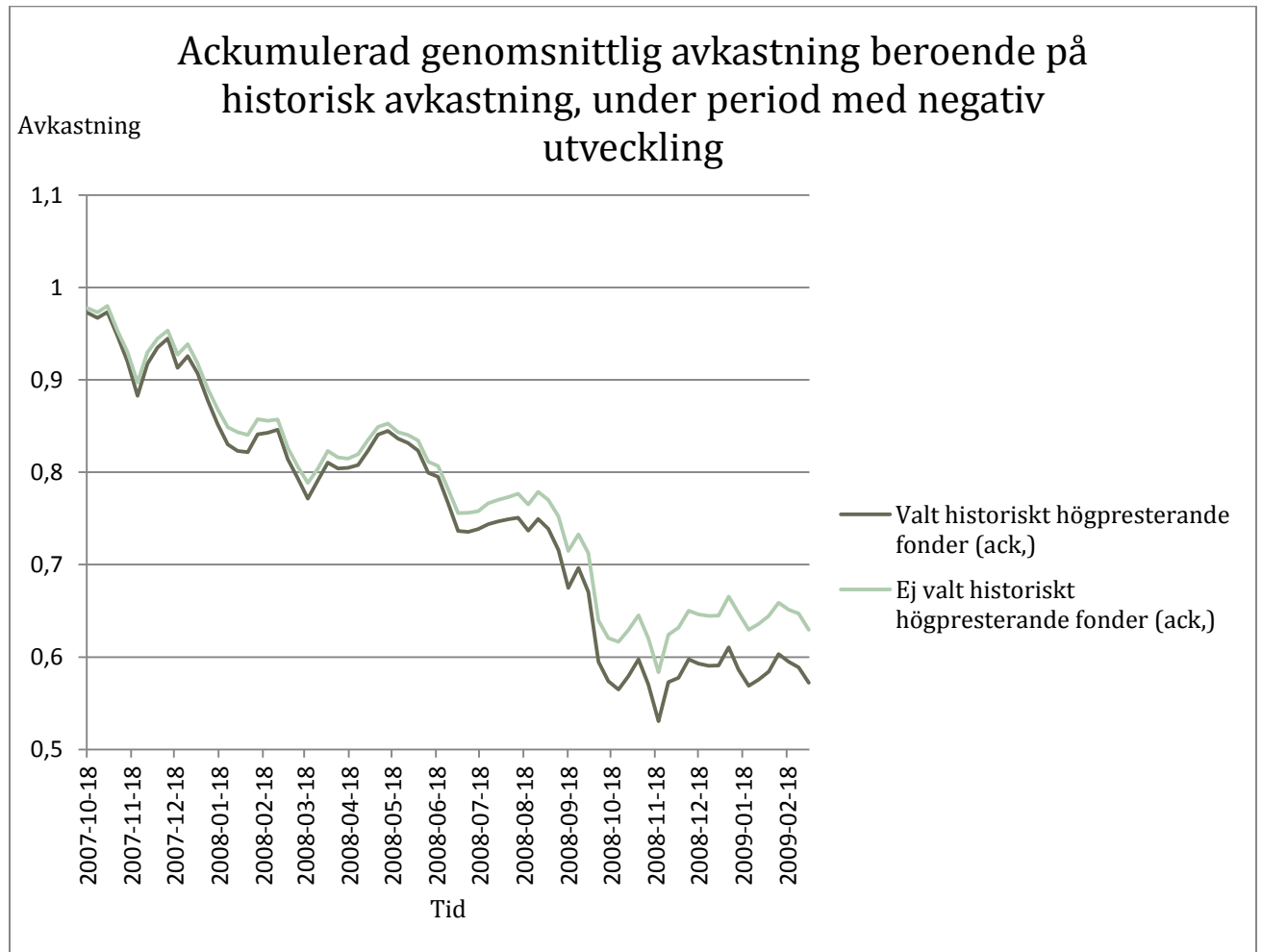
a. Grouping Variable: Historisk avkastning

H_0 kan förkastas. Personer som har använt historisk avkastning för att välja fonder har haft en bättre utveckling än genomsnittet under perioden.

Detta tyder på att det faktiskt är fördelaktigt att basera sitt pensionsfondsval på historisk avkastning. Det kan dock vara så att spararna som valt de fonder som historiskt sett presterat bäst även har annan typ av information som påverkar deras val. Att det är lönsamt anser vi bero på att Dahlquist et al. (2011) har rätt i sin tes att visa förvaltare är bättre än andra. Att effekten kan observeras anser vi beror på asymmetrisk information och att de svenska pensionsspararna inte till fullo är rationella, den effektiva marknadshypotesen verkar således inte vara tillämpbar på detta område.

Historisk avkastning under negativ utveckling

Under en period med negativ avkastning är dock förhållandet det omvända. Enligt Butlers och Domians (1993) slutsats att fonder med hög volatilitet så gott som alltid presterar bättre än fonder med låg volatilitet på lång sikt är detta fullt logiskt. Fonder med hög historisk avkastning ska då ha en högre volatilitet och i och med detta skall de prestera bättre under längre tidsperioder men den högre volatiliteten leder till kraftigare nedgångar under perioder med negativ avkastning.



Förhållandet mellan dessa siffror är inte statistiskt säkerställda. Vi kan således endast konstatera att linjerna i grafen ligger nära varandra under perioden med negativ avkastning.

Ranks

Historisk avkastning		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Har ej (enbart) de 25% bästa fonderna från föregående år	67660	43241,68	2,93E9
	Har endast de 25% bästa fonderna från föregående år i sin portfölj	18818	43231,66	8,14E8
	Total	86478		

Test Statistics^a

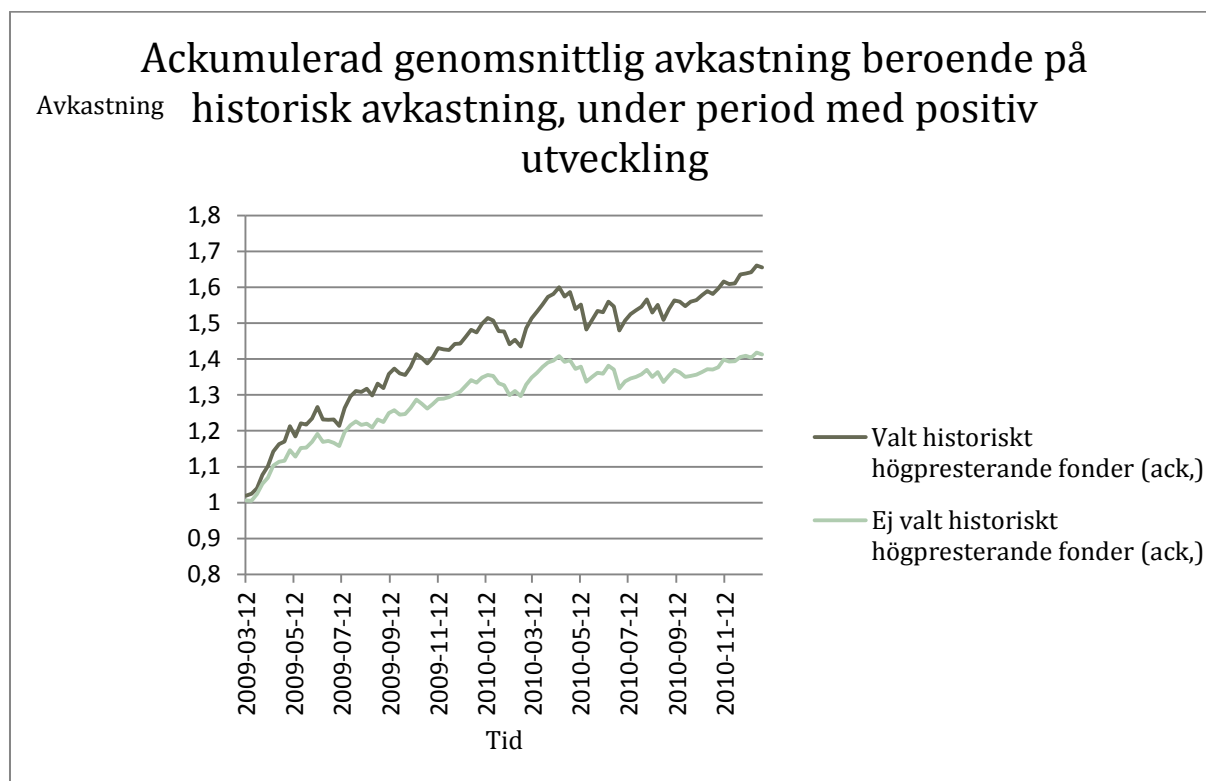
	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	6,365E8
Wilcoxon W	8,135E8
Z	-,049
Asymp. Sig. (2-tailed)	,961

a. Grouping Variable: Historisk avkastning

H_0 kan ej förkastas. Avkastningen hos de som baserat sina val på historisk avkastning är varken sämre eller bättre än genomsnittet i perioden med negativ utveckling.

Historisk avkastning under positiv utveckling

Som nämnts tidigare definieras fonder som presterar bäst på lång sikt bland annat av en hög volatilitet. Av samma anledning som att dessa fonder presterar sämst under perioder med negativ avkastning presterar de även bäst under perioder med positiv avkastning.



Under testets period med positiv avkastning hade de som valt fonder baserat på hög historisk avkastning en ackumulerad genomsnittlig avkastning på 66 procent vilket ger en genomsnittlig årlig avkastning på $1,66^{\frac{365}{658}} \Rightarrow 32\%$. För de som valt fonder baserat på annat än hög historisk avkastning var den genomsnittliga historiska avkastningen 41 procent, vilket renderar i en genomsnittliga årlig avkastning på $1,41^{\frac{365}{658}} \Rightarrow 20\%$.

Dessa siffror är inte statistiskt säkerställda, endast det inbördes förhållandet mellan dem är det enligt testet nedan.

Ranks

Historisk avkastning		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Har ej (enbart) de 25% bästa fonderna från föregående år	80013	49870,76	3,99E9
	Har endast de 25% bästa fonderna från föregående år i sin portfölj	21253	53505,07	1,14E9
	Total	101266		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	7,892E8
Wilcoxon W	3,990E9
Z	-16,110
Asymp. Sig. (2-tailed)	,000

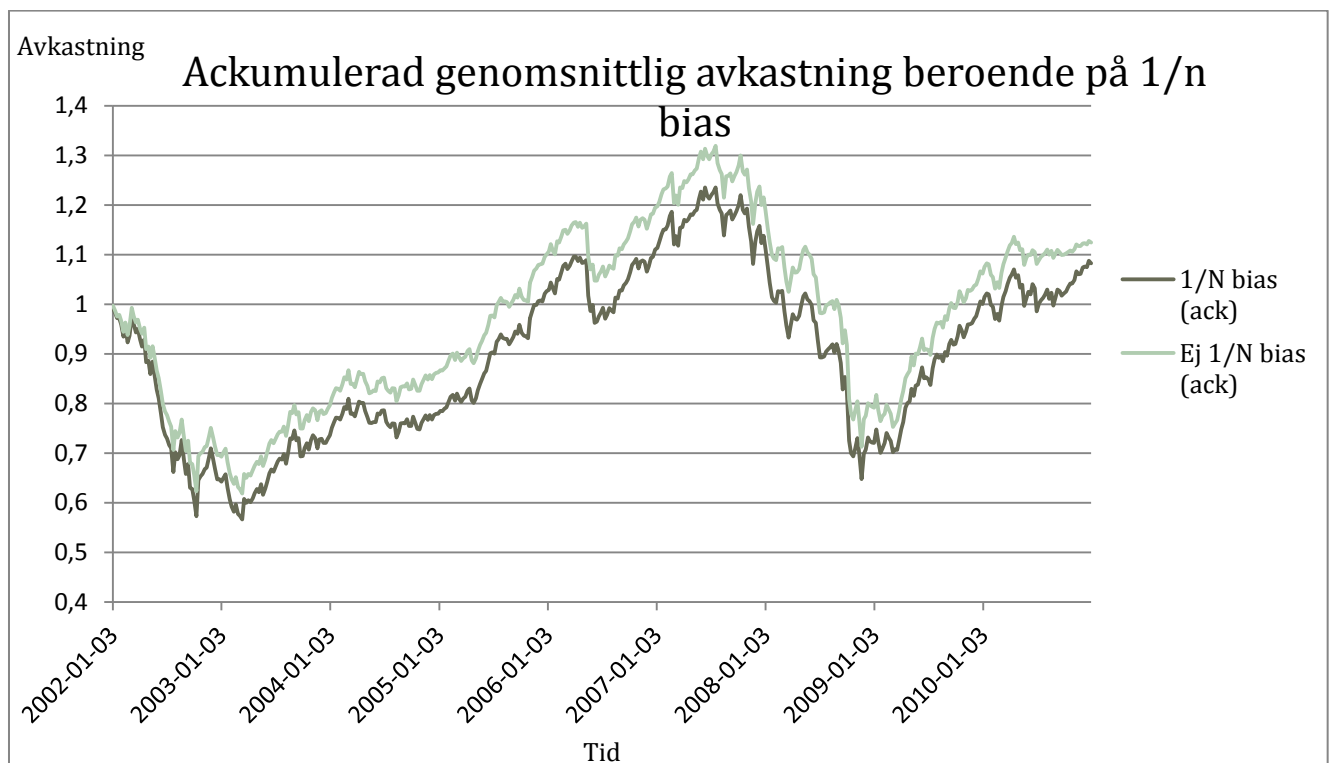
a. Grouping Variable: Historisk avkastning

H_0 kan förkastas. Personer som baserar sina val på historisk avkastning har högre avkastning i perioden med positiv avkastning.

Sammanfattningsvis anser vi att teorierna som hävdar att på lång sikt är det mer lönsamt att spara i fonder med hög volatilitet står sig väldigt starkt och får stöd i våra tre tester angående historisk avkastning som beslutsunderlag.

1/n bias under hela perioden

Skillnaderna mellan de som har 1/n bias och de som inte har det är mycket små och det finns ingen signifikans för testet perioden 2002 till 2010. Bernatzi och Thaler definierar 1/n fördelning som en form av naiv och icke fördelaktig diversifiering. Om det är så att sparare använder 1/n som ett sätt att undvika osäkerhet som litteraturen i ämnet föreskriver, verkar de enligt vårt test inte förlora särskilt mycket på detta. En optimal form av diversifiering (kovariansuträkning) tycks inte resultera i ett särskilt mycket bättre resultat än denna icke fördelaktiga form av diversifiering. Det skulle kunna förklaras av att dels består pensionsfonderna av många olika säkerheter som byts ut av förvaltaren vilket gör att dess förhållande till andra fonder förändras, och dels består fonderna av olika typer av säkerheter. Vi anser därför att det är orimligt att förvänta sig att den svenska befolkningen ska beräkna kovarians enligt Svedsätters et al's. test. Det går även ifrågasätta om det är värt att använda kovariansberäkning med tanke på detta resultat.



Dessa data saknar signifikans enligt tabellerna nedan och slutsatser kring denna graf är således subjektiva.

Ranks

	1/n bias	N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Har ej 1/n bias	60362	31669,99	1,91E9
	Har 1/n bias	2999	31902,70	95676188,50
	Total	63361		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	89847949,500
Wilcoxon W	1,912E9
Z	-,680
Asymp. Sig. (2-tailed)	,496

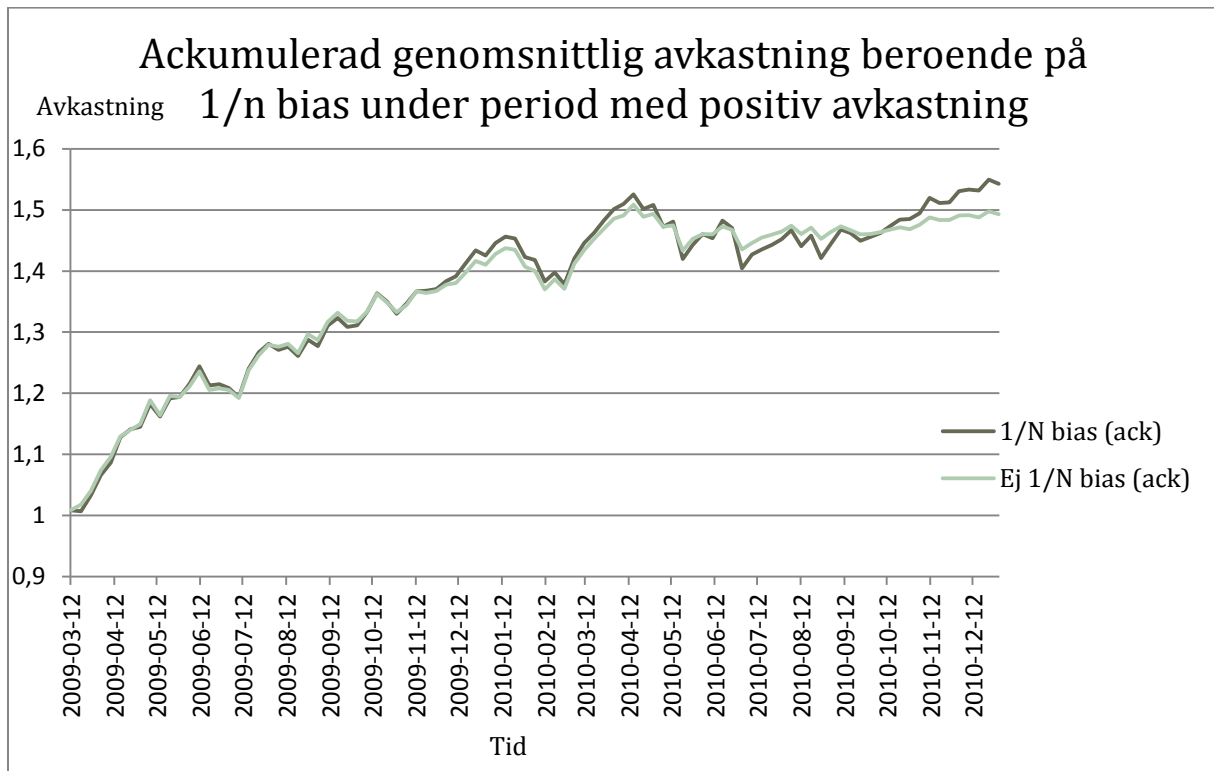
Grouping Variable: 1/n bias

H_0 kan ej förkastas. 1/n bias påverkar inte resultatet under den studerade perioden 2002-01-01 – 2010-12-30

1/n bias vid positiv avkastning

I perioden med positiv utveckling finns det en stark signifikans som säger att det är bättre att ha 1/n bias än att inte ha, dock är skillnaden även här mycket liten. Den genomsnittliga ackumulerade avkastning för de som har 1/n bias är 54 procent vilket ger en genomsnittlig årlig avkastning på $1,54^{\frac{365}{658}} \Rightarrow 27\%$. Motsvarande siffror för de som inte hade 1/n bias var 49 procent och $1,49^{\frac{365}{658}} \Rightarrow 25\%$.

Detta skulle kunna förklaras av att det finns en kovarians bland fonderna som är 1/n fördelade och att de av den anledningen har en något högre volatilitet., vilket leder till en något högre avkastning under en period då avkastningarna generellt är högre.



Signifikansen i detta var som sagt stark, dock är det bara det inbördes ordningen mellan de två grupper som stärks av denna signifikans, se tabeller nedan.

Ranks

	1/n bias	N	Mean Rank	Sum of Ranks
Veckoavkastning	Har ej 1/n bias	68763	35913,78	2,47E9
(inflationsjusterad)	Har 1/n bias	3204	37491,13	1,20E8
	Total	71967		

Test Statistics^a

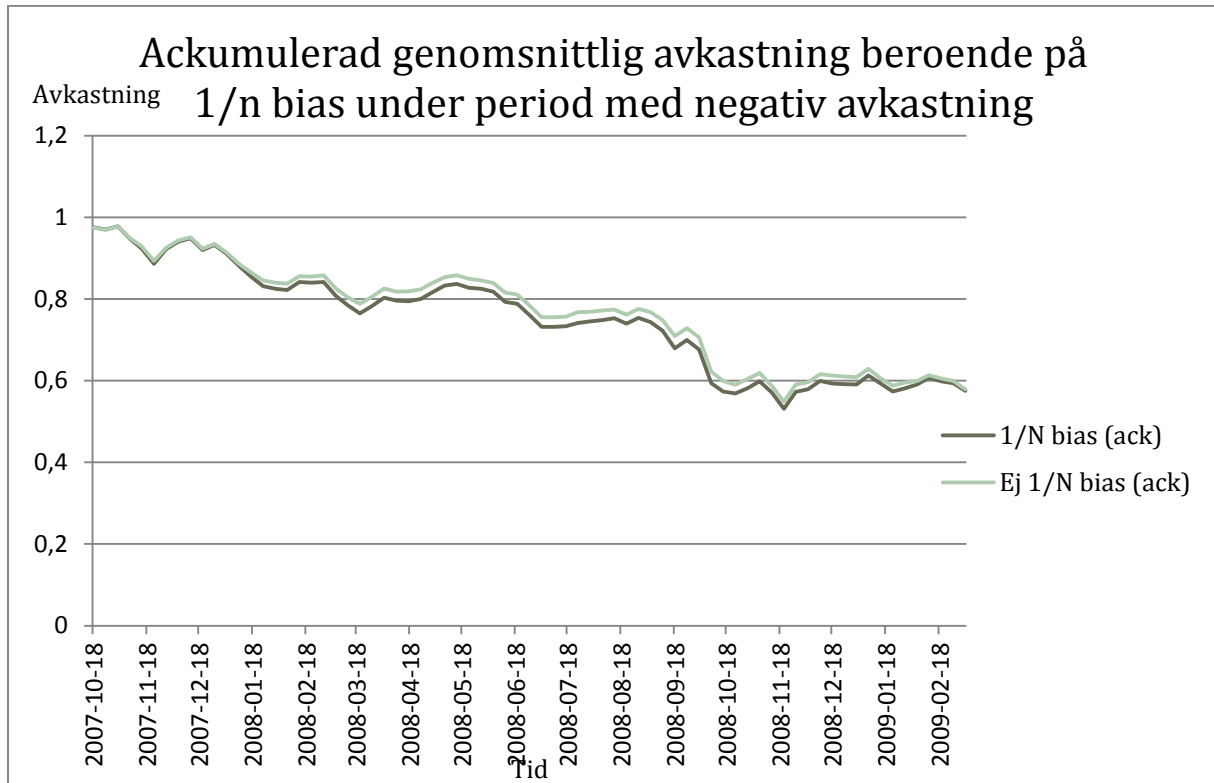
	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	1,053E8
Wilcoxon W	2,470E9
Z	-4,201
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: 1/n bias

H_0 kan förkastas. 1/n bias har ett bättre resultat än genomsnittet under perioden med positiv avkastning.

1/n bias vid negativ avkastning

Om vårt resonemang kring att en 1/n fördelning leder till en högre volatilitet (se föregående rubrik) skulle stämma skulle de som har 1/n ha en lägre avkastning under en period då avkastningarna generellt sett är negativa. Vårt test för denna period med negativ avkastning saknar dock signifikans och av den anledningen kan vi varken styrka eller förkasta vårt resonemang kring 1/n och volatilitet.



Detta test har låg signifikans, se tabeller nedan, därav kan inga slutsatser dras.

Ranks

1/n bias		N	Mean Rank	Sum of Ranks
Veckoavkastning	Har ej 1/n bias	51620	27081,74	1,40E9
(inflationsjusterad)	Har 1/n bias	2550	27161,59	69262057,00
Total		54170		

Test Statistics^a

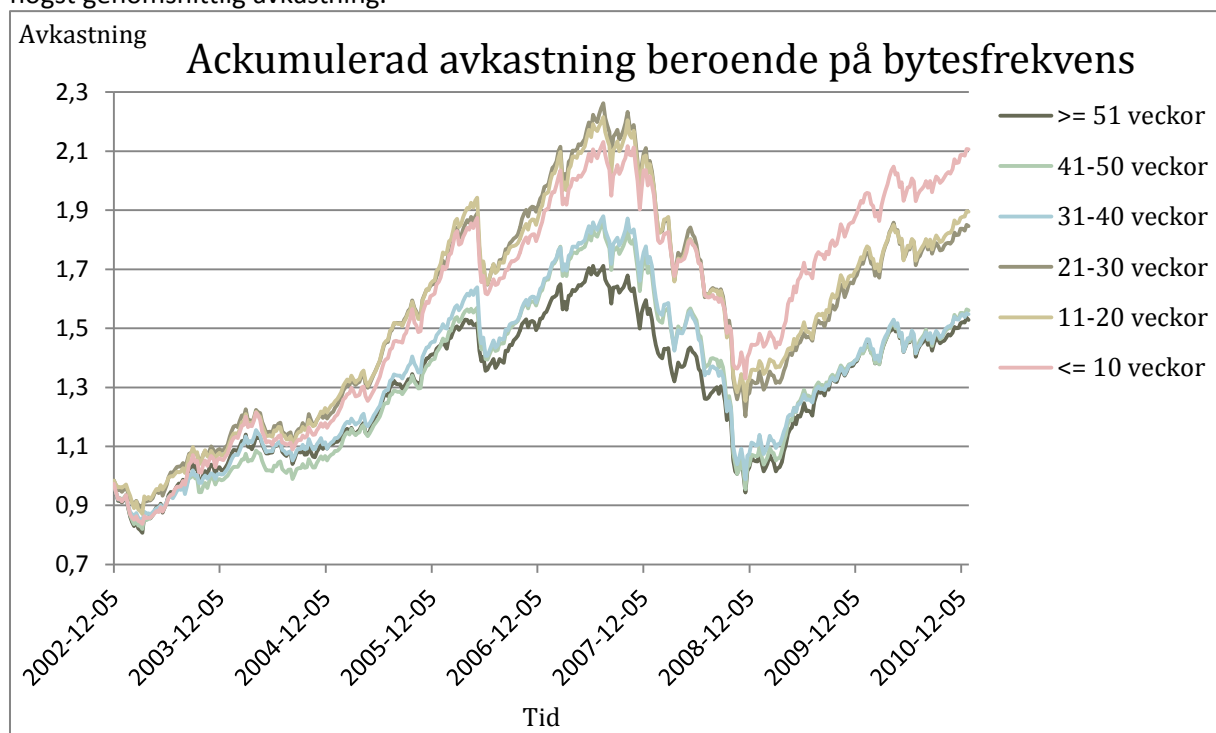
	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	65621468,000
Wilcoxon W	1,398E9
Z	-,252
Asymp. Sig. (2-tailed)	,801

a. Grouping Variable: 1/n bias

H_0 kan ej förkastas. 1/n bias har samma resultat som genomsnittet under perioden med negativ utveckling.

Effekten av hög aktivitet

Syftet med denna uppsats var att studera vilka parametrar mer än bytesfrekvens, som redan är bevisat av Dahlquist et al. (2011), som har signifikant betydelse för avkastningen för de svenska fondspararna och under vilka förhållanden. För att kunna göra detta är ett test av bytesfrekvensens betydelse under olika omständigheter under vår testperiod viktigt att utreda. Grafen nedan visar att även under vår testperiod har de som bytt fonder mest frekvent (mer frekvent än var tionde vecka) högst genomsnittlig avkastning.



(Tidslinjen för denna kurva är förskjuten jämfört med tidigare kurvor på grund av att 50 veckor måste passeras innan första observation kan göras).

Under perioden hade de som bytte fonder mer frekvent än var tionde vecka en ackumulerad genomsnittlig avkastning på 111 procent, vilket ger en genomsnittlig årlig avkastning på $\sqrt[3]{2,11} \Rightarrow 8,7\%$. De som bytte näst mest frekvent (mellan var elfte och var tjugonde vecka) hade en genomsnittlig ackumulerad avkastning på 89 procent vilket ger en genomsnittlig årlig avkastning på $\sqrt[3]{1,89} \Rightarrow 7,3\%$. Även de som bytte tredje mest frekvent (mellan var tjugoförsta och var trettionde vecka) hade en avsevärt högre genomsnittlig avkastning än medelvärdet som var 75 procent avkastning. De hade en genomsnittlig ackumulerad avkastning på 85 procent vilket ger en genomsnittlig årligavkastning på $\sqrt[3]{1,85} \Rightarrow 7,0\%$. De som byter fonder mer sällan än så har en lägre avkastning än genomsnittet.

Uppskattad ackumulerad respektive årlig genomsnittlig avkastning för respektive grupp.

Bytesfrekvens	Antal i urvalsgruppen	Ackumulerad avkastning	Årlig avkastning
<= 10 veckor	1016	110 %	8,7 %
11-20 veckor	724	89 %	7,3 %
21-30 veckor	651	85 %	7,0 %
31-40 veckor	499	55 %	5,0 %
41-50 veckor	473	56 %	5,0 %
>= 51 veckor	11457	53 %	4,8 %

Siffrorna är ej statistiskt säkerställda, endast deras inbördes storleksordning är säkerställd enligt nedan.

Ranks

Bytesvanor		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Har ej bytt fond de senaste 50 veckorna	51682	29087,52	1,50E9
	Har bytt fond de senaste 50 veckorna	6571	29437,55	1,93E8
	Total	58253		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	1,678E8
Wilcoxon W	1,503E9
Z	-1,589
Asymp. Sig. (2-tailed)	,112

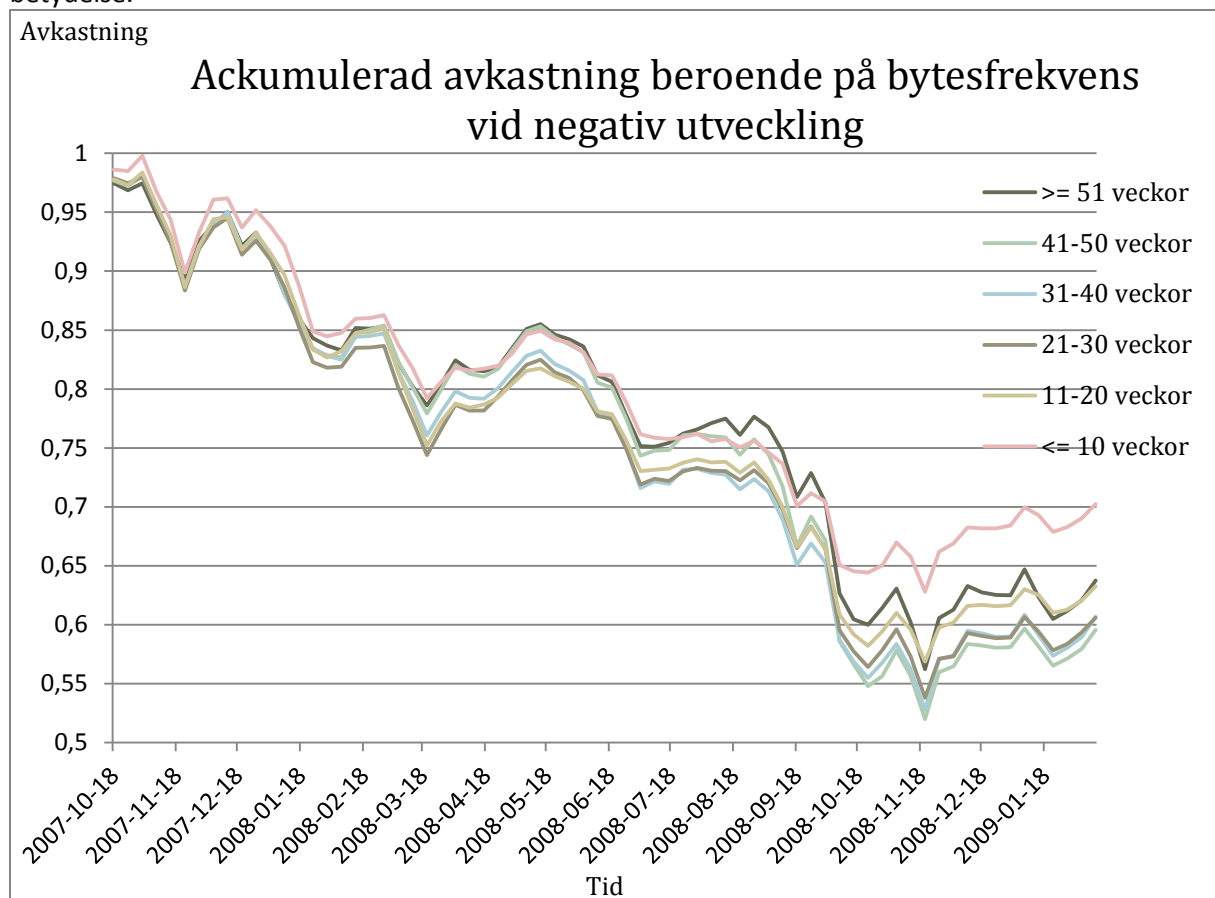
a. Grouping Variable: Bytesvanor

H_0 kan ej förkastas. Personer som aktivt byter fonder har samma resultat som genomsnittet under hela perioden.

Även om det inte är statistiskt säkerställt, är antagandet rimligt eftersom det ligger i linje med resultatet Dahlquist et al. (2011) kommit fram till. Vårt, precis som deras, test antyder att det är fördelaktigt att byta fonder så frekvent som möjligt.

Effekt av hög aktivitet vid negativ avkastning

Under testperiodens tid med negativ avkastning presterar de som byter fonder mest frekvent bäst. Deras ackumulerade genomsnittliga avkastning är -30 procent vilket ger en genomsnittlig årlig avkastning på $0,7^{\frac{365}{504}} \Rightarrow -23\%$. Medelvärdet för de olika bytesfrekvenserna var -37 procent ackumulerad genomsnittlig avkastning. Anmärkningsvärt i detta test är att endast de som byter mest frekvent presterar mer än marginellt bättre än så och att de som presterar näst bäst är de som byter mest sällan. Dahlquists et al. får till viss del rätt för sin tes att de svenska spararna inte är lyhörda och reagerar inte i tillräckligt stor utsträckning på förändringar, samtidigt visar vårt test att endast de som byter mest ofta (är mest lyhörda) presterar bättre än genomsnittet. Det är dock endast statistiskt säkerställt att de som byter mer ofta än var 51:a vecka presterar bättre än de som inte gör det. Grafen nedan antyder dock att det krävs ett väldigt frekvent bytande för att det ska ha någon betydelse.



Trots detta visar Mann-Whitney U testet att det är signifikant säkerställt att de som har bytt fonder under den närmsta 50-veckorsperioden presterar bättre än de som inte gjort det. Däremot är inte värdena i grafen säkerställda, endast den inbördes ordningen.

Uppskattad ackumulerad och årlig avkastning för respektive grupp.

Bytesfrekvens	Antal i urvalsgruppen	Ackumulerad avkastning	Årlig avkastning
<= 10 veckor	957	- 30 %	- 23 %
11-20 veckor	463	- 37 %	- 28 %
21-30 veckor	385	- 39 %	- 30 %
31-40 veckor	383	- 39 %	- 30 %
41-50 veckor	339	- 40 %	- 31 %
>= 51 veckor	12201	- 36 %	- 28 %

Ranks

Bytesvanor		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Har ej bytt fonder de senaste 50 veckorna	64265	37552,10	2,41E9
	Har bytt fond de senaste 50 veckorna	11089	38404,27	4,26E8
	Total	75354		

Test Statistics^a

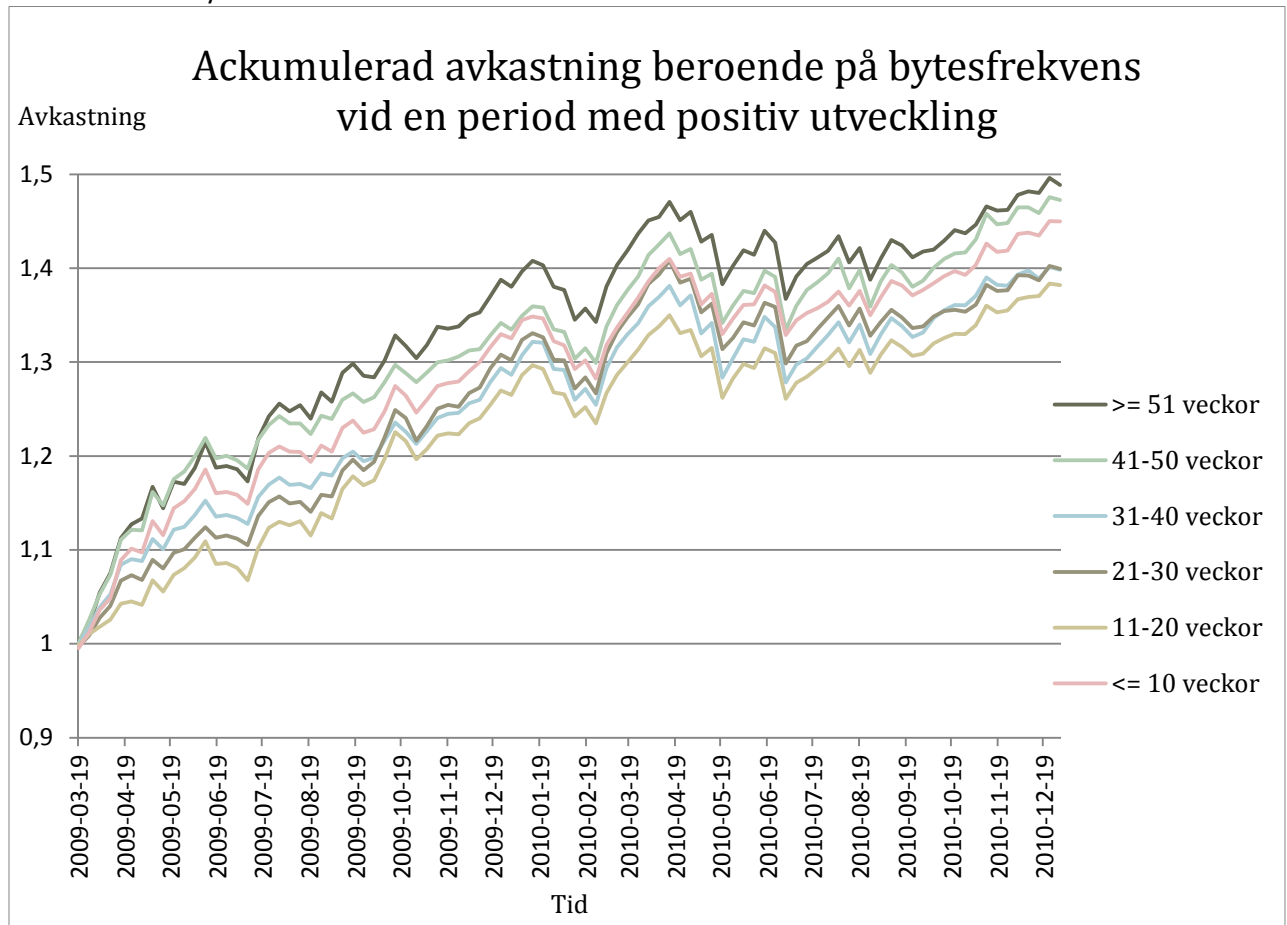
	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	3,483E8
Wilcoxon W	2,413E9
Z	-3,810
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: Bytesvanor

H_0 kan förkastas. De personer som aktivt byter fond under perioden med negativ utveckling har ett bättre resultat än genomsnittet.

Effekten av hög aktivitet under positiv avkastning

Under perioden med positiv avkastning i testet finns det ingen signifikans i huruvida det är bättre eller sämre att byta fonder ofta.



De som presterar allra bäst under denna period är de som byter mest sällan, de har en genomsnittlig ackumulerad avkastning på 49 procent vilket ger en genomsnittlig årlig avkastning på $1,49^{\frac{365}{658}} \Rightarrow 25\%$, att jämföras med de som byter mest ofta som hade en genomsnittlig ackumulerad avkastning på 45 procent vilket ger en genomsnittlig årlig avkastning på $1,45^{\frac{365}{658}} \Rightarrow 23\%$. Tabellen nedan visar ackumulerad och årlig avkastning för respektive grupp.

Under denna period är det mycket svårt att se några samband och mönster i avkastningen beroende på hur frekvent spararna byter sina fonder.

Uppskattat ackumulerad avkastning vid olika bytesfrekvens i perioden med positiv utveckling

Bytesfrekvens	Antal i urvalsgruppen	Ackumulerad avkastning	Årlig avkastning
<= 10 veckor	2095	45 %	23 %
11-20 veckor	1190	38 %	20 %
21-30 veckor	982	40 %	21 %
31-40 veckor	300	40 %	21 %
41-50 veckor	221	47 %	24 %
>= 51 veckor	10372	49 %	25 %

Tabellerna nedan visar att dessa resultat inte är statistiskt säkerställda, av den anledningen dras inga slutsatser av detta. Siffrorna i tabellen ovan ska endast ses som en förklaring av grafen.

Ranks

Bytesvanor		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Har ej bytt fond de senaste 50 veckorna	85203	50456,04	4,30E9
	Har bytt fond de senaste 50 veckorna	15614	50152,31	7,83E8
	Total	100817		

Test Statistics^a

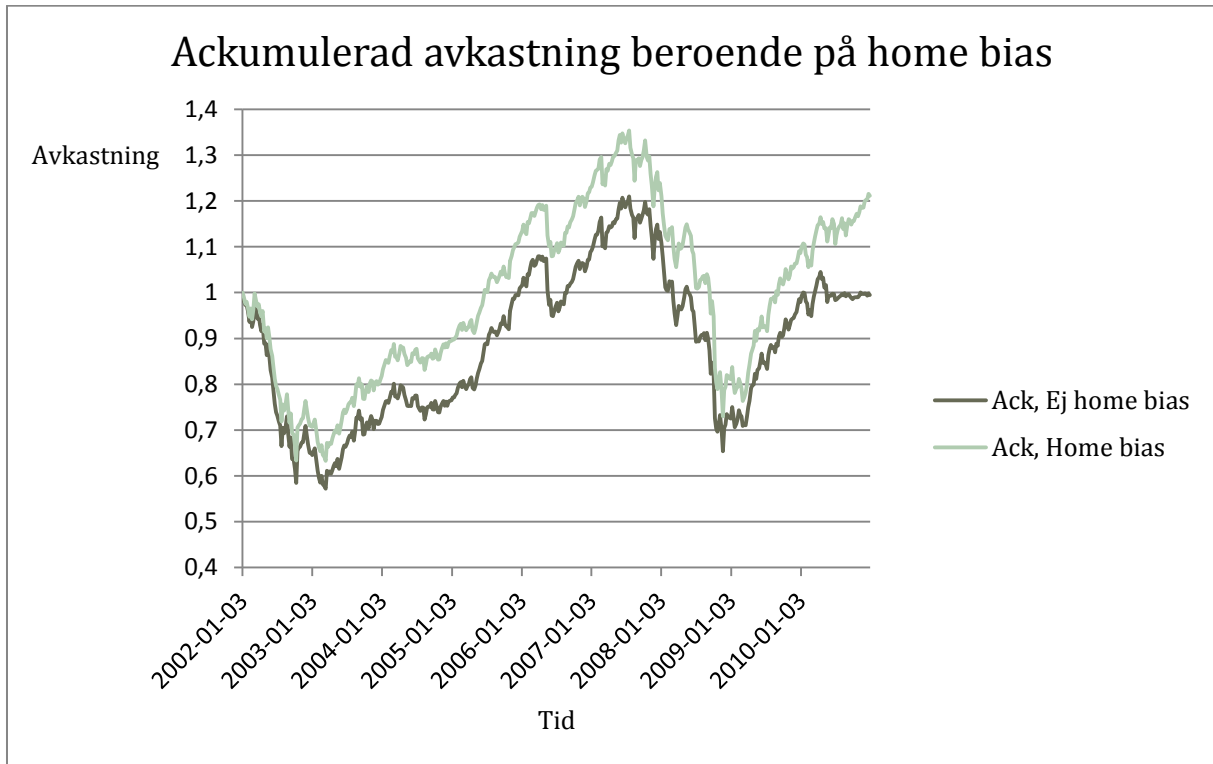
	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	6,612E8
Wilcoxon W	7,831E8
Z	-1,199
Asymp. Sig. (2-tailed)	.231

a. Grouping Variable: Bytesvanor

H_0 kan ej förkastas. Under perioden med positiv avkastning har personer som aktivt byter fonder samma avkastning som genomsnittet.

Home bias

Home bias är allmänt känt som ett psykologiskt bias som sänker prestationen i sparandet. I vårt experiment finner vi dock att de som i större utsträckning (mer än medelvärdet) väljer fonder i Sverige presterar bättre. Den ackumulerade avkastningen under perioden 2002 till 2010 för de som har mer home bias är 21 procent. Motsvarande siffra för de som har mindre home bias än genomsnittet är - 1 procent.



Den genomsnittliga årliga avkastningen för de som har home bias är $\sqrt[9]{1,2115} \Rightarrow 2\%$ och för de som har lägre home bias än genomsnittet är motsvarande siffra $\sqrt[9]{0,9948} \Rightarrow 0\%$.

OBS. de här procentsatserna är inte statistisk säkrade, endast den inbördes ordningen mellan dem är säkrad enligt testet nedan.

Att avkastningen är större vid home bias strider därmed mot det generella antagandet att home bias, något som författarna anser bero på främst två parametrar.

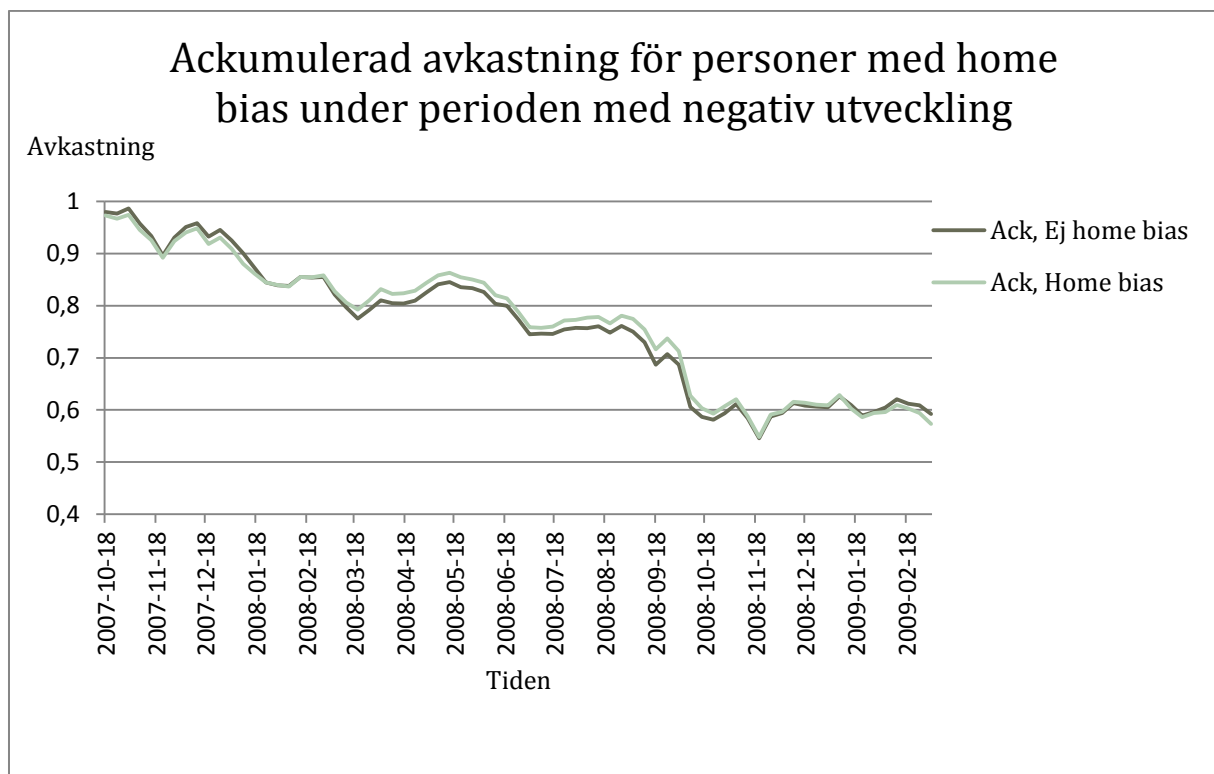
1. Internationell diversifiering har en bevisad effekt att sänka riskexponeringen i en portfölj. Enligt Nordén (2010) avtar dock denna effekt när antalet tillgångar i en portfölj överstiger tolv stycken. Eftersom denna studie är baserad på fonder är det ett rimligt antagande att fonderna innehåller mer än tolv olika tillgångar.
2. Vår studerade period (2002-2010) innehåller proportionellt längre perioder av positiv avkastning än negativ avkastning. Skillnaden i avkastning mellan att vara, eller inte vara, home bias är större under perioder med positiv avkastning (se nedan).

		Ranks		
Home bias		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Home bias under eller lika med genomsnittet	20053	31573,11	6,33E8
	Home bias över genomsnittet	43695	32012,82	1,40E9
	Total	63748		

Test Statistics ^a	
	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	4,321E8
Wilcoxon W	6,331E8
Z	-2,801
Asymp. Sig. (2-tailed)	,005

a. Grouping Variable: Home bias

H_0 kan förkastas. De som har ett home bias över genomsnittet har haft en bättre avkastning än genomsnittet under perioden.



Under perioder med negativ avkastning är skillnaden mellan de som är home bias och de som inte är det mycket liten. Vårt test innehåller perioden 2007-10-18 till 2009-03-05 där avkastningen var negativ för båda grupperna. De som hade home bias under denna period hade en ackumulerad avkastning på -43 procent vilket motsvarar en genomsnittlig årlig avkastning på $0,57^{\frac{365}{504}} \Rightarrow -33\%$. Den ackumulerade genomsnittliga avkastningen för de som inte är home bias under perioden är -41 procent vilket motsvarar en genomsnittliga årlig avkastning på $0,59^{\frac{365}{504}} \Rightarrow 32\%$.

Dessa procentsatser är inte statistisk säkrade, endast den inbördes ordningen mellan dem är säkrad enligt testet nedan.

Ranks

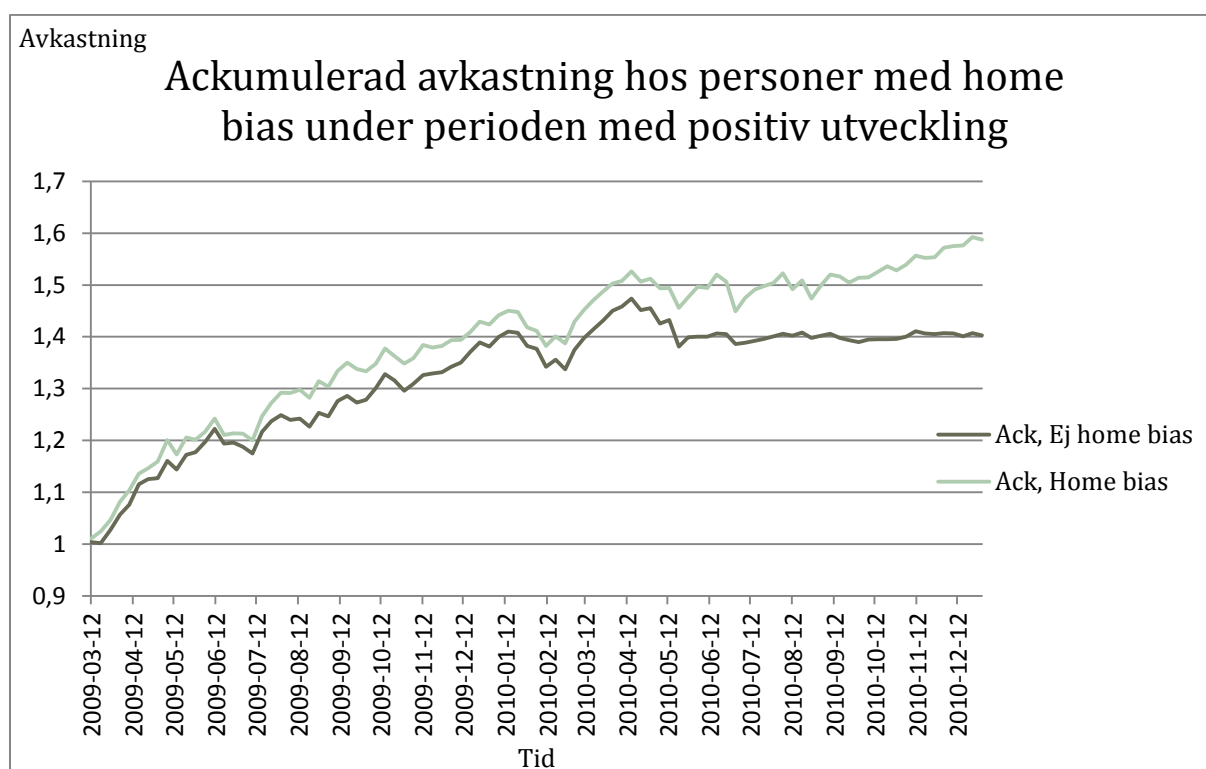
Home bias		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Home bias under eller lika med genomsnittet	16679	27226,10	4,54E8
	Home bias över genomsnittet	37161	26783,34	9,95E8
	Total	53840		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	3,048E8
Wilcoxon W	9,953E8
Z	-3,057
Asymp. Sig. (2-tailed)	,002

a. Grouping Variable: Home bias

H_0 kan förkastas. Personer som har ett home bias som är högre än genomsnittet klarar sig sämre än genomsnittet under perioden med negativ utveckling.



Under en period med positiv avkastning är skillnaden de som är home bias och de som inte är det större än under perioder som innehåller både positiv och negativ avkastning. Vårt test innehåller perioden 2009-03-12 till 2010-12-30 där avkastningen var positiv för både grupperna. De som hade home bias under denna period hade en ackumulerad avkastning på 59 procent vilket ger en årlig avkastning på $1,59^{\frac{365}{658}} \Rightarrow 29\%$. Den ackumulerade avkastningen för de som placerade sina PPM-pengar utomlands i större utsträckning var under denna period 40 procent, vilket renderar i en genomsnittlig årlig avkastning motsvarande $1,40^{\frac{365}{658}} = 21\%$.

Även testet av home bias påverkan på avkastningen ger ett visst stöd för den tidigare forskningen. Att det är bättre att vara home bias under perioden med positiv avkastning får stöd av att en persons investeringar är home bias är sämre diversifierad. Nordén (2010) hävdar dock att home bias kan innebära en högre riskexponering utan att det skulle leda till en förändrad avkastning. Något som stämmer väl in på en generellt underdiversifierad portfölj, dock visar vårt test inte på några sådana tendenser. Vårt test antyder snarare att den positiva effekten av home bias vid uppgång överstiger den negativa effekten vid nedgång. I det sammanhanget är det dock viktigt att ha i åtanke att endast den inbördes ordningen mellan att vara home bias och att inte vara det är säkerställt, inte de exakta siffrorna.

De här procentsatserna är inte statistisk säkrade, endast den inbördes ordningen mellan dem är säkrad enligt testet nedan.

Ranks

Home bias		N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Home bias under eller lika med genomsnittet	29309	33125,40	9,71E8
	Home bias över genomsnittet	42401	37742,64	1,60E9
	Total	71710		

Test Statistics^a

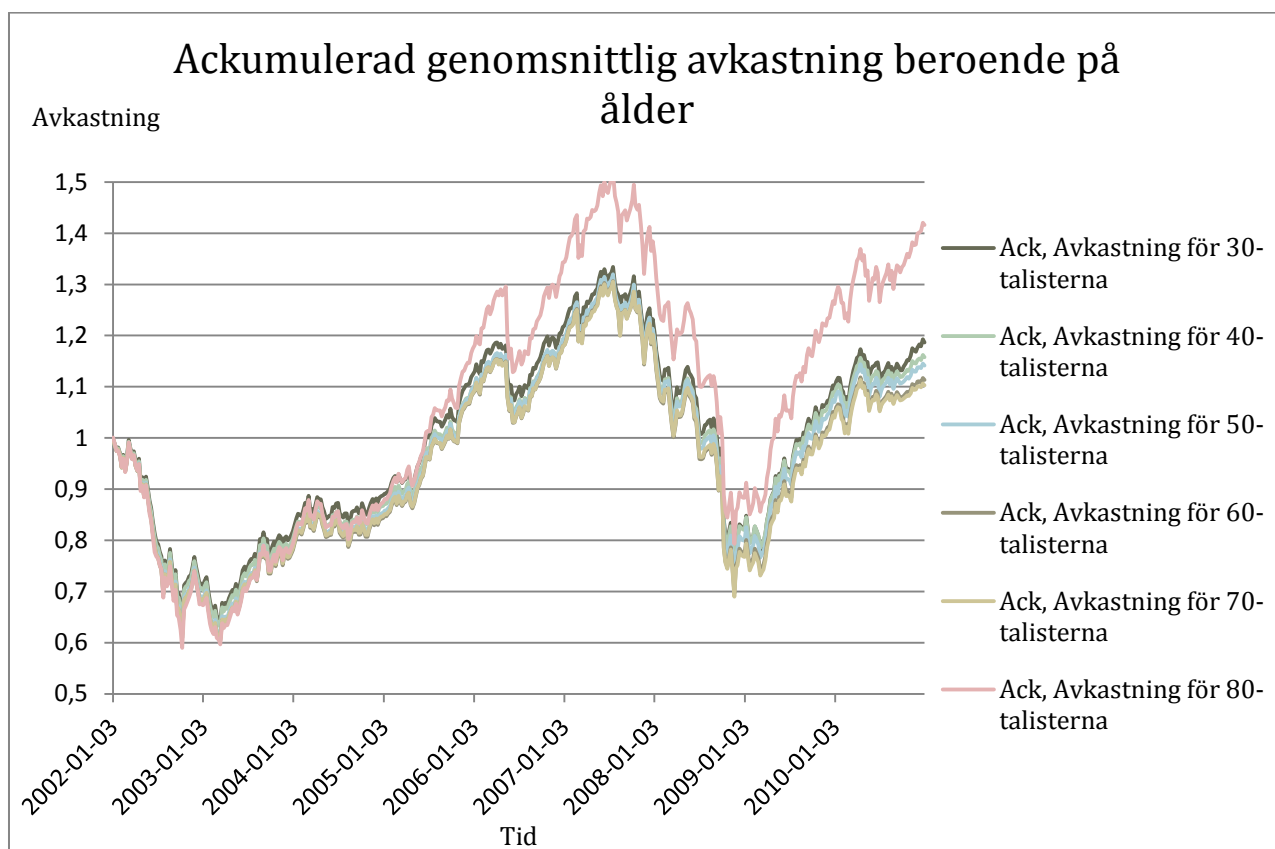
	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	5,413E8
Wilcoxon W	9,709E8
Z	-29,362
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: Home bias

H_0 kan förkastas. Personer som har home bias har ett bättre resultat än genomsnittet under perioden med positiv utveckling.

Avkastning för olika ålderskohorter

Enligt studier kring financial illiteracy för olika åldrar är det främst de yngsta och äldsta som är sämst på att fatta rationella finansiella beslut (Almenberg och Widmark, 2011). I vår testperiod finns ingen signifikans för att det skulle resultera i några skillnader i avkastning på pensionssparandet. De olika ålderskategorierna kovarierar olika starkt med andra parametrar (se rubrik *spearman korrelation mellan våra variabler*). Om slutsatsen kring financial illiteracy hålls som sann (att de yngsta och äldsta är sämst på att fatta finansiella beslut), tycks inte financial illiteracy ha någon påverkan på avkastningen på premiepensionen. Variationerna inom grupperna tycks vara väldigt stor (svaga signifikansnivåer), något som kan förklaras bland annat av att en tredjedel av vår finansiella beslutsförmåga bestäms av genetiska förutsättningar (Barnea et al., 2010). Möjligen är det rimligt att anta att många av de som klassas som finansiell illiterate inte är aktiva och därmed har sina pengar i sjunde AP-fonden. Effekterna skulle då bli mindre synliga.



Även om grafen ovan antyder att 80-talisternas avkastning skulle vara mer volatil än övriga ålderskategorier kan denna slutsats inte dras (se tabeller med Mann-Whitney U signifikans för respektive ålderskategori nedan).

30-talister

Ranks

	Född på 30-talet	N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Är ej född på 30-talet	62202	31884,85	1,98E9
	Är född på 30-talet	1546	31458,11	48634245,50
	Total	63748		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	47438414,500
Wilcoxon W	48634245,500
Z	-,901
Asymp. Sig. (2-tailed)	,368

a. Grouping Variable: Född på 30-talet

H_0 kan ej förkastas. 30-talister har samma avkastning som genomsnittet under perioden.

40-talister

Ranks

	Född på 40-talet	N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Är ej född på 40-talet	51181	31883,06	1,63E9
	Är född på 40-talet	12567	31839,65	4,00E8
	Total	63748		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	3,212E8
Wilcoxon W	4,001E8
Z	-,237
Asymp. Sig. (2-tailed)	,813

a. Grouping Variable: Född på 40-talet

H_0 kan ej förkastas. 40-talister har samma avkastning som genomsnittet under perioden.

50-talister

Ranks

	Född på 50-talet	N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Är ej född på 50-talet	50729	31844,04	1,62E9
	Är född på 50-talet	13019	31993,18	4,17E8
	Total	63748		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	3,287E8
Wilcoxon W	1,615E9
Z	-,825
Asymp. Sig. (2-tailed)	,409

a. Grouping Variable: Född på 50-talet

H_0 kan ej förkastas. 50-talister har samma avkastning som genomsnittet under perioden.

60-talister

Ranks

	Född på 60-talet	N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Är ej född på 60-talet	49160	31892,25	1,57E9
	Är född på 60-talet	14588	31814,68	4,64E8
	Total	63748		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	3,577E8
Wilcoxon W	4,641E8
Z	-,447
Asymp. Sig. (2-tailed)	,655

a. Grouping Variable: Född på 60-talet

H_0 kan ej förkastas. 60-talister har samma avkastning som genomsnittet under perioden.

70-talister

Ranks

	Född på 70-talet	N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Är ej född på 70-talet	50108	31871,35	1,60E9
	Är född på 70-talet	13640	31886,09	4,35E8
	Total	63748		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	3,416E8
Wilcoxon W	1,597E9
Z	-,083
Asymp. Sig. (2-tailed)	,934

a. Grouping Variable: Född på 70-talet

H_0 kan ej förkastas. 70-talister har samma avkastning som genomsnittet under perioden.

80-talister

Ranks

	Född på 80-talet	N	Mean Rank	Sum of Ranks
Veckoavkastning (inflationsjusterad)	Är ej född på 80-talet	55518	31866,16	1,77E9
	Är född på 80-talet	8230	31930,78	2,63E8
	Total	63748		

Test Statistics^a

	Veckoavkastning (inflationsjusterad)
Mann-Whitney U	2,280E8
Wilcoxon W	1,769E9
Z	-,297
Asymp. Sig. (2-tailed)	,766

a. Grouping Variable: Född på 80-talet

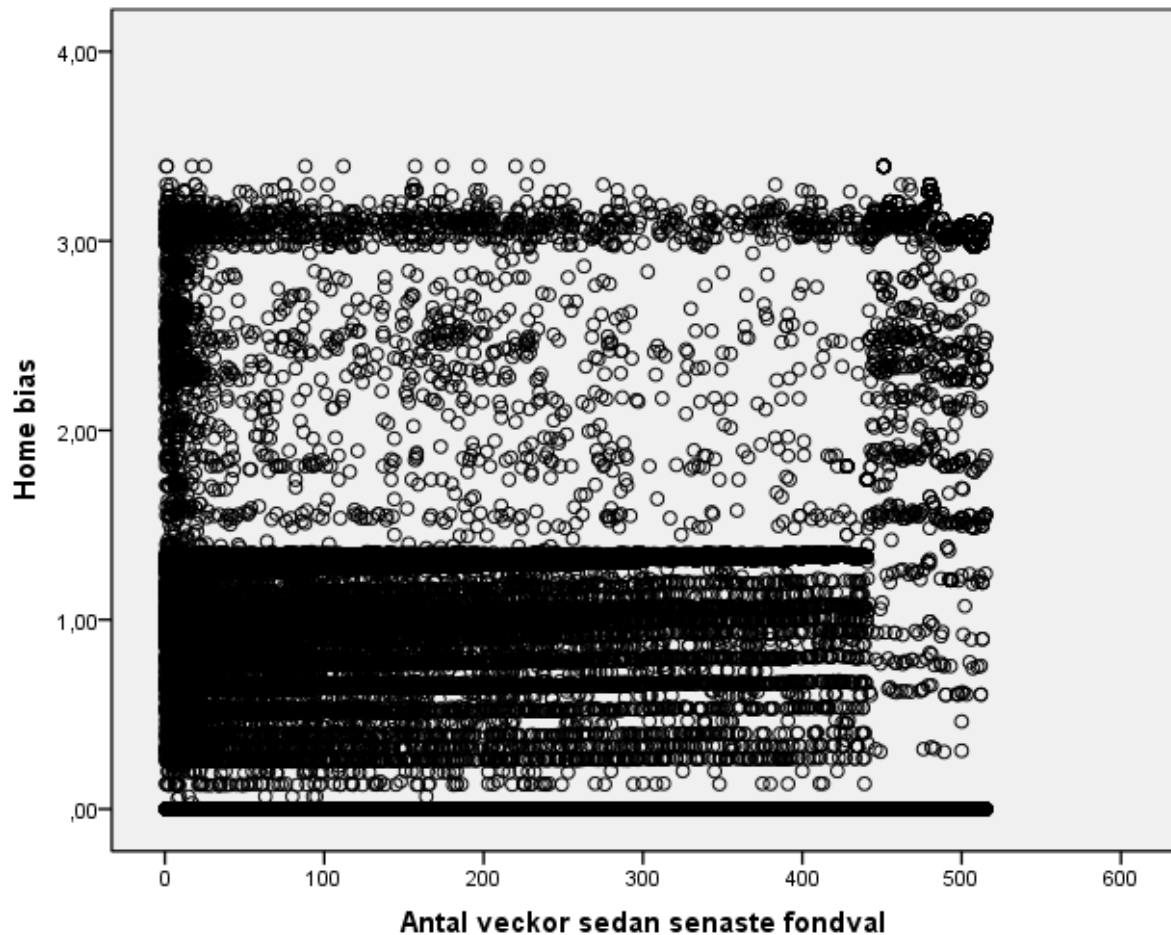
H_0 kan ej förkastas. 80-talister har samma avkastning som genomsnittet under perioden.

Spearman korrelation mellan våra olika variabler

			Correlations			
			Ålder	Home bias	Antal veckor sedan senaste fondval	Veckoränta (inflationsjusterad)
Spearman's rho	Ålder	Correlation Coefficient	1,000	-,041**	,112**	,003
		Sig. (2-tailed)	.	,000	,000	,429
		N	70628	70628	70628	70628
		Correlation Coefficient	-,041**	1,000	,388**	,022**
	Home bias	Sig. (2-tailed)	,000	.	,000	,000
		N	70628	70628	70628	70628
		Correlation Coefficient	,112**	,388**	1,000	,050**
		Sig. (2-tailed)	,000	,000	.	,000
	Antal veckor sedan senaste fondval	N	70628	70628	70628	70628
		Correlation Coefficient	,003	,022**	,050**	1,000
		Sig. (2-tailed)	,429	,000	,000	.
		N	70628	70628	70628	70628
Veckoränta (inflationsjusterad)	Correlation Coefficient	,003	,022**	,050**	1,000	
	Sig. (2-tailed)	,429	,000	,000	.	
	N	70628	70628	70628	70628	
	Correlation Coefficient	,003	,022**	,050**	1,000	
Sig. (2-tailed)	,429	,000	,000	.		
N	70628	70628	70628	70628		

** . Correlation is significant at the 0.01 level (2-tailed).

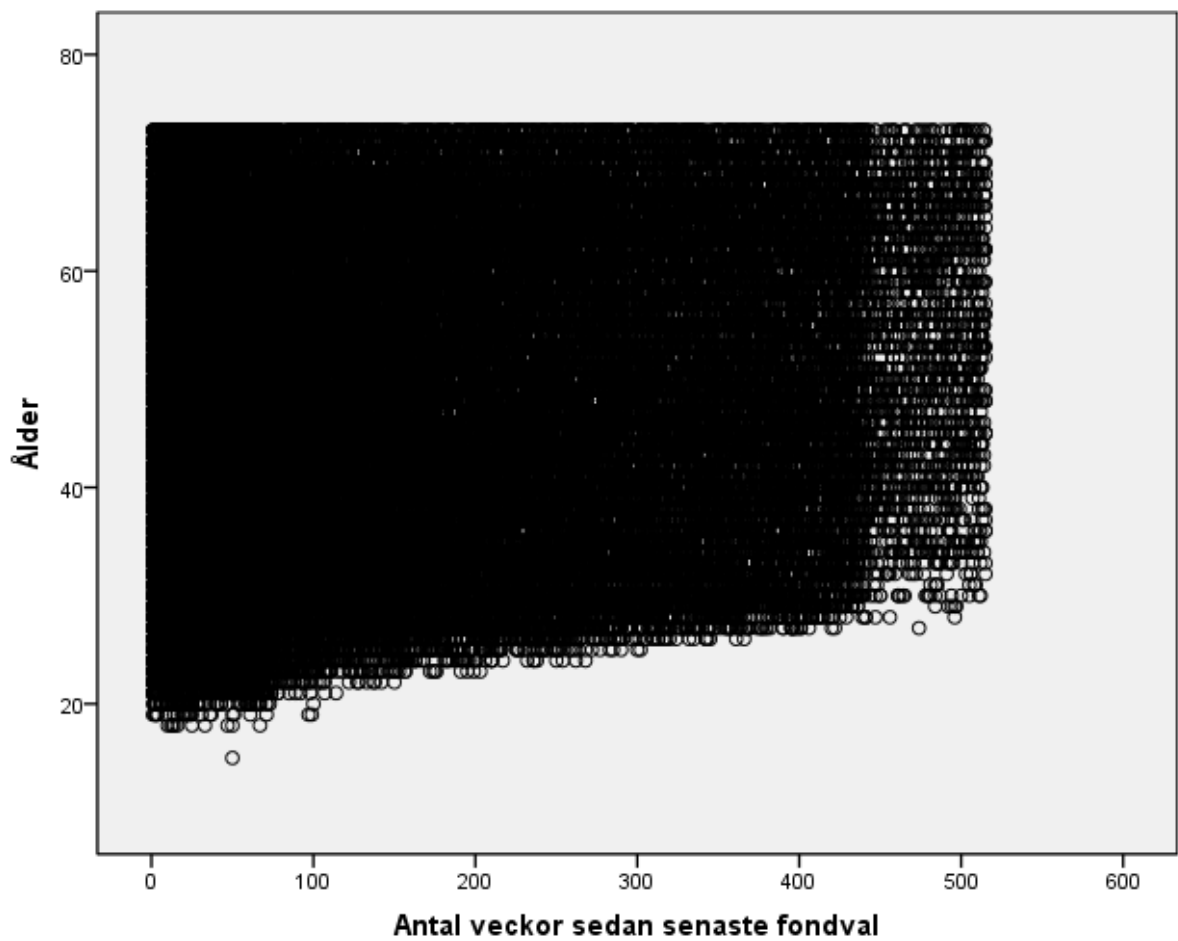
Tabellen ovan visar att bland de variabler vi hittat korrelation mellan är sambandet starkast (Spearman korrelation större än 0,1) mellan "Ålder" och "Veckor sedan senaste fondbytet" samt "Home bias" och "Veckor sedan senaste fondbytet". Endast dessa väljer vi att studera djupare.



Grafen ovan antyder att det finns de som aktivt väljer mycket svenska fonder men även en grupp som aktivt väljer att inte alls placera sina PPM-pengar i svenska fonder. Typfallet är att en sparare har cirka hälften av sina pensionspengar i svenska fonder. Korrelationen mellan dessa två (0,388) innebär att det finns en positiv korrelation, när den ena variabeln ökar gör även den andra det. Ju mindre frekvent en sparare byter ut sina fonder, desto mer home bias har denne. Detta samband är svårt att utläsa ur grafen, men det är samma typ av resultat som Nordén (2010) fann då han noterade att de investerare som har mer home bias tenderar att vara mer långsiktiga i sitt sparande. Nordén noterar även det motsatta; att personer som ofta byter sina fonder i högre grad också tenderar att vara mindre home biased, något han kallar *the competence effect*.

Grafen visar också att även bland de som mest frekvent byter ut sina fonder, finns det de som aktivt väljer att endast spara sina pengar i fonder som endast investerar i svenska värdepapper.

Ålder och veckor sedan senaste bytet



Korrelationen mellan ålder och antal veckor sedan senaste fondval (0,112) är förvisso stark, men vår tolkning av denna graf är att korrelationen beror på att de yngsta i urvalsgruppen inte varit inne i pensionssystemet tillräckligt länge för att kunna t varit inaktiva i 500 veckor. Av den anledningen dras inga slutsatser av detta resultat.

Slutsats

Syftet med denna uppsats var att utreda vilka parametrar som skiljer de aktiva spararna åt, vilket eller vilka beteenden som har störst påverkan på avkastningen från premiepensionen. Något oväntat var det ofta de som hade en, av tidigare forskning ansedd, ofördelaktig diversifiering som presterade bäst. Att använda hög historisk avkastning som beslutsvariabel i jakten på hög framtida avkastning verkar vara en bra strategi vid val av pensionsfonder, i synnerhet i högkonjunktur. Enligt mycket tidigare forskning skall det vara omöjligt att fatta rationella finansiella beslut på den typen av information. Vår slutsats kring detta är att, precis som Dahlquist hävdar, en hög historisk avkastning är ett tecken på en skicklig förvaltare.

1/n bias kritiserar mycket hårt i många forskningsrapporter och är allmänt betraktat som något väldigt negativt. Våra tester visar dock inga tecken på att de som har 1/n bias har sämre avkastning än de som inte har det.

Avkastningen mellan de som är aktiva skiljer sig väldigt mycket åt beroende på främst hur aktiva de är. I genomsnitt har de passiva spararna som har sina pengar i sjunde AP-fonden något sämre årlig avkastning än de som är aktiva. Bland de som är aktiva är spridningen på avkastningen väldigt stor. De som verkar prestera sämst är de som en gjort ett val och därefter aldrig brytt som om att följa upp det valet.

Starkast stöd ger våra resultat för den forskning som hävdar att det är fördelaktigt att på lång sikt placera sina pensionspengar i fonder med hög volatilitet. Trenden verkar vara att även om dessa fonder tappar mycket värde under perioder med negativ avkastning, ökar deras värde proportionellt sett ännu mer under perioder med positiv avkastning.

Det är lätt att föreställa sig att många av de som har kort tid kvar till pension flyttar sina pengar till mindre volatila fonder och att de som har väldigt långt kvar till sin pension har mindre incitament att vara aktiva. De ska dessutom enligt forskningen höra till dem som är sämst på att fatta rationella finansiella beslut. Detta borde innebära att de har en lägre avkastning än genomsnittet. Vårt test visar inga tendenser på att detta skulle vara sant.

Mycket forskning kring vilka beslutsregler som bör gälla på finansmarknaden bör skilja sig åt mellan olika marknader. Att handla med fonder skiljer sig mycket åt från handel med enskilda värdepapper som är oberoende av andra värdepapper.

Vi föreställde oss att vi skulle hitta starka samband mellan olika beteenden. Till exempel att de som byter sina fonder mest frekvent skulle vara de som visar minst tecken på olika psykologiska bias och att de som använder historisk avkastning som beslutsvariabel skulle vara överrepresenterade i 1/n kategorin. Den enda korskorrelation vi fann som vi tycker är intressant är den mellan home bias och bytesfrekvens. Det tycks finnas de som aktivt väljer bort svenska fonder helt och de som aktivt endast väljer svenska fonder.

Många forskare inom detta ämne kritiserar systemet och kommer med förslag på alternativ till hur det svenska pensionssystemet istället borde vara utformat. Vi upplever inte att våra kunskaper är tillräckligt stora för att komma med någon sådan rekommendation. Möjligtvis kan vi tycka att en inaktivitet längre än ett år borde leda till att sparandet per automatik flyttas till sjunde AP-fonden, och att spararen klassas som inaktiv.

Istället skulle vi vilja komma med en rekommendation baserat på våra tester:

- En sparare som vet med sig att hon inte kommer vara aktiv bör välja sjunde AP-fonden och stanna där.

En sparare som vill vara aktiv bör under perioder med uppgång välja de historiskt bäst presterande fonderna och vid nedgång flytta över sitt sparande till räntefonder alternativt sjunde AP-fonden.

Slutdiskussion

Vår undersökning har varit ganska explorativ till sin karaktär vilket gjort att vi löpande förändrat metoden under arbetets gång. När vi nu ser tillbaks på arbetet så är det främst två saker som bör förbättras.

Exakt data på alla fondavkastningar bör skaffas från Pensionsmyndigheten. Vi har inte haft några ekonomiska medel till vårt förfogande utan har arbetat med den data som myndigheten kunnat skicka utan att lägga ner extra tid.

Våra kunskaper i statistik har inte varit tillräckligt goda för att använda avancerade regressionsmetoder som GMM (general method of moments) en vanlig regressionsmetod inom econometri. Förmodligen skulle vi använda parvisa jämförelser av investeringar med likartad tidshorisont för att komma runt problemen med icke-linjäritet.

Ett intressant uppslag att testa vore att köra simulationer med hypotetiska portföljer som har olika former av bias. Portföljerna skulle i så fall motsvara verkliga portföljer med undantaget att fonderna valts slumpmässigt. Dessa hypotetiska portföljer skulle sedan utgöra en kontrollgrupp för att separera effekten av själva biaset och effekten från spararens smarta eller inte så smarta fondval.

Referenser

Almenberg Johan, Widmark Olof, 2011, *Räknefärdighet och finansiell förmåga*, Preliminära resultat från Finansinspektionens konsumentundersökning 2010.

Bali Turan G., Demirtas K. Ozgur, Levy Haim, Wolf Avner, 2009, *Bonds versus stocks: investors' age and risk taking*, Journal of Monetary Economics 56 (2009) pp. 817-830

Barnea Amir, Cronqvist Henrik, Siegel Stephan, *Nature or nurture: What determines investor behavior*, 2010, Journal of Financial Economics 98 (2010), pp. 583-604

Benartzi Shlomo, Thaler Richard H., 2007, *Heuristics and Biases in Retirement Savings Behavior*, The Journal of Economic Perspectives, ISSN 0895-3309, 07/2007, Vol. 21, No. 3, pp. 81-104

Butler Kirt C., Domian Dale L., 1993. *Financial Services Review*, 2(1): 41-49, JAI Press Inc. ISSN: 1057-0810

Cesarini David, Sandewall Örjan, Johannesson Magnus, 2003, *Confidence Interval Estimation Task and the Economics of Overconfidence*, SSE/EFI Working Paper Series in Economics and Finance, No 535

Cronqvist Henrik, 2006, *ADVERTISING AND PORTFOLIO CHOICE, THE UNIVERSITY OF CHICAGO, A DISSERTATION SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF BUSINESS IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY*

Dahlquist Magnus, Martinez José Vicente, 2011, *Investor Inattention: A Hidden Cost of Choice in Pension Plans*,

Dahlquist Magnus, Martinez José Vicente, Söderlind Paul, 2011, *Individual Investor Activity and Performance*,

De Long Bradford J., Shleifer Andrei, Summers Lawrence H. and Waldmann Robert J., 1990, *Journal of Political Economy*, Vol 98, No 4, pp. 703-738

Ekholm Anders, Pasternack Daniel, 2007, *Overconfidence and Investor Size*, Department of Finance, Swedish School of Economics and Business Administration

Hedeström Ted Martin, Svedsäter Henrik, Gärling Tommy, 2009, *Naive Diversification in the Swedish Premium Pension Scheme: Experimental Evidence*, Applied Psychology, 2009, 58 (3), 403-417

Jagric Timotej, Podobnik Boris, Kolanovic Marko, 2005, *Does the Efficient Market Hypothesis Hold?*, Eastern European Economics, Vol. 43, no. 4, July-August 2005, pp. 79-103

Newbold, P., 2007, *Statistics for Business and Economics*, Pearson Prentice-Hall, Upper Saddle River, New York.

Nordén Lars, 2010, *Individual home bias, portfolio churning and performance*, The European Journal of Finance, Vol. 16, No. 4, June 2010, 329-351

Palme Mårten, Sundén Annika, Söderlind Paul, 2005, *INVESTMENT CHOICE IN THE SWEDISH PREMIUM PENSION PLAN*, Center for Retirement Research at Boston College

Palomino Fredric, Sadrieh Abdolkarim, 2011, *Overconfidence and delegated portfolio management*, *Journal of Financial Intermediation* 20 (2011), pp. 159-177

Timmermann Allan, Granger Clive W.J., 2004, *Efficient market hypothesis and forecasting*, *International Journal of Forecasting* 20 (2004) pp. 15-27

Weaver Kent R., 2004, *Design and Implementation Issues in Swedish Individual Pension Accounts*, *Social Security Bulletin*, Vol. 65, No 4, pp. 38-57

<https://www.avanza.se/aza/fonder/fondguiden.jsp?orderbookId=363> hämtad 2012-12-20, klockan 18:17

<http://www.pensionsmyndigheten.se/FondArsstatistik2010.html>, hämtad 2012-01-08 klockan 12:36

Bilaga A - Systeminställningar och programkod

Hårdvara

Processor: Intel® Core™ i7 CPU 920 @ 2,67 GHz, 2.67GHz

Installerat minne (RAM): 16 GB

Mjukvara

IBM SPSS Statistics 19

MySQL Server 5.5

my.ini

```
[client]
port = 3306
socket = /tmp/mysql.sock
interactive-timeout = 86400
wait-timeout = 86400
```

```
[mysql]
default-character-set=latin1
```

```
[mysqld]
port = 3306
socket = /tmp/mysql.sock
interactive-timeout = 86400
wait-timeout = 86400
basedir="C:\Program Files\MySQL\MySQL Server 5.5\"
datadir="C:\ProgramData\MySQL\MySQL Server 5.5\data\"
character-set-server=latin1
```

```
back_log = 50
max_connections = 10
max_connect_errors = 10
table_open_cache = 2048
max_allowed_packet = 16M
```

```
binlog_cache_size = 1M
max_heap_table_size = 1000M
read_buffer_size = 2M
read_rnd_buffer_size = 16M
sort_buffer_size = 8M
join_buffer_size = 64M
thread_cache_size = 8
thread_concurrency = 8
query_cache_size = 64M
query_cache_limit = 2M
ft_min_word_len = 4
```

```
default-storage-engine = MYISAM
```

```
thread_stack = 192K
transaction_isolation = REPEATABLE-READ
tmp_table_size = 64M
log-bin=mysql-bin
binlog_format=mixed
slow_query_log
long_query_time = 2
server-id = 1
key_buffer_size = 32M
innodb_additional_mem_pool_size = 16M
innodb_buffer_pool_size = 2G
innodb_data_file_path = ibdata1:10M:autoextend
innodb_write_io_threads = 8
innodb_read_io_threads = 8
innodb_thread_concurrency = 16
innodb_flush_log_at_trx_commit = 1
innodb_log_buffer_size = 8M
```


innodb_log_file_size = 256M
innodb_log_files_in_group = 3
innodb_max_dirty_pages_pct = 90
innodb_lock_wait_timeout = 120

[mysqldump]
max_allowed_packet = 16M

[mysql]
no-auto-rehash

[myisamchk]
key_buffer_size = 512M
sort_buffer_size = 512M
read_buffer = 8M
write_buffer = 8M

[mysqlhotcopy]
interactive-timeout = 86400

[mysqld_safe]
open-files-limit = 8192

Programkod MySQL Stored procedures

init.sql

```
# Load data
source c:/Users/Andreas/Documents/init_database.sql
SELECT '**** init_database done ***';
INSERT INTO ppm.status_log VALUES ('Init database done.', 0);

# Add poorly formatted data from expired funds.
source c:/Users/Andreas/Documents/add_expired_funds.sql
SELECT '**** add_expired_funds done ***';
INSERT INTO ppm.status_log VALUES ('Add expired funds done.', 0);

# Transfrom the transaction log to one transaction per row.
source c:/Users/Andreas/Documents/find_transactions.sql
SELECT '**** find_transactions done ***';
INSERT INTO ppm.status_log VALUES ('Find transactions done.', 0);

# Calculate the yearly yields of funds
source c:/Users/Andreas/Documents/calculate_fund_yield.sql
SELECT '**** calculate_fund_yield done ***';
INSERT INTO ppm.status_log VALUES ('Calulate fund yeild done.', 0);

# Calculate top 25 funds
source c:/Users/Andreas/Documents/calculate_top_25.sql
SELECT '**** calculate_top_25 done ***';
INSERT INTO ppm.status_log VALUES ('Find top 25 funds done.', 0);

# Estimate weekly inflation
source c:/Users/Andreas/Documents/estimate_weekly_inflation.sql
SELECT '**** estimate_weekly_inflation done ***';
INSERT INTO ppm.status_log VALUES ('Estimate of weekly inflation
done.', 0);

CREATE INDEX kpi_w_index
  ON ppm.kpi (week)
  USING BTREE;

# Load procedure for AP7-Såfa
source c:/Users/Andreas/Documents/create_yield_for_ap7_safa.sql

# Load the procedure for calculating yields.
source c:/Users/Andreas/Documents/create_weekly_data.sql

# Calculate the yeild on the transactions on a weekly basis.
source c:/Users/Andreas/Documents/calculate_interest.sql
```

```
SELECT '**** calculate_interest done ***';
INSERT INTO ppm.status_log VALUES ('Calculation of weekly interest
done.', 0);
```

```
CREATE INDEX result_w_index
  ON ppm.result (week)
  USING BTREE;
```

```
CREATE INDEX result_p_index
  ON ppm.result (person_id)
  USING BTREE;
```

```
SELECT '**** result indexes created ***';
INSERT INTO ppm.status_log VALUES ('Result indexes created.', 0);
```

```
# Calculate yearly activity
source c:/Users/Andreas/Documents/find_frequent_traders.sql
SELECT '**** find_frequent_traders done ***';
INSERT INTO ppm.status_log VALUES ('Find frequent traders done.',
0);
```

```
# Calculate the home bias
source c:/Users/Andreas/Documents/calculate_home_bias.sql
SELECT '**** home bias calculation done ***';
INSERT INTO ppm.status_log VALUES ('Calculate home bias done.', 0);
```

```
# Move results back to disk.
alter table ppm.result engine=myisam;
SELECT '**** All done. Result committed to disk ***';
INSERT INTO ppm.status_log VALUES ('All done. Data committed to
disk', 0);
```

init_database.sql

```
DROP TABLE ppm.transaction_log;
CREATE TABLE ppm.transaction_log
(
    person_id BIGINT(10) UNSIGNED,
    fund_id MEDIUMINT(6) UNSIGNED,
    share TINYINT UNSIGNED,
    transaction_date DATE,
    age TINYINT(2) UNSIGNED
);

DROP TABLE ppm.fund_info;
CREATE TABLE ppm.fund_info
(
    fund_manager VARCHAR(100),
    fund_id MEDIUMINT(6) UNSIGNED,
    fund_name VARCHAR(100),
    currency CHAR(3),
    fund_fee DECIMAL(3, 2),
    ter DECIMAL(3,2),
    tka DECIMAL(3,2),
    fund_cathegory VARCHAR(100),
    fund_org_num CHAR(12),
    fund_in_fund CHAR(1),
    result_dependent_fee CHAR(1),
    delayed_value_report CHAR(1),
    has_spread CHAR(1),
    extended CHAR(1),
    start_date DATE,
    end_date DATE,
    status TINYINT UNSIGNED
);

DROP TABLE ppm.fund_choice;
CREATE TABLE ppm.fund_choice
(
    person_id BIGINT(10) UNSIGNED,
    age TINYINT(2) UNSIGNED,
    transaction_date DATE,
    fund_id1 MEDIUMINT(6) UNSIGNED,
    share1 TINYINT UNSIGNED,
    fund_id2 MEDIUMINT(6) UNSIGNED,
    share2 TINYINT UNSIGNED,
    fund_id3 MEDIUMINT(6) UNSIGNED,
    share3 TINYINT UNSIGNED,
    fund_id4 MEDIUMINT(6) UNSIGNED,
    share4 TINYINT UNSIGNED,
    fund_id5 MEDIUMINT(6) UNSIGNED,
    share5 TINYINT UNSIGNED
);
```

```

DROP TABLE ppm.fund_return;
CREATE TABLE ppm.fund_return
(
    week          DATE,
    fund_id       MEDIUMINT(6) UNSIGNED,
    r             DECIMAL(10, 9)
)
ENGINE=MEMORY;

```

```

DROP TABLE ppm.fund_return2;
CREATE TABLE ppm.fund_return2
(
    week          DATE,
    fund_id       MEDIUMINT(6) UNSIGNED,
    r             DECIMAL(10, 9)
)
ENGINE=MEMORY;

```

```

DROP TABLE ppm.temp_yield;
CREATE TABLE ppm.temp_yield
(
    week          DATE,
    r1            DECIMAL(10, 9),
    r2            DECIMAL(10, 9),
    r3            DECIMAL(10, 9),
    r4            DECIMAL(10, 9),
    r5            DECIMAL(10, 9)
)
ENGINE=MEMORY;

```

```

DROP TABLE ppm.fund_year_return;
CREATE TABLE ppm.fund_year_return
(
    year MEDIUMINT(4) UNSIGNED,
    fund_id MEDIUMINT(6) UNSIGNED,
    r      DECIMAL(12, 9)
);

```

```

DROP TABLE ppm.result;
CREATE TABLE ppm.result
(
    person_id          BIGINT(10) UNSIGNED,
    week               DATE,
    age                TINYINT UNSIGNED,
    l_n_bias           TINYINT UNSIGNED,
    over_confidence    TINYINT UNSIGNED,
    home_bias          FLOAT(7,4),
    num_trades         TINYINT UNSIGNED,
    active_choice      TINYINT UNSIGNED,
    weeks_since_choice INT          UNSIGNED,

```

```

    born_1930          TINYINT UNSIGNED,
    born_1940          TINYINT UNSIGNED,
    born_1950          TINYINT UNSIGNED,
    born_1960          TINYINT UNSIGNED,
    born_1970          TINYINT UNSIGNED,
    born_1980          TINYINT UNSIGNED,
    born_1990          TINYINT UNSIGNED,
    over_average       TINYINT UNSIGNED,
    over_default_option TINYINT UNSIGNED,
    r                  DECIMAL(13, 9)
)
ENGINE=MEMORY;

DROP TABLE ppm.top_25;
CREATE TABLE ppm.top_25
(
    year MEDIUMINT(4) UNSIGNED,
    fund_id MEDIUMINT(6) UNSIGNED
)
ENGINE=MEMORY;

DROP TABLE ppm.temp_kpi;
CREATE TABLE ppm.temp_kpi
(
    kpi_month DATE,
    kpi_change DECIMAL(4,2)
);

DROP TABLE ppm.kpi;
CREATE TABLE ppm.kpi
(
    week DATE,
    kpi DECIMAL(10, 9)
)
ENGINE=MEMORY;

DROP TABLE ppm.status_log;
CREATE TABLE ppm.status_log
(
    message VARCHAR(256),
    number INT
);

INSERT INTO ppm.status_log VALUES ('Missing fund yield', 0);

TRUNCATE TABLE ppm.temp_kpi;
LOAD DATA
LOCAL INFILE 'C:\\Users\\Andreas\\Downloads\\Fondbyte\\PR0101C5.txt'

```

```

REPLACE INTO TABLE ppm.temp_kpi
FIELDS TERMINATED BY ';'
IGNORE 4 LINES;

TRUNCATE TABLE ppm.transaction_log;
LOAD DATA
LOCAL INFILE 'C:\\Users\\Andreas\\Downloads\\Fondbyte\\Fondbyte.txt'
REPLACE INTO TABLE ppm.transaction_log
FIELDS TERMINATED BY '\t'
IGNORE 1 LINES;

TRUNCATE TABLE ppm.fund_info;
LOAD DATA
LOCAL INFILE
'C:\\Users\\Andreas\\Downloads\\Fondbyte\\Total_fondlista-2011-09-
12.csv'
REPLACE INTO TABLE ppm.fund_info
FIELDS TERMINATED BY ';'
IGNORE 2 LINES;

TRUNCATE TABLE ppm.fund_return;
LOAD DATA
LOCAL INFILE 'C:\\Users\\Andreas\\Downloads\\Fondbyte\\Avk_vecka
Aktuella.txt'
REPLACE INTO TABLE ppm.fund_return
IGNORE 1 LINES;

TRUNCATE TABLE ppm.fund_return2;
LOAD DATA
LOCAL INFILE 'C:\\Users\\Andreas\\Downloads\\Fondbyte\\Avk_vecka 5-
75.txt'
REPLACE INTO TABLE ppm.fund_return2
IGNORE 1 LINES;

LOAD DATA
LOCAL INFILE 'C:\\Users\\Andreas\\Downloads\\Fondbyte\\Avk_vecka 75
-100.txt'
REPLACE INTO TABLE ppm.fund_return2
IGNORE 1 LINES;

LOAD DATA
LOCAL INFILE 'C:\\Users\\Andreas\\Downloads\\Fondbyte\\Avk_vecka
500000.txt'
REPLACE INTO TABLE ppm.fund_return2
IGNORE 1 LINES;

CREATE INDEX fund_return_f_index
ON ppm.fund_return (fund_id)
USING BTREE;

CREATE INDEX fund_return_w_index
ON ppm.fund_return (week)
USING BTREE;

```

```
DROP TABLE ppm.home_bias_funds;
CREATE TABLE ppm.home_bias_funds AS SELECT fund_id FROM
ppm.fund_info WHERE fund_cathegory LIKE 'Sve%';
ALTER TABLE ppm.home_bias_funds ENGINE = MEMORY;
CREATE INDEX fund_home_bias_f_index
  ON ppm.fund_return (fund_id)
  USING BTREE;
```


add_expired_funds.sql

```
DELIMITER $$

DROP PROCEDURE IF EXISTS ppm.add_expired_funds$$

CREATE PROCEDURE ppm.add_expired_funds()

main: BEGIN

    # Delete any funds present here that are present in the main
    table.
    IF TRUE THEN
        remove_duplicate: BEGIN
            DECLARE done INT DEFAULT FALSE;
            DECLARE this_fund MEDIUMINT(6) UNSIGNED;
            DECLARE i INT DEFAULT 0;
            DECLARE fund CURSOR FOR
                SELECT DISTINCT fund_id
                FROM ppm.fund_return;
            DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

            OPEN fund;

            duplicates: LOOP
                FETCH fund INTO this_fund;

                IF done THEN
                    LEAVE duplicates;
                END IF;

                SET i = i + 1;

                SELECT i, this_fund;
                DELETE FROM ppm.fund_return2 WHERE fund_id = this_fund;

            END LOOP duplicates;

            CLOSE fund;

        END remove_duplicate;
    END IF;

    # Delete any date that occurs here that is not present in the main
    table.
    IF TRUE THEN
        remove_extra_dates: BEGIN
            DECLARE this_date DATE;
            DECLARE i INT DEFAULT 0;
            DECLARE hit, dud INT DEFAULT 0;
            DECLARE start_date DATE DEFAULT '1999-01-01';
            DECLARE end_date DATE DEFAULT '2011-12-01';

            date_loop: LOOP
```

```

SET this_date = start_date + INTERVAL i DAY;

IF DATEDIFF(end_date, this_date) <= 0 THEN
    LEAVE date_loop;
END IF;

SELECT COUNT(week) INTO hit FROM ppm.fund_return WHERE week =
this_date;
SELECT COUNT(week) INTO dud FROM ppm.fund_return2 WHERE week =
this_date;
# SELECT i, this_date, hit, dud;

IF hit = 0 AND dud > 0 THEN
    # SELECT 'Deleteing...';
    DELETE FROM ppm.fund_return2 WHERE week = this_date;
END IF;

SET i = i + 1;

END LOOP date_loop;
END remove_extra_dates;
END IF;

# Delete entries of the same fund the same week with different
yields.
# Implemented as a script instead of a delete since the join takes
so freaking long to resolve.

remove_double_yield: BEGIN
    DECLARE done INT DEFAULT FALSE;
    DECLARE w date;
    DECLARE f MEDIUMINT(6) UNSIGNED;
    DECLARE i INT DEFAULT 0;
    DECLARE fund CURSOR FOR
        SELECT week, fund_id FROM ppm.fund_return2 GROUP BY week,
fund_id HAVING COUNT(fund_id) > 1;
    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

    OPEN fund;

    double_loop: LOOP
        FETCH fund INTO w, f;
        SET i = i + 1;

        IF done THEN
            LEAVE double_loop;
        END IF;

        DELETE FROM ppm.fund_return2 WHERE week = w AND fund_id = f;

    END LOOP double_loop;
    SELECT i, 'double entries deleted...';

```

```
END remove_double_yield;

# Copy clean data into the main table.
INSERT INTO ppm.fund_return (week, fund_id, r) SELECT * FROM
ppm.fund_return2;
DROP TABLE ppm.fund_return2;

END main$$

DELIMITER ;

CALL ppm.add_expired_funds;
```

find_transactions.sql

```
TRUNCATE TABLE ppm.fund_choice;
DROP PROCEDURE IF EXISTS ppm.find_transactions;
DELIMITER $$

CREATE PROCEDURE ppm.find_transactions()
BEGIN
    DECLARE done INT DEFAULT FALSE;
    DECLARE f1, f2, f3, f4, f5 MEDIUMINT(6) UNSIGNED DEFAULT 0;
    DECLARE s1, s2, s3, s4, s5 TINYINT UNSIGNED DEFAULT 0;
    DECLARE total_share TINYINT UNSIGNED DEFAULT 0;
    DECLARE num_share TINYINT UNSIGNED DEFAULT 1;
    DECLARE p BIGINT(10) UNSIGNED;
    DECLARE f MEDIUMINT(6) UNSIGNED;
    DECLARE s TINYINT UNSIGNED;
    DECLARE d DATE;
    DECLARE a TINYINT(2) UNSIGNED;
    # Only one transaction per day is allowed so we get an ordered
list through:
    DECLARE log CURSOR FOR
        SELECT person_id, fund_id, share, transaction_date, age
        FROM ppm.transaction_log;
    # ORDER BY person_id, transaction_date;

    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

    OPEN log;

    # Traverse the data from top to bottom. It is ordered by age,
person_id, transaction_date
    en_person: LOOP

        FETCH log INTO p, f, s, d, a;

        IF done THEN
            LEAVE en_person;
        END IF;

        # SELECT p, f, s, d, a;

        CASE num_share
        WHEN 1 THEN
            SET f1 = f;
            SET s1 = s;
        WHEN 2 THEN
            SET f2 = f;
            SET s2 = s;
        WHEN 3 THEN
            SET f3 = f;
            SET s3 = s;
        WHEN 4 THEN
            SET f4 = f;
            SET s4 = s;
```

```

WHEN 5 THEN
    SET f5 = f;
    SET s5 = s;
ELSE
    # Should not happen. Debug:.
    SELECT p, a, d;
    SELECT f1,s1,f2,s2,f3,s3,f4,s4,f5,s5;
    SELECT num_share, total_share;
    LEAVE en_person;
END CASE;

SET total_share = s1 + s2 + s3 + s4 + s5;

IF total_share = 100 THEN
    # OK, we are done with this fund choice. Lets save it and reset
the state.
    INSERT INTO ppm.fund_choice
        VALUES( p, a, d, f1, s1, f2, s2, f3, s3, f4, s4, f5, s5);

    SET s1 = 0;
    SET s2 = 0;
    SET s3 = 0;
    SET s4 = 0;
    SET s5 = 0;
    SET f1 = 0;
    SET f2 = 0;
    SET f3 = 0;
    SET f4 = 0;
    SET f5 = 0;
    SET num_share = 1;
    ELSE
        SET num_share = num_share + 1;
    END IF;
END LOOP en_person;

CLOSE log;
END$$

DELIMITER ;

CALL ppm.find_transactions();

```

calculate_fund_yield.sql

```
DROP PROCEDURE IF EXISTS ppm.calculate_fund_yield;

DELIMITER $$

CREATE PROCEDURE ppm.calculate_fund_yield()
BEGIN
    DECLARE done INT DEFAULT FALSE;
    DECLARE f_id MEDIUMINT(6) UNSIGNED;
    DECLARE this_week, coming_week DATE;
    DECLARE this_r DECIMAL(10, 9);
    DECLARE year_yield DECIMAL(11, 9);
    DECLARE this_year INT;
    DECLARE fund CURSOR FOR
        SELECT DISTINCT fund_id
        FROM ppm.fund_return;
    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

    OPEN fund;

    loop_funds: LOOP
        FETCH fund INTO f_id;

        IF done THEN
            LEAVE loop_funds;
        END IF;

        # SELECT
        '*****';
        # SELECT f_id AS 'Processing fund_id:';

        calc_interest: BEGIN
            DECLARE done_calc INT DEFAULT FALSE;
            DECLARE this_fund CURSOR FOR
                SELECT DISTINCT week, r
                FROM ppm.fund_return
                WHERE fund_id = f_id
                ORDER BY week;
            DECLARE CONTINUE HANDLER FOR NOT FOUND SET done_calc = TRUE;

            OPEN this_fund;

            calc_loop: LOOP

                FETCH this_fund INTO this_week, this_r;

                IF done_calc THEN
                    LEAVE calc_loop;
                END IF;

                SET this_year = YEAR(this_week);
```

```

        # Find the first entry of a new year and start counting.
        IF DAYOFYEAR(this_week) <= 7 AND this_year >= 2002 AND
this_year < 2011 THEN
            SET year_yield = this_r;
            SET comming_week = ADDDATE(this_week, INTERVAL 1 WEEK);

            # Loop through weeks.
            week_loop: LOOP
                FETCH this_fund INTO this_week, this_r;

                IF done_calc THEN
                    LEAVE calc_loop;
                END IF;

                IF this_week != comming_week THEN
                    # Crap, hole in data. Skip this year.
                    SELECT 'Found a hole in data', f_id, this_week,
comming_week;
                    LEAVE week_loop;
                END IF;

                SET comming_week = ADDDATE(this_week, INTERVAL 1 WEEK);
                SET year_yield = year_yield * this_r;

                # Done with this year?
                IF YEAR(comming_week) > this_year THEN
                    # SELECT f_id, this_year, year_yield;
                    INSERT INTO ppm.fund_year_return VALUES (this_year,
f_id, year_yield);
                    LEAVE week_loop;
                END IF;

            END LOOP week_loop;
        END IF;
    END LOOP calc_loop;

    CLOSE this_fund;

END calc_interest;

END LOOP loop_funds;

CLOSE fund;

END$$

DELIMITER ;

CALL ppm.calculate_fund_yield();

```

calculate_top_25.sql

```
DROP PROCEDURE IF EXISTS ppm.calculate_top_25;

DELIMITER $$

CREATE PROCEDURE ppm.calculate_top_25()
BEGIN
    DECLARE y INT DEFAULT 2002;

    TRUNCATE TABLE ppm.top_25;

    WHILE y < 2011
    DO
        top_25: BEGIN
            DECLARE done INT DEFAULT FALSE;
            DECLARE last_value, i INT;
            DECLARE f MEDIUMINT(6) UNSIGNED;

            DECLARE fund CURSOR FOR
                SELECT fund_id
                FROM ppm.fund_year_return
                WHERE year=y ORDER BY r DESC;

            DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

            SELECT COUNT(*) INTO last_value FROM ppm.fund_year_return
WHERE year = y;

            SET last_value = last_value / 4;
            SET i = 1;

            SELECT y, last_value;

            OPEN fund;

            top_25_loop: LOOP

                FETCH fund INTO f;

                IF i > last_value THEN
                    LEAVE top_25_loop;
                END IF;

                SET i = i + 1;
                INSERT INTO ppm.top_25 VALUES (y, f);

            END LOOP top_25_loop;

            CLOSE fund;

        END top_25;

        SET y = y + 1;
    
```



```
END WHILE;  
END$$
```

```
DELIMITER ;
```

```
CALL ppm.calculate_top_25();
```

estimate_weekly_inflation.sql

```
DELIMITER $$

DROP PROCEDURE IF EXISTS ppm.estimate_weekly_inflation$$

CREATE PROCEDURE ppm.estimate_weekly_inflation()
BEGIN
  #
  # If the week is entirely within a month -> week_inflation =
  POW(month_inflation, 1/4)
  #
  # If the week lies within two months -> week_inflation =
  POW(weighted_average, 1/4)
  # Where: weighted_average = (month_1_inflation * days_in_month_1
  + month_2_inflation * days_in_month_2) / 7

  DECLARE week_end, week_start DATE;
  DECLARE done INT DEFAULT FALSE;
  DECLARE month_1_inflation DECIMAL(10, 9);
  DECLARE month_2_inflation DECIMAL(10, 9);
  DECLARE weighted_average DECIMAL(10, 9);
  DECLARE week_inflation DECIMAL(10, 9);
  DECLARE weeks CURSOR FOR
    SELECT DISTINCT week FROM ppm.fund_return WHERE week >=
2001-12-31;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

  TRUNCATE TABLE ppm.kpi;

  OPEN weeks;

  week_loop: LOOP
    FETCH weeks INTO week_end;

    IF done THEN
      LEAVE week_loop;
    END IF;

    SET week_start = DATE_SUB(week_end, INTERVAL 1 WEEK);

    IF MONTH(week_start) = MONTH(week_end) THEN
      SELECT 1 + kpi_change/100 INTO month_1_inflation
        FROM ppm.temp_kpi
        WHERE YEAR(kpi_month) = YEAR(week_end) AND
              MONTH(kpi_month) = MONTH(week_end);

      SET weighted_average = month_1_inflation;
    ELSE
      SELECT 1 + kpi_change/100 INTO month_1_inflation
        FROM ppm.temp_kpi
        WHERE YEAR(kpi_month) = YEAR(week_start) AND
              MONTH(kpi_month) = MONTH(week_start);
    END IF;
  END LOOP week_loop;
END
```

```

SELECT 1 + kpi_change/100 INTO month_2_inflation
FROM ppm.temp_kpi
WHERE YEAR(kpi_month) = YEAR(week_end) AND
      MONTH(kpi_month) = MONTH(week_end);

SET weighted_average = ((month_1_inflation * (7 -
DAYOFMONTH(week_end))) +
                        (month_2_inflation *
DAYOFMONTH(week_end))) / 7;

# SELECT week_start, week_end;
# SELECT month_1_inflation, month_2_inflation,
weighted_average;

END IF;

SET week_inflation = POW(weighted_average, 1/4);

INSERT INTO ppm.kpi VALUES (week_end, week_inflation);

END LOOP week_loop;

CLOSE weeks;

END$$

DELIMITER ;

CALL ppm.estimate_weekly_inflation();

```

create_yield_for_ap7_safa.sql

```
DELIMITER $$

DROP PROCEDURE IF EXISTS ppm.create_yield_for_ap7_safa$$

CREATE PROCEDURE ppm.create_yield_for_ap7_safa (IN ca TINYINT(2)
UNSIGNED,
                                                    IN cd DATE,
DATE,
                                                    IN nd
MEDIUMINT(6) UNSIGNED,
                                                    IN cf1
TINYINT UNSIGNED,
                                                    IN cs1
MEDIUMINT(6) UNSIGNED,
                                                    IN cf2
TINYINT UNSIGNED,
                                                    IN cs2
MEDIUMINT(6) UNSIGNED,
                                                    IN cf3
TINYINT UNSIGNED,
                                                    IN cs3
MEDIUMINT(6) UNSIGNED,
                                                    IN cf4
TINYINT UNSIGNED,
                                                    IN cs4
MEDIUMINT(6) UNSIGNED,
                                                    IN cf5
TINYINT UNSIGNED)
  IN cs5
#
# Special case if fund_id is 100010 (AP7 S afa) then the fund is a
# composite of
# 545541 (AP7 R antefond) AND 581371 (AP7 Aktiefond). Yay...
#
# The composition depends on age:
#
# age < 56:    100 % (AP7 Aktiefond) 0% (AP7 R antefond)
#
# age <= 56:   [(100 - 3 * (age - 55)) % AP7 Aktiefond] + [3 *
# (age - 55) % AP7 R antefond]
#

main: BEGIN
  DECLARE f_id MEDIUMINT(6) UNSIGNED;
  DECLARE w                                     DATE;
  DECLARE y                                     DECIMAL(10, 9) DEFAULT NULL;
# Yields over the time period.
  DECLARE y_ranta, y_aktie    DECIMAL(10, 9);
  DECLARE age                 INT;
  DECLARE i                   INT;
```

```

DECLARE data_year          INT DEFAULT 2011;
DECLARE ap7_aktie         MEDIUMINT(6) DEFAULT 581371;
DECLARE ap7_ranta        MEDIUMINT(6) DEFAULT 545541;
DECLARE ap7_safa         MEDIUMINT(6) DEFAULT 100010;

# SELECT 'ap7-safa', cd AS 'start date', nd AS 'end date';

# Fund choice 1.
IF cf1 = ap7_safa THEN
  fund1: BEGIN
    DECLARE done1 INT DEFAULT FALSE;
    DECLARE fund_yield1 CURSOR FOR
      SELECT week, fund_id, r
      FROM ppm.fund_return
      WHERE fund_id = ap7_aktie OR
             fund_id = ap7_ranta AND
             DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
             DATEDIFF(week, nd) <= 0    # the true starting
day.
      ORDER BY week;
    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done1 = TRUE;

    SET i = 0;
    SET y_ranta = NULL;
    SET y_aktie = NULL;

    OPEN fund_yield1;

    accumulator1: LOOP
      FETCH fund_yield1 INTO w, f_id, y;

      IF done1 THEN
        LEAVE accumulator1;
      END IF;

      SET age = ca - (data_year - YEAR(w));

      IF f_id = ap7_aktie THEN
        SET y_ranta = y;
      END IF;

      IF f_id = ap7_ranta THEN
        SET y_aktie = y;
      END IF;

      IF y_aktie IS NOT NULL AND y_ranta IS NOT NULL THEN
        SET i = i + 1;

        IF age < 56 THEN
          SET y = y_aktie;
        ELSE
          SET y = (y_aktie * (1 - 0.03 * (age - 55))) + (y_ranta
* 0.03 * (age - 55));

```

```

END IF;

# SELECT ca, YEAR(w), age;
# SELECT age, age - 55, y_aktie, y_ranta, y;

SET y_aktie = NULL;
SET y_ranta = NULL;

INSERT INTO ppm.temp_yield (week, r1) VALUES (w, y);
END IF;
END LOOP accumulator1;

CLOSE fund_yield1;
END fund1;
END IF;

# Fund choice 2.
IF cf2 = ap7_safa THEN
fund2: BEGIN
  DECLARE done2 INT DEFAULT FALSE;
  DECLARE fund_yield2 CURSOR FOR
    SELECT week, fund_id, r
    FROM ppm.fund_return
    WHERE fund_id = ap7_aktie OR
          fund_id = ap7_ranta AND
          DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
          DATEDIFF(week, nd) <= 0 # the true starting
day.
    ORDER BY week;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done2 = TRUE;

  SET i = 0;
  SET y_ranta = NULL;
  SET y_aktie = NULL;

  OPEN fund_yield2;

  accumulator2: LOOP
    FETCH fund_yield2 INTO w, f_id, y;

    IF done2 THEN
      LEAVE accumulator2;
    END IF;

    SET age = ca - (data_year - YEAR(w));

    IF f_id = ap7_aktie THEN
      SET y_ranta = y;
    END IF;

    IF f_id = ap7_ranta THEN

```

```

    SET y_aktie = y;
END IF;

IF y_aktie IS NOT NULL AND y_ranta IS NOT NULL THEN
    SET i = i + 1;

    IF age < 56 THEN
        SET y = y_aktie;
    ELSE
        SET y = (y_aktie * (1 - 0.03 * (age - 55))) + (y_ranta
* 0.03 * (age - 55));
    END IF;

    # SELECT ca, YEAR(w), age;
    # SELECT age, age - 55, y_aktie, y_ranta, y;

    SET y_aktie = NULL;
    SET y_ranta = NULL;

    INSERT INTO ppm.temp_yield (week, r2) VALUES (w, y);
END IF;
END LOOP accumulator2;

CLOSE fund_yield2;
END fund2;
END IF;

# Fund choice 3.
IF cf3 = ap7_safa THEN
    fund3: BEGIN
        DECLARE done3 INT DEFAULT FALSE;
        DECLARE fund_yield3 CURSOR FOR
            SELECT week, fund_id, r
            FROM ppm.fund_return
            WHERE fund_id = ap7_aktie OR
                fund_id = ap7_ranta AND
                DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
                DATEDIFF(week, nd) <= 0      # the true starting
day.
            ORDER BY week;
        DECLARE CONTINUE HANDLER FOR NOT FOUND SET done3 = TRUE;

        SET i = 0;
        SET y_ranta = NULL;
        SET y_aktie = NULL;

        OPEN fund_yield3;

        accumulator3: LOOP
            FETCH fund_yield3 INTO w, f_id, y;

```

```

IF done3 THEN
  LEAVE accumulator3;
END IF;

SET age = ca - (data_year - YEAR(w));

IF f_id = ap7_aktie THEN
  SET y_ranta = y;
END IF;

IF f_id = ap7_ranta THEN
  SET y_aktie = y;
END IF;

IF y_aktie IS NOT NULL AND y_ranta IS NOT NULL THEN
  SET i = i + 1;

  IF age < 56 THEN
    SET y = y_aktie;
  ELSE
    SET y = (y_aktie * (1 - 0.03 * (age - 55))) + (y_ranta
* 0.03 * (age - 55));
  END IF;

  # SELECT ca, YEAR(w), age;
  # SELECT age, age - 55, y_aktie, y_ranta, y;

  SET y_aktie = NULL;
  SET y_ranta = NULL;

  INSERT INTO ppm.temp_yield (week, r3) VALUES (w, y);
END IF;
END LOOP accumulator3;

CLOSE fund_yield3;
END fund3;
END IF;

```

```

# Fund choice 4.
IF cf4 = ap7_safa THEN
  fund4: BEGIN
    DECLARE done4 INT DEFAULT FALSE;
    DECLARE fund_yield4 CURSOR FOR
      SELECT week, fund_id, r
      FROM ppm.fund_return
      WHERE fund_id = ap7_aktie OR
      fund_id = ap7_ranta AND
      DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
      DATEDIFF(week, nd) <= 0      # the true starting
day.
      ORDER BY week;

```



```

DECLARE CONTINUE HANDLER FOR NOT FOUND SET done4 = TRUE;

SET i = 0;
SET y_ranta = NULL;
SET y_aktie = NULL;

OPEN fund_yield4;

accumulator4: LOOP
  FETCH fund_yield4 INTO w, f_id, y;

  IF done4 THEN
    LEAVE accumulator4;
  END IF;

  SET age = ca - (data_year - YEAR(w));

  IF f_id = ap7_aktie THEN
    SET y_ranta = y;
  END IF;

  IF f_id = ap7_ranta THEN
    SET y_aktie = y;
  END IF;

  IF y_aktie IS NOT NULL AND y_ranta IS NOT NULL THEN
    SET i = i + 1;

    IF age < 56 THEN
      SET y = y_aktie;
    ELSE
      SET y = (y_aktie * (1 - 0.03 * (age - 55))) + (y_ranta
* 0.03 * (age - 55));
    END IF;

    # SELECT ca, YEAR(w), age;
    # SELECT age, age - 55, y_aktie, y_ranta, y;

    SET y_aktie = NULL;
    SET y_ranta = NULL;

    INSERT INTO ppm.temp_yield (week, r4) VALUES (w, y);
  END IF;
END LOOP accumulator4;

CLOSE fund_yield4;
END fund4;
END IF;

# Fund choice 5.
IF cf5 = ap7_safa THEN
fund5: BEGIN

```

```

DECLARE done5 INT DEFAULT FALSE;
DECLARE fund_yield5 CURSOR FOR
    SELECT week, fund_id, r
    FROM ppm.fund_return
    WHERE fund_id = ap7_aktie OR
           fund_id = ap7_ranta AND
           DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
           DATEDIFF(week, nd) <= 0    # the true starting
day.
           ORDER BY week;
DECLARE CONTINUE HANDLER FOR NOT FOUND SET done5 = TRUE;

SET i = 0;
SET y_ranta = NULL;
SET y_aktie = NULL;

OPEN fund_yield5;

accumulator5: LOOP
    FETCH fund_yield5 INTO w, f_id, y;

    IF done5 THEN
        LEAVE accumulator5;
    END IF;

    SET age = ca - (data_year - YEAR(w));

    IF f_id = ap7_aktie THEN
        SET y_ranta = y;
    END IF;

    IF f_id = ap7_ranta THEN
        SET y_aktie = y;
    END IF;

    IF y_aktie IS NOT NULL AND y_ranta IS NOT NULL THEN
        SET i = i + 1;

        IF age < 56 THEN
            SET y = y_aktie;
        ELSE
            SET y = (y_aktie * (1 - 0.03 * (age - 55))) + (y_ranta
* 0.03 * (age - 55));
        END IF;

        # SELECT ca, YEAR(w), age;
        # SELECT age, age - 55, y_aktie, y_ranta, y;

        SET y_aktie = NULL;
        SET y_ranta = NULL;

        INSERT INTO ppm.temp_yield (week, r5) VALUES (w, y);
    END IF;

```

```
        END LOOP accumulator5;

        CLOSE fund_yield5;
    END fund5;
END IF;

END MAIN$$

DELIMITER ;
```

create_weekly_data.sql

```
DELIMITER $$

DROP PROCEDURE IF EXISTS ppm.create_weekly_data$$

CREATE PROCEDURE ppm.create_weekly_data(IN cp BIGINT(10) UNSIGNED,
                                        IN ca TINYINT(2)
                                        UNSIGNED,
                                        IN cd DATE,
                                        IN nd DATE,
                                        IN cf1 MEDIUMINT(6)
                                        UNSIGNED,
                                        IN cs1 TINYINT
                                        UNSIGNED,
                                        IN cf2 MEDIUMINT(6)
                                        UNSIGNED,
                                        IN cs2 TINYINT
                                        UNSIGNED,
                                        IN cf3 MEDIUMINT(6)
                                        UNSIGNED,
                                        IN cs3 TINYINT
                                        UNSIGNED,
                                        IN cf4 MEDIUMINT(6)
                                        UNSIGNED,
                                        IN cs4 TINYINT
                                        UNSIGNED,
                                        IN cf5 MEDIUMINT(6)
                                        UNSIGNED,
                                        IN cs5 TINYINT
                                        UNSIGNED)
main: BEGIN
  DECLARE a,i INT DEFAULT 0;
  DECLARE y1, y2, y3, y4, y5, y DECIMAL(10, 9) DEFAULT NULL; #
  Yields over the time period.
  DECLARE y_tot DECIMAL(13,9);
  DECLARE week_inflation DECIMAL(10,9);
  DECLARE w DATE;
  DECLARE over_conf_year INT DEFAULT 0;
  DECLARE l_n INT DEFAULT FALSE;
  DECLARE over_conf INT DEFAULT FALSE;
  DECLARE weeks_in_portfolio INT DEFAULT 0;
  DECLARE year_of_measure INT DEFAULT 2011;
  DECLARE year_of_birth INT;
  DECLARE g30, g40, g50, g60, g70, g80, g90 INT DEFAULT FALSE;
  DECLARE home_b FLOAT(7, 4);
  DECLARE temp_f MEDIUMINT(6) UNSIGNED;
  DECLARE active INT DEFAULT FALSE;
  DECLARE over_f MEDIUMINT(6) UNSIGNED;

  TRUNCATE TABLE ppm.temp_yield;
  #SELECT cp AS 'person_id', ca AS 'Age', cd AS 'Start date', nd AS
  'End date';
```

```

fund1: BEGIN
  DECLARE done1 INT DEFAULT FALSE;
  DECLARE fund_yield1 CURSOR FOR
    SELECT week, r
    FROM ppm.fund_return
    WHERE fund_id = cf1 AND
      DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
      DATEDIFF(week, nd) <= 0    # the true starting
day.
      ORDER BY week;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done1 = TRUE;

  IF cf1 IS NULL OR cf1 = 0
  THEN
    LEAVE fund1;
  END IF;

  SET a = a + 1;
  SET y1 = 1;
  SET i = 0;
  SET y_tot = 0;

  OPEN fund_yield1;

  accumulator1: LOOP
    FETCH fund_yield1 INTO w, y;
    IF done1 THEN
      LEAVE accumulator1;
    ELSE
      SET i = i + 1;
      SET y1 = y1 * y;
      INSERT INTO ppm.temp_yield (week, r1) VALUES (w, y);
    END IF;
  END LOOP accumulator1;

  IF i > 0 AND y1 IS NOT NULL THEN
    SET y_tot = y_tot + y1 * cs1 / 100;
    SET y1 = POW(y1, 1/i);
  ELSE
    SET y1 = Null;
    SET y_tot = Null;
  END IF;

  CLOSE fund_yield1;
END fund1;

fund2: BEGIN
  DECLARE done2 INT DEFAULT FALSE;
  DECLARE fund_yield2 CURSOR FOR
    SELECT week, r
    FROM ppm.fund_return
    WHERE fund_id = cf2 AND

```

```

                                DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
                                DATEDIFF(week, nd) <= 0      # the true starting
day.
                                ORDER BY week;
                                DECLARE CONTINUE HANDLER FOR NOT FOUND SET done2 = TRUE;

                                IF cf2 IS NULL OR cf2 = 0
                                THEN
                                    LEAVE fund2;
                                END IF;

                                SET a = a + 1;
                                SET y2 = 1;
                                SET i = 0;

                                OPEN fund_yield2;

                                accumulator2: LOOP
                                    FETCH fund_yield2 INTO w, y;

                                    IF done2 THEN
                                        LEAVE accumulator2;
                                    ELSE
                                        SET i = i + 1;
                                        SET y2 = y2 * y;
                                        UPDATE ppm.temp_yield SET r2=y WHERE week = w;
                                    END IF;
                                END LOOP accumulator2;

                                IF i > 0 AND y2 IS NOT NULL THEN
                                    SET y_tot = y_tot + y2 * cs2 / 100;
                                    SET y2 = POW(y2, 1/i);
                                ELSE
                                    SET y2 = Null;
                                    SET y_tot = Null;
                                END IF;

                                CLOSE fund_yield2;
                                END fund2;

                                fund3: BEGIN
                                DECLARE done3 INT DEFAULT FALSE;
                                DECLARE fund_yield3 CURSOR FOR
                                    SELECT week, r
                                    FROM ppm.fund_return
                                    WHERE fund_id = cf3 AND
                                DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
                                DATEDIFF(week, nd) <= 0      # the true starting
day.
                                ORDER BY week;
                                DECLARE CONTINUE HANDLER FOR NOT FOUND SET done3 = TRUE;

```

```

IF cf3 IS NULL OR cf3 = 0
THEN
  LEAVE fund3;
END IF;

SET a = a + 1;
SET y3 = 1;
SET i = 0;

OPEN fund_yield3;

accumulator3: LOOP
  FETCH fund_yield3 INTO w, y;

  IF done3 THEN
    LEAVE accumulator3;
  ELSE
    SET i = i + 1;
    SET y3 = y3 * y;
    UPDATE ppm.temp_yield SET r3=y WHERE week = w;
  END IF;
END LOOP accumulator3;

IF i > 0 AND y3 IS NOT NULL THEN
  SET y_tot = y_tot + y3 * cs3 / 100;
  SET y3 = POW(y3, 1/i);
ELSE
  SET y3 = Null;
  SET y_tot = Null;
END IF;

CLOSE fund_yield3;

END fund3;

fund4: BEGIN
  DECLARE done4 INT DEFAULT FALSE;
  DECLARE fund_yield4 CURSOR FOR
    SELECT week, r
    FROM ppm.fund_return
    WHERE fund_id = cf4 AND
      DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
      DATEDIFF(week, nd) <= 0      # the true starting
day.
    ORDER BY week;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done4 = TRUE;

IF cf4 IS NULL OR cf4 = 0
THEN
  LEAVE fund4;
END IF;

SET a = a + 1;

```

```

SET y4 = 1;
SET i = 0;

OPEN fund_yield4;

accumulator4: LOOP
  FETCH fund_yield4 INTO w, y;

  IF done4 THEN
    LEAVE accumulator4;
  ELSE
    SET i = i + 1;
    SET y4 = y4 * y;
    UPDATE ppm.temp_yield SET r4=y WHERE week = w;
  END IF;
END LOOP accumulator4;

IF i > 0 AND y4 IS NOT NULL THEN
  SET y_tot = y_tot + y4 * cs4 / 100;
  SET y4 = POW(y4, 1/i);
ELSE
  SET y4 = Null;
  SET y_tot = Null;
END IF;

CLOSE fund_yield4;

END fund4;

fund5: BEGIN
  DECLARE done5 INT DEFAULT FALSE;
  DECLARE fund_yield5 CURSOR FOR
    SELECT week, r
    FROM ppm.fund_return
    WHERE fund_id = cf5 AND
      DATEDIFF(week, cd) >= 7 AND # Because week date is
7 days after
      DATEDIFF(week, nd) <= 0 # the true starting
day.
    ORDER BY week;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done5 = TRUE;

  IF cf5 IS NULL OR cf5 = 0
  THEN
    LEAVE fund5;
  END IF;

  SET a = a + 1;
  SET y5 = 1;
  SET i = 0;

  OPEN fund_yield5;

  accumulator5: LOOP

```



```

    FETCH fund_yield5 INTO w, y;

    IF done5 THEN
        LEAVE accumulator5;
    ELSE
        SET i = i + 1;
        SET y5 = y5 * y;
        UPDATE ppm.temp_yield SET r5=y WHERE week = w;
    END IF;
END LOOP accumulator5;

IF i > 0 AND y5 IS NOT NULL THEN
    SET y_tot = y_tot + y5 * cs5 / 100;
    SET y5 = POW(y5, 1/i);
ELSE
    SET y5 = Null;
    SET y_tot = Null;
END IF;

CLOSE fund_yield5;

END fund5;

#
# Special case if fund_id is 100010 (AP7 Såfa) then the fund is a
composite of
# 545541 (AP7 Råntefond) AND 581371 (AP7 Aktiefond). Yay...
#
# The composition depends on age:
#
# age < 56:    100 % (AP7 Aktiefond) 0% (AP7 Råntefond)
#
# age <= 56:  [(100 - 3 * (age - 55)) % AP7 Aktiefond] + [3 *
(age - 55) % AP7 Råntefond]
#

IF cf1 = 100010 OR cf2 = 100010 OR cf3 = 100010 OR cf4 = 100010 OR
cf5 = 100010 THEN
    CALL create_yield_for_ap7_safa(ca, cd, nd, cf1, cs1, cf2, cs2,
cf3, cs3, cf4, cs4, cf5, cs5);
END IF;

#SELECT a, cf1, cf2, cf3, cf4, cf5;
#SELECT * FROM ppm.temp_yield;
#SELECT 'Total:', y1, y2, y3, y4, y5;
#SELECT y_tot;

#
# Is there 1/n bias?
#
# Model: cs1 = cs2 = cs3 = cs4 = cs5 = 20
#

```

```

SET l_n = (cs1 = 20) AND (cs2 = 20) AND (cs3 = 20) AND (cs4 = 20)
AND (cs5 = 20);
#SELECT '1/n done';

#
# Is there a active choice?
#
# Model: NOT (100% in Sjunde AP-fonden OR 100% pensions försäkring
OR 100% AP7-Såfa)
SET active = 1;

IF cs1 = 100 THEN
  IF cf1 = 110924 OR cf1 = 100008 OR cf1 = 100010 THEN
    SET active = 0;
  END IF;
END IF;

#
# Is there home bias?
#
# Sum of shares belonging to any of these fund cathegories:
# 'Svenska aktier och räntor',
# 'Sverige', 'Sverige index',
# 'Sverige kort',
# 'Sverige lång',
# 'Sverige realränta',
# 'Sverige småbolag'
#
# At a later stage the value is normalised replaced by:
# (trader's share swedish funds) / (market share of swedish
funds) '
#

SET home_b = 0;

IF cf1 IS NOT NULL THEN
  SELECT fund_id INTO temp_f FROM ppm.home_bias_funds WHERE
fund_id = cf1;
  IF temp_f IS NOT NULL THEN
    SET home_b = cs1;
  END IF;
END IF;

IF cf2 IS NOT NULL THEN
  SELECT fund_id INTO temp_f FROM ppm.home_bias_funds WHERE
fund_id = cf2;
  IF temp_f IS NOT NULL THEN
    SET home_b = home_b + cs2;
  END IF;
END IF;

IF cf3 IS NOT NULL THEN

```

```

    SELECT fund_id INTO temp_f FROM ppm.home_bias_funds WHERE
fund_id = cf3;
    IF temp_f IS NOT NULL THEN
        SET home_b = home_b + cs3;
    END IF;
END IF;

IF cf4 IS NOT NULL THEN
    SELECT fund_id INTO temp_f FROM ppm.home_bias_funds WHERE
fund_id = cf4;
    IF temp_f IS NOT NULL THEN
        SET home_b = home_b + cs4;
    END IF;
END IF;

IF cf5 IS NOT NULL THEN
    SELECT fund_id INTO temp_f FROM ppm.home_bias_funds WHERE
fund_id = cf5;
    IF temp_f IS NOT NULL THEN
        SET home_b = home_b + cs5;
    END IF;
END IF;
# SELECT 'Home bias:', home_b;

#
# Is there overconfidence?
#
# Model: (cf1 IN (TOP 25% yield past three years)) + (cf2 IN (TOP
25% yield past year)) ...
#

SET over_conf_year = YEAR(w) - 1;

IF over_conf_year >= 2002 AND over_conf_year <= 2010 THEN
    SET over_conf = TRUE;

    IF cf1 IS NOT NULL THEN
        SELECT fund_id INTO over_f FROM ppm.top_25 WHERE year =
over_conf_year AND fund_id = cf1;
        IF over_f IS NULL THEN
            SET over_conf = FALSE;
        END IF;
    END IF;

    IF over_conf AND cf2 IS NOT NULL THEN
        SELECT fund_id INTO over_f FROM ppm.top_25 WHERE year =
over_conf_year AND fund_id = cf2;
        IF over_f IS NULL THEN
            SET over_conf = FALSE;
        END IF;
    END IF;

    IF over_conf AND cf3 IS NOT NULL THEN

```

```

        SELECT fund_id INTO over_f FROM ppm.top_25 WHERE year =
over_conf_year AND fund_id = cf3;
        IF over_f IS NULL THEN
            SET over_conf = FALSE;
        END IF;
    END IF;

    IF over_conf AND cf4 IS NOT NULL THEN
        SELECT fund_id INTO over_f FROM ppm.top_25 WHERE year =
over_conf_year AND fund_id = cf4;
        IF over_f IS NULL THEN
            SET over_conf = FALSE;
        END IF;
    END IF;

    IF over_conf AND cf5 IS NOT NULL THEN
        SELECT fund_id INTO over_f FROM ppm.top_25 WHERE year =
over_conf_year AND fund_id = cf5;
        IF over_f IS NULL THEN
            SET over_conf = FALSE;
        END IF;
    END IF;
END IF;

#SELECT 'Over confidence done';

#
# Which generation do the person belong to?
#
# People born in december of the last year in a decade might be
placed in the wrong
# decade.. can't really tell since we don't know their birth
dates.
#
# generation = year_of_measure - age

SET year_of_birth = year_of_measure - ca;

IF year_of_birth >= 1930 AND year_of_birth < 1940 THEN
    SET g30 = TRUE;
END IF;

IF year_of_birth >= 1940 AND year_of_birth < 1950 THEN
    SET g40 = TRUE;
END IF;

IF year_of_birth >= 1950 AND year_of_birth < 1960 THEN
    SET g50 = TRUE;
END IF;

IF year_of_birth >= 1960 AND year_of_birth < 1970 THEN
    SET g60 = TRUE;
END IF;

```

```

IF year_of_birth >= 1970 AND year_of_birth < 1980 THEN
  SET g70 = TRUE;
END IF;

IF year_of_birth >= 1980 AND year_of_birth < 1990 THEN
  SET g80 = TRUE;
END IF;

IF year_of_birth >= 1990 AND year_of_birth < 2000 THEN
  SET g90 = TRUE;
END IF;

#
# Calculate yield on weekly investment. Skip weeks with incomplete
fund information.
#

save: BEGIN
  DECLARE done, got_null INT DEFAULT FALSE;
  DECLARE fund_yield CURSOR FOR
    SELECT week, r1, r2, r3, r4, r5
    FROM ppm.temp_yield;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

  OPEN fund_yield;

save_loop: LOOP

  FETCH fund_yield INTO w, y1, y2, y3, y4, y5;
  SET got_null = FALSE;
  SET y_tot = 0;

  IF done THEN
    LEAVE save_loop;
  END IF;

  #
  # Calculate yield and store result.
  #

  IF cs1 > 0 THEN
    IF y1 IS NOT NULL THEN
      SET y_tot = y_tot + y1 * cs1 / 100;
    ELSE
      SET got_null = TRUE;
    END IF;
  END IF;

  IF cs2 > 0 THEN
    IF y2 IS NOT NULL THEN

```

```

        SET y_tot = y_tot + y2 * cs2 / 100;
    ELSE
        SET got_null = TRUE;
    END IF;
END IF;

IF cs3 > 0 THEN
    IF y3 IS NOT NULL THEN
        SET y_tot = y_tot + y3 * cs3 / 100;
    ELSE
        SET got_null = TRUE;
    END IF;
END IF;

IF cs4 > 0 THEN
    IF y1 IS NOT NULL THEN
        SET y_tot = y_tot + y4 * cs4 / 100;
    ELSE
        SET got_null = TRUE;
    END IF;
END IF;

IF cs5 > 0 THEN
    IF y5 IS NOT NULL THEN
        SET y_tot = y_tot + y5 * cs5 / 100;
    ELSE
        SET got_null = TRUE;
    END IF;
END IF;

SELECT kpi INTO week_inflation FROM ppm.kpi WHERE week = w;

IF week_inflation IS NULL THEN
    SELECT 'week_inflation was null should not happen!', w;
    SET week_inflation = 1;
END IF;

SET y_tot = y_tot / week_inflation;

# Only save entry if we have yield information about all funds
the portfolio contains.
IF NOT got_null AND y_tot IS NOT NULL THEN
    #
    # How many weeks have passed since the fund choice was made?
Measure of activity.
    #

    SET weeks_in_portfolio = weeks_in_portfolio + 1;

    #
    # Save data.
    #

```

```

        INSERT INTO ppm.result (person_id, week, age, l_n_bias,
over_confidence, weeks_since_choice,
                                born_1930, born_1940, born_1950,
born_1960, born_1970, born_1980, born_1990,
                                home_bias, active_choice, r)
        VALUES (cp, w, ca, l_n, over_conf, weeks_in_portfolio,
                g30, g40, g50, g60, g70, g80, g90,
                home_b, active, y_tot);
ELSE
    # Record the dud
    UPDATE ppm.status_log SET number = number + 1 WHERE
message = 'Missing fund yield';
END IF;

END LOOP save_loop;

CLOSE fund_yield;

END save;
END main$$

DELIMITER ;

```

find_frequent_traders.sql

```
DELIMITER $$

DROP PROCEDURE IF EXISTS ppm.find_frequent_traders$$

CREATE PROCEDURE ppm.find_frequent_traders()
#
# Finds how many transactions a trader performs a given year.
# The result is stored in the ppm.result table.
#
main: BEGIN
  DECLARE done_person INT DEFAULT FALSE;
  DECLARE p_id BIGINT(10) UNSIGNED;
  DECLARE num_t INT;
  DECLARE transaction_year INT;
  DECLARE start_year INT DEFAULT 1999;
  DECLARE end_year INT DEFAULT 2010;
  DECLARE i INT DEFAULT 0;
  DECLARE person CURSOR FOR
    SELECT DISTINCT person_id FROM ppm.fund_choice;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done_person = TRUE;

  OPEN person;

  person_loop: LOOP
    FETCH person INTO p_id;

    IF done_person THEN
      LEAVE person_loop;
    END IF;

    SET transaction_year = start_year;
    year_loop: LOOP
      IF transaction_year > end_year THEN
        LEAVE year_loop;
      END IF;

      SELECT COUNT(*) INTO num_t FROM ppm.fund_choice
        WHERE person_id = p_id AND YEAR(transaction_date) =
transaction_year;

      IF num_t IS NOT NULL THEN
        UPDATE ppm.result SET num_trades = num_t
          WHERE person_id = p_id AND YEAR(week) = transaction_year;
      END IF;

      SET transaction_year = transaction_year + 1;
    END LOOP year_loop;

    SET i = i + 1;
    SELECT i, p_id;

  END LOOP person_loop;
```



```
CLOSE person;  
END main$$  
DELIMITER ;  
CALL ppm.find_frequent_traders();
```

calculate_home_bias.sql

```
DELIMITER $$

DROP PROCEDURE IF EXISTS ppm.calculate_home_bias$$

CREATE PROCEDURE ppm.calculate_home_bias()

# Normalize the home bias by dividing each individual share of
# swedish funds
# with the market average of swedish funds for that week.

main: BEGIN
  DECLARE done INT DEFAULT FALSE;
  DECLARE w DATE;
  DECLARE avg_bias FLOAT(7,4);
  DECLARE weeks CURSOR FOR
    SELECT DISTINCT week FROM ppm.result;
  DECLARE CONTINUE HANDLER FOR NOT FOUND SET done = TRUE;

  OPEN weeks;

  week_loop: LOOP
    FETCH weeks INTO w;

    IF done THEN
      LEAVE week_loop;
    END IF;

    SELECT AVG(home_bias) INTO avg_bias FROM ppm.result WHERE week =
w;
    UPDATE ppm.result SET home_bias = home_bias / avg_bias WHERE
week = w;

    SELECT w, avg_bias;

  END LOOP week_loop;

  CLOSE weeks;

END main$$

DELIMITER ;

CALL ppm.calculate_home_bias();
```